UNIVERSITÉ DU LUXEMBOURG

PhD-FSTM-2022-114
The Faculty of Science, Technology and Medicine

DISSERTATION

Defence held on 21/10/2022 in Esch-sur-Alzette

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN SCIENCES DE L'INGÉNIEUR

by

Daniele PROVERBIO
Born on 02 July 1994 in Borgomanero (Italy)

CLASSIFICATION AND DETECTION OF CRITICAL TRANSITIONS: FROM THEORY TO DATA

Dissertation Defence Committee:

Dr Jorge Gonçalves, dissertation supervisor
Professor, Université du Luxembourg

Dr Alexander Skupin, Chairman
Professor, Université du Luxembourg

Dr Rudi Balling, Vice Chairman
Professor, Universität Bonn

Dr Peter Ashwin
Professor, University of Exeter

Dr Marco Cosentino Lagomarsino
Professor, Università degli Studi di Milano

To my dear ones

“It is difficult to make predictions, especially about the future.”

- Niels Bohr -

Declaration

I hereby declare that, except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Daniele Proverbio
November 2022

Acknowledgements

Such a journey, and it's already ending. This thesis is a clean example that research - and science in general - is a social activity. Far from isolated ivory towers, very little would have been possible without the exceptional contribution of colleagues and friends, and without the constant support of those close to me. I owe you all my deepest gratitude!

First, I would like to thank my principal supervisor Jorge Gonçalves for his attentive guidance, his witty remarks and the intellectual flexibility he allowed me: coming from a control engineer, this was a double bliss! Many thanks to my co-supervisors Alex(ander) Skupin and Peter Ashwin: you gave me so many opportunities to grow as a researcher that I cannot count - projects, insights, travels. A special thank to Rudi Balling, the former director of the LCSB, for having created such a multidisciplinary and stimulating environment. Another special thank to Stefano Magni, almost an adjoint supervisor, to Johan Markdhal, bringing amoebas to other dimensions, to Françoise Kemp, a great partner in crime, and to Andreas Husch, who believed in a little physicist and suggested so many working challenges.

A big thank to my colleagues, I really enjoyed working with you. Thanks to whom I managed to collaborate with and to the others, always keen to share an idea, an insight, a cup of tea: Arthur, Atte, Beatriz, François, Gustavo, Inés, Jan, Jesùs, Laurent, Maria, Marino, Matias, Mehri, Meryem, Mila, René, Vladimir, Xin. Thanks to all members of other amazing groups: the CriTiCS DTU first and foremost, the ResearchLuxembourg COVID-19 taskforce WP6, the ICS group, the RSG Luxembourg, the dynamical group in Exeter, Matti from cold Finland. There are many other people, in and outside the LCSB, that made this journey unforgettable; you know you'll be stalked for extra projects and ideas!

Thanks to Sofia, Laurence, Sabine and all the secretaries, communication team and staff. Much wouldn't have been possible without your reliable support.

Thanks to my friends, old and new: you have been a mind-easing presence.

Thanks to my family, always interested, cheering and motivating. Without you, I wouldn't be what I am - for that, I am truly grateful.

And thanks to Chiara: you've been everything.

So long, and thanks for all the fish.

Abstract

From population collapses to cell-fate decision, critical phenomena are abundant in complex real-world systems. Among modelling theories to address them, the critical transitions framework gained traction for its purpose of determining classes of critical mechanisms and identifying generic indicators to detect and alert them ("early warning signals"). This thesis contributes to such research field by elucidating its relevance within the systems biology landscape, by providing a systematic classification of leading mechanisms for critical transitions, and by assessing the theoretical and empirical performance of early warning signals. The thesis thus bridges general results concerning the critical transitions field – possibly applicable to multidisciplinary contexts – and specific applications in biology and epidemiology, towards the development of sound risk monitoring system.

Table of contents

Glossary

AC(1)	Lag-1 Autocorrelation.
ARIMA	AutoRegressive Integrated Moving Average.
AUC	Area Under the (ROC) Curve.
CSD	Critical Slowing Down.
CT	Critical Transition.
EEG	ElectroEncephaloGram.
EKF	Extended Kalman Filter.
EWS	Early Warning Signal.
MAF	Min/Max autocorrelation factors.
MCMC	Markov Chain Monte Carlo.
MFPT	Mean First Passage Time.
ML	Machine Learning.
O-U	Ornstein-Uhlenbeck.
ODE	Ordinary Differential Equation.
PCA	Principal Component Analysis.
PDE	Partial Differential Equation.
PDF	Probability Density Function.
ROC	Receiver Operating Characteristic.
SDE	Stochastic Differential Equation.
SEIR	Susceptible-Exposed-Infectious-Removed.
SIR	Susceptible-Infectious-Removed.
SOC	Self-Organised Criticality.
SPDF	Stationary Probability Density Function.
SVD	Singular Value Decomposition.
WHO	World Health Organization.

Chapter 1

Introduction

1.1 Forewords

Imagine swinging on a chair. The equilibrium remains stable until a certain angle is accidentally hit, and you crash onto the ground. Suddenly losing stability and falling onto a new equilibrium is a hallmark of critical transitions. This example pertains to *static* equilibria; however, many natural systems are characterised by *dynamic* equilibria — repeated patterns over time — and transitions between them. The present thesis concerns the study of critical transitions in complex dynamical systems, with a specific focus on biological and epidemiological ones. Of particular interest is the taxonomy of recognised critical transitions and their systematic classification, followed by studies about the application of current knowledge for identification, detection and, potentially, prediction of critical transitions in monitored natural systems.

The present thesis has been developed in a highly interdisciplinary environment, with experimental biologists working hand-in-hand with data analysts and computational scholars, further paired up with physicists, mathematicians and engineers. They have different backgrounds and know-how. Hence, to set some common ideas and vocabulary that will be employed in the writing, this introductory section puts forward some notions about complex systems and models – mathematical, mainly. Later in the text, specific subsections will also summarise the main mathematical key points, to disseminate ideas and methods to other audiences and, hopefully, to stimulate interdisciplinary collaborations.

This chapter is concluded with an overview on the thesis objectives and on the contributions associated with the present PhD project.

1.2 Models (with focus on mathematical biology)

Primary goals of scientific activities are knowledge, prediction and control (Hepburn et al., 2021). To achieve such outputs, the scientific community has relied on method and methodologies, constantly updated and refined over the centuries and within disciplines. This section first presents an introductory overview, focusing on models and methods employed in biological and

mathematical-biology disciplines, and is concluded with two personal examples on the emerging role of mathematical and computational models as *in silico* testbeds for the natural sciences.

1.2.1 The scientific method and the "cycle for knowledge discovery"

> The sciences do not try to explain, they hardly even try to interpret; they mainly make models. By a model is meant a mathematical construct which, with the addition of certain verbal interpretation, describes observed phenomena. The justification of such a mathematical construct is solely and precisely that it is expected to work (J. von Neumann)

According to textbooks, the scientific method involves back-and-forth interplay between experiments, observations of natural phenomena and theoretical models, to guide our insights and contribute to building knowledge. The main purpose of the scientific method is, ultimately, to foster an objective understanding of observable phenomena. The whole process is guided by research questions that shape study designs and circumscribe the systems' properties that we care to focus on. By no means the scientific method is linear and uniquely characterised, although many scholars have, over the centuries, attempted to rationalise it. To make things more complicated, it might occur that different disciplines developed alternative or nuanced methods, in order to best exploit the instruments they have at disposal (Mazzocchi, 2019). Nonetheless, shared characteristics among the scientific method(s) can be recognised and a general skeleton depicted, to establish a common ground for further discussions.

To begin with, the method is an iterative, cycling process (see Fig. 1.1). Its end products are models – of potentially various kinds (see below). The cycle can be broken down in arbitrarily fine-grained steps, of which at least five are of paramount importance (Saetzler et al., 2011):

- Experiments. Foundation of scientific endeavours are experiments, to empirically observe natural phenomena.
- Data acquisition. Measuring the properties of systems observed during experiments and then curating, aggregating and analysing the corresponding data to systematize the collected information (Huang et al., 2013). If the data acquisition is expected to be hypothesis-driven (*cf.* Sec 1.2.3), this step can also help tuning the experimental procedures.
- Hypothesis formulation. Hypothesis about the functioning of observed phenomena are usually informed by acquired data and guided by mental processes such as induction (inference from repeated particular cases to the general case) or transformational reasoning (envisioning the effect of various transformation that the phenomenon might undergo, thus getting "a sense" of it, Simon (1996)).
- Model development. From hypothesis to modelling, this is the process of formalising the expected theory. Developing a model also helps to foster new hypothesis.
- Models are used to make predictions, which should be novel with respect of what's already known. They can be obtained by solving equations, simulating algorithmic models, or deducting properties from theoretical principles (Sherry, 2006). Depending whether the predictions hold against experiments, models can then be updated or even built anew (Kuhn, 1962). Every piece of knowledge so gained is not considered true in an absolute sense, but valid (or significant) until falsified by new contrasting evidences (Orzack, 2012).

- Back to experiments: predictions are contrasted against empirical observations and carefully tested. Confirming or falsifying the predictions (Popper, 2014; Oreskes et al., 1994) triggers updates in the underlying models and, eventually, guides the formulation of new hypothesis. When using statistical tools, this process is often associated with hypothesis testing, a form of inference to draw conclusions about model properties. Note the slight redundancy in terminology: here, "hypothesis" stands for alternative solutions from model predictions. In the other step, "hypothesising" encompasses the formulation of assumptions, modelling choices and preliminary pieces of novel theory.

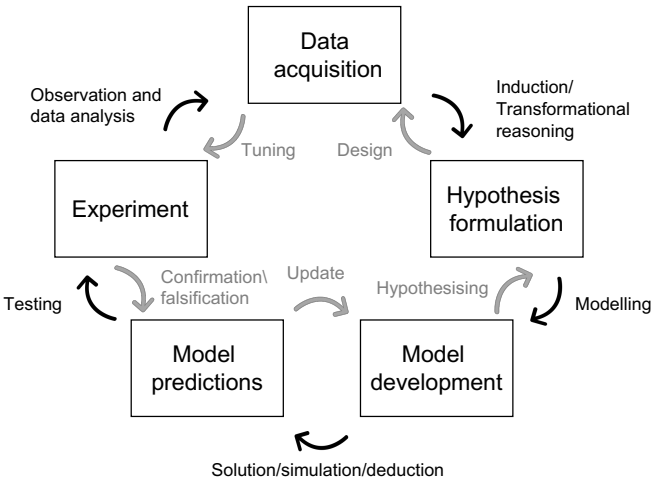


Fig. 1.1 Schematic illustration of the systems biology two-ways cycle for knowledge discovery, adapted and elaborated from Oreskes et al. (1994); Simon (1996); Saetzler et al. (2011); Huang et al. (2013). **Black path:** Careful observations of certain aspects of natural phenomena (experiments) produce data. Scientists reflect upon data in order to generate hypothesis, following various reasoning strategies including induction or "transformational reasoning". Iterating and distilling hypothesis eventually lead to develop models (*cf.* Sec. 1.2.2). Working upon models with analytical solutions, simulations, deduction etc. yields predictions. These are contrasted against new experiments, to check if predictions are confirmed. Accumulating such confirmations supports the verification or falsification of models and of their underlying hypothesis. **Grey path:** the inverse process of updating the "prior" steps in the black path, given the outcome of "posterior" steps. As the whole process is iterative, the two paths might be potentially equivalent in the outcome, but still differ in approach, methodology and procedures.

Entry points in this "knowledge cycle" are usually observations of natural phenomena (data-driven entry) or from hypothesis made about a certain system's behaviour being sufficiently analogous to that of another, well studied one (hypothesis-driven entry). As anticipated, the products are constantly refined models used for predictions, knowledge and control.

1.2.2 A non-exhaustive introduction to scientific models

[1]In general, a model is an idealised representation of a certain situation or phenomenon, according to the purpose of the modeller who selects the features to consider (often called "modelling choices"). Wind tunnels, for example, are models designed to research the effect of wind on automotive and airplane design. They enable researchers to simulate what happens when an object

[1]Part of this section has been adapted from "Models", in Proverbio D., Ley C. and Mangers P., *Coronavirus technical terms – explained by scientists from Luxembourg*, science.lu, 2020 (https://bit.ly/3qy02m5).

passes through the air in a controlled environment. To answer specific issues, science heavily relies on models (Jensen, 1998; Giere, 2004). Note that, under this framework, models can be idealized and controlled representations of any kind. In philosophy and mathematics, model theory is associated with the study of formal languages and their interpretation (Hodges, 2020), but in scientific methodologies as well as routine practice, models often play the role of archetypes or of substitutes for the object of study. Hence, there exist animal models, in vitro models, industrial models and mathematical models.

Mathematical models

Mathematical models are abstract representations, written in mathematical terms, whose elements are put in correspondence with the components of real-world phenomena. Once a plausible correlation has been established, the mathematical model is said to describe the given phenomenon. Mathematics is a formal discipline, with universal logical rules that can be derived to obtain results. This means that mathematical deductions and predictions can be proven and verified as correct or untrue, whereas other models have lower confidence levels.
Even mathematical models, despite their abstract character, have some limits. The bulk of simplifications, features, context and initial decisions form the model's assumptions. There is no model that can accurately represent the full reality. Attempting to include every possible real-world consequence would result in the most comprehensive description, but the mathematical complexity would be impossible to manage. Similarly, too simplified systems may become mathematically trivial and fail to accurately describe the original occurrence. In this spirit, A. Einstein is said to have written:

> Everything should be made as simple as possible, but not simpler.

Another witty quote, reflecting the degree of imprecision of mathematical models, but their nonetheless importance for knowledge, predictions and control, is:

> All models are wrong, but some are useful (G. Box).

In fact, mathematical models are extremely useful in multiple ways. First, for knowledge discovery. Studying the properties of a model is likely to shed light on our understandings, to answer certain questions and to inquire important mechanisms of the considered phenomenon. Second, to make predictions, *i.e.*, answering the question "what will happen"? Once an evolution or transformational rule has been established for a given system, it is possible to predict its development under multiple applications of such rule, given the initial conditions. Third, to inquire scenarios: "what would happen if...?" Mathematics as a formal language allows precise control over all involved variables. It is thus possible to modify some of them into alternative settings and perform counterfactual analysis, or to maintain the current settings in situations of larger uncertainty and investigate what is plausible to happen, should the situation be maintained. The uncertainty could be associated to scarce data about the initial conditions, openness of the system involving external and/or unpredictable inputs, intrinsic complexity of the system and so on. In this case, we speak of "projections" rather than "predictions", with corresponding degrees of confidence. Finally, the three points above can be used for system control: the predictions (or projections, if sufficient over the desired time horizon) of a suitably confirmed model can be combined with models of expected outcomes, to aptly steer a system of interest.

Classification of mathematical models

There is a whole spectrum of mathematical models, and their classification can be performed from different angles (Sherry, 2006; Gertsev et al., 2004). Models can be classified based on their object of study (distinguishing between simple and object-system models), on their fundamental operators (hence having linear, non-linear, algorithmic... models), on the way they treat parameters and variables (distinguishing between continuous or discrete-time/space models, static or dynamic, deterministic or non-deterministic, n-dimensional, qualitative or quantitative, etc.), on the primary modelling goal (are they descriptive, optimization or decisional), or on the implementation method (algebraic or approximate analytical models, or numeric vs simulated algorithmic ones) (Ashikhmin et al., 2004)[2].
A useful working classification, to navigate among the various modelling suggestions employed in mathematical[3]-, systems[4]-, synthetic[5]- or computational[6]-biology, is sketched in Fig. 1.2 and described below. The proposed classification is mainly based on object of study and modelling goal:

- "Pure" mathematical models. They focus on the study of abstract and formal properties of functions, sets, variables etc, as well as on method development in abstract settings. They do not have the primary purpose of reproducing data and observations, so they are placed outside the "spectrum" of scientific models for the natural sciences. Nonetheless, they constitute the foundation of all others.

- Physics-like (or prime-principled) models. They are based on first principles, tested and confirmed many times, to explain, characterise and predict the behaviour of a system. To date, they are the most powerful and accurate tool for studying physical and chemical systems. An overarching goal in other systemic disciplines is to develop similar predictive models for complex systems, although their verification is recognised to be extremely challenging (Oreskes et al., 1994).

- Phenomenological models. They try to reproduce observations (often, time evolutions) by referring to physical or biological/financial/ecological/... principles (depending on the discipline in which they are applied). They are known to be simplified and provide fewer causal explanations, but they allow many preliminary studies and plausible predictions. Among them, the so-called "minimal models" focus on a minimal set of mechanisms needed to produce a certain behaviour of interest.

- Optimization models. They can use either first principles (when there are any) or indicators, with the aim of reproducing and explaining a phenomenon in terms of "it is the result that optimises a certain function". They still have a degree of subjectivity due to the choice of optimisation criteria (which undergoes modelling choices as well), but they generate results

[2] Original version in Russian. An article translating and summarising the excellent contribution of this referenced book is available on Medium: Pospeev L., *Classification of mathematical models*, medium.com, Oct 13, 2019 (https://bit.ly/3FuZeD3).
[3] Jones et al., 2009; Hunt et al., 2008
[4] Kitano, 2002
[5] Benner et al., 2005
[6] Cahan et al., 2021; Fisher et al., 2006; Borshchev et al., 2004

which, it can be shown, are the best of all possible outcomes. In addition to knowledge development, the main end product of such models is the control of desired systems.

- Statistical models. They try to describe and reproduce observed data by means of objective indicators (mean, variance, etc.) and to reconstruct their evolution by means of "simple" functions (in the sense that they are combinations of basic functions, but not derived from first principles and subsequent manipulations). Statistical models are useful to capture trends in data and to extrapolate short-term predictions, but have usually little explanatory power.

- Indicators. They are simple tools for quantifying some quantity that, by common agreement, is judged potentially interesting. They are only descriptive and defined in a fairly subjective way. They have the advantage of enabling quantitative comparison between system elements.

The introduced classes are rather fuzzy and should not considered rigid but part of a spectrum (Fig. 1.2): models employed in everyday research are usually combinations to various degrees. For instance, regulatory network models combine statistical and optimization models to build something phenomenological that, if confirmed by controlled laboratory experiments, might develop into principle-based models. Same goes for many dynamical models to explain and reproduce the temporal behaviour of observed systems.

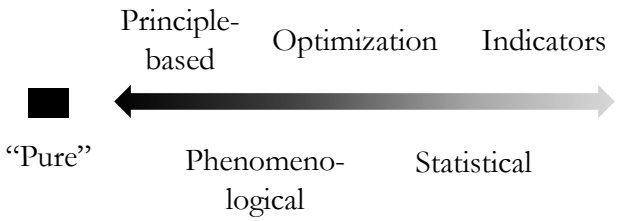


Fig. 1.2 Conceptual scheme representing the model classification described in the present section. "Pure" mathematical models (i.e. those whose main goal is to study mathematical models themselves) stand out but are basis of all others, to develop new models and verify the existing one. Then, a spectrum ranging from principle-based models to indicators accommodates the fuzzy classes described in the section.

Mathematical models as *in silico* testbeds: personal examples

Thanks to improved modelling techniques and better computation power, mathematical and computational models can be used to replace – up to a certain extent – the object of study. They could serve as abstract testbeds, in place of the real biological systems of interest, employed to pursue the goals described above. Two personal examples are shared below.

My Master thesis ("A Multi Agent System approach to complex micro-biological systems - In silico simulation of Dictyostelium Discoideum colonies") developed an agent-based model to reproduce swarming and gathering behaviours of *D. discoideum* colonies (social amoeboids). The model built upon biological knowledge to first reproduce their decentralised gathering patterns. The control mechanisms were fully accessible: it was thus possible to estimate the degree of robustness of the gathering process (Proverbio et al., 2020) and to test settings otherwise difficult

to design and obtain in in-vitro experiments. Overcoming their limitations was one among the model's benefits.

Another example is the ETH and Cambridge University thesis discussed by Master student J.L. Kruüsemann ("Modelling the behaviour of a synthetic gene circuit for controlled gene expression in Chlamydomonas reinhardtii"), whom I had the pleasure to co-supervise during my PhD. He developed a minimal dynamical model describing the regulation of ON-switches for controlled gene expression in *Chlamydomonas reinhardtii*. The model allowed predicting plausible parameter ranges for the ON-switch, to be further tested and verified in laboratory experiments. Among other insights, the model thus served as a first-pass tool to guide experiment design ad hypothesis generation, saving time and resources.

1.2.3 Methodologies

On top of the scientific method, scholars in various fields have developed structured methodologies, which include the values and justifications behind a particular characterization of scientific method (Hepburn et al., 2021). Such values include objectivity, reproducibility, simplicity, or past successes. In this sense, we could argue that society and culture do influence the practice of science, in addition to the self-corrective empirical method at its core.

Examples of methodological practices include the FAIR methodology for data management (Mayer et al., 2021), which is nowadays diffusing in the systems biology community. It stands for "findable, accessible, interoperable and reusable" and provides guidelines and principles to follow in case of data or code management and sharing. When applicable, the studies conducted for this thesis followed the FAIR methodology (e.g. compliance with data anonymization and referencing, code licencing and sharing, use of interoperable tools, etc.).

Another methodological distinction employed in this work is that between exploratory and confirmatory data analysis. Data analysis for natural samples, performed with qualitative and quantitative methods, can either aim at explore trends and patterns in data to generate hypothesis, or to test hypothesis and prediction, ideally confirming them. "Exploratory" researchers do not usually hold prior assumptions or hypotheses; such practice is sometimes disrespectfully called "fishing expedition" and is usually data-driven. Confirmatory data analysis, on the other hand, is hypothesis-driven and used to evaluate evidence, by challenging their assumptions about the data. Essentially, it serves to test a theory robustness by testing hypothesis and predictions against data from designed experiments. Depending on the type of analysis, research approaches and design might differ. Confusing the two methodologies might lead to poor scientific practices like HARKing – Hypothesizing After the Results are Known (Kerr, 1998; G. Butler et al., 2014).

1.3 Complex systems

This thesis addresses the issue of critical transitions in complex dynamical systems, whose general concepts are presented here.

A system is termed "dynamical" if it can be aptly modelled as a mathematical construct defined by a tuple $\{\Omega, \phi^t, \mu\}$, where Ω is a phase (or state) space, endowed with a family of evolution functions ϕ^t (trajectories) for each time point $t \in T$, and μ is a properly defined metrics

(see also Sec 2.2). In this thesis, the possibility of modelling the considered systems as *dynamical systems* is assumed, and the wording "are modelled as" will be often implied.

The term "complex systems" encompasses those systems whose behaviour is intrinsically difficult to tackle due to dependencies, interactions or relationships between their parts or between the given system and its environment (Mitchell, 2009; Ladyman et al., 2013). Epistemologically speaking, "the study of complex systems regards collective, or system-wide, behaviors as the fundamental object of study; for this reason, complex systems can be understood as an alternative paradigm to reductionism, which attempts to explain systems in terms of their constituent parts and the individual interactions between them" (Bar-Yam, 2002). In the biological sciences, complex systems theory includes disciplines and methodologies working towards "replacing black box geneticism with a larger model of a mutually interacting set of evolutionary/ developmental/communal/ecological dynamic processes" (Hooker, 2011), running at different sets of time- and spatial scales. Examples of such disciplines include dynamical systems theory, network theory (Milo et al., 2002) and control engineering (Iglesias et al., 2010). In general, systems that are deemed "complex" usually display some of the following properties (Hooker, 2011):

- Non-linear interactions and non-additivity. For some variable x, linearity means that the interaction law F is proportionately dependent to x: $F(kx) = kF(x)$ (with k a number). When this relationship breaks, the system is non-linear, and numerical combinations of solutions are in general not solutions (non-additivity). Non-linearity is a necessary[7] condition for complex dynamical behaviour.

- Irreversibility. In classical mechanics, processes can run forward and backward in time following the same laws. Real processes are instead dissipative and cannot run in reverse. A supply of energy from outside the system is necessary to maintain its processes (unless the system is very close to dynamical equilibrium). Living systems are topic examples of this kind. Irreversible processes might also be characterised by hysteresis (see Sec. 2.4.2).

- Constraints. Systems' rules of evolution are derived from a model of dynamical variable interrelationships, restrictions, and initial conditions (Cauchy problem). Constraints are limitations on a system's connections among variables that result from imposed physical conditions. Many complex systems have state-dependent limitations that aren't energy-conserving (non-holonomic).

- Static and dynamic equilibria. Some aspect A of a system is at equilibrium iff

$$\sum_i F_i(A) = 0 \quad \left(\text{or} \sum_i J_i(A) = 0 \right), \tag{1.1}$$

  that is, there is no net force (or flux) and no net 2^{nd} order rate of changes acting on that A. As a result, trends in A are conserved. Depending whether the time invariance concerns state parameters and variables (A = system state), or it concerns process parameters and rate variables (A = system processes), we talk of static or dynamic equilibrium. Dynamical equilibria in dissipative systems are usually maintained by sustaining energy from outside the

[7] But not sufficient: gravitational interactions are non-linear and non-additive, but the 2-bodies problem is not particularly complex.

system. Equilibria can be stable, unstable or meta-stable against some class of disturbances, depending whether they return to their equilibrium state after the disturbance (see Sec. 2.2.2 for additional information).

- Amplification. Stable equilibria that soon return to their resting state or its vicinity (basin of attraction) dampen the disturbances. When, instead, the system leaves the basin of attraction, or this is drastically transformed, the disturbance is amplified.

- Sensitivity to initial conditions. Due to non-linearities, small differences in system state can be amplified into large differences in subsequent system trajectory. Chaotic systems are one topical example (Ott, 2002). Systems living close or further to the edge of the basin of attraction might also undergo different trajectories under small disturbances, as well as those residing next to critical points (see Sec. 2).

- Path dependence (or "the importance of how you got there" (Ashwin et al., 2015)). In determining its evolution, not only the position in the state space might be important, but also the system rate of change and its historicity. They both can drive disturbance amplification and critical transitions.

- Finite deterministic unpredictability. As small uncertainties might be amplified, system predictability might be limited to knowledge of the initial state and soon fall into big uncertainties.

- Symmetry breaking. Symmetries are invariances under an operation, e.g. a rigid body is symmetric under translation along an axis. Existing symmetries can be disrupted (broken) either spontaneously or by external influences. Symmetry breaking often privileges some equilibria instead of others and is normally regarded as yielding emergence, order and complexity of many natural systems.

- Bifurcation. When structural instabilities (usually modelled as process parameter(s)) in a system lead to qualitative changes in its dynamical form (interpreted as changes of the attractor landscape), we speak of bifurcations. They can be local or global and often drive threshold phenomena next to criticality (see Sec 2.2.3).

- Self-organization. When some overall order emerges from an initially disordered system, due to local interactions of its parts, we speak of self-organization. The process is spontaneous, not needing control by external agents (Camazine, 2003). Biological systems are thought to have evolved following self-organising patterns, usually close to bifurcation points (criticality). This also implies the notion of order and organization, somewhere in between symmetric crystal structures and completely randomic dis-orders like those of gases. Through self-organising processes, natural systems (living ones in particular) may achieve autonomy (an internally organized capability for acquiring and directing ordered free energy from the environment to recover dissipated processes) and adaptation, adaptiveness and learning (the ability to maintain autonomy in dynamical environments).

- Emergence. When new qualitative behaviours are unexpectedly observed in complex systems, scientists might talk of "emergence". It is usually considered an epistemological,

not ontological, category. Two definitions are provided, quoting from O'Connor et al. (2015). *Predictive*: "Emergent properties are systemic features of complex systems which could not be predicted (practically speaking, or for any finite knower, or for even an ideal knower) from the standpoint of a pre-emergent stage, despite a thorough knowledge of the features of, and laws governing their parts". *Irreducible-Pattern*: "Emergent properties and laws are systemic features of complex systems governed by true, law-like generalizations within a special science that is irreducible to fundamental physical theory for conceptual reasons. The macroscopic patterns in question cannot be captured in terms of the concepts and dynamics of physics".

- Modularity and hierarchy. If a system consists of separate dynamical modules, its dynamics can be expressed as the convolution of inter- and intra-modules dynamics. Hence, a tweak on its components does not directly affect all the others. Modules can be arranged horizontally (modules at the same level), vertically (certain modules control others as an expression of hierarchical structures), or mixtures. Modules usually regulate each other through feedback mechanisms.

- Subdivision in levels/scales/scopes[8]. The overall processes observed in natural systems are often multi-scale (coordination from sub-cellular processes to organisms), multi-dimensional (involving organised interactions among many body parameters), and multi-phasic (asynchronous and non-stationary): they occur on many different timescales, often separated in terms of slowly vs fast varying variables and parameters. Whether such subdivision is epistemic (i.e. constructed by the observer) or real (stemming from differentiated hierarchical structures and modules) is still matter of debates (Ryan, 2007) and of modelling choices, for which hard discriminations do not exist and consensus has yet to be reached. The subject of critical transitions heavily involves timescale separations, which will be further discussed in Sec. 2.2.

In addition to system properties, other two characteristics influence the modelling of complex systems (Hooker, 2011): lack of analytical solutions for dynamical equations, resulting in the necessity of computational simulations to inquire their evolution; and condition-dependent laws, implying that – at least not at all scales and scopes – no simple set of universal laws can be deduced in advance (contrary to what happens in physics).

[8] "Level" is a somewhat blurry concept relating to observational levels of attention. "Scope" as a better term has been proposed by Ryan (2007), while "scale" is often used in the biological sciences in quotes like "bridging the scales" or"multiscale biology" (Eissing et al., 2011). Despite their subtle differences, whose full disclosure is beyond the scopes of this section, all three concepts refer to natural systems being modelled as layers of micro, meso and macro-states, and to modelling choices being guided by "zoom" and "resolution" set by the observer. The general idea is: contrary to "traditional" physical problems, it is hard if not impossible to identify systems' closures (where does it start and where does it end). However, some properties help guiding the definition of the system itself, i.e. functions holding up to certain dimensions or until certain time horizons. In this sense, similarly to photography, scope (or scale) is defined by both spatial and temporal boundaries, while resolution is defined as the finest distinction that can be observed between two alternative system configurations. The interplay between processes at different scales might lead to observational emergence (see above).

1.4 Thesis objectives and overview

This thesis focuses on the study of so-called Critical Transitions (CT) in complex dynamical systems, and on bridging mathematical theory and real-world problems. To begin with, the present work provides a comprehensive review of this rapidly developing research framework; then, it aims at assessing the power and reliability of theoretical results in the monitoring of biological and epidemiological systems. To this end, proposed indicators to detect and anticipate critical transitions (so-called Early Warning Signals, EWS) will be carefully analysed and tested both in analytical and simulated models, and on real-world data. Their sensitivity to modelling choices and dynamical settings, as well as their reliability in providing effective signals, will be carefully evaluated. This way, the present thesis contributes to the study of critical phenomena in natural (mainly biological and epidemiological) systems and sets solid foundations for the development of model-based risk assessment methods.

1.4.1 Thesis structure

The thesis chapter division is presented in Fig 1.3. It reflects the project structure (and how it developed) and is subdivided into three consecutive macro-blocks. The first one reviews literature results, suggests a coherent system out of the different contributions in various disciplinary fields and identifies open research questions. The second block contributes to theoretical and methodological results for EWS, both for general theory and for specific biological models. The third block pertains testing and applying theoretical results to empirical data (specifically, epidemiological data from the COVID-19 pandemic), as well as lessons learnt from alternative methods for real-time surveillance. The thesis is concluded with a Discussion chapter, summarising the project journey and portraying potentially interesting future directions.

A brief description of each chapter, as well as disclosure of the contributions of the PhD project, is provided below.

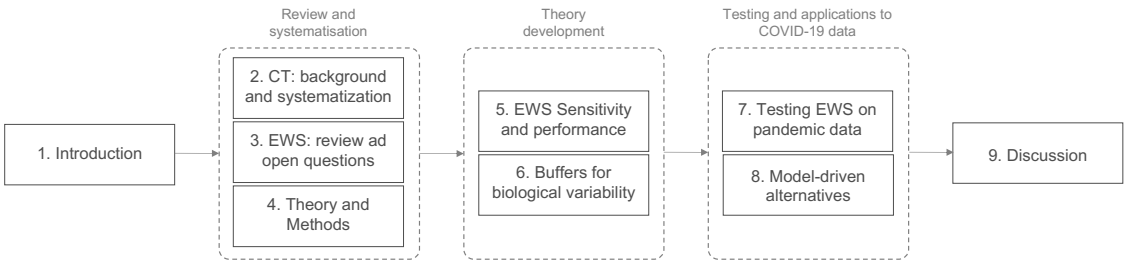


Fig. 1.3 Thesis structure and overview. After the introduction, three chapters are dedicated to reviewing and systematising literature results, open research questions, mathematical theory and methods for inquiry and analysis. Then, two other chapters present the contributions to theory development: studying the sensitivity and performance of EWS in various dynamical regimes; investigating how micro-biological cooperativity may buffer variability next to critical behaviour and how it modifies the trends of EWS. Two additional chapters address the use of EWS for epidemic surveillance: one tests theoretical predictions on COVID-19 data and studies their potential applications for epidemic alerting toolboxes; the other considers a model-based alternative for real-time monitoring, plus lessons learnt. The final discussion chapter includes ideas and potential future directions.

Introduction

- **Chapter 1** provides an introduction to methods and modelling vocabulary upon which this thesis is based, as well as an overview on defining characteristics of complex systems. The chapter aims at reflecting the interdisciplinary environment where the thesis project was developed, and to settle some common ground. It also contains the thesis overview and a list of contributions carried on during the PhD period.

Review and systematisation

- **Chapter 2** reviews the topic of critical transitions in dynamical systems, the underlying analysis framework and the main research topics. It provides an overview of the main results in the field and a systematic classification of known transitions. The chapter also overviews the connections between the CT framework, widespread concepts like system resilience, and other theoretical frameworks that had been investigating critical phenomena.

- **Chapter 3** investigates up-to-date results and open research questions regarding early warning signals, focusing on problems related to biological and epidemiological systems. To begin with, it inquires how to infer bistability properties from observed natural systems. Then, it reviews ongoing debates around critical phenomena, as well as challenges and guidelines to extract EWS from observational data. Finally, it connects the open research questions to the contributions presented in this thesis.

- **Chapter 4** puts together the mathematical theory underlying the study of critical transitions and the methods to investigate early warning signals. In addition, it groups methods for analysis of time series in search for EWS and towards risk assessment of systems' state, which will be employed in the subsequent analysis.

Theory development

- **Chapter 5** systematically studies the sensitivity of EWS to modelling assumptions and dynamical regimes, and their performance in those. Hence, the chapter assesses the observability of EWS in classes of systems and sets guidelines for their application, including optimal combinations of indicators to maximise the alerting performance. The study is initially carried out onto general models and then refined for a subset of biologically-relevant ones. Finally, the results are used to interpret experimental data.

- **Chapter 6** investigates how cooperativity mechanisms in cell regulation buffer variability next to critical behaviours, suggests a framework to extract resilience properties from data and assesses the performance of EWS when cooperativity modifies such resilience properties. It thus discusses potential applications and interpretations of such measures in biological studies. Its predictions are compared to experimental data.

Testing and applications on COVID-19 data

- **Chapter 7** tests and applies theoretical results to epidemiological data collected during the COVID-19 pandemic. It discusses practical challenges of pandemic monitoring and

describes how proposed EWS perform over worldwide collected data. The work provides an interpretation in light of CT theory and discusses several data analysis procedures to improve said performance.

- **Chapter 8** discusses a model-based approach for epidemic monitoring within a community. It introduces an extended Kalman Filter (EKF) to automatically and causally process SARS-CoV-2 RNA abundance in wastewater samples, which are complementary to detected positive COVID-19 cases. This work investigates the use of such mathematical tools to complement the pandemic surveillance toolbox, and overviews the lessons learnt towards real time monitoring of such a complex phenomenon. The goal is to provide insights for the future applications of real time monitoring.

Discussion

- **Chapter 9** concludes the thesis by summarising the current contributions and putting them in perspective with the open research questions outlined in early chapters. In addition, it suggests potentially interesting avenues that were not investigated during the project, but that will likely shape future research developments.

1.4.2 Contributions and dissemination

The present work resulted in a number of publications and contributions to scientific conferences.

Using the CT framework to inquire biological regulation patterns and their resilience was first discussed in the poster:

- D. Proverbio, J. Goncalves. *Critical transitions in systems biology - overview and perspectives*, 6th International Synthetic & Systems Biology Summer School – SSBSS 2019 Scuola Normale Superiore - Pisa, Italy, 2019,

where two preliminary case studies were introduced. One related to the autocatalytic feedback loop, deeper investigated in Chapter 6. The second case study was inspired by an early publication on swarming patterns (Proverbio et al., 2020) and, after an excursus on synchronization phenomena, is still open to further investigations.
The autocatalytic feedback loop case study, leveraged to analyse generic resilience properties, eventually resulted in a manuscript currently under review:

- D. Proverbio, A. N. Montanari, A. Skupin, J. Gonçalves, *Buffering variability in cell regulation motifs close to criticality* (under review). (Proverbio et al., 2022a)

The possibility of extending CT concepts to complex diseases (Bélair et al., 1995) was later discussed in the poster:

- D. Proverbio, J. Goncalves. *Dynamical modelling of complex diseases uncovers critical transitions in their phenotypes*. In Proceeding of INCOME - Integrative pathway modeling in systems biology and systems medicine, Berlin, Germany, 2019.

Advances on the systematic classification of critical transitions and on the investigation of the performance of their related EWS were presented as posters during two workshops:

- D. Proverbio. *Classification and detection of critical transitions*, 2nd CriTiCS Workshop - the challenges of multi-scale intergration in Biology, Belval, Luxembourg, 2019.
- D. Proverbio, J. Goncalves. *A systematic analysis of early warning signals to inquire the importance of the dynamic context*, Critics ITN Workshop, London, 2020.

The corollaries of such classification study in terms of distinction between smooth and abrupt transitions, and their relevance for the interpretability of diverse complex systems, were presented to conferences – as invited talk and poster – and as PsyArXiv working abstract:

- D. Proverbio. *Smooth or abrupt? How systems can change their state*. 5th Behaviour Change Science and Policy symposium: Complexity Science and Behaviour Change Interventions, Helsinki, 2020
- D. Proverbio, J. Goncalves. *Changing state: smooth or abrupt transitions?*, ENABLE Conference, Milan, 2021.
- M. T. J. Heino, D. Proverbio, K. Resnicow, G. Marchand, N. Hankonen. *Attractor landscapes: A unifying conceptual model for understanding behaviour change across scales of observation*. PsyArXiv, 2022 (Heino et al., 2022)

Systematising these results and the new outputs presented in Ch. 5 are expected to yield a new manuscript:

- D. Proverbio, A. Skupin, J. Gonçalves, *Systematic analysis and optimization of early warning signals for critical transitions* (in preparation).

Bridging critical transitions and observability properties from time-series data of complex systems was initiated in a manuscript, recently submitted:

- A. N. Montanari, L. Freitas, D. Proverbio, J. Gonçalves, *Functional observability and subspace reconstruction in nonlinear systems* (under review). (Montanari et al., 2022a)

Perspectives on linking critical transitions modelled on low-dimensional manifolds with network resilience were discussed as a talk:

- D. Proverbio. *Dimension reduction of complex networks to address critical transitions on low-dimensional manifolds*, BeNet conference, Namur, 2021.

The corpus on COVID-19 pandemic and its investigation as a critical transition is composed by talks, posters and recent publications, which will be further discussed in the corresponding chapters of this thesis:

- D.Proverbio, F. Kemp, S. Magni, and J. Goncalves. *Performance of early warning signals for disease emergence: a case study on COVID-19 data*. Plos Computational Biology, 18(3): e1009958, 2022. (Proverbio et al., 2022c)
- D.Proverbio, F.Kemp, S. Magni, L.Ogorzaly, H.M.Cauchie, J. Gonçalves, A.Skupin, and A. Aalto. *Model-based assessment of COVID-19 epidemic dynamics by wastewater analysis*. Science of the Total Environment, 827, p. 154235, 2022. (Proverbio et al., 2022b)

- F. Kemp, D. Proverbio, A. Aalto, L. Mombaerts, A.F. d'Herouel, A. Husch, C. Ley, J. Goncalves, A. Skupin, and S. Magni. *Modelling COVID-19 dynamics and potential for herd immunity by vaccination in Austria, Luxembourg and Sweden*. Journal of Theoretical Biology, 530, p. 110874, 2021. (Kemp et al., 2021)

- M. Burzyński, J. Machado, A. Aalto, M. Beine, J. Goncalves, T. Haas, F. Kemp, S. Magni,L. Mombaerts, P. Picard, D. Proverbio, A. Skupin, and F. Docquier. *COVID- 19 crisis management in Luxembourg: Insights from an epidemionomic approach*. Economics & Human Biology, 43, p. 101051, 2021. (Burzyński et al., 2021)

- D. Proverbio, F. Kemp, S. Magni, A. Husch, A. Aalto, L. Mombaerts, A. Skupin,J. Gonçalves, J. Ameijeiras-Alonso, and C. Ley. *Dynamical SPQEIR model assesses the effectiveness of non-pharmaceutical interventions against COVID-19 epidemic outbreaks*. Plos One, 16(5), p. e0252019, 2021. (Proverbio et al., 2021)

- F. Kemp, D. Proverbio, A. Aalto, C. Ley, J. Goncçalves, A. Skupin, S. Magni, *Dynamical modelling of COVID-19 pandemic*, 5th International Conference on Econometrics and Statistics (EcoSta 2022), Kyoto, 2022

- D. Proverbio, F. Kemp, S. Magni, J. Goncalves. *Evaluating the performance of EWS: a verification study on COVID-19 data*, Early Warning Signs for Abrupt Transitions workshop, Munich, 2021.

- D. Proverbio and F. Kemp. *Using mathematical models against epidemics*, Vrije Universiteit Brussel lectures, Bruxelles, 2020.

Extra "grey" literature pertaining COVID-19 studies and the activity within the Research Luxembourg COVID-19 Taskforce WP6 includes weekly reports to the local government (A. Aalto, S. Martina, D. Proverbio, F. Kemp, P. Wilmes, J. Goncalves, A. Skupin, *Update on the current situation in Luxembourg*, 2020-22), daily updates to the Ministry of Health, an in-house developed website for the estimation of $R(t)$[9] in Luxembourg (https://github.com/ResearchLuxembourg/restimator, together with R3 group at the LCSB) and dissemination articles for the general public:

- D. Proverbio, C. Ley, P. Mangers, *Coronavirus technical terms - explained by scientists from Luxembourg*, science.lu, 2020.

- D. Proverbio, *Getting back and preventing second waves: with which interventions?*, medium.com, 2020.

In general, mathematical modelling of complex systems can be applied to diverse case studies and contributes to our knowledge in various sub-fields. Hence, collaborations with colleagues and scientific partners led to additional publication less strictly related to CT. Those focusing to the robustness of multi agent systems are:

- J. Markdahl, D. Proverbio, L. Mi, and J. Gonçalves. *Almost global convergence to practical synchronization in the generalized Kuramoto model on networks over the n-sphere*. Communications Physics, 4(1), p. 1–9, 2021. (Markdahl et al., 2020)

[9]The time-dependent reproduction number, an important indicator that tracks the epidemic trend within a country. Further discussion is provided in later sections.

- J. Markdahl, D. Proverbio, and J. Goncalves. *Robust synchronization of heterogeneous robot swarms on the sphere*. In 2020 59th IEEE Conference on Decision and Control (CDC), p. 5798–5803, 2020. (Markdahl et al., 2021)

Towards the medical side of the institute's mission, contributions on the modelling of Deep Brain Stimulation[10] effects on neural tissue are covered in:

- M. Baniasadi, A. Husch, D. Proverbio, I. F. Arroteia, F. Hertel, J. Gonçalves, *Initialization of Deep Brain Stimulation Parameters with Multi-objective Optimization Using Imaging Data*, BVM conference, Heidelberg, 2022.

- M. Baniasadi, D. Proverbio, J. Goncalves, F. Hertel, and A. Husch. *Fastfield: An open-source toolbox for efficient approximation of deep brain stimulation electric fields*. NeuroImage, 223, p.117330, 2020. (Baniasadi et al., 2020)

- D. Proverbio and A. Husch. *Approxon: Heuristic approximation to the e-field-threshold for deep brain stimulation volume-of-tissue-activated estimation*. bioRxiv, 2019.

Finally, the successful participation to the *COVID-19 Lung CT Lesion Segmentation Challenge 2020*, powered by NIH, Nvidia and National Children's' Hospital, resulted in top 10 finalists worldwide and the output, together with the project group, of a working abstract:

- LCSBmedAI team: J. Sölter, D. Proverbio, M. Baniasadi, M. N. Bossa, V. Vlasov, B. G. Santa Cruz, and A. Husch, *Leveraging state-of-the-art architectures by enriching training information - a case study Contribution to the MICCAI COVID-19 Lung CT Lesion Segmentation Challenge*, orbilu.uni.lu, 2021.

[10]Deep Brain Stimulation (DBS) is a surgical treatment for several neurological disorders, including Essential Tremor (ET) and Parkinson's disease (PD). It aims at reducing the symptoms of the disease by stimulating specific brain regions with electrical impulses delivered by implanted electrodes. Modelling the effects of induced electric fields on the neural tissue requires balancing mean-field approximations and corrections for the tissue heterogeneity.

Chapter 2

Critical Transitions: a conceptual framework for critical phenomena

This chapter reviews and systematises the Critical Transitions (CT) framework. To begin with, it illustrates CT relevance for real-world problems and discusses motivations, aims and interconnections with other complex non-linear systems modelling paradigms. The chapter introduces the modelling background, classifies known transitions and puts them into perspective for theoretical and applied studies.

2.1 Systematic overview

Increasing evidence suggests that many real-world systems do not evolve smoothly and linearly, but rather exhibit rapid and hardly reversible shifts from one regime to another (Fig. 2.1). Examples are manifold and encompass numerous disciplines (Sornette, 2013; Scheffer, 2009; Feudel et al., 2018; Liu et al., 2021; Biggs et al., 2018). Non-smooth and irreversible changes were first observed in physical and chemical systems (Zeeman, 1979; Pathria et al., 2011). Ecosystems present several instances (May, 1977; Sheffer et al., 2001; Mills, 2004; Jiang et al., 2019), from abrupt eutrophication of shallow lakes (Carpenter et al., 1997; Wang et al., 2012a) to collapses of food webs (Van Nes et al., 2004; Paine, 1966; Sutherland, 1974) to desertification of grasslands and forests (Hirota et al., 2011; Demenocal et al., 2000) to bleaching of coral reefs (Wernberg et al., 2016; Petraitis et al., 2004). Numerous climate elements have been observed shifting or are suspected to (Ashwin et al., 2012; Thompson et al., 2011; Lenton et al., 2012), like the Atlantic meridional overturning circulation AMOC (Ritchie et al., 2019; Boers et al., 2021; Alkhayuon et al., 2019), the compost instability in peatlands (Wieczorek et al., 2011), and past climate changes (Livina et al., 2010); a list of highly sensitive mechanisms is routinely updated by the Intergovernmental Panel on Climate Change (Drijfhout et al., 2015). Social (Lade et al., 2013; Macy et al., 2021; Fraccascia et al., 2018) and financial systems also undergo abrupt changes like mini-crashes in stock markets (Diks et al., 2019; Dmitriev et al., 2017; Jurczyk et al., 2017; Kostanjcar et al., 2016). Psychology and behavioural sciences are active areas of research to identify impending mood or behavioural shifts (Stapelberg et al., 2018; Olthof et al., 2020; Foo et al., 2017) like sudden smoking cessation or depression (Tan et al., 2019; West et al., 2006; Van De Leemput et al., 2014). Biological systems are not exempt, and can undergo regime shifts

during cell fate decision (Mojtahedi et al., 2016; Ghaffarizadeh et al., 2014a), function regulation (Sharma et al., 2017; Hasty et al., 2000; Angeli et al., 2004a), or towards undesired disease states (Solé, 2003; Trefois et al., 2015; Cohen et al., 2022). So-called "dynamical diseases" have also been recognised (Mackey et al., 1987; Bélair et al., 1995), including sudden epileptic seizures (Meisel et al., 2012b; Jirsa, 2014), hearth fibrillation (Quail et al., 2015) and more. The rapid diffusion of epidemics within populations is another worrying phenomenon often exhibiting sudden outbreaks (Alonso et al., 2019; Brett et al., 2017; O'Regan et al., 2016; Horstmeyer et al., 2018; O'Regan et al., 2013). As many of these systems are interconnected, cascading regime shifts across scales are also feared and monitored (Dobson et al., 2007; Rocha et al., 2018; Klose et al., 2020).

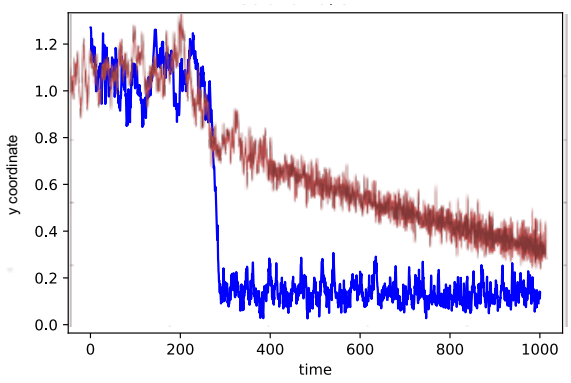


Fig. 2.1 A generic example of smooth (red) versus abrupt (blue) shift from one state (up) to another (down), as time progresses. Both simulations relate to stochastic normal forms (*cf.* Sec. A).

The common denominator among these systems – from small cells to people to societies to climate – is their complexity and non-linearity, yielding threshold points (often called "critical" or "tipping") at which abrupt regime shifts occur ("regimes" stand for dynamical equilibria). They share the following properties (adapted from Kuehn (2011)):

- A qualitative change in the dynamical equilibria;
- Such change occurs rapidly (abruptly) in comparison to the regular dynamics;
- The new state of the system is far away from its previous one (that is, it is clearly distinguishable by statistical analysis) and is stable.

Empirically speaking, neither of those three points can be *a priori* satisfied in any system. For instance, it is relatively easy to distinguish between cell phenotypes as different states; much more complicated is to tell between a "stable" and a "oscillatory" behaviour in a system whose "stable" state is already oscillatory, like for epileptic seizures. Similarly, assessing "rapidity" is a matter of comparison between scales, not an objective and universally recognized feature. Separating two equilibria with high confidence is also often hindered by data scarcity and measurements errors. Finally, the new state of the system should be at equilibrium; however, for excitatory phenomena (Wieczorek et al., 2011), sudden spiking activities are transient if compared with the global dynamics. Financial crashes also belong to the latter category.

Identifying, classifying, characterising and analysing such phenomena is thus a matter of scientific investigation and modelling (see also Sec. 3.1), which also determines the vocabulary used. "Regime shift" refers to any abrupt phenomenon observed empirically. If such phenomenon can be usefully modelled by a set of dynamic variables, whose mutual relations are defined by a set of parameters, points marking discontinuities and changes of equilibria are termed "tipping". Among the various classes of tipping mechanisms (see below), those emerging from discontinuous noisy bifurcations (Sec. 2.2) are termed "critical[1]", and their vicinity is often called "critical regime" or "criticality". A subset of those bifurcations, defined by mathematical properties as well as by their significant effects in the corresponding real-world systems, are sometimes called "catastrophic" (Thompson et al., 2011). A so-called critical transition (Scheffer, 2009) is thus the outcome of certain dynamical models. In this sense, the Critical Transitions (CT) framework is an umbrella for the study of critical events, linked to real-world phenomena by adequate modelling choices. It is a way of modelling systems, not a class of systems by itself. Its connections with other theories investigating critical phenomena is further investigated in Sec. 2.5.

A classical way of modelling systems' evolutions over time, possibly characterised by abrupt shifts, is by the theory of dynamical systems (Sec.2.2.1 and Katok et al. (1997); Wiggins et al. (1991)). It considers a system's state space – a manifold defined by all potential values of its defining variables, whose inter-relationships are given by parameters – and laws of evolution. Modelling choices determine what is aptly modelled as variable (usually, the degrees of freedom) or parameter (usually, what tunes the inter-relating functions). Often, complex dynamical systems do not visit the full state space with equal probability, but tend to reside in attractors, corresponding to more constrained states of stable equilibria (see Sec. 2.2.2). Their "depth", associated with properties of the defining dynamics, quantifies their intrinsic stability (La Salle, 1976). Depending on parameters' values, attractors can morph, disappear or collide. In addition, multiple attractors might populate the state space simultaneously. Dynamical manifolds are often rugged and ever-changing, like soft and morphing terrains, see Fig. 2.2 for an illustrative example. Considering evolving manifolds under the action of dynamical parameters could open new avenues for modelling and represents a change in perspective from the standard vision of static manifolds. A brief discussion, including the relevance for biological systems theory, is provided in Sec. 2.6.

On top of the deterministic laws of change, stochastic noise of various nature can be present, either due to random and/or thermal processes, or to lumping of fast degrees of freedom (Berglund et al., 2006; Freidlin et al., 1979). Attractors' stability is normally defined in relation to small to negligible noise intensity. Should the noise intensity be non-small, a system could be able to visit wider areas of the state-space and neglect the influence of shallow attractors (the system becomes ergodic in a wider confined sub-space, example in Fig. 2.2, left). "Mild" noise intensities, in turn, might cause the system to flicker between nearby attractors (Dakos et al., 2013). Inter-plays of scales thus matter to identify reasonable state spaces and adequate modelling, and to study systems' stability and potential critical transitions (Shi et al., 2016). Further discussion, using the mathematical concept introduced in Sec. 2.2, is provided in Sec. 2.4.

[1] "Critical" comes from the greek "krisis" ($\kappa\rho\iota\sigma\iota\zeta$) which in turn derives from the verb KRINO, "to separate, to discern". Thus, "crisis" refers to a moment (pointwise) that separates two different regimes; a turning point. In this context, it does not necessarily refer to a "state of emergency" nor to the medical meaning of "dramatic change".

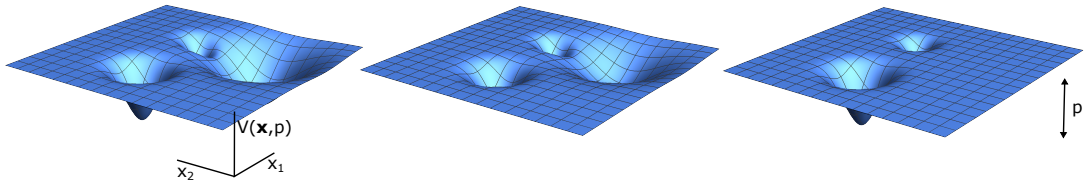


Fig. 2.2 Illustrative example of state space, defined by two variables, with multiple attractors (basins). This representation associated the z-axis with the potential (first integral of the laws of motion), when it is well-defined. Depending on the inter-actions of parameters (here, a single one p), the manifold can morph, with attractors becoming e.g. shallower (**centre**) or disappearing (**right**). Attractors are defined with respect to noise levels. Consider the **left** panel: if noise intensity is higher than the separatrix between the two attractors on the right side, they might be counted as one, since the system can visit both indefinitely. If noise intensity is higher than any separatrix, the system becomes fully ergodic and can visit the entire state space with equal probability. This conceptual figure is for illustrative purposes. Formal analysis is given in sections below and in Appendix A.

Systems' stability among multiple attractors in noisy environments is dual to the concept of resilience (Fisher, 2015; Liu et al., 2020a; Fraccascia et al., 2018; Bhamra et al., 2011), *i.e.*, the ability to withstand adversities and return to original states (also known as engineering resilience (Holling, 1996)), or not tip to a new regime (also known as ecological resilience (Dai et al., 2015)). In fact, "deeper" and "steeper" attractors make it easier for a system to cope with perturbations, while "shallower" and "gentler" ones reflect augmented fragility of the system state (Beisner et al., 2003). In turn, these concepts are closely related to robustness, i.e. system's ability to maintain its function (Kitano, 2004), a term often used in biology (Tsimring, 2014) and network science (Tejedor et al., 2017) – although, in the latter case, it may be linked to node removal and structural stability rather than to dynamics. These terms – stability, resilience, robustness – and their antonyms – instability, criticality, fragility – relate to the possibility that a system undergoes an abrupt regime shift due to combinations of control actions and perturbations. For the purposes of this thesis, they will be used interchangeably, with a slight abuse of vocabulary. Nonetheless, they correspond to quantitative nuances related to measurable properties of dynamical systems. Their exact correlation might not be established when multiple environmental drivers are changed simultaneously (Dai et al., 2015), so future works might consider exact terminologies. A more detailed description is provided later in the text, after introducing the mathematical tools that are useful to characterize them fully.

Relevant classes of critical transitions in open systems are closely related to robustness and loss thereof, in combination with different scales of noise intensity (see Sec. 2.4 and Sec. 2.3.2). They derive from systems becoming increasingly fragile and sensitive to perturbations, *e.g.* because of dis-balanced regulatory feedback mechanisms, variations of external controls or combinations thereof. In case multiple attractors co-exist, fragile systems can then tip onto alternative ones. Depending on the interplays between noise and attractors' properties, several classes of critical transitions can be recognised (Ashwin et al., 2012), *e.g.* those primarily driven from stochastic switching or those primarily characterised by bifurcation points. Nonetheless, as introduced in Sec. 1.3, validated mechanistic models are seldom available for many complex systems, including the climate, financial, biological ... ones mentioned above. The main endeavour of the CT framework is thus to propose, study and verify sets of mathematical indicators to alert against impending

regime shifts. Such indicators should be as generic as possible, with respect to modelling choices, to be applied to classes of critical transitions rather than being system-specific. They should therefore be based upon common dynamical mechanisms and stability properties, and tested against empirical data to verify their interpretability and adequateness.

When first proposed, these indicators and their relevant temporal patterns were termed "early warning signals" (EWS, see Scheffer et al., 2009; Drake et al., 2010). However, scholars soon realised the implicit determinism suggested by the terminology (Scheffer et al., 2015). "Early warning" could imply that the system is bound to keep becoming increasingly fragile and to eventually tip. Instead, the driving mechanisms can reverse and robustness be restored, or unexpected large deviations might drive the system onto alternative states earlier that potentially expected (Haragus et al., 2018). It was thus suggested (Scheffer et al., 2015; Dai et al., 2015) to employ the term "resilience indicators", arguably more precise in highlighting that the most used indicators are proxies of systems' resilience. This terminology would also remind that proposed indicators are risk-assessment rather than predictive tools, because of the strong relevance of noise (Zhang et al., 2015). Nonetheless, the name "early warning signals" – with its EWS abbreviation – proved to be good for communication and marketing, with the additional advantage of encompassing other systems-specific measurements as well (Clements et al., 2016a; Clements et al., 2018). It thus remained in the literature. Bearing those caveats in mind, the term "early warning signals" will also be used in the present thesis.

Over the years, several EWS have been proposed. A common denominator is their relationship with the Critical Slowing Down (CSD, *cf.* Sec.2.4.1) phenomenon, *i.e.* diverging response time in non-linear systems approaching a bifurcation (Scholz et al., 1987; Byrd et al., 2019; Scheffer et al., 2009; Dakos et al., 2008). This means that, close to criticality, the system takes increasingly longer times to buffer a perturbation and return to its original stable state, which also implies higher variability. Increasing variance, autocorrelation, skewness or recovery times (Dakos et al., 2015; Guttal et al., 2008; Lenton et al., 2012; Lade et al., 2012; Diks et al., 2019; Olde Rikkert et al., 2016; Dai et al., 2013), spectral reddening (Kleinen et al., 2003; Bury et al., 2020; Carpenter et al., 2011; Biggs et al., 2009), entropy-based (Lee et al., 2019; Brett et al., 2017) as well as measures for leading eigenvalues and combinations (Drake et al., 2010; Bury et al., 2021; Brett et al., 2020a; Williamson et al., 2016; Veraart et al., 2012) have thus been proposed and applied, with various degrees of success, to diverse systems. Application domains range from technology (Cotilla-Sanchez et al., 2012; Ren et al., 2015; Gopalakrishnan et al., 2016; Pavithran et al., 2021), to psychology (Dablander et al., 2021; Wichers et al., 2016; Heino et al., 2022), to epidemiology (Southall et al., 2021), ecology (Dakos et al., 2012a), climate science (Boers, 2021), biology (Navid Moghadam et al., 2020) etc. In addition to CSD-based indicators, other suggestions to target different critical transition classes include *ad hoc* metrics and flickering for primarily noise-induced shifts (Shi et al., 2016; Wang et al., 2012a; Drake, 2013; Olthof et al., 2020; Dakos et al., 2013; Xie et al., 2018). Other indicators have also been proposed specifically for networks (Aguirre et al., 2015; Chen et al., 2012b; Chen et al., 2019; Aihara et al., 2022; Mojtahedi et al., 2016; Matsumori et al., 2019; Kuehn et al., 2015; Loppini et al., 2019; Yang et al., 2018) and for spacially extended systems (Gowda et al., 2015; Kuehn et al., 2019; Rodríguez-Méndez et al., 2016; Chu et al., 2017).

Since this thesis is primarily focused on developing, interpreting and testing EWS for biological and epidemiological systems, after this general introduction and EWS derivation from

dynamical models in Sec. 2.4.2, a full chapter (Ch. 3) is dedicated to their current developments, open questions, promises and limitations.

2.2 Modelling background

One of the purposes of this thesis is interpreting the mathematical modelling of dynamical systems as a meaningful representation of observable phenomena, and to study their assumptions and properties. Consequently, this section provides a selected overview of mathematical constructs and concepts that will be used in subsequent analysis. In every subsection, a high-level overview is presented, along with its relevance for the framework of critical transitions. Formal definitions and theorems are instead included in the appendices. The introduction and disclosure of formal concepts will hopefully foster connections between mathematicians, modellers and experimentalists.

2.2.1 Modelling dynamical systems

The notion of dynamical systems mathematically formalizes the general concept of deterministic processes that evolve over time. The core motivation to develop such models is the possibility to determine the future state of a system given its past and some law of evolution. Appendix Sec. A details formal definitions and theorems.

In this perspective, a modeller first identifies a state (or phase) space Ω that embeds all possible states u in which the system can evolve, in terms of generalised coordinates $\mathbf{x}$. The generalised coordinates are an important modelling choice: in physics, position and momenta of moving bodies have been established as the smallest set of coordinates to describe motion, because of symmetry and conservation arguments. For other systems, different coordinates or combinations thereof might be better suited. To describe a cell's state, gene expression levels are often employed to construct the phase space (Aalto et al., 2020). It is important that a metrics can be expressed to define distances between states, in terms of the coordinates. Using terminology from statistical mechanics (Pathria et al., 2011), a specific configuration of state-space coordinates is called a micro-state (see Sec. A). Its corresponding macroscopic (possibly emerging) properties are termed "macro-state". For a living organism, a microstate could correspond to its genotype and a macrostate to its phenotype or its function (Kitano, 2002).

Another important ingredient to develop a dynamical system model is its law of evolution over time. It is expressed as a trajectory ϕ^t exploring the subset of Ω that is visited by the system at each time t. Reconstructing such law of evolution is among the primary scopes of mathematical sciences, as it expresses in a quantitative manner the system past, present and future. ϕ^t is usually expressed and analysed in terms of ordinary differential equations ODE, see Th. A.1.1. Should the system's evolution depend explicitly not only on time, but also on its state x, the law of evolution is better described as a partial differential equation PDE. A typical form for an ODE, in terms of variables $\mathbf{x}$ (identifying a state u) and parameters $\mathbf{p}$ (quantifying functional relationships between variables), is given by:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(\mathbf{t}), \mathbf{p}) \cdot A\,, \tag{2.1}$$

where $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is the system of (sufficiently smooth) functions governing the changes of each component of the variable vector $\mathbf{x}$ and A can be interpreted as an adjacent matrix reminding that elements of $\mathbf{x}$ can depend on one another. If $\mathbf{p}$ do not depend explicitly on time, the system is termed *autonomous*; if $\mathbf{p} = \mathbf{p}(t)$, the system is called *non-autonomous*. In the rest of the thesis, unless otherwise specified, autonomous systems will be mostly considered. Also, since the rest of the thesis will mostly address low-dimensional systems (cf. Sec. 2.3.2), a notation without the bold vector form will be employed.

To solve the dynamical system and reconstruct the full evolution, initial conditions (t_0, u_0) or boundary conditions $\phi(\mathscr{B})$ should be specified ($\mathscr{B}$ represents the boundary of the domain Ω). Boundary conditions can be expressed on the specific values of the solution (Dirichlet conditions) or on its first derivative (Neumann conditions). Combinations of the two make Cauchy boundary conditions. Finding the function that satisfied the differential equations and the constraints posed by initial or boundary value problems solves a Cauchy problem. A Cauchy problem for Eq. 2.1 is, in terms of initial values:

$$\begin{cases} \frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(\mathbf{t}), \mathbf{p}) \cdot A \\ \mathbf{x}(t_0) = \mathbf{x_0} \ . \end{cases} \tag{2.2}$$

Stating initial conditions is particularly important in case of multiple attractors (see later). In fact, they would determine the basin of attraction and the (possible) hysteresis properties.

Example: for a cell, its state can be defined by its enzymatic activity, quantified by the concentration $\mathbf{x}$ of relevant enzymes (in molar amounts). The initial condition u_0 is the first value measured. The set Ω is characterized by all possible values of enzyme concentrations; non-biological values are excluded and therefore constitute the exterior of the bounded set Ω. Then, $\mathbf{f}$ is the function describing the evolution of enzyme concentration over time, given different production or degradation rates expressed as parameters. Usually, $\mathbf{f}$ can be deduced from enzyme kinetics rules.

Dynamical models are of primary importance to investigate complex systems as they:

1. Suggest new experiments and guide them (hypothesis generation)
2. Test if all necessary components to explain a phenomenon are known (knowledge checking)
3. Combine mechanisms that are tested in reductionist experiments (synthesis)
4. Allow quantitative predictions for known systems during certain regimes (prediction)
5. Estimate qualitative features like stability and robustness (behaviour prediction)
6. Enable investigating counter-factual ("what if") analysis and scenarios.

2.2.2 Equilibria and stability

This section summarises the main concepts of equilibria and stability used for the study of critical transitions. The main motivation behind the analysis of equilibria and stability is to understand and predict the overall behaviour of a system, instead of looking for precise values. Questions like "is my system going to change infinitely or will it settle down to some limit value?", "has it reached a resting state?", "what about its oscillations?" are thus primarily addressed.

A dynamical system whose variable values do not vary any more in the phase space is said to have reached a steady-state equilibrium (formal definitions in Sec. A). Steady-state solutions might be reached after some transient time. Fixed points and sustained oscillations are examples of equilibria. The equilibria of a system can be usually[2] characterised by the properties of its Jacobian J s.t. $J_{ij} = \frac{\partial f_i}{\partial x_j}$: if its eigenvalues λ_i are all real and different from zero, then the equilibrium is called a hyperbolic fixed point. In addition, the sign of the eigenvalues determines the equilibrium stability. An equilibrium is said to be stable if $\lambda_i < 0 \; \forall i$ and all sufficiently small disturbances away from it damp down in time (see Th. A.4.2). A $\lambda_i > 0$ determines an unstable direction, where small perturbations get amplified and the trajectory diverges from the equilibrium. Imaginary eigenvalues characterise spiralling behaviours. The region around a stable equilibrium, defined by all points from which a trajectory converges to the equilibrium, is called an attractor.

Following the linear stability analysis presented above, one can extract information about stability and characteristics of its fixed ODE points by considering determinant, discriminant and trace of the system's Jacobian J (Strogatz, 2018). This is due to the relationship between matrix properties and its eigenvalues λ_i. The discriminant of the characteristic polynomial identifies the reality or complexity of the eigenvalues; trace $TrJ = \sum_{i=1}^{n} \lambda_i$; determinant $DetJ = \prod_{i=1}^{n} \lambda_i$. Consequently:

- If the discriminant of the characteristic polynomial (for second order polynomials, $\Delta = Tr^2J - 4DetJ$) is $\Delta > 0$, then $\lambda_i \in \mathbb{R}$ and fixed points are hyperbolic points (trajectories converge or diverge).
- If $\Delta < 0$, then $\lambda \in \mathbb{C}$ and fixed points are nodes (trajectories revolve around them).
- $DetJ < 0$ characterises a saddle point.
- If $DetJ > 0$ there are different possibilities , according to the trace:
  - Unstable point if $TrJ > 0$;
  - Stable point if $TrJ < 0$;
  - Hopf point (to periodic orbits) if $TrJ = 0$.

A graphical representation is provided in Fig. 2.3 with a Poincaré Diagram.

For 1-D systems $\dot{x} = f(x, p)$, the picture is even simpler. Suppose $f(x, p)$ is a continuously differentiable function. Suppose $\hat{x}(p)$ is an equilibrium point. Then, if $f'(\hat{x}) < 0$, then $\hat{x}$ is stable; if $f'(\hat{x}) > 0$, then $\hat{x}$ is unstable; if $f'(\hat{x}) = 0$, then it is necessary to assess its stability with other methods, like Lyapunov's functions. Graphically, it is immediate to check when $f(x, p)$ crosses the zero axis: these are fixed points. Moreover, if $f(x, p)$ goes up, the point is unstable; if it goes down, it is stable.

Here lies the power of a mathematical formulation: to experimentally verify that a real system is stable, one should in principle wait an infinite amount of time to check every response to slight perturbations. On the other hand, thanks to analytical manipulation of an ODE model, stability is easily checked as a global prediction for the considered system.

[2] There are cases in which a linear analysis is not sufficient and it is necessary to study system-specific associated Lyapunov functions.

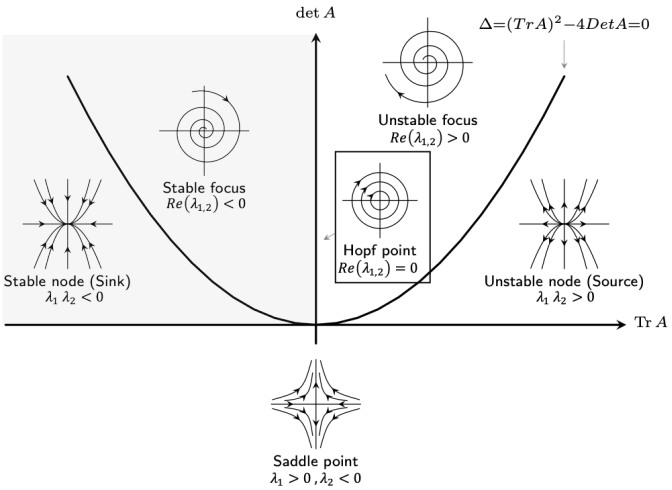


Fig. 2.3 Representing the information about the linear stability of the Jacobian in the $(DetA, TrA)$ plane is know as Poincaré diagram. In this case, example eigenvalues for a 2D system are also shown. Shaded region corresponds to stable fixed points. The figure is based upon Gernot Salzer, "Poincaré diagram: Classification of phase portaits in the $(DetA, Tra)$-plane", March 20, 2018, https://texample.net//tikz/examples/poincare/.

2.2.3 Bifurcations

Notably, hyperbolic points are structurally stable (robust): small perturbations may distort but do not change qualitatively the phase portrait near the equilibrium. Instead, non-hyperbolic points (with at least one eigenvalue with zero real part) are not robust: small perturbations can result in bifurcation of the equilibrium, marking distinct qualitative changes in the dynamics like disappearance of the equilibria, emerging instability etc. It is thus of great interest to study what happens when the leading eigenvalue(s) (the biggest negative λ_i) of a stable system approaches and reaches zero – that is, when a bifurcation occurs. Formal notions for the study of bifurcations in dynamical systems are presented in Sec. A. Here, some notable properties are highlighted.

Bifurcations can be local or global. In this thesis, local bifurcations will be considered. Consequently, recall that studies on critical transitions usually do not consider the full manifold spanned by dynamical systems, but concentrate on what happens in low-dimensional sub-regions characterised by the bifurcations of interest (Th. A.5.1 and Guckenheimer et al. (2013)). Since the bifurcation is driven by the leading eigenvalue(s), it usually occurs on low co-dimensions (the minimal number of parameters that elicit a bifurcation). In practice, the approach to a bifurcation can usually be described by a small number of control parameters or leading combinations of systems' parameters (Kuehn et al., 2021). By means of mathematical analysis, it is then possible to assess which is the minimal combination of controls driving such changes.

Another interesting property of local bifurcations is the existence of normal forms. Normal forms are simplified forms (often polynomial) that are topologically equivalent to the bifurcation form and therefore retain the same qualitative information about equilibria and stability (Def. A.5.7). In practice, normal forms are one-to-one associated to the main bifurcations and are used to study relevant features and associated metrics, since they are analytically tractable and general. Section A provides an overview of the main bifurcation normal forms analysed in this thesis, along with their most relevant properties.

Among bifurcation normal forms, some are particularly robust against small oddities and disturbances. They are called "generic" against unfolding (Golubitsky et al., 2012; Kuznetsov, 2013); more details in Sec. B). This property has two consequences. First, that the dynamical drivers of many systems can be understood in terms of local bifurcations and their unfolding, therefore fostering unifying modelling across disciplines and scales of observation. Second, that studying low-dimension and low-codimension bifurcations provide key insights on the qualitative behaviour of a vast array of complex systems. In fact, given two or more dynamical systems with a particular behaviour, they are topologically equivalent if they share similar equilibria and stability properties. In this case, they can be reduced to a subset of simple forms, used to predict e.g. statistical features. Hence, bifurcation studies are powerful mathematical tools to address different systems under the same lenses. Further quantitative refinements are subsequently needed to better identify and characterize specific systems and to connect them to experiments and observations.

A visual example of topological equivalence

When addressing research questions concerning the stability and resilience of dynamical systems, bifurcation and normal forms provide a unifying modelling tool (Kuehn et al., 2021). Here, one such example is provided. Consider two systems from diverse disciplines: a model (Scheffer, 2009) for harvested populations from ecology (further described along with Eq. A.9),

$$\frac{dX}{dt} = X\left(1-\frac{X}{K}\right) - c'\frac{X^2}{X^2+1}\,. \tag{2.3}$$

and a model for enzymatic activity from systems biology (Frigola et al., 2012; Strogatz, 2018), further described in Sec. 4.1.1:

$$\frac{dx}{dt} = K + \frac{cx^2}{1+x^2} - x\,. \tag{2.4}$$

Their global behaviour differ, but locally, next to a bifurcation point, they are topologically equivalent to a saddle-node normal form (Sec. A). One can visualize it in Fig. 2.4: the vector fields of the two realistic models (left and centre), when the leading parameter is changed, locally cross the x-axis in the same way as the fold normal form (right).

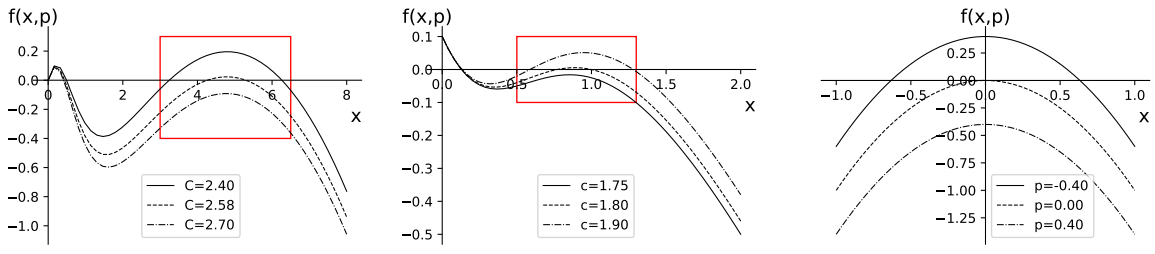


Fig. 2.4 Visual example of topological equivalence. Left: plot for $f(x,c')$ of the harvested population model Eq. 2.3. Centre: plot for $f(x,c)$ of the protein production model Eq. 2.4. Right: plot for $f(x,p)$ of the saddle node (fold) normal form $\dot{x} = -p - x^2$ (Sec. A). The red rectangle highlights the local region where the two realistic models are equivalent to the normal form, as they approach a bifurcation point ($f(x,p)$ crosses the x-axis) under the change of the parameter c' or c.

Bifurcation diagrams

A convenient way to represent bifurcations – especially when 1-dimensional and 1-co-dimensional – are bifurcation diagrams. They are formally defined in Sec. A and briefly presented here, along with their salient features and interpretation.

Consider the leading ("control", Strogatz (2018)) parameter driving a bifurcation, p, and the system equilibrium, $\tilde{x}$. Plot $\tilde{x} = \tilde{x}(p)$, that is, the set of values taken by $\tilde{x}$ for each value of p. Then, use a conventional visual identifier for stable and unstable branches. Common examples are solid or blue lines for stable manifolds and dashed or red lines for unstable ones. Critical points might be marked separately. Fig. 2.5 provides an example of diagram for the normal form of a saddle-node bifurcation: $\dot{x} = p + x^2$ (see Sec. A for more details).

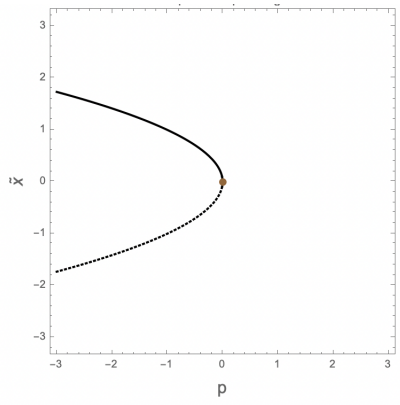


Fig. 2.5 Example of bifurcation diagram $(\tilde{x}, p)$ for a fold normal form. Solid line identifies the stable manifold, dashed line the unstable. The saddle-node point is marked in brown.

2.2.4 Slow-fast systems

As introduced in Sec. 1.3, complex dynamical systems often evolve across various time-scales: slow, medium and fast. Some processes are relatively faster than others, which unfold over longer time frames. When the time scale separation is significant, one can focus on fast degrees of freedom and approximate the slow ones as happening at quasi-steady states (that is, they are almost constant in time). Alternatively, one can concentrate on slower degrees of freedom while setting the faster ones at stationary values (they exhausted their transient trajectories and settled to equilibrium values already). Mathematically, these approximations are formalised by slow-fast systems models (*cf.* Sec. A and Kuehn (2011); Nyman et al. (2020) for details). A slow-fast system is expressed as:

$$\begin{aligned} \varepsilon \frac{dx}{dt} &= f(x,y) \\ \frac{dy}{dt} &= g(x,y)\,, \end{aligned} \tag{2.5}$$

where x contains fast degrees of freedom and y the slow ones. $0 < \varepsilon \ll 1$. A fast variable x can nonetheless be slower than another process variable z and so on, establishing a time-scale hierarchy. Often, variables are interpreted as fast degrees of freedom, $\dot{x} = f(x(t), p)$ while parameters are the slow ones, assuming quasi-steady state values $\dot{y} \equiv \dot{p} = 0$. The slow-fast systems interpretation

can thus be used to understand specific classes of critical transitions (Kuehn et al., 2011; Kuehn, 2013). In addition, this interpretation connects slow-fast systems to the concept of autonomous and non-autonomous systems (see above), so that extensions can be developed (O'Keeffe et al., 2020; Ashwin et al., 2017b).

Biologists, facing the enormous complexity of life, have tended to consider at least three time scales separately: physiology, development and evolution (Saunders, 2018). The Waddington landscape (Sec. 2.6) addressed some notable properties of development, interacting with the short-term physiological processes and affect the long-term evolution of organisms. Within physiological processes, additional time-scale hierarchies can be found, for instance in cell regulation: proteins have a "slow" mean life, whereas transcription happens quicker. Hence, for modelling purposes, it is costume to consider "fast" processes as "almost done, at quasi-steady state" and to concentrate on "slow" processes.

In epidemiology, the typical time scale for epidemic diffusion is dictated by infectivity rates and recovery/death periods. However, slowly changing behavioural habits, mutating viral variants or other long-term trends might be accounted for by considering quasi-steady state variations of model parameters. Sometimes, as observed during the recent COVID-19 pandemic, such changes happen instead on commensurable time scales, breaking down the slow-fast system approximation (see Chapter 7 for additional details and the effect on early warning signals).

As the slow-fast is a modelling approximation, other studies extend these simplifications to account for the role of time delays and finite-speed processes in the emergence of new patterns, e.g. in neural firing (Izhikevich, 2007).

2.2.5 The "pebble-down-the-hill" analogy

A first approximation neglecting inertia views dynamical systems as damped particles moving down a potential. In fact, given a law of motion in Lagrangian terms, we can obtain an equation in form of A.1 by setting to zero the second-derivate. Therefore, it is intuitive to represent a dynamical system as a pebble rolling down a hill while being immersed in a viscous medium that prevents significant acceleration. The hill represents the potential term that, for a scalar model such as A.3, is defined as $V(x)$ s.t $f(x) = -\frac{\partial V(x)}{\partial x}$. If there is noise acting onto the system, the ball might also wiggle and be pushed to perturbed directions. This illustration is often called "potential landscape". Adjoint potential landscapes can be formally calculated for noisy dynamical systems, therefore linking illustrations with actual models, see Sec. B.

The metaphor value of the "potential landscape" representation is that, it is possible to visualize the behaviour of a system as in Fig. 2.6. Each valley is associated with an attractor. Its depth and width determine the system stability and resilience properties (Dai et al., 2015): not only deeper valleys are more difficult to leave due to random fluctuations, but they also yield faster rates of convergence to the original equilibrium. Asymmetric potentials immediately relate to hysteresis: "climbing" a barrier on one sense may be easier than on the other sense. A slowly changing parameter morphs the potential's shape and depth, which become dynamic. Higher rates of parameter changes (no quasi-steady state assumption) might modify the associated statistical properties because the potential would morph faster than the system is able to settle back to the equilibrium and is therefore constantly "jiggled" around.

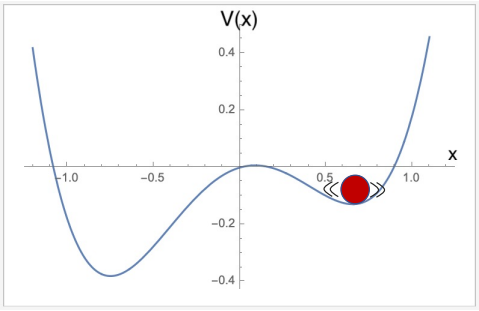


Fig. 2.6 The potential associated to a dynamical system in the form of $\dot{x} = a + bx - x^3$, where a and b are fixed values. The system state is represented by the red ball. Valleys represent attractors. In this case, if the ball rolls down it will stop at the attractor without oscillating, as it is overdamped (the associated system is said to be *dissipative*).

Note that not all systems are associated with a potential (depending to their analytical properties or their dimension) and that overdamped particles are a subset of a broader range of systems subjected to inertial effects and no dissipation (Ott, 2002). Hence, one should regard the "pebble-down-the-hill" as an analogy for certain classes of systems, not as a complete model.

2.3 Critical transitions in noisy dynamical system: systematic classification

Noise is ubiquitous in real-world systems. In addition to purely random processes like thermal fluctuations or nuclear decay, many other phenomena and fast degrees of freedom can be *modelled as* random processes, neglecting certain specific details but retaining key statistical properties (Namachchivaya et al., 1990). Noise in biological systems is extremely common (Su et al., 2019; Hasty et al., 2000; Tsimring, 2014), as well as in epidemiological processes, where it lumps behaviours, viral dynamics, transmission rates and so on (O'Regan et al., 2018; Brett et al., 2017). Formal modelling is achieved by the theory of stochastic processes (Gardiner, 1985; Papoulis et al., 2002) and its extensions to large deviations (Freidlin et al., 1998). A formal introduction is provided in Sec. A.

Combining noisy processes to dynamic bifurcations can be formalised as a multiple timescale slow-fast system, where $x \in \mathbb{R}^n$ are state variables, $p \in \mathbb{R}^m$ are system parameters, $\xi \in \mathbb{R}^l$ are stochastic variables (noise), and τ_i, $i = \{x, p, \xi\}$ are relative time-scales. Blending the framing introduced by Berglund et al. (2006) and Kuehn et al. (2011) with that of Shi et al. (2016) and Thompson et al. (2011), one can write the three-timescale systems as:

$$\begin{cases} \tau_x \frac{dx}{dt} = f(x, p, \xi) \\ \tau_p \frac{dp}{dt} = g(x, p, \xi) \\ \tau_\xi \frac{d\xi}{dt} = h(x, p, \xi) \end{cases} \tag{2.6}$$

Using this representation, systems can be classified according to their dimension n and codimension m, and framed into three main classes of critical transitions on the basis of relative

timescales (Ashwin et al., 2012):

$$\begin{aligned} &\text{B-tipping: } \tau_p \gg \tau_x \gg \tau_\xi \\ &\text{N-tipping: } \tau_p \gg \tau_x \simeq \tau_\xi \\ &\text{R-tipping: } \tau_p \simeq \tau_x \gg \tau_\xi \end{aligned} \tag{2.7}$$

On top of that, the timescale formalism allows to recognise when the system becomes ergodic and visits the full state-space uniformly without displaying transitions: $\tau_\xi < \tau_x$ (Shi et al., 2016).

For the analysis of such systems, stochastic differential equation (SDE) models or other formulations from the theory of stochastic processes are usually employed, see Sec. A. SDEs are usually written in the form:

$$dx = a(x)\,dt + b(x)\,dW\ . \tag{2.8}$$

Here, $a(x)$ is a deterministic drift term, $b(x)$ the diffusion function and dW the derivative of a Wiener process.

2.3.1 Considerations on dimensions and co-dimensions

Many real systems are usually multi-dimensional (Boccaletti et al., 2006), so loss of resilience and critical transitions sometimes need to be characterised considering the full set of variables and parameters, particularly when closely entangled. However, in many cases, low dimensional and low co-dimensional models, for which analytical solutions are often more accessible, can be aptly used to describe a vast array of situations.

A rather simple case happen when only one (or very few) observables are available, which are interested by critical transitions. Examples include epidemic outbreaks, measured in terms of new infections, cells states or phenotypes (e.g. ”cancerous” or healthy), or epileptic brain activity, assessed by unimodal EEG measurements.

Another solution is to lift (or embed) the nonlinear dynamics into a higher dimensional space where its evolution is approximately linear (Korda et al., 2018). The lifting can be achieved by using suitable operators or by changing the focus from “dynamics of states” to the “dynamics of observables” (Budišić et al., 2012).

An additional case considers the distinction between “microstates” and “macrostates” employed in statistical mechanics (see Sec. A and an illustration in Fig. 2.7). Macrostates are emergent properties (described by reduced amount of variables) from the activity of micro-constituents, like temperature emerging from atomic configurations and motion, or cells’ phenotypes emerging from interactions of genes in regulatory networks. Determining and analysing critical transitions in macrostates, given measures of their microstates, is subject to ongoing research (Pathria et al., 2011); for instance, several techniques for network systems are introduced in Sec. 4.7.

The final case distinguishes between local and global dynamics and makes use of the center manifold theorem (Th. A.5.1). In short, it states that complex manifolds can be, next to critical points, topologically equivalent to bifurcation forms. It is a local property (thus not valid in the whole state space) and it relates to topological and stability properties rather than to quantitative mechanisms; it nonetheless allows studying salient features of high-dimensional systems using low-dimensional bifurcations. As an example, Jirsa (2014) made use of this theorem to construct a phenomenological model to reproduce epileptic-like signals, without having to consider the

incredibly varied brain structure. Kuehn et al. (2021) further investigated the relevance of using center manifold theorem to study critical phenomena and to extract generic signals for CTs.

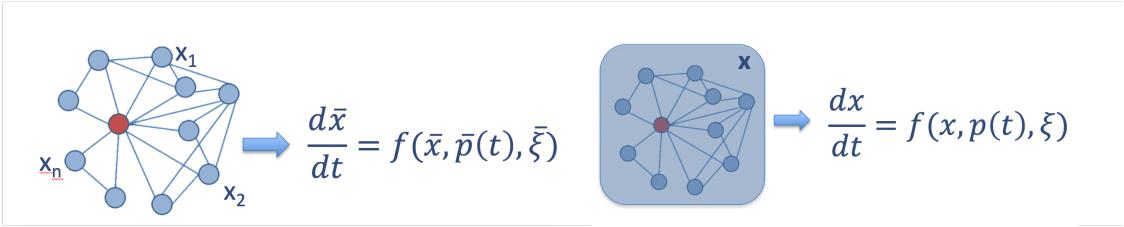


Fig. 2.7 An illustration for a network system composed by all its "microstates" (resulting in vector component), and the corresponding "macrostate" (lumping the micro-constituents into a single variable).

Critical transitions in high-dimensional systems are becoming increasingly studied, *e.g.* percolation-based (Rodríguez-Méndez et al., 2016), network (Iglesias et al., 2010; Drion et al., 2015b; Meng et al., 2020; Gao et al., 2016) or spacially-extended (Kiers et al., 2019; Gowda et al., 2015) transitions. However, much work is still open in this domain, including system-specific mapping to reduced systems (and conditions for it to hold) and performance of the center manifold theorem to describe real-world systems. As this thesis primarily focuses on noisy bifurcations, high-dimensional transitions will be only touched upon marginally.

2.3.2 Classification according to leading mechanism

Eq. 2.7 provides an elegant way of distinguishing between classes of critical transitions. Here, we provide a brief overview on them, bearing in mind that inferring whether a real-world system belongs in which class is subject to scientific investigation and is often challenging (see also Sec. 3.3). Over low-dimensional system descriptions, three classes of tipping effects can be recognised (Ashwin et al., 2012), according to time scale separation:

1. B-tipping: the abrupt shift is primarily driven by a bifurcation on a quasi-static attractor (slow passage through a bifurcation);
2. N-tipping: noisy fluctuations push the system away from the neighbourhood of a quasi-static attractor (transition due to random fluctuations);
3. R-tipping: the system fails to track a continuously changing attractor (non-autonomous system).

More challenging mechanisms occur when separation conditions do not hold and timescales mix up, or when big stepwise changes in external conditions happen, resulting in big forcing out of the attractor. Accommodating attractors collisions (Heagy et al., 1994) and other exotic cases in this or wider frameworks is demanded to future studies.

B-tipping

While slow changes happen and the parameter(s) is slowly (quasi-steady state) varied until a critical value, the system undergoes a sudden change in its set of equilibria. In this case, the leading eigenvalue slowly passes through zero across a bifurcation point and there is a transition

between hyperbolic manifolds (Kuehn et al., 2011). A condition for the correct application of B-tipping results is that "there is an adequate theoretical basis (or past evidence of threshold behavior) to show that there are parameters controlling the system that can be combined into a single control p for which there exists a critical control value p_c" (Thompson et al., 2011). Among existing bifurcations, only a finite subset is recognized as "critical" (Kuehn, 2011) or "dangerous" (Thompson et al., 2011), see Table 2.1 and Sec.A.

Table 2.1 Detailed discussion of thee most prominent bifurcations is provided in Sec.A.

List of critical bifurcations				
Local bifurcations				Global bifurcations
1-D, 1-cD	1-D, 2-cD	2-D, 1-cD	2-D, 2-cD	
Fold (saddle-node)	Cusp	Subcritical Hopf	Bautin	Homoclinic
Cyclic fold		Supercritical Hopf	Bogdanov-Takens	Boundary crisis
Subcritical pitchfork		Period dubling		
Transcritical				

Fold (point to point) and subcritical Hopf (point to cycle) are usually the most prominent bifurcations in the context of Critical Transitions (Scheffer, 2009), as they are abundantly observed in natural systems and are generic, *i.e.* they survive small imperfections and unfolding, see Def. A.5.10 and Thompson et al. (2011).

EXAMPLE: let us consider the normal form of a cusp bifurcation (Sec. A):

$$\dot{x} = f(x,p) = p_1 + p_2 x - x^3 \tag{2.9}$$

where $x \in \mathbb{R}$ is the system state. ODE 2.9, once p_2 is fixed at some value that allows bistability, has two stable equilibria ($x_{eq}^{(s)} = \pm 1$ after normalization). They correspond to the location of two stable wells. One additional unstable equilibrium $x_{eq}^{(u)}$ is located on the ridge between them. The mechanism that pushes the right-hand equilibrium from stable state to unstable is a local fold bifurcation: changing parameter p_1 alters the underlying potential until it vanishes. The system thus tips onto the second equilibrium (Fig. 2.8). Locally, this mechanism is typical of a fold normal form $\dot{x} = -p + x^2$, the local center manifold associated to Eq. 2.9.

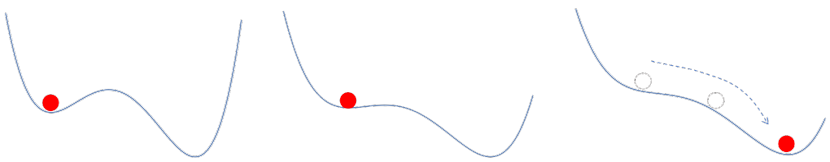


Fig. 2.8 "Pebble-down-the-hill" illustration of the fold B-tipping mechanism (for the "fold" subclass). As a parameter is varied, the bowl changes in shape, so that the original equilibrium gets destabilized (goes from stable to indifferent to unstable). The system "tips" when the equilibrium becomes unstable through a bifurcation.

N-tipping

Critical transitions far from bifurcations can happen due to noise. Random fluctuations can drive the system's state towards basin boundaries and drive it onto alternative attractors. Noise properties, *e.g.* multiplicative (Sharma et al., 2016a; Sharma et al., 2017) or coloured noise (Dutta et al., 2018; Kuehn et al., 2022), or in combinations with periodic forcing like in stochastic resonance case (Gammaitoni et al., 1998), can give rise to additional cases. The probability of tipping due to random occurrences strongly depends on stability and resilience system properties, and can be quantified using the classical theory of stochastic processes and associated measures like Kramers escape rate and mean first passage time (see Sec. B).

Due to noise, many authors observed that critical transitions cannot be predicted *strictu sensu*, but only alerted for, with associated probabilities (Zhang et al., 2015): the task is therefore more akin to detection of "dangerous" regimes and risk assessment, than to forecasting. When the noise component is most prominent, as in the case of N-tipping, some studies argue for the intrinsic unpredictability of such class of phenomena (Ditlevsen et al., 2010). Nonetheless, possible precursors for a change in regime and a settlement onto alternative states have been put forward and tested, like the "flickering" phenomenon (Dakos et al., 2015; Olthof et al., 2020). This is related to general fragility of shallow alternative attractors allowing the system to flicker back and forth, a signal that the original one lost resilience while the new one cannot yet "trap" the system.

EXAMPLE: let us consider again Eq. A.20. While B-tipping was driven by deterministic mechanisms, let us now add white noise, so to obtain a Stochastic Differential Equation:

$$dx = (p_1 + p_2 x - x^3)\,dt + \sigma dW \tag{2.10}$$

If σ is high in comparison to the barrier height, it can push the system state from one equilibrium up to the hill and then down to the second one. If σ is still high in comparison to the new ridge, the system can tip back, thus flickering between states until one equilibrium disappears completely (due to an underlying bifurcation) or the asymmetry overcomes the noise effects.

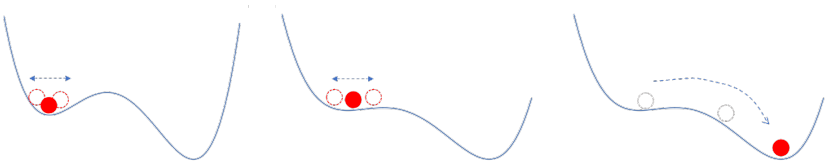


Fig. 2.9 Illustration for analogy of the N-tipping mechanism (additive noise): the system is already subjected to noise that makes it oscillate faster than the rate of parameter change (dashed arrow). If the bowl becomes less resilient, the fluctuations might drive the system over the tipping point, even though the equilibrium was still stable.

R-tipping

Let us consider a non-autonomous system in which $p = p(t)$. In this scenario, the attractor moves under the influence of a time-dependent parameter; if it moves "too fast", the system might lose its track and tip without having previously lost stability. Let us follow Ashwin's argument (Ashwin et al., 2012; Ashwin et al., 2017b) to construct a minimal model for R-tipping: consider the

generic fast-slow system Eq. 2.5. It can be re-written as a non-autonomous system:

$$\frac{dx}{dt} = F\left(x, p(t)\right) \tag{2.11}$$

Suppose that said system has a quasi-static equilibrium $\hat{x}(p)$ with a *tipping radius* $R > 0$. For some initial condition x_0 with $|x_0 - \hat{x}(p)|$, we assume that the evolution of x with time is given by:

$$\frac{dx}{dt} = M\left(x - \hat{x}(p)\right) \text{ for } |x - \hat{x}(p)| < R \tag{2.12}$$

where M is a fixed stable linear operator (recall that the system is linearized around $\hat{x}(p)$), that is, $|e^{Mt}| \to 0$ when $t \to 0$. If $|x(t) - \hat{x}(p(t))| < R$ we say that the system state $x(t)$ *traks* (adiabatically) the quasi-static equilibrium $\hat{x}(p)$. Otherwise, if there is a t_0 s.t. $|x(t) - \hat{x}(p(t))| = R$ then we say the solution *R-tips* at t_0.

Intuitively, one can think of a system with little inertia, under which the attractor is moving (Ritchie et al., 2016). By using the 'pebble in a bowl" metaphor, one can imagine moving the bowl fast enough so that the pebble cannot adjust immediately and rolls over (similar phenomena might happen when suddenly moving the bowl in an opposite direction as before: the pebble goes on rolling because of inertia (Haragus et al., 2018)). In a mechanical system, the very first experiment to identify R-tipping was described by Bonciolini et al. (2018). Alternative mechanisms involve potentials that vary periodically (periodic forcing) (Williamson et al., 2016) and other non-obvious thresholds (Perryman et al., 2014). For biological systems, one can think of an input condition that changes too fast for the system to adjust to it, so that it experiences an abrupt shift that wouldn't have happened in case of slow variation.Preliminary research show that speed dependency may concur in regulating cellular decision making (Nené et al., 2012), but additional studies are required on this direction.

EXAMPLE: the so-called "compost-bomb instability" (abrupt release of carbon dioxide due to fast variation of temperatures in Syberian peatlands) was described as a first example of R-tipping by Wieczorek et al. (2011). Fig. 2.10 provides an visual representation of the mechanism.

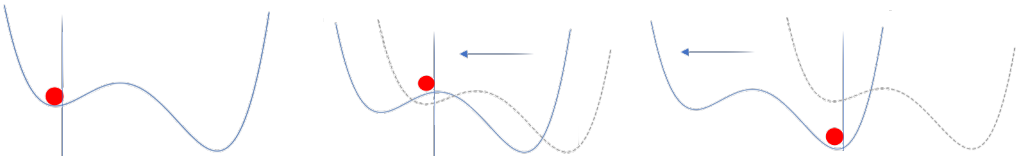


Fig. 2.10 Illustration for analogy of the R-tipping mechanism (in its simplest form): the bowl itself is moving because of time dependent parameter. If such movement is large enough with comparison to the inertia of the system, the latter is no more able to "track" the equlibrium and thus tips without the equilibrium losing stability. Solid lines identify the potential at its new positions, dashed lines to its "ghost" from time 0.

2.4 Noisy bifurcations

Among the classes above, noisy bifurcations (B-tipping with relatively small noise) are of particular interest to develop early warning signals based on dynamical systems theory. In fact, they are rather abundant in natural and societal systems, and they are associated with analytical results that can be applied on empirical measurements. The present thesis primarily focuses on this class of noisy bifurcations, the related phenomenon of critical slowing down and on the early warnings derived from it. This section thus overviews these concepts and their notable properties.

2.4.1 Critical slowing down and early warning signals

When a leading eigenvalue approaches zero, the response time to any perturbation diverges and the system becomes "slow" in returning to its equilibrium point. This effect is termed "critical slowing down" (CSD, further discussed in Sec. B) and is often looked for as a hallmark for impending regime shifts. CDS determines the expected trends of summary statistics indicators computed on fluctuations around an equilibrium – like variance, autocorrelation and other moments. As a consequence, they are often called "critical slowing down indicators" (Dakos et al., 2008; Dakos et al., 2012b; Van De Leemput et al., 2014).

The average rate of decay of perturbations to the equilibrium point may differ according to the bifurcation involved. This scaling, that propagates to CSD indicators as well, has been suggested as a feature to look for to distinguish bifurcations governing different real-world systems (Meijer, 2014; Bury et al., 2020). CSD scaling for each bifurcation normal form is listed in Sec. A.

Note that CSD is related to the vicinity to a non-hyperbolic point and is not unique to "critical" bifurcations. In addition, critical slowing down is a property that combines slow approaches to bifurcation points and the statistics of fluctuations around equilibria (see Stochastic Processes below). Hence, its theoretically expected properties might be more nuanced in real-world systems, yielding to modelling and interpretation challenges illustrated in Chapter 3.

The idea of measuring the "critical slowing down" phenomenon with a range of derived indicators, computed as summary statistics on time-series data, led to the development of "early warning signals" (EWS, Scheffer (2009)). As mentioned in the introduction, the term "resilience indicators" should sometimes be preferred: they are proxies of losses of resilience due to bifurcations; whether a transition occurs or not is further due to noise interplays.

Common statistical indicators are analytically derived from first-order expansions near the equilibrium manifold and from solving the corresponding Ornstein-Uhlenbeck (O-U) processes, see Sec. B. Variance, autocorrelation, spectral properties, entropy measures and so on are examples of such indicators. Their expected trends next to criticality can then be used as EWS: a rising autocorrelation could in fact anticipate a saddle-node-driven critical transition, see Fig. 2.11 for an example. Deriving the most common indicators and their trends is performed in Sec. B.

As they pertain to certain CT classes and near-equilibrium phenomena, EWS rest on assumptions, which should be careful tested before their routine application on real-world problems. Chapter 3 discusses known results and challenges. Later chapters in this thesis will be specifically devoted in estimating, quantifying and interpreting the performance of various EWS in general and on biological and epidemiological cases.

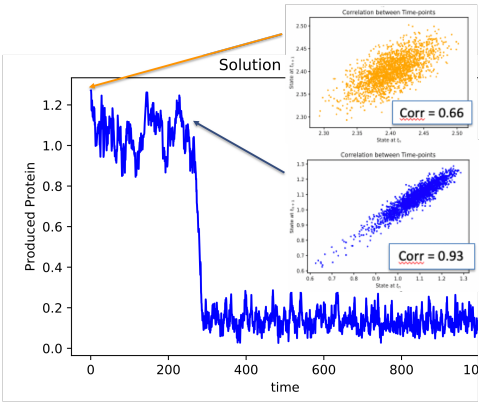


Fig. 2.11 Simulations from a biological system introduced in Sec. 4.1.1. A critical transition happens around $t = 300$ simulation steps. Autocorrelation ("corr") values increase when measured on windows closer to the transition. Such increase could be interpreted as an EWS.

2.4.2 Notable properties and visual representations

Due to the interplay of multiple attractors, noise and the local nature of early warning signals, local bistability is a fundamental property of systems that, in principle can be globally multistable (Sarkar et al., 2019; Feng et al., 2016; Lu et al., 2013). In fact, global multistability can be challenging to assess in the absence of valid models, whereas local bistability can be reduced to more generic bifurcation properties and associated with expected EWS. In addition, local bistability property is robust and conserved against noise up to certain levels (Shi et al., 2016). An illustrative example is in Fig. 2.12 (left, centre), with multistability along different degrees of freedom. The picture is simpler for "series" of local critical transitions in case of multistability along the same degree of freedom 2.12 (right). For these reasons, local stability will be the primary focus in the rest of the thesis. Note that the illustrations below mostly pertain saddle-node related critical transitions, as they are generic and well-represented in literature. Other bifurcations and related diagrams are considered in Sec. A.

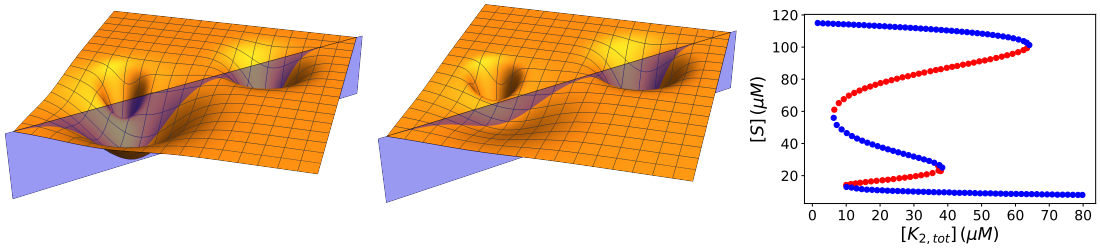


Fig. 2.12 Example of multistable system. **Left:** Imagine that noise intensity is such that the two attractors on the left are visited almost equally. Then, the preferred tipping direction is along the blue ridge. **Centre:** This effect may be accentuated when, under the action of control parameter(s), the original attractor becomes shallower and the system is driven to a bifurcation. The interplay between depth, eigendirections and noise intensity is to be treated quantitively with system-specific models. **Right:** multistability along one degreee of freedom (as if the attractors are aligned in series), viewed with a bifurcation plot. It results in a series of local saddle-node bifurcations. Figure adapted from Feng et al. (2016), Fig. 3ii.

As it is linked to local bistability, the low-dimensional "pebble-down-the-hill" analogy (Sec. 2.2.5) can be used to inquire local properties in various systems. Together with the bifurcation diagram representation (*e.g.* Fig. 2.5), it can be used to visualize several tipping classes (like in Sec. 2.3.2) and related properties. The two views are often complementary. For instance, Figs. 2.13 (left and right) correspond to Fig. 2.8 and Fig. 2.9, respectively. System's trajectories track the stable manifold (corresponding to equilibrium values) until either tipping at the bifurcation point (when the attractor disappears) or a bit before because noise intensity drives it past the unstable point (when the attractor becomes shallower and noise-induced transitions occur).

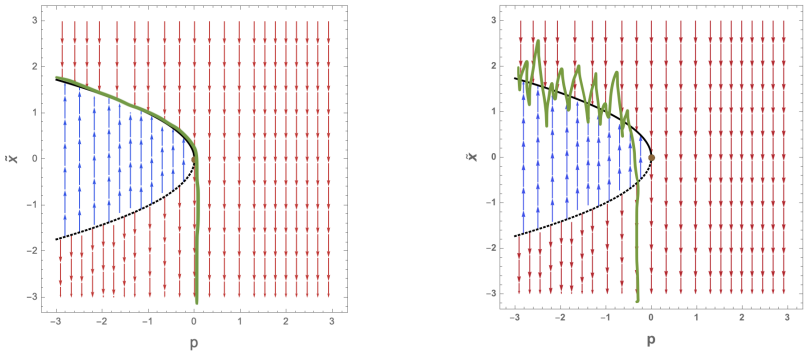


Fig. 2.13 Example of tipping classes viewed on bifurcation diagrams. **Left:** b-tipping. **Right:** n-tipping. System's trajectory (green) tracks the stable manifold (solid black). Red and blue arrows (representing vector fields) "push" the trajectory towards stability. A critical transition may occur either at the bifurcation point (attractor vanishes) or before, due to noise intensity driving the system to the unstable manifold and then away of the original attractor.

A notable property on non-linear bistable systems is hysteresis. It refers to asymmetric routes to tipping from one or the other attractor. If the asymmetry is too high, there can even be irreversibility, i.e. a system can tip on one direction but cannot return to the original state. This can be visualised on the bifurcation diagram, or by considering asymmetric potentials and particles wiggling into them.

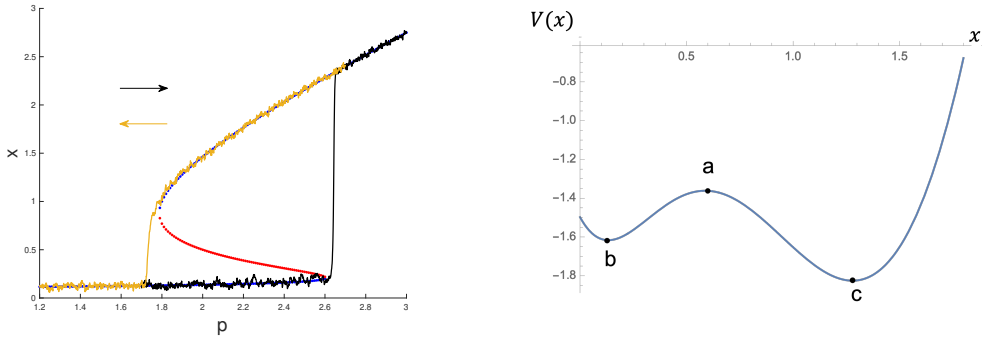


Fig. 2.14 Observing hysteresis. **Left:** Simulations of Eq. 4.1 over the bifurcation diagram. Decreasing parameters from left to right push the system to an "off" state through a CT (yellow route). To return to the original state, higher parameter values need to be reached to get t a second tipping point (black route).**Right:** The corresponding potential for a fixed p value. It derives from the adjoint stochastic potential (see Sec. B), which link dynamical models and the "landscape" illustration. In case of asymmetry between attractors, it might take longer to jump from one to another, either depending on random fluctuations or following the action of a parameter "lifting" the equilibrium, as $|V(a)-V(b)|<|V(a)-V(c)|$.

Other notable properties are system's stability and resilience. They are related to the illustrations provided by the landscape potential and the bifurcation diagram. Dai et al. (2015) link stability with the potential's depth and resilience to its "width", visualised in the bifurcation diagram as the amplitude of the fold parabola. Their interplay may yield different routes to CTs and different EWS performance. Chapter 6 quantitatively investigates these aspects. Formal definitions for resilience and stability, as well as for robustness, optimality and related concepts, can be traced back in literature. Their quantitative connection with complex feedbacks, their magnitude and timescales is still being recently investigated, see *e.g.* Frank et al. (2002); Franci et al. (2018); Kéfi et al. (2016).

Sensitivity to initial conditions is immediately observed next to criticality. Under the action of mild noise, small steps closer or further from the bifurcation point may signify remaining stable or tipping onto alternative equilibria.

Emergence can also be portrayed in such conceptualizations. In fact, quantitative changes, potentially deriving from combinations of conditions and influences, may push the system onto one stable regime or another, which can be regarded as the "emergent", "macroscopic" observables. Additional insights on these aspect might unravel modelling concepts and interpretations.

2.5 Connection with other "critical" frameworks

> The complexity of physical and biological scaling phenomena has been found to transcend the explanatory power of individual paradigmal concepts. The interaction between theoretical development and experimental observations has been very fruitful, leading to a series of novel concepts and insights (Marković et al., 2014).

The critical transitions framework sits among other research fields addressing critical phenomena from different angles. They are characterised by modelling choices about relevant separations of time scales (Marković et al., 2014) or by research objectives and focus. Cross-fertilization constantly occurs, to unravel fresh aspects of such important and sometimes elusive events.

Historically, statistical mechanics first tackled critical states during phase transitions, although keeping a rather "static" stance (Pathria et al., 2011; Binney et al., 1992). When the control parameter is allowed to slowly change in time, the so called "first-order" transitions (having discontinuities in the first derivative of the energy function) directly relate to critical transitions (Gondelfeld, 1992; Binder, 1987). First-order phase transitions have been observed in lasers and other physical devices (Scott et al., 1975), but have also been proposed as analytical tools to address cell development (Solé, 2003) or synchronization of coupled oscillators, including neural networks (Kopell et al., 1990). Recently studies started considering EWS in typical statistical physics systems like the Ising model (Barnett et al., 2013; Morales et al., 2015; Marinazzo et al., 2019).

Catastrophe theory (Thom et al., 1977; Zahler et al., 1977) tackled critical phenomena from a topological perspective, identifying "organising centres" like the cusp bifurcation. Key concepts were, *e.g.*, used to develop conceptual models of behaviour and perception, see Fig. 2.15. The endeavour of "explaining all critical phenomena" was nonetheless undermined by harsh criticism

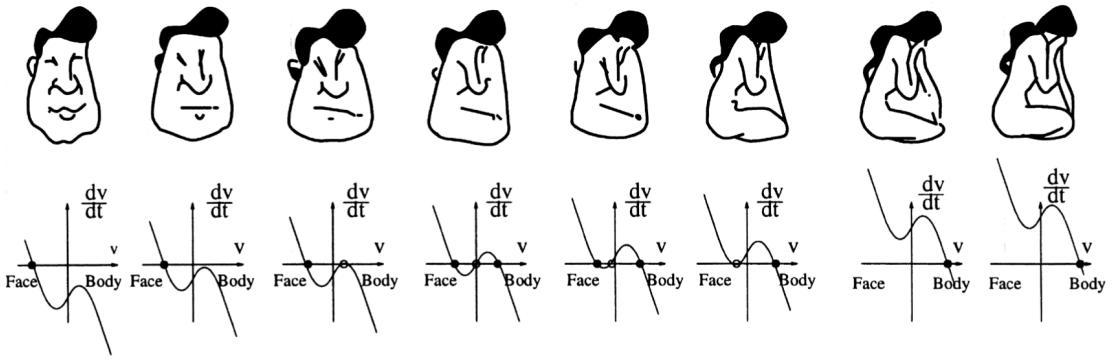


Fig. 2.15 "Gestalt" bistability in perception and its connection to the cusp model (compare with cusp behaviour in Sec. A). The picture, made by Attneave (1971), represents the biasing sequence for the Fisher (1967) figure.

of being abstract and pretentious, without openness to empirical investigation (Zahler et al., 1977; Kolata, 1977).

> No mathematical theory, no matter how elegant it may be, can serve as a substitute fo the hard work of learning facts about the world. (Zahler et al., 1977)

Nonetheless, analytical instruments like the unfolding of organizing centres now play an important role to develop more refined models, in particular in the field of mathematical neuroscience (Izhikevich, 1998). Here, notable efforts have been paid to connect topological concepts to conductance properties of excitable neurons (Drion et al., 2012; Drion et al., 2015a; Drion et al., 2018; Franci et al., 2012; Franci et al., 2014; Franci et al., 2013).

Another intriguing framework, primarily concerned about the emergence of power-law scaling in various natural phenomena, is self-organised criticality (SOC, Bak (1987); Bak (1996)). Starting from the sand-pile model, it was expanded to interpret many system, ranging from biology (Jensen, 1998), geology (Rundle et al., 2002), finance and so on (Gal et al., 2010; Gal et al., 2013). Its direct connection to first-order phase transitions was recognised and discussed by Dickman et al. (2003); Dickman et al. (2000); Di Santo et al. (2016). Both frameworks, augmented with noise influences, are also addressed by Sornette (2013).

Interest in critical phenomena also diffused in the engineering community, particularly in the field of systems control (Heinrich et al., 2010; Iglesias et al., 2010). In this case, the main questions regard the emergence of critical states from feedbacks (Franci et al., 2016; Franci et al., 2018; Drion et al., 2015b; Dethier et al., 2015; Bokes et al., 2018) and their relevance for control (Åström et al., 2010; Skogestad et al., 2005), as well as the identification of noisy precursors of ruptures (Jeffries et al., 1985).

The CT framework often implies a holistic, top down stance: assess to which CT class a system is likely to belong, then derive the suitable EWS. Ideally, this work should be complemented with reductionist, bottom-up approaches investigating the specific underlying mechanisms. Continuos

developments of systems-specific theoretical understanding would eventually make up for refined knowledge and better risk-assessment tools.

2.6 Paradigm shift and significance for systems biology

The critical transitions framework introduces a modelling shift: alternatively to the view of a fixed manifold in which equilibria are already pre-determined ("community perspective"), CT take an "ecosystem perspective" (Beisner et al., 2003) by considering the effect of external changes on the state of a system. The main difference between the community and ecosystem approaches is what is considered a variable and what is considered a parameter. In the community perspective, all degrees of freedom are variables; the landscape is fixed and the system can move solely under the influence of additional forces or noise. In the "ecosystem" perspective, state variables define the dynamical manifold, morphed by the action of parameters that model relationships between variables. The manifold itself is thus dynamic and accommodates other routes to tipping events. This re-conceptualisation thus opens new modelling and analysis avenues.

Developing from its decade-long debates, systems biology could profit from the "ecosystem" perspective to refine its conceptual archetypes, accommodate modelling strategies and unify theoretical and empirical observations.

A long-lasting conceptual paradigm has been the so-called Waddington's (or "epigenetic") landscape (Waddington, 1957). It is a metaphor for cell commitment. It is imagined as a rugged "energy potential", with systems' states depicted as marbles rolling down. The marbles sample the grooves on the slope, and arrest at the lowest points. Eventual cell fates are represented by such lowest points (Fig. 2.16 left). Ideally, the landscape shape is statically pre-determined by gene states and fixed (Fig. 2.16 right). Over the years (Ferrell, 2012; Ghaffarizadeh et al., 2014b), the Waddington's landscape has been associated to models of cell development and quantified from steady-states populations. It was thus framed as a "community perspective" manifold, multistable but static and sampled by noise (Wang et al., 2011; Kanamaru et al., 2013; Zhou et al., 2013; Sardanyés et al., 2018).

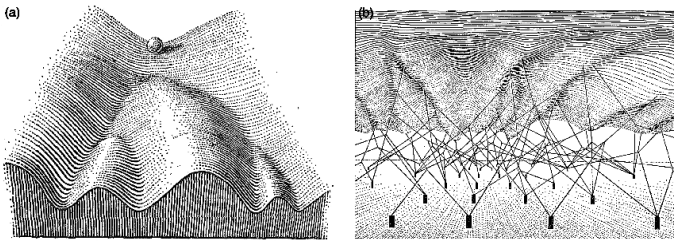


Fig. 2.16 Illustration of the Waddington's landscape, from its original publication (Waddington, 1957). **Left:** the landscape. **Right:** an illustration off genes shaping the above landscape, like rods on a tarpaulin.

More recently, modelling cells as dynamical systems has become widespread, including the analyses of their bifurcations towards criticality-driven cell decision (Pal et al., 2015; Duff et al., 2012; Kerr et al., 2019; Huang et al., 2007; Andrecut et al., 2011; Stanoev et al., 2021). In addition, cell lineage tools are making increasing use of bifurcation methods to predict discrete (and sometime discontinuous) commitment states (Richard et al., 2016; Eugenio et al., 2014).

Noise and bifurcations have often been treated separately as two drivers for cell differentiation, although some studies considering their combinations exist (Pfeuty et al., 2014).

These studies spurred calls to re-think the role of fixed landscapes crossed by continuous trajectories, and the search for re-conceptualizations of cell-fate decision mechanisms, both in terms of modelling and of theoretical understanding (Moris et al., 2016; Korolev et al., 2014). During the development of more refined dynamical systems-driven studies, however, the Waddington landscape was still used as a powerful metaphor, see e.g. Mojtahedi et al. (2016).

Building on what discussed above, it is quite natural to go beyond, and to look for the "ecosystem perspective" for dynamical systems evolving at different time scales. Identifying proper separations of state variables and control parameters, both dynamically evolving in time with respect to their typical time scales, makes up for a dynamic manifold, with attractors as states and transitions driven by both deterministic conditions (ramping parameters) and stochastic occurrences (noise). In the epigenetic interpretation, state variables may correspond to gene activation states; their trends and inter-relationships are in turn tuned by parameters representing dynamic controls. Dynamic multistability captures phenotypic plasticity. In addition to being potentially more flexible and versatile, this conceptual modelling accommodates various nuances of evolutionary theories, allowing adaptation in changing environments.

Complete gene regulatory networks are impractical to model. Instead, CT theories combine a holistic perspective and strive to connect it with a reductionist approach. For instance, they contribute to characterise dynamical pathways defined by common structural patterns ("network motifs", Alon (2019)). In this sense, they help connecting microstates (combinations of gene activation profiles) and macrostates (cell sates). Noisy bifurcations and other transition mechanisms may thus concur cooperatively in shaping cell development. Continuos as well as discontinuous transitions may occur, both driven by deterministic or stochastic processes. Heterogeneity can be accommodated by transients divergences from static equilibria, occurring within attractors. Asymmetries in cell-fate commitment (Mojtahedi et al., 2016) can be accommodated in unfolding mechanisms (Def. A.5.9 and Fig. 4.7), further tuned by noise. Dynamic landscapes can be determined using refined methodologies (many of which discussed later in the thesis) over time-series population data.

Determining which classes better represent certain cell lineages will thus occupy many future studies, alongside the progress of dynamical indicators to measure, detect and predict such fate transitions.

Chapter 3

Early warning signals: the explored and unexplored avenues

This chapter overviews the debate around the critical transitions framework and around the application and interpretation of theoretical early warning signals when extracted from real-world data. The evidence here presented reminds that observational hypothesis ought to be supported with statistical analysis, compared with model predictions, hopefully verified with controlled experiments (if available) and validated by iterating the "circle of scientific discovery" (*cf.* Fig. 1.1). Only by this careful analysis, it is possible to state that a system presents alternative stable states, that a certain phenomenon is indeed a critical transitions, and that measures associated with CTs can be used to construct an early warning system. In addition, this section lists known open questions and paves the motivations for Chapter 5.

3.1 Is an observed system bi- or multi-stable?

Necessary condition for observing signals related to critical transitions in bistable systems is that the observed system is (at least locally) bistable. Although tautological, this condition is often difficult to verify in unknown system, for which a mechanistic model is not available or is just being developed. A few methods for inferring bistable properties from data or scarce information have been developed: in this section, we review the ones often associated with the literature on critical transitions.

Bistability is a necessary condition to observe critical transitions, but it is not sufficient to determine whether a transition is bifurcation-driven or noise-driven. This issue is further discussed in Sec. 4.3.1 with mathematical examples. Multistability can be established using similar methods recursively.

3.1.1 Controlled experiments

Ideally, bistable properties of complex systems are assessed with controlled experiments, that track small and progressive modifications of control inputs and measure the response of the system output, in search for dependency on initial conditions, hysteresis, or to reconstruct the bifurcation diagrams. Such experiments are common in the physical sciences and in engineering applications, and many have been historically (Heller, 1967) and recently (Rosen et al., 2010)

performed to study critical properties in quantum, electrical and mechanical systems. In 2018, the first controlled experiment observing rate-induced tipping was carried out onto a laboratory-scale combustor subject to thermoacoustic instabilities in the presence of turbulence-induced noise (Bonciolini et al., 2018).

In order to check for critical transitions in ecological and biological systems (using controlled experiments), a few studies have been conducted only recently. A seminal work of Dai et al. (2012) performed a controlled study of population collapse in yeast colonies, reproducing the results predicted by May (1977) and, more recently, by Scheffer et al. (2009). The results of a whole ecosystem experiment were published by Carpenter et al. (2011), where researchers gradually added predators to a shallow lake food web, until its collapse. Drake et al. (2010) conducted an experiment with replicate laboratory populations of *Daphnia magna* (a small planktonic crustacean) controlled by deteriorating environmental conditions. Additionally, works addressing neural activity were performed by Meisel et al. (2015), who also investigated the scaling associated with different normal forms, as well as the possibility of using it to distinguish between leading bifurcation. These works have in common the careful control of stressing conditions and on the surrounding environment, so to adhere to the modelling assumptions of slow forcing, repeated measurements and white noise. Critical transitions and expected behaviours of statistical indicators were indeed observed, confirming their existence and evolution in various complex systems. Additional studies (Reyes et al., 2020) are recently suggesting experimental methods to detect bistability in chemical and biological networks.

Limitations and future research

This kind of study is arguably the most solid to verify the presence of critical transitions and to characterize their properties. Unfortunately, controlled experiments are often very complicated or unfeasible. They might be "just" extremely complicated when the system of interest is complex, open and subject to additional confounders that are very hard to identify and quantify. In case of climate or ecological systems, on the other hand, performing such experiments is impossible as no replicates are available or, in general, it is highly advisable not to wait until the critical transition to verify that it happened – think of mass extinctions or abrupt climate change. In all these cases, scholars mostly rely on models – more or less detailed – or on data from the past, hoping that future evolutions follow similar patterns and are thus amenable to extrapolation.

3.1.2 *A posteriori* analysis

Once observational data about the evolution of a system are available, it is possible to identify whether the system changed its condition abruptly. Here, the assumption is that the observations pertain the variable along which the transition might occur. Verifying regime shifts usually corresponds to identifying drastic changes in the properties of the observed time series.

When low-dimensional time series data are available, the first check is on the mean itself: already by eyeball visualization, it is often possible to notice rapidly changing trends. Such observations can be refined with statistical tools. Wang et al. (2012a) provide several checks to evidence bistability, critical transition and alternative states, using ecological data as case studies. These methods, which include checking for bimodality in the distribution density functions and clustering in the phase space, are discussed in details in Sec 4.3 along with others that we have

tested on other data sources. An interesting alternative is potential analysis (Livina et al., 2010), which aims at reconstructing the underlying dynamical potential. It is also discussed in Sec 4.3.

Other methods include fitting minimal bistable dynamical models to data (Lade et al., 2012) and comparing the goodness-of-fit with that obtained from smoother functions. Checking for transitions that involve period doubling or other periodicity modifications (e.g. those modelled by Hopf bifurcations) might require switching to the frequency domain to perform the analysis.

Although not considered later in this thesis, worth mentioning are also spatial data, where critical transitions are observable in the drastic deterioration of their components (e.g. grasslands shifting towards desertification) (Scheffer, 2009). Recent works are also considering the use of spatial indicators like Betti numbers (Storch et al., 2019b) to quantify the observed changes. Future work will likely build on top of recent image analysis techniques, e.g. extrapolating measures like Jaccard index or Dice score (Eelbode et al., 2020) for the observables of interest and follow their trajectory over time.

Finally, the trend of statistical indicators – those called "early warning signals" during detection and prediction tasks – can be used to support the evidence of alternative states in a *a posteriori* analysis. Several papers addressed this possibility in the past (Dakos et al., 2008; Lenton et al., 2012; Zhang et al., 2016; Wang et al., 2018; Van De Leemput et al., 2014; Burthe et al., 2016; Veraart et al., 2012; Chen et al., 2012a), but doubts exist whether such indicators are generic enough to reliably identify a critical transition and to distinguish it from other dynamical features, if no additional information and modelling of the system itself is available (see Sec. 3.4 for theoretical arguments, Sec. 3.3 for examples of ongoing debates, and Sec. 5 for systematic analysis). Analogous observations also apply to indicators employed on high-dimensional data, like dynamical network biomarkers (Chen et al., 2012b) or autocorrelation indexes (Mojtahedi et al., 2016; Yan et al., 2021).

Limitations and future research

Inferring information from observational data is often partial and subject to interpretation. While essential to construct hypothesis towards model development and system characterization, such statistical analysis is often not conclusive. For instance, depending on the sampling frequency and on the measurement choices, the transition from one state to another might look more or less abrupt, with consequent fuzziness of the interpretation in terms of critical transitions. Moreover, fitting piece-wise of smoother functions might be subject to real data seldom been sharply subdivided and could result to smooth high-order polynomials being preferred according to some cost function criterion, providing a false positive in system identification.
Fig. 3.1 shows a toy example for such a problem: observational data are potentially taken to be representative of a critical transition, and even a changepoint analysis (*cf.* Sec. 4.3.4) would suggest so. Nonetheless, this is a spurious effect induced by irregular sampling next to a gradual transition, along the smooth branch of an unfolded supercritical pitchfork bifurcation (*cf.* Sec 2.2) in the form

$$\frac{dX}{dt} = K + c \cdot (X-1) - (X-1)^3 \,, \tag{3.1}$$

with $K = 0.01$ and control parameter c spanning from -0.5 to 2, with critical value $c_0 = 0$. Simulations are performed by adding white noise with intensity 0.02. Although this made-up example may look arbitrary, similar situations happened several times during the monitoring of

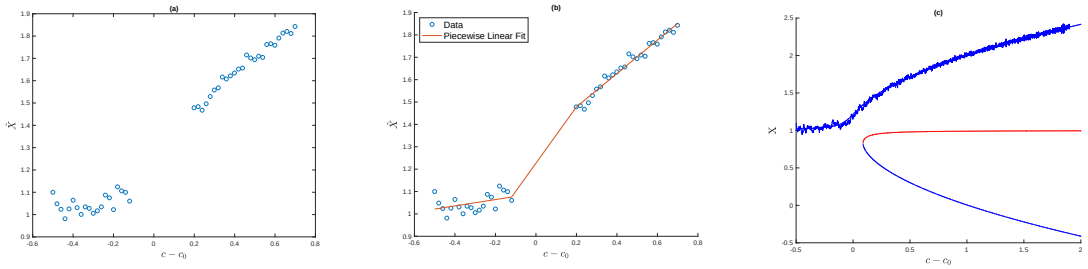


Fig. 3.1 Toy example: identifying a system as prone to undergoing critical transitions after analysing its time series. **(a)** Synthetic data ($\tilde{X}$) reproducing plausible observations: at a certain point during the measurement of a system, there is a gap in sampling; data are then clearly divided into two cluster. **(b)** Changepoint analysis (Sec. 4.3.4) reveals that the best fit distinguishes the two data clusters, suggesting a critical transition between the two regimes. **(c)** Nonetheless, the data derive from noisy simulations over the smooth branch of an unfolded subcritical bifurcation (Eq. 3.1, with c as control parameter with critical value c_0 and X as dynamical variable), which does not display critical transitions. Note that, with higher noise levels, a N-tipping transition could have been observed towards the alternative branch.

wastewater data against the COVID-19 pandemic (see reports on https://www.list.lu/en/covid-19/coronastep/). Reasons include working shifts, public holidays, and even an unprecedented flooding (https://bit.ly/3K95Dr1) of the region that disrupted the routine sampling frequency. As many EWS applications have been designed to target macro-scale real-world phenomena, such sampling issues should be taken into account and might induce biases in the interpretations. Other biases might e.g. result from repeated experiments performed at – ideally – constant parameter values, but without such values being sufficiently fine-grained. If the real system's evolution was well described by a supercritical pitchfork bifurcation (which is not critical (Kuehn et al., 2011)), an observer would get bimodal probability density functions but the interpretation would potentially be flawed into thinking about bimodality from criticality. Such questions might be raised to observational studies like e.g. Mojtahedi et al. (2016).

Another issue related to the use of piece-wise or smoother functions to fit observational data is, that real data are seldom "perfectly" abrupt, but often maintain some smoothness due to system inertia or noise. Sampling might also play a role in smearing transitions which could otherwise be modelled as "critical".

Regarding the use of statistical indicators to check for the criticality of a system form its time evolution, they are subject to the same questions encompassed by this thesis: are they unique, reliable and performing signals? Depending on the answers – and on which cases they hold –, the use of such indicators is more or less appropriate as supporting evidence for the existence of critical transitions in a given system. A potential risk is to fall into the "prosecutor fallacy", that is, to infer that the probability of observing an early warning signal given knowledge that a system is bistable is equal to the probability of system being correctly interpreted as bistable after the observation of an EWS (Boettiger et al., 2012a). Instead, the two conditional probabilities are weighted by the ratio between the probability of observing a trend associated to EWS (which can depend on other factors, like state-dependent noise), and that a transition occurs in general (which can also depend to random occurrences). See Sec. 3.4 for more about the fallacy and the interpretation of warning signals. In addition, if various data sources, analysis methods and

theoretical results concur in yielding contrasting results, debates might proceed without being solved by such analysis alone (see Sec. 3.3 for some examples).

Finally, a few words about bistability and abrupt transitions for high-dimensional systems: although some preliminary studies are being published in the very recent years (Weinans et al., 2019; Weinans et al., 2021), open questions still remain on how to collect evidence. A brief list of available techniques is provided in Sec. 4.3.6, while a broader overview of results ad future works is given in Sec. 9.3. For a recent review about EWS, see also (George et al., 2021).

3.1.3 *A prori* conditions for bistable networks

Without having curated time series at disposal nor complete models, but only relying on basic information about the system's structure and stationary data, is it possible to infer that such system is bistable? Differently phrased: how much information do we need to reliable infer bistability properties of an observed system? Such questions has been asked in particular about networked ecological systems (Kéfi et al., 2016) but, more in general, pertain any observed system described as a network, such as gene regulatory networks, metabolic pathways, food webs, climate feedbacks and so on.

The issue was tackled by Angeli et al. (2004a), where the authors established a necessary and sufficient conditions to infer bistable properties of a dynamical system, given a reduced amount of information about the system itself. The paper primarily focuses on bistability (or, more generally, multistability) in cell signalling, but the mathematical properties can be extended to any other complex system – more complex than very simple feedback systems, for which conditions already existed. The authors present a graphical method for deducing the stability behaviour and bifurcation diagrams for a class of feedback systems, mathematically deducing their stability properties when feedback is blocked.

The first take-home of this work is that, positive feedbacks are necessary but not sufficient conditions for a system to be bistable (different from what conjectured elsewhere, *cf.* Scheffer (2009)). The second contribution is a set of conditions to deduce the bistabiliy of networked systems for which a complete and mechanistic description is not available, let alone their complete evolutionary time series. The process works as follows:

1. Consider the system under study as an input-output (I/O) system, that is, identify what the control input might be and what to measure as output. This is part of modelling design.

2. For experimental studies: block the feedback from output to input. For theoretical studies: consider the open-loop system.

3. Obtain the Input/Output incidence graph[1] (Biggs et al., 1993) at steady state. It is not necessary to have the full network structure (including regulation functions), but only the "skeleton".

4. By following the theorem from Angeli et al. (2004b) (see below), deduce the presence of bistability.

Theorem-specific conditions to ensure bistability on the incidence graph are:

- Of a system, we can identify its input ω and output η and its labelled incidence graph. This way, we can recognize positive and negative effects. As a rule, activation in an incidence graph is marked by $+$ and inhibition with $-$. The sign along a full path on the graph is the product of the signs along that path (Gross et al., 2005).

- The action of each vertex of the incidence graph has to be monotonous (e.g. activation or inhibition, but not both).

- Conditions for "strongly monotone I/O system": (a) Every loop is positive; (b) all of the paths from the input to the output node are positive (meaning that all possible loops are positive); (c) there is a directed path from the input node to each node n_i; (d) there is a directed path from each n_i to the output node (the last condition ensures that the graph can be made).

- Well-defined I/O characteristic of the open-loop system: given any input ω, the asymptotically stable equilibrium exists and is unique for the open-loop system.

- The I/O characteristic is nonlinear ($\exists x^* s.t. \frac{\partial^2 \eta}{\partial \omega^2}|_{x^*} = 0$; a sigmoidal shape is ideal). It is thus not necessary to consider the full evolution (including transients) of the system, but only its stationary responses to disturbances.

Those requirements deduce that the system *can be* bistable. An additional requirement *guarantees* bistability:

- The $\eta = \omega$ diagonal intersects the I/O characteristic in $\hat{x}_i$ points. Then, stability of those points comes by looking at the slope s in $\hat{x}_i$: $s > 1$ highlights instability, $s < 1$ stability.

The result is general for n-component sytems. Complex networks can be analysed by making use of their modularity (bistable modules might build bistable systems). Examples, counterexamples and details can also be found in the referenced paper (Angeli et al., 2004a).

Under the lenses of dynamical theory, the proposed method assess the existence of a cusp bifurcation. In fact, the first requirements verify its presence, while the latter guarantees that the

[1] A graph is a triple (V,E,I), where V and E are finite sets, called the vertices and the edges, respectively, and $I : E \to V^2 \times 0,1$ is called the incidence function. The incidence function tells for each edge its end-vertices, and whether the edge is directed (1) or not (0). If an edge e is directed, then the first element and the second element of $I(e)$ denote the origin vertex and the destination vertex, respectively. This generality is required to represent all of undirected graphs, directed graphs, and mixed graphs, possibly combined with self-loops and multi-edges. A *labelled graph* is a graph together with labelling functions $L_V : V \to D_V$ and $L_E : E \to D_E$, where D_V and D_E are arbitrary sets, called the vertex labels and the edge labels, respectively (from http://kaba.hilvi.org/pastel-1.5.0/pastel/sys/incidence$_g$raph/incidence$_g$raph.htm and Gross et al. (2005)). Further reading in Sec. A

"pitchfork-like" parameter is properly tuned to make the system bistable. It follows that, if the system wants to get from one equilibrium to the other without changing its configuration, it has to undergo a local CT driven by a fold bifurcation (*cf.* Sec. 2.2). Hence, standard EWS apply and are suitable to be measured.

Application examples to ecological and physiological networks

Testing the above conditions on various graph models gives insights about their usefulness. In addition, it provides examples for systems where critical transitions and EWS can be expected and those where they shouldn't.

As a first example, consider the reconstructed incidence graph for an ecological system (a shallow lake) provided by Kéfi et al. (2016) (Fig. 3.2).

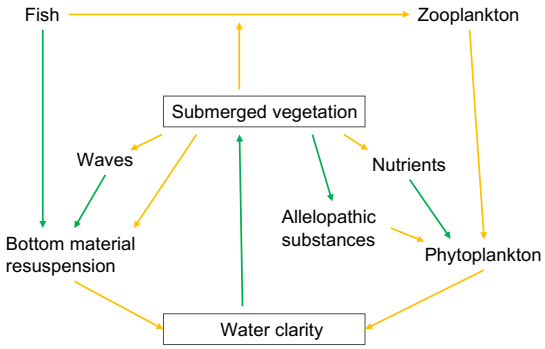


Fig. 3.2 The reconstructed labelled incidence graph for a shallow lake ecological system (adapted from Kéfi et al. (2016)). Orange arrows mean negative effects $-$ (depletion, predation, ...), green ones mean positive effects $+$ (enhancement, sustainment, ...).

The methodology introduced above can be applied as follows:

- Identify "Submerged vegetation" as input and 'Water clarity" as output.
- The incidence graph is already reconstructed from field observations (note that the mechanical properties of the interactions are unknown, meaning that we do not have the network but just the graph) and is properly labelled. Vertex are monotonous (Kéfi et al., 2016).
- Conditions (a) to (d) for the graph are respected (it is an easy exercise to check that all loops and paths from I to O are positive).
- It is known from field studies that the I/O response in non linear (Scheffer, 2009).

Hence, the system can be bistable. Unfortunately, Kéfi et al. (2016) do not specify the shape of the I/O characteristic, so it is not possible to say if the system *is* a priori bistable after applying the last condition. Nonetheless, for this specific system, bistability is observed *a posteriori* (Scheffer, 2009). This implies that the I/O response is actually behaving as expected from the theorem and that the regime shifts follow a fold bifurcation. As a consequence, the signals measured in these case studies from shallow lakes (Carpenter et al., 2011; Scheffer, 2009) are likely due to detection of the underlying fold bifurcation and not to other mechanisms at play, like increasing noise.

Let us consider a second example from physiology and medicine. The brain network in Fig. 3.3 reconstructs cortex and basal ganglia interactions and has been suggested to be relevant for studying the development of Parkinson's disease (Jones et al., 2014). Using the *a priori* methodology, we can check whether its alteration might contribute to a critical transition towards the disease state.

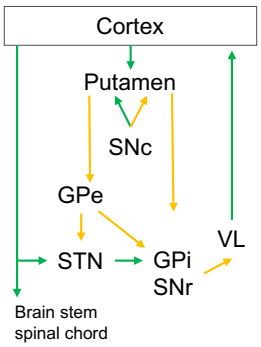


Fig. 3.3 The reconstructed incidence graph (labelled) for dopaminergic pathways between the basal ganglia and the cortex (adapted from Jones et al. (2014)). Graph variables are the activity of each brain region. Green arrows represent excitatory connections $(+)$; orange arrows represent inhibitory connections $(-)$. SNc: substantia nigra pars compacta; GPe: external segment of the globus pallidus; VL: ventrolateral nucleus of the thalamus; STN: subthalamic nuclus; GPi: internal segment of the globus pallidus; SNr: substantia nigra pars reticulata

In this case, the pathway is not monotonous and the I/O loops are not positive; therefore, the system cannot be bistable according to the theorem above. Consequently, either the pathway is still not complete and new elements might favour a different interpretation, or it is not possible to interpret the insurgence of Parkinson's disease as a critical transition towards a new state, after the reprogramming of the ganglia circuit.

As shown with two contrasting examples, this type of analysis paves the way for the correct application and interpretation of the theoretical results from the CT framework onto unknown systems.

Limitations and future research

Deducing in advance whether a system can be bistable or not is beneficial to correctly interpret and apply the results from the critical transitions frameworks. The theorem described and tested helps doing so with little information available. However, precise knowledge of the incidence graph and the I/O characteristic is necessary to fulfil the theorem requirements. Without those, it is only possible to draw preliminary and incomplete conclusions, like for the first example.

This promising avenue will greatly improve if new conditions are obtained for other experimental and/or modelling challenges, e.g. known missing links, multiplexes networks, time-scale separations among components, etc. In addition, an interesting direction would be to extend this methodology to simplicial complexes and hypergraphs (Bick et al., 2021). The underlying question would then be: how much information do we need to reliably infer bistability properties in a higher-order network?

3.1.4 Combining observational data and modelling

Observational data can suggest the existence of alternative states and shifts among them, but models are essential for good diagnosis (Scheffer et al., 2003). To begin with, models allow surpassing qualitative reasoning: as discussed in Sec. 3.1.3, positive feedbacks are necessary ingredients to observe regime shifts, but they lead to alternative states only if sufficiently strong with respect to other control mechanisms. Thus, quantitative reasoning provided by models is crucial. Moreover, bifurcations and other elements describing regime shifts are actually modelling ingredients: a bifurcation does not exist in real life, but is a useful concept to reproduce qualitative changes inn systems' state, driven by slowly varying forces. Finally, models – even minimal models (*cf.*Sec. 1.2.2) if more refined ones are not available – help distinguishing between transitions associated to bifurcations, noise or other mechanisms (Sec. 4.3.1).

In general, proving that a system is bistable and that it can undergo regime shifts is bound to follow the "cycle of knowledge discovery" (Fig. 1.1). It has been asked where does the burden of proof lie: instead of proving that a system does have alternative attractors, we might attempt to falsify this hypothesis. Unfortunately, this is anyway difficult in practice. Surely, inquiring complex systems with observations, experiments and models is a good way to obtain solid diagnosis on their dynamics, which in turn implies different views on monitoring and management options.

3.2 The challenges of extracting EWS from data

In systems that have been proven to undergo regime shifts between alternative states, extracting reliable early warning signals is still challenging. Here are listed several considerations from past literature and personal investigations, describing challenges for the extraction and accurate interpretation of EWS from measurement data. The review presented in this section will be expanded by original works reported in Chapters 5 and 7.

3.2.1 Different classes of models might display different EWS

After being originally associated with fold bifurcations only (Scheffer, 2009), the Critical Transitions framework now encompasses several classes of tipping mechanisms (*cf.* Sec. 2.3.2). A longed-for "generalised model" yielding model-free EWS is thus hard to foresee. Already Scheffer et al. (2012) observed that signals designed for fold models are not universally recorded. In general, CT classes might be associated with different early warning signals (Lade et al., 2012; Thompson et al., 2011) with varying detecting and alerting performance. More-so, non-bifurcation mechanisms for abrupt shifts like sudden shocks might not display EWS at all (Dakos et al., 2015). Other non-bifurcating classes, like noise-induced transitions, are potentially associated with other footprints like flickering (Dakos et al., 2012a), or might not show signals at all (Boettiger et al., 2013c; Ditlevsen et al., 2010). Hence, prior knowledge on the mechanism is required to adequately apply one indicator or another and to aptly intervene against potential falsely positive and falsely negative signals (Boettiger et al., 2013a; Ditlevsen et al., 2010).

Within the bifurcation-driven CT class, the evolution of statistical indicators associated with EWS might display differing scales: $\zeta(p,t) = f(p^{\alpha})$ with potentially different scaling factors

α for the parameter(s) p, for each indicator ζ (Meisel et al., 2015; Kuehn, 2013). This would in turn affect detection performances in real-world applications. Noise could also be a limiting factor in our capacity to observe meaningful signals. O'Regan et al. (2018) observed that intrinsic noise described by non-Markovian functions might change the expected trends of commonly used indicators; Sharma et al. (2016a) and Kuehn et al. (2022) extended similar observations to the case of multiplicative and coloured noise, suggesting a sort of "colour-blindness" that could impair the extraction of EWS from data. Noise has been recognised as a relevant issue in spatially extended systems, too. A modelling study on PDEs observed that statistic indicators are usually overwhelmed by noise effects (Gowda et al., 2015). In addition, recently explored techniques for dimension reduction of network systems have been shown to be sensitive to noise distribution (Weinans et al., 2019).

In addition to noise, another potentially confounding aspect is the separation of temporal scales: preliminary studies (Kuehn et al., 2011; Brett et al., 2017) suggested that commensurable time scales in the evolution of the main variables and of the control parameter might alter the expected EWS trends and shadow their associated alerts. Within the b-tipping class, further studies are then much needed about the interplay of noise and timescale effects, and how they affect the detection and prediction performance of various early warning signals.

3.2.2 EWS also precedes non-critical transitions

As discussed in Sec. 2.4, early warning signals are measures associated with the behaviour of common statistical indicators such as increasing variance or autocorrelation. It come with little surprise that such trends can be measured even in systems that do not exhibit critical transitions. Kéfi et al. (2013) reviewed transitions that are not "critical" per se but can still be preceded by EWS (e.g. continuous transitions in the first derivative like supercritical pitchfork-driven ones). The reason is that such EWS are associated with the phenomenon of critical slowing down, which also exists in other noisy bifurcations, including second-order ones. Similar remarks, from studies on continuous bifurcation, non-bifurcating models with coloured noise and other examples, have been put forward by recent literature (Boettiger et al., 2012b; Boettiger et al., 2012a; Boettiger et al., 2013a; Dutta et al., 2018). A simple scaling argument is also provided in Appendix A.

These remarks stress the importance of accounting for the likelihood of the system being bistable (or at least prone to transitions) in addition to the identification of a potential signal, in order to minimise false positive signals. An example of such potential fallacy is reported by Kaur et al. (2020), who observed rising variance at the beginning of the first COVID-19 epidemic wave. Nonetheless, epidemic first emergences are novelties followed by exponential growth, not zero-eigenvalue processes: as such, the increasing variance is more likely linked to increasing case numbers poorly coped with scarce tests, rather than signs of a bifurcation being crossed.

3.2.3 Single EWS are often not sufficient to predict abrupt shifts in real time series data

From minimal models to real data: single EWS are often not adequate to reliably predict regime shifts. An extensive survey performed by Burthe et al. (2016) shows that EWS in ecological real data series do not provide clear insights, while the use of EWS for low populations is not

recommended (Brett et al., 2017) as the ergodic hypothesis breaks down without a mean field modelling approach. In this case, there would be extreme sensitivity on individual variability and noise. Hence, scholars have suggested using different EWS at once or combining them for decision making (Biggs et al., 2009)

Combining EWS could potentially expand their detection and/or anticipation prowess. Such combinations can be explored with machine learning approaches for classification (thus solving optimization problems). However, they would necessarily be system-specific and require sufficient amounts of data. The advantage would be an improved interpretation of the machine learning output, in terms of dynamical theory. Examples exploring this possibility include Brett et al. (2020b), trying to anticipate epidemiological data with machine learning, the thesis "Dynamical modeling techniques for biological time series data" by L. Mombaerts, detecting epileptic seizures with combinations of statistical indicators, and Bury et al. (2021), applying deep learning anticipation algorithms, trained on normal forms, to various synthetic and real-world data.

3.2.4 Data curation procedures affect the extraction of EWS

Not only there are theoretical limitations for the immediate application of CT theory on observed data and for the extraction of associated early warning signals; data analysis challenges also add up to the complexity.

Data selection

Aptly selecting the considered data is already important. Complex real-world systems might exhibit transitions over one state-space direction and not on others, meaning that some variables are more likely to track transitions than others (Wang et al., 2012a). In this regard, modelling choices involve considering scalar quantities associated to e.g. phenotypes, networked variables, or anything measurable and subsequently analysed. Depending on the choice, different methods could then be applied, as well as different EWS, each with its own alerting performance. In general, data should derive from those measured quantities associated to variables of CT-prone models.

Studies on epidemiological models provide examples about the effect of selecting among two potentially useful sets of scalar data. Consider a simple, generic and normalised SIS[2] (Susceptible-Infectious-Susceptible) model (Kermack et al., 1927):

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI + \alpha I \\ \frac{dI}{dt} &= \beta SI - \alpha I \,. \end{aligned} \tag{3.2}$$

I represents the infectious compartment, S the susceptible one, α and β are transmission and recovery rates, respectively. In this case, the variable I, which can potentially undergo a transcritical bifurcation (Kuehn et al., 2011; O'Regan et al., 2013) for varying values of the parameter β, is associated with so-called prevalence data, representing active infectious cases over a population.

[2] Analogous remarks and results also hold for SIR (Susceptible-Infectious-Removed) models and related ones, see O'Regan et al. (2013) and Southall et al. (2021).

A popular alternative for epidemic monitoring, on the other hand, considers incidence data, i.e. daily new detected cases, without further following their progression along the disease. Incidence data are a differential quantity derived from prevalence data; instead of being directly captured by the SIS model, their dynamics is often modelled in terms of the underlying Poisson-driven transmission processes (Southall et al., 2021; O'Dea et al., 2019):

$$N_{t+\Delta t} = Poi\left(\int_t^{t+\Delta t} \lambda(s)ds\right) \simeq Poi(\Delta t\lambda(t)) \tag{3.3}$$

where N_t represent the counting process yielding incidence data, Poi is a non-homogeneous Poisson process and $\lambda(t) = T(I+1|I)$ represents incoming transmission probabilities in the hypothesis of no under-reporting. A common form, which links to Eq. 3.2, is $T(I+1|I) = \beta(t)S(t)I(t)$. The last approximation hold for small Δt.

It has been shown (Southall et al., 2020) that EWS calculated on synthetic prevalence or incidence data might display different detection performances. Similar results are observed in our studies on empirical data, described in Sec. 7. In addition, both data types could suffer from different systematic uncertainties (e.g. related to testing or tracking routines), which would consequently alter the associated noise distributions with undesired consequences (see above).

Data detrending

Data detrending methods can play a role, too. Recall that statistical indicators for EWS are derived from fluctuations around stationary equilibria: long-term moving trends thus need to be filtered from real-world time series. After that, variance and other statistics should be calculated on the detrended residuals instead of on the original data. The problem of detrending is a compelling one (Dakos et al., 2015; Dessavre et al., 2019): choosing the most appropriate detrending method (moving average, Gaussian, lowess, or other kernels filtering, ARIMA models) might yield better or worst detection and anticipation performances, depending whether the method is able to pick up spurious fluctuations in the data, long-term patterns and so on. Various authors have so far mostly used their preferred method (e.g. lowess (Bury et al., 2021), moving average (Southall et al., 2020; Kuehn, 2013)) and little consensus has so far been reached.
For a discussion on the mostly employed detrending methods, see Section 4.4; for an example of the impact of detrending method on the value of residuals, see Fig. 3.4.

Moving windows over time series

A procedure to carefully optimise is that of selecting appropriate moving windows to extract statistical indicators from time series data. Standard early warning signals are computed, under the assumption of quasi-steady state eigen-parameter, over stochastic repetitions of the same process, close to the transition. For a system to display analogous features in its time series observation, some caveats apply. First, the system should be ergodic, so that the statistics over the trajectory of a sample point are representative of those of multiple repetitions. Second, that data encompassed in one moving window are representative of such a trajectory. Kuehn et al. (2011) discussed this necessity on abstract models, while its testing on synthetic epidemiological data was performed by Brett et al. (2017); Dessavre et al. (2019). In general, tradeoffs are necessary: a window should be large enough to be representative of the process, but small enough to avoid incorporating past

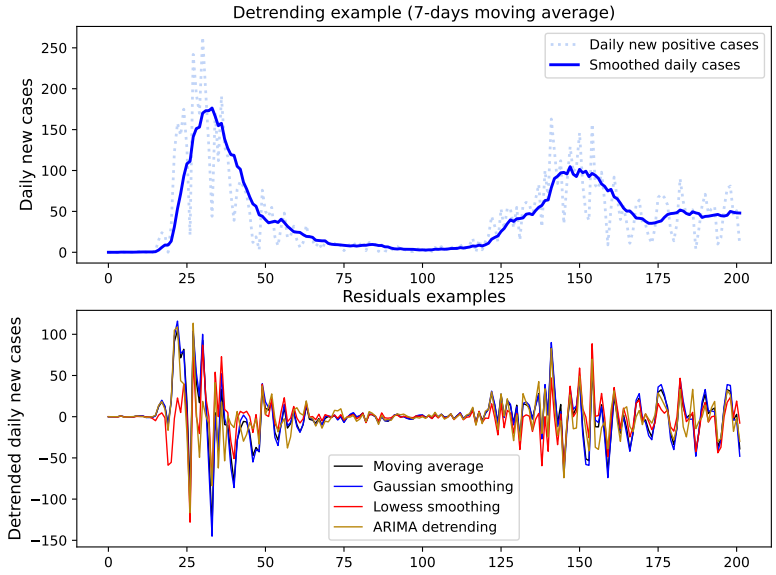


Fig. 3.4 Example of different detrended data, depending on method and kernel. Considered data are daily COVID-19 positively detected cases (incidence, linked to Eq. 3.3 in the discussion above) from late February to late August 2020. **Top:** an example of smoothing, using a commonly employed 7-days moving average. **Bottom**: residuals obtained with four different methods: 7-days moving average, smoothing with Gaussian kernel (non-centered, window size of 7 days, std=2 days), smoothing with *lowess* (fraction to estimate y values corresponding to 7 days), detrending with ARIMA(2,1,5). In the panel, i is possible to observe the slightly different residual values obtain with the various methods, which might induce biases or modified results in the later analysis.

trends unrelated to the approach to the transition and not to be too data-greedy.

One final aspect to recall: centring the window appropriately is necessary to correctly interpret the results. Centring the moving window means: which data point is associated with the corresponding value of a statistical indicator? Depending on the scopes, there might be several answers.

For purely retrospective detection purposes, where the full time series is already provided, centring the window to the central data point might suffice. It might even be desirable, in order to "peek into the future" and provide an unbiased and performing estimate of the transition point. One example for such feat is automatically recognising when an epileptic seizure starts on a fully recorded EEG.

For all other scopes, the window should be non-anticipative, i.e., the centring should be performed on the rightmost point (window "only looking into the past"). Studies focused on retrospectively searching for CSD-related signals could thus focus on what happens before the transition rather then across it, while research on anticipating CTs would properly mimic the data stream happening during on-line monitoring, which is obviously unaware of future values. Nonetheless, this practice is associated with two limitations. First, enough past data are necessary before being able to anticipate anything (or the window would be too short). Second, the same window should be used for smoothing and for centring the statistical indicators: hence, both processes should be handled with great care not to induce biases. In addition, the process might

reduce the lead time for prediction as it needs to wait for some extra data to be available. See Fig 3.5 for an example.

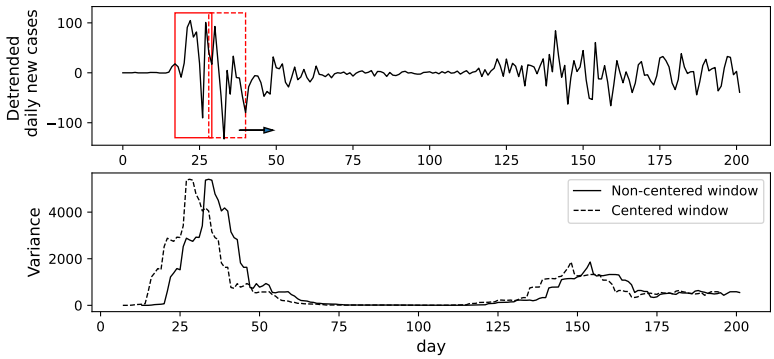


Fig. 3.5 Example of moving window to estimate the statistical indicators over a time series of detrended data. Depending if te window is centred or not, the associated value of the indicator might be shifted (blue and red pointer, respectively), with consequences on the interpretation (is the analysis non-anticipating or not) and on the studies about the lead time.

3.3 Inferring modelling frameworks from data: debates

Inferring systems' properties and identifying appropriate modelling of unknown complex systems is challenging and proceed with te accumulation of evidence and hints. Although such hints are suggested by theoretical considerations, debates might progress, depending on available data, the output of analysis procedures and theoretical interpretations. This section briefly illustrates some ongoing examples relative to critical phenomena.

Power laws are ubiquitous in nature and are associated with complex dynamics with long correlation times. They can be derived from self-organised criticality SOC models (Bak, 1987) and are often used to infer self-organisation in observed dynamical systems (Osorio et al., 2010; Sornette, 2013). Other explanations are, however, possible. Boffetta et al. (1999) analysed power laws between successive bursts of solar flares. They compared the observed scaling behaviour with that predicted by SOC models and by turbulence models, concluding that the latter reproduces the observations and thus has better explanatory power. The work spurred interest in modelling approaches targeting critical states and producing power laws. Furthermore, Dickman et al. (2000) noticed that: "is often hard to distinguish between SOC-like behaviour and other mechanisms for generating power laws. This task appears almost hopeless in situations where only limited data sets are available".

A similar debate is ongoing in neuroscience, whether self-organised criticality is a fundamental property of neural systems or something different is more likely to happen. Hypotheses and evidence for both stances are accumulating. Supporters of the SOC paradigm (Hesse et al., 2014; Brochini et al., 2016; Cocchi et al., 2017) bring about hypotheses, modelling results and observational evidence to strengthen their view. Other authors (Touboul et al., 2010; Touboul et al., 2017) observe that similar power laws can arise in the presence of turbulence or stochastic

dynamics, suggesting that power-law statistics are not sufficient to establish criticality. In addition, they argue that such power law heavily depends on the choice of the threshold that is used to define a neural spike: if such definition is varied, different distributions would possibly arise. As the brain is, arguably, more complex than solar activity, and data are more prone to uncertainties, the debate is likely bound to last for long.

Akin to questions about criticality in overall brain activity, another debate is: are epileptic seizures – sudden and uncontrolled electrical disturbances in the brain, see Fig. 3.6 – related to critical activities and self-triggering capacities? Not only this question regards the fundamental mechanisms of brain regulation (Stefanescu et al., 2012; Jirsa, 2014; Gollas et al., 2005; Mandelj et al., 2002; Ashwin et al., 2016); it also has practical implications for the predictability (including feature selection) and management of such a harmful condition (Navarro et al., 2002; Carney et al., 2011; Omidvarnia et al., 2018; Mormann et al., 2000; Schiff et al., 1994).

Many authors have recognised that epileptic activity is characterised by nonlinear dynamics at multiple time- and spatial levels and with long time- and space correlations (Parish et al., 2004; Dominguez, 2005; Meisel et al., 2012b; Aur, 2010; Chouzouris et al., 2018; Gal et al., 2010). However, it is not clear which macro-level theoretical framework – if any on its own – might better accommodate those observations. A number of scholars have suggested that epilepsy can be understood in terms of self-organised criticality of brain activity (Osorio et al., 2010; Osorio et al., 2009; Gal et al., 2013) or failure thereof (Meisel et al., 2012a). Slowly varying parameters through bifurcations or routes to chaos (Velazquez et al., 2011; Jirsa, 2014; Breakspear et al., 2006; Hoppensteadt et al., 1998) have also been called upon to tackle the issue, as well as the critical transitions framework (Kramer et al., 2012; Meisel et al., 2015; Chu et al., 2017). However, it is not clear whether the expected signals associated with critical slowing down can be reliably measured in EEG (Maturana et al., 2020) or not (Wilkat et al., 2019). Moreover, some spectral features observed in EEG (Chu et al., 2017) can be both related to noisy bifurcations or to turbulent travelling waves (Kuramoto, 1984). Future work, for various epilepsy classes (Aarabi et al., 2014), will build on micro-scale models of coupled oscillators to study their routes to anomalous synchronization (Pikovsky et al., 2001; Rodrigues et al., 2014; Boccaletti et al., 2006; Terry et al., 2012) or on travelling waves in neural excitable media (Ermentrout et al., 2001; Rubino et al., 2006; Balázsi et al., 2003; Winfree et al., 1984). Advanced data analysis will also be an asset, e.g. to distinguish noise from chaos (Rosso et al., 2007) and to refine power law scaling estimates. In the meanwhile, controversies keep the field moving (Frei et al., 2010).

Finally, an example from climatology illustrates the importance of inferring whether rapid transitions are triggered by stochastic disturbances or slow changes in environmental conditions. The Dansgaard-Oeschger events are a series of 25 rapid climate fluctuations that occurred during the last glacial period (Fig. 3.7). Various mechanisms have been proposed to explain their occurrence (Petersen et al., 2013), including fluctuations of the Atlantic meridional overturning circulation (AMOC) (Sima et al., 2004), corroborating hypothesis for their bistability ("sweet spots" for abrupt climate changes). However, it is still not clear whether rapid shifts between alternate states are triggered by critical slowing down mechanisms or by random occurrences (or by a combination of both). In fact, although several studies have looked for early warning signals associated with CSD (Livina et al., 2010; Lenton et al., 2012; Cimatoribus et al., 2013), others observe that the statistical distribution of the events would match Poisson processes (Ditlevsen

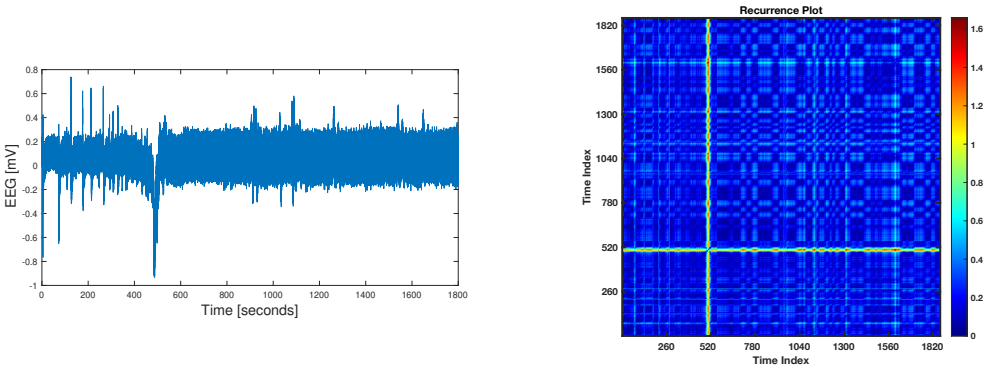


Fig. 3.6 **Left:** EEG recording (2000 Hz sampling frequency) of a Zebrafish larva's brain activity. The large spike is a clear symptom of an epileptic seizure triggered by PTX seizure-inducing drug. Data were collected by A. Oldano from the Integrative Cell Signalling partner group at the LCSB (headed by prof. Alex Skupin) and processed by L. Mombaerts in his thesis "Dynamical modeling techniques for biological time series data". **Right:** Recurrence plot (see Sec. 4.3.5) of the EEG signal, downsampled to one point every 1000 for calculation efficiency of the algorithm from Yang (2010). Colour scale reflects the euclidean distance between points in the embedding space. The vertical ad horizontal stripes clearly identify the epileptic spike. Other tartan-like patterns across the plot suggest additional trends (autoregressive, chaotic or of other kinds) that are likely not pure stochastic noise.

et al., 2007) and that, in general, no signals can be reliably recognised (Ditlevsen et al., 2010). Potentially, future studies based on more refined data-driven techniques (Bury et al., 2021) will contribute to mechanistic models in better interpreting paleoclimate data.

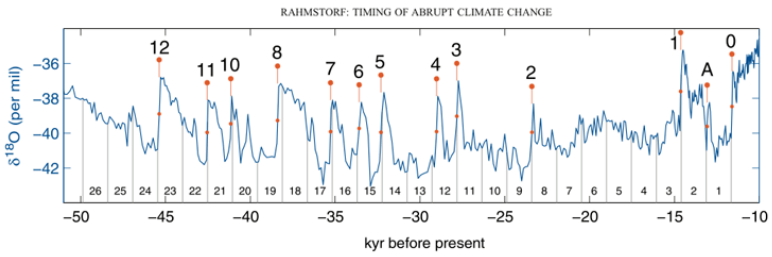


Fig. 3.7 Time series data from ice cores samples, displaying the oscillations occurred during the Dansgaard-Oeschger cycle. Figure from Wikimedia Commons, S. Rahmstorf, 2003. Geophys. Res. Lett. 30 1510-1514 (Creative Commons licence).

3.4 The "prosecutor's fallacy"

The prosecutor's fallacy is termed after courtroom misapplications of statistical reasoning. Famous cases associated to it are Sally Clark's[3], *People v. Collins 1968*[4], and Lucia de Berk's[5] (to name a few). The fallacy is associated with various misinterpretations of Bayes' theorem, describing the probability of an event, given prior knowledge of conditions related to that event:

$$P(A|B) = P(B|A)\frac{P(A)}{P(B)} \tag{3.4}$$

A typical example of the prosecutor's fallacy is ignoring the scaling on the conditional probabilities and equating the probability of being guilty, given the evidence, to the probability ob observing the evidence, given the person is guilt. On the contrary, one should also consider prior knowledge about the possibility of that person being really guilt (e.g. collected from other hits, trials or social conditions) and of finding similar evidences for other reasons (chance, other occurrences producing similar outputs and so on).

As EWS are symptoms of critical transitions, but not unique features to them, Boettiger et al. (2012a) elegantly illustrated how the fallacy might apply when interpreting signals extracted from data close to critical transitions. They reminded that the attempts to detect early warning signs for critical transitions follow from the assumption of a slowly changing parameter, which is a different class than stochasticity-driven transitions (as described in Sec.2.3.2). The signals associated to different transitions are also different. Without knowing *a priori* which category the transition belongs to, the risk of false positives increases. On the other hand, state-dependent noise or other stochastic effects might modify the trends of summary statistics, so that they might be confused for early warning signals. In this case, the fallacy argument pertains the – potentially dubious – inference of the existence of alternative states, given the prior observation of a transition and the measurement of EWS-like statistical trends.

An analogous problem also arises when using statistical trends, theoretically associated to early warning signals, to infer the approach of a critical transition. Using Eq. 3.4,

$$P(\mathrm{CT}|\mathrm{EWS}) = P(\mathrm{EWS}|\mathrm{CT})\frac{P(\mathrm{CT})}{P(\mathrm{EWS})}\,. \tag{3.5}$$

$P(\mathrm{EWS}|\mathrm{CT})$ is what's normally computed when assessing sensitivity and specificity in computational studies and constitutes the very first step towards real-world applications. However, it might not be immediately equal to $P(\mathrm{CT}|\mathrm{EWS})$ (the gold standard for monitoring): depending on the system, $P(\mathrm{CT})$ might be small thanks to redundancy protocols or design, while $P(\mathrm{EWS})$ might be large, e.g. due to state-dependent noise. The other way can also happen: $P(\mathrm{CT})$ being large

[3] In 1998, S. Clark was wrongly found guilt of murdering her two sons. The prosecution case relied on flawed statistical results provided by paediatrician prof. Meadow (https://en.wikipedia.org/wiki/Sally$_{\mathrm{C}}$lark).

[4] An American robbery trial, which became famous for misinterpretation of conditional probability and raised questions about the use of mathematics in courts, that tend to host few mathematically skilled attorneys (https://www.courtlistener.com/opinion/1207456/people-v-collins/ and Boettiger et al. (2012a))

[5] Lucia de Berk was mistakenly convicted for several murders, on the basis of the very low sheer probability of a nurse shift coinciding with many natural deaths in hospitals. The value was wrongly calculated (https://en.wikipedia.org/wiki/Lucia$_{\mathrm{d}}$e$_{\mathrm{B}}$erk)

due to stochastic disturbances and $P(\text{EWS})$ being small because the system has rapidly changing parameters or is prone to multiplicative noise. This creates ambiguity in case the dynamical context of the system is not appropriately known and calls for complementary studies to reliably assess the likelihoods of observing critical transitions in the real world.

3.5 Connecting my contribution to the open challenges

Unexpectedly large shifts may occur in dynamical complex systems. Theoretical results suggest that such shifts can be attributed to critical transitions between alternative stable states. Moreover, theory suggests that such transitions can be associated to specific behaviours of statistics indicators, that can serve as early warning signals during monitoring procedures. However, proposed EWS are not unique hallmarks of critical transitions and they are not sufficient nor necessary conditions for impending regime shifts (Boettiger et al., 2013a; Boettiger et al., 2013b). The blind application of EWS onto unknown systems has already been recognised as a potential source of incorrect conclusions (Boettiger et al., 2012b). Instead, EWS are mathematical constructs that are expected to provide meaningful information and good diagnosis if applied and interpreted appropriately, within each class of critical transitions, with respect to their underlying modelling assumptions.

To properly estimate the sensitivity and specificity of various EWS (and minimise the cases of prosecutor fallacies, *cf.* Sec. 3.4), a systematic analysis of the diverse dynamical contexts and tipping regimes is thus required. Such analysis is anticipated by the "taxonomical" classification of CTs illustrated in Sec. 2.3.2 and further carried out in Chapter 5 (from a theoretical perspective, later applied onto biological models and data, and in Chapter 7 (testing theoretical predictions and modelling assumptions on data from the COVID-19 pandemic).

The challenges described above and the subsequent analysis illustrated in Sec. 2.2 show that EWS can hardly play the role of universally applicable, "model free" indicators, nor are they forecasting tools per se (Dakos et al., 2015). Moreover, as current EWS are associated with local minimal models (see again Sec. 2.2), they are "weak" when it comes to quantitatively predict regime shifts. As a consequence, scholars have suggested not to look for specific thresholds in their values, but to consider them as "resilience indicators" to rank systems according to their fragility (Scheffer et al., 2015) or as "first pass tools" to quickly identify possibly-at-risk situations, to be then analysed with more advanced techniques (Clements et al., 2018). In this regard, Chapter 6 discusses how some biological mechanisms are able to buffer the variability next to critical regimes and quantifies how trends potentially associated to EWS may subsequently vary. Finally, Chapter 8 presents a refined model-driven method to track epidemic resurgence using complementary data, and the lessons learned.

Chapter 4

Models and methods to analyse critical systems

This chapter can be ideally subdivided into macro-blocks. To begin with, it presents two minimal models that will be used in the thesis and that retain several key characteristics observed in analysed data. Then, it discusses computational methods for simulation and analysis of stochastic dynamical systems. Later, it presents data analysis methods, useful for both synthetic and empirical data, including the computation of early warning signals and the performance assessment. Finally, it provides a brief review of techniques for multi-dimensional systems, to complement those mostly focusing on low dimensions.

4.1 Minimal models

The two minimal models described below are paradigmatic of critical transition behaviours in systems biology and epidemiology. They are used in the thesis for examples and illustrations. They will also be subject of deeper analysis in later chapters.

4.1.1 An illustrative biological dynamical model

Many biological systems self-regulate their functions through bistable circuits (Angeli et al., 2003). In particular, the action of positive feedback loops in regulating bistable cellular states have long been studied in systems biology (Smits et al., 2006; De Mot et al., 2016; Alon, 2019). They are recognised as key dynamical processes through which cells control fundamental functions like enzymatic activity or gene transcriptional changes during cell fate decisions (Hasty et al., 2000; Huang et al., 2007; Fiorentino et al., 2020). Understanding such control mechanisms is an active area of systems and synthetic biology (Acar et al., 2005; Veening et al., 2008; Pisarchik et al., 2014; Guinn et al., 2020). Autoactivating positive feedback loop, simple circuit motifs promoting bistability and fine regulation of dynamical states close to self-organised criticality, are particularly relevant (Frigola et al., 2012; Tyson et al., 2003). In these regimes, cellular heterogeneity, *i.e.*, random cell-to-cell variations (Komin et al., 2017) can further direct the transition (Kaern et al., 2005; Weber et al., 2013) and lead to regime shifts between alternative stable states of gene expression or of protein concentration. Some cells may, for example, live in the “off” expression state of a particular gene, whereas others are in the “on” expression state (Thomas et al., 2014).

Given the abundance of organisms self-regulating via positive feedbacks while living in noisy environments, understanding stochastic bistable regulation is a hot sub-topic in the field of systems biology and control (Lestas et al., 2008; Lestas et al., 2010; Norman et al., 2015). Known examples of stochastic positive feedback loops have been observed in the transcription networks of *E. coli*, (Milo et al., 2002) and in the regulation of β-*galactosidase*, which results from a sudden transition from low to high level states of the *lac operon* at a critical point of an inducer concentration (Ozbudak et al., 2004). Other mechanisms for stochastic switching exist; many of them can be immediately captured by positive feedback models (Norman et al., 2015).

A minimal rate-equation model of gene regulation, incorporating a closed-system auto-activating feedback loop, is the following (Strogatz, 2018; Alon, 2019):

$$\frac{dx}{dt} = K + c\frac{x^n}{1+x^n} - x + g(x)\eta(t)\,. \tag{4.1}$$

The (dimensionless) stochastic equation describes the dynamics of protein production, after RNA transcription. x models the protein concentration, K represents the basal expression rate, c the maximum production rate, n is the Hill's cooperativity index (Santillán, 2008) and the dissociation constant was normalised throughout the equation. The noise term $\eta(t)$ (integral representation of a Wiener process) has the following statistical properties: $\langle\eta\rangle = 0$, $\langle\eta(t)\eta(t')\rangle = 2\sigma\delta(t-t')$, where σ represents the noise intensity. $g(x)$ can be a multiplicative function to make the noise state-dependent. Here, only additive noise ($g(x) = 1$) is considered for introductory purposes. Forms of multiplicative noise and their interpretation will be provided in subsequent chapters.

A detailed derivation and study of Eq. 4.1 is provided by Weber et al. (2013) and Sharma et al. (2016a). A general overview is the following. Consider a circuit (sketched in Fig. 4.1) where a single gene activates, after a stream of transcription and translation processes, its own transcription factor TF. The TF is a regulatory protein, binding to DNA promoters in operator sites and constructing working homodimers. The circuit is controlled from outside stimuli because transcription is activated only by phosphorylated TF dimers; dimer phosphorylation depends on the activity of kinases and phosphatases, which bridge external control signals.

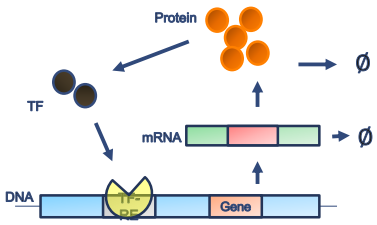


Fig. 4.1 Schematic picture of an autocatalytic positive feedback loop. Expression of the gene leads to the production of proteins that, directly after oligomerisation or mediated by the action of transcription factors, bind to their own promoter (in yellow), acting as a self-activator. In biological circuits, the expression of the protein results from mRNA translation, which is not considered in the simplified model as its action on the dynamics is mainly to tune time delays (Strogatz, 2018). Degradation of proteins and mRNA are denoted by the slashed circles.

There is a natural time-scale separation between fast reactions – binding and mRNA degradation, for instance, happening in seconds – and slow transcription. Fast reactions can thus be set at equilibrium, while slow reactions are here considered irreversible (Hasty et al., 2000). The mass

action kinetics can thus be written as (Sharma et al., 2016a):

$$\emptyset \underset{k_{\text{deg}}}{\overset{RV}{\rightleftharpoons}} X, \quad D + nX \underset{k_2}{\overset{k_1/V^n}{\rightleftharpoons}} D^*, \quad D^* \xrightarrow{k_p} D^* + X, \tag{4.2}$$

where X is the TF protein, D and D^* are the unbound and bound state of the promoter, $\emptyset$ denotes degraded states (obtained with rate k_{deg}), n denotes cooperativity in binding (how many TF need to cooperate to activate the promoter), R is the basal expression rate, V is the cellular volume, $k_{1,2}$ are binding and unbinding rate constants, and k_p is the protein production rate. For low copy numbers, individual jump processes should be considered and a discrete-state Master equation developed (Gardiner, 1985). However, if we consider a high concentration of cells, therefore a high number of relevant molecules, we can consider concentrations as dynamical variables[1]: $\{x,d,d^*\} = [\{X,D,D^*\}]$. The rate equation can thus be written as (Weber et al., 2013):

$$\frac{dx}{dt} = R + k_p(d+d^*)\frac{x^n}{k_2/k_1 + x^n} - k_{\text{deg}}x\,.$$

The single equation is achieved by assuming quasi-steady-state for the fast variables (see above and Frigola et al. (2012)). Finally, Eq. 4.1 is obtained by lumping some rate parameters, setting $K = R/(k_{\text{deg}}\sqrt[n]{K_d})$ and rescaling $x \leftarrow x/\sqrt[n]{k_2/k_1}$ and $t \leftarrow k_{\text{deg}}t$ into dimensionless variables.

The autoactivating feedback loop is bistable for values $0 < K < 1/3\sqrt{3}$ and monostable for other values (cusp-like behaviour, *cf.* Sec. A), provided that $n \geq 2$. $n \to \infty$ yields the logic approximation for the activating Hill function,

$$\lim_{n\to\infty} \frac{x^n}{1+x^n} = \Theta(x-1)\,,$$

where $\Theta(\cdot)$ is the Heaviside step function that renders system (4.1) a perfect toggle without bistability nor CSD. For many applications studying bistability in the system, n is therefore set to 2 without loss of generality (Sharma et al., 2016a; Weber et al., 2013).

Its vector field under the influence of c was shown in Fig. 2.4, along with a visual example of its topological equivalence to a saddle-node normal form close to the bifurcation point. The other critical point is also a saddle-node (Strogatz, 2018; Sharma et al., 2016a), see also the bifurcation diagram in Fig. 4.2.

4.1.2 Variations on the SIR epidemiological model

The description of diffusing epidemics inside of a populations has a long story dating back to the '20s (Kermack et al., 1927). Depending on the disease characteristics (time scale, infectiousness, incubation time, possibility of re-infection and so on), refined models have been developed over the decades, including distinctions between population-based models and individual (or agent) -based models (Frias-Martinez et al., 2011).

[1] This assumption is necessary to interpret the white noise term in Eq. 4.1 as a stochastic alteration of the "background" production of TF. Alternatively, one can use e.g. the Gillespie (2000b) formalism to derive a multiplicative noise term for the associated Langevin equation.

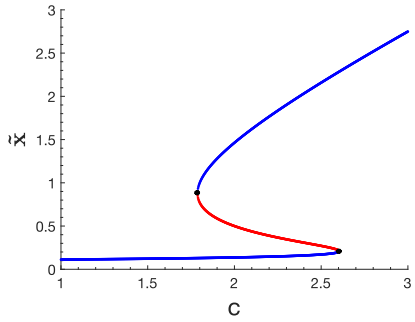


Fig. 4.2 **Left:** bifurcation diagram of Eq. 4.1 for $n = 2$. Blue lines are stable branches, red line marks the unstable manifold. Black dots highlight the fold points. Following one branch until the bifurcation point and coming back by following the other branch marks a hysteresis loop.

One paradigmatic minimal model is the continuous-time, mass conservative and compartment-based SIR model (Anderson et al., 1979), see Fig. 4.3, top. "SIR" stands for Susceptible → Infectious → Removed compartments. It assumes a homogeneously mixing population (or fully connected networks of individuals) and addresses the evolution of mean properties of the closed epidemiological system. The basic (deterministic) SIR is:

$$\begin{cases} \dot{\mathrm{S}} = -\frac{\beta \mathrm{SI}}{N} , \\ \dot{\mathrm{I}} = \frac{\beta \mathrm{SI}}{N} - \gamma \mathrm{I} , \\ \dot{\mathrm{R}} = \gamma \mathrm{I} , \end{cases} \tag{4.3}$$

with $S+I+R = N = \text{const}$ (N is the population size). Technically, this condition makes the third equation in (4.3) redundant. In this interpretation, S is the pool of individuals socially active and at risk of infection; I are individuals having developed the disease and being contagious; R lumps those that have processed the disease, being either recovered or dead. Transmission parameters from one compartment to another are β, the average contact rate, and γ, the inverse of mean contagious period. When focusing on infection dynamics rather than disease progression in patients (Kemp et al., 2021), the latter combines recovery and death rate (Syafruddin et al., 2012).

The SIR model can be subsequently enhanced to more realistic versions, a relevant one being the SEIR (Fig. 4.3, bottom) and its numerous extensions, see e.g. Yan et al. (2006); Arino et al. (2005); Gatto et al. (2020); Giordano et al. (2020); Sjoedin et al. (2020). Main feature of SEIR models is the inclusion of an "exposed" compartment E, accounting for viral incubation in individuals who have been infected but become contagious after some time (latent carriers). The flux to E is governed by an extra parameter α, the inverse of mean incubation period, whereas its outflux is dependent on γ as before. For their adequacy in describing epidemic diffusion after latency periods, SEIR models have been especially employed for the modelling of the COVID-19 pandemic, see Currie et al. (2020); Lai et al. (2020); De-Camino-Beck (2020); Anand et al. (2020); Anderson et al. (2020); Proverbio et al. (2021); Kemp et al. (2021) and many others.

For SIR models and their extensions, the essential control parameter is the basic reproduction number R_0, which quantifies the average number of new cases that an infectious individual might provoke, at the beginning of the pandemic (Legrand et al., 2007). For the basic versions of SIR and SEIR models, $R_0 = \beta/\gamma$ (Fraser et al., 2009). For more complex models, it is sufficient

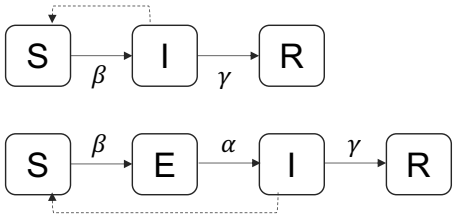


Fig. 4.3 **Top:** scheme of the SIR model. S, I and R are compartments, β and γ the rate parameters governing the transition among compartments. **Bottom:** scheme of the SEIR model, with its additional compartment E with influx rate α. Dashed arrows represent the feedback interaction between the S and I compartment, triggering further infections. More complex extensions can be found e.g. in Kemp et al. (2021).

to consider the eigenvalues of the transmission matrix to derive it (Diekmann et al., 2010; van den Driessche P., 2017; Blackwood et al., 2018). During the epidemic progression, isolation after diagnosis, vaccination campaigns and active mitigation measures are in action. Hence, the reproduction number may vary in time and we speak of "effective reproduction number" $\hat{R}(t)$ (Althaus, 2014). Estimating $\hat{R}(t)$ has practical purposes as a quantitative proxy for the main stressors on the healthcare system (Giordano et al., 2020; Sjoedin et al., 2020; Ferguson et al., 2020) and on the economic system (Alvarez et al., 2020; Atkeson, 2020; Altig et al., 2020; Burzyński et al., 2021).

Moreover, $\hat{R}(t)$ acts as a bifurcation parameter for the equilibria of epidemic dynamics: $\hat{R}(t) = 1$ marks a transcritical[2] bifurcation from disease exponential diffusion $\hat{R}(t) > 1$ to disease elimination $\hat{R}(t) < 1$ (O'Regan et al., 2013; Kuehn et al., 2011). Fig. 4.4 shows an example for a SEIR model from Proverbio et al. (2021): the fixed parameter β is tuned by ρ, with varying values, possibly over time. Consequently, $\hat{R} = \beta\rho/\gamma$. For values of ρ such that $\hat{R}(t) > 1$, we observe an exponential epidemic diffusion (which eventually fades due to depletion of the S compartment); when $\rho = \rho_i$, s.t. $\hat{R}(t) < 1$, the disease is immediately eliminated (Fig. 4.4, left). Fig. 4.4 (right) shows the bifurcation diagram between the equilibrium values at peak (both for I and I+E curves) versus ρ.

Interpreting epidemic diffusion in the critical transitions framework, in search for alerting signals of regime shift to complement the monitoring indicator toolboxes, is therefore an active area of research (Brett et al., 2017; Brett et al., 2018; Dessavre et al., 2019; O'Regan et al., 2016; Brett et al., 2020b; Southall et al., 2020). Likewise other system, it is paramount to verify the relevance of the slow-fast system hypothesis (see Sec. 2.2.4) and to properly consider the noise properties associated with the considered models, in order to assess and interpret the performance of EWS in real-world scenarios. Both tasks will be performed in dedicated chapters (7 and 8).

4.2 Simulate and analyse noisy dynamical systems

This section provides an overview of computational methods and resources to support the analysis, investigation and simulation of stochastic dynamical systems.

[2]Transcritical bifurcations emerge from minimal models. Additional compartments, coupling and dynamics on networks may yield saddle-node or backward bifurcations, or even richer dynamics (Zhang et al., 2016; Zarei et al., 2019; Widder et al., 2016; Horstmeyer et al., 2018; Herrera-Valdez et al., 2011).

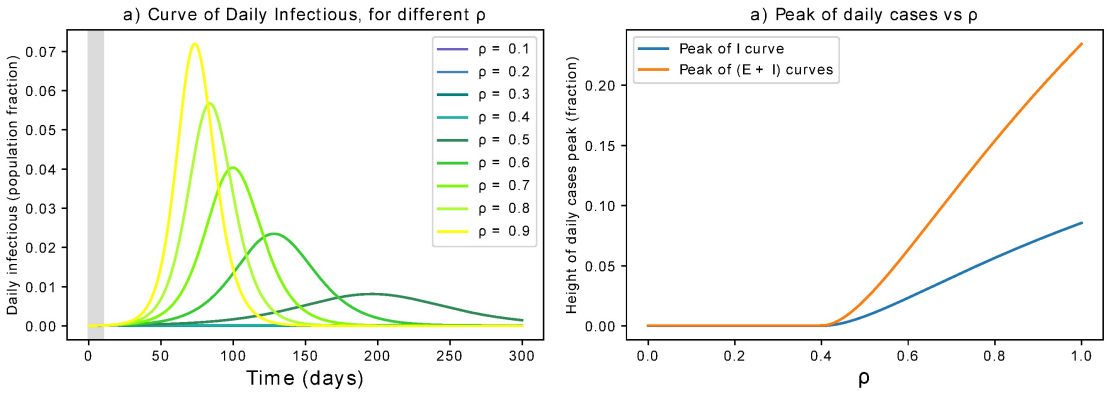


Fig. 4.4 **Left:** evolution of curves of infectious individuals I in a SEIR model, for different values of ρ (tuning of β, such that $\beta \to \rho\beta$ in Eq. 4.3). The grey area indicates the initial simulation phase, with $\rho = 1$. **Right:** equilibrium values at peak (both for I and I+E curves) versus ρ. In this case, $\rho = 0.4$ is a critical value, yielding $\hat{R}(t) = 1$ (the dependency over time is explicit only if $\rho = \rho(t)$, at least under slow-fast assumption). Other parameters values are fixed to COVID-19 relevant ones: $\beta = 0.85\,day^{-1}$; $\alpha = 0.2\,day^{-1}$; $\gamma = 0.34\,day^{-1}$, so that $R_0 = 2.5$ (Liu et al., 2020b; Wu et al., 2020). Refer to the original publication (Proverbio et al., 2021) for further details.

4.2.1 Bifurcation analysis

To determine bifurcation existence and properties, analytical methods check for critical values of control parameters and stability of the corresponding equilibria (see Appendix A). Graphical methods to inquire the behaviour of dynamical vector fields under the action of relevant parameters can also be employed (Sec. A). In addition, computer algorithms can support the analysis.

Custom algorithms can be programmed to repeat the steps of identifying equilibria and check their local stability. Robust and efficient softwares, also providing additional insights on parameter dependencies and branch analysis, are also available for the community:

- XPP/XPP-Aut is a program running under X11 and Windows to solve a vast array of differential equations and perform branch analysis. It is freely available here: http://www.math.pitt.edu/~bard/xpp/whatis.html.
- MATCONT is a continuation and bifurcation software that primarily comes as a toolbox for MATLAB and is powered by a GUI (Dhooge et al., 2003; Meijer, 2014; Ravnås, 2008).
- The Julia language bifurcation kit (https://rveltz.github.io/BifurcationKit.jl/dev/) recently got upgraded with a branch switching continuation library. It is available here: https://rveltz.github.io/BifurcationKit.jl/dev/branchswitching/. The automatic bifurcation diagram computation (https://rveltz.github.io/BifurcationKit.jl/dev/BifurcationDiagram/) is still under development at the time of writing.

4.2.2 Simulations

For numerical integration of deterministic systems (ODE), robust methods have been implemented in numerous software libraries. The simplest one is the first-order Euler method. Its basic idea is to approximate a continuous curve for the dynamical evolution with a series of small steps,

computable by finite-state machines. It truncates the Taylor expansion for a function at first order and employs a time discretization in any desired time interval $[0,T]$:

$$\begin{aligned} y_{n+1} &= y_n + hf(t_n, y_n) \\ t_n &= t_0 + nh\,. \end{aligned} \tag{4.4}$$

Here, every new value at step $n+1$ for the variable y is iteratively computed, updating its past value with a small increment of its evolution law. $h \ll 1$ is the size of every small step. It should be chosen such that the resulting simulations are independent to its variations (they are stable, Bathe et al. (1972)). The Euler scheme is rather basic. Further refinements have been developed over the years, including the 4th order Runge-Kutta integration methods and other correction methods to consider stiff, delayed, ... ODEs.

Numerical simulations of stochastic differential equations (SDE) derive from the original idea of the Euler method, of iteratively solving a truncation of an adequate expansion. The two most employed methods are the Euler-Maruyama and the Milstein (Bayram et al., 2018; Zahri, 2014), both based on running Monte-Carlo chains over truncated Ito-Taylor expansions. For the following explanation, consider a one-dimensional SDE in the form[3]:

$$\begin{aligned} dz &= f(t,z)dt + g(t,z)dW(t) \quad t_0 \leq t \leq T \\ z(t_0) &= z_0 \end{aligned} \tag{4.5}$$

The Wiener process dW satisfies the conditions of $W(0) = 0$, each increment is independent on distinct time intervals, and $W(t) - W(t') \sim \sqrt{|t-t'|}\mathcal{N}(0,1)$ (see Sec. A for further details). The Itô lemma (Gardiner, 1985; Bayram et al., 2018) allows to expand Eq. 4.5 as:

$$\begin{aligned} z(t) = xz(t_0) + f(z(t_0)) \int_{t_0}^{t} d\tau_1 + g(z(t_0)) \int_{t_0}^{t} dW(\tau_1) + \\ + \frac{1}{2} g(z(t_0)) g' z(t_0) \left\{ [W(t) - W(t_0)]^2 - (t - t_0) \right\} + rem \end{aligned} \tag{4.6}$$

Here, *rem* is the remainder containing higher orders of expansion. The numerical integration scheme employs a time discretization of the time interval $[0,T]$ into N steps of size $\Delta t = T/N$:

$$z(t_i + 1) = z(t_i) + f(z(t_i))\Delta t + g(z(t_i))\Delta W_i + \frac{1}{2} g(z(t_i)) g'(z(t_i)) \left[(\Delta W_i)^2 - \Delta t \right] + rem\,. \tag{4.7}$$

$\Delta W_i = W(t_i + 1) - W(t_i)$, for each $i = 0, 1, \ldots N-1$, are independent random variables from a distribution $\mathcal{N}(0, \Delta t)$. Initials conditions $z(t_0) = z_0$ should be specified.

As anticipated, the Euler-Maruyama scheme truncates at first order, thus requiring:

$$z(t_i + 1) = z(t_i) + f(z(t_i))\Delta t + g(z(t_i))\Delta W_i\,. \tag{4.8}$$

[3] Just here, z is used to distinguish between a generic stochastic variable and the deterministic variable y in Eq. 4.4. Later in the thesis, the usual x will be mostly employed.

The Milstein method considers the second term as well, resulting in:

$$z(t_i+1) = z(t_i) + f(z(t_)\Delta t + g(z(t_i))\Delta W_i + \frac{1}{2}g(z(t_i))g'(z(t_i))\left[(\Delta W_i)^2 - \Delta t\right];. \tag{4.9}$$

It better converges to the true Itô integral and was proven to have improved accuracy (Bayram et al., 2018). A gentle introduction to the method is provided on https://hautahi.com/sde$_s$imulation. Note that, when $g(z(t)) = const$ (only additive noise without state-dependency), the two methods are equivalent.

4.2.3 Fitting

Fitting a dynamical model to data essentially means finding the best parameter values such that the model reproduces the observations. To this end, many methods are available (Schittkowski, 2002). Here, two of them – used in publications – are briefly reviewed.

Least-square method

Fitting a single parameter formally means minimising the distance between simulated values and measured values. Such distance can be defined in terms of several functions, of which the root mean square is the most employed for practical and theoretical reasons (Taylor, 1997). In the example of a single parameter q, its best fit value $\hat{q}$ is such that

$$\hat{q} = \left\{ q' \mid RMS = \min_{q'} \sqrt{\frac{\sum_{i=1}^{N}(x(i)-\hat{x}(i))^2}{n}} \right\}. \tag{4.10}$$

$\hat{x}$ and x are, respectively, observed and model-predicted data, and N is the number of considered points.

Building on this, several libraries have been developed for various languages. Python hosts the *lmfit* library for nonlinear curve fitting; MATLAB has a Curve Fitting toolbox available and many other resources.

Bayesian Markov-Chain Monte-Carlo

A more advanced fitting technique is based on Bayesian inference implemented by means of a Markov Chain Monte Carlo (MCMC) approach. It allows to identify the probability distribution associated to each parameter, therefore computing their most likely values and uncertainties, as well as the constraints on their combinations in the parameter space. It is thus more suited in case of multiple parameters, whose identifiability (Gallo et al., 2021) is maybe uncertain.

First, defining a likelihood function for the MCMC is necessary. A common one is the sum of square residuals (SSR) like above::

$$SSR(\overline{\tau}|data) = \sum_{i=1}^{N}\left[x^{data}(i) - x^{model}(i,\overline{\tau})\right]^2. \tag{4.11}$$

$\overline{\tau}$ is the array of parameters, $i \in [1,N]$ counts the data points, x^{data} and x^{model} are measured and simulated data, respectively. The SSR is related to the reduced χ^2, both being ways to measure the

geometric distance between model simulation and data. In addition, a prior probability distribution, which will be iteratively updated into a posterior following Bayes' rule over the chain, should be included. Unless specific priors are known for some parameters (e.g. because of certain constraints or previous studies), flat priors are common choices.

Several sampling techniques can be employed to compute the chain and propose values to the Monte-Carlo sampler. They need to converge fast and to be robust. One employed in two recent publications (Proverbio et al., 2021 and Kemp et al., 2021) is the Delayed Rejection Adaptive Metropolis (DRAM) scheme (Haario et al., 2006), from the python package *pymcmcstat* (Miles, 2019). It combines two powerful ideas: adaptive Metropolis samplers – which adapts the covariance matrix of the proposal Gaussian distribution at specified intervals – and delayed rejection – which delays rejection by sampling from a narrower distribution. Robustness is ensured by running several chains in parallel for thousands of iteration steps (after some burn-in period). Convergence of the chain can be assessed by diagnostics such as the Gelman-Rubin's (Gelman et al., 1992; Brooks et al., 1998). It analyses variances within each chain set and between each chain set and returns values of the so-called "Potential Scale Reduction Factor (PSRF)". If such values are very close to 1, the chain had converged.

4.2.4 The Kalman filter

The Kalman filter is an efficient estimator for state variables. It is an algorithm using time-series measures of observed data, including noise and other inaccuracies, and outputs estimates for unknown variables. It tends to be more accurate than estimates based on a single measurement alone, as it calculates joint probability distributions over each variable and for each timeframe (Kalman, 1960).

The filter is based upon a dynamical model, known control inputs for the system of interest, and sequential measurements. Its goal is to estimate its state variables, updating their values as new data come in. In addition, it uses estimates of covariance to account for noisy data and uncertainties. That is used to weight the contribution of each output to the state estimate. The reciprocal weights are also termed the Kalman "gain". With a high-gain, the filter weights the most recent measurements more, and is thus more responsive to conform to them. Low-gains, on the contrary, produce smoother but less responsive results (higher "inertia" to changes). The algorithm is repeated recursively. An illustration is provided in Fig. 4.5.

Kalman filters based on linear equations are optimal estimators for linear systems and work well as signal processing tools. Further extensions have been developed which, even though not proven optimal, deal effectively with the problem of outputting model simulations that are representative of uncertainties-prone real system observations (Humpherys et al., 2012). One such extension, later used in this thesis, is the Extended Kalman filter (EKF), which allows to use non-linear dynamical models. The EKF has been a standard tool in systems theory, with many applications ranging from automated control (Stovner et al., 2018) to finance (Davis et al., 2013), bio-mechanics (Marchesseau et al., 2013), time series analysis (Harvey, 1990), epidemic reconstruction (Proverbio et al., 2022b) *etc.* It enables estimating state variables and parameters in a continuously updating fashion, as well as to perform predictions forward in time using the estimated parameters and stopping the filtering process. When coupling the filter with ODEs, they

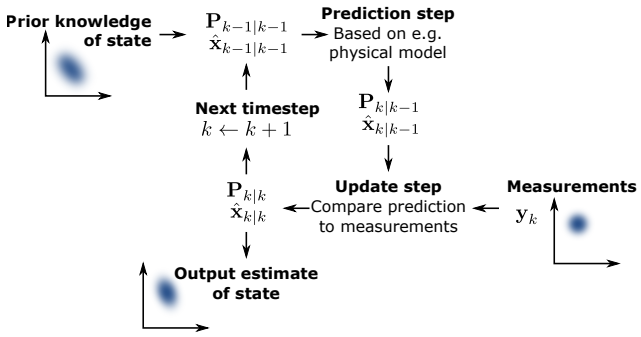


Fig. 4.5 Schematic illustration of the Kalman filter algorithm. It starts with some prior assumption for state variables, then makes a prediction; that is compared with measurements and, if necessary, an update step takes place. Predictions and comparisons are performed on the full joint probability distribution, so that the filter tracks not only the mean value of the state, but also the estimated variance. Image from P. Aimonen, https://commons.wikimedia.org/wiki/File:Basic$_c$oncept$_o$f$_K$alman$_f$iltering.svg (Creative Commons licence).

are usually discretised beforehand for ease of computation. An implementation of the Extended Kalman filter is provided in Chapter 8.

4.3 Evidence for bistability and abrupt transitions: methods

As introduced in Sec. 3.1.2, inferring whether a system is bistable or not (and might thus undergo critical transitions) from observational data might be challenging. Without repeating the limitations of this approach, already discussed in Sec. 3.1.2, this subsection reviews different methods that can be used to gain evidence for bistability and alternative states in time series data. As a toy example, some methods are illustrated with applications on simulated data from Eq. 4.1.

4.3.1 Increasing bimodality

If multiple observations of the variable of interest are available over the desired time period, observing the evolution of the probability density function over time would provide information of the system becoming increasingly bimodal (Wang et al., 2012a). This, in turn, is an insight about the system exploring different equilibria, separated by a potential barrier instead of being constrained in a single equilibrium of diffusing over the whole state space (Shi et al., 2016). An example of increased bimodality observed over time series is presented in Fig. 4.6.

Obviously, the same analysis can be carried out when repeated experiments, ideally covering one parameter value at a time, are available. In that case, it's not even necessary to assume that the windows over a time series are covering the same quasi-steady state parameter values. Common methods are Gaussian mixture models (Trefois, 2014). Articles employing the distribution method to inquire bifurcation properties of cell-differentiating systems are, for instance, Mojtahedi et al. (2016); Eugenio et al. (2014).

Bifurcation-driven transitions in bi-modal systems

Although necessary, observing bimodality with this method is not sufficient to conclude that a system might undergo a *critical* transition driven by a fold. In fact, the system behaviour may be

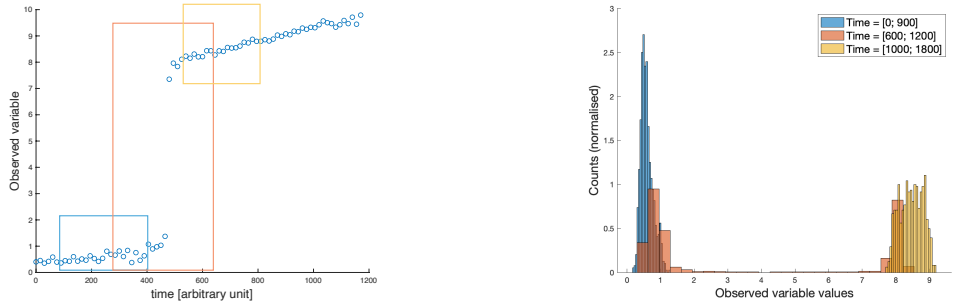


Fig. 4.6 **Left:** One example of simulated time series data, from Eq. 4.1, sampled every 30 time points. σ = 0.1. Simulations are performed from low to high parameter values, including the critical one. In different colours, the time windows used to estimate the density functions on the left. **Right:** Example of observing bistability arising at different times. The density functions (normalised as probability) are computed over windows at different times (see right figure). We can then appreciate a mono-modal distribution becoming bi-modal and then mono-modal again when the system settles into a new equilibrium.

well represented by smooth bifurcations exhibiting, at most, noise-induced transitions between their branches. Fig. 4.7b shows such a case, comparing noise-induced and bifurcation-induced transitions in a supercritical or subcritical pitchfork bifurcation (whose forms are reported in Fig. 4.7a).

Fig. 4.7c displays the theoretical evolution of the probability density function, for each quasi-steady-state value of the control parameter and for both bifurcations. If sufficient and sufficiently fine-grained data are available, fitting them with one predicted distribution or the other and performing a model selection procedure could help determining which bifurcation is more probable. One such preliminary step, distinguishing between a null model and a cusp, was performed by (Eugenio et al., 2014). Future works might test whether a subcritical pitchfork bifurcation (which is discontinuous also at the end of the first stable point) better explains certain differentiation processes.

4.3.2 Potential analysis

Related to the search for bimodality, potential analysis strives to reconstruct the stochastic potential landscape for the system of interest. It can be based on reconstructing the dynamical framework (Livina et al., 2010) or on entropy features (Wang et al., 2010), or on combinations and further methods (Zhou et al., 2012; Nolting et al., 2016).

As the present thesis is based on dynamical theory, let us sketch the basic idea of the dynamical approach – next to criticality. Assuming a Langevin dynamics like:

$$dz = -U'(z)dt + \sigma dW \tag{4.12}$$

where $U(z)$ is a potential function of the state variable z and U' its derivative, σ is the noise intensity and W a standard Wiener process. Its corresponding Fokker-Planck equation for the probability density function $P(z,t)$ is given by:

$$\partial_t P(z,t) = \partial_z[U'(z)P(z,t)] + \frac{1}{2}\sigma^2\partial_z^2 P(z,t)\ , \tag{4.13}$$

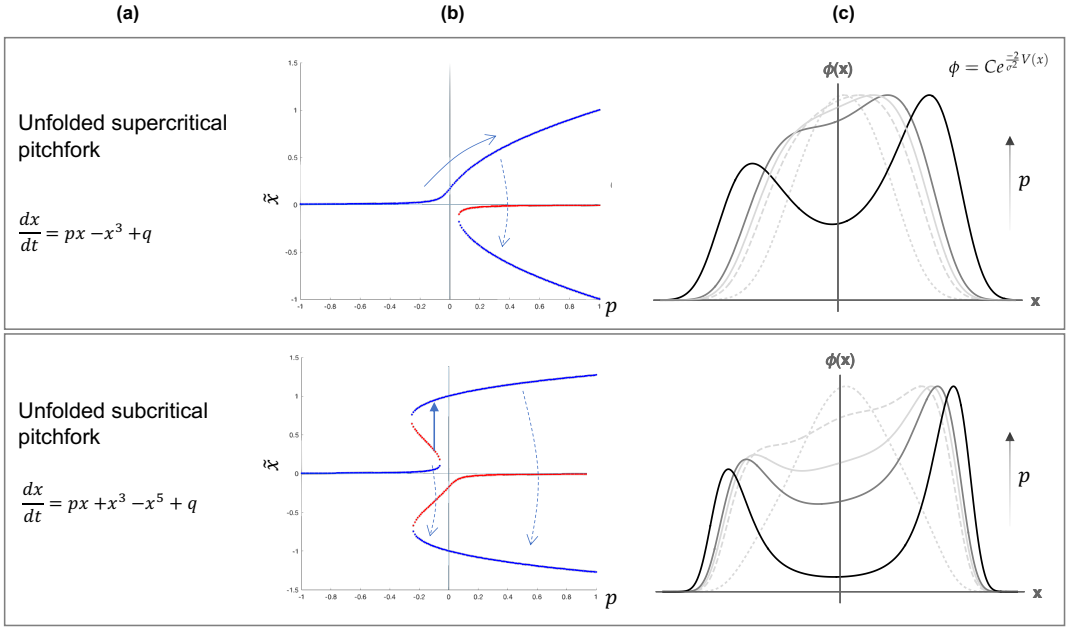


Fig. 4.7 Two contrasting transitions mechanisms in bimodal systems, following a supercritical (top) or subcritical (bottom) bifurcation. **(a)** Naming and normal form of the two considered bifurcations. Both present an unfolding term (q) to represent asymmetries and "accidents" that can be encountered in real systems. **(b)** Bifurcation diagrams, with $q = 0.1$ (blue: stable; red: unstable). The arrows indicate possible paths for a transition. Solid: smooth (on upper branch of supercritical); solid bold: critical (fold, towards upper branch of subcritical); dashed: noise-induced. **(c)** Probability density functions $\phi(x)$, for increasing quasi-steady-state parameter p values, assuming Gaussian noise added to the normal forms. $\phi(x)$ is given by the formula on top-right corner, see Sec. B and Eugenio et al. (2014), and for fixed noise intensity.

which has a stationary solution (Gardiner, 1985):

$$P(z) \sim e^{\frac{-2U(z)}{\sigma^2}} \tag{4.14}$$

If the probability density function can be empirically reconstructed using the data (call it p_d), inverting Eq. 4.14 yields:

$$U(z) = -\frac{\sigma^2}{2} \log p_d \tag{4.15}$$

Beyond the basic sketch, Livina et al. (2010) used a more refined polynomial fit to determine p_d and $U(z)$, assuming gradient functions, while Zhou et al. (2012) reviewed computational methods to reveal the landscape in non-gradient systems. They also suggest a new decomposition of vector fields to compute a quasi-potential function that is consistent with Freidlin-Wentzell's theory (Freidlin et al., 1979) but is not limited to two attractors.

Potential analysis is becoming widely applied in the biological sciences and would reveal if multiple attractors exist and, ideally, quantify their depth and stability. Under the assumption that they can be dynamic under parameter changes, reconstructed bistable potentials are basic conditions to observe critical transitions.

4.3.3 Sequential analysis

Sequential analysis (Wald, 1945) is a type of statistical analysis where the sample size is not fixed in advance but progresses as new data come in. At a certain point, the sampling might be stopped according to a certain criterion to define a significant result. Wang et al. (2012a) used a Student's t-test with $p \leq 0.01$ to identify if and when a breakpoint happens in shallow lakes time series data.

4.3.4 Changepoint detection

This method is somewhat related to sequential analysis as it involves looking for clustering conditions, specifically on time series data. Assuming that the time series can be described as a piece-wise function, several algorithms have been designed to search if and when a change-point occurs, based on similarity criteria (Truong et al., 2020).

In python, at least two packages are dedicated to changepoint analysis:

- The *rupture* library[4] was adapted from a similar package in R. It uses various search methods, including Pruned Exact Linear Time (PELT) (Wambui et al., 2015), Binary segmentation and Window-based. In a nutshell, those methods use different algorithms to find if splits exist in the time series, either consider it wholly or by computing discrepancies between adjacent bucks of points. This library is particularly suitable for off-line detection of change points.

- The *changefinder* package[5] applies a sequentially discounting autoregression time series modelling (SDAR) (Saaid et al., 2012). It extends autoregressive models, where old data weight less (are "discounted"), and searches for discrepancies between predicted and observed data, outputting anomaly scores. The authors suggest that this method is particularly suitable for on-line detection.

MATLAB also has a few libraries available:

- *findchangepts*[6] detects when a statistical property of a signal changes abruptly, by minimising the residual errors from a parametric global method. It allows choosing among various properties (mean, standard deviation, spectral characteristics etc.) and is very versatile for off-line detection.

- *ischange*[7] follows a split-search method to subdivide the time series in piece-wise functions, based on thresholding upon the variance. It is slightly simpler than *findchangepts*, but more easily implemented for on-line detection. Fig. 4.8 displays two examples of change point detection using *ischange*: one, using simulated data from Eq. 4.1, the second from real data[8] of SARS-CoV-2 abundance in wastewater samples, proportional to active positive cases, whose dynamics is associated with a passages through a transcritical bifurcation (O'Regan et al., 2013).

[4] https://github.com/deepcharles/ruptures
[5] https://github.com/shunsukeaihara/changefinder
[6] https://nl.mathworks.com/help/signal/ref/findchangepts.html
[7] https://nl.mathworks.com/help/matlab/ref/ischange.html
[8] Data are provided by the Research Luxembourg COVID-19 initiative CORONASTEP (researchluxembourg.lu/coronastep) and are aggregated at https://gitlab.lcsb.uni.lu/SCG/cowwan.

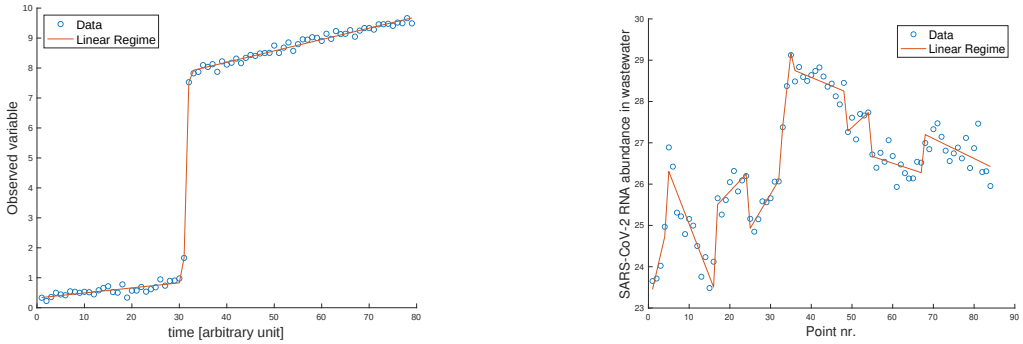


Fig. 4.8 Using MATLAB's *ischange* algorithm on different data sets. **Left:** Simulated data from Eq. 4.1; $\sigma = 0.1$. Threshold for ischange is set equal to 3σ. The metod correctly distinguishes between the two regimes and identifies the change point. **Right:** Real data of SARS-CoV-2 abundance in wastewater samples from Luxembourg (log scale). Data points are ordered sequentially and no in time, as the sampling frequency might differ over the year. Threshold for *ischange* optimised to reproduce the first two waves of infection and then kept fixed for later data. The algorithm identifies changes in dynamics (decreasing, increasing, stable) associated with passages through transcritical bifurcations of the underlying epidemic dynamics (*cf.* Sec. 4.1.2).

In a similar spirit to SDAR modelling, Wang et al. (2012a) looked for changes between the observed time series and the predicted evolution using an ARIMA model. When the discrepancy becomes significant, the authors identify the transition.

Despite the power and usefulness of changepoint algorithms, care should be maintained in pre-processing the data and in distinguishing between smooth and abrupt transitions, as exemplified in Fig. 3.1.

4.3.5 Embedding and recurrence plots

Reconstructing the state-space where the dynamics evolves can provide additional insights. Takens's theorem (Takens, 1981) provides conditions to reconstruct a delay embedding of the original attractor from observations made with a generic function. In this case, the embedding is a geometrical object that preserves important morphological properties of the original attractor, but does not preserve the geometric shape of the structures.

The time-delay embedding theorem provides a convenient (though not unique) representation of the embedded attractor in terms of delay coordinates. In this case (assuming discrete sampling of the observed system), a delay vector is:

$$\mathbf{y}(k) = [y(k), y(k-\tau), \ldots, y(k-(d_e-1)\tau)]^T \,, \tag{4.16}$$

where k is each time point considered during the reconstruction, τ is a chosen delay and d_e is the embedding dimension.

This popular method has been used to unveil and classify data patterns in various natural systems (Sugihara, 1994; Sugihara et al., 1999) and, more recently, to inform non-linear dynamics forecasting (Ma et al., 2014; Shi et al., 2021). The use of Takens's theorem to detect – and possibly anticipate – critical transitions primarily resides in reconstructing a "typical" attractor space and searching for trajectories that diverge from it. Although appealing, this method can be limited in

practice. The biggest challenge is how to derive an effective embedding dimension d_e. Indications and conjectures point to the use of diagnostic measures like entropy and nearest-neighbours of embedding, but they might not be unique or require additional meta-parameters. In addition, natural noise could make many algorithms fragile and poorly reproducible.

After the embedding is obtained, several methods and metrics can be used to identify diverging trajectories or changing patterns. Metrics include distance metrics and other dynamical indicators like cross-entropy and Lyapunov's exponents. However, preliminary results on epileptic EEG data were not able to extract reliable detection signals because of the low signal-to-noise ratio (see the thesis "Dynamical modeling techniques for biological time series data"). Future studies might better quantify the detection performance of such metrics in different systems.

For oscillatory systems like heart rhythms (Babinec et al., 2005), a common method is using recurrence plots (Eckmann et al., 1995). The plot is a graph of $\mathbf{y}(i) \sim \mathbf{y}(j)$ for every pair of time points i and j. It shows, for each moment i, the times j at which a trajectory visits a neighbourhood of the phase space as at time j. Instead of immediately defining a threshold δ for the neighbourhood, such that

$$R(i,j) = \begin{cases} 1 & ||\mathbf{x}(i) - \mathbf{x}(j)|| < \delta \\ 0 & \text{otherwise,} \end{cases} \tag{4.17}$$

for each plot pixel $R(i,j)$, modern algorithms display colourmaps for $R(i,j)$, colour-coded according to the distance $||\mathbf{x}(i) - \mathbf{x}(ij)||$. An example is shown is Fig. 3.6.
Recurrence plots are useful to visualize periodic patterns (diagonal stripes, spaced according to the harmonics), chaotic data (with point "islands" depending on the route to chaos (Ott, 2002)), multiple trends at different timescales (visual structures), etc. Regions that clearly stand out from the rest help to visually identify qualitative changes in the oscillatory behaviours and might indicate dynamical transitions (although noise might confound the plots which should therefore be carefully interpreted).

4.3.6 Methods for multi-dimensional systems

The methods described above mostly apply to one-dimensional data evolving over time. When several variables are measured at a time, and it is not clear which is of primarily interest, or suspects exist that transitions might happen to combinations of variables, other methods could be employed. As such methods have not been deeply investigated during the present project, this subsection presents a rapid list.

- **Phase space analysis:** If more than one variable is available (in multi-dimensional or networked systems, for instance), looking for clusters in their respective phase space (or PCA-ed spaces), and confirming that there is a temporal flow along those clusters, might help uncovering stable states and transitions between them. One such example is provided by Wang et al. (2012a).

- **Weighted multidimensional recurrence network:** suggested by Heino et al. (2021), this method observes whether variable states can be wired to the same states occurring at later

time, thus observing the recurrence of the same configurations over and over, providing evidence for attracting states.

- **Attractor reconstruction**: there are several methods available to reconstruct an attractor using high-dimensional observational data, e.g. Takens's theorem (see above and, e.g., Noakes (1991)) or landscape reconstruction methods.

- ***Ad hoc* indexes from dynamical theory**: a line of research is developing indexes, based on dynamical considerations and expert knowledge about specific biological systems, whose trend is expected to provide evidence for the existence of critical states in high-dimensional systems. The underlying idea is to obtain combinations of EWS (from low-dimensional system), based on considerations about the network structure of the considered systems. After that, magnify the signal associated with a critical state by using expert knowledge. For further details, we refer to the original publications (Aihara et al., 2022; Chen et al., 2012b; Mojtahedi et al., 2016; Yan et al., 2021).

- **Dimension reduction and mono-dimensional techniques:** the methods described above can be employed after mono-dimensionalising the data. This can be achieved using e.g. PCA or MAF techniques (Weinans et al., 2019), that is, looking for the phase space direction that retains most of the variance or of the autocorrelation. Alternatively, some of the methods described in Sec. 4.7 can also be employed.

- **Eigenvalues of the covariance matrix:** instead of reducing the dimension and then computing indicators associated with critical transitions, one might directly employ the eigenvalues of the covariance matrix (if it is possible to infer it from multiple time series). The eigenvector associated with the largest eigenvalue helps uncovering the direction of lowest resilience, while its evolution might provide indications about a potentially upcoming transition (Chen et al., 2019).

4.4 Methods for detrending data

As introduced in Sec 3.2.4, different methods to detrend the original time series have been proposed and employed in studies on critical transitions. In this section, some of the most commonly used methods are listed, presented and discussed. Although some analysis about their impact in the signals performance have been recently published (Lenton et al., 2012; Kuehn, 2013; Dessavre et al., 2019; Dakos et al., 2008; Proverbio et al., 2022c), there is still little consensus on which should be preferred. Moreover, several methods include meta-parameters to be adjusted: when repeated studies are possible, such parameters might be viable to optimization; otherwise, educated guesses are expected to guide their selection. For discussion about challenges in data detrending and in using moving windows over time series, refer to Sec 3.2.4.

4.4.1 Moving average

A commonly used method is to filter the data with an uniform moving average, so that a data point belonging to the smoothed time series, $z(t)$, is given by the unweighted mean of k original data

points x. k defines the width of the rolling window used to filter the original time series. In case of a non-anticipative rolling window, the smoothed point derives from past data:

$$z(t) = \frac{1}{k} \sum_{t'=t-k}^{t} x(t') \tag{4.18}$$

This ensures that the smoothing is non-anticipative, but might induce shifts in time (variations in the mean are not aligned with the variations in the data). In this case, the window size k is usually a meta-parameter to tune. Previous literature studies (Kuehn, 2013; Dessavre et al., 2019) suggest choosing window sizes that are not too big (which would require much of the available data), but not too small either. In the latter case, the ergodic assumption might not hold anymore; in addition, it might miss important trends for noisy indicators like AC(1).

4.4.2 Gaussian filtering

It uses a filter with a Gaussian impulse response, and corresponds to convoluting the original signal with a kernel given by a Gaussian function like:

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-x^2}{2\sigma^2}} \tag{4.19}$$

where σ is the standard deviation. In computer applications, the Gaussian kernel is usually truncated within rectangular moving windows. Both the size of the window and σ are meta-parameters to be adjusted. Similar considerations to the moving average case apply. Most filtering engines in popular languages allow to centre the kernel or not: the choice depends on whether the study is performed in a non-anticipative setting (mimicking an on-line mode) or not.

4.4.3 Lowess/loess smoothing

Lowess stands for "locally weighted scatterplot smoothing"; loess for "locally estimated scatterplot smoothing". These methods imply a non-parametric local non-linear regression. The difference lies in that loess relies on weighted quadratic least squares regression, while lowess uses a weighted linear least squares regression. In recent computer implementations, they might be treated as synonyms (or as current and predecessor).
The lo(w)ess method (Cleveland et al., 1990) implies evaluating the value of $f(x)$ (for each x of the time series) by using its k neighbouring samples,determined by a K-Nearest-Neighbours (KNN) algorithm. The "bandwidth" k is a meta-parameter determining how much of the data is used to fit each local polynomial; in practice it tunes statistical bias and variance. The most common weighting among neighbouring points is given by the tri-cube weight function:

$$w = \begin{cases} (1-|x|^3)^3 & |x| < 1 \\ 0 & |x| \geq 1, \end{cases} \tag{4.20}$$

aptly normalized by its highest value. Then, the algorithm performs a linear (lowess) or quadratic (loess) regression among the k-identified points, using the weighting above.
This method is mostly employed as an interpolant. Using it for on-line studies might require enforcing the use of past data only, especially when using public libraries that would otherwise

use all locally available points. An option is to superimpose an additional window, that truncates the time series at "horizon" values, and then run the lo(w)ess on this subset data.

4.4.4 ARIMA detrending

The ARIMA is an alternative class of interpolant (also used for forecasting) that uses a more global algorithm than the loess'. ARIMA stands for AutoRegressive Integrated Moving Average (Hillmer et al., 1982) and is widely used to separate short- and long-term trends in non-stationary time series data. Each element of the model is encoded by a meta-parameter (p,d,q). Hence, several ARIMA models exist depending on the parameters' combinations, and are identified as ARIMA(p,d,q). Any parameter can be set to 0 to get simpler ARMA, AR, I or MA models (Brownlee, 2017).
In a complete ARIMA(p,d,q) model, the *autoregressive* part, i.e. linear regression using past data to anticipate future ones, is modulated by p, which counts how many lagged observations are to be taken in:

$$x_t = c + \sum_{i=1}^{p} a_i x_{t-i} + \varepsilon_t \ , \tag{4.21}$$

where the subscript t denotes a discretization of data points, ε corresponds to the assumed white noise and a_i are the regression parameters.
To eliminate trends and stabilize the mean, the model performs *differencing* of raw observations (hence, the overall *integration*); the number of times that the raw observations are differenced is tuned by d. The first order considers:

$$z_i = x_i - x_{i-1} \ ; \tag{4.22}$$

the second order goes to

$$z_i = (x_i - x_{i-1}) - (x_{i-1} - x_{i-2}) \tag{4.23}$$

and so on. $d = 1$ addresses linear trends, $d = 2$ quadratic trends etc.
Finally, q is the size of the moving average window used to compute the residuals.

Optimising (p,d,q) means choosing parameter combinations such that i) the filtered data are stationary and ii) the residual mean and standard deviation are minimised. The first point fixes d, while the second point helps fixing the other two parameters and can be assisted by additional diagnostic tools such as Partial AutoCorrelation Functions (PACF). ARIMA toolboxes and libraries are present in the most common programming languages that support statistical analysis.
In general, the ARIMA is very sensitive to various degrees of inner trends contained in the data and can be a good tool to inquire time scale separations. On the other hand, they are more computationally expensive and require optimization steps necessitating abundant data availability. The remarks about using it for on-line purposes are closely related to those of loess (see above).

4.4.5 Model-based detrending

Theoretically, if a complete model is available, the use of generic early warning signals is undermined (*cf.* Ch. 1). However, scholars might still want to calculate such estimators as a validation step of the model choice, or to derive alternative signals to strengthen the observations,

or to derive other useful statistics. In this case, it is possible to fit[9] the underline dynamical model, estimate its parameters and get the residuals as:

$$r(t) = x(t) - z(t) \,, \tag{4.24}$$

where $r(t)$ identify the residuals, $x(t)$ are the original data points and $z(t)$ are the values estimate using a dynamical model. To apply the procedure routinely, as soon as new data come in, a Kalman filter (Kalman, 1960) can also be employed, see Sec. 4.2.4.

4.5 Calculate statistical indicators for early warning signals

Estimating the statistical indicators to extrapolate potential early warning signals is rather straightforward, as they correspond to statistical measures often extracted from time-series or distributional data (Dakos et al., 2012a).

For time series, it is first necessary to detrend the time series, as explained above. Subtracting the obtained moving average, which should be representative of the deterministic trend, from the original data returns the "residuals" (or detrended fluctuations), upon which one can calculate the desired indicators. To perform a non-anticipating analysis, the indicators should be computed on backward sliding windows, such that the associated time point is the rightmost one. This way, all estimates are agnostic of future values. Caveats about detrending and moving windows have been anticipated in Sec. 3.2.4. Derivation and interpretation of statistical indicators, as well as the underlying mathematical assumptions – including ergodic hypothesis – are described in Sec. B.

All EWS indicators are then estimated on the residuals. For instance, the variance is given by:

$$\mathrm{Var}^{\mathrm{time}}{}_{i,t} = \frac{1}{M-1} \sum_{s=t_0}^{t} (A_{i,s} - \hat{A}_{i,s})^2 \tag{4.25}$$

for any time point i with data A, over a sliding window with size $t - t_0$ including M time points. $\hat{A}$ is the moving average.

Similarly, one can estimate any other moment $\mathbb{E}[x^n]$, as well as autocorrelation at desired lag, power spectral density and entropy features. Alternatively to detrending and subsequent analysis, if one suspects that multiple time-scales are convoluted in the data, processing the data with Fourier or Wavelet transforms (Rosso et al., 2006; Coifman et al., 1992; Wittenberg et al., 1999; Abdel-Hamid et al., 2016; Meisel et al., 2012b; Kirby, 2005), and then estimating the indicators at each basis frequency, could be an option. Although potentially useful to look for fine-grained patterns in the data, this could come with reduced interpretability and transferability to new and/or unknown contexts.

When distributions, associated to each slowly varying value of the control parameter(s) are available, the ergodic hypothesis is easily satisfied and it is not necessary to resort to sliding windows. Instead, all statistical indicators can be immediately calculated from the distribution.

[9] Fitting methods can range from least-square to more advanced maximum likelihood methods informed by Markov Chain Monte Carlo samplers, see Sec. 4.2.3.

This can be the case for data from e.g. single-cell techniques (Kolodziejczyk et al., 2015), in the assumption that they represent distributions around mean values, for each sampling time-point (at least, until mono-modality is maintained). For instance, variance is:

$$\mathrm{Var}_j^{\mathrm{distr}} = \frac{1}{N-1}\sum_{r=1}^{N}(B_{j,r} - \hat{B}_j)^2 \tag{4.26}$$

for any point j corresponding to a single parameter value, with N data B distributed around a mean value $\hat{B}$. Other statistical moments and indicators can be computed similarly (Dessavre et al., 2019).

4.5.1 Quantification of theoretical trends and EWS

Statistical indicators *per se* are not early warning signals. In fact, it is their theoretically expected increasing trends, associated with the CSD phenomenon, that could provide EWS (see also Sec. B). Similarly, it is the reddening of the power spectrum that may provide a warning. It is therefore necessary to go beyond simple visualization of trends in indicators and to quantify them. Two methods have been commonly suggested, especially for time-series data: looking for an indicator to pass a number of σ_{Ind}s beyond its basal standard deviation (Dablander et al., 2022; Dessavre et al., 2019), or to look for the Kendal"s τ score (Boettiger et al., 2012a; Southall et al., 2020; Brett et al., 2017; Clements et al., 2016b; Wang et al., 2018). An additional method, akin to checking for deviations from basal σ_{Ind} and especially useful for distributional data, involves looking for significant p-values between "basal" and subsequent distributions. Below, a brief presentation for each method is provided.

- Looking for σ_{Ind} thresholds ("Ind" stands for "Indicator", to distinguish from noise level σ). Normal systems' activity (far from bifurcation points) is associated with variance and other indicators fluctuating to low values. Such small fluctuations are associated with an average trend and a standard deviation σ_{Ind} around it. Take that average value as reference. Then, consider when an indicator increases past $a \cdot \sigma_{Ind}$, where a can either be defined *a priori* using common statistical knowledge (e.g. $a = 2$ or 3), or it can be determined *a posteriori* to optimise some detection performance like AUROC (see below). An example is provided in Fig. 4.9, left. This method is quite intuitive and helps associating warnings wit absolute values of indicators; on the other hand, it can be sensitive to fluctuations. In addition, it could be complicated to transfer the optimal threshold to different cases.

- The Kendall's τ score is a non-parametric coefficient of monotonicity. It is defined as (Kendall, 1938):

$$\tau = \frac{\#\text{concordant pairs} - \#\text{discordant pairs}}{M(M-1)/2}. \tag{4.27}$$

  M is the number of considered time points. Two generic points (t_1, x_1) and (t_2, x_2) are said to be a concordant pair if, for $t_1 < t_2$, $x_1 < x_2$, and a discordant pair otherwise. A constant trend is expected to have $\tau = 0$. After calculating each indicator on a moving window, the Kendall's τ score is normally estimated, for each timepoint, on additional sliding windows (example provided in Fig. 4.9, centre). Such additional windows to compute thee Kandall's score should also be employed for distributional data. As a reference, values far from the

bifurcation are normally employed. This method is more robust than the σ thresholds against fluctuating scores, but requires the tuning of an extra window and is less intuitive.

- In certain case, like from biological experiments performed with single-cell RNA sequencing, distributions are available for sparse values of the control parameter. Hence, we can compare the full distributions (a "basal" one and others suspectedly closer to the transition) and check whether they are significantly separated towards increasing values. The significance can be assessed e.g. using p-value scores. See Fig. 4.9 (right) for an example. This method can still be sensitive to fluctuating scores, but it has the advantage of relying on a-prioristic values. On the other hand, it requires abundant distributional data and is less suitable for time-series analysis with single repetitions.

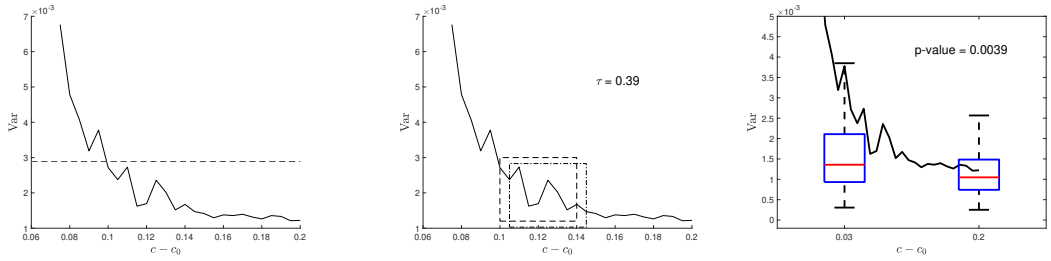


Fig. 4.9 **Left:** Example of looking for trends past σ_{Ind} thresholds – in this case, $2\sigma_{Ind}$ (dashed line). **Centre:** Example of Kendall's τ estimation. Compare the trends within two sliding windows. If the new one is monotonously increasing with respect to the old one, $\tau > 0$ (the steeper the trend, the higher τ). **Right:** Example of p-value between two distributions corresponding to different parameter values. Each distribution corresponds to an average value of the statistical indicator (superimposed and shifted for visualization purposes). p-values's significance can be checked with standard statistical methods.
All figures use variance computed from simulations of Eq. 4.1, with $n = 2$, $K = 0.1$ and $\sigma = 0.02$. c_0 is the critical value for bifurcation point.

4.5.2 Publicly available packages

Despite the ease of self-implementing the statistics needed to extrapolate early warning signals, a few packages are also publicly available and ready to be used:

- For the R language, the Early Warning Signal toolbox is available on cran at: https://cran.r-project.org/web/packages/earlywarnings/index.html. It has been developed and curated by Vasiliis Dakos and co-workers of the Sparcs program (https://sparcs-center.org/).
- For python, the *ewstools* package has been recently published: https://ewstools.readthedocs.io/en/latest/. Credits go to Thomas Bury.

4.6 Assessing the performance of early warning signals

Quantifying the performance of EWS means quantifying their adequacy to detect regimes with losing resilience. Casting the problem as a detection challenge (Zhang et al., 2015; Drake, 2013; Zhang et al., 2016) enables the usage of tools to quantify sensitivity and specificity of various

indicators. This way, one could aim at detecting the transition itself, or the detect resilience levels associated to increasing probabilities of tipping.

A common tool to assess detection performance is the Receiver Operator Characteristics (ROC) analysis. It returns the ROC curve, a parametric plot of sensitivity and specificity of a classification method, plotted as a function of the detection threshold (Brett et al., 2020a; Fawcett, 2006). In practice, the ROC analysis sets a classification criterion ($Ind > a \cdot \sigma_{Ind}$, for instance, see above) and checks how many true positives (sensitivity) and false positives (1-specificity) are observed. Then it changes a and repeats. The final curve is composed by points identified by the coordinates (1-specificity(a); sensitivity(a)). A generic example is shown in Fig. 4.10. The overall detection performance of each EWS is quantified by the area under the ROC curve (AUC, which is sometimes called AUROC – area under the ROC curve – to tell it from other "areas under"). A value AUC$= 0.5$ means that the statistics detection performance is as good in classifying as randomly guessing. A good indicator should have AUC close to 1, see also Fig. 4.10.

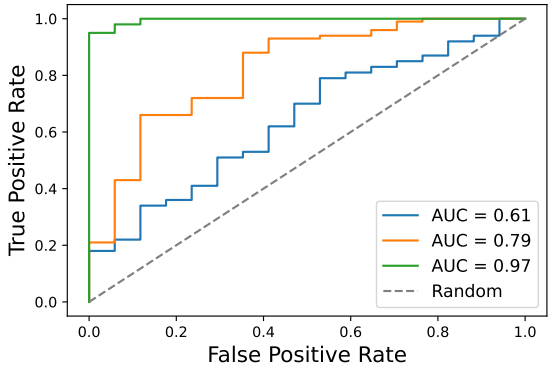


Fig. 4.10 Example of ROC curve (made up). Each point returns the proportion of true detections vs false ones, for each classification threshold. Curves moving closer to the upper-right corner correspond to best performance (highest AUC): any threshold correctly identifies the true events, so setting the minimal one suffices to minimise the false positives. Curves closer to the diagonal line display slightly better performance than a random classifier. Curves in the middle call for decisions about the optimal thresholds, to set tradeoffs between true and false positives.

4.7 Dimension reduction for critical states

What is discussed above refer to mono-dimensional systems and data collections, which are the focus of the current project. Nonetheless, multi-dimensional systems are receiving increasing interest in very recent years. One way of approaching multi-dimensional systems with scarce information and incomplete models is to perform dimension reduction, and the to apply critical transitions theory for low dimensional systems. Depending on the research question, these methods can be used to collect evidence about the existence of alternative states and abrupt shifts, or to process the data before looking for early warning signals (see workflow sketch in Fig. 4.11). Some of the recently developed techniques are discussed below.

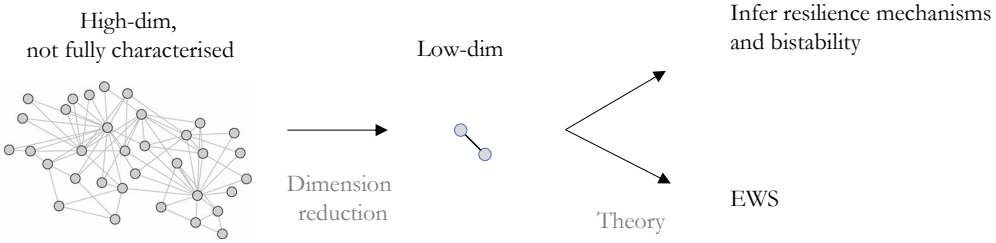


Fig. 4.11 Workflow sketch for the use of dimension reduction techniques in combination of critical transitions theory, either to look for bistability and mechanisms for criticality, or to pave the way for the application of EWS.

4.7.1 Changing basis

Principal Component Analysis (PCA) is the most employed technique to change the basis to describe a system and to identify the leading eigen-directions in the phase space. It has been discussed by Weinans et al. (2019) as a way of identifying the direction of lowest resilience in multivariate systems, thereby informing about the variables combinations more at risk of undergoing a critical transition. The same paper also introduces the Mix/Max autocorrelation factors (MAF) as an alternative method, to reduce the system according to its autocorrelation.

PCA computes the principal components of a collection of points (a sequence of unit vectors that form an orthonormal basis that best fits the data) and projects the data points onto only the first few principal components, while preserving as much of the data's variation as possible. Any data analysis software contains functions to compute PCA. Care should be taken when cropping the dimensions after running diagnostic tools: if the first of firsts dimensions retain low percentages of variability, using a single leading direction is hard to justify (data might be quite evenly scattered across eigen-directions).

MAF is another transformational algorithm, detecting the direction of maximum autocorrelation (Desbarats et al., 2000). The output factors are not orthogonal but provide an ordering from highest to lowest autocorrelation, which is in turn related to lowest to highest resilience. It can be easily implemented by following the process described by Weinans et al. (2019).

The main advantage of such methods is that they are model-free and can be readily implemented with very little knowledge about the system. Nonetheless, both methods are observed to be highly sensitive to noise type, intensity and distribution over variables. MAF is slightly more robust, whereas PCA is a better alternative in case data are sparsely collected. Additional caveats are required when interpreting the results: whether the system's recovery is slow or fast, whether local information is sufficiently representative of global behaviours (e.g. spiralling returns to stable states), and whether the directions of highest variance or autocorrelation match 1:1 the directions of lowest resilience. Other limitations involve the detection power of EWS calculated after applying such techniques (Weinans et al., 2021). It has been observed that signals might be too low, shaded by combinations of fluctuations (also observed in Lever et al. (2020)), or could provide false positives. A systematic exploration of high-dimensional parameters, dynamical regimes and techniques is further required to assess the benefits of these methods.

4.7.2 Weighted dimension reduction

Under the assumption of homogeneous networks, Gao et al. (2016) developed a dimension reduction technique weighting the contribution of each node by its degree and thus obtaining a "effective" dynamics over one dimension. A sketch of the idea is the following.

Consider a system consisting of N coupled components (nodes), whose activities $\mathbf{x} = (x_1, \ldots, x_N)^T$ follow the dynamics:

$$\dot{x_i} = F(x_i) + \sum_{j=1}^{N} A_{ij} G(x_i, x_j) \,, \tag{4.28}$$

where $F(x_i)$ describes the self-dynamics of each component, $G(x_i, x_j)$ describes the interaction laws between component i and its interacting neighbours and A_{ij} is a positive connectivity matrix. The effective state x_{eff} of the system is obtained by the average nearest neighbours activity, assuming that A_{ij} has little degree correlation:

$$x_{\text{eff}} = \frac{\mathbf{I}^T A \mathbf{x}}{\mathbf{I}^T A \mathbf{I}} = \frac{\langle s^{\text{out}} x \rangle}{\langle s \rangle} \,, \tag{4.29}$$

where $\langle . \rangle$ is the averaging operation, s^{out} is the out-degree for every node and $\langle s \rangle = \langle s^{\text{out}} \rangle = \langle s^{\text{is}} \rangle$ is the average weighted degree. $\mathbf{I}$ is the unit vector. The same procedure is also performed to obtain an effective coupling parameter, i.e. a macroscopic parameter that the matrix A_{ij} collapses to:

$$\beta_{\text{eff}} = \frac{\mathbf{I}^T A \mathbf{s^{in}}}{\mathbf{I}^T A \mathbf{I}} = \frac{\langle s^{\text{out}} s^{\text{in}} \rangle}{\langle s \rangle} \,, \tag{4.30}$$

This way, the changes in β_{eff} lump al microscopic details of the perturbation on the microscopic system (node/link removal, weight reduction or any combination). One can thus use x_{eff} and β_{eff} to reduce the multi-dimensional equation 4.28 into a 1D equation:

$$\dot{x_{\text{eff}}} = F(x_{\text{eff}}) + \beta_{\text{eff}} G(x_{\text{eff}}, x_{\text{eff}}) \tag{4.31}$$

The authors show that this method is able to reconstruct a "universal resilience function", that is, a bifurcation diagram x_{eff} vs β_{eff} including the transition point from one stable state to another. Interestingly, the transition point is fully determined by the dynamics, independent of the network topology A_{ij}.

The strength of this method relies on its simple assumptions and its ability to reconstruct and "effective" dynamics that can immediately inform about bistability properties. However, as observed by other authors, the simple weighting over node degree might not be accurate in case off heterogeneous networks. In addition, the uncertainties associated with this method might hinder subsequent EWS signals and limit the precision for estimating the critical value.

This paper spurred the interest for other weighting techniques, that are currently being developed. They are getting increasingly more accurate in determining the resilience function and the critical effective-parameter value, at the cost of requiring more information about the system network, which is an important tradeoff to consider in real-world applications.

Laurence et al. (2019) developed a spectral dimension reduction technique, based on weighting the nodes activity onto eigenvalues of the adjacent matrix, to target homogeneous but non-random networks. Tu et al. (2021) extended the idea to node-specific dynamics, obtaining a polynomial reduction. Pan et al. (2020) tested phase diagrams and extended them to simplicial complexes. Other methods, that nonetheless require advanced knowledge of the network topology and dynamics, are reviewed in Cheng et al. (2021).

Similar in spirit, Jiang et al. (2018) applied a degree- and eigenvalue-weighted averaging with the goal of obtaining a two variables system (not 1-D). They claimed it was better representative of the specific class of mutualistic ecological network they considered. As a result, they obtained resilience functions exhibiting tipping points and analysed the role of several network properties on the reduced model. Notably, what mostly played a role in this case was not the method itself (both degree and eigenvalue weighting performed well on synthetic data) but the modelling choice of conserving two dimensions instead of one, thereby discussing the role of "objective" techniques versus expert-informed ones.

Chapter 5

EWS sensitivity and performance

5.1 Motivation

To address systems lacking bottom-up mechanistic models and abundant data, dynamical systems theory has been employed to derive generic early warning signals, that can be extracted from data. However, they are measures that rest upon assumptions. Whether EWS may perform as expected, thus providing reliable signals of resilience loss, or fail to do so, largely depends on their sensitivity to the underlying assumptions and whether they are valid in empirical systems. Failure to match modelling assumptions and system properties may result in confounding effects (McNamee, 2003) and in incorrect interpretations of estimated signals.

This chapter aims at systematically investigating the effects that relaxing some key assumptions have on the behaviour of EWS. In particular, the chapter complements existing literature on the role of noise properties and expands the investigation to non-slow-fast systems. The results provide insights as to why applying standard EWS on different complex systems may provide outputs that contrast with data-driven studies. The theoretical results are then applied to models of biological regulation, where they are employed to detect and anticipate critical switches. Optimal EWS are also estimated. Then, they are contrasted with data-driven results from experiments about mitochondrial activity. Consequently, this chapter provides a set of guidelines to choose which indicator(s) of resilience is more suitable and best performing for particular systems.

5.2 Introduction

Regime shifts after tipping points are abundant in many real-world complex systems, which often lack bottom-up mechanistic models and abundant data pertaining different conditions (Chapter 2). Particularly for such cases, the critical transitions framework aims at developing generic indicators to detect and alert for impending regime shifts. The most employed method to extract EWS is to exploit proxies of critical slowing down near noisy bifurcations (Sec. 2.4). This mechanism is sufficiently generic to capture key properties of certain classes of CTs (see Sec. 2.3.2), but relies upon several assumptions. Hence, it is not trivial to assess *a priori* whether EWS can be observed in systems of interest, and what is their expected performance.

To begin with, one assumes that system reduction to closures around bifurcating manifolds is possible (Sec. A). This presumes a timescale separation (Sec. 2.3) between the main observ-

ables, usually embedded in low dimensional systems, the slow (quasi-steady-state) degrees of freedom, modelled as control parameters, and the remaining fast degrees of freedom, addressed by considering stochastic terms. The second assumption pertains the stochastic terms themselves, which are assumed to be white in time. Whether EWS can be extracted from data regardless of the underlying assumptions, and how their detection and alerting performance is impacted, is still matter of debates, see chapter 3. In fact, both assumptions may fall short in certain systems of great relevance.

The influence of noise on statistical patterns prior to bifurcations was observed in ecological, epidemiological and climate models. For instance, Dakos et al. (2012b) showed that variance may increase or decrease prior to population collapse, depending on how environmental noise is associated to model parameters. O'Regan et al. (2018) showed that intrinsic (internal) and extrinsic (environmental) noise may drastically alter EWS trends. Analysing several normal forms[1], a Susceptible-Infectious-Susceptible (SIS) epidemiological model[2] and an ecology model, the authors warned against process-specific effects and potential misinterpretation of EWS due to noise properties. Kuehn et al. (2022) considered coloured noise in non-Markovian (*i.e.,* non-memory-less) systems, which model long-term historic fluctuations producing relevant autocorrelation effects over long time scales. As an example, the authors discuss box models for the Atlantic Meridional Overturning Circulation (AMOC). Forests and coral reefs may also be impacted by long-term climate memory (Bolt et al., 2018). The authors concluded that observed discrepancies between data analysis and EWS predictions may result from colour-blindness: coloured noise can make classical warning indicators to fail and be misinterpreted. In these cases, if noise type in data is not known, predictability becomes very difficult.

In addition to modelling results, estimation of EWS from empirical data showed that, if one or both assumptions are incorrect, classical EWS perform poorly to detect and anticipate transitions in epidemic dynamics, see Proverbio et al. (2022c) and chapter 7. Providing possible explanations for discrepancies in predictability is thus necessary to apply EWS correctly.

Noise and commensurable time scales in threshold processes occur both in systems biology and epidemiology. Understanding how such dynamical features impact EWS trends is crucial to interpret those estimated from data and to guide the selection of the best indicators to tackle each system.

In systems biology, the CT framework has been suggested to provide useful insights and methods (Korolev et al., 2014; Trefois et al., 2015), and statistical indicators have been applied directly (Dai et al., 2012) or within composite indicators (Li et al., 2014; Mojtahedi et al., 2016; Aihara et al., 2022; Richard et al., 2016; Yang et al., 2018; Li et al., 2021b) to extract early warnings. However, biological systems are known to exhibit various dynamical regimes depending on specific processes. Microscopic processes, internal regulation and environmental disturbances can result in white as well as multiplicative (state-dependent) noise even in homeostatic processes (Kepler et al., 2001; Lestas et al., 2008; Lestas et al., 2010; Bruggeman et al., 2018). Fluctuations are also known to yield stochastic switching (Norman et al., 2015), a regulation mechanism better described by the n-tipping class (Sec. 2.3.2) and complementary to deterministic switching

[1] See Sec. A for details.
[2] See Sec. 4.1.2 for an introduction to compartment models in epidemiology.

through bifurcations. In this case, depending on noise properties, statistical indicators may only detect just-on-time occurring shifts, rather than anticipating them, or fail completely. For what concerns the slow-fast system assumption, many driving mechanisms are assumed at quasi-steady state, like mRNA transcription versus protein degradation (Segel, 1988; Yasemi et al., 2021). However, comparable mean life-times of chemical compounds may be observed in other control mechanisms (Del Vecchio et al., 2016; Gunawardena, 2014) or in hormone action (Bradford et al., 1994). Hence, understanding how relaxing such assumption influences temporal patterns of EWS is necessary to explain empirical observations.

In epidemiology, the impact of non-white noise has been studied by Brett et al. (2017); O'Regan et al. (2018) and Southall et al. (2020) among the others. The authors observed a diminished capability of detecting epidemic outbreaks or elimination in several example models of multiplicative noise, linked *e.g.* to small infectious numbers, sampling protocols or population behaviours. In addition, Brett et al. (2017) computationally inquired the influence of rapid parameter ramping in certain models, observed a delayed or absent response of several indicators, compared to stationary theoretical predictions.

In the last years, these observations started a call for caution when using EWS in routine monitoring practices. Hence, this chapter first theoretically investigates the reason behind such concerns, assessing the sensitivity of early warning signals to different dynamical contexts and assessing which ones are expected to best perform in commonly observed settings.

Then, for a specific case of gene regulatory mechanisms, the present analysis derives sets of indicators that, taken together, optimise the anticipating power of EWS applied to data. The optimisation problem is solved for different combinations of noise properties, providing guidelines to tackle realistic measurements.

Overall, the chapter expands the theoretical foundations for the application of EWS in unknown systems, and sets the stage for the analysis carried out in subsequent chapters, which focus on testing the theoretical predictions onto empirical data.

5.3 Theory and Methods

This study focuses on early warning signals associated to bifurcation-driven loss of stability (b-tipping, Sec. 2.3.2). Hence, we hereby consider minimal forms and biological models whose transitions are characterised by noisy bifurcations. The background theory, as well as notable results for stationary processes with white noise, is detailed in Sec. 2.3 and Appendix A and B. In what follows, the most salient points are briefly recalled.

5.3.1 Generic EWS from normal forms

Among the critical bifurcations listed in Table 2.1, the 1-dimensional, 1-codimensional ones – the most common in CT studies – are studied using their corresponding normal forms (*cf.* Sec. A). Focusing on normal forms means considering local behaviours of all dynamical systems that display such modulating mechanism. Hence, they allow to extract analytical and generic results for wide classes of systems (Kuehn et al., 2021), at the price of neglecting homeostatic dynamics far from tipping points. To characterise the behaviour of statistical indicators and their scaling

next to criticality – and thus derive solutions for the corresponding EWS – normal forms are the most appropriate modelling constructs.

Normal forms are characterised by dynamical equations in the form

$$\begin{aligned} \frac{dx}{dt} &= f(x,p) \\ \frac{dp}{dt} &= \varepsilon g(x,p) \end{aligned} \tag{5.1}$$

where x is the variable of interest and p is the control parameter, with a critical point $p = 0$. If $\varepsilon \to 0$, the dynamics of p is considered at quasi-steady state with respect to x (see an introduction to slow-fast systems in Sec. 2.2.4).

As a reminder, the normal forms of interest are the following. Saddle-node bifurcation, often associated with ecological collapses (Scheffer, 2009) or biological state transitions (Alon, 2019), are defined by $f(x,p) = p - x^2$. At $p = 0$, a stable ($\tilde{x}_s = \sqrt{p}$) and unstable ($\tilde{x}_u = -\sqrt{p}$) branch collide and vanish, resulting in a critical transition to an alternative branch (if it exists). Transcritical bifurcations $f(x,p) = px - x^2$ are characteristic, for instance, of epidemic outbreaks, see Sec. 4.1.2. Here, $x_1 = 0$ and $x_2 = p > 0$ meet at $p = 0$ and exchange stability. Finally, the family of pitchfork bifurcations $f(x,p) = px + l\,x^3$ describe branching processes from one to two states (or viceversa); $l > 0$ identifies subcritical bifurcations, associated to critical transitions, while $l < 0$ defines the supercritical case, with a continuous transition over mean values. This mechanism is identified in cell regulation processes (Moris et al., 2016).

Noisy bifurcations, from which EWS can be extracted (Berglund et al., 2006), are characterised by the presence of additional stochastic terms, see Eq. 2.8. Recall that, for the fast variable, stochastic differential equations for 1D systems can be written, in Itô form, as

$$dx = f(x,p)dt + h(x,p)dW\ , \tag{5.2}$$

where dW is a Wiener process and $f : \mathbb{R} \times \mathbb{R} \to \mathscr{R}$ is a suitable normal form from those described above. Systems and processes may be characterised by different stochasticity, which enters the SDE models as additive ($h(x,p) = 1$) or multiplicative terms to the Wiener processes, with variance σ (Allen, 2010). In the second case, if the microscopic kinetics is known, one can derive the exact form of the noise term from the Master equation using Gillespie formalism (Gillespie, 2000a). Alternatively, one can use a diffusion approximation (Allen, 2010; Van Kampen, 1992) to derive a noise term proportional to system state ($h(x,p) = x$) or to the deterministic part of Eq. 5.2 ($h(x,p) \propto f(x,p)$). In the latter case, O'Regan et al. (2018) considered intrinsic noise with $h(x,p) = \sqrt{f(x,p)}$; here, $h(x,p) = f(x,p)$ is also studied, to reflect modelling of biological regulatory circuits as described by Hasty et al. (2000).

Following the procedure detailed in Sec. B, Eq. 5.2 is analysed by solving the slow dynamics, linearising around a trajectory inside the stable (attracting) manifold and changing the coordinates to highlight the residuals around the linearization. This procedure gives

$$dy = \partial_x f(\tilde{x}_s(t),t)y\,dt + \sqrt{h^2(x)}dW \tag{5.3}$$

where $\tilde{x}_s$ corresponds to the attracting part of the critical manifold. The linearised drift term corresponds to the leading eigenvalue of the deterministic normal form. Its magnitude $|\partial_x f(\tilde{x}_s(t),t)|$ is the asymptotic decay rate of a perturbation. It corresponds to the concept of engineering resilience (Holling, 1996). A change of notation $|\partial_x f(\tilde{x}_s(t),t)| = k$ makes explicit that Eq. 5.3 corresponds to a (possibly non-autonomous) Ornstein-Uhlenbeck process (O-U). It is a well-studied problem in stochastic processes theory, with analytical solutions for its statistics in different regimes (Allen, 2010; Gardiner, 1985). Such solutions constitute the basis to derive EWS in different dynamical contexts, including the possibility of $p(t) \neq$ const (small ramping in the control parameter, not at steady-state).

Note that Eq. 5.3 can be regarded as a first order autoregressive model. However, its derivation from normal forms allows more nuanced interpretation: rather than being hypothesised as a statistical model to capture simple relationships, it is general for all models that can be reduced to normal forms. It is therefore linked to phenomenological models rather than purely statistical ones (see Fig. 1.2 and Kuehn et al. (2021)). In this respect, $|\partial_x f(\tilde{x}_s(t),t)| = k$ hides the dependencies of the drift term to the control parameter p, depending on different normal forms. This may yield diverse scaling for the summary statistics, which will be explicitly considered in the results below. A linear relationship $k = p$ is characteristic of (locally) linear systems used, *e.g.*, in engineering, and corresponds to the approach to a transcritical bifurcation. If k is a nonlinear function of p, it accounts for additional non-linearities in the fast-slow system relationship.

5.3.2 Simulations: optimal performance and sensitivity to noise

To investigate the interplay of bifurcation-driven loss of resilience and noise levels, analysing minimal analytical models may not be sufficient, due to their underlying approximations. Computer simulations are thus used to address a representative system, the autoactivating feedback loop motif described in Sec. 4.1.1.

The simulated model

The model describes protein concentration dynamics embedded in a feedback loop. Its adimentionalised Langevin equation is given by Eq. 4.1. The system is bistable, with sudden shifts between equilibria due to fold bifurcations driven by the parameter c (also verified using XPP-Aut, see Sec. 4.2.1). To consider uncertainties associated to distributional data, on top of mean trends, and to reproduce data analysis protocols, early warning signals are estimated from system simulations. Since the quasi-steady state assumption is generally accepted for such systems (Del Vecchio et al., 2016), the current analysis focuses on noise effects, whether yielding n-tipping or possibly skewing statistical indicators due to multiplicative nature. This is done by setting $g(x)$ (from Eq. 4.1) appropriately, as discussed above. Note that $g(x)$ from Eq. 4.1 corresponds to $h(x)$ in Eq. 5.3. The nomenclature was chosen to highlight generic or specific systems, but the same results apply.

In all simulations, $K = 0.1$ to set bistability. The analysis concentrates on the upper stable branch of the bifurcation diagram (Fig. 4.2) to compare with white noise results. In this case, multiplicative noise corresponds to intrinsic regulatory mechanisms (Hasty et al., 2000; Norman et al., 2015) rather than stochasticity due to small numbers (Gillespie, 2000b). Simulations are performed in MATLAB (R2021b) using the Milstein method (Sec. 4.2.2) with a time step of 0.01

(arbitrary units). For quasi-steady state simulations, distributional data for each c from far to close the bifurcation point are computed upon stable values of system's state, after a transient.

Computed indicators

Statistical indicators, whose analytical counterpart is derived from normal forms, are first computed as described in Sec. 4.5. Following theoretical considerations and practice, those selected are: variance and higher-order moments, autocorrelation at lag 1, coefficient of variation (std/mean), index of dispersion (var/mean). Shannon entropy H_s, for which the analytical derivation is particularly challenging in some settings (see Sec. B), is also estimated. All indicators are computed using their corresponding MATLAB functions.

Test EWS sensitivity to noise levels and n-tipping

Higher noise levels may alter the significance of increasing trends, as uncertainties associated with statistical indicators may increase accordingly. In addition, higher noise levels may induce n-tipping events as system trajectories may accidentally collide with unstable branches (see an illustration on Fig. 2.13). Hence, understanding the robustness and performance of EWS for different noise levels bridges studies about b-tipping and n-tipping, and the associated robustness and performance of EWS.

Various noise levels are considered, corresponding to low to medium, $\sigma \in [0.01; 0.05]$. Higher values are not considered as they would yield "pure" n-tipping. For simulations using additive white noise, $g(x) = 1$; a relevant case for multiplicative noise is, instead, $g(x) = x$.

To quantify the increasing trends of the indicators listed above, and whether they can provide an early warning, the test for significant Welch's p-value is used (*cf.* Sec. 4.5.1). In this way, the study can inform biological experiments, which are often performed over distributional data from few time points. It also allows to estimate the sensitivity to noise intensities and the expected lead time for detection or anticipation of critical transitions.

For the latter purpose, the analysis extracts at which value of the control parameter c the p-value crosses a significance threshold. Using standard significance values from biological literature, $p_{sig} = 0.05$ is chosen. When p-value $< p_{sig}$, it means that an indicator has significantly increased more than the baseline, triggering a warning signal. Consider all c_i tested during the simulations, $i = 1..N$ with $N = (c_{max} - c_{min})/0.002$; c_{max} and c_{min} are two arbitrary values greater and lower than the bifurcation value c_0, within the bistable region. Out of all c_i, estimate $c_{sig} = c_j$, where j is the first index at which p-value$_j < p_{sig}$ stably, *i.e.*, without considering small fluctuating values (for that, a smoothing is employed). This is performed for each indicator $\mathscr{I}$ and each noise level σ. Hence, the analysis estimates

$$c_{sig}^{\mathscr{I}}(\sigma) = c_j \text{ s.t. p-value}_j(\mathscr{I}) < p_{sig} \wedge \min(j) \,. \tag{5.4}$$

Complementing the analysis of the lead parameter requires understanding how many noise-induced tipping events occurred before it. This quantification assesses whether the increasing indicators are primarily picking losses of resilience or transitions that already happened. This

analysis is crucial to interpret warning indicators as "anticipating" or "just-on-time detecting" the tipping events.

To do so, for each parameter value c and for each noise level σ, count how many trajectories, out of all repeated simulations, tipped onto the alternative stable manifold. Call the counter $\mathscr{C}$. A plot $\mathscr{C}(c,\sigma)$ (normalised to be interpreted as probability) gives an indication of how much n-tipping becomes relevant as noise levels increase, and allows estimating what happens at each $c_{sig}^{\mathscr{I}}(\sigma)$.

Optimise EWS according to lead parameter

Which is the best combination of statistical indicators to maximise the lead parameter, *i.e.* how much before the bifurcation point a significant signal is triggered?

To answer this question, consider a linear combination of indicators, which were previously tested singularly:

$$\mathscr{S} = \sum_k w_k \mathscr{I}_k \,. \tag{5.5}$$

Call a set of weights $\mathbf{w} = \{w_k\}$, with the constraint $\sum_k w_k = 1$. Extract its corresponding $c_{sig}^{\mathscr{S}}(\sigma)$ by measuring when p-values of the newly constructed distribution mark a significant increase, adapting Eq. 5.4 accordingly. Finally, solve an optimization problem by finding the optimal $\hat{\mathbf{w}}$ as

$$\hat{\mathbf{w}} \text{ s.t. } \max_{\mathbf{w}} \left[\sum_l c_{sig}^{\mathscr{S}}(\mathbf{w}, \sigma_l) \right] \,, \tag{5.6}$$

that is, maximising the scores S derived from all c_{sig} as σ is increased. This way, the optimal parameter is also checked for its robustness against noise. The optimization is performed using a grid search method.

5.3.3 Experimental data

Analytical and simulation results are used to interpret experimental data about mitochondrial activity in HeLa cells, described, reported and analysed in the PhD thesis of Trefois (2014). Refer to it for experimental protocols and analysis methods. Data correspond to critical transition experiments of mitochondrial membrane potential $\Delta\Psi_m$ (state variable) driven by combination of chemical toxins, namely Oligomycin A ($C_{45}H_{74}O_{11}$, an antibiotic inhibiting mitochondrial respiration) and Antimycin A ($C_{28}H_{40}N_2O_9$, a compound depleting mitochondrial ATP).

5.4 Results

5.4.1 Analytical results on normal forms

From Eq. 5.3, standard approaches (Gardiner, 1985; Allen, 2010) derive the behaviour of statistical indicators in different dynamical contexts (slowly ramping parameter and various forms of multiplicative noise). Sec. B details the derivation of statistical moments, autocorrelation, spectral density and Shannon entropy in the stationary, white noise case (corresponding to the common setting). In the case of multiplicative noise for stationary O-U processes, analytical

results for variance, autocorrelation and spectral density, the three main used indicators, are easily obtained from literature (Allen, 2010; Sharma et al., 2016a). The lag-τ autocorrelation function does not depend on the O-U variance $h^2(x,p)$ but only on $|\partial_x f(\tilde{x}_s,p)|=k$:

$$AC = e^{-k\tau}. \tag{5.7}$$

Hence, the common indicator lag-1 autocorrelation only depends on the dampening rate:

$$AC(1) = e^{-k}. \tag{5.8}$$

Instead, the power spectrum of the Fourier transforms and the variance, two common indicators, explicitly depend on the O-U variance:

$$S(\omega) = \frac{h^2(\tilde{x},p)}{k^2+\omega^2} \tag{5.9}$$

$$Var = \frac{h^2(\tilde{x},p)}{2k}. \tag{5.10}$$

Consequently, the coefficient of variation also depend on $h^2(\tilde{x},p)$ as it is defined by

$$CV = \frac{\sqrt{Var}}{\tilde{x}_s}. \tag{5.11}$$

Entropy-based indicators, on the other hand, are more challenging to derive in case of multiplicative noise and ramping parameters, as their defining integrals may not be solvable. The derivation in case of white noise is described in Sec. B; for the other cases, their behaviour is estimated using computer simulations (see sections below).

In all cases, the analytical results for each normal form can be obtained by substituting the corresponding dependency of the drift term to the control parameter: for the fold $k=2\sqrt{p}$, for the transcritical $k=p$ and for the pitchforks $k=2p$. In Fig. 5.1, the effect of multiplicative noise on the trends of common indicators is shown using $h(\tilde{x},p)=x$ and $h(\tilde{x},p)=x^2$.

Analogous results can be obtained for slowly changing resilience function, linearly ramped as $k=k_0-at$, where k_0 is any initial condition, a is a small rate parameter (not enough for r-tipping) and the ramping stops at the critical value $k=0$. Both parameter values are set to 1 to represent commensurable time scales. Here, only white noise is considered. This is a particular case of an inhomogeneous process[3], for which statistical moment solutions exist (Gardiner, 1985) in the form

$$\langle y(t)\rangle = e^{-\int_0^t k(t')dt'} \tag{5.12}$$

$$\langle y(t)y(t')\rangle = \frac{\sigma^2}{2k}e^{-2\int_0^t k(t")dt"} + \sigma^2\int_0^t e^{-2\int_{t'}^t k(s)ds}dt' \tag{5.13}$$

[3]This formalism can also serve as a basis to calculate statistics for generic multi-dimensional inhomogeneous processes (Gardiner, 1985); however, they are system-specific and, for the reasons discussed previously, outside the scopes of the present thesis.

Derived statistics are calculated analogously. Fig. 5.1 shows how trends of common indicators are changed by the time dependency of the resilience function, after solving the above equations with the Mathematica software, which particularly helps to calculate the rightmost integral yielding the non-elementary Error function Erf(t). Note that, since such functions are involved, a linear scaling is also used for the fold bifurcation (one can think of it as a first-order approximation, or directly as $p = p_0 - a't$). Otherwise, combining Erf(t) and a square root scaling yields non-analytical results.

Fig. 5.1 shows expected trends of common statistical indicators, for the three main normal forms and in different dynamical contexts. Although the scaling differs, the qualitative trends are conserved across the bifurcations. This aspect should raise an alert: it is extremely difficult to infer the existence of a critical transitions by simply analysing statistical indicators extracted from systems' data. In fact, the same trends associated to critical slowing down can be observed in supercritical bifurcations (which are not critical in the current sense, *cf.* Sec. 2.3). Studies about scaling (Meisel et al., 2015) and in search of bistability (Sec. 4.3) are first required to infer whether a system undergoes CTs.

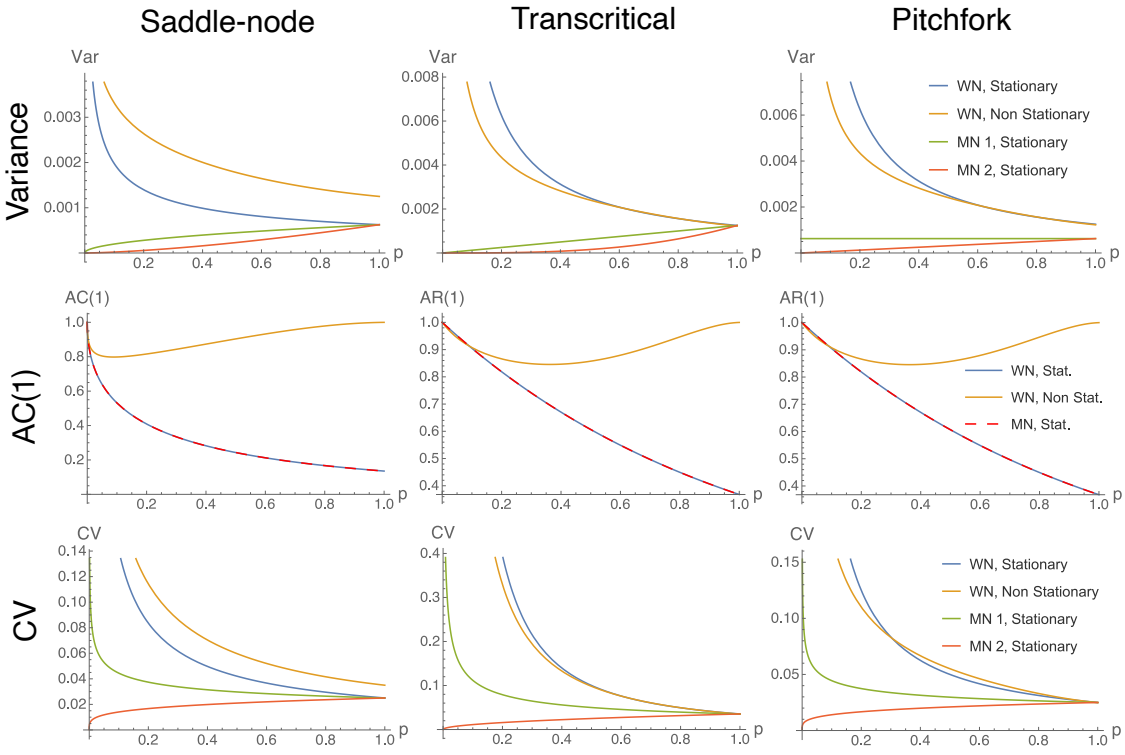


Fig. 5.1 Trends of common statistical indicators (Var, AC(1) and CV) for saddle-node (fold), transcritical and pitchfork bifurcations, in different dynamical contexts (combinations of noise characteristics and stationarity for the control parameter, as explained in the main text). MN 1: multiplicative noise where $h(\tilde{x}, p) = x$; MN 2: multiplicative noise where $h(\tilde{x}, p) = x^2$. As the autocorrelation is independent on noise, only MN 1 is considered, to show that it overlaps with the simple case.

For what concerns the dependency of statistical indicators and EWS on the dynamical context, the present analysis complements what was suggested in literature. Depending on the type of multiplicative noise, the expected increasing trends may be altered or completely disrupted,

resulting in no early warnings prior to tipping points. In addition, commensurable time scales in non-stationary processes may modify the expected patterns, so that raising reliable alerts becomes more challenging.

The dynamical context influences the considered indicators differently. In particular, the autocorrelation is robust against changing noise properties, whereas the variance is more sensitive to them, but maintains its expected trends even in case of ramping parameters. The coefficient of variation is also robust in case of commensurable time scales and copes well in case of certain types of multiplicative noise. Overall, what matters is the competition between changes in noise and changes in resilience: depending on which one is more rapid, the indicators and their associated EWS may perform as expected or fail to anticipate an impending critical transition.

The power spectrum is also altered by multiplicative noise, following Eq. 5.10. Fig. 5.2 shows one example for the case $h(\tilde{x}, p) = x$ (right) compared to the case of white noise (left), for quasi-steady state parameter values. Note that, as substituting $k \propto p$ only induces a different scaling, the drift parameter k is directly employed. Like above, multiplicative noise alters the expected behaviours. Hence, extracting reliable early warnings associated to spectral reddening becomes more challenging and their performance is expected to drop.

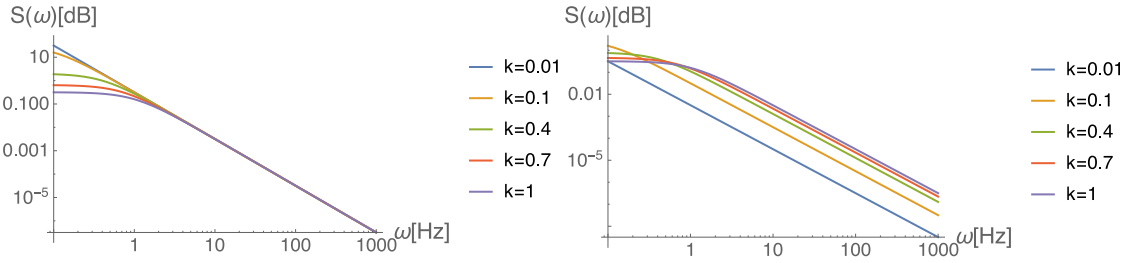


Fig. 5.2 **Left:** power spectral density with white noise, for different values of the drift parameter k. **Right:** power spectral density with multiplicative noise.

Measurement uncertainties

The analytical results from above are further used to study the effect of measurement uncertainties and how they may skew the statistical indicators related to process noise. For this, consider process standard deviation (square root of Eq. 5.10) and measurement standard deviation σ_m as uncertainties, and perform error propagation by summing them in quadrature (Taylor, 1997). This results in:

$$\sigma_{tot}^2 = Var + \sigma_m^2 . \tag{5.14}$$

Here, σ_{tot}^2 represents the total expected (stationary) variance resulting from inner processes and measurement uncertainties.

To derive the autocorrelation, combine its definition Eq. B.21 with Eq. 5.14. In principle, we can explicitly consider multiplicative noise like above. However, the goal in this case is to compute if notable discrepancies exist between ideal measurements (no uncertainty) and realistic measurements (with some uncertainty, that can be filtered to correspond to white noise). Hence,

only the case of white process noise is currently considered. The result is:

$$AC(1)_{\mathrm{m}} = \frac{\frac{\sigma^2}{2k}e^{-k}}{\sqrt{\left(\frac{\sigma^2}{2k}+\sigma_{\mathrm{m}}^2\right)^2}}. \tag{5.15}$$

Obviously, $\lim_{\sigma_{\mathrm{m}}^2 \to 0} AC(1)_{\mathrm{m}} = AC(1)$.

The comparison between statistical indicators computed from ideal measurements and realistic measurements is shown in Fig. 5.3 (for the variance) and in Fig. 5.4 (for AC(1)). The analysis is performed on the fold bifurcation; for the others, it suffices to change the scaling $k \propto p$.

As expected, measurement uncertainties only induces a shift in the absolute values of variance, but do not change its trends (compare Fig. 5.3 left and centre and their scale, as well as the orange an red lines in Fig. 5.3, right). This indicates that the associated EWS are robust against measurement uncertainties and guarantees that they can be extracted from empirical data. Analogous results can be obtained for the CV, that is directly dependent on Var.

Fig. 5.3 also displays the dependency of Var on the process noise σ. Clearly, increasing noise yields increasing variance even far from the bifurcation point $p = 0$. Without proper knowledge or control on noise levels, this might induce spurious signals, not associated to the vicinity of a tipping point.

Finally, Fig. 5.3 shows that increasing noise may change the lead time for triggering early alerts. For higher noise levels, the "elbow" of the Var(p) graph occurs for higher values of p, thus further from the bifurcation point. However, to trigger reliable alerts, any increase should be statistically significant (*cf.* Sec. 4.5.1). Hence, it is necessary to consider the mean trends (as displayed here) as well as the associated uncertainties. This aspect will thus be further investigated in the later section, using simulations to extrapolate uncertainties on statistical indicators from distributional data.

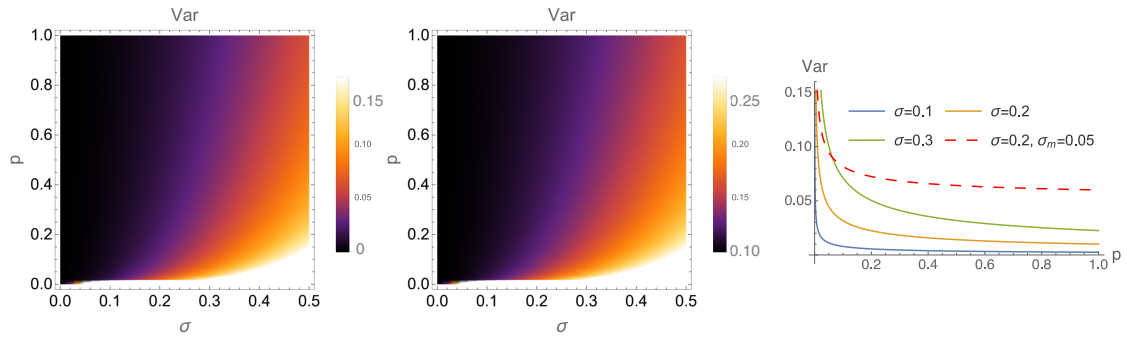


Fig. 5.3 **Left:** Dependency of variance on control parameter p and O-U noise level σ, without measurement uncertainties. **Centre:** Dependency of variance on control parameter p and O-U noise level σ, with measurement uncertainty $\sigma_{\mathrm{m}} = 0.1$. **Right:** Examples of "sliced" graphs from the previous plots (three with increasing σ and one – dashed – with σ_{m} as well), to allow comparison.

Measurement uncertainties have little impact on AC(1) as well, as shown in Fig. 5.4: small uncertainties $\sigma_{\mathrm{m}} = 0.1$ do not modify the trends (compare Fig. 5.4 left and centre and see blue and yellow lines in Fig. 5.4, right). Only relatively high measurement uncertainty levels change

the absolute values of expected lag-1 autocorrelation, but maintain the increasing patterns close to critical points (Fig. 5.4, right).

As already discussed, the autocorrelation does not explicitly depend on noise. Increasing process noise does not alter its expected values, as observed in Fig. 5.4 (left and centre). Autocorrelation is therefore a robust indicator for when system noise is unknown. However, what discussed above also applies here: for a trend to trigger significant warnings, not only the mean trends should be considered, but also their associated uncertainties. They help assessing the performance of indicators and the associated lead times. Such analysis is performed using computer simulations, in the following sections.

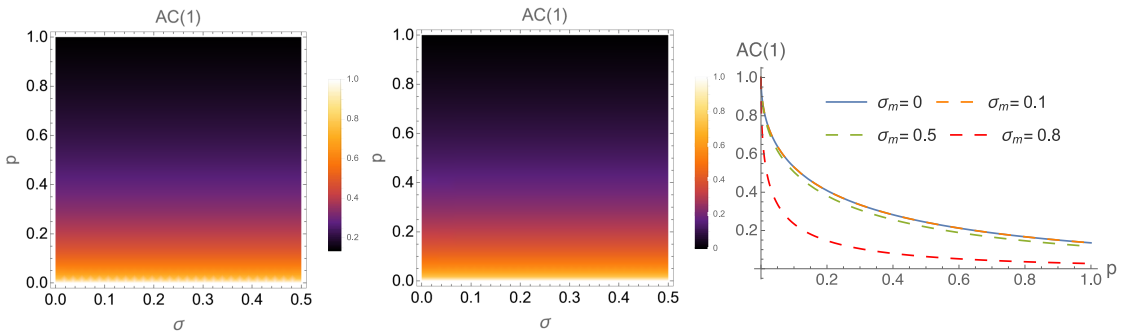


Fig. 5.4 **Left:** Dependency of lag-1 autocorrelation on control parameter p and O-U noise level σ, without measurement uncertainties. **Centre:** Dependency of AC(1) on control parameter p and O-U noise level σ, with measurement uncertainty $\sigma_m = 0.1$. **Right:** Examples of "sliced" graphs from the previous plots (without measurement uncertainty or with three different values), to allow comparison.

CV for "realistic" fold

To subsequently compare theoretical predictions with empirical estimations, let us focus on the fold normal form, which corresponds to the bifurcation observed on experimental data (see Trefois (2014), Fig. 5.15, and below). In this case, the coefficient of variation (CV. eq. 5.11) depends on the mean expected value $\tilde{x}_s$, tracking the stable manifold. On empirical data, it is difficult to normalise by the critical value and fix the normal form around $p_0 = 0$ and $\tilde{x}_s(p) = 0$, since such critical values are largely unknown.

Hence, instead of computing $\tilde{x}_s(p) \to 0$ like on perfectly reconstructed normal forms, one often computes $\tilde{x}_s(p) \to x'_0$, where x'_0 corresponds to the first value where the system tips, but which is unknown a priori. This can be modelled as $\tilde{x}_s(p) = x'_0 + \sqrt{p}$, to be used in the denominator of CV. Hence, Eq. 5.11 becomes

$$CV_r = \frac{\sqrt{Var}}{x'_0 + \sqrt{p}} \tag{5.16}$$

Its effect is displayed in Fig. 5.5. In case of white noise, the expected increasing trend is conserved. However, in contrast to what is observed in Fig. 5.1, other multiplicative noise forms are able to alter the behaviour and shadow possible early warnings.

Similar effects have been observed on computer simulations of epidemiological models (Southall et al., 2020): even thought CV is expected to be robust and well-performing (see also Fig. 5.1), when dynamical mean values induce additional uncertainties, it may perform poorly. This

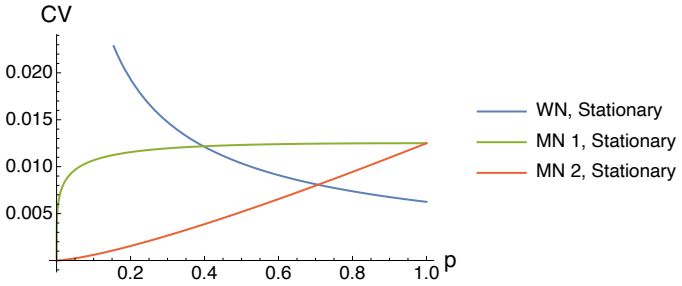


Fig. 5.5 Graphs of Eq. 5.16 for white noise and the two multiplicative noise forms used in Fig. 5.1. Without loss of generality, $x_0' = 1$.

fact is also observed on real-world epidemiological data, see chapter 7. Overall, these observations suggest that CV may be system-specific and less generic than suggested in early literature studies.

Discussing skewness and kurtosis

For certain simulated systems, the third statistical moment (skewness) has been suggested to provide useful early warnings (Guttal et al., 2008). However, experimental results (Dai et al., 2012) were not able to confirm the expectations, estimating a flat and fluctuating trend before a tipping point. Analytical results are thus required to settle the issue, as skewness has continued to be tested among other indicators in different systems (Southall et al., 2020).

For a stochastic process with quasi-steady state parameter, its statistical moments can be expressed as

$$\langle y^n \rangle - \langle y \rangle^n = \int_{-\infty}^{\infty} (y' - \mu)^n P(y') dy' \tag{5.17}$$

where $P(y')$ is the associated probability density function (Sec. B) and μ is the expected average value.

For odd n, if $\mu = 0$ and $P(y')$ is symmetric, the integral equals 0 by definition. Symmetric probability density functions are associated, for instance, with quadratic potentials that are typical of bifurcation normal forms under white noise, for which (Gardiner, 1985)

$$P(y) = \sqrt{\frac{k}{\pi\sigma^2}} Exp\left[-\frac{2}{\sigma^2} U(y)\right] = \sqrt{\frac{k}{\pi\sigma^2}} Exp\left[-\frac{ky^2}{\sigma^2}\right] \tag{5.18}$$

since the quadratic potential $U(y)$ is expressed as $ky^2/2$. Consequently, the normal forms – in particular, the saddle-node – considered above are expected to display a flat skewness.

On the other hand, the integral 5.17 may be non-zero, and even dependent on the drift parameter k, if $\mu \neq 0$ or if $P(y)$ is asymmetrical. In the first case, solving Eq. 5.17 yields

$$\text{Skew} = -\frac{\mu(3 + 2k\mu^2)}{2k}. \tag{5.19}$$

In this case, as $k \to 0$, the skewness is expected to increase, potentially providing an early warning (Fig. 5.6).

On the other hand, an asymmetric potential can be obtained in case of multiplicative noise (Gardiner, 1985; Sharma et al., 2016a); see also Sec. B. Depending on the specific form, it may be

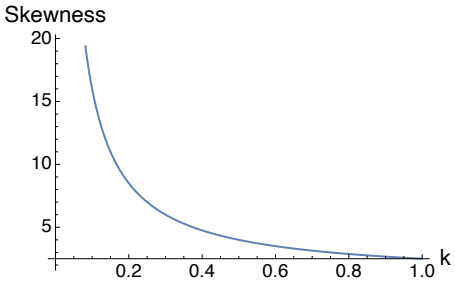


Fig. 5.6 Graphs of Eq. 5.19 for $\mu = 1$

possible to observe increasing trends associated to EWS, but they may be system-specific and not generalisable. In this sense, there is no ambiguity between the results of Guttal et al. (2008) and Dai et al. (2012): they were studying systems with different properties, using an indicator that is not particularly performant and generalisable.

Finally, recall that, for symmetric potentials, all even moments depend on the variance, including the kurtosis (the fourth moment). In case of $\mu = 0$ (typical white noise), kurtosis $= 3\text{Var}^2$. This can be obtained either by solving Eq. 5.17 or by using the well-known Wick decomposition for polynomial averages (Bloch et al., 1958). If $\mu \neq 0$ (or for other exotic noise forms),

$$\text{Kurt} = \frac{3 + 4k\mu^2(3 + k\mu^2)}{4k^2}, \tag{5.20}$$

whose leading term for $0 < k < 1$ still equals Var^2. Hence, variance is already representative of higher moments, which do not improve EWS, unless system-specific noise and drift forms are considered.

A final note: Eq. 5.19 and 5.20 are valid as long as $\text{Re}[k] > 0$, which is true in case of b-tipping. In both cases, noise level σ was normalised to 1 for ease of notation.

5.4.2 Simulation results

To go beyond average trends, computer simulations are used as described in Sec. 5.3.2, considering both white and multiplicative noise. System Eq. 4.1 at quasi-steady state is used.

White noise

To appreciate how indicators evolve with system Eq. 4.1 subject to additive white noise, refer to Fig. 5.7. It shows the eight considered indicators, with their mean values and standard deviations over repeated experiments. Patterns before (right of dash line) and after (left of dashed line) the transition are considered, to study the alerting and/or detecting performance.

For low noise level ($\sigma = 0.02$ used in the figure), the uncertainty region is well constrained around the mean, but higher σ may increase it and therefore modify the significance of increasing trends. Variance, AC(1) and H_S peak at the critical value, whereas others peak slightly later as they are more sensitive to the mean (including that of noise-induced transitions). Visually, skewness and kurtosis do not provide much early warning, but rather detect when all system's repetitions

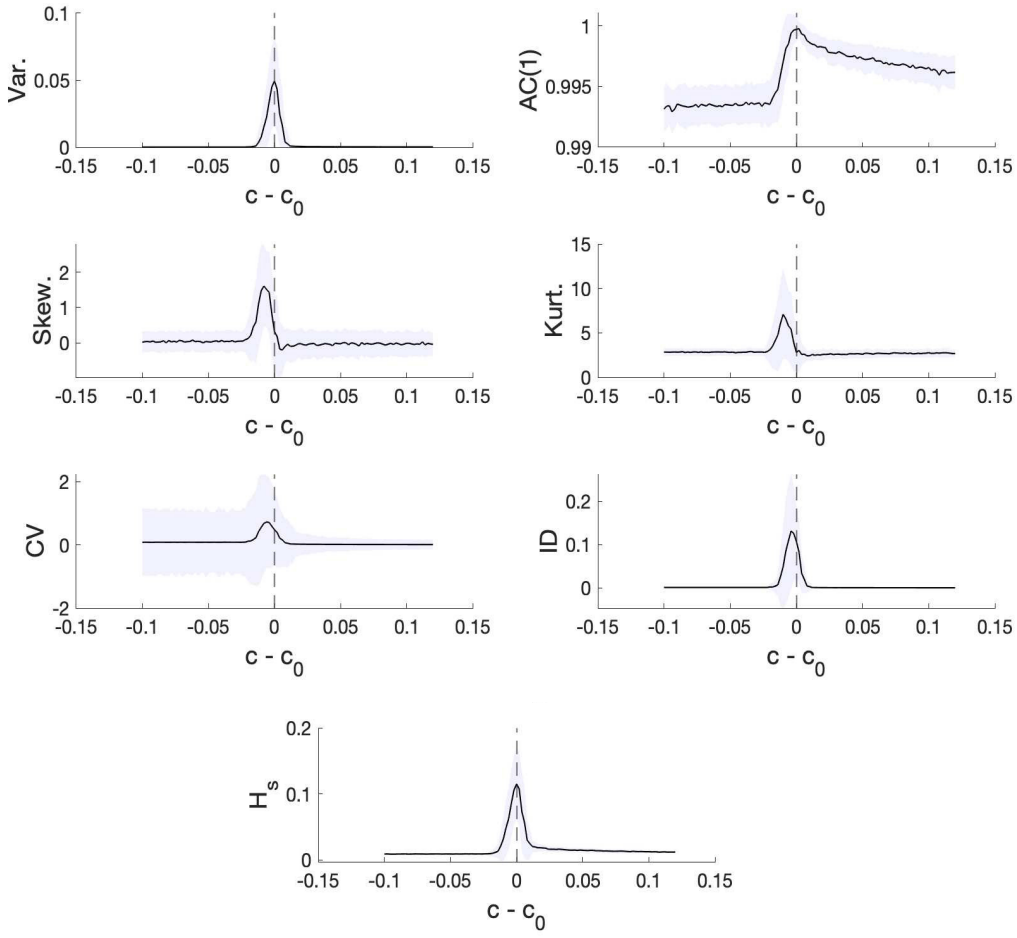


Fig. 5.7 Example of evolution of considered indicators, for each value of the control parameter c. The x-axis indicates the distance of the control parameter to its critical value c_0. The bifurcation point is marked with a dashed line. As the analysis concentrates on the top branch of Fig. 4.2, c is decreased from high to lower values. Hence, the approach to the bifurcation is from right to left. Black lines: mean values over repeated realization. Shaded area: one standard deviation.

tipped. CV and ID display the highest uncertainty bounds, especially after the critical point, as they are the most influenced by mean values (recall that, after the transition, the stable regime is at values close to zero, therefore more sensitive to noise).

To estimate the lead parameter, measuring how much in advance an increasing trend becomes significant (thus including associated uncertainties), the procedure detailed in Sec. 5.3.2 is performed. Fig. 5.8 shows the results, for every considered noise level and indicator.

The most stable indicators – not fluctuating across σ values – are Var, AC(1) and Shannon entropy H_S. CV performs averagely for low noise, then becomes sensitive to large deviations in the mean and its associated c_{sig} drops. Higher order moments and ID behave poorly: they fluctuate considerably at different σ values and often become significant after the bifurcation point (below $c_{sig} = 0$, dashed line).

These results, on top of quantifying the performance of various indicators and assessing their robustness, also inform the subsequent optimization procedure. To construct the set of considered indicators, only the robust and significant ones – variance, AC(1) and Shannon entropy H_S – will be considered.

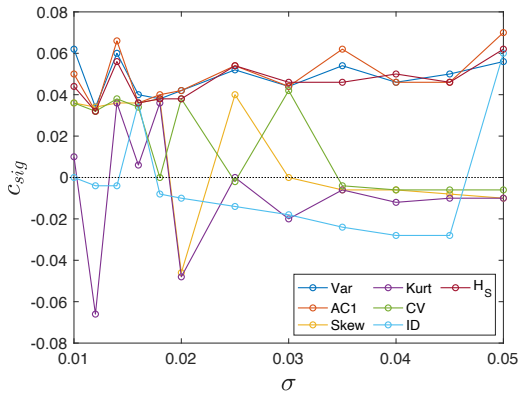


Fig. 5.8 Dependency of c_{sig} (defined in Eq. 5.4) to σ, for each considered indicator $\mathscr{I}$.

Finally, the effect of n-tipping on indicators performance is assessed. In principle, their increasing patterns (associated to EWS) can be due to loss of resilience, but also to tipping events happening during simulation repetitions. For systems biology studies, these would correspond to cells already switching phenotype, while the "average" state of the colony is still the original one. If such n-tipping occurs, increasing trends should be interpreted as "just-on-time" detecting critical transitions in multi-component systems, rather than "anticipating" them *tout court*.

Noise-induced transitions are more likely for higher noise levels (Gardiner, 1985); hence, counting n-tipping events allow to critically evaluate the EWS performance from Fig. 5.8: for instance, the sharp improvement in lead parameter for ID at $\sigma = 0.05$ may derive from several n-transitions happening already and ID picking up the change in mean. Instead, since CV relies on different scaling between variance and mean that produces high uncertainties, its performance becomes poor for large σ.

Fig. 5.9 reports the counts $\mathscr{C}$ of n-tipping events, for each parameter value c and at increasing noise levels σ. Normalising $\mathscr{C}$ by the total number of simulated trajectories allows estimating the probability of n-tipping. For low σ, such probability significantly increases only very close to the tipping point. Increasing statistical indicators are thus expected to relate to losses of resilience. However, for higher noise levels, n-tipping is much more probable when the system is still distant from the bifurcation point. In particular, for $\sigma > 0.04$, some noise-induced events start occurring around $c = 0.05$, which is when EWS may be triggered (see Fig. 5.8). Consequently, for high signal/noise settings, resilience indicators can just-on-time detect critical transitions, but not anticipate them much earlier. For low noise levels, on the other hand, EWS correctly pick up resilience loss, before transitions driven by combinations of noise and bifurcations.

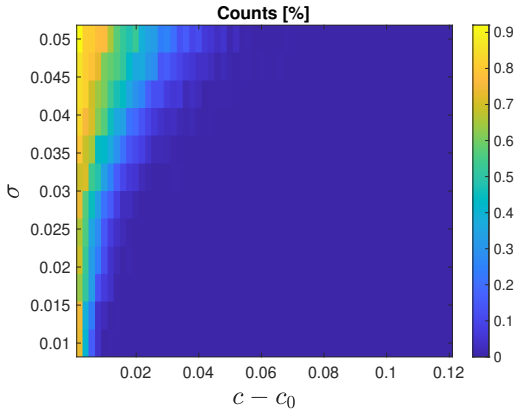


Fig. 5.9 Counter $\mathscr{C}$, counting how many tipping events occur at each parameter value $c - c_0$ and for each noise level σ.

Multiplicative noise

Similar analysis as above is performed by including multiplicative noise in the system. In this case, $g(x) = x$ (from Eq. 4.1). The linear dependency is the simplest to reproduce state-dependency; other nuanced and system-specific types can be considered in future studies. The results, obtained and interpreted as above, are shown in Fig. 5.10.

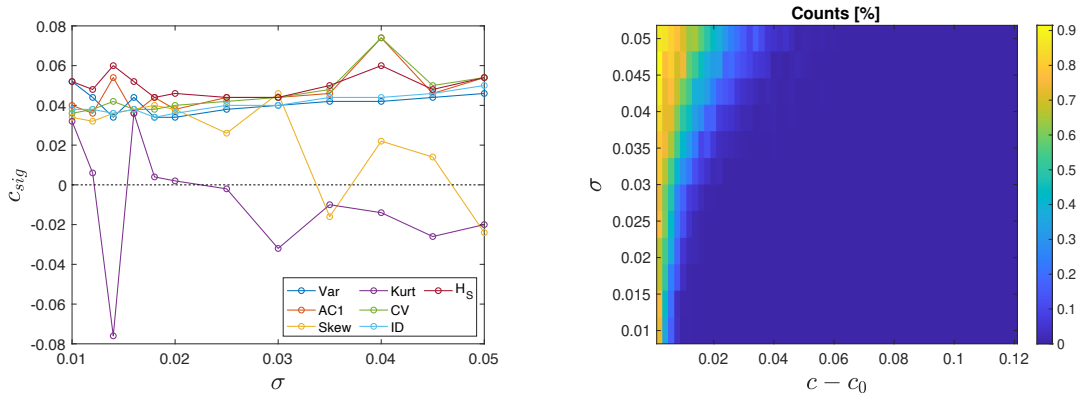


Fig. 5.10 c_{sig} **(left)** and $\mathscr{C}$ **(right)** for the multiplicative noise $g(x) = x$ case. See Fig. 5.8 and 5.9, respectively, to compare with the white noise case.

Like for the white noise case, significant increases can be observed for Var, AC(1) and H_S as the system loses resilience. ID and CV becomes more reliable: due to the interference of multiplicative noise, contributions from Var become more prominent than from mean, so ID and CV follow Var more closely and retain similar performance. On the other hand, high-order moments confirm their poor ability to anticipate and detect the transition.

Like above, for higher σ, the indicators should be interpreted as "just-on-time" detecting resilience loss rather than anticipating it: the interplay of increasing noise and loss of resilience makes for several tipping events occurring much before the critical value, at values comparable with c_{sig}. Once again, this analysis reveals the importance of considering dynamical settings like noise intensity to interpret the signals correctly.

Overall, this analysis performed on multiplicative noise guarantees that, for the specific considered system, the proposed indicators are sufficiently robust and performant. Note that, because of the present focus onto the upper stable branch of the system, we can associate the results to the mean-field approximation still being rather valid. Other studies (Sharma et al., 2016a), looking at the lower branch where small fluctuations play a bigger role, warned that some expected trends may be further altered. Prior knowledge and minimal information about the system under study are thus essential to design experiments and data analysis, in search for meaningful signals.

5.4.3 Optimising EWS

Following the procedure described in Sec. 5.3.2, let us look for the best combination of indicators, that can provide the highest lead parameter c_{sig} to alert for losses of resilience. Note that, in this case, the overall performance over all σ is considered. Hence, the interpretation in terms of "anticipating" or "just-on-time detecting" gets blurred once again. For that, system-specific analysis with curated data is recommended.

Consistently with the rest of the section, both white noise (Fig. 5.11, left) and multiplicative noise (Fig. 5.11, right) cases are subject to the optimisation procedure. Fig. 5.11 shows the scores S from Eq. 5.6, for the various combinations of weights $\mathbf{w}$ (called "Combination index #"). The problem is non convex due to inter-dependencies of indicators. Hence, a grid search with stride 0.1 is used. To interpret Fig. 5.11, note that H_S increases monotonously from left to right, Var follows a saw-tooth pattern (highest value to lowest, with highest = $1, 0.9, 0.8, \ldots$ from left to right) and AC(1) completes the combinations.

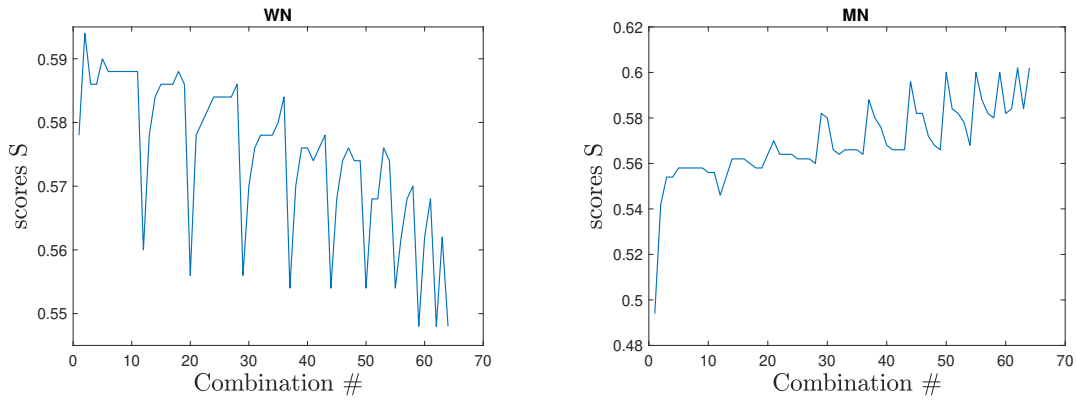


Fig. 5.11 Score S for white (**left**) and multiplicative ($g(x) = x$, **right**) noise, depending on weight different combinations, indexed on the x-axis.

For the white noise case, the maximum score is achieved for $\hat{\mathbf{w}} = [0.9, 0.1, 0]$, that is, high weight for variance, low for autocorrelation, and no contribution from Entropy. In general, the high plateau on the left corresponds to $w_{H_s} = 0$. As Shannon entropy, in this case, is directly dependent on variance (see Sec. B), it does not contribute significantly to improving the overall c_{sig}. Since significant increases depend both on mean behaviours and on uncertainties, preferring simple indicators with low uncertainties is a winning strategy.

For multiplicative noise, the picture changes. Here, including entropy with high w_{H_s} provides the best performance. Note that calculations for entropy in Sec. B do not hold here, but require other analytical derivations currently not solved. Mixing low contributions from Var and AC(1) do not particularly help. The maximum is reached for $\hat{\mathbf{w}} = [0,0,1]$. Looking at Fig. 5.10, this preference for entropy is mostly driven by cases with low σ.

In the multiplicative noise case, on top of mean trends and uncertainties, the interplay between noise and dynamics alters the performance of early warning signals. Hence, interpreting the results becomes more complicated and, for real-world systems, may be further optimised with data-driven procedures.

Overall, the optimisation analysis confirms that, depending on dynamical features like noise properties, one or another indicator may be more performant. If one or the other noise type can be modelled *a priori*, then the indicators for EWS can be readily selected. In realistic settings, however, mixed noise types may concur in adding memory to the system or altering the dynamical responses. The section below address those cases.

Other multiplicative noise classes

The analysis above shows that multiplicative noise requires different optimal combinations of indicators to provide early warning signals. Is this consistent for other classes of multiplicative noise, or the optimal combination depends on the specific noise form?

To test this question, other two types of multiplicative noise are used, which are common in literature for biological systems. The first type assumes $g(x) = x^2$, *i.e.*, it is proportional to the deterministic part of the corresponding normal form (O'Regan et al., 2018) and largely amplifies the fluctuations. The second type assumes $g(x) = \frac{x^2}{1+x^2}$, *i.e.*, it is proportional to the deterministic part of the original system (Hasty et al., 2000).

Both cases return consistent results with the above multiplicative noise case. Although c_{sig} and $\mathscr{C}$ change due to different random perturbations, the optimization procedure returns Shannon entropy H_S as the best indicator to detect resilience loss before transitions begin. In fact, for both additional noise cases, $\hat{\mathbf{w}} = [0,0,1]$.

This means that, in cases where multiplicative noise adds a layer of complexity to the mean field regime, entropy is in general the best performing indicator. Hence, when facing a system of interest belonging to the same class as the one considered here, it suffices to determine the type of noise – white or multiplicative, if the system allows a binary choice – to decide whether to use variance or entropy to monitor its resilience.

The present analysis concentrates on the upper branch of system Eq. 4.1. On the lower branch, intrinsic and extrinsic noise effects become more prominent due to low molecule concentrations. Nonetheless, Sharma et al. (2016a) observed preliminary result that are in line with the present ones, so one could expect that a similar optimization procedure could output consistent results.

Combination of additive and multiplicative noise

Gaussian additive noise is usually considered an adequate diffusion approximation of many small amplitude Poisson noises, but experimental results show that gene transcription very often occurs

in a bursty, unpredictable and intermittent manner (Chen et al., 2017). These events are mostly due to "intrinsic" stochasticity, *i.e.*, determined by structure, reaction rates, timescales and species concentrations of the underlying biochemical processes. The resulting noise distributions are described as heavy-tailed and are poorly captured by diffusion approximations (Pan et al., 2010). Instead, differences between cells after measurements, often called "extrinsic noise" (Pan et al., 2010), are supposedly well captured by Gaussian noise models.

Instead of using simple white noise, several studies showed that correlated noise is more realistic (Zhang et al., 2012; Sidney et al., 2010), as it aptly combines properties of intrinsic and extrinsic noise, and fits experimental data better than purely additive processes (Wang et al., 2012b). In addition, correlated noise is what is observed from dynamical single-cells experiments (Pedraza et al., 2008; Dunlop et al., 2008). One way of modelling it is to consider time-delayed or coloured noise (Wang et al., 2015; Wang et al., 2016). Alternatively to coloured noise, combining additive and multiplicative noise is a valid modelling alternative (Liu et al., 2009). For some systems like genetic toggle-switches, it is actually a better modelling choice, as the concentration of reactants is non-negative (Chen et al., 2017).

Hence, this section investigates optimal combinations of indicators for EWS if both additive and multiplicative noise are present in the system and reverberate onto the data. To do so, take the system Eq. 4.1 equipped with noise terms $g(x) = \alpha + (1-\alpha)x$, where α weights the white or multiplicative noise component ($\alpha = 1$ corresponds to purely additive white noise, $\alpha = 0$ to purely multiplicative). Note that, following the above results about the generic performance of EWS indicators for different types of multiplicative functions, the simple linear state-dependent relation is used as a reference.

The optimisation procedure is repeated like previously. In this case, the optimal combinations $\hat{\mathbf{w}}$ may further differ. Fig. 5.12 summarises the results. For different α, $\hat{\mathbf{w}}$ changes. If the additive noise content is higher, Var is preferred as an indicator to maximise the lead parameter. For cases with prevalence of multiplicative noise, entropy is better. For intermediate realistic cases, a mixture, with AC(1) also taken into account, optimises the performance.

Note that, since the problem is non-convex, there are cases where more than one combination solves the optimization. If that is the case, Fig. 5.12 reports the two alternatives (same colour, but dashed line). For the same reason, there is not a unique law but $\hat{\mathbf{w}}$ may slightly fluctuate. However, the trends described above remain clear.

This analysis suggests that, in realistic settings with correlated noise due to intrinsic and extrinsic stochasticity, combining indicators yields higher lead parameters and allows earlier detection of fragile states.

In particular, choosing "sub-optimal" states leads to lower S, which translates in lower and more uncertain c_{sig}. Changing $\hat{\mathbf{w}}$ by about 20% induces a relative drop in S of 4%, on average over the various α, with greater drops for $\alpha > 0.5$. As a comparison, setting the weights on "non-optimal" values (e.g. using full variance in case of multiplicative noise) results in relative drops of S up to 18%. Thereby, on average, the loss of resilience is not detected immediately, but the EWS react to critical transitions that already happened. Hence, depending on the study tolerance levels and the experimental uncertainties, selecting sub-optimal $\hat{\mathbf{w}}$ may still yield acceptable results, in particular for realistic combinations of additive and multiplicative noise; however, choosing

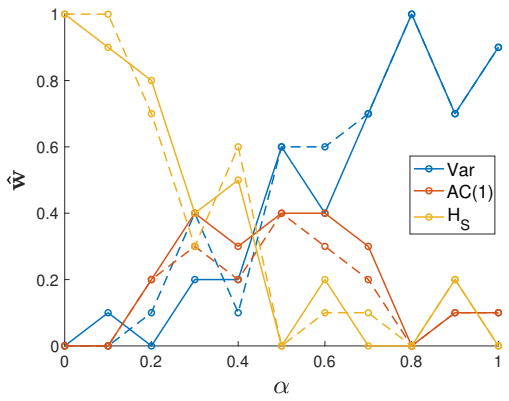


Fig. 5.12 Optimal weight sets $\hat{\mathbf{w}}$ for each α. $\alpha = 1$ corresponds to pure white noise, $\alpha = 0$ to multiplicative. Dashed lines correspond to alternative $\hat{\mathbf{w}}$.

non-optimal weights could results in significant drops of performance which should be improved with subsequent iterations.

5.4.4 Comparison with experimental data

Consider the system described in Sec. 5.3.3 and its derived experimental data. $\Delta\Psi_m$ measures mitochondrial membrane potential (state variable) and Antimycin A concentration is the control parameter.

After some transient time, Antimycin A impairs mitochondrial functions and drives the system into a bistable region, followed by critical transitions. Fig. 5.13 (right) shows the bifurcation diagram, at steady-state (all cells evaluated at time=22h), corresponding to mean values for each mitochondrial condition measured by $\log(\Delta\Psi_m)$, estimated with Gaussian mixture models (see Trefois (2014) for details). For the experiments, Antimycin A was increased from low to high concentrations (in nano-molars); Oligomycin A is kept constant at 2 g/ml (refer to Table 5.2 of Trefois (2014)). Bistability is observed for Antimycin concentrations between 8.7 and 15.5 nM.

Note that the experiments were performed only by increasing Antimycin concentrations, without reducing it. Hence, observing values on the lower branch means that the system was able to switch before reaching the tipping point, a likely sign of noise-induced transitions. The blue point in the middle of the imputed unstable branch can be interpreted similarly – a value sustained by random fluctuations.

After assessing bistability, the coefficient of variation (CV) was estimated as a measure of variability, to test the emergence of early warnings. Fig. 5.13 (left) shows CV values at steady state (all cells evaluated at time=22h, after transients), as function of different conditions corresponding to increasing concentrations of Antimycin A, from 8.7 nM to 18 nM. It is rather flat until condition 16, where it has a small drop followed by an increasing trend. According to Table 5.2 of Trefois (2014), condition 16 is right before the system enters its imputed region of bistability.

As expected, CV does not display notable patterns before entering the bistable region. On the contrary, its subtle dip may initially trigger a spurious signal and can be interpreted as a noisy effect. Then, CV values strongly increase from the 15th to the 36th condition under steady-state

assumption, as bimodality unfolds. However, it does so much before reaching the bifurcation point. This aspect, together with the observation from above about likely n-tipping occurring in the system, suggests that CV increases are strongly driven by the mean suddenly switching to alternative, low values. Augmented variability thus amplifies and smoothens the transitions to alternative means ($\langle \Delta\Psi_m \rangle$), but it is unlikely to anticipate the critical regime. In this case, CV works fine for just-on-time detection (it identifies when few cells begin tipping) but not for global anticipation.

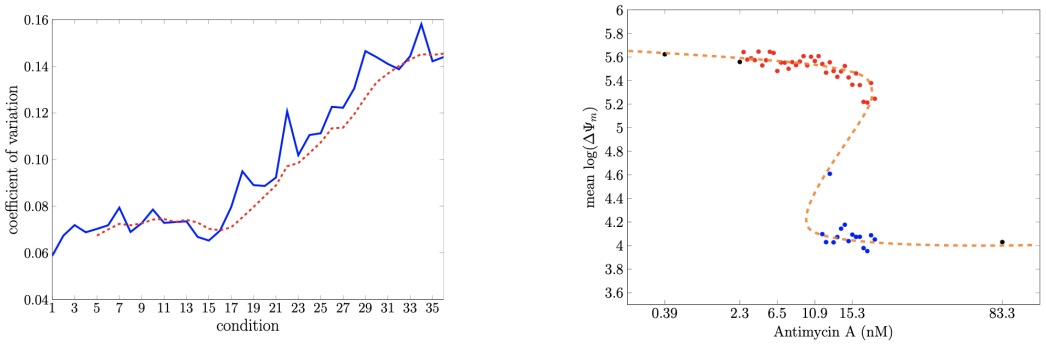


Fig. 5.13 **Left:** Increasing CV at steady-state, for the mitochondrial experiments. Blue line: CV over all conditions under steady-state (t=22h). Red line: smoothing with a 5-step window. Figure corresponding to Fig. 5.9 of Trefois (2014). **Right:** bifurcation diagram at steady-state (t=22h). Red dots: estimated high state $\Delta\Psi_m$ averages, μ_1. Blue dots: low state $\Delta\Psi_m$ averages, μ_2, for conditions greater than 22. Black dots: average $\Delta\Psi_m$ for concentrations of 0.39, 2.3 and 83.3 nM Antimycin A (two states before the bimodal region and one after it). Figure corresponding to Fig. 5.15 of Trefois (2014). All figures reproduced with permission from the author.

5.5 Discussion

The present chapter presents a systematic analysis on the expected behaviour and performance of CSD-related indicators and associated EWS. The results are significant for several reasons.

First, they complement previous literature on the dependency of EWS to dynamical contexts and inform subsequent investigations on performance, observability and interpretability of EWS. In particular, the chapter proves that noise effects and commensurable time scales may alter EWS behaviours and drastically diminish their performance. Hence, the current results support existing suggestions (Dutta et al., 2018) that EWS are actually not generic *in toto*; nonetheless, *sub-classes* of CTs exist, defined by the considered dynamical contexts, where EWS are robust and generically applicable to numerous systems. In addition, the results illustrate the importance of considering realistic settings – measurements with uncertainties, non-normalizations, etc. – when making predictions to contrast against data, that should be properly collected and curated. Finally, the chapter alerts that extracting some patterns in statistical indicators is not sufficient to infer the existence of critical transitions. Overall, the chapter charts potentials and limitations for the application of EWS in real-world systems.

Second, the analytical and general results allow to interpret numerous observations from simulation and experimental studies, in various research fields. Some examples have already been provided in the chapter, *e.g.* the arguments between Guttal et al. (2008) and Dai et al. (2012) on the role of skewness, or the observations of Brett et al. (2017) about commensurable time scales and of Southall et al. (2020) about the performance of CV. Another interesting point is the interpretation of experimental estimation of CV trends from critical transitions in mitochondrial membrane potential: theoretical results help understanding the role of noise in driving the transitions and modify the behaviour of EWS.

Overall, this analysis helps understanding EWS trends in different contexts. For future works, re-interpreting other existing arguments (see chapter 3) in light of the current results may provide additional insights.

Third, the analysis allows to assess whether the indicators are "anticipating" or "detecting" the loss of resilience. According to the interplay of approaches to bifurcation points and noise levels, proposed EWS may succeed in detecting resilience loss before tipping events, therefore anticipating critical transitions within systems of interest, or simply detecting the very first tipping events. Such analysis calls for caution against claims and overoptimistic applications of proposed EWS in real-word monitoring protocols. Even when looking at the optimal combination of statistical indicators to provide early warning signals, the existence of underlying tipping events cannot be ruled out. In the example of the genetic toggle switch, the analysis reveals that many indicators become significant when relevant numbers of cells already changed phenotype. Keeping in mind this epistemic distinction is important in real-world applications, as further discussed in Sec. 8.5.4.

Finally, it is possible to provide a set of practical guidelines for using certain resilience indicators to address specific systems. When the system is known (or suspected) to evolve over time scales that are not "slow-fast", it is best to use variance and derived indicators. Variance is also the best indicator (possibly, combined with autocorrelation) in case of slow-fast systems subject to white noise, and should be preferred to other ones. In case of multiplicative noise of various types, however, autocorrelation is more robust, while increasing entropy is in general the best performing EWS. In case of realistic mixture of additive and multiplicative noise, however, combining multiple indicators improves the results obtained with univariate strategies.

These results have immediate consequences in routine applications. For instance, two contrasting indexes were proposed to evaluate cell differentiation stages, one based on variance (Chen et al., 2012b) and another on autocorrelation (Mojtahedi et al., 2016). If cell systems are expected to evolve slowly (w.r.t. their typical time scales) but are characterised by combinations of intrinsic and environmental noise (Bruggeman et al., 2018), the second strategy should be preferred. Moreover, developing additional methods based on entropy, derived from cell distributions, will further improve the detection and anticipation capabilities. Future studies, improving the present analytical results and considering additional contexts, will shed light onto their potential use for online alerting.

By carefully considering EWS robustness and sensitivity to dynamical features, this analysis observes when confounding and spurious signals may appear. When bistability properties or

tipping points are not known, relying only on early warning signals from the CT framework may yield false positives (an increase is observed, but not associated to impending critical transitions) or false negatives (no significant increase is observed because it is shadowed by noise or commensurable time scales). Other dynamical contexts, not considered here, may further skew the expected signals and trick our interpretation, in particular during online monitoring. Overall, EWS performance for detection and early alert needs to be refined for classes of systems and tested against empirical data. Building upon the current results, this task is performed in the subsequent chapters.

Code availability

The code to reproduce the results is available at https://github.com/daniele-proverbio/EwsSensitivityPerformance

Chapter 6

Resilience properties and buffers for biological variability

Adapted from: D. Proverbio, A.N. Montanari, A. Skupin, J. Gonçalves. *Buffering variability in cell regulation motifs close to criticality*. Physical Review E, 106: L032402, 2022 (Proverbio et al., 2022a)

Authors Contribution

DP: Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Roles/Writing - original draft, Writing - review & editing. ANM: Formal analysis, Investigation, Roles/Writing - original draft, Writing - review & editing. AS: Investigation, Supervision, Writing - review & editing. JG: Funding acquisition, Project administration, Supervision, Writing - review & editing.

6.1 Motivation

Previous chapters and empirical literature (Dai et al., 2015; Radchuk et al., 2019) suggest that interplays between resilience and stability properties may concur in determining how systems cope with fluctuations and how performant EWS against impending transitions may be. To inquire these aspects, a tractable model of high biological relevance is here considered. The model is a self-activating positive feedback loop, a common regulatory motif found in viruses (Hasty et al., 2000) and cell circuits (Alon, 2019) producing critical bifurcations with different properties.

This study has two main purposes. First, to identify and characterise buffering mechanisms, *i.e.*, suppressors for fluctuations and noise-induced regime shifts, close to criticality. Second, to uncover system's resilience properties from data, when the regulatory mechanism is not known in details. EWS derived from the CT theory are thus tested in situations where the positive feedback loop fold bifurcation is altered by extra parameters. This analysis demonstrates the usefulness of single metrics for multivariate systems. Since it relies on minimal models, the present framework can be used to extract robustness and variability properties of other more complex models and empirical data close to criticality.

6.2 Introduction and modelling background

Many biological systems self-regulate their functions through bistable circuits, which have been associated to genetic (Angeli et al., 2004b; Kheir Gouda et al., 2019) as well as growth feedbacks (Deris et al., 2013). As introduced in Sec. 4.1.1, positive feedback loops are key regulatory mechanisms deserving paramount interest by systems biologists. Close to criticality, the dynamical motif has reduced resilience and the system can exhibit augmented variability in response to noise (Mojtahedi et al., 2016; Sharma et al., 2016a). This is typical of non-linear systems approaching a critical bifurcation (Scholz et al., 1987; Scheffer et al., 2009; Byrd et al., 2019) and corresponds to an augmented sensitivity of the system to random perturbations. The phenomenon corresponds to critical slowing down (CSD, see Sec. 2.4.1). Mechanisms to buffer variability while maintaining the critical state are thus necessary to finely control desired transitions (Frank et al., 2002; Qian, 2012), or to better cope with undesired shifts (Scheffer, 2009). In fact, cells and other biological systems are thought to live close to criticality to quickly respond to changing environmental conditions (Mora et al., 2011), but they should not respond to random changes (noise), in order to maintain their evolutionary fitness. Hence, strategies to control drastic switches through b- or n-tipping (Sec. 2.3) and to buffer variability should be recognised and deeply studied

To this end, two main dynamical strategies can cooperate: moving the system state away from the bifurcation point, or deepening the basin of attraction to avoid random fluctuations pushing the system state to undesired attractors. These strategies correspond to changes in different environmental or regulatory conditions in cellular systems (Dai et al., 2015), allowing organisms to exploit different mechanisms to buffer variability close to criticality. For instance, the "sponge" action of molecular compounds (Ebert et al., 2012; Siciliano et al., 2013) is a known biological buffer, helping to hamper fluctuations around equilibrium states and preventing noise-induced transitions. Phase separation within cells is also known to provide such effects (Klosin et al., 2020), as well as temporal relay of signalling molecules (Lestas et al., 2010). However, dynamical mechanisms close to critical states are less studied. This work identifies one such mechanism to buffer variability; specifically, the role of cooperative interactions that tune the activation function of positive feedbacks.

To illustrate this, consider a simple an well-known a-dimensional model for stochastic autoactivating positive feedback (Santillán, 2008; Strogatz, 2018):

$$\dot{x} = K + c\frac{x^n}{1+x^n} - x + \eta(t)\ . \tag{6.1}$$

The model describes protein concentration dynamics after RNA transcription by Michaelis-Menten kinetics around equilibrium reactions (Frigola et al., 2012; Weber et al., 2013). Sec. 4.1.1 discusses the model in details. Let us recall the key information. Variables and parameters are: x for the protein concentration, with steady state $\tilde{x}$, *i.e.*, $\dot{x}|_{\tilde{x}} = 0$, K for the basal expression rate, and c for the maximum production rate. The deterministic terms can be grouped into a single function $f(x)$. The critical value c_0 for the control parameter marks a saddle-node bifurcation. The dissociation constant (in denominator of Hill function) was normalised to 1 without loss of generality (Smolen et al., 1998). However, it can be re-included as an additional constant q, yielding a modified activation function $x^n/(q+x^n)$.

The noise term $\eta(t)$ accounts for intrinsic stochasticity of biological processes (Hasty et al., 2000). Only additive noise is considered here, to approximate the fast degrees of freedom in a mean-field regime (Berglund et al., 2006). Noise has the following statistical properties: $\langle \eta \rangle = 0$, $\langle \eta(t)\eta(t') \rangle = 2\sigma\delta(t-t')$, where σ represents its intensity. Multiplicative (state-dependent) noise, introduced and discussed in chapter 5, is not primarily investigated here, to concentrate on the effects of extra parameters on stability properties. In fact, the interplay of cooperativity and multiplicative noise adds a layer of complexity; it is briefly discussed in Sec. 6.5 and is suggested to future studies. For an introduction on simple motifs, see Sharma et al. (2016a).

The Hill coefficient n is a non-linear term modelling cooperative binding mechanisms. It is usually interpreted as the number of transcription factors (TF) that cooperatively promote transcription (Santillán, 2008). Literature suggests that n is associated to measured noise distributions (Maienschein-Cline et al., 2010), but it is less clear how n concurs in controlling fluctuations and transitions close to critical states. To have bistability in the circuit, $n \geq 2$ is necessary. When $n \to \infty$, the system is modelled using the logic approximation for the activating function:

$$\lim_{n\to\infty} \frac{x^n}{1+x^n} = \Theta(x-1)\,. \tag{6.2}$$

Here, $\Theta(\cdot)$ is the Heaviside step function that turns system (6.1) into a perfect toggle without bistability nor CSD.

In this chapter, two main investigations are conducted on Eq. 6.1.

The first one inquires the dependence of stability properties on the cooperativity index n. To begin with, stability properties are quantified by the stochastic potential and the stationary probability density function (Sec. B) associated to states close to criticality. They show that increasing cooperativity values do not alter the underlying bistability, but modify the depth of the potential and increase the separation of alternative states. Then, system sensitivity to noise close to critical points is investigated in terms of changing variability measures – variance and autocorrelation. This way, it is tested how cells can keep production rates c close to their critical values and nonetheless increase resilience and buffer variability using other regulation mechanisms.

Since the presence of non-linear binding interactions often leads to analytically intractable models (Cao et al., 2018), a methodology that combines analytical and geometrical analysis on normal forms (Sec. A) is employed. They are used to derive approximate closed-form solutions for variance and autocorrelation, which measure systems' variability close to criticality.

The second investigation considers the impact of changing system resilience to EWS. The question is: since varying values of n alter some quantitative bifurcation properties, can we still rely on CT-based early warning signals to detect an impending tipping point? The goal is to assess in which parameter range the CSD-based early warning signals correctly indicate an impending regime shift. To better capture noisy effects in b-tipping, which might not be included in approximated analytical results, numerical simulations are employed for this part of the study. They are detailed in the sections below, together with the justification for the choice of parameter values.

6.3 Stability properties

Increasing values of n modify the bifurcation diagram for Eq. 6.1, as shown in Fig. 6.1, left. Hence, they do not alter the existence of bifurcations, like K would (Fig. 6.1, right); however, they modify their quantitative properties and, possibly, the associated system's resilience and the way fluctuations are buffered. Note that the following analysis concentrates on the upper branch of bistable systems (to be consistent with the mean-field assumption) and moving leftwards to the saddle-node bifurcation point. Consequently, K is always chosen appropriately (*cf.* Fig. 6.1; $K = 0.1$ is used without loss of generality).

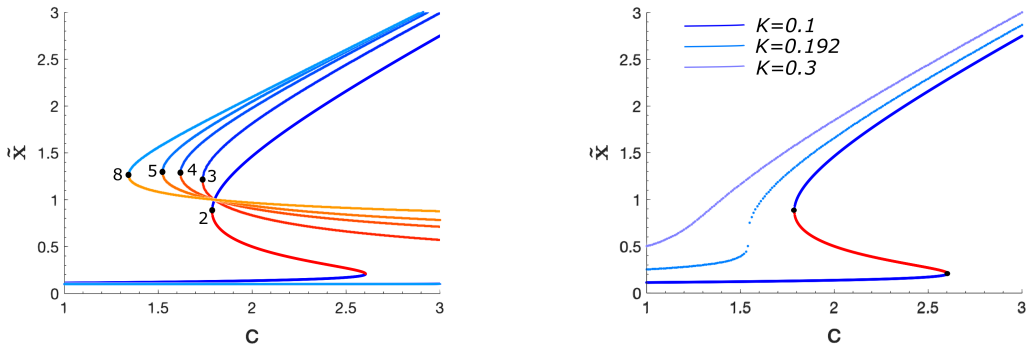


Fig. 6.1 **Left:** Bifurcation diagrams $(\tilde{x}, c)$ of Eq. (6.1) for different Hill coefficients n. Stable and unstable branches are represented by blue and red colours, respectively. Black dots identify the bifurcation (saddle-node) points c_0. Each value of n is displayed close to its corresponding diagram. **Right:** Bifurcation diagrams $(\tilde{x}, c)$ of Eq. (6.1) for different K, which acts as a cusp parameter (Sec. A). It yields bistability or not depending on its value range. For $n = 2$, bistability occurs for $0 < K < 1/(3\sqrt{3})$ (Weber et al., 2013).

6.3.1 Probability density function

To characterise system stability properties, stationary potentials and probability density functions (PDF) are analysed in their dependency on n, in analoogy to previous works (Friedman et al., 2006; Kumar et al., 2014). These characteristics are also useful to derive probability transitions rates and mean passage times (Sec. B and Zheng et al. (2011)).

Consider the forward Fokker-Plank equation (Sec. B) for the probability density function $P(x,t)$ associated to Eq. (6.1):

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[f(x)P(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[\sigma P(x,t)\right]. \tag{6.3}$$

Here, $f(x)$ lumps the deterministic terms of Eq. (6.1), *i.e.*,

$$f(x) = K + c\frac{x^n}{1+x^n} - x$$

The stationary solution $P_s(x)$ is biologically interpretable as the PDF after rapid transients. Its takes the form (Sec. B)

$$P_s(x) = N_c e^{-\phi(x)}, \tag{6.4}$$

$$\phi(x) = \frac{1}{2}\ln\sigma - \frac{1}{\sigma}\int^x f(x')dx'. \tag{6.5}$$

Here, $\phi(x)$ describes the adjoint stochastic potential and its depth is related to system resilience, *i.e.*, its ability to recover after a perturbation. N_c is a normalization constant, so that $\int_\Omega P_s(x) = 1$ (Ω is the domain). It is computed with corresponding functions of the Mathematica software.

Fig. 6.2 shows the dependency of $\phi(x)$ and $P_s(x)$ on n when the system is either in an "off" state far away from criticality (Fig. 6.2a,b), close to the criticality (Fig. 6.2c,d) or beyond it, where the "on" state is favoured (Fig. 6.2e,f). Fig. 6.3 provides a complementary view for the stationary PDF, that is, its phase portrait $(P_s(x), \partial_x P_s(x))$. Each loop identifies a peak in $P_s(x)$. The bistable region is therefore characterised by two loops, of various "radius" and position depending on n.

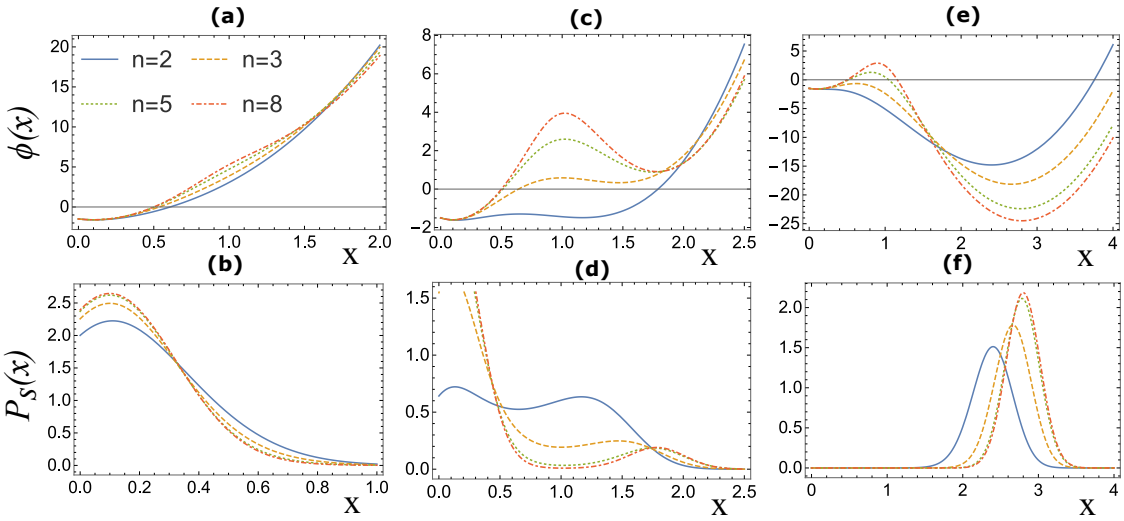


Fig. 6.2 Effect of the Hill coefficient n on the stochastic potential $\phi(x)$ (left column) and on the stationary probability density function $P_s(x)$ (right column), when only additive white noise is present. **(a, b)** $c = 0.8$ ("off" state); **(c, d)** $c = c_0(n) + 0.05$ (multistable region); **(e, f)** $c = 2.7$ ("on" state favoured). In all cases, $K = 0.1$ and $\sigma = 0.05$.

Consider Fig. 6.3. Increasing the cooperativity index n does not alter the underlying bistability, but modifies the depth of the potential and increases the separation of alternative states For the "off" and "on" states, the corresponding equilibria exhibit significantly deep attractor basins with only minor dependence on n, as also indicated in the bifurcation diagram (Fig. 6.1) and by $P_s(x)$.

In the bistable region close to critical points, the situation changes. The potential $\phi(x)$ displays two commensurable wells, which are more evident and symmetric for larger n, suggesting that both states become equally occupied in noisy environments. For increasing n, $P_s(x)$ displays sharper peak separation between the bistable states: the system diffuses less to intermediate states and is more constrained around single equilibrium values, as anticipated due the steeper potential

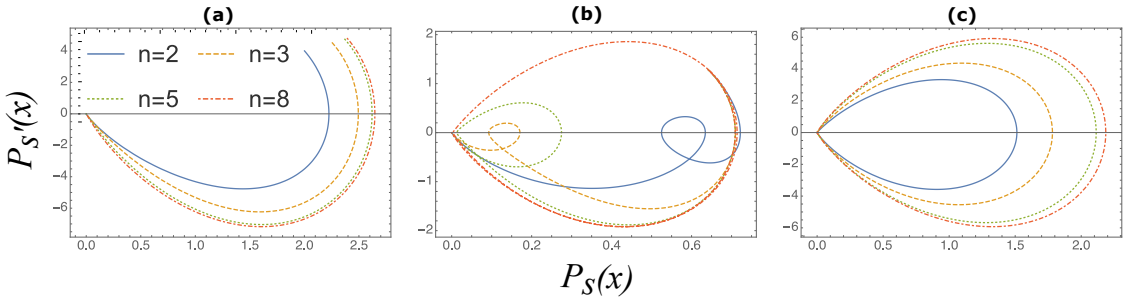


Fig. 6.3 Phase plane of the stationary probability density function ($P_s(x)$ versus its first derivative). Each loop identifies a peak, corresponding to a stable steady state. Panels correspond to: **(a)** Monostable region, on lower branch ("off" state). The loop is not complete as values below 0 for $P_s(x)$ are not biologically consistent. **(b)** Bistable region, close to criticality. Two loops exist for each n, identifying the two possible solutions. **(c)** Monostable region, on upper branch ("on" state), with a single loop highlighting the only stable solution. Panels (a,b,c) correspond to panels (a,b), (c,d), (e,f) in Fig. 6.2.

barriers in $\phi(x)$. Random deviations are thus suppressed faster and transitions from one state to another are sharper and are therefore more robust against noise.

6.3.2 Variability measures

The close question is: close to criticality, how does n influence variability measures like variance and autocorrelation? Analytical derivations for variance and autocorrelation are known for minimal models (Gutierrez et al., 2009), but they become more challenging for high n. Nonetheless, the present goal is to use a generic framework, that can be easily extended to other resilience patterns in complex networks (Gao et al., 2016) or to other complex activation functions. Hence, a useful strategy is to focus on local analysis close to bifurcation points and to employ a geometrical methodology.

To derive generic results for critical manifolds, using their local topological equivalence to bifurcation normal forms is useful (Sec. A and Kuehn et al. (2021)). For saddle-node bifurcations like in Fig. 6.1, the associated normal form is

$$\dot{x} = p - x^2 \,, \tag{6.6}$$

with two equilibrium manifolds $\tilde{x}_{1,2} = \pm\sqrt{p}$, one stable $(+)$ and the other unstable $(-)$. Note that the normal form corresponds to a parabola. To study the behaviour of stochastic solutions near the stable manifold, consider the evolution of its first-order perturbation, $y = \delta x|_{\tilde{x}_1}$ exposed to the same additive white noise $\eta(t)$ as in Eq. (6.1) (Kuehn et al., 2011). Now, perform a simple change of notation:

$$k = 2\sqrt{p} \,. \tag{6.7}$$

Eq. 6.7 corresponds to the distance of the control parameter p from its critical value $p_0 = 0$. This way, k is proportional to $c - c_0$ from the original system. The corresponding Langevin equation, to describe mean-field fluctuations around the stable equilibrium, is given by a typical

Ornstein–Uhlenbeck (OU) process with exact solutions for statistical moments (Gardiner, 1985):

$$\dot{y} = -ky + \eta(t)\,. \tag{6.8}$$

For simple systems with known mechanistic model, one can directly use the analytical Eq. A.8; in the present case, the calculation is performed in Sec. 6.3.3. However, it would be impossible for incomplete models or for studies based on experimental data. Hence, another method is proposed.

To connect the quantitative effects of n with the more qualitative topological form Eq. (6.6), recall that n widens or narrows the local parabolic shape of the original bifurcation diagram for Eq. (6.1) (see Fig. 6.1). From basic geometric considerations, Eq. (6.6) thus needs to be augmented with an additional term ρ, that modifies the focal width of its parabolic stable manifold. The focal width is the width of the parabola at focal level. For a parabola $(y-y_0) = a(x-x_0)^2$, the focal width FW is equal to a. Basic geometrical calculations allow to derive it from parabolic equations. Including the term ρ, Eq. (6.6) is modified into

$$\dot{x} = p - \rho x^2\,. \tag{6.9}$$

In this formulation, ρ corresponds exactly to the focal width of the normal form. As anticipated, Sec. 6.3.3 contains analytical derivations for the approximation of system (4.1) to the normal form (6.9), and its relationships with the geometrical results. Propagating ρ into Eq. (B.9) adds a tuning term to the bifurcation parameter, $k \to \sqrt{\rho}k$. Hence, the corresponding O-U process for a semi-quantitative saddle-node normal form is

$$\dot{y} = -\sqrt{\rho}ky + \eta(t)\,. \tag{6.10}$$

Spectral properties and statistical moments can be calculated immediately, provided that ρ is constant as assumed here. Among them, the primarily interest is for quasi-steady-state variance (Var) and lag-1 autocorrelation (AC(1)), measures of system variability close to criticality. Recall that they have been proposed as proxies for system resilience and as the most common EWS for impending bifurcation points (see chapter 5 and, *e.g.*, Scheffer et al. (2009); Trefois et al. (2015)). Following the present methodology, future studies might inquire other indicators and their behaviour under changing n and additional conditions. Analytical solutions for Var and AC(1), derived from the O-U process (6.10), are (Gardiner, 1985):

$$\mathrm{Var} = \frac{\sigma}{\sqrt{\rho}k}, \tag{6.11}$$

$$\mathrm{AC1} = e^{-\sqrt{\rho}k}\,. \tag{6.12}$$

Eqs. 6.11 and 6.12 are generic for noisy bifurcations whose resilience properties are altered by increasing focal widths, modelled as multiplicative parameters like ρ. They thus constitute a general result. Fig. 6.4 plots them as functions of ρ and k.

To connect with the original autoactivating feedback system, it is necessary to extrapolate the focal width of the bifurcation diagrams for each n. This is done by fitting a parabolic form $c = \alpha\tilde{x}^2 + \beta\tilde{x} + \gamma$ to the data points of each bifurcation diagram in the vicinity of the saddle point. Using MATLAB Curve Fitting toolbox also provides uncertainties over $\vec{\theta} = [\alpha, \beta, \gamma]$, resulting

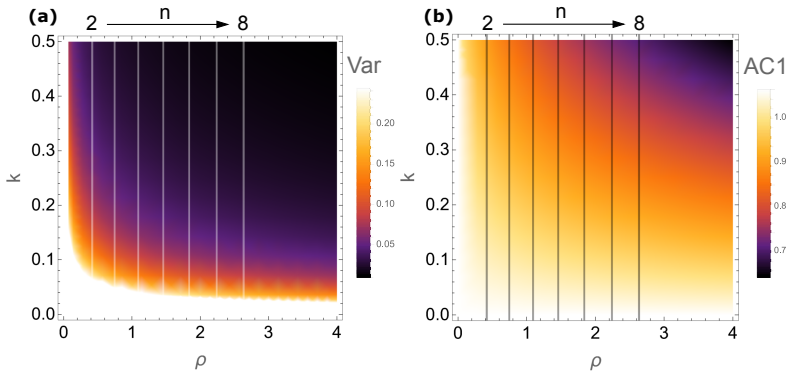


Fig. 6.4 **Left:** Theoretical dependence of Var on ρ (related to focal width) and k (distance measure from critical parameter values). Solid vertical lines represent slices for fixed values of $\hat{\rho}$ (Eq. 6.15) corresponding to the mean FW shown in Fig. 6.5. The associated n values increase from left to right. **Right:** same, but for AC1.

from small deviations from a perfect parabolic shape. By definition, the fitted focal width is

$$FW = 2|\tilde{x}(c_F) - \tilde{x}_F|\,, \tag{6.13}$$

where $(\tilde{x}_F,\ c_F)$ are the coordinates of the parabolic focus, extrapolated from the fit. To get a reasonable estimate of the corresponding uncertainties, the associated standard deviation is derived from the fitted parameter uncertainties $\mathrm{std}(\bar{\theta}_i)$ using a first-order approximated propagation method (Taylor, 1997):

$$\mathrm{std}(FW) = \left[\sum_i \left(\frac{\partial FW}{\partial \theta_i}\mathrm{std}(\theta_i)\right)^2\right]^{\frac{1}{2}}. \tag{6.14}$$

The relationship between FW and n is plotted in Fig. 6.4c, with the corresponding $\mathrm{std}(FW)$ from Eq. 6.14. A best fit study is also performed. The pattern decreases quadratically, thereby marking a rapid decrease followed by almost plateauing. Hence, bounded and relatively small cooperativity indexes are in principle sufficient to effectively buffer variability close to criticality. This fact may have important practical implications for synthetic biology studies.

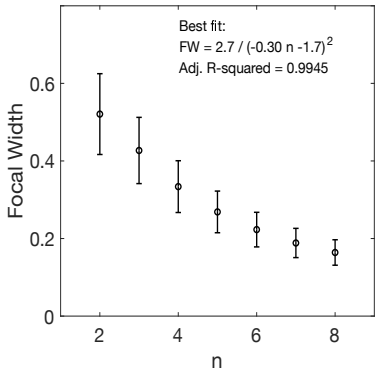


Fig. 6.5 Relationship between FW (Eq. (6.13)) and corresponding values of n, with best fit. Error bars correspond to one standard deviation from Eq. (6.14).

The estimated $\hat{\rho}$ values from fitted focal widths, for $n = 2$ to $n = 8$, are obtained as

$$\hat{\rho} = \xi (FW)^{-1}, \tag{6.15}$$

where ξ is a tuning parameter proportional to the Hill function (see Sec. 6.3.3). Mean $\hat{\rho}$ values are marked in Fig. 6.4a,b with solid vertical lines. Consistently with the trend observed in Fig. 6.4c, the mean values spread as n increases (from left to right). Low n values yield higher sensitivity to noise, as both Var and AC1 show substantially higher values for small cooperativity indices n, even when k is large (*i.e.*, further away from the critical point, but still within the bistable region, *cf.* Fig. 6.1). Thus, values of ρ can belong to two regions: one, where the values for both metrics are high for all k (left side of Fig. 6.4a,b), or another one where both metrics maintain low values for most k and increase rapidly close to criticality (right side of Fig. 6.4a,b). The region $\hat{\rho} \to \infty$ corresponds to the logic approximation (6.2) with $n \to \infty$, where Var and AC1 also change abruptly in a step-wise manner.

The ultra-sensitive region $\hat{\rho} \to 0$ is spanned by increasing dissociation constants (in the denominator of the Hill function, *cf.* Fig. 6.6), and potentially by changing other parameters, here not explicitly considered, or by different activation functions describing, for example, wild-type vs mutant organisms Hasty et al., 2000. Other pathways like growth feedbacks Deris et al., 2013 will likely correspond to additional regions in the parameter space. These investigations are left to future studies.

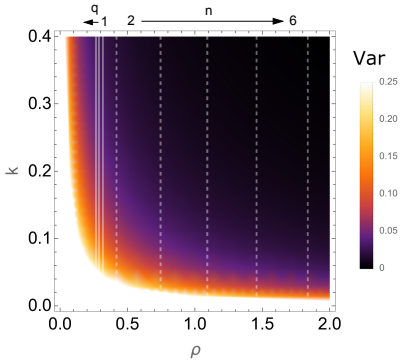


Fig. 6.6 The figure is analogous to Fig. 6.4, with $\hat{\rho}$ corresponding to increasing n represented as dashed lines (also, notice the change of scale). Solid lines, instead, correspond to different values of the dissociation constant q in the Hill function $\frac{x^n}{q+x^n}$. Values are increasing from right to left in $[2, 2.5, 3]$, towards the ultra-sensitive region $\rho \to 0$. For these, $n = 2$ and $K = 0.1$.

6.3.3 Analytical derivation of normal form results

The normal form associated to a dynamical system $\dot{x} = f(x,c)$ can be analytically derived from a Taylor expansion around the bifurcation point (Sec. A):

$$\dot{x} = f(x,c) = f(\tilde{x}_0, c_0) + \left.\frac{\partial f}{\partial x}\right|_{(\tilde{x}_0,c_0)} (x - \tilde{x}_0) + \left.\frac{\partial f}{\partial c}\right|_{(\tilde{x}_0,c_0)} (c - c_0) + \frac{1}{2} \left.\frac{\partial^2 f}{\partial x^2}\right|_{(\tilde{x}_0,c_0)} (x - \tilde{x}_0)^2 + \dots . \tag{6.16}$$

Here, $\tilde{x}_0$ identifies the critical value on the equilibrium manifold $x = \tilde{x}$, c_0 is the critical value for the control parameter, and $(\tilde{x}_0, c_0)$ is a saddle-node bifurcation point, *i.e.*, $f(\tilde{x}_0, c_0) = 0$ and

$\partial f/\partial x|_{(\tilde{x}_0,c_0)}=0$. Hence, the normal form Eq. 6.6 can be derived by truncating higher-order terms, yielding

$$\dot{x} \approx \left.\frac{\partial f}{\partial c}\right|_{(\tilde{x}_0,c_0)} (c-c_0) + \frac{1}{2}\left.\frac{\partial^2 f}{\partial x^2}\right|_{(\tilde{x}_0,c_0)} (x-\tilde{x}_0)^2. \tag{6.17}$$

Nonetheless, this procedure is not always possible. Besides the assumption that f is sufficiently differentiable, it requires prior knowledge of the system model $f(x,c)$, which may not be available in experimental setups. This justifies the introduction of the geometric methodology in Sec. 6.3.2, which can be directly applied to data.

Let us discuss the relationships of geometrical and analytical results, when an analytically tractable model is available. Let $f(x,c) = K + c\frac{x^n}{q+x^n} - x$ be the deterministic terms in Eq. 6.1, where c is the production rate (control parameter) and q is the dissociation constant. The Taylor expansion

$$\dot{x} \approx \underbrace{\frac{\tilde{x}_0^n}{q+\tilde{x}_0^n}(c-c_0)}_{p} + \underbrace{\left(-\frac{c_0 nq\tilde{x}_0^{n-2}}{2(q+\tilde{x}_0^n)^2} - \frac{c_0 n^2 q\tilde{x}_0^{n-2}(\tilde{x}_0^n-q)}{2(q+\tilde{x}_0^n)^3}\right)}_{-\rho}(x-\tilde{x}_0)^2 \tag{6.18}$$

is directly related to the saddle-node normal form Eq. 6.6, with a shifted equilibrium $\tilde{x}_0$. Note that p (and, hence, $k=2\sqrt{p}$) is proportional to the distance $(c-c_0)$ of the control parameter from its critical value, tuned by the Hill function $\xi = \frac{\tilde{x}_0^n}{q+\tilde{x}_0^n}$. Around the bifurcation point $(\tilde{x}_0, c_0)$, where $\dot{x}=0$, the parabolic form $c = \alpha\tilde{x}_0^2 + \beta\tilde{x}_0 + \gamma$ can be analytically determined from Eq. (6.18) as $\alpha = \rho/\xi$, $\beta = -2\rho\tilde{x}_0/\xi$, and $\gamma = \rho\tilde{x}_0^2/\xi + c_0$. Hence, the focal width FW is related to ρ according to Eq. 6.15.

The conclusions drawn from numerical fitting in Fig. 6.4, regarding the focal width convergence and ρ divergence, can also be verified analytically. The critical parameter c_0 can be derived from the tangency condition for a saddle-node bifurcation ($\partial f/\partial x|_{(\tilde{x},c_0)}=0$), yielding

$$c_0 = \frac{(x_0^n+q)^2}{nqx_0^{n-1}}. \tag{6.19}$$

Substituting ρ and c_0 into Eq. 6.15 of main text leads to

$$FW = \frac{\xi}{\rho} = \frac{2\tilde{x}_0^{n+1}}{(n+1)\tilde{x}_0^n + (1-n)q}, \tag{6.20}$$

which tends to zero in the limit of $n\to\infty$, if $\log(\tilde{x}_0) > 0$, as seen in Fig. 6.5 for $q=1$. Likewise, $\lim_{n\to\infty}\rho = \infty$. From $f(x,c)$, note that $\lim_{n\to\infty}\tilde{x} = K+c$, so the premise $\log(\tilde{x}_0) > 0$ holds if $K+c_0 > 1$, which is satisfied since $K=0.1$ and $c_0 = q = 1$ in the limit of the logic approximation Eq. 6.2. Values obtained from Eq. 6.20 can also be used to doublecheck the consistency of fitting results.

The focal width convergence to zero implies that the parabolic bifurcation diagrams represented in Fig. 6.1 collapses to a line for $n\to\infty$, losing the bistability property of the system as in the logic approximation Eq. 6.2. This same analysis can also be used to show that ρ decreases as the dissociation constant q increases for $q > 1$, as seen in Fig. 6.6.

6.4 Generality of EWS

After studying stability properties, the study focuses on performance of EWS against impending bifurcation points. The motivation is to consider complex systems lacking validated mechanistic models; in our case, this would translate to a scenario where n—or even the precise activation function—of an eukaryotic cell is poorly identifiable (Hasty et al., 2000). This consideration leads to questioning if we can identify statistical signals, computed on empirical data, that provide reliable information about the system's loss of resilience. Increasing trends of Var and AC1 have been widely suggested to work as EWS Scheffer et al., 2009; Trefois et al., 2015 but their robustness remains elusive. The present goal is to study how generic they are in the identified parameter range, despite the influence of an extra parameter ρ altering system resilience.

To account for mean trends and uncertainties, as well as to deeply study the interlinks of loss of resilience and noise, the system 4.1 is numerically integrated using the Euler-Maruyama method (see Sec. 4.2.2). To mimic cell populations slowly evolving close to equilibrium, we sample 10^4 time points over 200 repeated experiments in dependence of c. This leads to a distribution of statistical indicators, an approach discussed in 4.5. Fig. 6.8a, inset, provides an example.

To distinguish between bifurcation-induced transitions, anticipated by loss of resilience, and noise-induced transitions, measuring the scale between the distance to bifurcation point and the noise level is done by means of Kramers' escape rate (Sec. B):

$$\tau = \frac{e^{\frac{U(\tilde{x}_2)-U(\tilde{x}_1)}{\sigma}}}{2\pi\sqrt{|U''(\tilde{x}_1)U''(\tilde{x}_2)|}} \tag{6.21}$$

For any saddle-node bifurcation manifold 6.9 equipped with additive noise,

$$U(\tilde{x}_2) - U(\tilde{x}_1) = \frac{4}{3}\frac{k^3}{\sqrt{\rho}} \quad \text{and} \quad |U''(\tilde{x}_{1,2})| = 2\sqrt{p\rho}\,.$$

Hence,

$$\tau = \frac{\pi}{k\sqrt{\rho}} \exp\left[\frac{4}{3}\frac{k^3}{\sqrt{\rho}\sigma}\right]. \tag{6.22}$$

Technically, τ depends on ρ. However, the difference is very nuanced and practically negligible with respect to noise variance. To set comparable simulations, the generic (Kuehn et al., 2011)

$$\tau \simeq \mathscr{O}(\exp[k^3/\sigma]) \tag{6.23}$$

is thus considered.

Fig. 6.7 displays τ from Eq. 6.23 as a function of σ and k. Values lying at the exponential boundary of Fig. 6.7 (between orange and blue colour) provide comparable ranges of control parameters and noise levels for all simulations with different n. They distinguish two regimes, one where few noise-induced transitions might occur ($\tau \lesssim 2$, Fig. 6.8a,b) and another regime primarily determined by bifurcation-driven resilience loss ($\tau \gtrsim 2$, Fig. 6.8c,d). For the considered $\sigma = 0.02$, the system is very close to critical points.

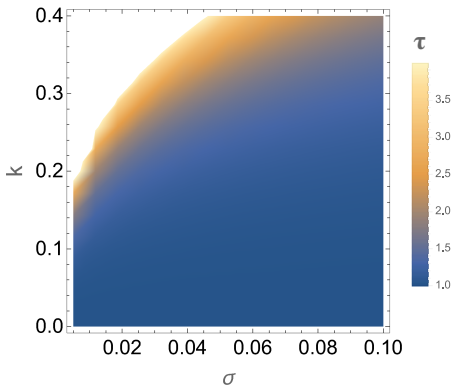


Fig. 6.7 Generic escape rate τ as a function of noise level σ and k, from Eq. (6.23).

Results from computer simulation experiments are shown in Fig. 6.8a–d (average values shown). Since the original system can be explored, $c - c_0$ is used instead of the more generic k. Like above, $K = 0.1$ without loss of generality.

When the dynamics is mostly characterised by the bifurcation (Fig. 6.8c,d), both Var and AC1 display patterns consistent with those predicted in Fig. 6.4. In addition, increasing n better buffers variability – absolute values for variance and autocorrelation are reached when n is low. When the noise level becomes comparable to the potentials' depth (Fig. 6.2c,d), Vars for different n become very close to one another due to the more prominent role of noise uncertainties. AC1s (Fig. 6.8a,b) remain separated due to lower sensitivity to noise (*cf.* Eq. 6.12), but with less marked – and, therefore, harder-to-detect – trends. That the variance is potentially more informative in these kinds of regimes, as its trends are more easily distinguishable, is similar to observations from real-world data, see Proverbio et al. (2022c) and Chapter 7.

6.4.1 EWS quantified

As explained in Sec. 4.5.1, EWS are associated to increasing trends of statistical indicators. For online applications (*i.e.,* as new data come in and without future knowledge of the system evolution), it is necessary to quantify whether an observed increasing trend is statistically significant, assessing whether it corresponds to a EWS or some spurious fluctuation (Boettiger et al., 2012a). Several strategies can be employed to do so (*cf.* 4.9). Since distributions at each c are available, this study looks for significant p-values between the computed distributions close to criticality and those far from the bifurcation (reference). An example is displayed in the inset of Fig. 6.8a.

Sec. 4.5.1 describes the details. In short, once fixed the reference distribution, all the others are compared with a t-test and the corresponding p-value registered. Notably, all p-values from different n are mutually consistent with each other: despite the different absolute values, the increasing trends are conserved. Fig. 6.9 show patterns of p-values from Var and AC1 (top and bottom, respectively, with the corresponding standard error), averaged over all n.

The p-values are computed closer to the bifurcation in Fig. 6.9a ($c - c_0$ corresponding to the parameters in Fig. 6.8a,b) and farther from bifurcation in Fig. 6.9b ($c - c_0$ corresponding

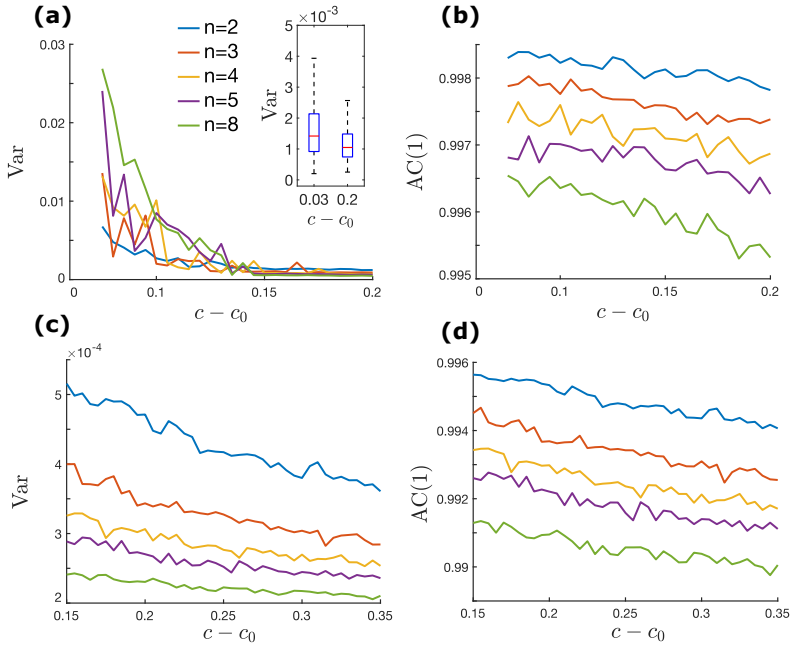


Fig. 6.8 Average trends of **(a)** Var and **(b)** AC1 as a function of c close to bifurcation c_0, in the regime where some noise-induced transitions might occur (blue area of Fig. 6.7). Simulations are presented over 200 realizations for different n. For each c, the indicators spread into distributions; example for two c values are in inset (a). **(c,d)** Average trends of Var and AC1 farther from bifurcation point c_0.

to the parameters in Fig. 6.8c,d). After an initial transient where the distributions are poorly distinguishable, the p-value decreases below significant levels. Whether 0.1 or 0.05 (two common thresholds, dashed lines in Fig. 6.9) are considered significant determines how much in advance the impending loss of resilience can be detected. For, $c - c_0$ even closer to 0, p-value < 0.01. Hence, this analysis certifies the potential use of proposed EWS to detect approaching bifurcation points in biological motifs.

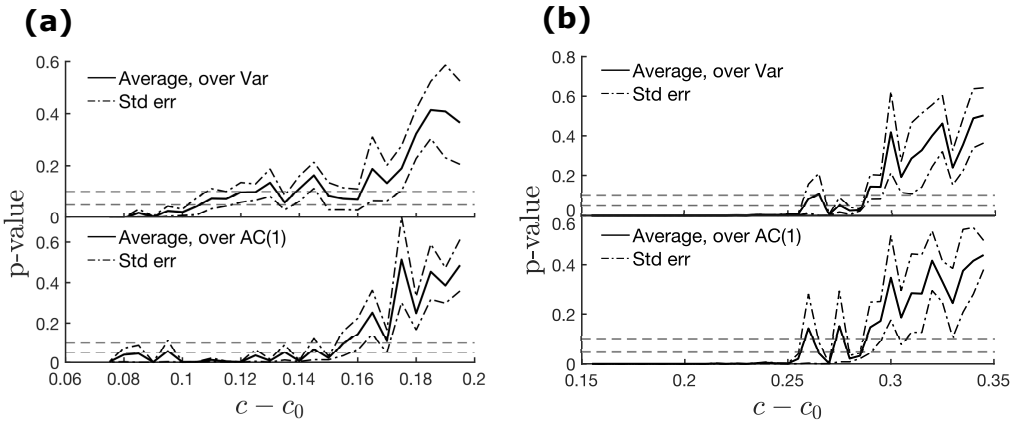


Fig. 6.9 Evolution of p-values between Var and AC1 distributions at each $c - c_0$ and the reference distribution. The latter corresponds to **(a)** $c - c_0 = 0.2$ (starting of "closer to bifurcation") and **(b)** $c - c_0 = 0.35$ ("farther from bifurcation"). Dashed lines represent typically used p-values in biological experiments.

6.4.2 Application to empirical data

The theoretical framework can be applied on experimental data to interpret and predict the relationship between resilience properties, bifurcation diagrams and EWS. As a case study, consider publicly available data[1] from Dai et al. (2015). They performed an experimental study, on yeast colonies subject to various degradation conditions towards collapse, to illustrate how such conditions yield different values for the measured coefficient of variation, used as EWS. The authors carefully reconstructed the bifurcation diagram for the yeast population, whether stably living or collapsing, as function of a dilution factor or of sucrose concentration. The population undergoes a fold bifurcation for critical values of each stressor (Dai et al., 2012; Dai et al., 2013).

Experimental data show that the two bifurcation diagrams for each stressor display different focal widths (higher for sucrose experiments), that the authors associate to resilience properties (see (Dai et al., 2015), Fig. 2A,B). Moreover, different increasing trends for the coefficient of variation, used as early warning signals of impending bifurcation points (see Dai et al. (2015), Fig. 2C,D,E), were observed. In particular, CV calculated on dilution factor experiments show more substantial increase. The authors hypothesised that such differing trends are associated to altered resilience properties, due to different environmental drivers.

Using publicly available data from the same study, the empirical bifurcation diagrams are reproduced in Fig. 6.10 (right) and the coefficients of variations are shown in Fig. 6.11 (right). In both cases, only average values are reported and rescaling is performed to allow better comparison between experimental data and theoretical predictions. It consists in extracting the distance from bifurcation point instead of using absolute values for the control drivers. To do so, the x-axis is rescaled to DL − Dilution factor, where DL stands for the fold bifurcation point, extrapolated from parabolic fit, and, respectively, to S − [Sucrose]. In addition, only points close to critical values are analysed.

The geometrical framework developed above is applied to the empirical data. First, a parabolic form is fitted to the data points and the corresponding focal width is extracted. That is used to extract values for empirical $\hat{\rho}$, using Eq. 6.15. In this case, the tuning parameter ξ is unknown. As the present goal is to compare two fitted parameters for the same population, $\xi = 1$ is assumed. In addition, uncertainties are presently not propagated, to allow focusing on the proof of concept results. Since $\hat{\rho}_{DL} = 2.1\text{e}{-7}$ and $\hat{\rho}_S = 9.8\text{e}{-11}$ differ by four orders of magnitude, discrepancies on ξ and propagated uncertainties are expected to play minor roles in the results.

Having extracted $\hat{\rho}$ geometrically, it can be inserted in Eq. 6.9 to derive a semi-quantitative saddle-node normal form, that explicitly takes into account system resilience. Fig. 6.10 (left) shows the corresponding bifurcation diagrams, which consistently reproduce the features of empirical ones (compare in Fig. 6.10). Estimated values of $\hat{\rho}$ place them in the sensitive subspace, leftward in Fig. 6.4.

It is now possible to predict the behaviour of early warning signals as the system approaches its critical values. To compare with results from Dai et al. (2015), the coefficient of variation is

[1] Dai, L., Korolev, K.S., Gore, J. (2015), Data from: Relation between stability and resilience determines the performance of early warning signals under different environmental drivers, Dryad, Dataset, https://doi.org/10.5061/dryad.k30v3.

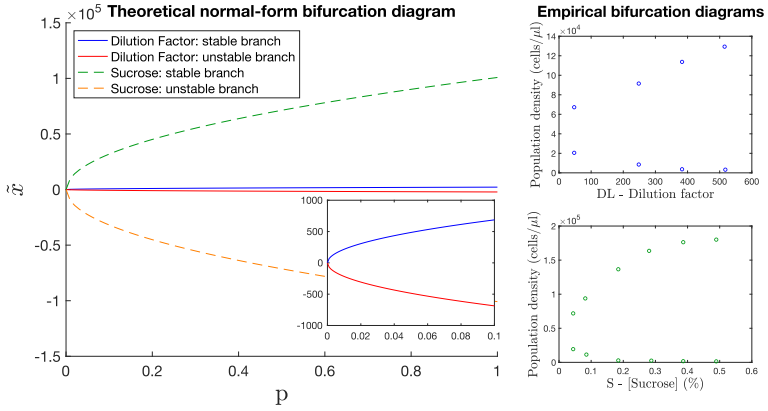


Fig. 6.10 Bifurcation diagrams. **Left:** for theoretically reconstructed normal forms, following Eq. 6.9. Inset: zoom-in for the bifurcation diagram related to dilution factor experiments. **Right:** for experimental data, from (Dai et al., 2015). The diagram plots stable (upper branch) and unstable (lower branch) population density values as functions of two environmental drivers leading too population collapse. Scaling of the x-axis, to better compare with theoretical diagrams, are performed as explained in the main text.

calculated as

$$\mathrm{CV} = \frac{\sqrt{\mathrm{Var}}}{\mathrm{Mean}} = \frac{\tilde{\sigma}\rho}{2p^{3/4}}\,. \tag{6.24}$$

Here, we can directly use p from Eq. 6.9 instead of k from Eq. 6.7 to make clear that another system is considered. $\tilde{\sigma}$ is a term for noise intensity, which is unknown. However, it corresponds to a scaling factor for the absolute values of CV, not altering its trends nor the associated EWS. Hence, w.l.o.g, $\tilde{\sigma} = 0.1$.

After fixing $\rho = \hat{\rho}_{DL}$ or $\rho = \hat{\rho}_S$, it is possible to predict the trends of CV as a function of p, the distance to critical point. The results are shown in Fig. 6.11 (left) and can be compared with empirical CV extracted from data, Fig. 6.11 right. Apart from absolute values (linked to $\tilde{\sigma}$ and propagated uncertainties), they lead to the same conclusions: both environmental drivers elicit increasing trends in CV as the bifurcation is approached, but dilution stress entails more substantial and earlier increase. Even though EWS can be observed in both cases, as the system is drifting to a critical transition characterised by critical slowing down, the response of yeast populations to different environmental stress yields more or less substantial trends and lead times.

The analytical framework developed in this chapter was therefore verified using empirical data. It was also useful to interpret experimental data as influenced by additional parameters linked to focal widths, which influence resilience properties and early warning signals performance. Hence, the proposed framework allows to link observations and complex systems to quantitative models, to foster new results in this direction.

6.5 Introduction to multiplicative noise

Gene regulation is known to be characterised by state-dependent noise (Bruggeman et al., 2018), depending on the value of the dynamical variables, which enters the describing equations in a multiplicative way. Most often, combinations of additive (white) and multiplicative (state-dependent) noise are present, describing extrinsic cell-cell variation or intrinsic fluctuations in

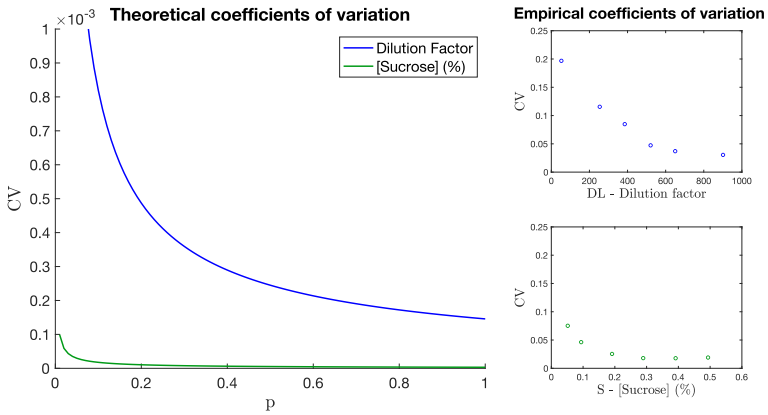


Fig. 6.11 Coefficients of variation. **Left:** theoretical CV, obtained from Eq. 6.24 for te corresponding $\hat{\rho}_{DL}$ and $\hat{\rho}_S$. **Right:** Average values for the empirical CV, from (Dai et al., 2015). Like in Fig. 6.10, the x-axis was rescaled.

regulation, respectively (Pan et al., 2010). To go beyond the pure white noise approximation, consider the case of pure multiplicative noise. Depending on the relative strength of each type of noise, the two cases considered in this work thus correspond to two extreme cases. Intermediate values are possible and are discussed in Sharma et al., 2016b for the case $n = 2$.

To model multiplicative noise, Eq. 6.1 from the main text such that is modified such that:

$$\eta(t) \to \eta(x,t) = c\frac{x^n}{1+x^n}\eta'(t)\,, \tag{6.25}$$

where $\eta'(t)$ is still an integral of a Wiener process and $\eta(x,t)$ depends on the state variable as described in Hasty et al., 2000. The resulting system is then simulated using a Milstein method (Sec. 4.2.2) and analysed using the same protocol described in the main text. The results are shown in Fig. 6.12. The AC(1) behaves similarly to what observed for the case of white noise (compare with Fig. 6.8b), while the variance displays different trends. Both behaviours are expected from Eq. 6.11 and 6.12: Var depends on the Ornstein-Uhlenbeck variance σ, which is altered by the multiplicative noise form, while AC(1) does not. This supports the use of reduced normal forms to study fluctuations near critical points. Let us focus on Var: apart from point-wise deviations associated to random switches, it maintains slowly increasing linear trends rather than significantly diverging as in Fig. 6.8a. This suggests that multiplicative noise can be another mechanism to buffer fluctuation amplitudes but not correlation. Future works may further study this interesting evidence. Overall, Var is anyway suppressed to smaller absolute values by higher n, in particular $n > 5$.

These findings are reflected into the EWS capabilities of statistical indicators, as quantified by the p-value analysis. In fact, AC(1) trends are significantly increasing like in the white noise case, thereby providing warning signals that a critical regime is approaching. On the contrary, trends associated with Var are less consistent, as reflected by oscillating p-values around significant levels. In this case, signals may be spuriously interpreted, thus calling for caution when using Var as an EWS indicator in systems strongly affected by state-dependent noise (as anticipated in chapter 5).

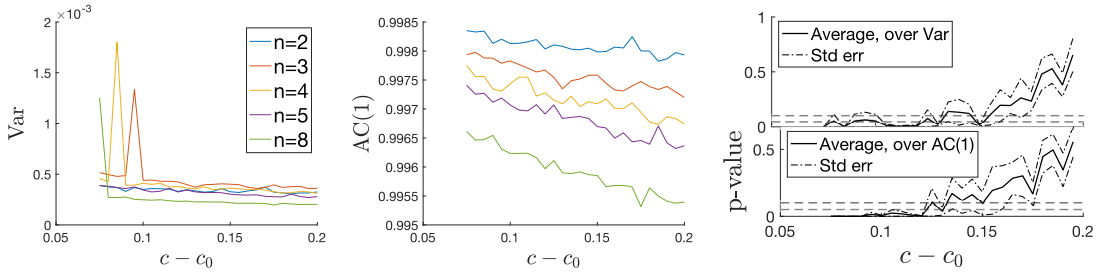


Fig. 6.12 Results for the case of state-dependent noise (from left to right: evolution of Var, AC(1) and p-values). Compare the figure with Fig. 6.8a,b and Fig. 6.9a.

6.6 Discussion

Understanding how cells regulate variability and cope with fluctuations is an active area of research (Snijder et al., 2011). This study focused on how dynamic patterns buffer systems' variability in critical regimes. Primary attention has been dedicated to the role of the Hill coefficient in shaping system resilience via alterations of the driving fold bifurcation.

Stability properties were first quantified by means of master equation formalism, associated to cooperativity values for the positive feedback loop motif. In addition, parameter ranges, where both variance and autocorrelation display low relative sensitivity to noise, were determined. In other ranges, however, the system poorly buffers its variability. Investigating whether they could correspond to other dynamical mechanisms is demanded for future studies, which may unravel alternative ways by which cells regulate their states or support the hypothesis of a self-organised fine-tuning in "safe" parameter spaces.

Sate-dependent noise has also been found to buffer variability amplitude near criticality. Addressing the interplay of noise properties and other regulation mechanisms captured by dynamical parameters will unravel control strategies employed by biological systems to improve their fitness.

In addition to assessing buffering mechanisms, a second result links them to EWS from the theory of critical transitions. This analysis shifted the focus from solutions of known mechanistic models to inference of system resilience from data. To this end, the proposed framework can be easily applied on different dynamical models and experimental setups. The methodology was based on the centre manifold theorem A.5.1 for bifurcation normal forms and allows to extrapolate resilience measures from simulated and empirical data. The methodology and the corresponding results are local, focusing on the vicinity of critical points rather than being globally applicable in the full dynamical range of cell functions. Future studies are necessary to test it on other systems, both theoretical and empirical. Incorporating multiplicative noise and other dynamical regimes is also left for future studies.

Among EWS proposed in literature, this study first concentrated on increasing variance and autocorrelation, since several indicators have been subsequently developed upon, to detect cell-fate decisions (Mojtahedi et al., 2016) and to possibly anticipate undesired shifts to e.g. cancerous states (Yang et al., 2018; Aihara et al., 2022). This study supplemented analytical results with numerical simulations of stochastic feedback loop motifs, to better investigate the interplay be-

tween cooperativity and noise and their effect in the identification of EWS. Hence it assessed their sensitivity to regulation mechanism, to interpret and apply warning indicators correctly. In the considered parameters' range, EWS are sufficiently generic to detect resilience loss. However, their use should be treated carefully, if other quantitative correlations between noise (including multiplicative noise) and bifurcations exist, that could shadow theoretical trends. Following our methodology, future studies might inquire other indicators and their behaviour under changing n and additional conditions.

This chapters developed an analysis framework to extract stability and resilience properties of complex systems. The framework was tested on two case studies: an abstract cell regulatory motif and experimental data about yeast populations. Theoretical predictions were verified in both cases and enabled interpretations of observations still not covered by quantitative modelling, thus closing the loop between experimental observations, hypotheses and model-based deduction.

Code availability

The code to reproduce the results is available at https://github.com/daniele-proverbio/Buffer-bio-variability

Chapter 7

Testing EWS on pandemic data

Adapted from[1] **:** D.Proverbio, F. Kemp, S. Magni, and J. Goncalves. *Performance of early warning signals for disease emergence: a case study on COVID-19 data*. PLOS Computational Biology, 18(3): e1009958, 2022 (Proverbio et al., 2022c).

Authors Contribution

DP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Roles/Writing - original draft, Writing - review & editing. FK: Formal analysis, Investigation, Software, Roles/Writing - original draft, Writing - review & editing. SM: Investigation, Supervision, Writing - review & editing. JG: Funding acquisition, Project administration, Supervision, Writing - review & editing.

7.1 Motivation

As overviewed previously, recent research have recommended using the theory of dynamical systems to derive generic early warning signals (EWS) to tackle critical transitions. In particular, they have been put forward to expand the set of alerting indicators against epidemic onsets. Despite substantial theoretical research, their empirical performance has yet to be thoroughly appraised. This chapter contributes to such testing over empirical data, specifically by analysing succeeding COVID-19 waves in a number of nations. It shows that, EWS can be beneficial in detecting new outbreaks when some basic modelling assumptions are met; however, they may be unable to detect quick or noisy modifications in epidemic dynamics. As a result, thanks to country-specific dynamical characteristics and careful data collection methods, it is possible to evaluate potential and limitations of such indicators.

[1] **Preliminary analysis also adapted from:** D. Proverbio, F. Kemp, S. Magni, A. Husch, A. Aalto, L. Mombaerts, A. Skupin,J. Gonçalves, J. Ameijeiras-Alonso, and C. Ley. *Dynamical SPQEIR model assesses the effectiveness of non-pharmaceutical interventions against COVID-19 epidemic outbreaks*. PloS one, 16(5), p. e0252019, 2021. (Proverbio et al., 2021)
and from: F. Kemp, D. Proverbio, A. Aalto, L. Mombaerts, A.F. d'Herouel, A. Husch, C. Ley, J. Goncalves, A. Skupin, and S. Magni. *Modelling COVID-19 dynamics and potential for herd immunity by vaccination in Austria, Luxembourg and Sweden*. Journal of Theoretical Biology, 530, p. 110874, 2021. (Kemp et al., 2021)
and from: M. Burzyński, J. Machado, A. Aalto, M. Beine, J. Goncalves, T. Haas, F. Kemp, S. Magni,L. Mombaerts, P. Picard, D. Proverbio, A. Skupin, and F. Docquier. *COVID- 19 crisis management in Luxembourg: Insights from an epidemionomic approach*. Economics & Human Biology, 43, p. 101051, 2021. (Burzyński et al., 2021)

7.2 Introduction

Epidemics pose relevant and enduring threats to human societies (Hsiang et al., 2020). The recent COVID-19 pandemic, triggered by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Pedersen et al., 2020, is a vivid example. It obliged governments to impose lockdowns and other non-pharmaceutical interventions, with various degrees of effectiveness (Proverbio et al., 2021). It disrupted social and economic life in unprecedented manners, and triggered changes in working habits and lifestyle (see Fig. 7.1), often difficult to cope with (Burzyński et al., 2021). It also spurred rapid re-arrangements of value chains and industrial production, suddenly oriented to produce and distribute massive doses of vaccines. Mathematical modelling and quantitative analysis of pandemic patterns become key tools to track its diffusion and inform data-supported decision making[2]. Similarly, developing reliable models focusing on various stages of disease progression and vaccination strategies allowed to quantitatively estimate plausible scenarios, linking conceptual epidemiological studies and contingent needs (Kemp et al., 2021).

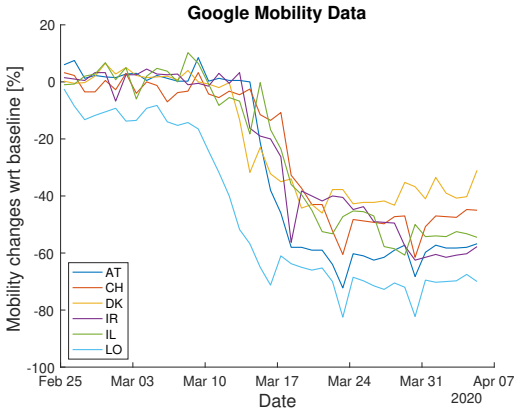


Fig. 7.1 Mobility trends according to Google Mobility Report, for several countries (COVID-19 Community Mobility Reports (https://www.google.com/covid19/mobility/), Accessed 14 Jun 2020). Dates correspond to first lockdowns. Each line corresponds to mobility change averaged across Grocery & Pharmacy, Retail & Recreation, Workplaces and Transit stations. Generalised reductions of mobility can be observed just after the first non-pharmaceutical measures were issued. Abbreviations: AT = Austria, CH = Switzerland, DK = Denmark, IL = Israel, IR = Ireland, LO = Lombardy.

In similar spirit, researchers suggested to aim for rapid and early detection of disease emergence by developing quantitative tools, towards science-based risk assessment (Horstmeyer et al., 2018). Ideally, detailed mechanistic models could forecast epidemic diffusion. However, their development is often hampered by combinations of noise, non-linearity and lack of curated data sets. It comes with little surprise that, despite their potential impact, only a few agent-based models were successfully developed and updated during the pandemic period, e.g. (Thompson et al., 2021; Kerr et al., 2021; Cuevas, 2020; Wilmes et al., 2020). Consequently, monitoring is usually based on theoretically-informed indicators like the reproduction number R, the average number of secondary infections from a single contagious case in a susceptible population Legrand

[2]Refer e.g. to weekly reports to the local government (https://www.researchluxembourg.org/wp-content/uploads/2022/04/Update$_{C}$ovid$_{A}$ssessment$_{2}$02200331.pdf) or to international reports, widespread during various pandemic phases, like N. M Ferguson et al. *Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand*. Imperial College, London. DOI: https://doi. org/10.25561/77482, 2020.

et al., 2007. As the pandemic progresses and is impacted by viral variants, non-pharmaceutical interventions, vaccination campaigns, social behaviours and so on, R might vary in time. It can thus be converted in the effective time-dependent reproduction number $R(t)$ Althaus, 2014. It can be estimated, with various methods and levels of precision, from epidemiological data, and lumps together various mechanisms that drive the epidemic progression. However, $R(t)$ is sensitive to data collection protocols and uncertainties, in addition to be constructed upon models based on several assumptions (Proverbio et al., 2021).

In recent years, studies have proposed to expand the toolbox of alerting indicators with different methods, that could detect shifts in epidemic dynamics without relying on detailed mechanistic models (Brett et al., 2017; Brett et al., 2020a; Miller et al., 2017; Phillips et al., 2020; O'Regan et al., 2013). These methods build on the critical transitions theory and look for noisy precursors of regime shifts. In fact, within epidemiological models, the reproduction number $R(t)$ acts as a control parameter driving the epidemic equilibria through bifurcations, *cf.* Sec. 4.1.2. Specifically, re-emergence of infectious diseases is associated with a transcritical bifurcation at $R(t) = 1$ (O'Regan et al., 2013). In principle, it is reasonable to expect that EWS could detect impending epidemic re-emergence.

Several studies have investigated EWS performance on theoretical or computational epidemiological models (Brett et al., 2020a; Southall et al., 2020; Brett et al., 2018; O'Dea et al., 2019). However, There has been so far limited testing on empirical data. First observations were reported by Harris et al. (2020), while Brett et al. (2020b) applied data-driven approaches. Liu et al. (2021) tested a different approach on COVID-19 data, derived from bifurcation theory on networks. Southall et al. (2021) recently published a review. Despite the growing corpus of studies, practical application of such early warning signals is still debated: it may be system-specific, or may depend on the interplay of empirical observed dynamics and modelling predictions. For routine surveillance procedures, it is required to perform more tests and characterise potential confounders, which could affect the expected signals.

Here, the aim is to perform controlled tests, to assess the detection performance of EWS against re-emergence of observed epidemics, and to interpret it according to dynamic features observed in the data and modelling assumptions. The idea is not to screen all possible EWS on all empirical datasets, like in Dablander et al. (2021); O'Brien et al. (2021). On the contrary, it is to test whether some EWS work when they are expected to, what happens in other cases, and why. Thereby, this work is similar to the strategy of "natural experiments" (Diamond et al., 2010): construct a data set including relevant time series data; elucidate possible confounders, i.e. deviations from assumptions or dynamical features that might alter the expected signals; evaluate the performance of EWS and interpret it according to theoretical results.

To this end, the data set is based on curated worldwide COVID-19 data, extracted from so-called "second waves" during summer 2020. Such second waves were associated to $R(t)$ re-crossing 1 from below Xu et al., 2020; Cacciapaglia et al., 2020. Here, the first outbreaks will not be considered: they were characterised by an exponential diffusion following a sudden emergence, not by a bifurcation crossing. Subsequent waves will not be considered either, because countries began issuing combinations of rapid and rarely synchronised interventions, which added complexity and potential confounders. Analysis of time series data allow to carefully consider dynamical features associated to modelling assumptions, like rate of evolution of $R(t)$ and noise

properties Dakos et al., 2012a. These features allow to interpret the empirically derived EWS and their performance in different contexts.

To allow comparison with expected EWS behaviour, this chapter first recalls theoretical results from literature. Then, it describes construction and analysis of the test set. The study of empirical EWS behaviour and performance will follow, together with their interpretations in light of dynamical characteristics associated with modelling assumptions. The discussion puts the results into context, and highlights their relevance or theory and applications.

7.3 Theory and Methods

7.3.1 Mathematical theory and EWS

When consistent with mean-field approximations, the dynamics of COVID-19 infectiousness is well described by SIR-like models (Anderson et al., 1979; Proverbio et al., 2021). Refer also to Sec.4.1.2 and Eq. 4.3. Among the various extensions of SIR models, we follow the one described by O'Regan et al. (2013). It explicitly considers a time-varying suppression parameter that, if lifted, could drive the leading parameter $R(t)$ past its threshold value 1, across a transcritical bifurcation. To avoid notation confusion with Rs and other symbols, convert $S \to X$, $I \to Y$ and $\to Z$ in the typical model.

So, we have a homogeneous populations of susceptible individuals (X); they can become infectious (Y) and are eventually removed by death or recovery (Z). Transition rates are, as in Eq. 4.3, the infection rate β and the removal rate γ, which lumps recovery and death rate. There is abundant literature to trace empirical values for COVID-19, e.g. (Wu et al., 2020; Proverbio et al., 2021).To model open systems, consider influx rate μ' and outflux μ'', of susceptible people moving inside or outside the considered regions. We can assume that such fluxes are small and balanced ($\mu' = \mu'' = \mu$), since the year 2020 was characterised by many travelling restrictions preventing large commuting[3]. On top of those, consider an additional influx rate η: this corresponds to random entry of new cases, which can trigger subsequent disease outbreaks. Finally, consider a probability p that some susceptible individuals are isolated and protected due to intervention measures. This can happen through "physical" interventions (e.g. non-pharmaceutical interventions limiting social interaction or changes in people's behaviour) or by vaccination (Peng et al., 2020). The extended SIR model is:

$$\begin{aligned} \dot{X} &= \mu(1-p) - \beta XY - (\eta+\mu)X \\ \dot{Y} &= \beta XY + \eta X - (\gamma+\mu)Y \\ \dot{Z} &= \mu p + \gamma Y - \mu Z\,. \end{aligned} \tag{7.1}$$

The total population size N is assumed to be constant, that is $X+Y+Z=1$ since variables are normalized. The control parameter R associated to Eq. 7.1 is given by:

$$R = \frac{\beta}{\gamma+\mu}(1-p) \tag{7.2}$$

[3]For a dataset of government measures, refer to the ACAPS website (https://bit.ly/3nFFqUS).

and gets to its critical value 1 when $p^* = 1-(\gamma+\mu)/\beta$, where a transcritical bifurcation occurs on the (Y,p) bifurcation diagram (O'Regan et al., 2013). The choice of p being the driver for $R(t)$ is specific to this example case, which does not undermine the generality because, eventually, it is $R(t)$ to determine the bifurcation. Other specific parameters governing $R(t)$ can act upon other compartments and correspond to other interventions, see Proverbio et al. (2021). The basic reproduction number for COVID-19, at the beginning of the pandemic and without any intervention, was estimated within $2 < R < 4$ (*cf., e.g.,* Liu et al. (2020a)).

Assuming that $p(t)$ changes slowly over time, the SIR model 7.1 approaching the transition can be expressed as a slow-fast system:

$$\begin{cases} \text{Eq. 7.1} \\ \dot{p} = \varepsilon f(X,Y,p)\,. \end{cases} \tag{7.3}$$

Here, $0 < \varepsilon \ll 1$ and f is a ramping function. $\varepsilon \to 0$ is a limit case of quasi-steady state change. Under this condition, the dynamical shift can be interpreted as a slow crossing through a bifurcation point. From this, the summary statistics can be computed as introduced in Chapter. 4. Often (O'Regan et al., 2013; Ashwin et al., 2012), the function f is assumed to be a constant $f = \tilde{p}$. Hence, p is considered a linear function of time:

$$p = p_0 + \tilde{p}t\,. \tag{7.4}$$

Following Eq. 7.2, $R \to R(t)$ is linear as well. If $\tilde{p} > 0$, then R gets reduced (towards elimination if R was above 1). If $\tilde{p} < 0$, R increases (towards a new emergence, if it was below 1). COVID-19 data only allow to investigate the second case, since most countries implemented suppression measures very rapidly (Flaxman et al., 2020), rendering Eq. 7.3 not satisfied in the first case.

Stochastic fluctuations are not negligible when the transition is approached from below, especially if daily cases are few. So, the system can be inscribed in the framework of noisy bifurcations (Sec. 2.4) and early warning precursors can be in principle computed. In the presence of noise, the transitions described above should be modelled by a stochastic master equation. O'Regan et al. (2013) and Alonso et al. (2007) already reduced the stochastic master equation to Eq. 7.3 and a Fokker-Plank equation for the fluctuations. Recalling the procedure is useful to highlight the assumptions underlying the derivation of early warning signals prior to the transition.

The system 7.1 can be reduced to its first two equations using the condition $\dot{N} = 0$. Hence, it suffices to consider the transitions in and out X and Y. Then, the quasi-steady state p is constant for each small time step dt.

Random jump processes describe the transitions in and out epidemiological states, modelled as (X,Y), $(X-1,Y)$, $(X,Y+1)$ and so on. The "stating" state is described by $\alpha = (X,Y)$, any other state by $\bar{\alpha}$. $T_i(\alpha,\bar{\alpha})$ is the probability of transitioning between states, depending on transition rates. $P(X,Y,t) = \text{Prob}(X(t),Y(t) = (x,y))$ is the probability that the state vector is equal to some pair of non negative integer numbers (x,y). Jumps in and out states are denoted by the index i. Depending on the considered system, O'Regan et al. (2013); Southall et al. (2020) provide examples of $T_i(\alpha,\bar{\alpha})$. The master equation for the stochastic process is thus:

$$\frac{dP(\alpha,t)}{dt} = \sum_{\alpha\neq\bar{\alpha}} T_i(\alpha|\bar{\alpha})P(\bar{\alpha},t) - \sum_{\alpha\neq\bar{\alpha}} T_i(\bar{\alpha}|\alpha)P(\alpha,t) \tag{7.5}$$

In general, it is a nonlinear equation. The van Kampen expansion (Van Kampen, 1992) can be used to approximate discrete random variables with continuous random variables. This allows to derive analytical solutions about average behaviours and associated fluctuations. Large N is required for the expansion: considering the population of medium to big countries (not villages, for instance) suffices. To leading order, the expansion of Eq. 7.5 is equivalent to Eq. 7.3. To quantify the fluctuations at next-to-leading order, use the corresponding Fokker-Plank equation (Sec. B). For Eq. 7.5, it is equivalent to the system of SDEs (Alonso et al., 2007):

$$\begin{aligned} \frac{d\sigma}{dt} &= b_{11}\sigma(t) + b_{12}\zeta(t) + \Gamma_1(t) \\ \frac{d\zeta}{dt} &= b_{21}\sigma(t) + b_{22}\zeta(t) + \Gamma_2(t)\ , \end{aligned} \tag{7.6}$$

where Γ_j are white noise processes. The elements of the $B = \{b_{kl}\}$ matrix are functions of transition rates. Eq. 7.6 thus connects the epidemic stochastic description with SIR-like models like Eq. 7.1, coupled with a noise term.

Eq. 7.6 can be analysed with its Fourier transform. Variance, autocorrelation and other statistical moments can be obtained by integrating the power spectrum, calculated on the fluctuations around the infectious state. Specific values depend on the eigenvalues of matrix B and of the covariance matrices of Γ_j. Fig. 7.2 shows the trends of variance and autocorrelation next to the transition, calculated similarly to O'Regan et al. (2013), for a fast and slow approach to $R(t) = 1$.

If the noise is relatively small with respect to the deterministic trend and normally distributed, the trend of these summary statistical indicators on the fluctuations is expected to increase next to the transition. Such increases are often quantified by Kendall's τ (*cf.* Sec. 4.5). If this analysis outputs reliable signals before the transition, they would constitute an early warning.

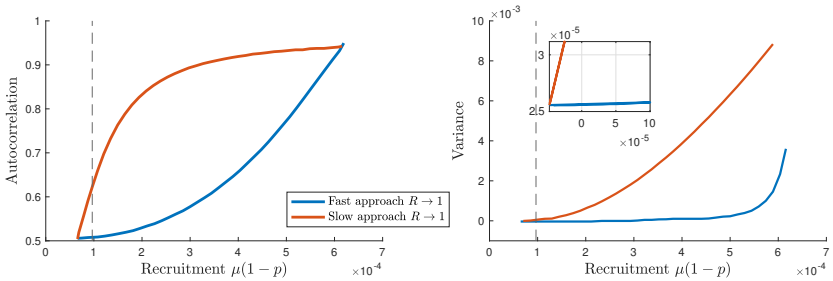


Fig. 7.2 Theoretical EWS for epidemic re-emergence. **Left**: AC(1) increases before the transition if the approach of $R(t)$ to 1 is slow (slow-fast assumption satisfied, red); otherwise it increases after the transition (bifurcation delay, blue). **Right**: variance increases before the transition (dashed grey line) if the approach of $R(t)$ to 1 is slow (slow-fast assumption satisfied, red; the inset magnifies the effect). Otherwise, it increases after the transition (bifurcation delay, blue). Both plots are derived from the analytical results of O'Regan et al., 2013 and reproduce its Fig. 9, d—e. On the x-axis: values for the recruitment rate [day^{-1}] of new infectious; they could trigger a re-emergence.

Finally, recall that the results on topological equivalence and normal forms (Sec. A) also apply in this case. Close to a transcritical bifurcation, the system can be reduced to

$$\dot{\theta} = q\theta - \theta^2 \tag{7.7}$$

where q is the bifurcation parameter and θ the variable of interest (in this specific case, Y from Eq. 7.1). Here, when q reaches its critical value, the extinction state and positive steady state coalesce and exchange stability. In the presence of noise ξ, the same linearization procedure around the equilibrium $\tilde{\theta}$, described in Sec. B, can be employed to obtain a Langevin equation (Berglund et al., 2006; Kuehn, 2011):

$$d\xi = -\frac{\partial f}{\partial \theta}|_{\tilde{\theta}}\xi + \sqrt{\sigma^2 g^2(\tilde{\theta})}dW \tag{7.8}$$

where dW is a Wiener process, σ^2 is the noise level and g is the diffusion coefficient from the associated Fokker-Planck equation. Eq. B.8 is an Ornstein–Uhlenbeck process(Allen, 2010) when $|\partial f/\partial\theta|_{\tilde{\theta}}| = k$ is linear; it has known statistical moments (Sec. B). Hence, the theoretical results described above can be complemented with those from the theory of noisy precursors of normal forms (Berglund et al., 2006; Kuehn, 2011; O'Regan et al., 2018 and chapter 5). Important remarks and assumptions are: 1) Critical transitions are local phenomena. EWS are thus not global measures, but are expected to work in the vicinity of the regime shift. 2) It is required that the epidemic dynamics is expressed in terms of a fast-slow system like Eq. A.3. When approximating $R(t) \to 1$ as a linear pattern, the modelling assumption Eq. A.3 is satisfied if the regression coefficient (the slope of the linear trend) is small. Otherwise, the expected trends will be either distorted or will not occur (Brett et al., 2017; O'Regan et al., 2013). 3) Multiplicative noise can modify the indicators trend. For instance, decreasing variance was observed in case of non-white multiplicative noise (O'Regan et al., 2018). So, the closer random fluctuations are to be additive noise, the more robust the performance of EWS is expected to be. 4) In case of non-equilibrium combinations of non-white noise and of non-fast-slow description, bifurcation delays may occur. They are changes of the system state (and of its indicators) that lag behind the theoretical bifurcation, yielding warning signals to emerge much later than the epidemic re-emergence. 5) If the transition is triggered by large random fluctuations – n-tipping, see Sec. 2.3.2 – no EWS is expected to be observed (Ditlevsen et al., 2010).

On top of theoretical studies, the performance of proposed CT-based indicators and related EWS has been investigated in computational studies. Increases in variance were found to be the best indicators of re-emergence, in terms of signal-to-noise ratio and of detection performance (Brett et al., 2017; Brett et al., 2020a). However, it is still not entirely clear when modelling assumptions are relaxed. This aspect is going to be investigated in the following sections.

7.3.2 Data collection and curation

Recall that the present aim is to verify whether EWS work when they are expected to, and explain why they might malfunction otherwise. Consequently, to construct the data set, only countries that faced a re-emergence of COVID-19 cases between beginning of March (starting of wide viral diffusion) and mid-September 2020 will be considered. Later data points are not collected: many

countries began issuing new and punctual social measures that rapidly impacted the epidemic trends, which could hinder the careful analysis of confounders.

When possible, prevalence data are employed, i.e. active cases A over the whole population of a regional area. This complies with SIR-like models' variables and with what was suggested in literature (Southall et al., 2020). For Luxembourg, active cases are directly retrieved from the government website (COVID19.public.lu/fr/graph). They are sampled over the whole population, with large scale testing (Wilmes et al., 2021) and careful control of the hospital system. For other countries, prevalence data are often not available. Hence, they are estimated by the proxy (Giordano et al., 2020):

$$A = C - D - \tilde{R}. \tag{7.9}$$

C indicates the cumulative positive cases, D is the number of registered deaths and $\tilde{R}$ counts the number of recovered patients. Data for each country are obtained from public repositories of confirmed detected, deceased and recovered cases: the John Hopkins University collection Dong et al., 2020 and the European Centre for Disease Prevention and Control database (https://www.ecdc.europa.eu/en/COVID-19/data). Italian data from the Veneto region are also employed, as an example of regional data with an identifiable second wave during the considered time interval. They are retrieved from the Github repository of the Italian "Dipartimento della Protezione Civile - Emergenza Coronavirus" (https://github.com/pcm-dpc/COVID-19). All databases were accessed up to 15/09/2020.

Data quality, particularly about $\tilde{R}$ cases, is relevant to obtain a robust estimator for A. To enhance data quality, several selection criteria are designed (see below). Consequently, the original publication (Proverbio et al., 2022c) focuses on prevalence data to compare the results with interpretations from various literature sources. Similarly, the investigation on the use of incidence data will be briefly reported at the end of this chapter. It uses registered daily new cases from the same sources listed above. Beware that incidence data might as well be influenced by testing bias, like lower testing over weekends and other factors. The second analysis complements that on prevalence data, by investigating EWS performance on real-world monitoring protocols.

As anticipated, selection criteria ae initially performed to best curate the database. To begin with, time series with very few active cases are rejected: there, intrinsic transmission stochasticity and measurement noise likely dominate the deterministic behaviour captured by SIR-like models. Another quality check is on the share of positive cases over performed tests: if it is $> 5\%$ around the transition, WHO guidelines suggest possible undertesting (see WHO reports such as https://bit.ly/3dARcy1). Corresponding information is obtained from the OurWorldInData curated dashboard Max Roser Hannah Ritchie et al., 2020 and is reported as a summary in Tab. 7.1. Since EWS from critical transitions are based on mean-field homogeneous SIR-like models, whole countries with clear spatial heterogeneity like Italy (Rivieccio et al., 2020) are also discarded. Regional data can be used instead, if available. Finally, it was necessary to discard some public time series that behave clearly differently from the typical epidemiological curves described, e.g. in Giordano et al. (2020); Reno et al. (2020); Reiner et al. (2020); Kemp et al. (2021). The latter case is mostly due to data management and sharing, or to different reporting frequencies for recovered cases. This would result in "sawtooth" curves for active cases, where the detrended

fluctuations are associated to reporting standards rather than being representative of intrinsic dynamics. Some examples are displayed in Fig. 7.3.

The list of countries eventually selected, together with relevant characteristics, is in Tab. 7.1. Corresponding curves of active cases, their smoothing and the associated $R(t)$ are displayed in Fig. 7.4.

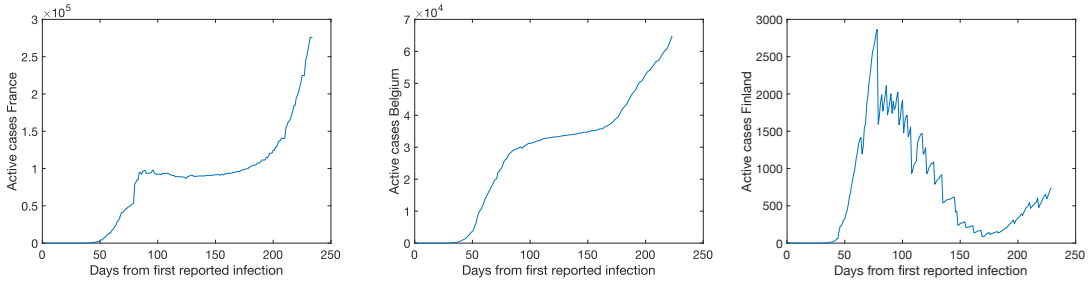


Fig. 7.3 Examples of discarded time series, following some of the criteria explained above. Among others, France **(left)** and Belgium **(centre)** had epidemic curves that differ from the typical "bell-shaped" SEIR-like behaviour. This is related to reporting protocols for recovered and dead patients. Finland **(right)** is an example of "sawtooth" evolution, due to recovered cases being reported with different frequencies than daily cases. Data are from Dong et al. (2020).

Country	Population	Area [km^2]	Test ratio	Share tests +	Date re-emerg.
State of Victoria (AUS)	6.681	227.444	1.8	<0.1%	27/06/2020
Austria (AUT)	8.917	83,879	0.6	[0.1; 0.6]%	01/07/2020
Denmark (DNK)	5.831	42,933	2.0	[0.1; 1]%	03/08/2020
Israel (ISR)	9.217	22,145	1.9	[0.3; 2]%	01/06/2020
Japan (JPN)	125.8	377,975	0.1	[1; 3.9]%	28/06/2020
Korea, South (KOR)	51.78	100,210	0.15	[0.3; 0.7]%	13/08/2020
Luxembourg (LUX)	0.632	2,586	10	[0.1; 0.4]%	29/06/2020
Nepal (NPL)	29.14	147,516	0.2	[0.3; 4.7]%	29/07/2020
Singapore (SGP)	5.686	728.6	5.1	[1.7; 3.7]%	25/07/2020
Veneto (VEN)	4.906	18,345	0.7	[0.4; 2.7]%	29/07/2020

Table 7.1 Information about the selected countries: population (in millions inhabitants), area (in km^2), average number of daily tests per 1000 inhabitants (test ratio), performed during the considered period (March - August 2020), share of positive (+) tests (in the same considered period, in percentage range from min to max values) and re-emergence date. The last two indicators derive from Max Roser Hannah Ritchie et al. (2020),where the full time evolution is reported. Following the large scale testing strategy, Luxembourg stands out for its higher number of tests performed per inhabitant. Sec. 7.3.3 describes how the most likely date of COVID-19 re-emergence is estimated.

7.3.3 Analysis of dynamical features

Before analysing EWS next to a transition, it is necessary to identify such transition *a posteriori* and get a "ground truth" date of re-emergence. To this end, one option is using a data-driven

estimation of the time-dependent $R(t)$ (recall that $R(t) = 1$ marks the transition point[4]. Here, it is done with Bayesian inference by means of a Markov Chain Monte Carlo (MCMC) method, similarly to Brett et al. (2020a); Systrom et al. (2020); Abbott et al. (2020). This method estimates the probability of observing a certain value of $R(t)$, by calculating the likelihood of seeing k new cases given the candidate $R(t)$, following a Poisson transmission process. When modelling "arrivals" of discrete-state stochastic processes, Poisson transmissions are widely employed, e.g. for Ebola (Ajelli et al., 2016) and Influenza (Liu et al., 2018) models. Pre-processing with a Gaussian 7-days window smoother is performed to avoid fitting spurious bumps. Note that, since this is only used for a retrospective analysis, it does not modify the non-anticipating scheme for the EWS.

A previous implementation from Systrom et al. (2020) was adapted for the estimates (also refer to it for further details). Given an average rate of λ new cases per day, the probability of seeing k new cases is distributed according to the Poisson distribution:

$$p(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}. \tag{7.10}$$

In turn, λ depends on R as (Bettencourt et al., 2008):

$$\lambda = k_{t-1} e^{\gamma(R-1)} \tag{7.11}$$

for all time points. γ is the reciprocal of the serial interval, which is close to 4 days for COVID-19, original variant (Du et al., 2020; Park et al., 2020). Since the estimate comes with uncertainties, γ is treated as a random sample from a Gaussian distribution, centred at 4 days with an assumed standard deviation of 0.2. Consequently, the probability of observing a time series $x = \{x_t\}$ for each t between t_0 (first report date) and any T, discretised by small steps δ, is given by:

$$p(x|R) = \prod_{t=t_0}^{T-\delta} p(k_{t+\delta}|\lambda_t) \tag{7.12}$$

Following Bayes' rule, the posterior distribution of $R(t)$, for each time point, is given by (up to a normalization constant):

$$p(R|x) \propto p(x|R)q(R), \tag{7.13}$$

where $q(R)$ is a prior distribution. For each time point after $t_0 + 1$, the prior equals the preceding posterior. We follow the implementation of Systrom et al., 2020 to generate thousands of MCMC samples with the Metropolis Hastings algorithm, starting from a Gaussian prior $\mathcal{N}(R, \sigma)$, with $\sigma = 0.15$ (assumption).

The method outputs a posterior distribution $p(R|\text{data})$, with a most likely $R(t)$ and credible intervals. To derive a transition date with high confidence, $p(R|\text{data})$ is integrated to obtain the probability that $R(t)$ is indeed greater than 1:

$$\mathscr{P}(R(t) > 1) = \int_1^\infty p(R|\text{data})dR\,. \tag{7.14}$$

[4]With a slight abuse of notation, both the theoretical and data-based reproduction number are identified with the same symbol $R(t)$. This, to stress their close connection and to ease the notation, as only the data-based one will be employed in the analysis.

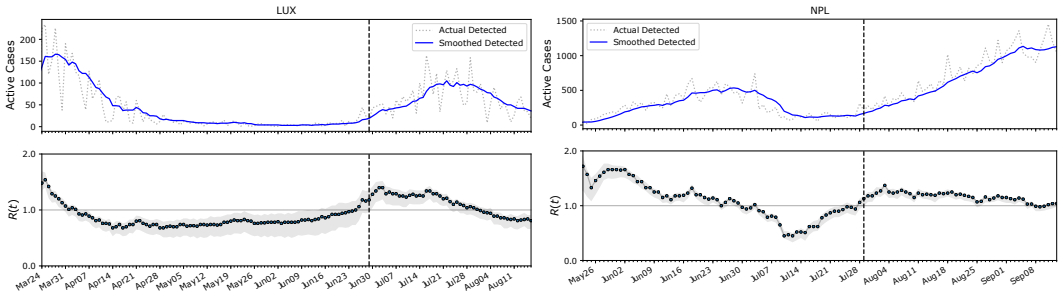


Fig. 7.4 Curves of active cases for two example countries with their associated $R(t)$ (median values and 50% credible intervals). The vertical dashed line identifies the day marked for the transition, see Tab. 7.1. Examples for Luxembourg **(left)** and Nepal **(right)** are shown here; all plots are available on Proverbio et al. (2022c), supplementary material.

Since $R(t) > 1$ is associated with an exponential increase of infectious cases after a transcritial bifurcation, $\mathscr{P}(R(t) > 1)$ can be interpreted as the probability of actually seeing an epidemic outbreak. The most likely day of transition occurrence, t_{em}, corresponds to the first day when $\mathscr{P}(R(t) > 1)$ from Eq. 7.14 reaches its maximum value of 1. t_{em} is assumed as the "ground truth". Fig. 7.4 shows the results of the Bayesian $R(t)$ estimation (median values and 50% credible intervals) for the considered countries.

After calculating the outbreak date, the modelling assumptions of normally distributed fluctuations and of slow approach to the critical transition are also tested.

The global distribution of stochastic fluctuations, filtered from the complete time series with a 7-days moving Gaussian kernel (Dessavre et al., 2019; Lenton et al., 2012), is analysed to check the additive noise assumption. The window size reflects typical cycles of data reporting and of COVID-19 fluctuations (Ricon-Becker et al., 2020). Out of the distribution, skewness and kurtosis ae computed to measure deviations from Gaussian noise, characterized by skewness=0 and kurtosis=3 (Taylor, 1997).

The rate of approach of the control parameter to its critical value is instead measured to test the assumption of slow approach to the transition. For this, compute the time-dependent $R(t)$ like above. Then, consistently with the fast-slow system description of Eq. A.3, fit a linear function:

$$R(t) = a + b \cdot t \tag{7.15}$$

in the interval $t \in [\tilde{t}, t_{em}]$. Here, t_{em} is the day associated with novel disease emergence as explained above; $\tilde{t}$ is the day associated with the minimum of $R(t)$ after the first wave. The control parameter ramping speed is captured by the regression coefficient $b \pm \sigma_b$, along with its uncertainty. As an indication, $R(t)$ is said to be "slowly evolving" if it goes from its minimum value to 1 in much more time than the COVID-19 serial interval (about 4 days (Nishiura et al., 2020)); that is a proxy of the disease time scale. The fitting is performed with the *scipy* Python library, considering as uncertainty the 50% credible interval from the distribution of $R(t)$. The goodness of fit is evaluated with the reduced χ^2 score: $\chi^2_{\text{red}} < 1$ guarantees the goodness of the linear fit.

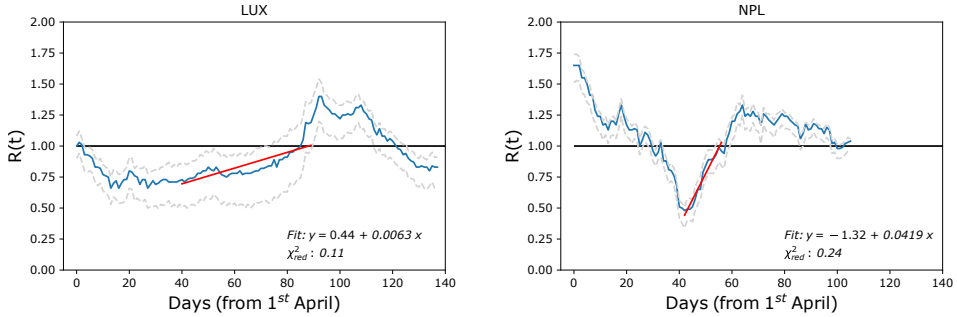


Fig. 7.5 Estimate of transition rate to R critical value. $R(t)$ evolution (blue line; dashed lines are $\pm 50\%$ CI) is fitted with a linear trend (Eq. 7.15) before the transition, to estimate the rate of approach to the threshold value 1. The fit starts around the minimum of $R(t)$ (excluding small fluctuations) and proceeds until the median value crosses 1 (horizontal line). The best fit curve is in red. $\chi^2_{\text{red}} < 1$ guarantees that a good fit is achieved. The regression coefficient b is a proxy for the rate of approach to 1. Examples for Luxembourg **(left)** and Nepal **(right)** are shown here; all plots are available on Proverbio et al. (2022c), supplementary material.

7.3.4 Estimation and quantification of EWS

Estimation of early warning signals from time series data is performed as explained in Sec. 4.5: detrend the time series, obtain the residuals and compute the statistical indicators associated to each point with a backward sliding window, i.e. one where the associated time point is the rightmost one. As anticipated in Sec. 4.4, there are several concurrent detrending methods. To corroborate the current analysis with an investigation of their effects, three methods are used and compared: a uniform moving mean, a Gaussian kernel, and ARIMA models. The latters are specifically fit for each country by tuning the (p,q,d) parameters appropriately. All detrending methods are non-anticipating. Hence all estimates are agnostic of future values and reflect real-world practices: calculations are performed as soon as a new data point becomes available.

Estimated indicators are variance, suggested to be the most robust indicator for epidemic re-emergence, lag-1 autocorrelation AC(1), coefficient of variation (CV) and skewness. Consistently with Sec. 4.5, the variance is:

$$\text{Var}_{i,t} = \frac{1}{M-1} \sum_{s=t_0}^{t} (A_{i,s} - \hat{A}_{i,s})^2 \tag{7.16}$$

for any time point i with active cases A, over a sliding window with size $t - t_0$ including M time points. $\hat{A}$ is the moving average. The others are computed similarly, over the same sliding window. Due to 1-day sampling frequency of COVID-19 data, power spectrum reddening (Biggs et al., 2009) and sample entropy (Brett et al., 2017) cannot be estimated. All indicators are computed with MATLAB functions. Note that the estimation of $\mathscr{P}(R(t) > 1)$ is done a posteriori, that is, once the complete time series is available. Instead, EWS indicators are calculated a priori, without knowing in principle if a transition is approaching.

The increasing trends of each indicator are quantified with the Kendall's τ coefficient of monotonicity (see Sec. 4.5). Constant trends are expected to have $\tau = 0$. This is a "null" value, compared with the τ scores from time series with identified transitions. After calculating each

indicator on a moving window (its size is discussed later in the text), the Kendall's τ score is calculated, for each timepoint, on windows of the same size, over an overall period $-30 < t_{em} < 5$ days around the transition – the positive data set. $t > -30$ is chosen to avoid significant overlaps with the first epidemic wave, $t < 5$ to account for possible small bifurcation delays (Kuehn, 2013). For the negative data set, τ values are taken way before the transition onset, i.e. $t < -30$.

The ROC analysis (Sec. 4.6) is used to classify each time point as either before or after disease re-emergence. To do this, determine whether the estimated τ is higher or lower than some threshold value at each considered timepoint and determine whether each time series is classified correctly by that threshold. This gives a proportion of true positives and false positives. Various values for $0 < \tau < 1$ are compared to those belonging to the positive and negative data set, for each country. Specificity and sensitivity results over all countries time series are grouped, to obtain the final ROC curve. The overall detection performance of each EWS is quantified by the area under the ROC curve (AUC). AUC$= 0.5$ means that the statistics detection performance is as good as a random classifier. A good indicator should have AUC close to 1, to identify the transition by its increasing trend. AUC close to 0 indicates good classification, although resulting from a decreasing indicator. This does not correspond to the predetermined theoretical prediction and can be regarded as a spurious signal.

7.4 Results

7.4.1 Analysis of country-wise dynamical characteristics

After constructing the dataset (Tab. 7.1), the dynamical characteristics are investigated. Fig. 7.6b shows various ranges of noise distribution, measured by skewness and kurtosis. Those of Austria, Luxembourg, Nepal, Singapore and Veneto are close to Gaussian. This is not true for the other countries, possibly due to social dynamics or imperfect data reporting (Bauchner et al., 2020).

The rate of approach of $R(t)$ to its critical value also differs, as indicated in Fig. 7.6c by the regression coefficients (recall Eq. 7.15). State of Victoria, Austria, Luxembourg, Singapore and Veneto display a slow approach to the critical value; they are better suited to be appropriately described as slow-fast systems like Eq. A.3. Japan and South Korea show intermediate values, while other countries - Denmark, Israel and Nepal - have a faster evolution of the control parameter, which does not satisfy the assumption of slow evolution.

This analysis allows to subdivide the considered countries into two test sets, based on whether the assumptions of slow approach and whit-ish noise distribution are satisfied. Group clustering is employed to make the subdivision. State of Victoria, Austria, Luxembourg, Singapore and Veneto are grouped together, under the assumption of slow rate. Except for Australia, their noise distribution is also close to Gaussian. They thus form the test set $\mathscr{Y}$, used to further assess the performance of EWS. On the other hand, Denmark, Israel and Nepal display higher rates of approach to $R(t) = 1$ and large deviations for Gaussian noise distribution and are grouped in a set $\mathscr{N}$. $\mathscr{N}$ is used to interpret the performance of EWS in settings that diverge from theoretical models and represent possible limitations of the predetermined predictions. South Korea and Japan are more ambiguous when clustering over the slope of $R(t)$; hence, they are split into $\mathscr{Y}$ and $\mathscr{N}$, respectively, based on their relative vicinity to Gaussian noise distribution.

Luxembourg is peculiar: it satisfies the modelling assumptions and is the closest to being a "controlled experiment". From literature and practical experience, we know that the country is small, homogeneous population-wide interventions were in place, and a Large Scale Testing (LST) strategy was implemented, reaching more than 70.000 tests per week over a population of about 600.000, thus allowing extensive and frequent monitoring. Hence, Luxembourg is employed as an initial sample to test the theoretical predictions about the local behaviour of EWS.

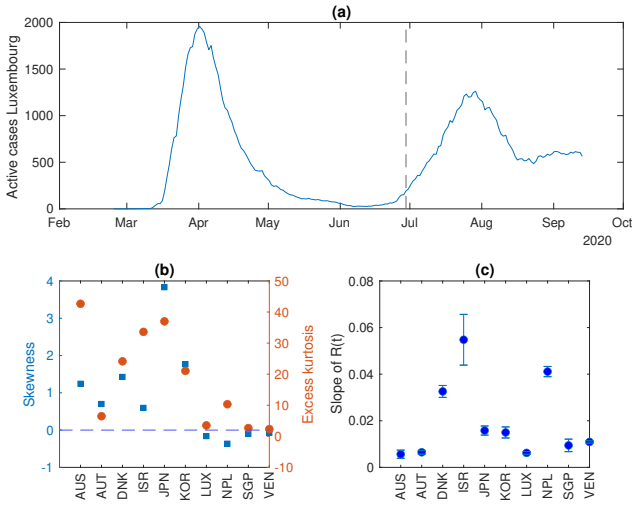


Fig. 7.6 Analysis of the dynamical characteristics of the countries included in the data set. **(a)** An example of an epidemiological curve of active cases from Luxembourg. The dashed line indicates the transition, measured by $R > 1$. **(b)** Skewness μ indicates the symmetry of the distribution, whereas kurtosis γ indicates the relevance of its peak with respect to the tails. Large deviations from $\mu = 0$ (dashed line) and $\gamma = 3$ are associated with non-normal distributions, so we display the excess of kurtosis $\gamma - 3$. **(c)** Regression coefficient of $R(t)$ and its associated uncertainty, obtained from the linear fit Eq. 7.15.

7.4.2 Local trends on controlled data and impact of detrending methods

Now, focus on Luxembourg, which displays the best data in terms of curation of prevalence data (see Methods) and of satisfaction of theoretical assumptions (see above). Theoretical predictions about the local behaviour of common EWS can be immediately tested.

The effect of detrending methods in generating residuals is first tested. To do so, compare the fluctuations around the deterministic trend obtained with a Gaussian kernel smoothing, a moving average filtering and an ARIMA(2,1,3) model. Fig. 7.7 (left) shows the time evolution of the residuals obtained with the three methods, as well as their mutual correlations. Gaussian and moving average filtering have similar output (Pearson correlation coefficient $\rho = 0.95$). This is likely related to the Gaussian bandwidth of 7 days, used to reflect known weekly fluctuations related to testing routines. Consequently, the Gaussian kernel smoothing is used in the rest of the analysis. However, the ARIMA method returns residuals that are less correlated with the previous ones ($\rho = 0.23$); its effect on EWS thus needs further investigation.

Then, study the behaviour of the variance (theoretically, the most robust EWS Brett et al., 2020a) next to the transition point. Fig. 7.7 (left) shows its increase prior to the transition, as expected from theoretical studies. The trend is independent of the moving window size and on the

detrending method. The corresponding Kendall's τ measure of monotonous increase is similar for both methods and all window sizes (*cf.* values reported in Fig. 7.7, right). The lead time is instead slightly advanced for shorter window sizes. Larger window sizes produce smaller fluctuations but a visually reduced absolute increase; in addition, they might capture old decreasing trends, that should not be analysed. On the contrary, smaller window sizes yield less smoothed curves but larger absolute variance values. Nonetheless, they might not include enough data points to capture the trends in more noisy estimators like AC(1) (Dessavre et al., 2019).

For the rest of the analysis, a window of 14 days will be used as a reasonable trade-off: it collects enough data to be robust without being over-dependent on past history. In addition, the ARIMA residuals produce a visually clearer increase in variance, but with comparable Kendall's τ (even slightly lower). So, to make for a complete analysis, both detrending methods will be considered to study EWS empirical performance. These findings confirm that, in a controlled setting that satisfy the modelling assumptions (Luxembourg), the transition to disease emergence can be anticipated by increasing variance, used as EWS.

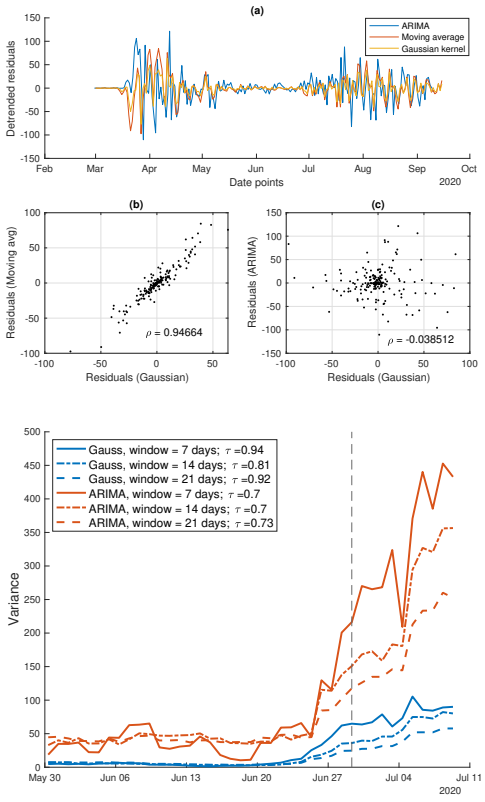


Fig. 7.7 **Left:** Analysis of the residuals from different detrending methods (case study from Luxembourg shown). a) The detrended fluctuations time series. b) Correlation between residuals obtained from Gaussian or moving average filtering. c) Correlation between residuals obtained from Gaussian or ARIMA filtering. **Right:** Analysis of the variance in the Luxembourg setting. Its increase is evident prior to the transition (dashed vertical line). The variance is computed over the residuals from Gaussian filtering and ARIMA detrending. The increasing trend during the considered time window is quantified by the associated τ values, which is little sensitive to the sliding window size.

7.4.3 Global trends of EWS

After confirming the local expected behaviour of the variance, using Luxembourg as a highly controlled setting, let us widen the analysis to the global performance of other EWS. To do this, set the "zoom" far from the bifurcation and look for different countries from the pre-defined dataset (Tab. 7.1). This way, theoretical predictions and EWS potential use in more general contexts can be investigated. Lag-1 autocorrelation (AC(1)), skewness and coefficient of variation (CV) are considered as common alternatives to variance. The moving window size is set to 14 days as discussed above. To compare the trend of EWS with the approach to the bifurcation, $\mathscr{P}(R(t) > 1)$ (Eq. 7.14) is also calculated and reported.

From the test set $\mathscr{Y}$, Fig. 7.8 shows the results for Luxembourg, Austria, State of Victoria. It also includes Israel, which does not satisfy the EWS assumptions (*cf.* Fig. 7.6), to inspect a deviant case. The figure considers EWS trends after the first wave, up to about a month after the second epidemic insurgence. The left column refers to indicators estimated after Gaussian filtering; the right one refers ARIMA detrending. All the other graphs, for other countries from Tab. 7.1 can be found in supplementary material of (Proverbio et al., 2022c).

In Luxembourg and Austria, the variance trend is in accordance with the theoretical increasing one: it has a small but visible increase prior to the transition and a subsequent monotonous trend along the second wave. In Austria it however displays some fluctuations after the relaxation of the first wave. The coefficient of variation CV is similar; in fact, it depends on the variance and on a stable equilibrium in infectious numbers. Instead, AC(1) shows an increasing trend very close to the transition point, but gives possibly spurious signals during the global time series. In turn, the skewness does not display immediately detectable relevant trends. This could have been expected from computational studies (Southall et al., 2020) and might be related to noise properties, as suggested by Guttal et al. (2008).

On Australian data, when processed by eye, Variance and CV start increasing close to the transition, and the trend becomes more pronounced around the 7^{th} of July. This might be a hallmark of bifurcation delays, often associated with deviations from Gaussian noise (Kuehn, 2013; O'Regan et al., 2018); it could also relate to delays in tests results reporting or symptoms onsets. Israel is an interesting case study. It diverges from the theoretical assumptions (Fig. 7.6): its transition to epidemic re-emergence is rapid, and the noise distribution is far from being Gaussian. As a consequence, theory predicts that EWS trends might be disrupted – which is what happens. In fact, the variance remains flat around the transition (it slightly decreases, too); CV and skewness slightly decrease, while AC(1) is not associated with informative patterns. A delay occurs more than 20 days after the transition, as abrupt as the exponential increase in infectious data . It is clear that applying early warning signals on appropriate contexts is crucial to obtain reliable information.

Expand to ARIMA-related indicators: similarly to what observed in Fig. 7.7, the variance trends are similar, and so are those of CV and AC(1). On the other hand, the skewness looks different. For instance in Austria, it increases when the detrending is performed with Gaussian kernel, but it decreases after ARIMA (compare the two panels of Fig. 7.8). It is known that the skewness is very sensitive to noise distribution (Guttal et al., 2008): small changes in the residuals, due to the different filtering procedure, might suffice to modify its trend.

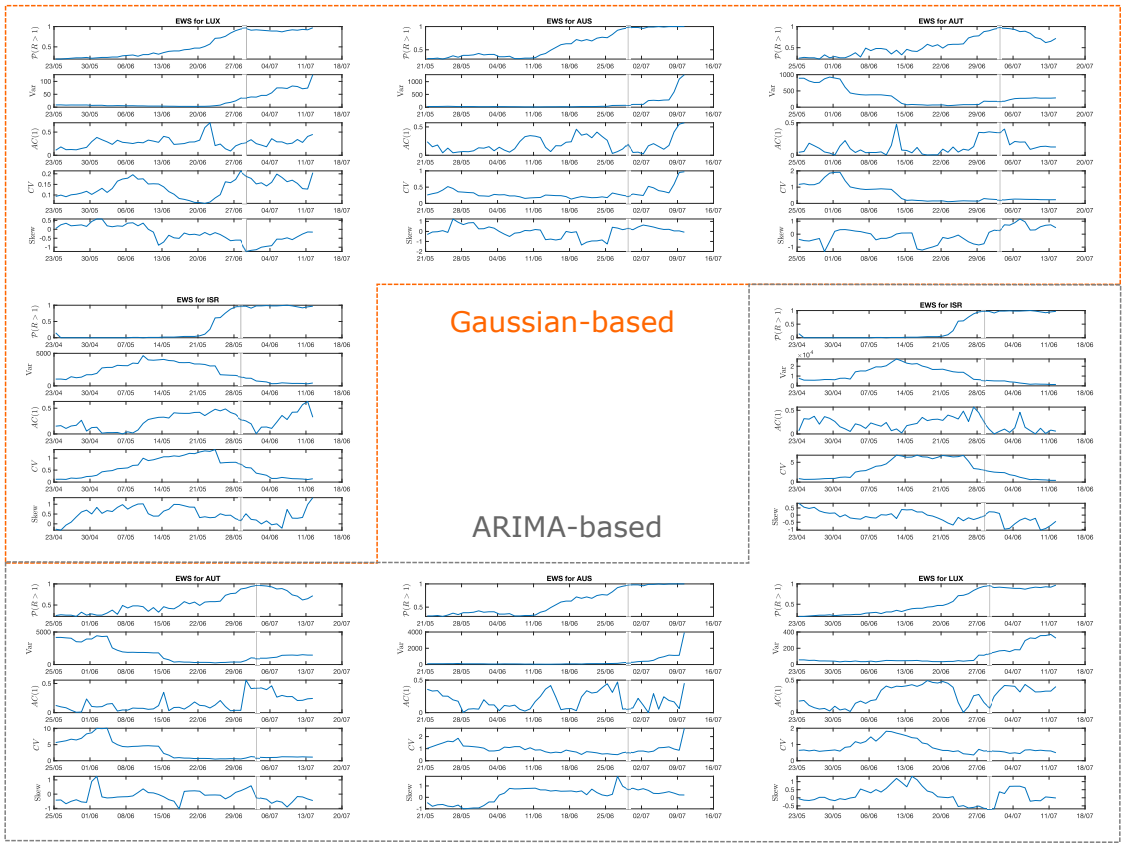


Fig. 7.8 Evolution of EWS far from the transition point. Four example countries are shown: Luxembourg and Austria, with controlled features; State of Victoria (Australia), with small deviations from controlled features; and Israel that does not satisfy theoretical conditions. Considered EWS are the most common ones (variance, lag-1 autocorrelation, coefficient of variation, skewness). In addition, to mark the approach to the transition, $\mathscr{P}(R(t) > 1)$ from the Bayesian estimation (see Eq. 7.14) is displayed. The vertical line reports the transition date. **Orange box:** Gaussian filtering. **Grey box:** ARIMA detrending.

7.4.4 ROC analysis of EWS performance

Consider the problem of online detection of incoming re-emergence. Distinguishing between robust increases and spurious fluctuations is crucial to optimise the true positive signals and minimise the false negatives. A retrospective analysis of time data is often not sufficient and is only useful for offline detection. This is why non-anticipative methods have been used so far, and why quantitative estimations of EWS performance in robustly detecting the transition is now performed.

Like before, the increasing trend is quantified with Kendall's τ score, for each indicator and over the same 14 days window. Then, assess which values are associated with a passage through the transition point and estimate the ROC curve and AUC values, as described in Sec. 7.3.4. Fig. 7.9 shows the ROC curves for the considered indicators, grouping all countries in $\mathscr{Y}$. Fig. 7.9a is for ROC curves derived from Gaussian filtering; Fig. 7.9b focuses on ARIMA detrended data. Tab. 7.2 reports the corresponding AUC values, for both methods.

Variance consistently performs better than a random classifier, while the lag-1 autocorrelation seldom performs slightly better than that. This agrees with aforementioned results from literature,

e.g. Southall et al. (2020); Brett et al. (2020a). The skewness does not improve detection performance, probably because of its fluctuations around the 0 value, in turn associated with noise distribution of original data. Interestingly, the coefficient of variation is overall the worst performer. This could depend on its sensitivity to data fluctuations: they are often non negligible even in countries belonging to the test set $\mathscr{Y}$ (*cf.* Fig. 7.6). Note that our findings are sensitive to the estimated time of re-emergence. Since it comes with uncertainties, it also complicates the estimation of the lead time, which is best for Luxembourg (5 days), a setting that is close to the analytical assumptions.

ARIMA detrending method yields overall better performances. AC(1) is an exception: both methods return similar values close to random classifiers. ARIMA estimates finer trends at different time scales, thus returning more accurate fluctuations: this might explain why ARIMA residuals yield higher AUC. On the contrary, Gaussian filtering might be slightly more rough: it considers average time scales and returns approximated estimates for the fluctuations. This argument might explain the improved performance of skewness: over fine-grained time scales, the ARIMA may pick the slight asymmetry in residual distributions, yielding skewness-related signals. In general, this analysis highlights the importance of choosing the right detrending method, to increase the detection performance of various indicators.

The same analysis is performed over the set $\mathscr{N}$. Their AUC values are reported in Tab. 7.2: the considered indicators are not able to detect the transitions, performing worse than random classifiers. This supports what was already noticed for Israel in Fig. 7.7: disrupted trends can occur, contradicting what is expected and thus returning false negative signals. For variance and CV, AUC values close to 0 indicate that the transition is well detected by decreasing trends. This contradicts the theoretical predictions. Such features are possibly linked to non-complete relaxation of the indicators after first waves, or to delays. Most likely, it is an instance of spurious signal, to be carefully interpreted. Hence, if a system is not known or there is difficulty in determining the type of data, incorrect conclusions could be drawn when interpreting the time series trend. Care should therefore be maintained.

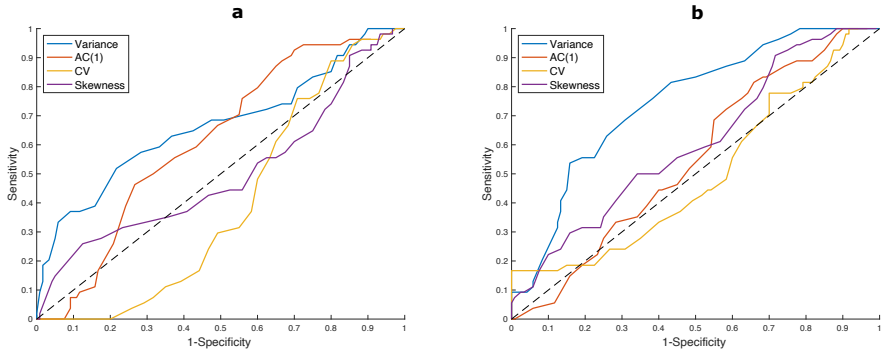


Fig. 7.9 ROC curves for each considered indicator, with sensitivity and specificity calculated on each timepoint for all countries in $\mathscr{Y}$. Each point corresponds to a test value for τ, to define if the detection is positive. Diagonal line corresponds to the ROC of a random classifier. Curves above it imply better performance. **(Left)** Computed on Gaussian filtered data; **(Right)** Computed on ARIMA detrended data.

	Gaussian det.		ARIMA det.	
Indicator	over $\mathscr{Y}$	over $\mathscr{N}$	over $\mathscr{Y}$	over $\mathscr{N}$
Variance	0.6671	0.1981	0.7123	0.0934
AC(1)	0.5258	0.4995	0.5182	0.2840
CV	0.3968	0.1043	0.3626	0.0368
Skewness	0.4664	0.2925	0.5482	0.4609

Table 7.2 AUC scores for different indicators, over $\mathscr{Y}$ and $\mathscr{N}$ datasets, after Gaussian or ARIMA detrending methods.

7.5 Incidence data

The analysis on prevalence data is complemented with that on incidence data (daily new cases). The latter avoids potential bias induced by the estimation of recovered patients, but might be more sensitive to systematic fluctuations associated with testing routines. The two analysis can thus be regarded as representative of real-world monitoring capacities.

In accordance with preliminary studies that investigated incidence data for re-emergence of infectious diseases (Southall et al., 2020; O'Dea et al., 2019), the results are similar and consistent with those for prevalence data. However, there is a number of differences worth stressing.

The distribution of fluctuations around the average trend is different than what observed in Fig. 7.6. In the case of incidence data, the deviation from Gaussian noise distribution is more pronounced (see Fig. 7.10b), reflecting the larger fluctuations associated with testing protocols, like reduced weekend testing (see Fig. 7.10a). Interpreting the subsequent results as consequences linked to the noise distribution is therefore more challenging. In addition, Japan and South Korea have different skewness and excess of kurtosis than in Fig. 7.6. Therefore, as the rate of approach to $R(t) = 1$ is conserved (Fig. 7.10c), the initial criterion of placing one country in the test set $\mathscr{Y}$ and one in $\mathscr{N}$ still holds. However, they are swapped: Japan is placed in $\mathscr{Y}$ and South Korea in $\mathscr{N}$. This is done for consistency with the previous analysis; it is anyway verified that the following results are little sensitive to this choice. Other countries are grouped identically.

As for the trends of statistical indicators, both next to the transition and far from it, the results are comparable with those from prevalence data. Gaussian-based and ARIMA-based results are even closer (Fig. 7.11). However, there are some interesting differences. To begin with, variance and CV on the $\mathscr{N}$ set are less associated with decreasing trends close to the transition, but are closer to random classifiers (see Tb. 7.3). This might suggest that, as discussed above, the result on prevalence data was likely spurious and related to relaxations from the first wave. Daily new cases are not delayed because of recovered cases; hence, such effect is less marked.

The second, more striking difference regards the skewness. As shown in Fig. 7.12, and quantified in Table 7.3, the skewness is particularly good in detecting the transition to disease emergence on the test set $\mathscr{Y}$. This differs from Southall et al. (2020)'s results but is more in line to what was suggested by Guttal et al. (2008). Since there are only few studies connecting observed EWS to noise distribution, making a conclusive interpretation is very challenging. Speculatively, the detection performance might be linked to skewness sensitivity to noise distributions (as already introduced in Guttal et al. (2008) themselves): potentially, the correct combination of

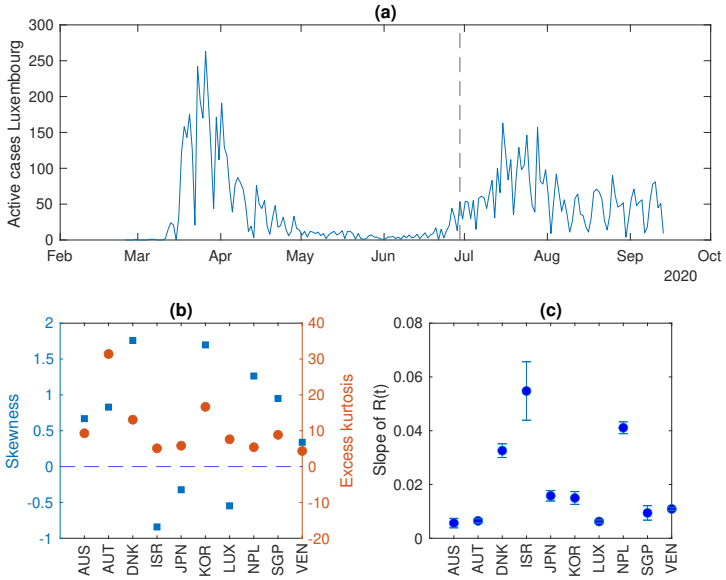


Fig. 7.10 Analysis of the dynamical characteristics of the countries included in the data set, for incidence data. **(a)** An example of an epidemiological curve of daily new cases from Luxembourg. The dashed line indicates the transition, measured by $R > 1$. **(b)** Measures of the distribution of data fluctuations. Skewness μ indicates the symmetry of the distribution, whereas excess of kurtosis $\gamma - 3$ indicates the relevance of its peak with respect to the tails. Large deviations from $\mu = 0$ (dashed line) and $\gamma = 3$ are associated with non-normal distributions. **(c)** The regression coefficient of $R(t)$ and its associated uncertainty.

fluctuations and approach to $R(t) = 1$ might have yielded the observed result. Further analytical and computational studies might better investigate this issue, which is solely reported here. In any case, the skewness is again no better than a random classifier on the $\mathscr{N}$ set.

	Gaussian det.		ARIMA det.	
Indicator	over $\mathscr{Y}$	over $\mathscr{N}$	over $\mathscr{Y}$	over $\mathscr{N}$
Variance	0.6718	0.4681	0.7334	0.2561
AC(1)	0.5234	0.3132	0.2624	0.6021
CV	0.3370	0.2019	0.4380	0.1813
Skewness	0.7826	0.5005	0.6805	0.4991

Table 7.3 AUC scores for different indicators, over $\mathscr{Y}$ and $\mathscr{N}$ datasets, after Gaussian or ARIMA detrending methods.

This complementing analysis on incidence data provides additional insights and questions, to be further compared with theoretical studies. It highlights the importance of choosing observable data, as anticipated in Sec. 3.2.4. It also stresses that, as the considered indicators rely on a number of assumptions, we are justified in using them when such assumption are satisfied, but this is not always the case in real-world settings. Otherwise, the EWS sensitivity to such assumptions might yield spurious signals and hinder our capability to extend them in uncertain contexts.

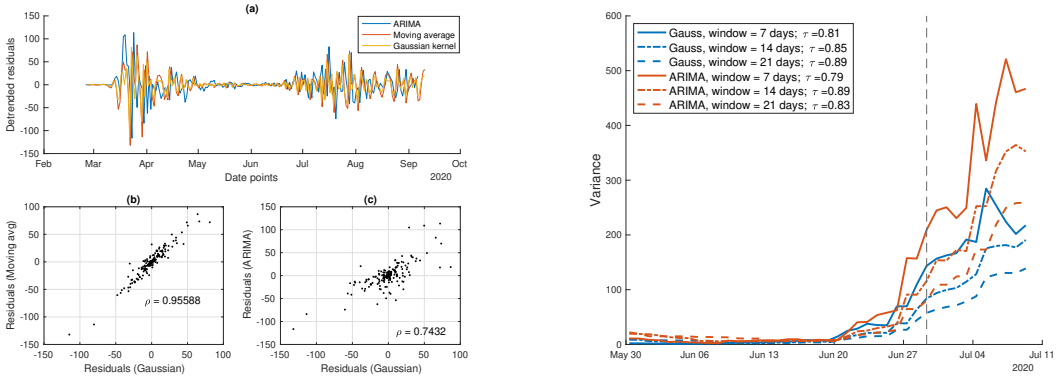


Fig. 7.11 **Left:** Analysis of the residuals from different detrending methods (case study from Luxembourg shown). **Right:** Analysis of the variance in the Luxembourg setting. This figure is computed on incidence data; compare with Fig. 7.7 and refer to its caption for explanation.

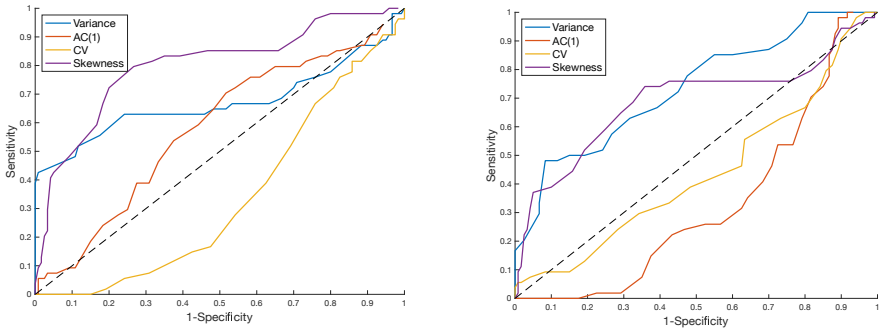


Fig. 7.12 ROC curves for each considered indicator, with sensitivity and specificity calculated on each timepoint for all countries in $\mathscr{Y}$. Each point corresponds to a test value for τ, to define if the detection if positive. The diagonal line corresponds to the ROC of a random classifier. Curves above it imply better performance. **Left**: Computed on Gaussian filtered data. **Right**: Computed on ARIMA detrended data.

7.6 Discussion

This study focused on verifying and interpreting empirical analysis according to theoretical assumptions. It is based on a curated dataset, to perform observational studies and check for possible confounders. Like $R(t)$ and other indicators (Adam, 2020), EWS are estimates that rest on assumptions; hence, the study screened the dataset to assess the matching of empirical features and theoretical assumptions. It also investigated the impact of different detrending methods on the final EWS performance, informing researchers about the importance of data processing methods. Its results suggest that some EWS from the CT framework are able to detect the transition to disease re-emergence when necessary theoretical assumptions like normal distribution of data fluctuations and slow change rates are satisfied. On the contrary, noise and commensurable time scales may obscure the early warning signals. These observations call for caution when interpreting monitoring outputs. Future studies might relate the dynamical features observed to social behaviours, political strategies or monitoring protocols. Such information would significantly impact the reliability of EWS for online epidemic surveillance.

This study comes with limitations that might be overcome when new and better curated data sets become available. First, data quality could be a limiting factor, despite being representative of real world monitoring capacities. In particular, the proxy of active cases A relies on reliable estimates of recovered and dead patients. Moreover, the official numbers of positives might neglect undetected asymptomatic cases. Both cases are investigated by considering incidence data. Such analysis supports earlier results, but highlights again the importance of quality data and recall that, from a practical point of view, looking at a measure or another might make a difference for the monitoring efficacy.

Secondly, the used estimators come with uncertainties. The empirical $R(t)$ is an estimate of the true reproduction number, based on the assumption of homogeneous dynamics. Moreover, the use of moving windows might produce odd behaviours of EWS (Dessavre et al., 2019) which can contribute to poor signals; window size should also be optimised for every epidemic and sampling protocols.

Third, the current definition of "ground truth" transition date is somewhat conservative: it is requested to get maximum probability of $R(t)$ to be greater than 1. In the real world, appropriate detection thresholds are conditional to the costs of late outbreak alerts. They require careful assessment by public health authorities, which could modify the estimated lead time - how much in advance a re-emergence can be predicted. Because of this, quantification of lead time is not performed here. Challenges related to lead time assessment are also discussed in Chapter 8. The interpretation of "bifurcation delays" is closely connected: depending on the definition of "ground truth", they might be less severe than what discussed before. COVID-19 latency periods and reporting delays could be alternative explanation for observed delays; this aspect can be elucidated when extra data are available, using nowcasting methods (Wu et al., 2020).

Real epidemics might behave differently that what is usually modelled; however, the present analysis supports the use of minimal dynamical models to predict relevant aspects of complex epidemics. While detailed and complete multivariate models are being developed, insights and indicators to detect epidemic re-emergence can be provided by macro-scale models based on complex systems theory. The current results corroborate theoretical literature findings and their basic assumptions and warn against naive applications of summary statistics as EWS: if not correctly applied, they could return possibly misleading spurious signals. These findings also call for developing forecasting techniques based on pattern recognition, in different dynamical regimes. For instance, Machine Learning-based algorithms (Brett et al., 2020b) could be built, augmented and interpreted using validated EWS. Epidemic monitoring and control will likely take advantage of such dual synergies.

Code availability

The code to reproduce the analysis is available at https://github.com/daniele-proverbio/EWS_epidemic.

Chapter 8

A model-driven alternative to pandemic monitoring

Adapted from: D.Proverbio, F.Kemp, S. Magni, L.Ogorzaly, H.M.Cauchie, J. Gonçalves, A.Skupin, and A. Aalto. *Model-based assessment of COVID-19 epidemic dynamics by wastewater analysis*. Science of the Total Environment, 827, p. 154235, 2021. (Proverbio et al., 2022b)

Authors Contribution

DP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Roles/Writing - original draft, Writing - review & editing. FK: Methodology, Writing - review & editing. SM: Methodology, Writing - review & editing. LOy: Data curation, Funding acquisition, Writing - review & editing. HMC: Data curation, Funding acquisition. JG: Conceptualization, Funding acquisition, Supervision, Writing - review & editing. AS: Conceptualization, Funding acquisition, Supervision, Writing - review & editing. AA: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Roles/Writing - original draft, Writing - review & editing.

8.1 Motivation

The previous chapter introduced the importance of real-time epidemic monitoring. It highlighted the potential and limitations of signals derived from the CT framework and discussed the challenges on quantifying lead time. One option to continuously early alerting systems is to consider additional indicators from complementary measures. A similar line of reasoning, that resonates with practical needs in real-world contexts, was for instance suggested by Clements et al. (2018) in ecology. To this end, data from viral RNA load in wastewater samples are used and coupled with mathematical tools, to reconstruct the epidemic diffusion within a population and to provide warnings that could anticipate epidemic curves from personal testing procedures. This chapter also illustrates the challenges to connect ideal mathematical methods with real-world data retrieval protocols and additional procedures.

8.2 Introduction

Accurately estimating epidemic dynamics is essential for targeted and effective interventions. In particular, attentive monitoring is required to detect – and possibly prevent – subsequent outbreaks, even in situations of reduced active RT-PCR or antigen testing. While this is true for any pandemic, it is of paramount importance for the current COVID-19 one, likely the most impactful pandemic of the current century as elucidated in Chapter 7.

A cost-effective alternative to population-based testing is the analysing SARS-CoV-2 abundance in wastewater (Farkas et al., 2020; Larsen et al., 2020): it studies viral load produced by infected patients, whether symptomatic or not, and is largely independent of personal decisions to get tested or seek for healthcare (Peccia et al., 2020). In principle, it does not depend on symptom onset delays or waiting time for testing results. During long-term epidemic surveillance, it is thus potentially faster and more reliable for alerting indications (Weidhaas et al., 2021; Wurtzer et al., 2020; Randazzo et al., 2020). At the time of writing, more than 60 countries and 270 universities have activated wastewater surveillance systems against COVID-19, for a total of more than 3,300 active registered sites[1] (Naughton et al., 2021) and many practical uses (Ahmed et al., 2021; Quilliam et al., 2020; Reeves et al., 2021). In Luxembourg, such analysis was performed within the Research Luxembourg COVID-19 taskforce, for our local government.

Much still needs to be done to sensitively improve the method. Experiments and data processing are constantly being refined, but modelling procedures should, too (Daughton, 2020; Bandala et al., 2021; Lahrich et al., 2021). In particular, it is necessary to go from qualitative understanding of epidemic trends to quantitative estimates of the shedding population size, considering data uncertainties and allow inference of future outcomes (Tiwari et al., 2021; Zhu et al., 2021). These methods would also allow to quantify the expected lead time, which is essential for alerting systems.

Some methods exist in the literature. Some build upon qualitative and semi-quantitative retrospective studies of lagged correlations (Nemudryi et al., 2020; Kumar et al., 2020). Others employ regression models (Cao et al., 2021; Vallejo et al., 2021; Li et al., 2021a; Hasan et al., 2021; Huisman et al., 2021). An alternative study (McMahan et al., 2021) uses an SEIR model, supported by data-driven estimates of viral evolution within individuals. All models are country-specific and highlight the challenges listed above, in particular regarding noisy data (potentially biased by detected case numbers) and varieties of individual infection periods. In addition, they do not reconstruct the full epidemic trajectory nor make extrapolations: hence, they are limited in studying the alerting time.

The work described here suggests an alternative, automated and causal-based approach, termed CoWWAn (COVID-19 Wastewater Analyser). It couples a SEIR model (see Sec. 4.1.2) with an extended Kalman filter (EKF, see Sec. 4.2.4). This way, the method produces simulation results, representative of real system observations, that account for uncertainties. Hence, it is possible, for each time step, to estimate and interpret a selected number of free parameters from the SEIR model. Population-wide epidemic diffusion can thus be quantitively and causally reconstructed, interpreted in light of the epidemiological model, and projected forward in time, overcoming the extrapolation limitations of correlation-based statistical approaches (Cao et al., 2021; Li et al., 2021a). The method is also adaptable to different sampling routines and regional areas. On a

[1] A curated dashboard can be accessed at: https://www.covid19wbec.org/covidpoops19

side note, the Kalman filter was applied on wastewater viral data by Cluzel et al. (2022) and textCourbariaux et al., 2022. However, a linear first-order autoregressive model was used there, making it a signal processing tool primarily. Although good to smooth noise from wastewater measurements, it lacks the other features mentioned above.

In particular, CoWWAn allows to compare the timing of wastewater-based alerts with that from population-testing-based alerts. It was indeed suggested (Cao et al., 2021), D'Aoust et al., 2021; Kumar et al., 2021 COVID-19 resurgence in a community could be anticipated with wastewater analysis. This idea should be investigated beyond retrospective analysis: for real-time detection of impending epidemic resurgence, it is in fact imperative to distinguish between fluctuations and robust increases. This observation is connects to the ROC analysis described in the previous chapter and aims at optimising the true positive signals and minimise the false ones. CoWWAn uses EKF-based predictions to capture robust trends in the epidemic dynamics and compare early alerts from case-based monitoring and wastewater-based predictions. Several lessons can thus be learned from this study.

8.3 Materials and Methods

Before proceeding to make predictions and use them for early alerts, performing the quantitive reconstruction is required. Several data sources, as well as mathematical integration of the aforementioned models, are needed.

8.3.1 Data

At least three types of data are necessary to fully calibrate the pipeline: COVID-19 RNA load in wastewater, detected cases from the area spanned by the sewage infrastructure, and possibly a proxy of the true case numbers, like the ratio off detected versus true, at a certain moment of time. The latter can be estimated, for instance, with seroprevalence studies.

The pipeline was tested and applied onto several countries, also outside Luxembourg (for which routine calculations are still performed at the time of writing). Some criteria were applied to construct the dataset. The list of wastewater sampling projects comes from the COVID19 Poops Dashboard (Naughton et al., 2021). Among them, only a few have readily accessible databases. From those, we opted for time series ranging from beginning 2021 up to at least six months later, to allow proper calibration. We also required that the sampling was at least weekly (on average), and that the corresponding case numbers were available. Smoothed data were discarded, to prevent bias and avoid breaking the causality of projections. If a single regional database had time series from multiple treatment plants, two representative ones were selected; usually they corresponded to the largest served population.

The corresponding detected case numbers, confirmed with positive RT-PCR tests, corresponded to officially reported numbers from the same databases. When available, region-specific seroprevalence studies like Snoeck et al. (2020); Pollán et al. (2020) were used to better tune the pipeline. Data from Luxembourg sewage sampling were supplied by the Research Luxembourg COVID-19 initiative CORONASTEP (researchluxembourg.lu/coronastep); detected case numbers and R_{eff} were obtained from the Luxembourg Ministry of Health website (COVID19.public.lu/fr/graph).

The selected datasets are reported in Tab. 8.1, together with their official sources, experimental protocols, units of measure for wastewater data, associated equivalent population, and detected case numbers. All datasets are updated up to August 2021. Note that different normalization protocols are considered: this illustrates the generality of the present method, but hinders quantitative comparison between fitted parameters. The equivalent populations is a proxy of the number of people who contributed to the wastewater load: it is the ratio between the sum of the pollution load collected during 24 hours by sewage facilities and services, and the individual pollution load in household sewage produced by one person in the same time.

Summary investigations are performed onto the collected data, to inform the development of the model. Some counties display peculiarities: Raleigh county rounded the data, after normalization: up-scaling might yield additional uncertainties. Netherlands reports countrywide averages over weeks and daily data from communal treatment plants: we use the latter to improve the temporal resolution, averaging over samples from the same day. Kitchener displays a sudden jump on May 17, 2021 during a time when case numbers remain stable. Scaling subsequent by a factor of 0.4 considerably improved CoWWAn's estimates, suggesting possible sudden changes in testing or sampling strategies. In the original publication (Proverbio et al., 2022b), both scaled and unscaled results are presented in the supplementary material. Following results rely on un-scaling.

For all countries, correlation analysis between tested positive case numbers and of RT-qPCR wastewater data was performed. An example is reported in Fig. 8.1; all other plots are available in supplementary material of Proverbio et al. (2022b). The correlation is rather close but not perfect: this aspect stresses the usefulness of wastewater data for epidemic monitoring, but reminds that more developed models based on complex epidemiological dynamics are necessary for quantification. Such correlation values, along with good but not perfect linear fitting, also justify the inclusion of a scaling parameter in the cost function used for parameter fitting (Eq. 8.10).

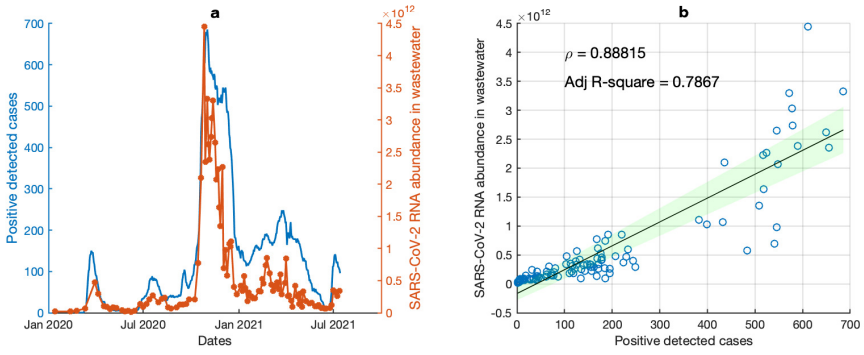


Fig. 8.1 **Left:** time series of detected cases (blue) and of SARS-CoV-2 RNA abundance in wastewaters, from Luxembourg data. **Right:** correlation between the two, together with adjusted R-square diagnostic of linear fitting (straight line).

8.3.2 The SEIR stochastic model

As anticipated, a modified SEIR model is used as basis for the Extended Kalman filter. They accurately describe COVID-19 epidemic dynamics (Proverbio et al., 2021; He et al., 2020). This simple yet descriptive model structure was chosen in stead of more complex ones – more difficult

Country	City/area	Population served	Time range yy/mm/dd	Units of measure	Source
Spain	Barcelona Prat de Llobregat	2,000,000	20/08/24 – 21/07/26 (11 months)	SARS-CoV-2 RNA copies/day	doi:10.5281/zenodo.4147073 https://sarsaigua.icra.cat/
Canada	Kitchener	242,000	21/01/11 – 21/08/22 (7 months)	SARS-CoV-2 RNA copies /PMMoV copies	https://www.regionofwaterloo.ca/en/health-and-wellness/covid-19-wastewater-surveillance.aspx
Slovenia	Kranj	40,000	20/10/08 – 21/08/02 (10 months)	SARS-CoV-2 RNA copies/PMMoV copies	https://github.com/sledilnik/data
Switzerland	Lausanne	240,000	20/10/01 – 21/08/04 (10 months)	SARS-CoV-2 RNA copies/day/10^5 equivalent inhabitants	https://sensors-eawag.ch/sars/lausanne.html
Slovenia	Ljubljana	280,000	20/09/10 – 21/08/02 (11 months)	SARS-CoV-2 RNA copies/copies PMMoV	https://github.com/sledilnik/data
Luxembourg	Luxembourg	610,000	20/02/25 – 21/08/02 (17 months)	SARS-CoV-2 RNA copies/day/10^5 equivalent inhabitants	https://www.list.lu/en/covid-19/coronastep/
USA	Milwaukee	615,934	20/08/25 – 21/08/04 (11 months)	MGC/person/day (gene copies per person per day)	https://www.dhs.wisconsin.gov/covid-19/wastewater.htm
Netherlands	Netherlands	17,178,109	20/09/07 – 21/07/30 (11 months)	SARS-CoV-2 RNA copies/day/10^5 equivalent inhabitants	https://data.rivm.nl/covid-19/COVID-19$_$rioolwaterdata.csv
USA	Oshkosh	78,300	20/09/15 – 21/08/19 (11 months)	MGC/person/day (gene copies per person per day)	https://www.dhs.wisconsin.gov/covid-19/wastewater.htm
USA	Raleigh	460,000	21/01/03 – 21/08/08 (7 months)	SARS-CoV-2 RNA copies/day/10^5 equivalent inhabitants	https://covid19.ncdhhs.gov/dashboard/wastewater-monitoring
Spain	Riera de la Bisbal	100,000	20/07/06 – 21/07/26 (13 months)	SARS-CoV-2 RNA copies/day	doi:10.5281/zenodo.4147073 https://sarsaigua.icra.cat/
Switzerland	Zrich	450,000	20/10/01 – 21/08/04 (10 months)	SARS-CoV-2 RNA copies/day/10^5 equivalent inhabitants	https://sensors-eawag.ch/sars/zurich.html

Table 8.1 Considered regions, associated equivalent population served by sewage facilities (according to official sources), time ranges of data, units of measure of wastewater data, and their sources. The sources listed here also include the case numbers, detected in the same area covered by the sewage system of interest.

to calibrate and identify (Roda et al., 2020; Kemp et al., 2021) – to reliably estimate community incidence from noisy data. As introduced in Sec. 4.1.2, SEIR models considers Susceptible $S(t)$, Exposed $E(t)$, Infectious $I(t)$ and Removed $R(t)$ compartments, and population flows governed by rate parameters. Recall that the standard interpretation of E is of individuals who have been exposed and infected, but who are not yet infectious due to incubation lag (Lai et al., 2020). The transition to becoming contagious is accommodated by an additional mean incubation period α^{-1} parameter. Properties include conservation of mass, $S(t)+E(t)+I(t)+R(t)=N$ (with constant N) and no possibility of re-infection ($R \not\to S$) during each period of infection transmission. In fact, although waning immunity has been registered, it usually occurs over months (Goldberg et al., 2021).

The original SEIR model is augmented into a stochastic version, to model intrinsic stochasticity in transmission processes and viral shedding. To do this, associate each transition between compartments with a random process. Similarly to what described in Chapter 7, the SEIR model assumes that each susceptible person has probability $\beta(t)I(t)/N\,dt$ to become infected on an infinitesimal time interval $[t,t+dt)$, and that infection events are independent. The number of new infections at $[t,t+dt)$ is thus a random variable from a binomial distribution $\mathscr{B}(n,p)$ with $n=S(t)$ and $p=\beta(t)I(t)/N\,dt$. Assuming high enough number of cases and stationary rate parameters over a time interval $\Delta t=1$ day (Gillespie, 2000a), the binomial distribution can be well approximated by a normal distribution with mean $\beta(t)S(t)I(t)/N\,dt$, and variance $\beta(t)I(t)/N\,dt(1-\beta(t)I(t)/N\,dt)S(t)=\beta(t)S(t)I(t)/N\,dt+\mathscr{O}(dt^2)$. The same reasoning can be repeated for all other transitions between compartments. As a result, the stochastic SEIR model is given by:

$$\begin{cases} \frac{d}{dt}S(t) = \frac{-\beta(t)S(t)I(t)}{N} - \sqrt{\frac{\beta(t)S(t)I(t)}{N}}w_1(t) \\ \frac{d}{dt}E(t) = \frac{\beta(t)S(t)I(t)}{N} - \alpha E(t) + \sqrt{\frac{\beta(t)S(t)I(t)}{N}}w_1(t) - \sqrt{\alpha E(t)}w_2(t) \\ \frac{d}{dt}I(t) = \alpha E(t) - \tau I(t) + \sqrt{\alpha E(t)}w_2(t) - \sqrt{\tau I(t)}w_3(t) \\ \frac{d}{dt}R(t) = \tau I(t) + \sqrt{\tau I(t)}w_3(t) \end{cases} \tag{8.1}$$

where w_j are mutually independent white noise processes. The β parameter is allowed to be time-varying, to reflect changes in mitigation measures (masks, vaccines, *etc.*), social interaction, and varying infectivity of emerging viral variants. $\beta(t)$ is considered a state variable to be estimated by the Kalman filter.

The viral flows into wastewater is modelled by a companion variable $A(t)$. It models effective number of active shedding cases producing virions to wastewater. A stochastic process is incorporated like above, to obtain a dynamics:

$$\frac{d}{dt}A(t) = \frac{\beta(t)S(t)I(t)}{N} - \gamma A(t) + \sqrt{\frac{\beta(t)S(t)I(t)}{N}}w_1(t) - \sqrt{\gamma A(t)}w_4(t)\,. \tag{8.2}$$

The A compartment is parallel to E, I, and R, so $S(t)+E(t)+I(t)+R(t)=N$ still holds. The influx to the A compartment is the same as that to the E compartment, while the outflux lumps together the dynamics of viral production (which is known to follow some kinetic trajectory in the hosts' body (Néant et al., 2021)), the decay rate of SARS-CoV-2 RNA in water (Gundy et al., 2009; Sala-Comorera et al., 2021), and inertia in abundance dynamics due to mixing in wastewater

collecting pools. Since the parameter γ lumps together properties of the virus and details on wastewater sampling; it is separately fitted for each region. Delays associated with in-sewer travel time are not considered, as they were estimated to be significantly lower than the transmission time scales (median of 3.3h versus 1 day (Kapo et al., 2017)). The A compartment thus allows to better follow the time evolution, including potential decaying inertia, and to consider explicitly the uncertainties associated to the shedding mechanism instead of the disease progression. Together, Eq. 8.1 and Eq. 8.2 form the combined SEIR-WW system.

Real-world measurements are compared with the following outputs from the model: number of daily detected cases, and virion abundance in wastewater. The number of detected cases on day $t \in \mathbb{N}$ is assumed to be a share of people passing the incubation period on that day:

$$y_c(t) = c_t \int_{t-1}^{t} \alpha E(s) ds \,, \tag{8.3}$$

where $c_t \in [0,1]$ is the share of detected cases out of all cases. This parameter accounts for under-testing and asymptomatic cases (see Sec. 8.4.1 and Eq. 8.8 for further discussion). c_t might depend on the day of the week, since there often are some weekday-dependent fluctuations in testing. The virion abundance in wastewater is assumed to be linearly dependent on A,

$$y_w(t) = \nu A(t) \,. \tag{8.4}$$

Here, ν is a tuning parameter to reflect the incubation, production and shedding of viral load from infected people (Néant et al., 2021; Wölfel et al., 2020; Miura et al., 2021) and normalisation of the wastewater data. Explicit corrections linked to precipitations or other environmental factors are not considered: previous studies evaluated them to be poorly correlated with RT-qPCR observations (Vallejo et al., 2021; Li et al., 2021a). An implicit tuning is nonetheless included in the fitting, *cf.* Eq. 8.10.

8.4 Developing the complete SEIR-WW-EKF model

The inputs for the EKF algorithm[2] (see Sec. 4.2.4) are the update function $f(x)$ to implement Eq. 8.1, the observation matrices $C(t)$ (for case numbers and wastewater data), the state noise Q (uncertainty on estimated variables), and the measurement error covariance $U(t)$ (uncertainties on empirical data). Then, the method evaluates the set of variables of interest $(x_{1\ldots6}(t) = [S,E,I,A,D,\beta](t))$ and their associated uncertainty matrix P. As introduced in Sec. 4.2.4, embedding a continuos-time dynamical system like Eq. 8.1 in the EKF is better performed after time-discretising it. The explicit Euler method is a reasonable option:

$$x(t+\Delta t) = x(t) + f(x(t))\,\Delta t + w(t) \,. \tag{8.5}$$

An auxiliary dynamic state variable $D(t)$ is also defined, related to the modelled number of daily new infections on each day:

$$\begin{cases} D(t) = 0, & \text{for } t \in \mathbb{N}, \\ \frac{d}{dt} D(t) = \alpha E(t) \,. \end{cases}$$

[2]The algorithm implementation was performed primarily by the co-author A. Aalto.

Practically, $D(t)$ is the differential counterpart of $y_c(t)$ and allows direct comparison with monitoring data. It is reset every day to keep track of current new infections.

Including the auxiliary variable, the state space has 6 dimensions: $x_{1\ldots6}(t) = [S,E,I,A,D,\beta](t)$. $R(t)$ is redundant because N is conserved; it is therefore omitted. Eq. 8.5 is complemented with the resetting of $x_5(t)$ to zero once per day. The function $f(x)$ can be de-coupled into a stoichiometric matrix B and a reaction function $r(x)$:

$$f(x) = \begin{bmatrix} -x_1x_3x_7/N \\ x_1x_3x_7/N - \alpha x_2 \\ \alpha x_2 - \tau x_3 \\ x_1x_3x_7/N - \gamma x_4 \\ \alpha x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1x_3x_7/N \\ \alpha x_2 \\ \tau x_3 \\ \gamma x_4 \end{bmatrix} := Br(x)\,.$$

As argued in the previous section, the state noise $w(t)$ can be well approximated as normally distributed with mean zero and covariance

$$Q(x) = \kappa^2 \Delta t B \,\mathrm{diag}(r(x)) B^\top + \Delta t Q_\beta$$

arising from the stochastic model Eq. 8.1; each white noise process w_j for $j = 1,\ldots,4$ in (8.1) corresponds to its respective reaction $r_j(x)$. The coefficient κ accounts for modelling errors, like deviations from the SEIR's assumption of homogeneous and perfectly mixed population. It can also be interpreted as a sensitivity tuning parameter: lower κ leads to higher sensitivity but noisy estimates; higher κ decreases sensitivity but increases robustness against noise. The parameter β has no dynamics through $f(x)$, but it is updated by the Kalman filter; its contribution to the covariance is in the matrix Q_β. This is zero everywhere, except for the element $Q_\beta(6,6) = q_\beta$; it acts as a tuning parameter controlling the magnitude of change of $\beta(t)$ in one day.

The measurements from the model are either detected cases on a given day and/or wastewater sampling. To this end, possible observation matrices are defined:

$$C_c(t) = [0\ 0\ 0\ 0\ c_t\ 0], \quad C_w = [0\ 0\ 0\ \nu\ 0\ 0], \quad \text{and } C_b(t) = \begin{bmatrix} C_c(t) \\ C_w \end{bmatrix}, \tag{8.6}$$

where the sub-indices refer to cases (c), wastewater (w), and both (b). Recall that c_t is the share of detected cases on a day t. It is a coefficient to reflect the testing strategy, which often depends on the day (reduced testing on weekends and on public holidays). Empirical measurements are assumed to come with an additive, normally distributed noise with mean zero and covariance $U(t) = \mathrm{diag}(U_c(t), U_w)$ (or just $U(t) = U_c(t)$ or $U(t) = U_w$ if only one of the measurements is available). The variance of observed cases, $U_c(t)$, is obtained by assuming that cases are detected independently with probability c_t. This leads again to a Binomial distribution for detected cases with mean $c_t D(t)$ ($D(t)$ is the number of new infections on day t). This is unknown; so, a smoothed estimate $D(t) = \bar{y}_c(t)/\bar{c}_t$ is used (barred variables stand for 7-days moving averages). The variance of the Binomial distribution is given by $U_c(t) = D(t)c_t(1-c_t)$. For Raleigh, recall that rounding off case numbers might yield additional uncertainties: 23^2 is thus added to the variance $U_c(t)$ to account for the (independent) uncertainty (23 is the largest possible rounding error, $N/20{,}000$).

The complete SEIR-WW-EKF algorithm is reported in Algorithm 1. It is used to calculate three different state estimates: $\hat{x}_c(t)$ using only case number data ($C(t) = C_c(t)$); $\hat{x}_w(t)$ using only wastewater data ($C(t) = C_w$ on days when wastewater sampling is done, otherwise Kalman update is skipped); $\hat{x}_b(t)$ using both case and wastewater data ($C(t) = C_b(t)$ on days when wastewater sampling is done, $C(t) = C_c(t)$ otherwise). These were then used to estimate the data that were excluded from the state estimation, initially for calibration and reconstruction of daily cases, and then to perform the desired predictions. Hence, we calculated $\hat{y}_w(t) := C_w(t)\hat{x}_c(t)$ and $\hat{y}_c(t) := C_c(t)\hat{x}_w(t)$ (C_i are the Kalman filter observation matrices, Eq. 8.6).

A simple outlier saturation for the wastewater data complements the Kalman filter. The model-predicted value for a wastewater measurement is given by $C_w\tilde{x}$, with prediction error variance $C_w\tilde{P}C_w^\top + U_w$. If the measurement differs from the model-prediction by more than four standard deviations, the measurement is replaced by the saturated value $C_w\tilde{x} \pm 4(C_w\tilde{P}C_w^\top + U_w)^{1/2}$.

.

Set $P_0 \in \mathbb{R}^{6\times 6}$ and $\hat{x}(0)$;
for $t = 1, ..., T$ **do**
  set $\tilde{x} = \hat{x}(t-1)$ and $\tilde{x}_5 = 0$;
  set $\tilde{P} = P_{t-1}$ and $\tilde{P}_{j,5} = \tilde{P}_{5,j} = 0$ for $j = 1, ..., 6$;
  for *i=1,...,M* **do**
    $\tilde{P} = (I + \Delta t J_f(\tilde{x}))\tilde{P}(I + \Delta t J_f(\tilde{x})^\top) + Q(\tilde{x})$;
    $\tilde{x} = \tilde{x} + \Delta t f(\tilde{x})$;
  end
  Measurement error covariance: $S = C(t)\tilde{P}C(t)^\top + U(t)$;
  State update: $\hat{x}(t) = \tilde{x} + \tilde{P}C(t)^\top S^{-1}(y(t) - C(t)\tilde{x})$;
  Error covariance update: $P_t = \tilde{P} - \tilde{P}C(t)^\top S^{-1}C(t)\tilde{P}$;
end

Algorithm 1: The Extended Kalman filter for the SEIR-WW model with time step $\Delta t = 1/M$ (we use $M = 10\ \mathrm{d}^{-1}$). J_f is the Jacobian of the function $f(x)$, obtained from the Jacobian of the reaction function by $J_f = BJ_r$. The algorithm is standard, but the prediction step consists in solving a time-discretised ODE. The observation matrix $C(t)$ is chosen from the three possibilities described in (8.6). Note the resetting of $D(t) = \tilde{x}_5$ before the prediction loop.

The final workflow is illustrated in Fig. 8.2. It combines a model of a dynamical system with measurements (case numbers y_c, wastewater measurement y_w or both) obtained from the real system that is being modelled. At each time step, CoWWAn first predicts the next state $x_{1...6}(t)$ by propagating the old state estimate using the underlying model. From the predicted state estimate, the predicted measurement is calculated using the measurement model. Finally, the state estimate is updated based on the discrepancy of the true measurement and the model-predicted measurement. The model's state estimate reflects the state of the real system, and it can be used to predict the system's dynamics in the future.

8.4.1 Model parameters

CoWWAn has some free parameters, to be fixed upon data or with educated assumptions. Sensitivity of model estimates upon assumptions is performed and displayed in Fig. 8.3.

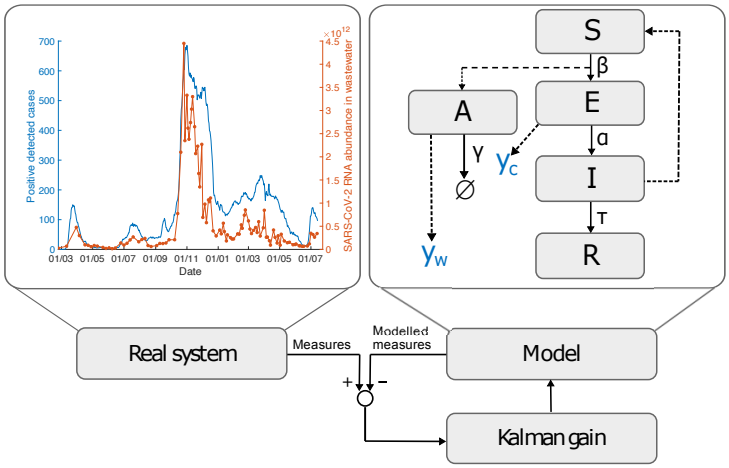


Fig. 8.2 Model workflow. The Kalman filter combines measurements from the real system with predictions from the dynamical model, which extends a SEIR model. Empirical data are daily positive cases, shown in blue as the smoothed moving average, and wastewater sampled data, shown in orange with unit of measure of RNA copies/day/100,000 equivalent inhabitants (example for Luxembourg). Details of the SEIR blocks are described above.

Most time series data begin after the pandemic already diffused within a region, so the initial sizes for the E and I compartments are automatically computed from the data by

$$E(0) = \frac{\eta_0}{\alpha}\left(1+\sum_{t=1}^{5}\frac{y_c(t)}{5}\right) \qquad \text{and} \qquad I(0) = \frac{\eta_0}{\tau}\left(1+\sum_{t=1}^{5}\frac{y_c(t)}{5}\right). \tag{8.7}$$

Here, α and τ are the transition rates $E \to I$ and $I \to R$, respectively. Their inverses are the average duration of staying in the original compartment (see also Ch. 7). η_t is the average ratio of total and detected cases at day t. 5 data points are a trade-off between approximating values on the first day and considering sensitivity to noise. The model is little sensitive to this choice, *cf.* Fig. 8.3.

η_t should also be estimated to link the measurements of population testing with those of wastewater analysis (ideally objective and insensitive to testing capacities). For the first wave in Luxembourg until June 1, 2020, $\eta_t = 3$ following the results of early prevalence studies (Snoeck et al., 2020). Later, $\eta_t = 1.8$; this choice was cross-validated with an independent SEIR model fitted to Luxembourg data (Kemp et al., 2021). The reduction is linked to a large scale testing campaign (Wilmes et al., 2021), and to overall increased testing activity.

Other regions, are more challenging: the available prevalence studies usually consider early epidemic stages, when there ae little to no available wastewater data. Large changes between first and subsequent waves are expected; hence, estimates from these prevalence studies are not usable for later stages. In the absence of additional reliable values, $\eta_t = 1.8$ is maintained for all other regions and its impact on modelling results is assessed in Fig. 8.3. It is possible to further calibrate such values with new and tailored prevalence studies. In principle, one initial estimate for η_t is sufficient, to match virion abundance in wastewater with total case numbers. Once the calibration is done, CoWWAn estimates the total number of infections, including both detected

and undetected cases. To obtain an estimate of detected cases, a potentially time-varying estimate of η_t is necessary.

Daily ratios c_t of detected and total cases are obtained from η_t modulated by a weekly testing rhythm. c_t is automatically estimated, by first averaging over five weeks, and then by a moving average over three weeks:

$$\tilde{c}_t = \begin{cases} \frac{35}{5} \frac{\sum_{j=0}^{4} y_c(\mathrm{mod}(t-1,7)+1+7j)}{\sum_{s=1}^{35} y_c(s)} & \text{for } t \leq 35\,, \\ \frac{21}{3} \frac{y_c(t-7)+y_c(t-14)+y_c(t-21)}{\sum_{s=t-20}^{t} y_c(s)} & \text{for } t > 35\,. \end{cases}$$

These values are normalised by the weekly moving average:

$$c_t = \frac{7\tilde{c}_t}{\eta_t \sum_{s=t-6}^{t} \tilde{c}_s}\,. \tag{8.8}$$

The procedure for the first five weeks is not causal, but some data is anyway needed for model calibration. Later, values for c_t are causally determined from data. To obtain final values on public non-weekend holidays, c_t is reduced by a factor of 4 from the value given by Eq. 8.8 to account for reduced testing. In case the weekly rhythm is not regular, manual tuning could help improving the performance. Alternatively, one can estimate c_t based on number of performed tests, for example.

The variance of the wastewater measurements U_w is also estimated from data, by

$$U_w = K \left(\mathrm{median} \left| y_w(t_j) - \frac{1}{5} \sum_{i=j-2}^{j+2} y_w(t_i) \right| \right)^2 . \tag{8.9}$$

Each t_i is the time point when wastewater sampling is done. The scaling factor K is either $1/10$ when wastewater data is used alone and $K = 1$ when both case and wastewater data are used, like for the outlier detection. In the plots of wastewater data reconstruction, $K = 1$ is used to plot the uncertainty envelope.

When optimising the model to reproduce the observations, there is a final detail to consider: the dependency of wastewater measurement to the number of detected cases is not perfectly linear due to dilution, non-mixing environment and other factors, (*cf.* Vallejo et al. (2021) and Fig. 8.1). A simple power transformation is applied to the wastewater samples, for which the exponent ε is regarded as a tuning parameter of slight nonlinearity. ε and the other proportional parameters γ and ν are fitted by calculating the Kalman filter state estimate using the wastewater data, and then minimising the cost function

$$\begin{aligned} &\min_{\gamma,\nu,\varepsilon} \sum_{t=1}^{M} \left(y_c(t) - C_c(t) \hat{x}_w(t;\gamma,\nu,\varepsilon) \right)^2 \\ &\text{such that } \gamma \in [0.2, 4],\ \varepsilon \in [0.4, 1]\,. \end{aligned} \tag{8.10}$$

This minimises the error in estimating the case numbers by the EKF state estimate using only wastewater data.

All model parameters, either fixed by literature or fitted from Eq. 8.10, are reported in Tab. 8.2. Note that, due to different wastewater data normalisations, ν parameters are not comparable be-

tween regions. Similarly, ε parameters might depend on the used techniques. Data from different laboratories may contain significant differences (Cluzel et al., 2022).

Symbol	Explanation	Value	Source
α	Rate $E \to I$	0.44 d^{-1}	(Kemp et al., 2021)
τ	Rate $I \to R$	0.32 d^{-1}	(Kemp et al., 2021)
$\beta(0)$	Initial infectivity	0.44 d^{-1}	—
Δt	Time step length	0.1 d	—
$q_{\beta,1}$	Variance of $\beta(t+1)-\beta(t)$ when $t \leq 30$	0.05^2 d^{-2}	—
$q_{\beta,2}$	Variance of $\beta(t+1)-\beta(t)$ when $t > 30$	0.005^2 d^{-2}	—
κ	EKF sensitivity parameter	4	—
N	Population size	regional	see Supplementary Tab. 1
γ	Rate $A \to \varnothing$	regional	Fitted by Eq. 8.10
ν	Ratio of y_w/A	regional	Fitted by Eq. 8.10
ε	Exponent in nonlinear mapping of WW data	regional	Fitted by Eq. 8.10
U_W	Measurement error variance of wastewater data	regional	Estimated by Eq. 8.9
$E(0)$	Initial size of E-compartment	regional	Estimated by Eq. 8.7
$I(0)$	Initial size of I-compartment	regional	Estimated by Eq. 8.7
$\mathrm{var}(E(0))$	Uncertainty of $E(0)$	regional	$(E(0)/2)^2$
$\mathrm{var}(I(0))$	Uncertainty of $I(0)$	regional	$(I(0)/2)^2$

Table 8.2 Model parameters: description. Parameter symbols, descriptions, values, and their sources. The parameter q_β, controlling the allowed change of $\beta(t)$ in one day, is changed after 30 days. This is done to allow rapid changes in the beginning of the pandemic, when a strict lockdown quickly suppressed its propagation and to account for errors in initial $\beta(0)$. d stands for "days". When the source is not indicated, the parameter values is first initiated as an educated guess and then tested with sensitivity analysis (see Fig. 8.3).

Fig. 8.3 assessed the sensitivity of model performance on assumed parameter values, with Luxembourg as a reference. It demonstrates the robustness of the model and justifies the current parameter choices. Sensitivity analysis involves varying the reference parameters up to $\pm 50\%$ of their original value. In general, the projections are consistent and slightly vary for values very far from reference ones. The model is most sensitive to the parameter c_t, which is usually estimated with independent methods. The minimal error corresponds to the reference value, while deviations induce larger errors. Changes in c_t are normally compensated in the pipeline by a likewise change in ν. This observation justifies the differing fitted values reported in Tab. 8.3 for each region. It also recall the relevance of accurate seroprevalence studies to reduce projection errors. In fact, projection errors grow slower for overestimated c_t than for underestimated c_t. Therefore, in case a good estimate of eta_t is missing, using a possibly overestimated rather than underestimated value is advised for short-term projections. However, this may lead to higher overshoots in long-term projections due to overestimation of the susceptible population size.

8.4.2 Analysis of model outputs

To obtain variables of epidemiological interest, state estimates outputted by the CoWWAn model are further analysed. Note that two estimates using only wastewater data are computed: one without interpolating data between sampling days (WW) and one with linear interpolation (ipWW).

Parameter	Barcelona	Kitchener	Kranj	Lausanne	Ljubljana	Luxembourg
N	2,000,000	242,000	40,000	240,000	280,000	634,730
γ	0.20 d^{-1}	4.00 d^{-1}	1.43 d^{-1}	3.05 d^{-1}	3.21 d^{-1}	1.62 d^{-1}
ν	$4.86 \cdot 10^{-2}$	2.73	1.38	$6.67 \cdot 10^{10}$	2.77	$6.40 \cdot 10^{4}$
ε	0.40	0.40	0.40	1.00	0.526	0.613
U_W	656	2.68	58.5	$3.43 \cdot 10^{23}$	511	$1.75 \cdot 10^{12}$
$E(0)$	1,824	379	17	156	76	8
$I(0)$	2,527	525	24	203	105	11

Parameter	Milwaukee	Netherlands	Oshkosh	Raleigh	Riera	Zurich
N	615,934	17,178,109	68,000	460,000	100,000	450,000
γ	4.00 d^{-1}	0.368 d^{-1}	4.00 d^{-1}	4.00 d^{-1}	1.20 d^{-1}	0.547 d^{-1}
ν	0.134	185	3.73	$2.58 \cdot 10^{3}$	12.7	$1.87 \cdot 10^{7}$
ε	1.00	0.500	0.866	0.434	0.400	0.789
U_W	1.03	$2.50 \cdot 10^{11}$	50.7	$6.00 \cdot 10^{8}$	$1.69 \cdot 10^{3}$	$3.44 \cdot 10^{17}$
$E(0)$	298	4,412	166	1,815	4	238
$I(0)$	413	6,112	231	2,514	6	330

Table 8.3 Model parameters: fitted values, for each region. Initial values for SEIR compartments are in units of equivalent inhabitants.

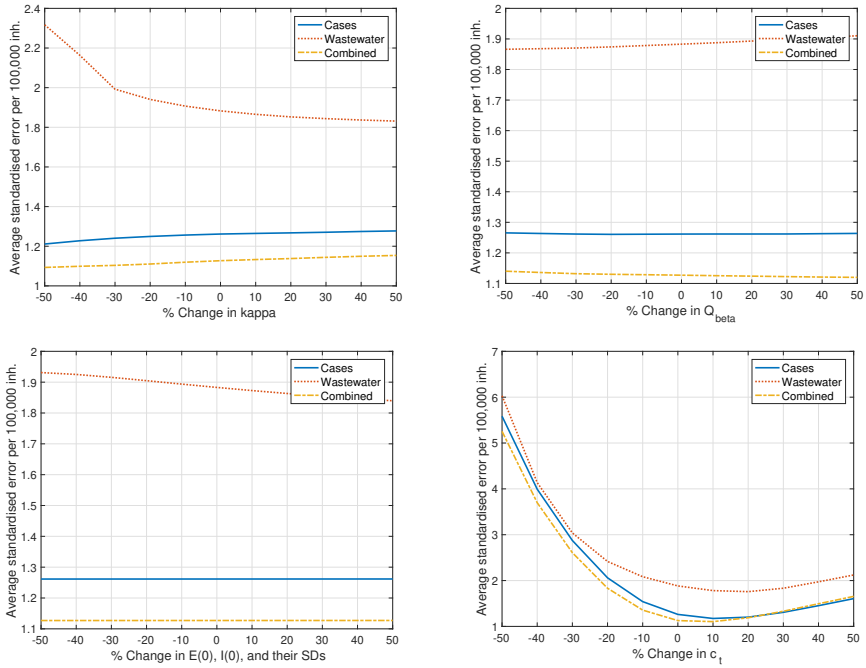


Fig. 8.3 Sensitivity analysis against changes in hand-picked parameters. The plots show the changes in 7-day window prediction performance for Luxembourg when a parameter is changed. The no-change numbers (0%) are those shown for Luxembourg in Tab. 8.3. The "normalised error" corresponds to the measure employed throughout the analysis and defined in Eq. 8.12. Changes in c_t (the share of detected cases) is compensated by changes in ν by the same amount. Note that other parameters are kept fixed. Some changse could be compensated by re-running the parameter tuning pipeline. Note that +50% change in c_t corresponds to a situation where 83% of total cases are detected.

The time-varying effective reproduction number R_{eff} (*cf.* Ch. 7[3]) is directly extrapolated as (Kemp et al., 2021):

$$R_{\text{eff}} = \frac{\beta(t)}{\tau}\frac{S(t)}{N}, \tag{8.11}$$

where $\beta(t)$ and $S(t)$ are state estimates. For N and τ, see Tab. 8.2.

Short and mid-term projections are possible at any time t_0: stop the Kalman filtering and simulate the model forward in time, starting from the latest state estimate and keeping the infectivity parameter constant ($\beta(t) = \widehat{\beta}(t_0)$). The effect of uncertainty in the parameter estimate $\beta(t_0)$ can be also quantified: simulate envelopes using $\widehat{\beta}(t_0) \pm 2\sqrt{P_{t_0}(6,6)}$; for every t, $P(6,6)$ represents the variance associated to β in the Kalman filter update, as discussed in Algorithm 1. Other uncertainties are omitted in these simulations; therefore, the envelope of short-term projections might be underestimated.

As introduced in Sec. 8.2, quantifying the quality of short-term projections using either case data only, wastewater data only, or both provides more reliable estimates of the epidemic unfolding over short time horizons. So, at each time step when wastewater data is available, the Kalman filter state estimation is stopped, and the SEIR-WW model is simulated T days forward without taking into account any new data. The total number of observed cases from the projection is calculated and compared with the actual number of observed cases during the same time horizon. Their absolute difference constitutes the prediction error. The prediction errors are standardised by the square root of the true number of cases. The latter represents the standard deviation estimate assuming case numbers on a given time are binomially distributed. Thereby, standardised scores are averaged over all time points on which the prediction is made, obtaining an overall average normalised error. This is scaled per 100,000 equivalent inhabitants to enable comparison between countries. Overall, the scaled average standardised prediction error ξ is:

$$\xi = \frac{1}{M}\sum_{i=1}^{M}\frac{|x_i - \tilde{x}_i^j|}{\sqrt{x_i}}\frac{100{,}000}{N}. \tag{8.12}$$

Here, i is the index of each point in any time horizon $[t_0, T]$ with M points in total; j is an index that considers the original type of data used for projections, *i.e.* $j = \{c, w, b\}$ for case data only, wastewater data only, or both combined. In the state estimate using combined data, the wastewater data are not interpolated. Tilde-ed variables are the Kalman projections while non-tilde-ed variables correspond to measured data. N is the equivalent population of interest (*cf.* Tab. 8.3).

8.5 Results

As introduced, the implemented pipeline aims at quantitatively mapping wastewater data to cases, reconstruct and project epidemic dynamics, and assess predictions and early-warning capabilities to assist healthcare management.

[3] Note the slightly different notation, to highlight the different contributions to the term. Time dependence is implied.

8.5.1 Inference of case numbers and epidemic indicators

To begin with, the model is calibrated to test cases via parameter fitting, whose results are in Tab. 8.3. After that, CoWWAn infers state variables and parameters, and quantitatively reconstructs the time evolution of observed cases from wastewater data. Reconstruction results are in Fig. 8.4a. Thanks to the model structure and the interpretation of he underlying SEIR model, the results can be compared with the true number of detected cases. They are shown in Fig. 8.4a, red and black lines, for Luxembourg. The global magnitude of the shedding population is in Fig. 8.4a, blue line, from mean estimates. The latter is an extrapolation from CoWWAn estimates and information about the "dark number" of undetected cases (provided as a model parameter, see above); further independent studies to estimate this quantity help to fine-tune the results.

The results are overall better than those from simpler linear regression models (even after data curation to reduce the noise, like Vallejo et al. (2021)): CoWWAn's inferences achieve consistently higher correlation (Fig. 8.4b, blue and red sets), with $\rho_{\text{Pears}} \in [0.7;0.9]$ for all considered regions. Additional details concur in further improving the reconstruction performance. On the data collection side, frequent sampling yields $\rho_{\text{Pears}} = 0.91$ and 0.95, respectively for Luxembourg and Milwaukee with two or more samples/week, while Barcelona has $\rho_{\text{Pears}} = 0.7$ with one sample/week (Fig. 8.4d). Here, interpolating wastewater data points before the EKF estimation can improve the reconstruction. In general, an adequate sampling rate is recommended to improve the EKF performance, since it improves its predictions as new data points are available. In addition, unnoticed changes in the share of detected cases or in testing/sampling strategies yield discrepancies, as well as augmented positivity rates during high case numbers (Max Roser Hannah Ritchie et al., 2020). On a side note, detecting such discrepancies can provide additional evidence about under-testing and could guide targeted scaling of population tests. In general, the Extended Kalman filter improves its predictions as new data points are available, so an adequate sampling rate is recommended to improve its performance.

Using reconstructed data and Eq. 8.11, one can derive the effective reproduction number R_{eff} to characterise diffusion trends within a community (Huisman et al., 2021). Fig. 8.4a reports an example for Luxembourg: the inference is consistent with the indicator reported by the Ministry of Health on its website (see Section 8.3.1). Importantly, it highlights the same noteworthy trends that, in practical use, are often more useful than the exact values. Observe the three waves in 2020 (March, June and late October), a small rebound in March 2021 attributed to the emergence of the alpha variant and one wave in late June 2021, attributed to the gamma and delta variants; all waves are characterised by $R_{\text{eff}} > 1$. Supplementary material of Proverbio et al. (2022b) reports values for all other countries.

8.5.2 Short-term predictions of epidemic trends

After case numbers reconstruction, the mechanistic model underlying CoWWAn allows interpretable and causal-based predictions. This represents a plus with respect to correlation or extrapolation-based alternatives. To make such predictions, it is sufficient to simulate the model forward at any desired time, starting from the latest state estimate and keeping the transmission parameter β constant (*cf.* Section 8.4.2).

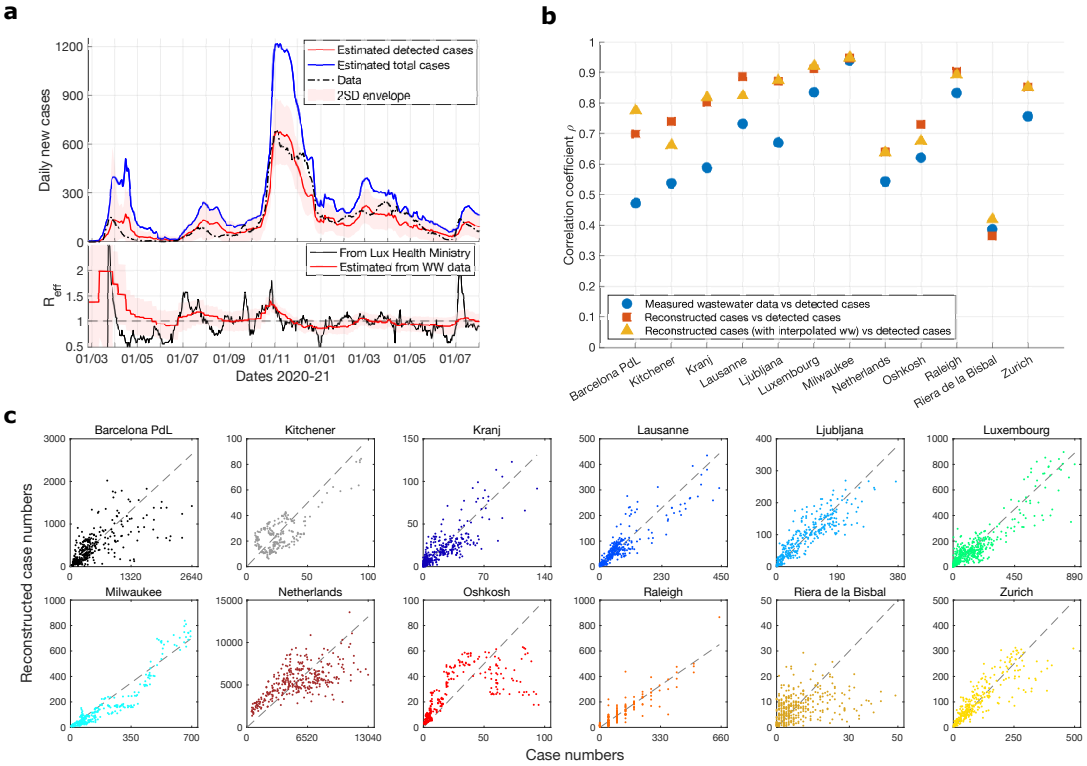


Fig. 8.4 Reconstruction of case numbers and inference of epidemic indicators. **(a):** Reconstruction example for Luxembourg. Top: Comparison of case numbers: official detected data (black line), reconstructed from wastewater data (red) including the 2 Standard Deviations $\simeq 95\%$ confidence interval (shadowed region), and total positive cases inferred by CoWWAn (blue). Bottom: R_{eff}, estimated (red, with its associated 2 SD shadowed region) or officially reported by the Luxembourg Ministry of Health. (b): Pearson's correlation coefficients ρ_{Pears} from linear regression between detected cases and measured wastewater data (blue), ρ_{Pears} between detected cases and CoWWAn-reconstructed case numbers from wastewater data (red, corresponding to correlation values from panels c), and ρ_{Pears} between CoWWAn-reconstructed case numbers from wastewater data (after interpolating wastewater data) and detected cases (yellow). **(c):** Reconstruction results for all considered regional areas, compared with detected case numbers. The dashed line represents equal values.

For short-term periods, predictions with CoWWAn are good. For Luxembourg, Fig. 8.5a shows an example of such 7-days predictions for each day of wastewater sampling, where the number of detected cases (blue) is compared with the predicted numbers derived from wastewater data or from case number data. Wastewater-based short-term predictions are well correlated both with case-based projections ($\rho_{\text{Pears}} = 0.95$) and with true case numbers ($\rho_{\text{Pears}} = 0.94$). For all other countries, similar figures and summary tables can be found in supplementary material of Proverbio et al. (2022b). They display consistent results.

The prediction performance is determined by Eq. (8.12), as the average standardised prediction error. As one could expect, performances of the wastewater-based pipeline are usually slightly lower than those from pure case numbers, as the former reconstructs the case numbers themselves before making the predictions. Nonetheless, all regional estimates lie within one standard deviation of the 1:1 (equal performance) line (Fig. 8.5b), thus bearing acceptable performance. The only exceptions are Oshkosh, probably due to under-testing during late 2020, and Kranj, whose low

case numbers are subject to larger uncertainties. In general, the largest discrepancies are observed when case numbers plateau or decline after a rapid increase. These are associated with $\beta \neq$ cost and yield potential overshoots of the predictions (see also Fig. 8.5a). This effect is associated to large changes in social activities during epidemic waves and rapid implementations of stricter restrictions; they are not explicitly included in the model but implicitly learned from the epidemic curve by the EKF, with some delay.

The standardised error grows quite linearly with increasingly long prediction horizons (Fig. 8.5c). There, wastewater predictions, usually less susceptible to daily fluctuations, are more "stable" than those based on case numbers: in fact, their uncertainty grows slower for longer prediction horizons. This aspect allows quantifying and comparing the precision for different horizons.

One final practical remark: CoWWAn's EKF-based approach enables integrating different types of data to further improve the quality of predictions. It can thus include both wastewater and case data together, slightly but systematically improving the prediction accuracy (Fig. 8.5c). This further suggests that wastewater data contains independent information about the epidemics state, as previously suggested by Fernandez-Cassi et al. (2021).

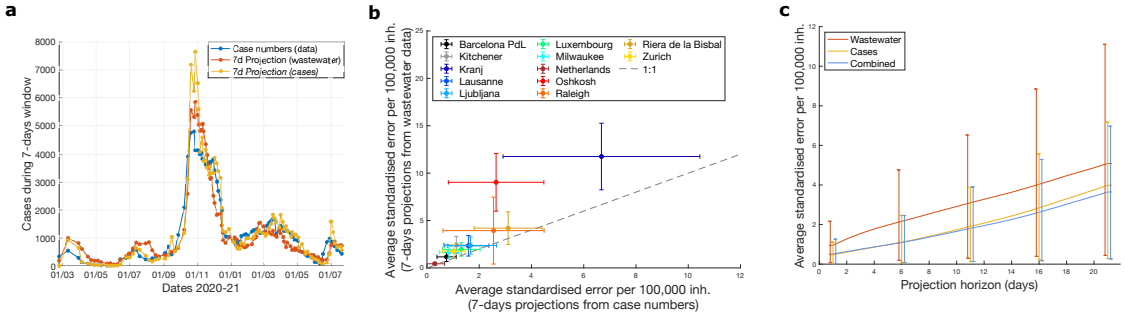


Fig. 8.5 Predictions of future epidemic trends using CoWWAn. **(a):** Prediction examples for Luxembourg, comparing predictions over the 7-days ahead of each point (either estimated from case numbers or wastewater data) with the true detected cases in the same time period. **(b):** Comparison of wastewater-based and cases-based predictions. The performance is evaluated in terms of average standardised error, normalised to equivalent population. The dashed line represents equal values. Error bars correspond to one standard deviation. **(c):** Predictions performance for different time horizons (mean and 80th percentiles over the considered regions) for three inputs: case numbers, wastewater data, or both data combined. For all panels, "inh." stands for inhabitants.

8.5.3 Long-term projections of epidemic scenarios

Typically, heterogeneous and evolving adaptations of population behaviour and institutional measures make epidemic forecasts only meaningful for relatively short time horizons. In fact, the assumption of constant β might not hold any-more; moreover, it is known that small uncertainties for short-term predictions are amplified over longer periods and the precision drops, similarly to what happens in weather forecasts (Petropoulos et al., 2020). Over long-term time horizons, we thus speak of "projections" of contrasting scenarios. They assume no changes in infection dynamics and allow counterfactual analysis of differing unfoldings from current social or pharmaceutical measures and/or changed viral infectivity (Fig. 8.6a,b). They can also be used to investigate plausible scenarios, by artificially modifying the model parameters.

As for other models applied to complex systems, projection uncertainties increase with longer time horizons (Fig. 8.6b), reflecting the set of potential changes of conditions. Nonetheless, projections based on case numbers or on wastewater data are consistent with each other within error bounds, supporting the possibility of using wastewater data for consistent what-if analysis. In addition, the mechanistic-based model allows assessing the changes in desired precision. Models applied in quickly changing conditions are known to be uncertain (Santosh, 2020). Similarly, our projections' precision varies depending on whether they are conducted during a rapid increase of case numbers or during stable trends, calling for caution in interpreting these results as plausible projections rather than forecasts.

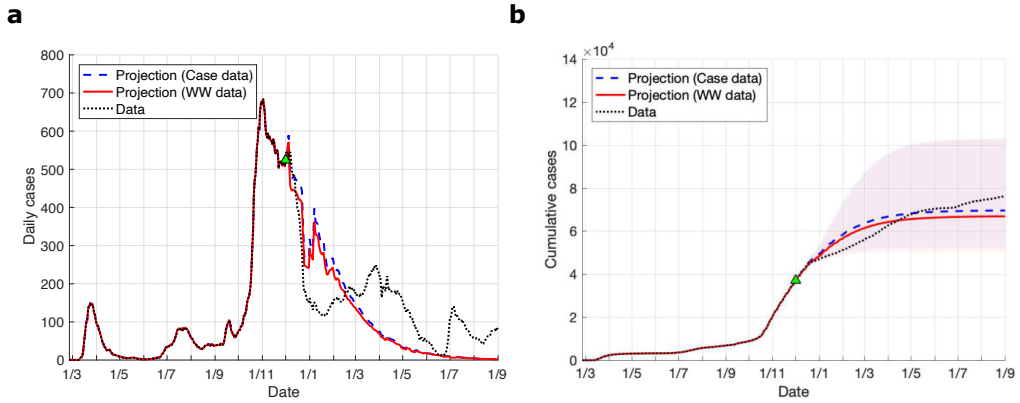


Fig. 8.6 Long term projections using CoWWAn. **(a):** Long-term projected curves of daily cases compared with daily detected case numbers. **(b):** Long-term projected curves of cumulative cases compared with cumulative detected case numbers. Blue and red ribbons represent ± 2 SD error bounds (SD stands standard deviation); note that the ribbons might overlap. Both panels (a) and (b) report examples for Luxembourg data, with projections starting at the date marked by the green triangle.

8.5.4 Modelling assesses wastewater warning performance

Finally, consider the assessment of wastewater data early-warning capabilities. To begin with, let us clarify what we mean by "early". In this specific case, the pandemic is unfolding, and wastewater data capture viral load shed by already infected individuals. So, anticipating re-emergences *tout-court* is not the scope. Instead, the aim is to give immediate alerts and possibly anticipate the curves of *detected* positive cases, which often lag up to 7-10 days beyond real infections (Wu et al., 2020). In turn, this would constitute an early alert for hospitalizations and healthcare burdens, which usually follow by another two-three weeks (Kemp et al., 2021).

Cao et al. (2021), D'Aoust et al. (2021) and Kumar et al. (2021) had suggested that wastewater analysis could provide early warnings for COVID-19 resurgence in a community. They employed a retrospective analysis to identify peaks in wastewater data that occurred before or close o peaks in detected case numbers. However, for the real-time detection of impending epidemic resurgence, distinguishing between fluctuations and robust increases is crucial to optimise the true positive signals and minimise the false negatives. What if, in real time, the marked peak was instead considered spurious? Or, if a peak that had been considered significant at day $\hat{t}$ is instead judged irrelevant by some retrospection performed at day $\hat{t}+$ some weeks? Fig. 8.7 provides an example on a dummy variable Y. At time $\hat{t} = 20$, only the blue time series is available. Is $Y(\hat{t})$ the first hallmark of an increase? Three scenarios are possible: a noteworthy subsequent

increase (orange), a stabilization on slightly higher values (green) or a spurious jump followed by "standard" dynamics (red). Only observers at time $\hat{t}+5$ (or more) can confidently assess the value of $Y(\hat{t})$. However, at $\hat{t}$, one can test the likelihood of different scenarios using causal-based models.

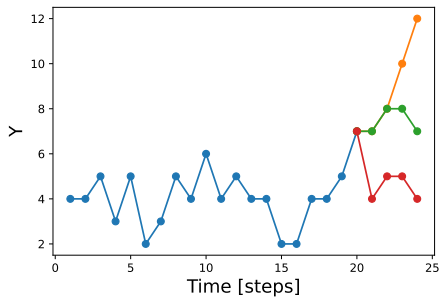


Fig. 8.7 Example time series of dummy data Y, to illustrate how observers can get dilemmas when new data points become available. Is $Y(\hat{t})$ a signal of a true increase? Three scenarios are instead possible.

To evaluate the alerting power of on-line (real time) systems, it is thus not sufficient to compare developed time series with a retrospective analysis. CoWWAn addresses this challenge by the EKF-based predictions that, thanks to the causal structure, capture robust trends in the epidemic dynamics. The shot-term predictions introduced in Sec. 8.5.2 do the job, as they well characterise spurious or significant trends. CoWWAn thus allows to see if case-based and wastewater-based predictions provide early warnings of COVID-19 resurgence.

Fig. 8.8 displays the 7-days predictions about pandemic trends, obtained from wastewater data (red) and from detected case numbers (yellow), and compare them with the true observed evolution (blue). Do red and yellow curves correctly track the increasing trend of the blue curve? Observe that, overall, the prediction curves accurately increase when a new COVID-19 wave is observed in a region. However, the timing might slightly differ depending, *e.g.*, on the testing frequency[4]. On one hand, this analysis demonstrates the potential of wastewater data to detect incoming increasing trends; on the other, it quantitatively verifies the recent calls by Bibby et al. (2021) for cautious interpretation: alerts based on wastewater analysis might be just-on-time or even lagging slightly behind the true infection waves.

Nonetheless, wastewater-based reliable alerts are often more advanced than those based on case numbers alone, *e.g.* for Kitchener or Raleigh. As a result, wastewater-based monitoring could indeed be an effective method to detect new waves of infection, but the lead time should be carefully assessed case-by-case, according to the sampling frequency and other characteristics of the wastewater-analysis pipeline. In short, reliable warnings can be triggered, but properly verifying how "early" is still demanded to future and tailored studies.

8.6 Discussion

This chapter discussed a mathematical model to leverage alternative data to construct a monitoring system, ideally capable of anticipating detected epidemic insurgence. The whole pipeline, termed CoWWAn, combines a SEIR model with an extended Kalman filter to process wastewater data in

[4]Other examples, for the remaining increases observed in the considered countries, are reported as supplementary material for Proverbio et al. (2022b).

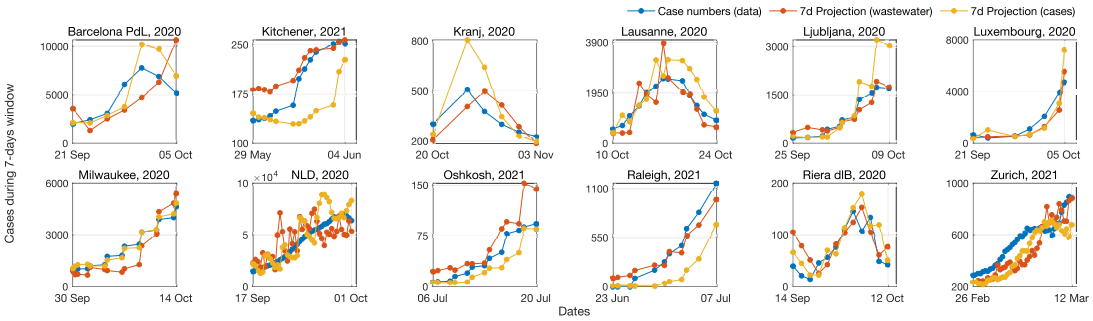


Fig. 8.8 Zoom into the epidemic resurgences visually recognised in the considered regions. Short-term projections used to identify robust trends in epidemic resurgence, for different examples (one per region; other examples in Supplementary Fig. 16). We compare 7-days projections from case numbers and from wastewater data with the true detected case numbers.

an automated and mechanistic-based manner. The model supports the reconstruction of infection curves from wastewater data and allows short-term prediction and long-term projections of future trends. Such predictions can provide early and robust information, particularly relevant for healthcare management since hospital admission is downstream of the susceptible-exposed-infectious flow (Kemp et al., 2021).

The model itself is interesting and useful in practice. In fact, CoWWAn can be easily applied and extended to different areas, and it overcomes some of the limitations of previous studies. In addition to mapping wastewater data to case numbers, it allows improving R_{eff} estimation. Previous studies attempted that (Huisman et al., 2021), but did not achieve such good performances. In addition, CoWWAn allows quantitative analysis of early-alert capabilities of wastewater data, building an additional step for the estimation of lead times in real-world settings. Overall, it connects mathematical theory and its recent developments with real-world capabilities and needs.

Obviously, the approach comes with some limitations. To begin with, it provides information on a community level but does not single out the infected individuals. Hence, it does not relate to contact tracing or to detailed information like age distribution or infection clusters. This is typical of wastewater-based monitoring.

The reconstruction of case numbers depend on the mean-field SEIR approximation: although meaningful when concentrating on average epidemic trends (Kollepara et al., 2021), it might yield uncertainties in case of heterogeneous behaviours like clusters. In addition, tailoring region-specific model parameters is recommended to fine-tune the performance and reduce the uncertainties over the estimates. Usually, the parameters can be estimated with independent methods or educated prior information, in particular concerning seroprevalence. The current set of proposed parameters might not be complete for all countries and some system-specific tuning is necessary despite the generality of the method.

The projections are somewhat sensitive to the ratio of total and detected cases, see Fig. 8.3. Long-term projection are more influenced, due to potential errors in the estimated level of natural immunity in the population. Short-term projections are less sensitive, since any error in the estimated size of the susceptible population is compensated by the infectivity parameter estimate. Note that this sensitivity is a shortcoming of every model-based projection. Reliable estimates for

the ratio of true versus detected cases are thus necessary during model calibration. Seroprevalence studies may provide best estimates, especially if they can distinguish between antibodies from previous infection and vaccination (Suhandynata et al., 2021).

Data quality and sampling frequency are key for good predictions: escalating the sampling precision and rates improves the model estimations. Future studies could optimise the data needs to obtain desired precision levels, depending on the costs associated.

Interestingly, the model can be used for additional studies. To begin with, it is generic enough to be expanded to other epidemic contexts. Future discrepancies between wastewater-based estimates and detected cases might be used as indications about changes in the share of detected cases and could be used to trigger a warning against potential undertesting.

Finally, the present study provides several lessons learned for on-line early warning systems. It allows investigating the conceptual difference between "predictions" and "projections", which apply depending on the assumptions of constant parameters hold or until the uncertainty estimates become significant and overcome the desired confidence. The study also allows making sense of short-term predictions associated with noisy data and probabilistic estimates. Depending on the nature of complex systems under study, their associated uncertainties and dynamic characteristics, such "predictions" and early alerts might pertain more to the field of dynamic forecasts or of risk assessment. This study thus paves the way to future studies, that further quantify lead times, precision/recall and real-time implementations.

Code availability

The code to reproduce the analysis is available at gitlab.lcsb.uni.lu/SCG/cowwan.

Chapter 9

Discussion

This thesis addresses the topic of critical transitions in natural systems and the challenge of providing reliable warnings against them. To verify or falsify theoretical predictions and claims, the thesis bridges mathematical theory with real-word data, thus setting solid foundations for the applications of warning signals.

The present work provides a systematic classification of critical transitions and of corresponding dynamical regimes, reviews state-of-the-art ideas, methods and EWS, highlights their strengths and open research questions, and tests hypotheses and performances on biological and epidemiological systems. To this end, a combination of techniques from various fields have been employed, such as non-linear dynamical systems theory, stochastic processes, systems control, time series analysis, statistics and signal processing, as well as knowledge about biology, epidemiology and risk assessment. In addition to deriving modelling results, this combination of approaches allows to interpret their significance within scientific paradigms.

Below, the main findings and significance of previous chapters are briefly summarised and discussed, together with potential future directions of research.

9.1 Conceptual remarks

To begin with, let us consider the conceptual improvements brought about by the CT framework, as well as their relevance for specific disciplines.

Interdisciplinary outlook

The CT modelling framework focuses on abrupt regime shifts in dynamical complex systems. As such, it employs a variety of mathematical approaches to contribute to knowledge discovery and model-based predictions (see Chapters 1 and 4). In addition, the CT framework is inherently interdisciplinary, as it combines mathematical results with domain-specific knowledge. In fact, its baseline modelling concepts – e.g., dynamical attractors and noisy bifurcations – are, in principle, generically applicable to many systems satisfying certain basic conditions (Chapter 2). This explains the rapid adoption in many research fields, from climatology, to ecology, to psychology, epidemiology and so on. In addition to main chapters, the Appendices provide formal definitions and mathematical instruments to foster interdisciplinary and transversal collaborations.

Model development, as well as interpretation of derived results and applications, requires domain-informed inputs. This is particularly true for the flagship contribution of the CT framework: early warning signals against impending regime shifts (see Chapter 3; see also Sec. 2.1 for remarks on terminology). Although ideally generically applicable to a wide variety of noisy bifurcation-driven phenomena, they are measures that rest upon assumptions. As such, it is necessary to carefully verify EWS validity, in each system of interest, before applying, interpreting and routinely employing them.

Summarising what discussed in details in Chapters 2, 3 and 5, three defining elements should be combined: alternative regimes, EWS, and "dynamical context", *i.e.*, the mix of system's dimension and co-dimension, noise properties and reciprocal time scales. Given two of them, the third element can be inferred. Verifying a single element may not be sufficient to rule out spurious signals or to support modelling choices. Overall, EWS should be used carefully within risk assessment toolboxes, as they could be associated with spurious signals. They can be aptly used for detection or as "first pass tools" (Clements et al., 2018) to trigger scholars' attention towards specific systems and regimes. In these cases, false positives are not problematic if compared with complementary signals, but the risk of false negatives should be minimised using combinations of approaches.

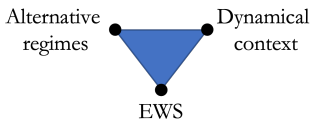


Fig. 9.1 A simplex is formed by the presence of alternative regimes (*e.g.* due to bistability), resilience indicators (EWS) and dynamical context (the combination of system's dimension and co-dimension, noise properties and reciprocal time scales; sometimes they constitute modelling assumptions, other times they can be assessed from data).

Deterministic vs stochastic changes

A conceptual benefit of the CT framework is to put determinism back into perspective. Stochastic processes have been long studied for their role in overcoming potential barriers and driving systems onto alternative attractors. In systems biology in particular, the Waddington's landscape, a long-standing illustrative paradigm, is implicitly based on the assumption that a cell is "bound" to evolve under some potential action, but that its fate choices are primarily driven by noise (*cf.* Sec. 2.6). On the contrary, introducing the effect of dynamical bifurcations, driven by control parameters, adds a layer to the picture. The potential is not static any more, but gets modified by the action of deterministic changes. The interplay between deterministic resilience and noise-induced shifts makes up for a richer modelling framework to accommodate and describe real-world phenomena.

Making this paradigm shift explicit in this thesis contributes to the sparse but insisting calls in systems biology communities to go beyond the Waddington's landscape picture, looking for alternative or complementary modelling tools that could spur new testable predictions and methods. Thereby, time series analysis strategies like those presented in Chapter 4 and Appendix A, as well as additional derived tools, may be employed in biology and other disciplines towards better systems' understanding, prediction and control.

9.2 Summary of main results

This thesis develops, refines and tests warning indicators against critical transitions. Specific attention is dedicated to EWS derived from the theory of noisy bifurcations. Old and new analytical and data-analysis methods have been combined and employed, see Chapter 4, not only to propose new indicators, but also to verify their validity and to assess their performance on real-world data. In addition to general theoretical settings, two domain-specific case studies have been considered: biological regulatory circuits and epidemiological dynamics.

In Chapter 5, the observability of EWS is systematically analysed, depending on theoretical modelling assumptions . The analysis demonstrates that trends of main indicators may be altered or disrupted if the time-scale separation assumptions breaks down, or if noise is non-white. In the first case, variance-based indicators are particularly robust, while autocorrelation-based indicators should be preferred in the second case. When noise types are mixed, which is common in realistic settings, combining indicators – instead of using single ones – optimises the performance. These findings thus guide the application and interpretation of resilience indicators for detection or anticipation of regime shifts. This allows interpreting experimental findings and informs subsequent studies, thereby supporting the "cycle for knowledge discovery" (Fig. 1.1).

Chapter 6 studies the ability of cell regulatory mechanisms to hamper fluctuations and noise-induced transitions. This analysis constitutes a system-specific verification of EWS generality when certain modelling assumptions are loosened and quantifies EWS performance in detecting impending cell-function decisions from distribution data. The findings provide novel insights onto control mechanisms in cellular processes, as cooperativity is shown to concur in fine-tuning system resilience close to criticality. The analysis provides quantitative information to experimental observations about bistable systems' resilience (Dai et al., 2015) and paves the way for systematic monitoring of cell states, towards desired functions or away from harmful ones.

Chapter 7 focuses on testing theoretical predictions and analysis methods on real-world epidemiological data, and on assessing the corresponding performance of EWS. Worldwide data from the COVID-19 pandemic are analysed, characterised by different dynamical regimes (rate of approach of control parameter $R(t)$ to its critical value 1 and noise distribution). Such regimes influenced the ability of CT-derived warning signals to consistently detect and anticipate impending epidemic waves. Furthermore, EWS performance was shown to be sensitive to data processing like detrending methods. These findings verify the theoretical prediction from literature and from previous chapters, and provide useful information about their applicability to real-world monitoring programs.

Chapter 8 provides a model-driven alternative to epidemic monitoring, using complementary data about COVID-19 diffusion within communities – SARS-CoV-2 RNA abundance in wastewater samples. The approach uses an extended Kalman filter, coupled with epidemiological modelling, to reconstruct patterns of epidemic diffusion and to perform causal short-term predictions and long-term projections. For many countries, the pipeline work well and outperforms statistical-based methodologies in complementing individual testing campaigns, estimating quan-

titative future epidemic developments and computing the early warning provided by wastewater data. The study provides several lessons for real-time monitoring of complex systems and suggests a proof-of-concept for the use of Kalman filters for risk assessment of critical phenomena.

9.3 Future perspectives

Studies of complex systems, and of critical transitions in particular, call for many future investigations. Several open research questions are discussed in Chapter 3; as many were triggered during the PhD thesis. Each Chapter already contains limitations and perspectives of specific studies. In this section, a selection of broad and inspiring future directions is presented and discussed.

Equilibrium and non-equilibrium

Abrupt regime shifts occur over different time-scales and dynamical contexts. Further including non-equilibrium and non-quasi-steady-state modelling will provide deeper capabilities to understand natural phenomena and to develop more refined and performing EWS. Additional insights will come from considering exotic noise-related effects and their interplay with non-equilibrium phenomena. These studies would go beyond what presented and discussed in Chapters 5 and 7, and would embrace other CT classes described in Sec. 2.3.2. Mathematical advancements, physical understanding and model-based predictions are necessary; identifying natural phenomena and experimental setups to test theoretical results is also crucial.

Data-driven inference of CT classes

Often, there is little consensus about the driving mechanisms of critical transitions (*cf.* Sec. 3.3). Bury et al. (2021) recently started developing a Machine Learning-based classification of diverse bifurcations (within b-tipping), with the goal of comparing EWS performance. Instead, inferring the class of critical transition (Sec. 2.3.2) with data-driven methods is still unexplored. Distinguishing between b-tipping and n-tipping (to begin with) is crucial not only to make sense of data and foster better models, but also to infer whether certain transitions are mostly driven by deterministic or stochastic factors. This would have immediate consequences in the correct interpretation of *e.g.* cell fate decision or glacial shifts (Ditlevsen et al., 2010). Developing such tools would thus be a valuable asset.

Deep learning classifiers, based on suitable network architectures, could do the job. Preliminary training and validation data would come from simulated time series of normal forms and known models, undergoing various bifurcations and/or noise-induced transitions, in different dynamical contexts. They would provide a sufficiently large library of examples to develop proof-of-concept studies. Normal forms are blueprints of diverse transitions, irrespective of full mechanistic models. Subsequently, the proof of concept could, in principle, be tested on real-world data like atrial fibrillation or epilepsy time series, which are present in our institute.

Distinguishing leading bifurcations in cell differentiation processes

Aforementioned developments in non-linear dynamics theory find domain-specific applications. Congruously with the rest of the thesis, let us focus on systems biology first.

Determining which dynamical process leads differentiation routes in various cell lineages is important to develop better models, to test existing ones and, in general, to better understand such key developmental processes (Moris et al., 2016). Having refined and precise models allow developing gene expression dynamics inference methods (Eugenio et al., 2014; Aalto et al., 2020) to reveal the epigenetic landscape (Li et al., 2015) and to foster better comprehension of the underlying bistability mechanisms (Pomerening, 2008).

Assessing which type of transition drive differentiation processes close to criticality is a delicate issue. Possibly depending on the cell type and developmental stage, the process might be primarily driven by deterministic mechanisms (following a b-tipping scheme), or by noise and other disturbances (n-tipping), or even by rate-dependent mechanisms (r-tipping), see Sec.2.3.2. In addition, transitions might be smooth or abrupt, with different routes towards the alternative stable state and with different associated EWS. When deterministic routes are known, preliminary steps may test which bifurcation better represents the differentiating dynamics. A starting point is to fit their predicted probability density functions (see Sec. 4.3.1 and Chapter 6) to experimental data and to contrast the corresponding goodness of fit. One readily available dataset to test is provided by Eugenio et al. (2014); others may also be considered.

Online risk assessment

In addition to knowledge discovery, the CT field aims at developing reliable warnings against impending shifts. However, EWS still need advances and refinements before being applied. Several thesis chapters elucidated their promises and limitations, including data collection protocols, consensus on data analysis methods (detrending schemes, moving windows and so on) and compliance with the "detection simplex" (Fig 9.1). For both low-dimensional and high-dimensional systems, future studies will confirm whether online (*i.e*, real time) risk assessment using EWS upon real-world problems (or, a least, on well-defined subsets) is a matter of refinements and theoretical developments, or wishful thinking (Ditlevsen et al., 2010).

Additional strategies may complement or supplement existing ones. One such example is topological analysis of embedding state spaces; an application on financial crashes has been carried by Gidea et al. (2018); Gidea et al. (2020). Alternatively, one can exploit the recently hypothesised link between singularities and observability properties of complex systems (Montanari et al., 2022a), and develop consistent indicators like SVD minimum coefficients, after time series embedding.

New indicators can also be tested on time series data, after changing basis. An interesting case is the wavelet entropy. It was first used, among other features, to detect transitions to seizure activity, and looked promising (*cf.* thesis "Dynamical modeling techniques for biological time series data"). Given a signal s and its decomposition $(s_i)_i$ in an orthonormal basis, an entropy measure can be defined as a functional satisfying the additivity property. The orthonormal basis can be constructed using Fourier decomposition or, instead of working in the frequency domain only, the wavelet transform (Burrus, 1997). A corresponding (non-normalised) entropy measure is thus defined as (Abdel-Hamid et al., 2016; Coifman et al., 1992):

$$E_w = -\sum_i s_i^2 \log s_i^2 \,, \tag{9.1}$$

with the convention $0 \log 0 = 0$. This is often called a "Shannon"-type entropy since it has the same properties of distance between the signal and the decomposition, and because we can interpret $||P_H s||^2 = s_i^2$ as a "quantum-mechanics"-like probability (Coifman et al., 1992). Here, P_H is an orthogonal decomposition operator to the subspace H. For instance, the Poisson wavelet is used to reproduce the Fourier power spectrum (Kirby, 2005) and enhance interpretability. This or alike indicators may foster new avenues, when frequent and abundant data are available.

Machine learning (ML) techniques may be potential allies. Bury et al. (2021) started developing generic proof-of-concept algorithms, but applications on real-world data are so far system-specific, *e.g* on epilepsy (Rasheed et al., 2020; Usman et al., 2017; Mombaerts, 2019 (thesis)) or atrial fibrillation (Ebrahimzadeh et al., 2018; Hill et al., 2019; Olier et al., 2021; Gavidia, 2022). They also require abundant data, which are seldom available. Nonetheless, ML algorithms may extract features related to approximations of order greater than one (hence, beyond O-U processes) to improve performances. Alternatively, theory-based indicators may guide feature selection in ML models, see Chapter 5 and Brett et al. (2020b), enhancing their interpretability.

An appealing alternative for online risk assessment could come from extended Kalman filters, coupled with non-linear models (when available) or with generic normal forms. They may provide reasonable predictions and contrasting scenarios, linking mathematical models and noisy measurements. However, normal forms are only valid locally, in the vicinity of critical points. Hence, tradeoffs between recognising impending bifurcations and refining the analysis with EKF-like methods are still necessary.

Critical transitions in multi-dimensional systems

So far, low-dimensional systems have been primarily considered. High-dimensional ones are nonetheless of paramount importance, as many real-world systems have been aptly modelled as networks of various orders and multi-varied data are becoming increasingly available.

Studying critical transitions on networks is current ongoing work. Some research focuses on graph properties like node removal or centrality (Morone et al., 2019; Loppini et al., 2019; Moutsinas et al., 2020); others consider the dynamical effects of evolving link parameters (Meng et al., 2020; Bhandary et al., 2021). Future works within network theory might unravel other general mechanisms, like those introduced by Gao et al. (2016) and discussed in Sec. 4.3.6 and Sec. 4.7. For instance, Kuehn et al. (2015) proposed methods to obtain the explicit centre manifold A.5.1 on simple network systems escaping through a saddle node. Looking for early warning signals in networks is another open research direction (see again Sec. 4.3.6 and Sec. 4.7, as well as Rodríguez-Méndez et al. (2016); Chen et al. (2019); Matsumori et al. (2019)).

Complementary researches focus on conditions for certain transition classes to happen in high dimensional systems. Kiers et al. (2019) discussed conditions for r-tipping, while looking for evidences of multistability and alternative states in high-dimensional data sets is under development (*cf.* Sec. 3.1). Extending necessary and sufficient conditions for bistability to higher-order networks and simplicial complexes (Bick et al., 2021) is another intriguing idea, as

well as checking for the possibility of critical transitions and associated warning patterns. As of now, no publication on this direction is known.

Spatially extended systems are also of great interest. The study of critical transitions and EWS in PDEs counts a few works (Gowda et al., 2015; Kuehn et al., 2019). Moreover, understanding transitions and warning indicators in systems that stretch over wide areas is of environmental and ecological relevance. Although the problem is known and recognised, only recently quantitative measures are being developed, including the use Betti numbers to characterise topological connectedness (Storch et al., 2019a). ML-based image segmentation measures (Taha et al., 2015) like Dice-score or Jaccard-index may potentially add value in providing univariate measures of spatial changes. Alternatively, one can reconstruct the state space using embedding techniques (see Sec. 4.3.6) and then predict new time points (Ma et al., 2014) or detect topological changes using *e.g* persistent homology (Zomorodian et al., 2005; Gidea et al., 2020) or observability properties (Aparicio et al., 2021; Montanari et al., 2022a).

Bridging theoretical and computational studies with real-world data is a necessary next step. A recent work on this direction applies dimension reduction techniques (Sec. 4.7) on behavioural data from the COVID-19 pandemic (Fan et al., 2020), to characterise rapid changes in people's perception and compliance to public rules, in turn modelled as correlation networks. Many other studies within different domains are nonetheless required.

Cascading regime shifts

As complex systems evolve on different time-scales and a different levels, it is immediate to ask whether tipping on short timescales could be connected to, or restricted or enslaved by tipping on long timescales. In literature, investigations on this direction are still limited (Rocha et al., 2018; Zhong et al., 2019; Ashwin et al., 2017a). One possible way of developing related theories could consider one "fast" variable that is tipping as the "slow variable" of another system (i.e. its parameter). So, the tipping on one system might produce large pushes on the second. If that is sufficiently large, a cascade effect occurs. On the contrary, one may consider similar time scales and consecutive tipping due to network effects. Dedicated studies are recommended in this promising research direction, to develop better understanding, derive warning indicators (if any) and identify natural systems that may be prone to tipping across scales.

System resilience

In general, many studies dedicated to critical transitions are dual to those addressing systems' resilience, a long-recognised property of complex systems and networks (Liu et al., 2020a). Many new avenues exist. Improving theoretical definitions, conceptualization and analysis is a necessary first step. Then, there lies the problem of uncovering resilience properties from time series data, which is particularly complicated when mechanistic models and parametrizations are not available. Indicators for loss of resilience and related phenomena should therefore be continuously developed and tested, even beyond the CSD umbrella.

Another crucial step is bridging the gap between theoretical studies and empirical testing on real-world systems. This can be done by refining the classification of system properties and

dynamical contexts and by improving modelling and data-driven methods to infer the class of empirical systems. In addition, methods to cope with incomplete information – hence, towards the engineering side of the CT framework – are recommended to identify the functional observability space (Montanari et al., 2022b).

Finally, after understanding and prediction, it comes the control of tipping-prone systems, in the identification of safe operating spaces and in the apt steering of their behaviour towards desired transitions or away from dangerous ones.

9.4 Conclusion

Critical transitions in complex dynamical systems are a fresh topic of great relevance for scientific research and practical applications. It has connections with various fields and strives to address many open questions. It brings many results for knowledge discovery and provides novel insights upon systems' variability in critical regimes. In addition to improving theoretical results, this thesis analysed and verified the ideas and hypothesis about application of derived indicators within monitoring systems.

Based on theory and analysis of empirical data, the current results assessed that such application should be done carefully to avoid spurious signals, and should be potentially coupled with model-based methods. For offline detection, EWS have the potential to complement other analysis methods; for online prediction, though, they still require refinements, in particular when critical points are not known at all. In both cases, knowledge about the dynamical context is crucial, to properly frame the CT class that is being observed.

Overall, this thesis provides solid foundations for future studies on critical transitions – both from a theoretical and an applied side. As a wise man said[1], "There are more things on heaven and earth, Horatio, than are dreamt of in your philosophy". Many exciting things lie ahead.

Code availability

The code to reproduce the figures not associated to specific papers is available at https://github.com/daniele-proverbio/thesis.

[1] W. Shakespeare. *Hamlet*, 1603.

Appendix A

Formal Definitions

A.1 Dynamical systems and notable features

This subsection provides formal definitions for dynamical systems and notable features introduced in Sec. 2.2.1. Most of definitions and theorems come from and are detailed in Strogatz, 2018; Kuznetsov, 2013; Golubitsky et al., 2012; Berglund et al., 2006; Thompson et al., 2011; Dobrushkin, 2014. A dynamical systems is defined as follows:

Definition A.1.1. *A **dynamical system** is a triple $\{\Omega, \phi^t, \mathscr{T}\}$, where $\mathscr{T}$ is a time set, Ω is a (Banach) state (or phase) space with a properly defined metrics on it and $\phi^t : \Omega \to \Omega$ is a family of evolution operators parametrized by $t \in \mathscr{T}$ and satisfying:*

i. $\phi^0 = id$

ii. $\phi^{t+s} = \phi^t \circ \phi^s$ *($\circ$ is the convolution operator).*

Definition A.1.2. *An **orbit** (or **trajectory**) starting at u_0 is an ordered subset of the state space Ω,*

$$Or(u_0) = \{u \in \Omega : u = \phi^t u_0, \forall t \in \mathscr{T} \text{ such that } \phi^t u_0 \text{ is defined}\}$$

Definition A.1.3. *The **phase portrait** of a dynamical system is a partitioning of the state space into orbits.*

For analysis purposes, one can explicit a dynamical system as a set of differential equations where algebraic operations are well defined. In fact, the main result of the following theorem is that a system of ordinary differential equations can be regarded as a dynamical system[1].

Theorem A.1.1. *Consider a system of **ordinary differential equations***

$$\dot{u} = f(u, p, t), \quad u \in \mathbb{R}^n, \quad p \in \mathbb{R}^m \tag{A.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is smooth in an open region $U \subset \mathbb{R}^n$. Then there is a unique function $u = u(t, u_0)$, $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ that is smooth in (t, u_0) and satisfies, for each $u_o \in U$, the following conditions:

[1] Another way to represent dynamical systems whose time is discretized are *maps*. However, they are not of primary interest in this thesis.

i. $u(0,u_0) = u_0$

ii. there is an interval $\mathscr{I} = (-r_1(u_0), r_2(u_0))$*, where* $r_{1,2} > 0$*, such that, for all* $t \in \mathscr{I}$*,*

$$y(t) = u(t,u_0) \in U$$

and

$$\dot{y} = f(y(t))';.$$

The function $u(t,u_0)$ *produces a* ***solution curve***

$$Cr(u_0) = \{(t,u) : u = u(t,u_0), t \in \mathscr{I}\} \subset \mathbb{R} \times \mathbb{R}^n$$

and an orbit

$$Or(u_0) = \{u : u = u(t,u_0), t \in \mathscr{I}\} \subset \mathbb{R}^n\,.$$

Note that the orbit is the projection of the solution curve onto the state space $\mathbb{R}^n$*. The evolution operator* $\phi^t : \mathbb{R} \to \mathbb{R}$ *can thus be written*

$$\phi^t u_0 = u(t,u_0)$$

As a consequence, we can regard the system of ordinary differential equations like Eq. A.1 as a dynamical system.

When ODE A.1 does explicitly depend on time, it is termed *nonautonomous*; when it does not, it is termed *autonomous*. u are state variables that univocally define the state of the system. n is the system *dimension*, namely the minimum number of state variables (in physics, it is also called the "number of degrees of freedom"). m represents the number of free parameters that specify the evolution function. The function(s) $f(u,p,t)$ can depend linearly or non-linearly to its variables.

A.2 Slow-fast systems

Many systems of non-linear ODEs are characterized by different time scales. When such time scales are well-defined mathematically, or when we can assume such separation by observing the desired phenomenon, the system can be written as a *slow-fast ODE* in the form:

$$\varepsilon \frac{dx}{dt} = f(x,y) \tag{A.2}$$

$$\frac{dy}{dt} = g(x,y)\,, \tag{A.3}$$

where $0 < \varepsilon \ll 1$ is a small parameter. x contains fast degrees of freedom and y the slow ones. Alternatively, one can write the system in terms of fast time $s = t/\varepsilon$ as

$$\frac{dx_s}{ds} = f(x_s,y_s) \tag{A.4}$$

$$\frac{dy_s}{ds} = \varepsilon g(x_s,y_s)\,. \tag{A.5}$$

There are different methods to verify the time scale separation in models, the most notable ones involve spectral theory and large deviation theory (Berglund et al., 2006). Empirically, one can consider the rates (or times) in which a process and its coupled ones occur and compare e.g. their orders of magnitude.

Definition A.2.1. *A **slow manifold** is a set of equilibria onto which the system is projected if an asymptotically stable equilibrium $\hat{x}(y)$ is admitted for each quasi-steady-state y.*

By doing so, the complexity of the system is largely reduced and analytical results can be more easily obtained (Strogatz, 2018). As a matter of fact, as y is almost fixed, one can interpret it a the usual slowly varying parameter p.

A.3 Equilibria, cycles, invariant sets

Definition A.3.1. *A point $\hat{u} \in \Omega$ is called an **equilibrium point** (fixed point) if $\phi^t \hat{u} = \hat{u}$ for all $t \in \mathscr{T}$.*

Note a slight abuse of wording when referring to "equilibrium". In fact, this mathematical term refers to the physical notion of "steady-state", given by the no-change condition $\frac{du}{dt} = 0$ in the associated ODE. It is different from the mechanical definition of equilibrium as balance of forces acting on the system $\sum_i F_i = 0$ or the thermodynamic requirement of absence of net flux of energy though the system $\sum_i J_i = 0$ (adiabatic regime).

Definition A.3.2. *A **cycle** is a periodic orbit, that is, a nonequilibrium orbit Cy such that each point $\tilde{u} \in Cy$ satisfies $\phi^{t+\tau}\tilde{u} = \phi^t \tilde{u}$ for all $t \in \mathscr{T}$. The minimal $\tau > 0$ with this property is called the period of the cycle.*

Definition A.3.3. *A cycle of a continuous-time dynamical system, in a neighbourhood of which there are no other cycles, is called a limit cycle.*

Definition A.3.4. *An **invariant set** of a dynamical system is a subset $\mathscr{S} \subset \Omega$ such that $\tilde{u} \in \mathscr{S}$ implies $\phi^t \tilde{u} \in \mathscr{S}$ for all $t \in \mathscr{T}$.*

An equilibrium point is an invariant set, so is a cycle.

Definition A.3.5. *An equilibrium $\hat{u}$ is called **hyperbolic** (or generic) if there are no eigenvalues of $\partial_u f(u)|_{\hat{u}}$ on the imaginary axis. Otherwise, the equilibrium is called **elliptic**.*

A.4 Stability

We get now into the definition of stability, of primary importance to assess how sensitive a system is to small perturbations.

Theorem A.4.1. *An invariant set $\mathscr{S}$ is stable if the following conditions hold:*

i. for any sufficient small neighbourhood $U \supset \mathscr{S}$ there exist a neighbourhood $V \supset \mathscr{S}$ such that $\phi^t u \in U$ for all $u \in V$ and all $t > 0$;

ii. there exist a neighbourhood $U_0 \supset \mathscr{S}$ such that $\phi^t u \to \mathscr{S}$ for all $u \in U_0$, as $t \to +\infty$.

Condition (i) is often called "Lyapunov stability", whereas condition (ii) is referred to as "asymptotic stability". A stable invariant set is often called an **attractor**. Fixed points and limit cycles are classified as (linearly) stable or unstable depending on the eigenvalues of their associated Jacobian.

Theorem A.4.2. *Consider a dynamical system A.1, where f is smooth. Suppose that it has a fixed point $\hat{u}$. Then $\hat{u}$ is (linearly) stable if all eiganvalues $\lambda_1, ..., \lambda_n$ of $\partial_u f(\hat{u})$ satisfy* $\mathrm{Re}(\lambda) < 0$.

Extracting stability information from determinant, discriminant and trace of the Jacobian is described in Sec. 2.2.1, along with their graphical representation with a Poincaré Diagram, Fig 2.3.

A.5 Bifurcations

Mathematically speaking, when the behaviour of a system changes, we encounter a bifurcation. Hence, it is necessary to formalize the notion of "behaviour" and that of "change" in terms of the elements presented so far. Note that, once the phase space is defined by the state variables, its properties are then given by the parameters introduced in Eq. A.1.

Definition A.5.1. *A dynamical system $\{\mathscr{T}, \mathbb{R}^n, \phi^t\}$ is called **topologically equivalent** to a dynamical system $\{\mathscr{T}, \mathbb{R}, \psi^t\}$ if there exist a homeomorphism $h : \mathbb{R}^n \to \mathbb{R}^n$ mapping orbits of the first system to orbits of the second system while preserving the direction of time.*

Definition A.5.2. *A dynamical system $\{\mathscr{T}, \mathbb{R}^n, \phi^t\}$ is called **locally topologically equivalent** near an equilibrium $\hat{u}$ to a dynamical system $\{\mathscr{T}, \mathbb{R}, \psi^t\}$ near an equilibrium $\hat{y}$ if there exist a homeomorphism $h : \mathbb{R}^n \to \mathbb{R}^n$ that:*

i. is defined in a small neighborhood $U \in \mathbb{R}^n$ of $\hat{u}$;

ii. satisfies $\hat{y} = h(\hat{u})$;

iii. maps orbits of the first system in U onto orbits of the second system in $V = h(U) \subset \mathbb{R}^n$ preserving the direction of time.

Definition A.5.3. *The appearance of a topologically non-equivalent phase portrait under variation of parameters is called a **bifurcation**.*

Definition A.5.4. *The **codimension** of a bifurcation system A.1 is the difference between the dimension of the parameter space and the dimension of the corresponding bifurcation boundary.*

Definition A.5.5. *A **bifurcation diagram** of the dynamical system is a stratification of its parameter space induced by the topological equivalence, together with representative phase portraits for each stratum.*

Definition A.5.6. *The **normal form** of a dynamical system is a simplified form (in scalar system, a function) that can be useful in determining the system's behaviour. All systems exhibiting a certain type of bifurcation are locally (around the equilibrium) topologically equivalent to the normal form of the bifurcation.*

Formally speaking (Kuznetsov, 2013), consider a dynamical system

$$\dot{x} = f(x, p') \ , x \in \mathbb{R}^n \ , p' \in \mathbb{R}^n \tag{A.6}$$

and a polynomial system

$$\dot{\zeta} = g(\zeta, p; \beta) \ , \zeta \in \mathbb{R}^n \ , p \in \mathbb{R}^k \ , \beta \in \mathbb{R}^l \tag{A.7}$$

having dimension n, codimension k and polynomial order l. Without loss of generality, we can assume that the tipping point occurs at $(x, p) = (0, p_0)$. A translation would do it. Then:

Definition A.5.7. *System A.7 is called a* ***topological normal form*** *for the bifurcation if any generic system A.6 with the equilibrium $x = 0$ satisfying the same bifurcation conditions at $p' = 0$ is locally topologically equivalent near the origin to A.7 for some values of the coefficients β_i.*

Here, *generic* means that the system satisfies certain conditions around the critical point: $\frac{\partial^j f}{\partial \varphi^j}|_{(0,p_0)}$, where j is the derivate order and $\varphi = \{x, p\}$. In particular, $\frac{\partial^j f}{\partial x^j}$ are called *nondegeneracy conditions* and are related to the "criticality" of a bifurcation (Kuehn, 2011), while $\frac{\partial^j f}{\partial p^j}$ are called *trasversality conditions* and govern the bifurcation unfolding (see below) and thus its genericity.

Bifurcations can be divided into two principal classes:

- Local bifurcations, which can be analysed through changes in the local stability properties of equilibria, periodic orbits or other invariant sets as parameters cross through critical thresholds;
- Global bifurcations, that cannot be detected purely by a stability analysis of the invariant set. They are harder to detect because they involve large regions of the phase plane rather than just the neighbourhood of a single fixed point.

One important concept to reduce dynamical systems to local, low-dimensional normal forms is the **center manifold** (Haragus et al., 2010; Crawford, 1991). According to the linearized systems and its stability properties (see above), a dynamical system has invariant manifolds of which one is the center manifold, having the same dimension of the stable manifold (attractor) and tangent to the center subspace at every equilibrium point (Guckenheimer et al., 2013). A center manifold corresponds to a slow manifold (Def. A.2.1) when the eigenvalues of the center subspace are all precisely zero, rather than just real part zero.

Theorem A.5.1. *The* ***center manifold emergence theorem*** *says that a neighbourhood around an equilibrium point may be chosen so that all solutions of the system staying in that neighbourhood tend exponentially quickly to some solution $y(t)$ on the center manifold.*

The theorem allows to reduce complex multi-dimensional dynamics to low dimension dynamics, locally described by bifurcation normal forms. This guarantees that, for stability analysis and related locally "qualitative" features, it is sufficient to study bifurcation normal forms and their properties, to which the system reduces and that represent universal (local) routes to phase changes. Generalising centre manifold theorems and techniques to non-autonomous systems (Pötzsche et al., 2006) unleashed, among the others, studies of rate-induced tipping points (Sec. 2.3.2).

Models are idealized assumptions of reality (Chapter 1). When proved wrong or inadequate, they can be improved by adding small terms that were initially neglected and that would model e.g. impurities or asymmetries in real-world systems.

Definition A.5.8. *A model obtained by adding small parameters to a given system is called an* ***unfolding*** *of the original system.*

Unfolding is nicely described in singularity theory (Golubitsky et al., 2012; Kuznetsov, 2013), to which the reader is referred for formal studies involving advanced topological structures whose definition is beyond the scopes of this appendix. Let us now consider those systems with finite codimension with respect to their topological equivalence, and let us limit our attention to the equilibria solutions. Hence, let us focus on bifurcation unfolding.

Definition A.5.9. *An* ***unfolding*** *of a dynamical system under static equivalence is one that exhibits all possible bifurcations of the equilibrium (rest) points, up to topological equivalence of the set of equilibria.*

In other terms, it investigates what happens when small terms are added to the original bifurcation, mimicking extra parameters, small offsets or "impurities".

Definition A.5.10. *A (local) bifurcation is said to be* ***generic*** *if it results from the unfolding of other bifurcations and does not result in other bifurcations if unfolded.*

A.6 Critical normal forms

As anticipated in Def. A.5.7, a normal form characterises a topological equivalence to bifurcations observed in complex systems. Hence, it locally retains the qualitative features (stability, sensitivity to small perturbations, etc.), in a more compact and analytically tractable form. In addition, according to the centre manifold theorem A.5.1, bifurcation normal forms are adequate descriptors of multi-dimensional dynamical systems near equilibrium points. In this regard, they have often been called "universal patterns" to critical and explosive phenomena (Thom et al., 1977; Kuehn et al., 2021).

Reducing a system to its (local) normal form may involve different techniques. direct and general center manifold calculation (Guckenheimer et al., 2013; Kuehn et al., 2021). A simpler technique is Taylor expansion around $x \sim x_c$ and $p \sim p_c$ up to second or third order (Strogatz, 2018):

$$\dot{x} = f(x,p) = f(x_c,p_c) + (x-x_c)\left.\frac{\partial f}{\partial x}\right|_{(x_c,p_c)} + (p-p_c)\left.\frac{\partial f}{\partial p}\right|_{(x_c,p_c)} + \frac{1}{2}(x-x_c)^2\left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_c,p_c)} + \dots, \tag{A.8}$$

where subscript c identifies "critical" values. However, this procedure is not always possible and it is often necessary to hypothesise the bifurcation to decide which order to truncate in the expansion. Graphical inspection might help crafting the hypothesis by checking how the diagram $(f(x,p),x)$ changes while varying p and comparing its local properties next to the x-axis with those of normal forms (reported below). An example is provided in Fig. A.1. It analyses a harvested population

model from Scheffer (2009):

$$\frac{dX}{dt} = X\left(1 - \frac{X}{K}\right) - c'\frac{X^2}{X^2+1}, \tag{A.9}$$

where X is population density, K is the carrying capacity and c' is the maximum harvest rate.

Another common option is the spectral analysis of the Jacobian (Sec. A) and the verification of specific conditions (listed below) defining each normal form (Strogatz, 2018). A web service curated by prof. A. J. Roberts (http://www.maths.adelaide.edu.au/anthony.roberts/sdenf.php) provides assistance by solving the computer algebra for a range of finite-dimensional systems.

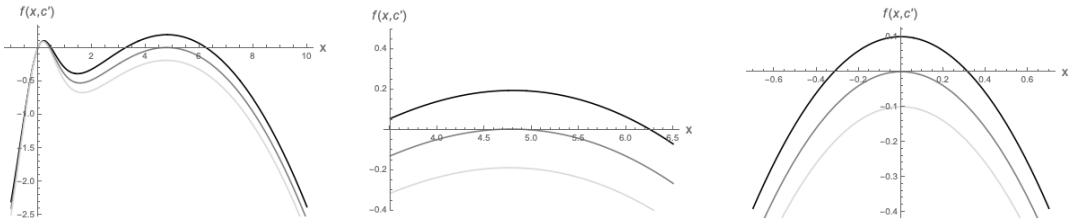


Fig. A.1 **Left**: graph representation of $f(x,c')$, at different values of c', for model A.9. **Centre:** zoom-in the the previous plot, around $x \simeq 0.47$, to compare with the next plot. **Right:** sketch of fold vector field (*cf.* Fig.A.2, right) for different values of c'. Compare it with the centre plot: when c' is varied, both graphs move from black to light gray. Although $(X_0, c_0) \neq (0,0)$ due to translations and specific quantitative properties of the realistic model, the qualitative behaviour around the bifurcation is the same.

The remaining of this section lists and briefly discusses the normal forms (Def. A.5.7) of the main low-dimensional bifurcations. Each normal form is presented as $\dot{x} = f(x,p)$ and accompanied by figures about $f(x,p)$ and its bifurcation diagram[2] $(\tilde{x},p)$. We also list the critical sowing down (CSD) scaling of each bifurcation, anticipated in Sec. 2.2 and further discussed in Sec .B. In addition to critical bifurcations, the supercritical pitchfork bifurcation has been added for its importance as a counterexample, for its historical connection to second-order phase transitions and for its potential relevance in the biological disciplines, *cf.* Sec. 2.6. Further reading in Kuznetsov (2013); Strogatz (2018).

A.6.1 Fold (saddle-node)

This is the basic mechanism by which fixed points are created and destroyed. A fold (or saddle-node, or "blue-sky") bifurcation is characterized by the following conditions for a dynamical system $f(x,p)$:

$$\begin{aligned} &\frac{\partial f}{\partial x}(0,p_0) = 0 \qquad \frac{\partial f}{\partial p}(0,p_0) \neq 0 \\ &\frac{\partial^2 f}{\partial x^2}(0,p_0) \neq 0 \end{aligned}$$

[2]Plots for bifurcation diagrams and vector fields are adapted from Suba Thomas, "Bifurcation Diagrams with Flow Fields" http://demonstrations.wolfram.com/BifurcationDiagramsWithFlowFields/, Wolfram Demonstrations Project, Published: March 7 2011.

The fold is a 1-dimension, 1-codimension bifurcation. Its normal form is:

$$\dot{x} = p + x^2 \,. \tag{A.10}$$

For $p < 0$, there are two equilibria $\hat{x}_{1,2} = \pm\sqrt{-p}$. $\hat{x}_1$ is stable and $\hat{x}_2$ is unstable. For $p > 0$ there are no equilibria in the system. When $p = 0$ the two equilibria collide and disappear, so $p_0 = 0$ is the critical value for the parameter.

Remark A.6.1. *The system $\dot{x} = p - x^2$ can be considered the same way, as well as any system mapped with $x \to -x$ and $p \to -p$.*

Lemma A.6.1. *Near the origin, the system $\dot{x} = p + x^2 + O(x^3)$ is locally topologically equivalent to the system $\dot{x} = p + x^2$.*

To visualize how $f(x,p)$ is varied under the action of p, see Fig. A.2 (left). The fold mechanism is given by a translation, driven by p, of $f(x,p)$ across the 0-axis, that destroys (or creates, depending on the direction) the equilibria. In addition, Fig. A.2 (centre) shows the evolution of the associated analytical potential $V(x,p)$ s.t. $f(x,p) = -\frac{\partial V}{\partial x}$ while p is varied. It it straightforward to associate it to the "pebble-down-the-hill" analogy (see Sec. 2.2.5). Fig. A.2 (right) shows the fold normal form bifurcation diagram, that displays the picture at once. For each value of the parameter there are different equilibria. Solid lines: stable equilibria; dashed lines: unstable equilibria. $(x,p) = (0,0)$ is the critical point.

CSD SCALING:

$$u' = \partial_x (p + x^2)|_{\hat{x}=\sqrt{-p}} u = -2\sqrt{-p}u = \mathscr{O}(p^{1/2})u \,. \tag{A.11}$$

Therefore the recovery exponent is $\alpha = 1/2$, that is, perturbations get recovered with rate proportional to the square root of the parameter. If $p \to 0$, the speed of recovery goes to zero as well, hence the "Slowing Down" phenomenon.

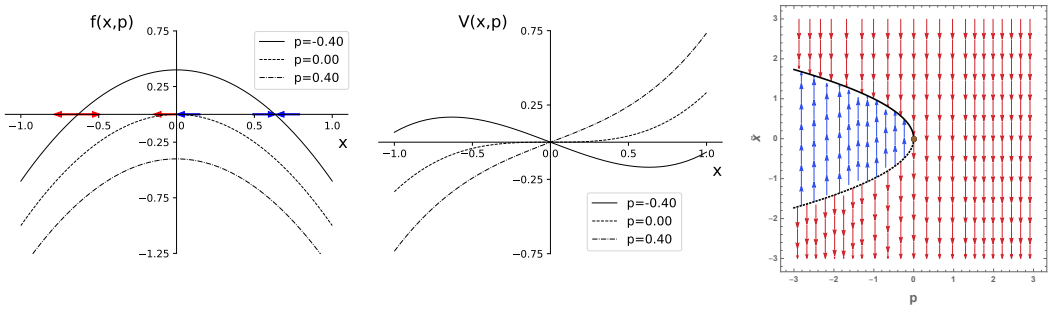


Fig. A.2 **Left:** evolution of $f(x,p)$ for a fold normal form, when p changes from negative to positive values. Red diverging arrows indicate the direction of instability; blue converging arrows that of stability. For $p = p_0 = 0$ a saddle node is created, with a stable and an unstable direction. **Centre:** quasi-steady state potential $V(x,p)$ for different values of p. For $p < 0$, a stable steady-state is present (valley); that becomes flat when $p = 0$ and unstable for $p > 0$. In the latter case, the "pebble" will roll down towards another, far away equilibrium. **Right:** bifurcation diagram of the fold normal form A.10. The equilibrium manifold is in black (solid line: stable; dashed: unstable).Vector field lines indicate whether the system is pushed upwards or downwards.

A.6.2 Transcritical

There are certain situations when a fixed point always exist (e.g. populations with zero individuals will never grow even thought the growth rate increases); however, such equilibrium changes stability. Transcritical bifurcations does the job by following the conditions:

$$\begin{array}{ll} \frac{\partial f}{\partial x}(0,p_0)=0 & \frac{\partial^2 f}{\partial x \partial p}(0,p_0)\neq 0 \\ \frac{\partial^2 f}{\partial x^2}(0,p_0)\neq 0 & \frac{\partial f}{\partial p}(0,p_0)=0 \end{array}$$

The associated normal form is:

$$\dot{x} = px - x^2 \,. \tag{A.12}$$

It has one fixed equilibrium at $\hat{x}_1 = 0$ that switches stability according to the sign of p, while the other equilibrium is $\hat{x}_2 = p$ and changes stability in reverse order as compared to $\hat{x}_1$. The evolution of $f(x,p)$ with respect to the parameter looks as in Fig. A.3. The bifurcation diagram is depicted in Fig. A.4.

CSD SCALING:

$$u' = \partial_x(px - x^2)|_{\hat{x}=0} u = \mathscr{O}(p)u \;,\; \alpha = 1 \,. \tag{A.13}$$

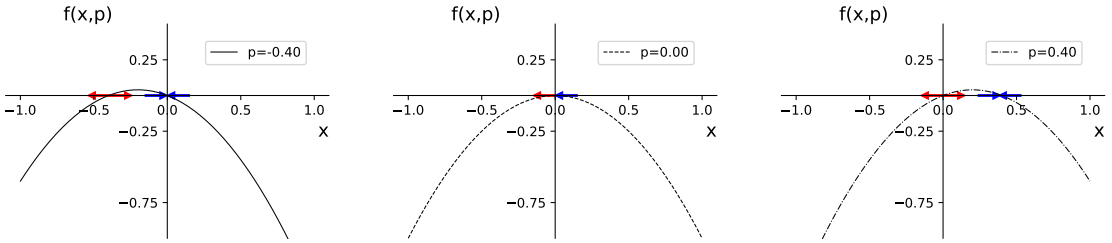


Fig. A.3 Evolution of $f(x,p)$ for a transcritical normal form when p changes from negative to positive values. $f(x,p)$ moves in a tilting fashion. Red diverging arrows indicate the direction of instability; blue converging arrows that of stability. $p = 0$ is the critical value.

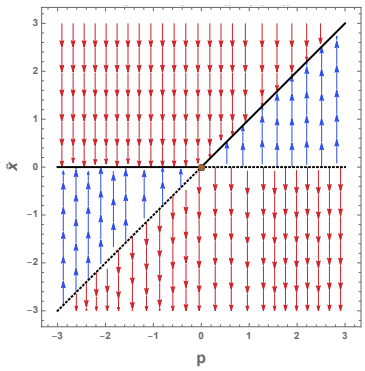


Fig. A.4 Bifurcation diagram of the transcritical normal form A.12. The equilibrium manifold is in black (solid line: stable; dashed: unstable).Vector field lines indicate whether the system is pushed upwards or downwards.

A.6.3 Subcritical pitchfork

This is common in problems where there is symmetry (e.g. spatial symmetry) so that the system can choose in which equilibrium to converge. Generic conditions are:

$$-f(x,p) = f(-x,p)\,(\text{symmetry}) \qquad \frac{\partial^2 f}{\partial x \partial p}(0,p_0) \neq 0$$
$$\frac{\partial f}{\partial x}(0,p_0) = \frac{\partial^2 f}{\partial x^2}(0,p_0) = \frac{\partial f}{\partial p}(0,p_0) = 0 \qquad \frac{\partial^3 f}{\partial x^3}(0,p_0) > 0$$

Its normal form is:

$$\dot{x} = px + x^3\,. \tag{A.14}$$

It has three equilibria for $p < 0$, of which $\tilde{x}_3 = 0$ is stable and two symmetrical $\tilde{x}_{1,2} = \pm\sqrt{-p}$ are unstable. After p_0, the unstable points collide with the stable one and vanish, while $\tilde{x}_3$ becomes unstable. Thus, trajectories diverge to infinity. The evolution of $f(x,p)$ is depicted in Fig. A.5. Fig. A.6 shows its bifurcation diagram.

CSD SCALING

$$u' = \partial_x(px + x^3)|_{\hat{x}=0} u = \mathscr{O}(p)u\;,\; \alpha = 1\,. \tag{A.15}$$

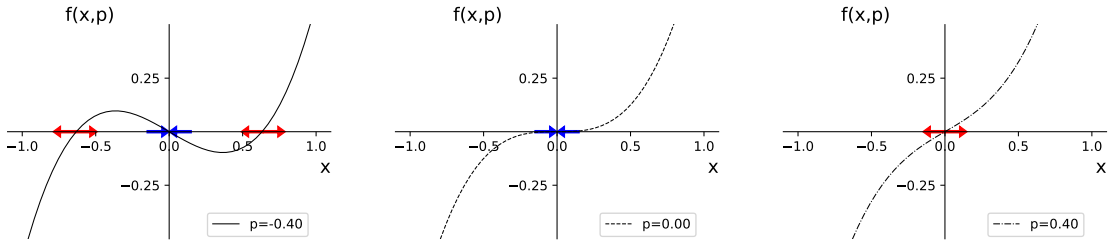


Fig. A.5 Evolution of $f(x,p)$ for a subcritical pitchfork normal form when p changes from negative to positive values. Red diverging arrows indicate the direction of instability; blue converging arrows that of stability. $p = 0$ is the critical value.

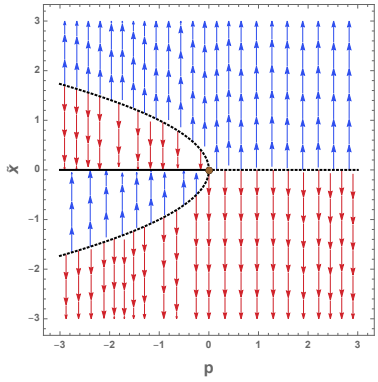


Fig. A.6 Bifurcation diagram of the subcritical pitchfork normal form A.14. The equilibrium manifold is in black (solid line: stable; dashed: unstable).Vector field lines indicate whether the system is pushed upwards or downwards.

When zooming out the purely local behaviour of real systems, the explosive instability is often counterbalanced by the stabilizing influence of higher order terms. To respect the symmetry conditions, $\mathscr{O}(x^5)$ is then inserted in Eq. A.14 to obtain the canonical form of a global subcritical

pitchfork:

$$\dot{x} = px + x^3 - x^5 \,. \tag{A.16}$$

Coefficients of higher order terms can be considered $= 1$ after proper non-dimensionalization. Fig. A.7 (left) displays $f(x,p)$ for different values of p and unravels local mechanisms analogue to systems A.14 (close to $x = 0$) and A.10 (close to $x \sim \pm 0.8$). The bifurcation diagram of system A.16 is shown in Fig. A.7 (right) after numerical estimation. A subcritical pitchfork bifurcation is a key mechanism to produce tri-stability (co-existence of three stable points) and sharp transitions between attractors. Fig. A.7 (right) shows hysteresis as effect of non-reversibility of the process. Imagine starting at $(p,\tilde{x}) = (-1,0)$ and increasing the parameter. When $p = p_0 = 0$ the system jumps on the upper or lower branch, mostly depending on random fluctuations. Then, to go back to its original state, it is necessary to decrease the parameter values *below* $p = 0$, finally getting to a fold point and having a second abrupt shift towards the origin. The evolution of its associated potential is reported in Fig. A.7 (centre). We can observe a vanishing tri-stability when $p = 0$, which then settles into bistability. That, however, it reached in a critical manner, not smoothly as for a supercritical pitchfork bifurcation (see below).

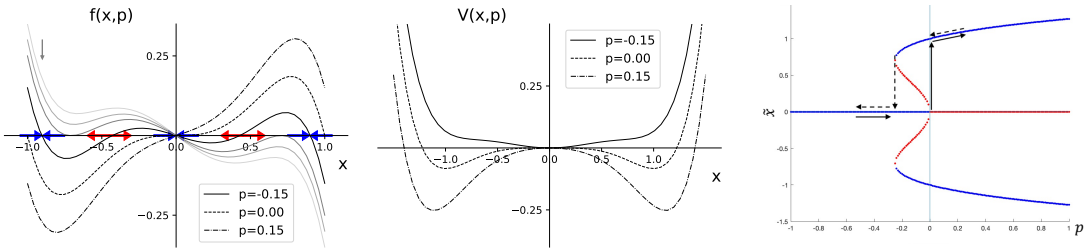


Fig. A.7 **Left:** evolution of $f(x,p)$ for a global subcritical pitchfork form, when p changes from negative to positive values. Red diverging arrows indicate the direction of instability; blue converging arrows that of stability, in the case of $p = -0.15$. Close to $x = 0$, the behaviour is the same as in Fig. A.6. Further from it, the local behaviour is topologically equivalent to that of a saddle-node. To observe that, follow the light grey plots along the direction pointed by the arrow. Close to $x = -0.8$ and $x = 0.8$, the "hunchbacks" approach the x-axis with a similar mechanism to that in Fig. A.2. They correspond to the local fold bifurcations marked in the bifurcation diagram. **Centre:** quasi-steady state potential $V(x,p)$ for different values of p. For $p < 0$, a stable steady-state is present (valley); that becomes flat when $p = 0$ (while other two stable states exist) and unstable for $p > 0$, where only the two extreme stable states reman. **Right:** bifurcation diagram of the global subcritical pitchfork normal form A.16, estimated with custom numerical methods. Red branches: unstable; blue: stable. Solid arrows indicate the directions followed when p is increased. Dashed lines: directions followed when p is decreased.

A.6.4 Supercritical pitchfork

Although not strictly critical in the sense of CTs (Kuehn et al., 2011), the supercritical pitchfrok bifurcation is worth being studied for several reasons. First, it provides an example of smooth transition in the first derivate of the potential (whereas critical transitions are discrete); hence, it allows to directly compare its notable features with those of "more" critical bifurcations. This helps in distinguishing necessary and sufficient conditions for the application of EWS (Kéfi et al., 2013; Dutta et al., 2018). Second, it is historically relevant as it connects to second-order

phase transitions studied in statistical mechanics and thus provides mutual insights (Pathria et al., 2011). The conditions determining a supercritical pitchfork bifurcation are:

$$\begin{array}{ll} -f(x,p) = f(-x,p)\,\text{(symmetry)} & \frac{\partial^2 f}{\partial x^2}(0,p_0) = 0 \\ \frac{\partial f}{\partial x}(0,p_0) = 0 & \frac{\partial f}{\partial p}(0,p_0) = 0 \\ \frac{\partial^2 f}{\partial x \partial p}(0,p_0) \neq 0 & \frac{\partial^3 f}{\partial x^3}(0,p_0) < 0 \end{array}$$

The last condition distinguishes the supercritical from the subcritical case, in a way that makes the former case stable and the latter unstable and thus critical (Kuehn, 2011). Following the conditions, its normal form is:

$$\dot{x} = px - x^3\,. \tag{A.17}$$

It is characterized by one stable point $\tilde{x}_3 = 0$ for $p < 0$ that becomes unstable at p_0, while two other stable points $\tilde{x}_{1,2}$ appear. It is the main mechanism that continuously creates bistability, in a way that one single attractor splits in two as the control parameter is varied, thus allowing different configurations for the system. However, this transition is not abrupt and no hysteresis is present. As mentioned, the mechanism is associated with second-order phase transitions in Statistical Mechanics Strogatz, 2018.

Fig. A.8 shows the evolution of $f(x,p)$ and Fig. A.9 (left) its associated bifurcation diagram. In addition, Fig. A.9 (right) reports the evolution of the associated potential.

It is interesting to compare it with that of the subcritical pitchfork bifurcation (Fig. A.9, right), in that both potentials evolve towards a two-state equilibrium, but the mechanism (around the bifurcation values) is strikingly different when considering the smoothness of the transition, its reversibility and the co-existance of multiple states. These characteristics are also reflected when small perturbations are considered (see section below) and inform the modelling of processes with similar asymptotic states but different evolution mechanisms.

CSD Scaling

$$u' = \partial_x(px - x^3)|_{\hat{x}=0} u = \mathcal{O}(p)u\ ,\ \alpha = 1\,. \tag{A.18}$$

Which is indicative for the fact that also non-critical (CT-wise) transitions have CSD and thus can exhibit Early Warning Signals, as discussed in Sec. 3.2.2.

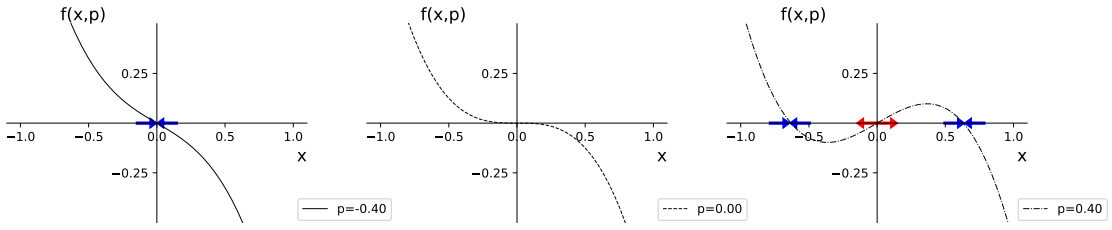


Fig. A.8 Evolution of $f(x,p)$ for a supercritical pitchfork normal form when p changes from negative to positive values. Red diverging arrows indicate the direction of instability; blue converging arrows that of stability. $p = 0$ is the critical value.

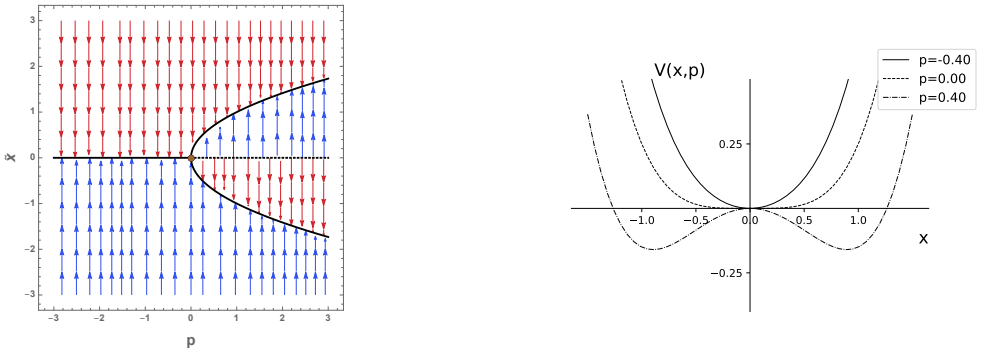


Fig. A.9 **Left:** bifurcation diagram of the supercritical pitchfork normal form A.17. The equilibrium manifold is in black (solid line: stable; dashed: unstable).Vector field lines indicate whether the system is pushed upwards or downwards. **Right:** evolution of $V(x,p)$ for a subcritical pitchfork normal form when p changes from negative to positive values. Notice that the final steady state is with two minima, but the mechanism to reach this state is different from that of a subcritical pitchfork shown in Fig. A.7, centre.

A.6.5 Cusp

The cusp is a dimension-1, codimension-2 bifurcation, for which is generic. According to R. Thom's classification of catastrophes Thom et al., 1977, the cusp is the main "organising centre" in one dimension. It is defined according to the conditions:

$$\begin{aligned}&\frac{\partial f}{\partial x}(0,p_0)=\frac{\partial^2 f}{\partial x^2}(0,p_0)=\frac{\partial f}{\partial p}(0,p_0)=\frac{\partial^2 f}{\partial p\partial x}(0,p_0)=0\\&\frac{\partial^3 f}{\partial x^3}(0,p_0)\neq 0\end{aligned} \tag{A.19}$$

The cusp normal form is:

$$\dot{x}=a+bx-x^3\,. \tag{A.20}$$

It can be regarded as an "intersection" of supercritical pitchfork and fold bifurcations (see Fig. A.10, left), or as an unfolding (Def. A.5.8) around a supercritical pitchfork bifurcation. The projection of its stable region in the (a,b) plane is shown in Fig. A.10, right. Appendix C reports a historical example of a simple toy model displaying cusp-like behaviour, that can be (and was, during the thesis) studied experimentally.

In this case, CSD is dictated by the direction of evolution, whether it is along the fold or pitchfork eigenvector.

A.6.6 Subcritical Hopf

It is a critical 2-dimensions, 1-codimension bifurcation from stable point to cycle. Its normal form is:

$$\begin{aligned}\dot{x}_1&=px_1-x_2+lx_1(x_1^2+x_2^2)\\\dot{x}_2&=x_1+px_2+lx_2(x_1^2+x_2^2)\,.\end{aligned} \tag{A.21}$$

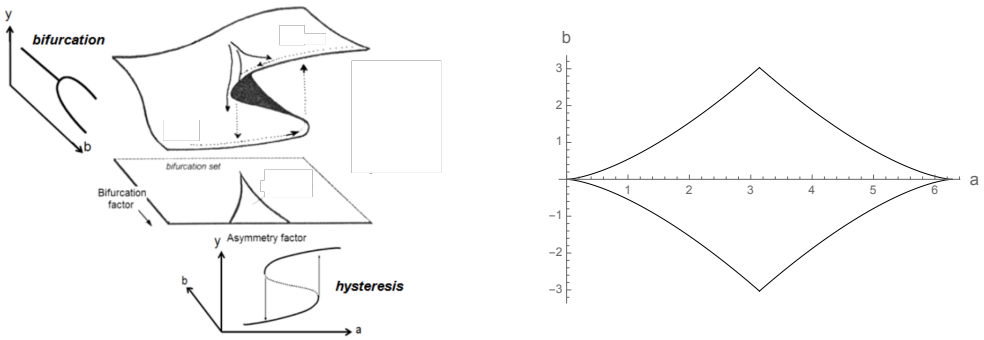


Fig. A.10 **Left:** artistic representation of a cusp stable manifold, its projection on the (a,b) plane and the "intersection" of supercritical pitchfork and fold bifurcations, according to the direction marked by one or the other parameter. Figure from Stamovlasis (2014). **Right:** projection of its stable region in the (a,b) plane. Values from Catastrophe Machine (*cf.* Appendix C).

The meta-parameter l is called *Lyapunov coefficient*. Its sign identifies the Hopf criticity ($l > 0$ for subcritical, $l < 0$ for supercritical).

Its qualitative properties can be obtained by inquiring the subcritical pitchfork, as there exist a change of coordinates that connects the two.

Proof. Equations A.21 are symmetrical with the $(x_1^2 + x_2^2)$ term that closely resembles a radius. Thus it is natural to ask for a change of coordinates from planar to polar:

$$(x_1, x_2) \to (\rho, \vartheta)\,. \tag{A.22}$$

So that:

$$\begin{cases} x_1 = \rho\cos\vartheta \\ x_2 = \rho\sin\vartheta \end{cases} \tag{A.23}$$

whose differential is:

$$\begin{cases} \dot{x}_1 = \dot{\rho}\cos\vartheta - \rho\sin\vartheta\,\dot{\vartheta} \\ \dot{x}_2 = \dot{\rho}\sin\vartheta + \rho\cos\vartheta\,\dot{\vartheta}\,. \end{cases} \tag{A.24}$$

Substituting Eq. A.23 and A.24 into Eq. A.21 gives:

$$\begin{cases} \dot{\rho}\cos\vartheta - \dot{\vartheta}\rho\sin\vartheta = \cos\vartheta(p\rho + \rho^3) - \rho\sin\vartheta \\ \dot{\rho}\sin\vartheta - \dot{\vartheta}\rho\cos\vartheta = \sin\vartheta(p\rho + \rho^3) - \rho\cos\vartheta\,. \end{cases} \tag{A.25}$$

Which is valid if:

$$\begin{cases} \dot{\rho} = p\rho + \rho^3 \\ \dot{\vartheta} = 1\,. \end{cases} \tag{A.26}$$

That is, if the evolution of the system along the radial parameter follows a subcritical pitchfork bifurcation and the angular speed is constant. □

CSD SCALING Using the previous result, CSD comes from the very same calculation as that of the pithcfork:

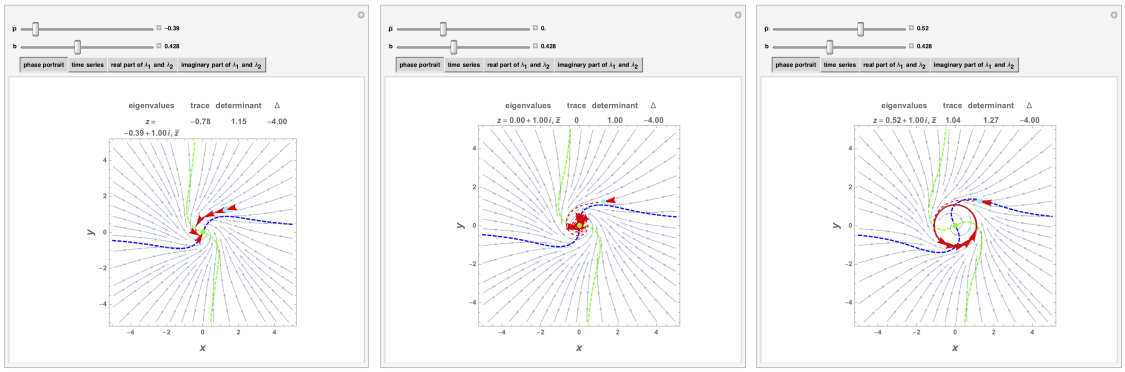
$$[u' = \partial_\rho(p\rho + \rho^3)|_{\hat{\rho}=0}u = \mathscr{O}(p)u\,,\ \alpha = 1\,. \tag{A.27}$$

A.6.7 Supercritical Hopf

It is the "non-abrupt" counterpart of the subcritical case. It emerges when complex eigenvalues become zero in their real part. It let a stable point become a fixed cycle in a continuous fashion. Its radial evolution is connected to the supercritical pitchfork by following similar reasoning to that of the subcritical Hopf (see above). Its normal form is:

$$\begin{aligned} \dot{x}_1 &= px_1 - x_2 + lx_1(x_1^2 + x_2^2) \\ \dot{x}_2 &= x_1 + px_2 + lx_2(x_1^2 + x_2^2)\,. \end{aligned} \tag{A.28}$$

The meta-parameter l is called *Lyapunov coefficient*. Its sign identifies the Hopf criticity ($l > 0$ for subcritical, $l < 0$ for supercritical).



$p < 0$ $\qquad$ $p = 0$ $\qquad$ $p > 0$

Fig. A.11 Evolution of the trajectories for a Hopf pitchfork normal form when p changes from negative to positive values, displayed on the phase portrait along $(x_1, x_2) = (x, y)$. $b = -l$ is the (negative) Lyapunov coefficient for subcritical Hopf. For $p < 0$, only fixed point solutions are stable. For $p = 0$ there is a Hopf point, leading to limit cycles for $p > 0$. Discriminant, trace and determinant (*cf.* Sec A) are also reported. The figures are generated with code adapted from Suba Thomas, "Bifurcation Diagrams with Flow Fields" http://demonstrations.wolfram.com/BifurcationDiagramsWithFlowFields/, Wolfram Demonstrations Project, Published: March 7 2011.

CSD SCALING Using the previous result, CSD comes from the very same calculation as that of the pithcfork:

$$[u' = \partial_\rho(p\rho + \rho^3)|_{\hat{\rho}=0} u = \mathscr{O}(p)u\ ,\ \alpha = 1\,. \tag{A.29}$$

A.6.8 Other - possibly critical - bifurcations

Higher-dimension and -codimension bifurcations exist and can be critical depending on their parameter settings. However, they have not been further considered in this thesis. For completeness, the most relevant ones for continuous systems (Kuehn, 2011) are listed below.

- Bautin (generalized Hopf): 2D, 2c-D (plus Lyapunov parameter l that defines its criticity depending on the sign). Similarly to the Hopf bifurcation, it describes shifts from fixed

points to limit cycles. Its normal form is:

$$\begin{aligned}\dot{x}_1 &= p_1 x_1 - x_2 + p_2 x_1(x_1^2+x_2^2) + l x_1 (x_1^2+x_2^2)^2 \\ \dot{x}_2 &= x_1 + p_1 x_2 + p_2 x_2 (x_1^2+x_2^2) + l x_2 (x_1^2+x_2^2)^2 .\end{aligned} \tag{A.30}$$

- Bogdanov-Takens: 2D, 2c-D (plus Lyapunov parameter s that defines its criticity depending on the sign). It emerges from the coalescence of a saddle-node and a Hopf. Its normal form is:

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= p_1 + p_2 x_2 + x_1^2 + s x_1 x_2 .\end{aligned} \tag{A.31}$$

- Hopf-fold and Hopf-Hopf (3D, 2 c-D) may be critical depending on parameter settings.

A.7 Stochastic Processes

The way scientists model random phenomena (or phenomena that involve random or supposedly randomic processes) is through the theory of stochastic processes. It is a discipline that combines kinetics, dynamical systems, probability theory and information theory to study noise. This section lists a few important concepts that will be used in the thesis. Most of the definitions, as well as more extensive presentations of the discipline and related problems can be found in (Freidlin et al., 1998; Gardiner, 1985). This section assumes that the reader is familiar with basic concepts of probability and statistics (suggested further reading in Papoulis et al. (2002)).

Definition A.7.1. *A **random** or **stochastic** variable $\check{x}$ is a number associated to the outcome of a random phenomenon ξ.*

$$\xi \xrightarrow{\check{x}} \check{x}(\xi) \tag{A.32}$$

Definition A.7.2. *Given a fixed value x and having defined the concept of "probability"[3], we define the **distribution function** $F_x(x)$ as*

$$F_x(x) = \mathscr{P}\{\check{x} \leq x\} \tag{A.33}$$

Notable property: $F_x(x) \geq 0$ and is monotonously increasing.

It records all the probabilities implying that a random variable assumes values that are lower than a certain number. Another (more common) way to represent the same concept is by means of the probability density function.

Definition A.7.3. *The **probability density function** is defined as*

$$f_x(x) = \frac{dF_x(x)}{dx} \tag{A.34}$$

Notable property: $f_x(x) \geq 0$. Moreover, $f(x) = \sum_i P_i \delta(x - x_i)$ for discrete cases.

[3] Defining probability and discussing its theorems is beyond the scope of this appendix. Refer to Papoulis et al., 2002 for advanced explanation.

Definition A.7.4. *A **stochastic process** is a rule to associate values to the outcome of a random phenomenon that evolves in time:*

$$sp : \xi(t) \to \check{x}(\xi, t) \tag{A.35}$$

There are discrete or continuos stochastic processes, depending on the domain of t and of possible values of $\check{x}$, with different properties. As this thesis focuses on the assumption of variable continuum, only continuous stochastic processes will be considered below. In order to study the evolution of such processes, the most used equations are the associated Langevin equations and their corresponding Fokker-Planck equations, described below. Connecting Langevin equations to fine-grained descriptions like Master equations can be done, e.g. according to Gillespie's formalism (Gillespie, 2000b), but will not be discussed further.

Definition A.7.5. *A **Wiener process** is a Gaussian process, continuos in time. It models a Brownian motion (random motion with independent increases). It is identified as the integral of a white noise process (stochastic process with zero mean).*

Definition A.7.6. *An equation that takes into account the deterministic driving force of a process and its stochastic component (often in the form of white noise) is called the associated **Langevin equation** to that process. Its general form is:*

$$\dot{x}_i = \frac{dx_i}{dx} = a_i(\bar{x}) + \sum_{m=1}^{n} b_i^m(\bar{x}) g_m(t) \tag{A.36}$$

where the superscript $\check{}$ was dropped for simplicity. a_i and b_i are generic functions. $g_m(t)$ are random functions, typically the derivative of a Wiener process dW. In terms of differential forms, a Langevin equation reads:

$$du = a(x)\,dt + b(x)\,dW \tag{A.37}$$

The general form of a Langevin equation can be more complicated and can involve differently "coloured" noise, or noise that is state-dependent or has other properties. It can be solved with different techniques and its moments can also be estimated if the distribution of every $g_m(t)$ is known.
The Langevin equation describes the evolution of the process. The evolution of its associated probability density function (meaning the rule to transit from one state to another), on the other hand, is given by the associated Fokker-Planck equation.

Definition A.7.7. *The associated **Fokker-Planck equation** to a stochastic process is a diffusion equation for its probability density function P that locally approximates an Itô stochastic differential equation. In its simplest form it reads:*

$$\partial_t P(x,t|x_0,t_0) = -\partial_x \left[A(x,t)P(x,t|x_0,t_0)\right] + \frac{1}{2}\partial_x^2 \left[B(x,t)^2 P(x,t|x_0,t_0)\right] \tag{A.38}$$

where $P(x,t|x_0,t_0)$ is the "probability to go to value x at time t given the value x_0 at time t_0, $A(x,t)$ is a "drift" term that accounts for the deterministic forces acting on the system and $B(x,t)$ is a diffusion term.

Obviously, Eq. A.38 can be extended to its multivariate case. As a PDE, it can be rarely solved analytically (most of the time by means of path integral techniques) unless one considers

its stationary values ($\partial_t P = 0$). Boundary conditions need to be specified. In case of multiplicative noise ($b(u) \neq$ const in Eq. A.37), A and B can be combinations of a and b (Risken et al., 1996).

Definition A.7.8. *Let us consider stationary processes in which* $a(x) = -\frac{\partial V(x)}{\partial x}$ *(case of a potential well) and* $b(x,t) = \sqrt{2D}$*. An* ***Ornstein-Uhlenbeck process*** *is then defined as*

$$du = -k\,u\,dt + \sqrt{2D}\,dW \tag{A.39}$$

or, alternatively, by its Fokker-Planck representation

$$\partial_t P = \partial_x(kxP) + D\,\partial_x^2 P\,. \tag{A.40}$$

Linking stochastic processes (driven by pure noise) to slow-fast systems (A.3) is not trivial. Khasminskii was among the first to suggest that the effect of fast degrees of freedom on the slow variable dynamics could be approximated as noise term (Khas' minskii, 1966). Later on, various scholars (Namachchivaya et al., 1990) assessed the equivalence of stochastic averaging and stochastic normal forms. Nowadays, it is a widely employed first approximation (Berglund et al., 2006) to model fast fluctuations as stochastic processes on top of the slowly varying degrees of freedom. The "double nature" encompassed by such modelling (genuine random realizations and fast dynamics) can be tackled by specifying functional forms for the noise terms, which might be described by non-Gaussian distributions or non-Markovian processes.

A.8 Graphs and Networks

Definitions are adapted from Gross et al. (2005); Boccaletti et al. (2006).

Definition A.8.1. *A* ***graph*** *is a triple* (V,E,I)*, where* V *and* E *are finite sets, called the vertices and the edges, respectively, and* $I : E \to V^2 0,1$ *is called the incidence function. The incidence function tells for each edge its end-vertices, and whether the edge is directed (1) or not (0).*

If an edge e is directed, then the first element and the second element of $I(e)$ denote the origin vertex and the destination vertex, respectively. This generality is required to represent all of undirected graphs, directed graphs, and mixed graphs, possibly combined with self-loops and multi-edges.

Definition A.8.2. *A* ***labeled graph*** *is a graph together with labeling functions* $L_V : V \to D_V$ *and* $L_E : E \to D_E$*, where* D_V *and* D_E *are arbitrary sets, called the vertex labels and the edge labels, respectively*

Definition A.8.3. *A* ***network*** *is the realistic counterpart of a mathematical graph. It may have an associated dynamics, links can have associated functions for activation and so on.*

As long as we bear in mind that "graph" is a topological representation whereas "network" is also involving dynamics, they can be used pretty much interchangeably. Just recall the following terms (Barabási et al., 2016):

Network Science	Graph Theory
Network	Graph
Node	Vertex
Link	Edge

A.9 Microstate and macrostate

"Macrostate" and "microstate" are terms derived from Statistical Mechanics. An intuitive outline that extends the more formal definition from physics (Pathria et al., 2011) is:
Definition A.9.1. *A microstate is a specific configuration of individual constituents that, once coupled, give rise to emerging macrostates.*

In thermodynamics, given a system with $\mathscr{D}$ degrees of freedom and its phase space $\Omega = (q_i, p_i)$, the microstate is specified by a single point in the phase space. q_i are generalised coordinates and p_i the generalised momenta. When $\mathscr{D} \to \infty$, the phase space can be divided into cells of size $\Delta_0 = \Delta q_i \Delta p_i$, each treated as a microstate. Similar reasoning can be put forward for other systems, where the generalised coordinates relate to other quantities. For instance, gene regulatory networks can be described in gene state spaces (Wu et al., 2020), where each combination of gene expression levels can be treated as a microstate.

Definition A.9.2. *A macrostate of a system refers to its macroscopic (possibly emergent) properties that account for its function, such as temperature or density.*

Overall, a macrostate is what results from interactions of single components of a complex system. It is an "emerging" property. In statistical mechanics, different configurations of molecule speeds (its microstate) create a single temperature value (associated macrostate). Any given macrostate may be associated with many different microstates.

Example: for a biological system, we can think of a macrostate as a functional state. Importantly, it can be observed and measured. We conjecture that gene expression profiles can be related to microstates in a biological context. These are often inferred. If a circuit is known, we can think of its control parameters as microstates, as different combinations may produce the same outcome (e.g. the same function, or macrostate).

Appendix B

Notable results for stochastic dynamical systems

This appendix puts together notable theoretical results for stochastic dynamical systems, that are of crucial importance to study systems' resilience and early warning signals. It is introduced by revising and exemplifying the concept of generic bifurcation, for their importance in the study of B-tipping.

B.1 Generic bifurcations

As introduced in Def. A.5.10, a bifurcation is generic if it is not mapped onto another one after an unfolding (slight additive perturbation). Among low dimension bifurcations, fold and Hopf bifurcations are generic for codimension one (Kuznetsov, 2013). In fact, the real part of an eigenvalue can reach zero (vanish) in two ways: either $\lambda = 0$ (appearance of a fold bifurcation), or a complex conjugate pair of eigenvalues reach the imaginary axis, $\lambda = \pm\omega_0 i$ (appearance of a Hopf bifurcation). The cusp bifurcation is generic for codimension two. Pitchfork and transcritical bifurcations are generic only if additional constraints about symmetry and asymptotic behaviour are considered, otherwise their unfolding maps them into a fold.

Envisioning realistic applications, when small perturbations or deviations from an ideal model are always present, the standard way of considering B-tipping is through the analysis of fold bifurcations and their associated early warning signals. Unless other types are considered for different scopes, the fold bifurcation is thus a good candidate to address local analysis and prediction of regime shifts driven by B-tipping mechanisms. To visualize this, consider a subcritical bifurcation (Fig. B.1) and observe how its critical neighbourhood is locally mapped onto a fold after a small unfolding introduced by an additive parameter q, such that Eq. A.14 becomes $\dot{x} = px + x^3 + q$.

B.1.1 Example: the Euler buckling problem

An example of unfolding is the "Euler buckling problem". The swiss mathematician was the first to pose the following problem: imagine a slender, straight column sustaining a burden. What is its behaviour when the burden increases in weight? It was soon found that there exist a critical

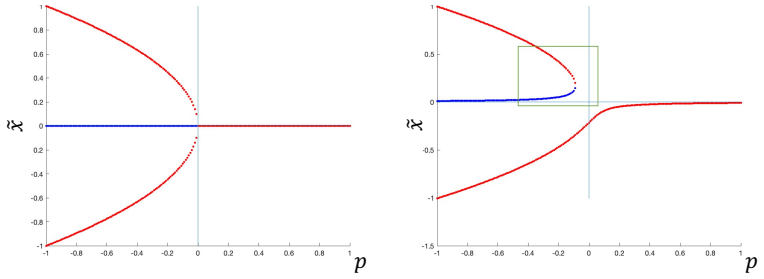


Fig. B.1 **Left:** bifurcation diagram of a subcritical pitchfork bifurcation A.14. For better visualization, the vector field has been removed and the fixed equilibrium manifold coloured according to the stability of its points. Blue = stable; red = unstable. **Right:** bifurcation diagram of an unfolded subcritical pitchfork bifurcation $\dot{x} = px + x^3 + q$, where q is the unfolding parameter. In this case, its value is fixed to $q = 0.01$ for visualization purposes. Note that locally (region highlighted by the green square) the bifurcation corresponds to a fold (*cf.* Sec. A).

weight value that makes the column buckle. This behaviour, known as "Euler's buckling", is well known in engineering, including control engineering (Venkadesan et al., 2007). For idealized slender columns, it is described with a subcritical pitchfork bifurcation Thompson et al., 2011; Kuehn, 2013. θ (angle of the column with respect to upright position) is the only degree of freedom available given the rotational invariance; $\mathscr{F}_\int$ is the control parameter and is viewed as the force applied to the column. The model is (Venkadesan et al., 2007):

$$\dot{\theta} = p_1(\mathscr{F}_\int - p_2)\theta + p_3\theta^3 - p_4\theta^5\,, \tag{B.1}$$

where p_i are additional parameters to be fixed.

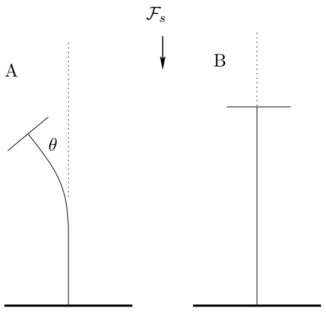


Fig. B.2 Sketch of a column subjected to pressing force $\mathscr{F}_\int$ (Kuehn, 2013), eventually buckling.

When considering realistic columns, it is necessary to consider slight imperfections. This breaks the symmetry of the column, resulting in asymmetric buckling. Model-wise, these imperfections are inserted in the normal form as additive terms. The modified model is:

$$\dot{\theta} = p_1(\mathscr{F}_\int - p_2)\theta + p_3\theta^3 - p_4\theta^5 + q\,. \tag{B.2}$$

q is the unfolding parameter that destroys the bifurcation in a similar fashion to Fig. B.1, right.

B.2 Ornstein-Uhlenbeck processes and critical slowing down

When studying critical transitions close to tipping points, Ornstein-Uhlenbeck (O-U) processes (Eq. A.39) are of utmost importance to approximate their dynamics subject to small fluctuations. In fact, they correspond to continuous-time counterparts of autoregressive models, which are in turn the archetype of critical transition-prone systems, Scheffer (2009). Overall, O-U processes describe the motion of a particle moving in a quadratic potential with an escape barrier – related to fold bifurcation normal form (*cf.* Sec. A) – and subject to noise, that can model either random disturbances or the action of fast degrees of freedom (Berglund et al., 2006). To recall the working hypothesis, they are:

1. The system is evolving along the leading eigenvector of the bifurcation, that is, we are hypothesising a centre manifold setting. This modelling simplification holds as long as the system approaches a tipping point characterized by a shrinking manifold along the leading direction. This is also valid as long as we define a single macroscopic variable that accounts for a global mean field.

2. Slowly varying parameter allowing the system to follow its changes adiabatically (slow-fast system and equilibrium hypothesis).

3. Noise level is low, and a high signal-to-noise ratio is maintained. To begin with, let us consider additive white noise.

To obtain the theoretical results leading to the existence and identification of early warning signals, hypothesis 1-3 must be satisfied. Relaxing one or more might yield different conclusions in the expected behaviour of EWS (see Chapter 5).

Historically, Ornstein-Uhlenbeck processes have been deeply studied and many results and demonstrations are available in books covering stochastic processes (e.g. Allen (2010); Gardiner (1985); Berglund et al. (2006); Risken et al. (1996); Papoulis et al. (2002)). Hence, this section concentrates on results that are of direct relevance for the current development of the CT framework, while further reading is directed to the books just mentioned.

B.2.1 Studying how perturbations evolve in B-tipping driven by a fold

Recall the normal form of a generic fold bifurcation, Eq. A.10:

$$\dot{x} = p + x^2, \tag{B.3}$$

It has two steady states:

$$\hat{x_1} = -\sqrt{-p-0} \text{ (stable)} \tag{B.4}$$

$$\hat{x_2} = -\sqrt{-p-0} \text{ (unstable)} \tag{B.5}$$

where the term "-0" explicits the distance from the saddle point that, from the properties of the normal form, is located in $x_c = 0$. Let us now "sit" in a neighborhood of the attractor (stable fixed point) $\hat{x_1}$ to see what happen after small perturbations. Hence, we perform a local linearization by

considering $\delta x = (x - \hat{x}_1)$. Thus:

$$\frac{d\delta x}{dt} \simeq f(\hat{x}_1) + \frac{\partial f}{\partial x}|_{\hat{x}_1} \delta x + \mathrm{O}(\delta x^2)\,. \tag{B.6}$$

So, using Eq. B.3 and Eq. B.4, we obtain:

$$\frac{d\delta x}{dt} \simeq 2\sqrt{-p}\delta x\,. \tag{B.7}$$

Let us now convert this deterministic form into a stochastic one by adding a Wiener process with diffusion term σ. This modelling choice converts the family of ODEs into SDEs. Here, we skip some mathematical formalism, which has been covered during the years by Khas' minskii (1966), Namachchivaya et al. (1990) and Berglund et al. (2006) among the others, and concentrate on the physical interpretation. In addition, a change $\delta x \to y$ makes the notation lighter into:

$$dy = 2\sqrt{-p}y\,dt + \sigma dW\,. \tag{B.8}$$

The equation describes a system evolving under small noise in a neighbourhood of the stable equilibrium, when this is not far away from the saddle node (linearization). An illustrative sketch is presented in Fig. B.3.

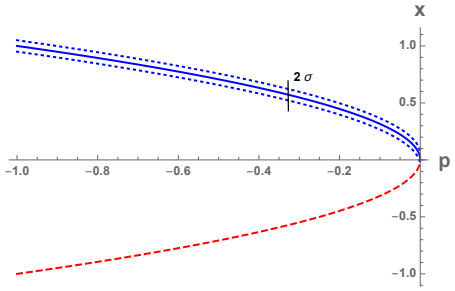


Fig. B.3 Bifurcation diagram of the saddle-node, with uncertainty band given by the noise level

The term $(\hat{x}_1 - 0) = \sqrt{-p}$ is precisely the distance of the stable equilibrium from the saddle node and depends on the leading parameter p. From now on, we rescale it to a new variable $-k$.

$$dy = -k\,y\,dt + \sigma dW \tag{B.9}$$

The sign "$-$" in "$-k$" is inserted to be consistent with the following interpretation: Eq. B.9 is the associated Langevin equation to a Ornstein-Uhlenbeck process (compare with Eq. A.39 and see Papoulis et al. (2002); Gardiner (1985)).

Consequently, we can interpret the term that multiplies the deterministic drift as $-\frac{\partial V}{\partial x}$ where $V(x)$ is the potential governing the drift of the particle subjected to random noise. In our case, thanks to the choices made,

$$V = \frac{1}{2}k y^2\,, \tag{B.10}$$

that is, a quadratically shaped adjoining potential typical of an overdamped oscillator under noise, of which k represents the depth. The quadratic potential can be termed differently, according to the points of view of different sub-fields: "effective potential" in nonequlibrium phase transitions

(Zaikin et al., 2001), "quasi-potential" in large deviation theory (Freidlin et al., 1998; Nolting et al., 2016; Zhou et al., 2012). The working hypothesis is that boundary of the ideal potential V can grasp the boundary of the attracting basin of the original model after sufficiently long time. Hence, it proves to be analytically tractable and presents the necessary steps to understand the main qualitative features of more complicated critical transitions. However, it requires ad hoc extensions when studying system-specific quantitative details like observability boundaries and lead times.

Given the relationship $k = 2(0+\sqrt{-p})$ from Eq. B.8, when the equilibrium approaches a saddle point (p goes towards zero from negative values) the associated analytical potential gets flatter as k approaches zero, see also Fig. A.2 (centre). Intuitively, recall the "pebble-down-the-hill" analogy (Sec. 2.2.5). The potential governs the acceleration under which a rolling ball moves; hence, it is connected to the recovery time after a perturbation that moves the ball away from the equilibrium. When the system approaches its saddle node, the recovery time slows down. Thus we get the phenomenon of Critical Slowing Down CSD. The scaling of slowing down depends on the specific dependence of k to p. For a fold, it goes as a square root; or other bifurcations, it might vary (see sections above).

Thanks to the discussed methodology, we can now study saddle-node driven B-tipping using the formalism of Stochastic Processes, where the potential depth (at quasi-steady state) plays the role of the control parameter. In the remaining of this section, the general distance-to-critical-point k will be employed. For each specific bifurcation, it will suffice to substitute the dependence $k = k(p)$ (see also Kuehn et al. (2011)) to get the correct scaling of each notable characteristic (statistical indicators, power spectrum, etc.). Estimating the scaling of observed regime shifts has been proposed as a notable feature to distinguish the (potentially) underlying bifurcations (Meisel et al., 2012b; Bury et al., 2020).

B.2.2 Critical Slowing Down (CSD)

In statistical mechanics, CSD is a phenomenon connected to second-order phase transitions. At the equilibrium point, decays of small perturbations are slower, as they follow lower order polynomials Strogatz, 2018. First-order phase transitions, which are in turn related to critical transitions, are also affected by CSD when approaching a tipping point. Let us first concentrate on critical slowing down in fold transitions (around saddle-node points). According to centre manifold theory, it is sufficient to study the behaviour of the vector field in a neighbourhood of the tipping point in order to inquire the relaxation coefficient (often called Lyapunov coefficient, Kuehn et al. (2011)).

Similarly to the subsection above, we begin from the normal form Eq. B.3 and concentrate around the stable equilibrium point $\hat{x}_1$. Let us consider small deviations y around $\hat{x}$. For sufficiently small deviations, they are described by a linearization $\dot{y} = f(\hat{x}+y) - f(\hat{x})\, y \simeq (D_x f)|_{\hat{x}}\, y$, where $f(x)$ is the vector field (right hand term in Eq. B.9). Therefore, for the fold we get:

$$u' = -2x|_{\sqrt{-p}}\, u = -2\sqrt{-p}u \tag{B.11}$$

Hence, any perturbation decays as $O(-p^{1/2})$ back to the equilibrium, with a relaxation coefficient is $\alpha = 1/2$. Determining the relaxation coefficients for other bifurcations allows assessing their

CSD scaling (reported in Sec. A or each normal form).

B.2.3 Stochastic indicators towards tipping points

In an Ornstein-Uhlenbeck process, the drift coefficient (distance to tipping) can be interpreted as the potential depth, and is thus immediately related to system's resilience. As the trends of statistical indicators of O-U processes is known to depend on the drift term, it is also immediate to study how such trends change when the leading parameter is varied. The corresponding results have been proposed to provide early warning signals for an impending shift. Let us express Eq. B.9 using another typical notation for stochastic processes, $y_t = y(t)$ and $\sigma = \sqrt{2D}$.

$$dy_t = -k\,y_t\,dt + \sqrt{2D}dW_t \tag{B.12}$$

The statistical indicators of the stationary process[1] Eq. B.12 are derived as (Papoulis et al., 2002; Gardiner, 1985):

- Covariance (temporal self-correlation): $\langle y_t y_{t'}\rangle = \frac{D}{k}e^{-k|t-t'|}$.
- Variance: $\langle y_t^2\rangle - \langle y_t\rangle^2 = \frac{D}{k}(1-e^{-2kt})$. Hence, for $t\to\infty$,

$$Var \xrightarrow{t\to\infty} \frac{D}{k} \tag{B.13}$$

- Autocorrelation: $e^{-k|t-t'|}$. Hence,

$$AC(1) = e^{-k}. \tag{B.14}$$

- In general, all other statistical moments can be expressed as:

$$\langle y^n\rangle - \langle y\rangle^n = \int_{-\infty}^{\infty}(y'-\mu)^n P(y')dy' \tag{B.15}$$

  where $P(x)$ is the probability density function, a Gaussian in case of white noise or derived from the Fokker-Plank equation (see Sec. B), and μ is the mean expected value.

- Power spectral density:

$$S(\omega) = \frac{D}{\pi(k^2+\omega^2)} \tag{B.16}$$

- Shannon entropy:

$$H_s(y) = \frac{1}{2}\log(2\pi\mathrm{Var}) + 1 \tag{B.17}$$

Additional entropy and auto-information measures can be further derived for other processes obeying other evolution laws (Frank, 2004; Heseltine et al., 2016; Heseltine et al., 2019; Chapeau-Blondeau, 2007) or being analysed with diverse orthogonal basis like wavelets (Coifman et al., 1992; Kirby, 2005; Abdel-Hamid et al., 2016).

[1] A stochastic process is said to be *stationary* if its mean does not depend on time (mean $= \eta =$ const) and its covariance only depends on time differences: $\langle x(t)x(t')\rangle = \langle x(t)x(t+\tau)\rangle$, where $\tau = t'-t$.

Proof. Proof for Variance, Covariance and Autocorrelation. Equation B.12 is solved with an Itô-like integration and the method of constants for differential equation solution :

1. Ansatz of solution: $z = y_t\, e^{kt}$
2. Itô differentiation:

$$\begin{aligned} d\left(e^{kt}y_t\right) &= k\, e^{kt} y_t\, dt + e^{kt} dy_t \\ &= k e^{kt} y_t\, dt + e^{kt}\left(-k\, y_t\, dt + \sqrt{2D}\, dW_t\right) \\ &= e^{kt}\sqrt{2D}\, dW_t\,. \end{aligned}$$

3. Integration and re-substitution for z

$$\begin{aligned} e^{kt}y_t - y_0 &= \sqrt{2D}\int_0^t e^{sk} dW_s \\ y_t &= y_0 e^{-kt} + \sqrt{2D}\int_0^t e^{k(s-t)} dW_s\,. \end{aligned} \tag{B.18}$$

   This is the stochastic trajectory.

If the initial condition is deterministic or Gaussian, it is straightforward to estimate mean and time correlation function:

$$\begin{aligned} \langle y_t\rangle &= \langle y_0 e^{-kt} + \sqrt{2D}\int_0^t e^{k(s-t)} dW_s\rangle \qquad \textbf{Mean} \\ &= \langle y_0 e^{-kt}\rangle + \langle\sqrt{2D}\int_0^t e^{k(s-t)} dW_s\rangle \\ &= y_0 e^{-kt} + \sqrt{2D}\int_0^t e^{k(s-t)}\langle dW_s\rangle \\ &= y_0 e^{-kt} + 0 \xrightarrow{t\to\infty} 0 \end{aligned}$$

$$\begin{aligned} \langle y_t y_{t'}\rangle &= y_0^2 e^{-k(t-t')} + \left(\sqrt{2D}\right)^2 \int_0^t\int_0^{t'} e^{k(s-t)} e^{k(s'-t')}\langle dW_s\, dW_{s'}\rangle \qquad \textbf{Covar.} \\ &= y_0^2 e^{-k(t-t')} + 2D\, e^{-k(t+t')}\int_o^t\int_0^{t'} e^{k(s+s')}\delta(s-s')\, ds\, ds' \end{aligned}$$

where we used the property of Wiener processes $\langle dW_s\, dW_{s'}\rangle = \delta(s-s')\, ds\, ds'$. $\delta(s-s')$ is the Dirac delta. Then:

$$\begin{aligned} \langle y_t y_{t'}\rangle &= y_0^2 e^{-k(t+t')} + 2D\, e^{-k(t+t')}\left.\frac{e^{2k-min(s,s')}}{2k}\right|_0^{min(t,t')} \\ &= y_0^2 e^{-k(t+t')} + \frac{D}{k} e^{-k(t+t')}\left(e^{2k-min(t,t')} - 1\right) \\ &= y_0^2 e^{-k(t+t')} + \frac{D}{k} e^{-k|t-t'|} - \frac{D}{k} e^{-k(t+t')}\,. \end{aligned}$$

By neglecting the rapid initial transient, we finally get:

$$\langle y_t y_{t'} \rangle \xrightarrow{t,t' \to \infty} \frac{D}{k} e^{-k|t-t'|} \tag{B.19}$$

Finally, to estimate the **variance**, use its definition:

$$\langle y_t^2 \rangle - \langle y_t \rangle^2 = \langle y_t y_{t'}|_{t'=t} \rangle - \langle y_t \rangle^2 = \frac{D}{k}\left(1 - e^{-2kt}\right) \tag{B.20}$$

The **autocorrelation** follows as:

$$Autoc = \frac{Cov(x(t)x(t'))}{\sqrt{Var(x(t))Var(x(t'))}} = e^{-k\cdot|t-t'|} \text{ for } t,t' \to \infty \tag{B.21}$$

□

Proof. Proof for the power spectral density $S(\omega)$.

In a stationary stochastic process, its spectral density is defined as the Fourier transform of the Covariance function Papoulis et al., 2002:

$$S(\omega) = \mathscr{F}\left[E\left\{y(t+\tau)y(t)\right\}\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} R(\tau)e^{-i\omega t}dt \tag{B.22}$$

where $R(\tau)$ is the covariance function and the dimension of the spectral density is $[S(\omega)] = \text{dB}$. In case $t,t' \to \infty$ we already have the expression for the covariance function (Eq. B.19), so the calculation is straightforward as long as we perform a change of coordinates $|t'-t| \to |t|$ which is valid for stationary processes:

$$\begin{aligned} S(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega t}\frac{D}{k}e^{-k|t|} \\ &= \frac{D}{2\pi k}\left(\int_{-\infty}^{0} e^{t(-i\omega+k)}dt + \int_{0}^{\infty} e^{-t(i\omega+k)}dt\right) \\ &= \frac{D}{2\pi k}\left(\frac{1}{-i\omega+k} + \frac{1}{i\omega+k}\right) \\ &= \frac{D}{\pi k}\frac{k}{k^2+\omega^2} \end{aligned}$$

So the final expression is:

$$S(\omega) = \frac{D}{\pi(k^2+\omega^2)} \tag{B.23}$$

Alternatively, we can consider the O-U process as a stochastic input process with $R_{xx}(\tau) = q\delta(\tau)$ (white noise input) and linear transfer function $h(\tau)$ given by $h'(\tau) + ch(\tau) = \delta(\tau)$ (linear response given by potential). Then, in Fourier space,

$$S_{yy} = q\frac{1}{k^2+\omega^2}$$

where $q = D/\pi$ and $c = k$ (from the potential). □

Proof. Proof for the Shannon entropy of a stationary process subject to Gaussian noise.

Consider a Gaussian distributed variable $y \sim \mathcal{N}(\mu, \text{Var})$. Its entropy is:

$$
\begin{aligned}
H_s(y) &= -\int p(y') \log p(y') dy' = \\
&= -\mathbb{E}\left[\log \mathcal{N}(\mu, \text{Var})\right] = \mathbb{E}\left[\log\left[\frac{1}{\sqrt{2\pi\text{Var}}} \exp\left(-\frac{1}{2\text{Var}}(x-\mu)^2\right)\right]\right] = \\
&= \frac{1}{2}\log\left(2\pi\text{Var}\right) + \frac{1}{2\text{Var}}\mathbb{E}\left[(x-\mu)^2\right] = \\
&= \frac{1}{2}\left(\log(2\pi\text{Var}) + 1\right)
\end{aligned}
$$

□

Eq. B.19 and B.20 express how covariance and variance evolve when a quasi-steady parameter shapes the potential towards a tipping point. The limit $t, t' \to \infty$ is representative of system timescales that are much longer than that of the noise (the system relaxes to its equilibrium). It is possible to recognize an increase in both measurements when $k \to 0$. This phenomenon, illustrated in Fig. B.4, was suggested by Scheffer et al. (2009) as an early warning signal (EWS) to detect the approach towards a tipping point. Since the spectral density also depends on k, it can also be studied to suggest other early warning signals, as well as entropy measures.

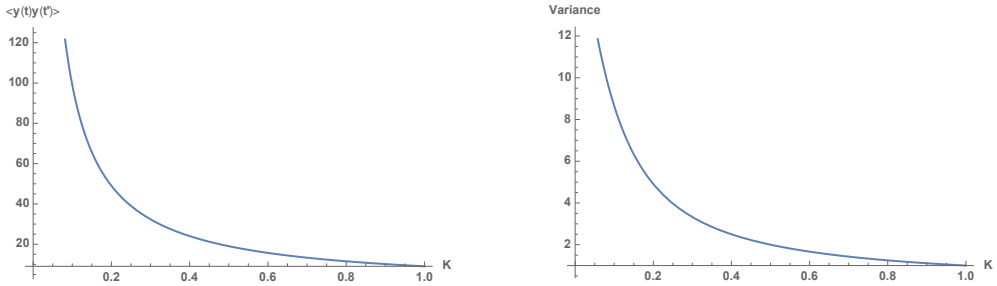


Fig. B.4 **Left:** Covariance, with $\delta t = 1$, vs parameter k. **(Right)** Variance vs k. Note that both measures increase when $k \to 0$. Such an increase has been suggested to provide an early warning signal.

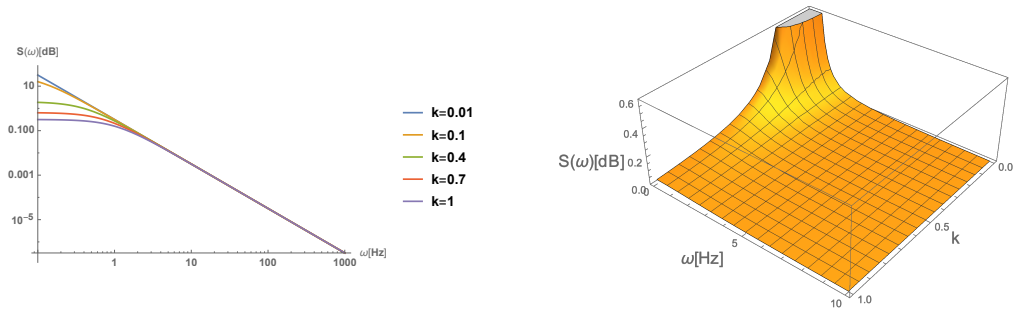


Fig. B.5 **Left:** Power spectral density for frequencies ω, at different values of the parameter k. The reddening (increasing power on low frequencies) happening for decreasing k has been associated to potential EWS (Thompson et al., 2011; Biggs et al., 2009). **Right:** Full dependency of $S(\omega, k)$ on both ω and k, provided that the latter is varied at quasi-steady sate and hat the process is overall stationary.

From the theoretical derivation of early warning signals, one immediately recognise some important features that have been studied more in detail in recent computational works, i.e. (Kuehn,

2013; Dessavre et al., 2019; Brett et al., 2017; O'Regan et al., 2018) and assessed against empirical data in (Proverbio et al., 2022c). The variance is generally more "stable" than covariance and autocorrelation against the choice of a moving window, as it does not depend on a $\delta t = |t - t'|$. As shown in Fig. B.6, for higher δt, the time correlation remains flat for wider intervals of k, but then increases more rapidly. This induces a tradeoff between its sensitivity (how big its differential change is, which is related to how well we can discriminate between a real signal and a sampling of noise) and the predictability lag (how far from the critical value it is possible to determine an increase). On the other hand, the autocorrelation does not depend explicitly on the noise distribution and is thus insensitive to its details, contrary to the variance.

These characteristics hold as long as we have repeated measurements over an ergodic distribution. The assumption of ergodicity[2] might be true if we allow the system to relax back to its equilibrium, that is, k is almost fixed with comparison to the system timescale and the measurement rate. However, a sliding window technique might yield invalid estimates if the ergodic assumption is not valid.

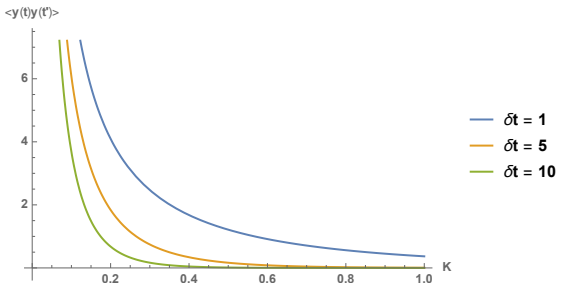


Fig. B.6 Temporal correlation vs k at different δt: $\langle y(t)y(t+\delta t)\rangle$.

B.3 Fokker-Plank equation and stationary probability density functions

The Fokker-Planck equation A.38 describes the evolution of the probability density function $P(x,t|x_0,t_0)$ that a process takes the value x at time t, from its previous state. Let us recall it here:

$$\partial_t P(x,t|x_0,t_0) = -\partial_x \left[A(x,t)P(x,t|x_0,t_0)\right] + \frac{1}{2}\partial_x^2 \left[B(x,t)^2 P(x,t|x_0,t_0)\right] \tag{B.24}$$

From now on, the terms "$|x_0,t_0$" will be implied. Solving the Fokker-Planck provides information about the full evolution of $P(x,t)$, which is paramount to estimate statistical moments for the whole process as (Papoulis et al., 2002):

$$\mathbb{E}[h(x)] = \int_x^\infty h(x')P(x')dx' \tag{B.25}$$

where $\mathbb{E}[\cdot]$ denotes the expected value.

Solving the Fokker-Plank equation might be complicated, depending on the specific shape of $A(x,t)$ and $B(x,t)$. For stationary processes and for differentiable and integrable functions,

[2] An ergodic system is such if the mean over repeated measurements equals the mean over time.

it is instead possible to estimate the stationary probability density function (SPDF) by solving $\partial_t P(x,t) = 0$ as (Risken et al., 1996):

$$P_s(x) = \frac{N_c}{B(x)} \exp\left(\int^x \frac{A(x')}{B(x')} dx'\right). \tag{B.26}$$

N_c is a normalization constant to interpret $\int P_s(x')dx'$ as a probability. $P_s(x)$ can also be put in the form:

$$P_s(x) = N_c e^{-\phi(x)}. \tag{B.27}$$

Here, $\phi(x)$ is called the stochastic potential of the system. This function maps the equilibria of the stochastic dynamical system and their basins of attraction, allowing an interpretation of "pebble down the hill" for the asymptotic behaviour of the system. It thus reconstructs the "adjoint"/"effective"/"quasi-" potential introduced before, starting from the PDE associated to the considered system. Application examples include (Sharma et al., 2016a; Su et al., 2019).

B.4 Kramers' escape rate and mean first passage times

Essentially, the idea behind early warning signals for critical transitions is: given a stochastic process, how do its statistical indicators change when the underlying potential is varied under the influence of a leading parameter?

As a consequence, one can use the manifold results in stochastic processes theory, analyse their dependence on parameters and interpret them accordingly. Two additional measures that provide insights onto the behaviour of stochastic dynamical processes are the mean first passage time (MFPT) and the Kramers' escape rate. In escape problems from a potential well, they respectively measure how long a system is expected to jump onto an alternative equilibrium for the first time, or how often the system is expected to reach the top of the potential barrier. Since they are directly related to probabilities of shifting equilibria, both the MFPT and the Kramers' rate have been proposed as measures of system's resilience (Sharma et al., 2016a; Arani et al., 2021). Like above, this section provides an overview of salient features for both measures. Further reading and proofs can be found in Van Kampen (1992) and Gardiner (1985).

Consider a double-well potential like that in Fig. B.7. Assume it is (locally) quadratic near the attractor and (locally) quadratic around the barrier tip. Also assume we work in the damped regime and thus we can write the stochastic dynamics as a Langevin equation. Because of noise, the system can fluctuate away from the equilibrium (point a) and reach the top of the barrier (point b) by chance. From there, it can rapidly fall into a new attracting state. As above, we work with the ergodic hypothesis that the average time evolution of a system equals its statistical repetitions. Another assumption is that the well is deep compared to noise level. The rate at which many copies of the system can reach the potential top, in a stationary process at equilibrium under Gaussian noise and in time-independent potential, is solved by Kramers' escape problem (Van Kampen, 1992). A clear derivation, accounting for each assumption presented above, is provided by Ritchie Jr (2016). As a bottom line, the Kramers' escape rate, for any left(right)-adjointed

potential, is given by:

$$\langle T(y)\rangle \simeq 2\pi \frac{e^{\frac{V(b)-V(a)}{D}}}{\sqrt{V''(a)|V''(b)|}} \tag{B.28}$$

where V'' denotes second derivate and a and b are, respectively, the point of potential minimum and maximum. Recall that, in this section, the variable y denotes the local linearization next to a quadratic potential.

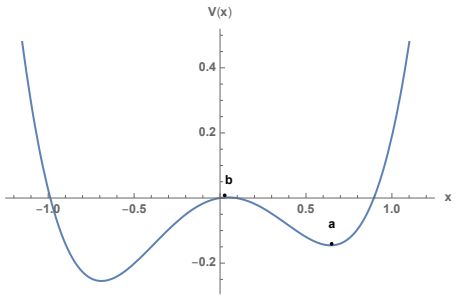


Fig. B.7 Representation of a double-well potential.

In case of O-U process, associated to a linearised approach to saddle-node, evaluating the mean Kramers' escape rate $\langle T(y)\rangle$ vs the parameter k is straightforward. k represents both the potential height difference and its second derivative nn the equilibrium points (see Eq. B.10 and substitute Eq. B.5 into it.). Thus, by substituting k into Eq. B.28, we get:

$$\langle T(y)\rangle \simeq \frac{2\pi}{k} e^{\frac{-k}{D}} \tag{B.29}$$

where the absolute value is dropped since $k \in \mathbb{R}_+$. Fig. B.8 shows how the escape rate changes according to $k \to 0$ and for different levels of noise intensity D. The mean escape rate of particles flowing from the original well increases rapidly as the tipping point is approached. With high noise, $\langle T(y)\rangle$ is already non-zero even for large k, which is indicative of the N-tipping phenomenon. Therefore, provided that abundant measures and prior knowledge of noise levels are available, the mean escape rate can be used as a proxy for the potential depth and thus for the system's resilience.

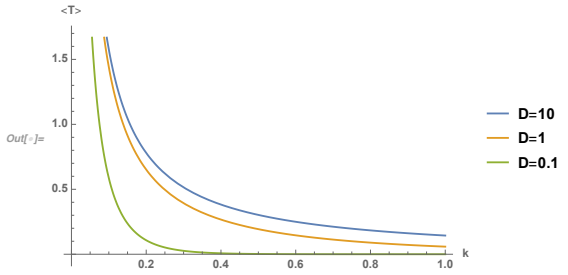


Fig. B.8 Kramers' escape rate $T(y)$ vs k, at different noise levels D from Eq. B.12.

Instead, considering the full potential $U(x) = -px + 1/3x^3$ generated by a fold normal form $\dot{x} = p - x^2$. It yields an excursion time (Freidlin et al., 1979; Berglund et al., 2006)

$$\tau_E = \mathscr{O}\left(\exp\left(\frac{2\Delta U}{\sigma^2}\right)\right), \tag{B.30}$$

with $\Delta U = U(\sqrt{p}) - U(-\sqrt{p}) = 4/3p\sqrt{p}$. Excursions are very rare if the noise level is small compared to the time scale separation, $0 < \sigma \ll \sqrt{\varepsilon} \ll 1$. If this is not true, random excursions are much more likely and n-tipping might occur with high chance.

Finally, consider a generic double-well potential elicited by a generic bistable function

$$dx = f(x)dt + \sigma dW\,. \tag{B.31}$$

Call $\hat{x}_l$ and $\hat{x}_r$ the left and right stable steady-states separated by an unstable potential barrier $\hat{x}_b$. A moving particle subject to noise can exit from its potential well (say, from $\hat{x}_r$) and cross the barrier. When does the first passage occur (on average)? The question is addressed by the exit time. When the first-passage time is averaged over many realizations, we get the mean first-passage time MFPT. Let us call $\mathscr{T}(x)$ the MFPT towards $\hat{x}_b$ starting in a generic $x > \hat{x}_b$ (around a, if we keep the same image from Fig. B.7). $\mathscr{T}(x)$ satisfies the following diffusion equation (Gardiner, 1985):

$$f(x)\frac{\partial \mathscr{T}(x)}{\partial x} + \frac{1}{2}\sigma\frac{\partial^2 \mathscr{T}(x)}{\partial x^2} = -1 \tag{B.32}$$

given boundary conditions $T(\hat{x}_b) = 0$ and $T(x_R) = 0$, where x_R denotes the rightmost boundary of the potential well. The MFPT is obtained by solving Eq. B.32.The average passage time from the basin of attraction of $\hat{x}_r$ towards state $\hat{x}_b$ is given by:

$$\begin{aligned}\mathscr{T}(\hat{x}_r) &= 2\int_{\hat{x}_b}^{\hat{x}_r}\frac{1}{\psi(y)}dy\int_y^{x_R}\frac{\psi(z)}{\sigma}dz \\ \psi(x) &= \exp\left(\int_{x_0}^{x}\frac{2f(x')}{\sigma}dx'\right)\end{aligned} \tag{B.33}$$

To estimate the MFPT from the leftmost attractor, it suffices to set the boundary conditions appropriately. It is also possible to extend the calculation to non-homogeneous processes and to multi-dimensional domains (Gardiner, 1985; Arnold et al., 1983).

Appendix C

The Catastrophe Machine

The Zeeman's catastrophe machine is a historically important example to illustrate critical transitions governed by a cusp bifurcation. It can be easily constructed and allows to showcase the effect of control parameters towards abrupt shifts. Since it is governed by classical laws of mechanics, it is also possible to reproduce its dynamics and to predict its stability behaviour. In addition, the Catastrophe Machine is a powerful tool for teaching: it is among the simplest (closed) systems that exhibit tipping and hysteresis, which can be shown simply and intuitively. More advanced machines can be used to illustrate different critical transitions (Woodcock et al., 1976) or the appearance of chaotic behaviour when multiple nonlinear systems are coupled (Nagy et al., 2013).

Because of its benefits (didactics, intuition, manipulation, modelling), the catastrophe machine is a nice tool to have. Thanks to the support of the LCSB maintenance team, I constructed a simple machine (described in details below) and a more advanced one, that shows routes to chaos. The photographs in this section depict our catastrophe machine during practical exercises.

C.1 What is the Catastrophe Machine

The Catastrophe machine was invented by E. C. Zeeman as a toy for illustrating the at-that-time new Catastrophe Theory (Poston et al., 1973; Zeeman, 1979). It consists on a wheel governed by a couple of elastics, connected to the wheel by a pivot; one elastic is fixed to a pin, while the other can be controlled by the user. Everything is set on a board, see Fig. C.1.

The machine empirically exhibits sudden shifts: if the second elastic is stretched past a threshold, the wheel turns abruptly. Hysteresis is also observed when the second elastic is relaxed: to tip back, it must shrink past the point of the first tipping (Fig. C.2a–d). The machine shows the same qualitative behaviour regardless of wheel radius, un-stretched elastics lengths and distance from pin to wheel centre. This is a"structural stability" property: changes in parameters make no essential qualitative difference (Poston et al., 1979). Hence, the catastrophe machine is a perfect toy model to study critical transitions: it is controllable, it is possible to calculate its potential and

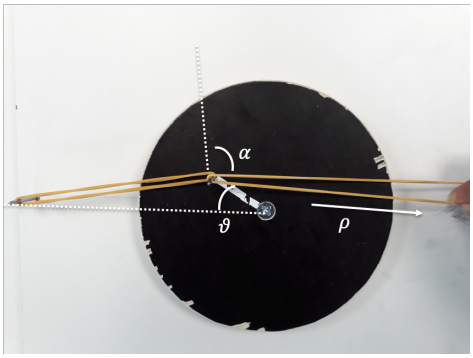


Fig. C.1 The catastrophe machine is composed by wheel and elastics, one of which is fixed by a pin while the other is held by the user. The angle θ is considered as state variable; α represents the reciprocal position of the two elastics; ρ is the force exerted by the user (forcing); it also accounts for the friction, that is always opposing the movement. In order for the machine to move, the friction should not be too high (Poston et al., 1979)).

to study different configurations leading to similar qualitative behaviours (Thom et al., 1977).

C.2 Solving the stability problem

As anticipated above, the state variable is identified by the angle θ, while the parameter α represents the ratio between the angle spanned by the second elastic w.r.t. the x axis and θ; ρ is the forcing (see Fig. C.1). To study the stability of the machine, it is sufficient to look at its potential. It is possible to study the Lagrangian (Nagy et al., 2013) or, in the simple static case, to consider the forces acting on the system (see Fig. C.3) and then integrate. We follow the second approach. Let us define:

- OP = 1. It is the wheel radius; its value is set to unitary measure without loss of generality.
- L_1 and L_2: lengths at rest of the two elastics.
- F_{el1} and F_{el2} are, respectively, the elastic forces stored by elastic 1 and 2; they follow Hook's law $F_{el} = -k_{\text{el}}\,\Delta L$. k_{el} is the elastic constant.
- T_m represents the machine's torque.

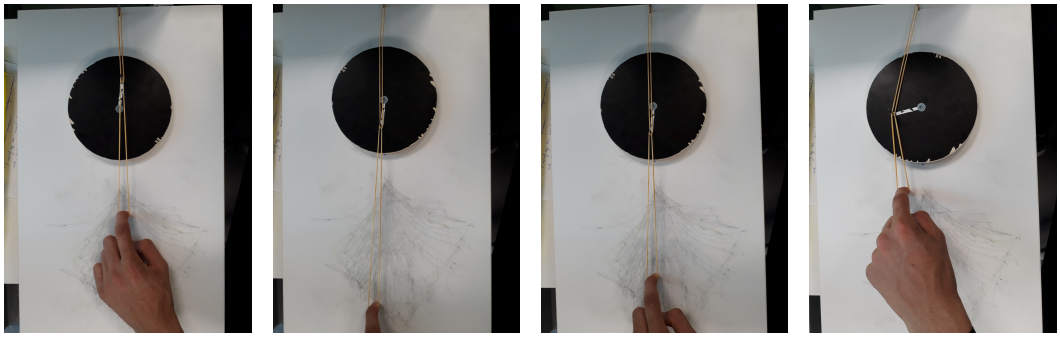
Let us also employ the long distance approximation: AB is almost parallel to the x axis (to select order of expansions for θ).

By applying simple mechanic considerations, we first estimate the potential term related to the elastic forces at rest. Start by:

$$F_{\text{el1}} = -k_{\text{el}} \cdot CD = k_{\text{el}}(1-\cos\theta)\,.$$

Considering the small angles approximation helping to truncate the Taylor expansion up to $O(\theta^3)$,

$$T \simeq F_{\text{el1}}\,\sin\theta = k_{\text{el}}(1-\cos\theta)\sin\theta \simeq k_{\text{el}}\left(1-1+\frac{\theta^2}{2}\right)\theta = k_{\text{el}}\,\frac{\theta^3}{2} + O(\theta^4)\,.$$



(a) (b) (c) (d)

Fig. C.2 Tipping and hysteresis can be observed in the evolution of the system (from left to right). **(a):** The machine is at equilibrium: a small forcing is not enough to move it. **(b):** the forcing overcomes a threshold, and the wheel turns rapidly to a new equilibrium. **(c):** such equilibrium is kept stable by the friction; if we let the elastic shrink, the wheel does not return to the initial position by passing through the previous point. **(d):** when a second threshold is passed, the wheelturns back to the first equilibrium. A video animation is provided, see link in Appendix D

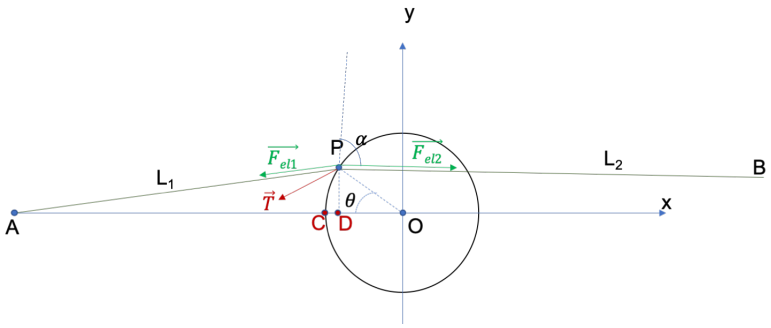


Fig. C.3 Sketch of the forces acting on the machine.

Hence, the potential is

$$V(\theta) \propto k_{\text{el}}\,\theta^4 + \text{perturbations}\,.$$

The "perturbations" are added to account for whatever moves the configuration from its resting state. We thus estimate the perturbative terms with respect to additional parameters:

- α quantifies the shift from the "$AB \parallel x$" assumption. By applying a Taylor expansion to the elastic force[1] we get (in the simplest, linear case) $F_{\text{pert}} \propto \alpha\theta$.
- ρ is directly the exerted force minus the friction.

As a consequence, the potential is:

$$V(\theta) \propto \theta^4 - \alpha\theta^2 + \rho\theta\,. \tag{C.1}$$

[1] $F(\theta+\Delta\theta) \simeq F(\theta) + \frac{dF}{d\theta}(\theta - 0) + O(\theta^2)$; if we assume (simplest form) that F is linearly dependent on θ (ansatz, for small angles) then we get $F(\theta+\Delta\theta) \simeq F(\theta) + \alpha\theta$

We can now play with this potential which is, up to numerical constants, the same one obtained using Langrangian methods (Poston et al., 1979). Numerical constants are neglected for two reasons: a) to concentrate on the effect of the parameters on the stability landscape, b) we here focus on qualitative predictions, therefore it is unnecessarily complex to take constants into account as the final result is independent on such details (Thom et al., 1977).

C.2.1 Analysis of the potential

Depending on the parameters values, the potential takes different shapes underlying different stability properties. This section studies the effect of bending parameter α and tilting parameter ρ on the potential and on its associated vector field $f(\theta,\bar{p}) = -\frac{dV(\theta,\bar{p})}{d\theta}$, $\bar{p}$ being the parameter set.

Bending parameter α

First, focus on the effect of α alone. It is called "bending parameter", as it describes how the second elastic is bended compared to θ. In addition, α has a bending effect on the potential (see Fig. C.4, right). Set $\rho = 0$ and consider :

$$V(\theta) = \theta^4 - \alpha\theta^2 \tag{C.2}$$

and

$$f(\theta) = 4\,\theta^3 - 2\,\alpha\theta\,. \tag{C.3}$$

α is akin to the control parameter of a supercritical pitchfork bifurcation (Sec A) governing the bistability and the degree of hysteresis of the system, namely the possibility to develop smooth (low alpha) or critical transitions (high alpha). In fact, Eq. C.3 represents the normal form of a vector field associated to a supercritical pitchfork bifurcation (also compare Fig. C.4 left with Fig. A.8).

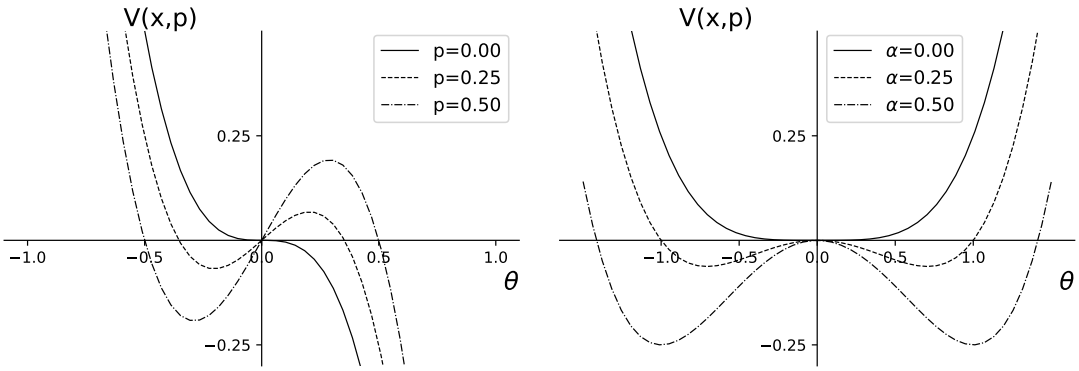


Fig. C.4 **Left:** $f(\theta)$ for various steady-state values of α. If the bending parameter is augmented from zero to any other value, the vector field is deformed from being monotonous to having three different regions of monotonicity. **Right:** when α increases, the potential goes from a single equilibrium point to three, of which two are stable and the latter is unstable. The system is develops bistability: it is now susceptible to undergo critical transitions. The process is exactly analogous to a pitchfork bifurcation, Sec A.

Tilting parameter ρ

ρ is termed "tilting" because it tilts the potential, see Fig. C.5. To empathize what happens to the equilibria, let us fix $\alpha = 1$ to be in a bistable configuration. Then, study:

$$V(\theta) = \theta^4 - \theta^2 + \rho\theta \tag{C.4}$$

and

$$f(\theta) = 4\theta^3 - 2\theta + \rho\,. \tag{C.5}$$

Increasing values of ρ cause the critical transition (provided that the system was already bistable), as one equilibrium suddenly loses stability and the other is pushed onto an alternative one. This follows the typical phenomena related to a fold catastrophe (saddle-node bifurcation, Sec A). In fact, in the neighbourhood of the second equilibrium, the vector field can be expanded in the normal form of a fold ($f_{\text{local}} \simeq -\rho - \theta^2$).

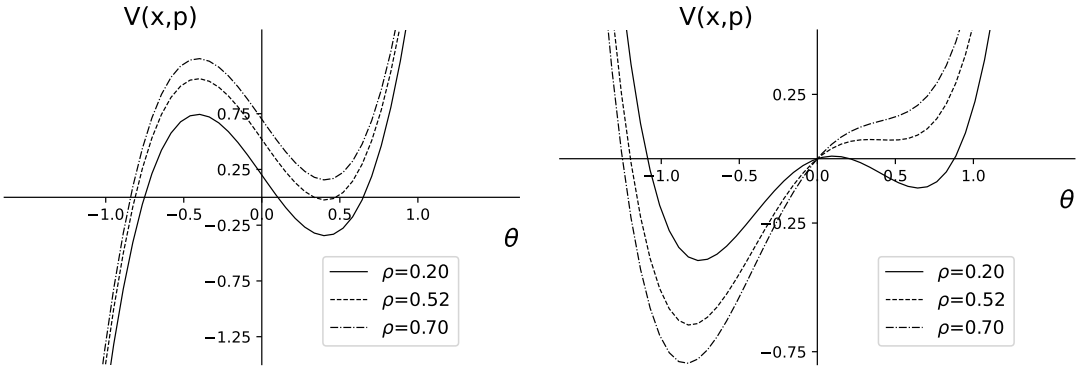


Fig. C.5 **Left:** The vector field is translated upwards under increasing ρ. Initially, three equilibria exist, then only two (of which one is stable and the other unstable). Finally, one equilibrium is destroyed and only one survivor hosts the system state. **Right:** the tilting parameter deforms to the potential as if it was tilted. One equilibrium goes from stability to instability and then disappears. Hence, any system state lying on the right branch will eventually transit critically towards the left-hand equilibrium. The process is exactly analogous to a fold bifurcation, Sec A.

Combining parameters

What happens, then, if we combine the effect of the two parameters? The potential, Eq. C.1, is bended and tilted at the same time (Fig. C.6), yielding a) asymmetric stable configurations (solid line), b) asymmetric bistable configurations (dashed), c) configurations with critical transitions happening at lower forcing ρ than before (dash-dotted). In general, the two parameters combined set the degree of hysteresis of the system and the possibility of observing a critical transition.

C.2.2 Analytical calculation

Since the explicit potential is known, it is possible to analytically calculate the values for which the tipping happens, as well as the functional form of the parameters relationship leading

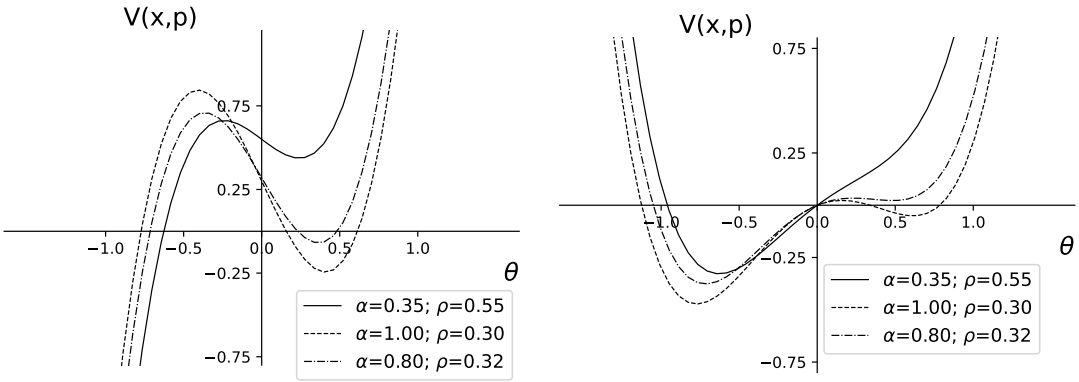


Fig. C.6 **Left:** Different $f(\theta)$ coming from different parameter combinations.**Right:** different $V(\theta)$ associated with different parameter combinations. The shape of the potential (ruled by α) is necessary to set the susceptibility of a system to tip, but the intersection with the zero axis is what drives the CT. So, both parameters should be taken into account while studying similar systems.

to critical transitions. The critical transition point is given by the intersection of the vector field manifold and of its tangent space. We can verify this fact in two ways:

1. Formally, consider the unfolding of the associated bifurcation, and study its associated vector field. Look for steady states and their stability. From this, identify the fold point and the stability of the associated manifold (Kuehn, 2013). The process is akin to solving the following system:
$$\begin{cases} f = 0 \\ D_x f|_{C_0} = 0 \end{cases} \tag{C.6}$$
where C_0 represents the fold manifold.

2. Intuitively, we need to check when a minimum in the potential coalesces into a flexus point (this is the condition for indifferent equilibrium). Hence, the system to be solved is:
$$\begin{cases} \frac{dV}{dx} = 0 \\ \frac{d^2V}{dx^2}|_{\hat{x}} = 0 \end{cases} \tag{C.7}$$
where the first equation finds the minima $\bar{x}$ and the second calculates a flexus point in the minima.

Obviously, the two formulations are equivalent in $1D$ (our case) thanks to the equivalence of the derivate operator and the manifold algebra (and do not forget that, in $1D$, $\frac{dV}{dx} = f$).

Critical Transition

Let us assume that the system is already at a bistable configuration (set $\alpha = 1$ w.l.o.g.). Then, estimate the critical value of ρ for which the system undergo a CT. Let us solve Eq. C.7. In our case,
$$\begin{cases} 4\hat{x}^3 - 2\hat{x} + \rho = 0 \\ 12\hat{x}^2 - 2 = 0\,. \end{cases} \tag{C.8}$$

The solution is:

$$\begin{cases} \hat{x}_{1,2} = \pm \left(\frac{1}{6}\right)^{1/2} \\ \rho_{CT}|_{\hat{x}_1} = \frac{4}{3}\left(\frac{1}{6}\right)^{1/2} \end{cases}. \tag{C.9}$$

$\hat{x}$ identifies the steady state points and ρ_{CT} is the critical value of the parameter for which a critical transition occurs. Those values have no real quantitative meaning when compared to the actual machine, as the potential was already rescaled to avoid constant dependencies on empirical values; it is difficult to get them from an artisanal machine, unless under perfectly controlled conditions. Nonetheless, this depicts the general procedure to study critical values in a model.

Rising bistability followed by Critical Transition

When both parameters are considered, we have to look at combinations for which a CT happens. Hence, we search for a set (α, ρ) for which the system is bistable and tipping. There are two parameters to consider, so we can look at the reduced curve (Kuehn, 2013) defined by $\rho(\alpha)$. Following the same procedure as before, thus solving system C.7 again, we eventually get, from $\hat{x}_1$:

$$\rho = \left(\frac{2}{3}\alpha\right)^{\frac{3}{2}} \tag{C.10}$$

as a solution. The other three solutions are given by $\hat{x}_2$ and by considering the case where $\theta \to \theta - \pi$ (corresponding to the opposite side of the pivot in the Catastrophe Machine). By plotting all four relationships $\rho = f(\alpha)$, we obtain the diamond-shaped plot shown in Fig. C.7 (left). This is the hallmark of a cusp bifurcation, see Sec. A and Fig. A.10, left. Fig. C.7 (left) predicts the set of points (qualitatively speaking) that is obtained empirically by playing with the machine and by pinpointing with a pencil when it tips (see Fig. C.7, right).

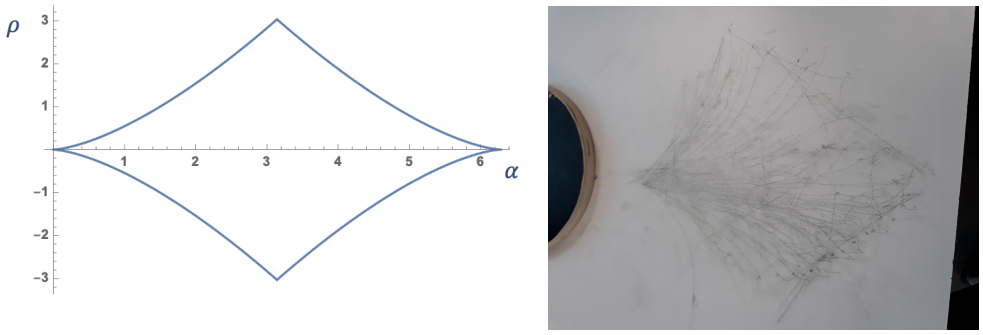


Fig. C.7 **Left:** The shape of $\rho = f(\alpha)$.**Right:** The diamond-shaped drawing that was empirically obtained by following the tipping behaviour of the machine with a pencil. It is not perfectly symmetric because of the friction and of elastic properties.

Two side notes on the analysis just performed. First, I wrote "qualitative" for the following reason: we are asking "does the machine exhibits CT? If so, under which conditions?" rather than "What is the exact value of the forcing after which the machine tips over"? Second, the present analysis was possible thanks to an available mechanistic model, developed from first principles of Newtonian dynamics. On the other hand, it might not be conclusive to observe a diamond shape on a board and infer that it was produced by a potential like C.1.

C.3 Limitations and challenges

Limitations: the catastrophe machine can be analysed thoroughly thanks first principles yielding a closed-form potential. Then, the emphasis on the two parameters controlling the system inscribes this study on the theory of catastrophes (Thom et al., 1977), that works fine for those systems whose potential is known. On the other hand, as several authors pointed out (Zahler et al., 1977; Kolata, 1977), it is not fruitful nor informative to speculate about the existence of such potential in any system that behaves vaguely similarly. It is the proponent who, in principle, should have the burden of proof, starting from experiments ad developing models. Using Catastrophe Theory as a mathematics-based philosophy could be misleading, as it detaches the modelling attempts from the experimental practice. On the other hand, such theory - and the Catastrophe Machine in particular - has already proven to be useful in controlled contexts and as an heuristic guide for experiments and speculation (hypothesis formulation). Further developments will also include diminishing the friction and coupling two wheels to obtain a chaotic behavior (Nagy et al., 2013). The latter case, which was initiated with a new machine (Fig. C.8) but was not further studied during the PhD period, will help to assess conditions for chaoticity rising from coupled nonlinear systems.

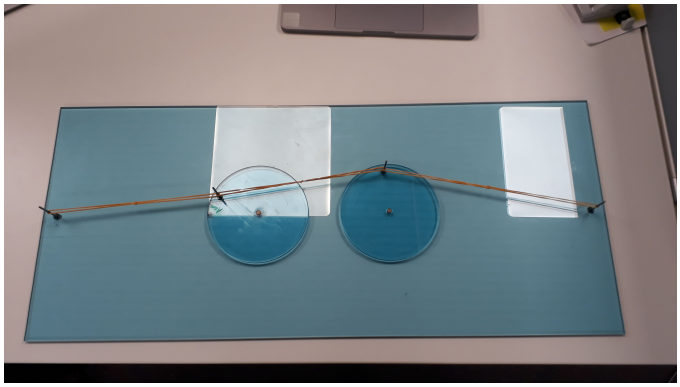


Fig. C.8 The upgraded Catastrophe Machine. It consists of two wheels, with increased distance from the pivots (to make the infinity approximation better), of which one can be adjusted in different positions. Moreover, friction was consistently reduced thanks to the material (plastic instead of wood).

Appendix D

Animations

The study of dynamical systems involve movies, rather than static pictures. Hence, a collection of animations and videos about normal forms and example systems accompanies this thesis. They are stored in the following Dropbox folder as supplementary material:

https://www.dropbox.com/sh/ryw36uz92nj04gs/AAAf-x2K-9YJ0tL6f23tflTMa?dl=0

Fig. D.1 QR-code linking the shared folder with supplementary animations.

List of contents

- cat_mach_tipping: video of tipping behavior exhibited by a catastrophe machine (Sec. C).
- cusp_vector-field: how the normal form vector field of a cusp bifurcation (Eq. A.20) changes under the effect of varying ρ and α.
- hopf_vector-field: how the normal form vector field of a supercritical Hopf bifurcation (b>1, corresponding to l in Eq. A.28) changes under the effect of varying a (corresponding to p in Eq. A.28).
- protein-production_vector-field: how $f(x,p)$ from the protein production model (Eq. 4.1, without noise) changes with varying p.

- saddle-node_bif-diag: exploring the bifurcation diagram of a saddle-node normal form (Eq. A.10) when p is increased; the point corresponds to system's state. Compare with Fig. A.2, right.
- saddle-node_vector-field: how the normal form vector field of a saddle-node bifurcation changes under the effect of varying p. Compare with Fig. A.2, left.
- saddle-node_potential: how the normal form potential of a saddle-node bifurcation changes under the effect of varying p. Compare with Fig. A.2, centre.
- subcritical-pitchfork_bif-diag: exploring the bifurcation diagram of a subcritical pitchfork normal form (Eq. A.14) when p is increased; the point corresponds to system's state. Compare with Fig. A.6.
- subcritical-pitchfork_vector-field: how the normal form vector field of a subcritical pitchfork bifurcation changes under the effect of varying p. Compare with Fig. A.5.
- subcritical-pitchfork_ potential: how the normal form potential of a (global) subcritical pitchfork bifurcation changes under the effect of varying p. Compare with Fig. A.7, centre.
- supercritical-pitchfork_bif-diag: exploring the bifurcation diagram of a supercritical pitchfork normal form (Eq.A.17) when p is increased; the point corresponds to system's state. Compare with Fig. A.9, left.
- supercritical-pitchfork_vector-field: how the normal form vector field of a supercritical pitchfork bifurcation changes under the effect of varying p. Compare with Fig. A.8.
- supercritical -pitchfork_ potential: how the normal form potential of a supercritical pitchfork bifurcation changes under the effect of varying p. Compare with Fig. A.9, right.
- transcritical-pitchfork_bif-diag: exploring the bifurcation diagram of a transcritical normal form (Eq. A.12) when p is increased; the point corresponds to system's state. Compare with Fig. A.4.
- transcritical_vector-field: how the normal form vector field of a transcritical bifurcation changes under the effect of varying p. Compare with Fig. A.3.
- transcritical_ potential: how the normal form potential of a transcritical bifurcation changes under the effect of varying p.

Acknowedgments

Animations were made with Wolfram Mathematica. Plots of phase planes are based on: Suba Thomas "Bifurcation Diagrams with Flow Fields",
http://demonstrations.wolfram.com/BifurcationDiagramsWithFlowFields/, Wolfram Demonstrations Project, Published: March 7 2011

References

Aalto, A. et al. (2020). "Gene regulatory network inference from sparsely sampled noisy data". In: *Nature communications* 11.1, pp. 1–9. DOI: 10.1038/s41467-020-17217-1 (cit. on pp. 22, 177).

Aarabi, A. and B. He (2014). "Seizure prediction in hippocampal and neocortical epilepsy using a model-based approach". In: *Clinical Neurophysiology* 125.5, pp. 930–940. DOI: 10.1016/j.clinph.2013.10.051 (cit. on p. 57).

Abbott, S et al. (2020). "Temporal variation in transmission during the COVID-19 outbreak". In: *Wellcome Open Res* 5.112 (cit. on p. 138).

Abdel-Hamid, L. et al. (2016). "Retinal image quality assessment based on image clarity and content". In: *Journal of Biomedical Optics* 21.9, p. 096007. DOI: 10.1117/1.jbo.21.9.096007 (cit. on pp. 79, 177, 206).

Acar, M., A. Becskei, and A. Van Oudenaarden (2005). "Enhancement of cellular memory by reducing stochastic transitions". In: *Nature* 435.7039, pp. 228–232. DOI: 10.1038/nature03524 (cit. on p. 61).

Adam, D. (2020). "A guide to R—the pandemic's misunderstood metric". In: *Nature* 583.7816, pp. 346–348 (cit. on p. 149).

Aguirre, J. and S. Manrubia (2015). "Tipping points and early warning signals in the genomic composition of populations induced by environmental changes". In: *Scientific Reports* 5, pp. 1–7. DOI: 10.1038/srep09664 (cit. on p. 21).

Ahmed, W. et al. (2021). "Wastewater surveillance demonstrates high predictive value for COVID-19 infection on board repatriation flights to Australia". In: *Environ. Int.* P. 106938. DOI: https://doi.org/10.1016/j.envint.2021.106938 (cit. on p. 152).

Aihara, K. et al. (2022). "Dynamical network biomarkers: Theory and applications". In: *Gene* 808, p. 145997. DOI: 10.1016/j.gene.2021.145997 (cit. on pp. 21, 76, 88, 127).

Ajelli, M. et al. (2016). "Spatiotemporal dynamics of the Ebola epidemic in Guinea and implications for vaccination and disease elimination: a computational modeling analysis". In: *BMC medicine* 14.1, pp. 1–10. DOI: 0.1186/s12916-016-0678-3 (cit. on p. 138).

Alkhayuon, H. et al. (2019). "Basin bifurcations, oscillatory instability and rate-induced thresholds for Atlantic meridional overturning circulation in a global oceanic box model". In: *Proceedings of the Royal Society A* 475.2225, pp. 1–31. DOI: 10.1098/rspa.2019.0051 (cit. on p. 17).

Allen, L. J. S. (2010). *An introduction to stochastic processes with applications to biology*. CRC press (cit. on pp. 90, 91, 93, 94, 135, 203).

Alon, U. (2019). *An introduction to systems biology: design principles of biological circuits*. CRC press (cit. on pp. 41, 61, 62, 90, 111).

Alonso, D., A. J. McKane, and M. Pascual (2007). "Stochastic amplification in epidemics". In: *Journal of the Royal Society Interface* 4.14, pp. 575–582. DOI: 10.1098/rsif.2006.0192 (cit. on pp. 133, 134).

Alonso, D., A. Dobson, and M. Pascual (2019). "Critical transitions in malaria transmission models are consistently generated by superinfection". In: *Philosophical Transactions of the Royal Society B* 374.1775, pp. 1–19. DOI: 10.1098/rstb.2018.0275 (cit. on p. 18).

Althaus, C. L. (2014). "Estimating the reproduction number of Ebola virus (EBOV) during the 2014 outbreak in West Africa". In: *PLoS Currents* 6. DOI: 10.1371/currents.outbreaks.91afb5e0f279e7f29e7056095255b288 (cit. on pp. 65, 131).

Altig, D. et al. (2020). "Economic uncertainty before and during the COVID-19 pandemic". In: *Journal of Public Economics* 191, p. 104274. DOI: 10.1016/j.jpubeco.2020.104274 (cit. on p. 65).

Alvarez, F. E., D. Argente, and F. Lippi (2020). *A simple planning problem for COVID-19 lockdown*. Tech. rep. National Bureau of Economic Research (cit. on p. 65).

Anand, N. et al. (2020). "Predicting the Spread of COVID-19 Using SIR Model Augmented to Incorporate Quarantine and Testing". In: *Transactions of the Indian National Academy of Engineering* 5.2, pp. 141–148 (cit. on p. 64).

Anderson, R. M. and R. M. May (1979). "Population biology of infectious diseases: Part I". In: *Nature* 280.5721, pp. 361–367. DOI: https://doi.org/10.1038/280361a0 (cit. on pp. 64, 132).

Anderson, R. M. et al. (2020). "How will country-based mitigation measures influence the course of the COVID-19 epidemic?" In: *The Lancet* 395.10228, pp. 931–934. DOI: https://doi.org/10.1016/S0140-6736(20)30567-5 (cit. on p. 64).

Andrecut, M. et al. (2011). "A general model for binary cell fate decision gene circuits with degeneracy: Indeterminacy and switch behavior in the absence of cooperativity". In: *PLoS ONE* 6.5, e19358. DOI: 10.1371/journal.pone.0019358 (cit. on p. 40).

Angeli, D. and E. D. Sontag (2003). "Monotone control systems". In: *IEEE Transactions on Automatic Control* 48.10, pp. 1684–1698. DOI: 10.1109/TAC.2003.817920 (cit. on p. 61).

Angeli, D., J. E. Ferrell, and E. D. Sontag (2004a). "Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems". In: *Proceedings of the National Academy of Sciences of the United States of America* 101.7, pp. 1822–1827. DOI: 10.1073/pnas.0308265100 (cit. on pp. 18, 47, 48).

Angeli, D. and E. D. Sontag (2004b). "Multi-stability in monotone input/output systems". In: *Systems and Control Letters* 51.3-4, pp. 185–202. DOI: 10.1016/j.sysconle.2003.08.003 (cit. on pp. 48, 112).

Aparicio, A. et al. (2021). "Structure-based identification of sensor species for anticipating critical transitions". In: *PNAS* 118.51, e2104732118 (cit. on p. 179).

Arani, B. M. et al. (2021). "Exit time as a measure of ecological resilience". In: *Science* 372.6547. DOI: 10.1126/science.aay4895 (cit. on p. 211).

Arino, J. et al. (2005). "A multi-species epidemic model with spatial dynamics". In: *Mathematical Medicine and Biology* 22.2, pp. 129–142. DOI: https://doi.org/10.1093/imammb/dqi003 (cit. on p. 64).

Arnold, L. and R. Lefever (1983). *Stochastic Nonlinear Systems in Physics, Chemistry, and Biology*. Springer. DOI: 10.1007/978-3-642-12601-7 (cit. on p. 213).

Ashikhmin, V. N. et al. (2004). *Introduction to the mathematical modelling'*. Limited Liability Company "Logos" (cit. on p. 5).

Ashwin, P. et al. (2012). "Tipping points in open systems : bifurcation , noise-induced and rate-dependent examples". In: *Philosophical Transactions of the Royal Society A* 370.1962, pp. 1166–1184. DOI: 10.1098/rsta.2011.0306 (cit. on pp. 17, 20, 30, 31, 33, 133).

Ashwin, P. and A. Zaikin (2015). "Pattern selection: The importance of "how you get there"". In: *Biophysical Journal* 108.6, pp. 1307–1308. DOI: 10.1016/j.bpj.2015.01.036 (cit. on p. 9).

Ashwin, P., S. Coombes, and R. Nicks (2016). "Mathematical Frameworks for Oscillatory Network Dynamics in Neuroscience". In: *Journal of Mathematical Neuroscience* 6.1, pp. 1–92. DOI: 10.1186/s13408-015-0033-6 (cit. on p. 57).

Ashwin, P., J. Creaser, and K. Tsaneva-Atanasova (2017a). "Fast and slow domino regimes in transient network dynamics". In: *Physical Review E* 96.5, p. 52309. DOI: 10.1103/PhysRevE.96.052309 (cit. on p. 179).

Ashwin, P., C. Perryman, and S. Wieczorek (2017b). "Parameter shifts for nonautonomous systems in low dimension: Bifurcation- and rate-induced tipping". In: *Nonlinearity* 30.6, pp. 2185–2210. DOI: 10.1088/1361-6544/aa675b (cit. on pp. 28, 33).

Åström, K. J. and R. M. Murray (2010). *Feedback systems*. Princeton university press (cit. on p. 39).

Atkeson, A. (2020). *What will be the economic impact of COVID-19 in the us? rough estimates of disease scenarios*. Tech. rep. National Bureau of Economic Research (cit. on p. 65).

Attneave, F. (1971). "Multistability in perception". In: *Scientific American* 225.6, pp. 62–71 (cit. on p. 39).

Aur, D. (2010). "The Physical Mechanism in Epilepsy - Understanding the Transition to Seizure". In: *Comparative and General Pharmacology* (cit. on p. 57).

Babinec, P, M Kučera, and M Babincová (2005). "Global characterization of time series using fractal dimension of corresponding recurrence plots: from dynamical systems to heart physiology". In: *Harmon. Fractal Image Anal.* 1, pp. 87–93 (cit. on p. 75).

Bak, P. (1996). *How nature works*. Springer Science \& Business Media (cit. on p. 39).

Bak, P. (1987). "Self-organized criticality:An explanation of the 1-f noise". In: *Physical Review Letters* 59.4, pp. 381–384. DOI: 10.1103/PhysRevLett.59.381 (cit. on pp. 39, 56).

Balázsi, G., A. H. Cornell-Bell, and F. Moss (2003). "Increased phase synchronization of spontaneous calcium oscillations in epileptic human versus normal rat astrocyte cultures". In: *Chaos* 13.2, pp. 515–518. DOI: 10.1063/1.1567652 (cit. on p. 57).

Bandala, E. R. et al. (2021). "Impacts of {COVID-19} pandemic on the wastewater pathway into surface water: A review". In: *Science of the Total Environment*, p. 145586. DOI: https://doi.org/10.1016/j.scitotenv.2021.145586 (cit. on p. 152).

Baniasadi, M. et al. (2020). "FastField: An open-source toolbox for efficient approximation of deep brain stimulation electric fields". In: *NeuroImage* 223.February, p. 117330. DOI: 10.1016/j.neuroimage.2020.117330 (cit. on p. 16).

Bar-Yam, Y. (2002). "General features of complex systems". In: *Encyclopedia of Life Support Systems (EOLSS), UNESCO, EOLSS Publishers, Oxford, UK*, p. 1 (cit. on p. 8).

Barabási, A.-L. and Others (2016). *Network science*. Cambridge university press (cit. on p. 198).

Barnett, L. et al. (2013). "Information flow in a kinetic Ising model peaks in the disordered phase". In: *Physical review letters* 111.17, p. 177203. DOI: 10.1103/PhysRevLett.111.177203 (cit. on p. 38).

Bathe, K. J. and E. L. Wilson (1972). "Stability and accuracy analysis of direct integration methods". In: *Earthquake Engineering & Structural Dynamics* 1.3, pp. 283–291. DOI: 10.1002/eqe.4290010308 (cit. on p. 67).

Bauchner, H., R. M. Golub, and J. Zylke (2020). "Editorial concern—possible reporting of the same patients with COVID-19 in different reports". In: *Jama* 323.13, p. 1256. DOI: 10.1001/jama.2020.3980 (cit. on p. 141).

Bayram, M., T. Partal, and G. O. Buyukoz (2018). "Numerical methods for simulation of stochastic differential equations". In: *Advances in Difference Equations* 2018.1, pp. 1–10. DOI: 10.1186/s13662-018-1466-5 (cit. on pp. 67, 68).

Beisner, B. E., D. T. Haydon, and K. Cuddington (2003). "Alternative stable states in ecology". In: *Frontiers in Ecology and the Environment* 1.7, pp. 376–382. DOI: 10.1890/1540-9295(2003)001[0376:ASSIE]2.0.CO;2 (cit. on pp. 20, 40).

Bélair, J. et al. (1995). "Dynamical disease: Identification, temporal aspects and treatment strategies of human illness". In: *Chaos* 5.1, pp. 1–7. DOI: 10.1063/1.166069 (cit. on pp. 13, 18).

Benner, S. A. and A. M. Sismour (2005). "Synthetic biology". In: *Nature Reviews Genetics* 6.7, pp. 533–543. DOI: 10.1038/nrg1637 (cit. on p. 5).

Berglund, N. and B. Gentz (2006). *Noise-induced phenomena in slow-fast dynamical systems: a sample-paths approach*. Springer Science \& Business Media (cit. on pp. 19, 29, 90, 113, 135, 181, 183, 198, 203, 204, 213).

Bettencourt, L. M. A. and R. M. Ribeiro (2008). "Real time bayesian estimation of the epidemic potential of emerging infectious diseases". In: *PloS one* 3.5, e2185. DOI: 10.1371/journal.pone.0002185 (cit. on p. 138).

Bhamra, R., S. Dani, and K. Burnard (2011). "Resilience: the concept, a literature review and future directions". In: *International journal of production research* 49.18, pp. 5375–5393. DOI: 10.1080/00207543.2011.563826 (cit. on p. 20).

Bhandary, S. et al. (2021). "Network resilience of FitzHugh-Nagumo neurons in the presence of nonequilibrium dynamics". In: *Physical Review E* 103.2, pp. 1–12. DOI: 10.1103/PhysRevE.103.022314 (cit. on p. 178).

Bibby, K. et al. (2021). "Making Waves: Plausible Lead Time for Wastewater Based Epidemiology as an Early Warning System for {COVID-19}". In: *Water Research*, p. 117438. DOI: https://doi.org/10.1016/j.watres.2021.117438 (cit. on p. 169).

Bick, C. et al. (2021). "What are higher-order networks?" In: *arXiv preprint arXiv:2104.11329* (cit. on pp. 50, 178).

Biggs, N., N. L. Biggs, and B. Norman (1993). *Algebraic graph theory*. Vol. 67. Cambridge university press (cit. on p. 48).

Biggs, R., S. R. Carpenter, and W. A. Brock (2009). "Turning back from the brink: detecting an impending regime shift in time to avert it". In: *Proceedings of the National academy of Sciences* 106.3, pp. 826–831. DOI: 10.1073/pnas.0811729106 (cit. on pp. 21, 53, 140, 209).

Biggs, R., G. D. Peterson, and J. C. Rocha (2018). "The Regime Shifts Database: a framework for analyzing regime shifts in social-ecological systems". In: (cit. on p. 17).

Binder, K. (1987). "Theory of first-order phase transitions". In: *Reports on Progress in Physics* 50, pp. 783–859. DOI: 10.1088/0034-4885/50/7/001 (cit. on p. 38).

Binney, J. et al. (1992). *The theory of critical phenomena - an introduction to renormalization group*. New York: Oxford University Press (cit. on p. 38).

Blackwood, J. C. and L. M. Childs (2018). "An introduction to compartmental modeling for the budding infectious disease modeler". In: *Letters in Biomathematics* 5.1, pp. 195–221. DOI: 10.1080/23737867.2018.1509026 (cit. on p. 65).

Bloch, C. and C. De Dominicis (1958). "Un développement du potentiel de gibbs d'un système quantique composé d'un grand nombre de particules". In: *Nuclear Physics* 7, pp. 459–479. DOI: 10.1016/0029-5582(58)90285-2 (cit. on p. 100).

Boccaletti, S. et al. (2006). "Complex networks: Structure and dynamics". In: *Physics reports* 424.4-5, pp. 175–308. DOI: 10.1016/j.physrep.2005.10.009 (cit. on pp. 30, 57, 198).

Boers, N. (2021). "Observation-based early-warning signals for a collapse of the Atlantic Meridional Overturning Circulation". In: *Nature Climate Change* 11.8, pp. 680–688. DOI: 10.1038/s41558-021-01097-4 (cit. on p. 21).

Boers, N. and M. Rypdal (2021). "Critical slowing down suggests that the western Greenland Ice Sheet is close to a tipping point". In: *Proceedings of the National Academy of Sciences of the United States of America* 118.21, pp. 1–7. DOI: 10.1073/pnas.2024192118 (cit. on p. 17).

Boettiger, C. and A. Hastings (2012a). "Early warning signals and the prosecutor's fallacy". In: *Proceedings of the Royal Society B* 279.1748, pp. 4734–4739. DOI: 10.1098/rspb.2012.2085 (cit. on pp. 46, 52, 59, 80, 122).

— (2012b). "Quantifying limits to detection of early warning for critical transitions". In: *Journal of the Royal Society Interface* 9.75, pp. 2527–2539. DOI: 10.1098/rsif.2012.0125 (cit. on pp. 52, 60).

Boettiger, C., N. Ross, and A. Hastings (2013a). "Early warning signals: The charted and uncharted territories". In: *Theoretical Ecology* 6.3, pp. 255–264. DOI: 10.1007/s12080-013-0192-6 (cit. on pp. 51, 52, 60).

Boettiger, C. and A. Hastings (2013b). "From patterns to predictions". In: *Nature* 493.7431, pp. 157–158. DOI: 10.1038/493157a (cit. on p. 60).

— (2013c). "No early warning signals for stochastic transitions: Insights from large deviation theory". In: *Proceedings of the Royal Society B* 280.1766, pp. 10–12. DOI: 10.1098/rspb.2013.1372 (cit. on p. 51).

Boffetta, G. et al. (1999). "Power laws in solar flares: Self-organized criticality or turbulence?" In: *Physical Review Letters* 83.22, pp. 4662–4665. DOI: 10.1103/PhysRevLett.83.4662 (cit. on p. 56).

Bokes, P., Y. T. Lin, and A. Singh (2018). "High Cooperativity in Negative Feedback can Amplify Noisy Gene Expression". In: *Bulletin of Mathematical Biology* 80.7, pp. 1871–1899. DOI: 10.1007/s11538-018-0438-y (cit. on p. 39).

Bolt, B. van der et al. (2018). "Climate reddening increases the chance of critical transitions". In: *Nature Climate Change* 8.6, pp. 478–484. DOI: 10.1038/s41558-018-0160-7 (cit. on p. 88).

Bonciolini, G. et al. (2018). "Experiments and modelling of rate-dependent transition delay in a stochastic subcritical bifurcation". In: *Royal Society Open Science* 5.3, p. 172078. DOI: 10.1098/rsos.172078 (cit. on pp. 34, 44).

Borshchev, A. and A. Filippov (2004). "From System Dynamics and Discrete Event to Practical Agent Based Modeling: Reasons, Techniques, Tools". In: *Simulation* 66.11, pp. 25–29 (cit. on p. 5).

Bradford, K. J. and A. J. Trewavas (1994). "Sensitivity thresholds and variable time scales in plant hormone action." In: *Plant Physiology* 105.4, p. 1029. DOI: 10.1104/pp.105.4.1029 (cit. on p. 89).

Breakspear, M. et al. (2006). "A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis". In: *Cerebral Cortex* 16.9, pp. 1296–1313. DOI: 10.1093/cercor/bhj072 (cit. on p. 57).

Brett, T. et al. (2020a). "Detecting critical slowing down in high-dimensional epidemiological systems". In: *PLoS computational biology* 16.3, e1007679. DOI: 10.1371/journal.pcbi.1007679 (cit. on pp. 21, 82, 131, 135, 138, 142, 146).

Brett, T. S., J. M. Drake, and P. Rohani (2017). "Anticipating the emergence of infectious diseases". In: *Journal of The Royal Society Interface* 14.132, p. 20170115. DOI: 10.1098/rsif.2017.0115 (cit. on pp. 18, 21, 29, 52–54, 65, 80, 89, 109, 131, 135, 140, 210).

Brett, T. S. et al. (2018). "Anticipating epidemic transitions with imperfect data". In: *PLoS computational biology* 14.6, e1006204. DOI: 10.1371/journal.pcbi.1006204 (cit. on pp. 65, 131).

Brett, T. S. and P. Rohani (2020b). "Dynamical footprints enable detection of disease emergence". In: *PLoS biology* 18.5, e3000697. DOI: 10.1371/journal.pbio.3000697 (cit. on pp. 53, 65, 131, 150, 178).

Brochini, L. et al. (2016). "Phase transitions and self-organized criticality in networks of stochastic spiking neurons". In: *Scientific reports* 6.1, pp. 1–15. DOI: 10.1038/srep35831 (cit. on p. 56).

Brooks, S. and A. Gelman (1998). "General Methods for Monitoring Convergence of Iterative Simulations". In: *J. Comput. Graphi. Stat.* 7, pp. 434–455. DOI: 10.1080/10618600.1998.10474787 (cit. on p. 69).

Brownlee, J. (2017). *Introduction to Time Series Forecasting With Python*. Machine Learning Mastery, p. 367 (cit. on p. 78).

Bruggeman, F. J. and B. Teusink (2018). "Living with noise: on the propagation of noise from molecules to phenotype and fitness". In: *Current Opinion in Systems Biology* 8, pp. 144–150. DOI: 10.1016/j.coisb.2018.02.010 (cit. on pp. 88, 109, 125).

Budišić, M., R. Mohr, and I. Mezić (2012). "Applied koopmanism". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 22.4, p. 47510. DOI: 10.1063/1.4772195 (cit. on p. 30).

Burrus, C. S. (1997). "Introduction to wavelets and wavelet transforms: a primer". In: *Englewood Cliffs* (cit. on p. 177).

Burthe, S. J. et al. (2016). "Do early warning indicators consistently predict nonlinear change in long-term ecological data?" In: *Journal of Applied Ecology* 53.3, pp. 666–676. DOI: 10.1111/1365-2664.12519 (cit. on pp. 45, 52).

Bury, T. M., C. T. Bauch, and M. Anand (2020). "Detecting and distinguishing tipping points using spectral early warning signals". In: *Journal of the Royal Society Interface* 17.170, p. 20200482. DOI: 10.1098/rsif.2020.0482 (cit. on pp. 21, 35, 205).

Bury, T. M. et al. (2021). "Deep learning for early warning signals of tipping points". In: *Proceedings of the National Academy of Sciences of the United States of America* 118.39, e2106140118. DOI: 10.1073/pnas.2106140118 (cit. on pp. 21, 53, 54, 58, 176, 178).

Burzyński, M. et al. (2021). "COVID-19 crisis management in Luxembourg: Insights from an epidemio-nomic approach". In: *Economics and Human Biology* 43, p. 101051. DOI: 10.1016/j.ehb.2021.101051 (cit. on pp. 15, 65, 129, 130).

Byrd, T. A. et al. (2019). "Critical slowing down in biochemical networks with feedback". In: *Physical Review E* 100.2, p. 22415. DOI: 10.1103/PhysRevE.100.022415 (cit. on pp. 21, 112).

Cacciapaglia, G., C. Cot, and F. Sannino (2020). "Second wave COVID-19 pandemics in Europe: a temporal playbook". In: *Scientific reports* 10.1, pp. 1–8. DOI: 10.1038/s41598-020-72611-5 (cit. on p. 131).

Cahan, P. et al. (2021). "Computational Stem Cell Biology: Open Questions and Guiding Principles". In: *Cell Stem Cell* 28.1, pp. 20–32. DOI: 10.1016/j.stem.2020.12.012 (cit. on p. 5).

Camazine, S. (2003). *Self-organization in biological systems*. Princeton University Press (cit. on p. 9).

Cao, Y. and R. Francis (2021). "On forecasting the community-level COVID-19 cases from the concentration of SARS-CoV-2 in wastewater". In: *Sci. Total Environ.* 786, p. 147451. DOI: https://doi.org/10.1016/j.scitotenv.2021.147451 (cit. on pp. 152, 153, 168).

Cao, Z. and R. Grima (2018). "Linear mapping approximation of gene regulatory networks with stochastic dynamics". In: *Nature communications* 9.1, pp. 1–15. DOI: 10.1038/s41467-018-05822-0 (cit. on p. 113).

Carney, P. R., S. Myers, and J. D. Geyer (2011). "Seizure prediction: Methods". In: *Epilepsy and Behavior* 22.SUPPL. 1, S94–S101. DOI: 10.1016/j.yebeh.2011.09.001 (cit. on p. 57).

Carpenter, S. R. et al. (2011). "Early warnings of regime shifts: A whole-ecosystem experiment". In: *Science* 332.6033, pp. 1079–1082. DOI: 10.1126/science.1203672 (cit. on pp. 21, 44, 49).

Carpenter, S. R. and M. L. Pace (1997). "Dystrophy and eutrophy in lake ecosystems: implications of fluctuating inputs". In: *Oikos*, pp. 3–14. DOI: 10.2307/3545794 (cit. on p. 17).

Chapeau-Blondeau, F. (2007). "Autocorrelation versus entropy-based autoinformation for measuring dependence in random signal". In: *Physica A: Statistical Mechanics and its Applications* 380.1-2, pp. 1–18. DOI: 10.1016/j.physa.2007.02.077 (cit. on p. 206).

Chen, D.-q. and L.-z. Cheng (2012a). *Multiplicative Noise* (cit. on p. 45).

Chen, L. et al. (2012b). "Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers". In: *Scientific Reports* 2, pp. 18–20. DOI: 10.1038/srep00342 (cit. on pp. 21, 45, 76, 109).

Chen, S. et al. (2019). "Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems". In: *Scientific Reports* 9.1, pp. 1–14. DOI: 10.1038/s41598-019-38961-5 (cit. on pp. 21, 76, 178).

Chen, X., Y. M. Kang, and Y. X. Fu (2017). "Switches in a genetic regulatory system under multiplicative non-Gaussian noise". In: *Journal of Theoretical Biology* 435, pp. 134–144. DOI: 10.1016/j.jtbi.2017.09.010 (cit. on p. 106).

Cheng, X. and J. M. A. Scherpen (2021). "Model Reduction Methods for Complex Network Systems". In: *Annual Review of Control, Robotics, and Autonomous Systems* 4.1, pp. 425–453. DOI: 10.1146/annurev-control-061820-083817 (cit. on p. 85).

Chouzouris, T. et al. (2018). "Chimera states in brain networks: Empirical neural vs. modular fractal connectivity". In: *Chaos* 28.4, pp. 1–11. DOI: 10.1063/1.5009812 (cit. on p. 57).

Chu, H. et al. (2017). "Predicting epileptic seizures from scalp EEG based on attractor state analysis". In: *Computer Methods and Programs in Biomedicine* 143, pp. 75–87. DOI: 10.1016/j.cmpb.2017.03.002 (cit. on pp. 21, 57).

Cimatoribus, A. A. et al. (2013). "Dansgaard-Oeschger events: Bifurcation points in the climate system". In: *Climate of the Past* 9.1, pp. 323–333. DOI: 10.5194/cp-9-323-2013 (cit. on p. 57).

Clements, C. F. and A. Ozgul (2016a). "Including trait-based early warning signals helps predict population collapse". In: *Nature Communications* 7, pp. 1–8. DOI: 10.1038/ncomms10984 (cit. on p. 21).

— (2016b). "Rate of forcing and the forecastability of critical transitions". In: *Ecology and Evolution* 6.21, pp. 7787–7793. DOI: 10.1002/ece3.2531 (cit. on p. 80).

— (2018). "Indicators of transitions in biological systems". In: *Ecology Letters* 21.6, pp. 905–919. DOI: 10.1111/ele.12948 (cit. on pp. 21, 60, 151, 174).

Cleveland, R. B. et al. (1990). "STL: A seasonal-trend decomposition". In: *J. Off. Stat* 6.1, pp. 3–73 (cit. on p. 77).

Cluzel, N. et al. (2022). "A nationwide indicator to smooth and normalize heterogeneous {SARS-CoV-2 RNA} data in wastewater". In: *Environment International* 158, p. 106998. DOI: https://doi.org/10.1016/j.envint.2021.106998 (cit. on pp. 153, 162).

Cocchi, L. et al. (2017). "Criticality in the brain: A synthesis of neurobiology, models and cognition". In: *Progress in neurobiology* 158, pp. 132–152. DOI: 10.1016/j.pneurobio.2017.07.002 (cit. on p. 56).

Cohen, A. A. et al. (2022). "Synchrony of biomarker variability indicates a critical transition: Application to mortality prediction in hemodialysis". In: *iScience* 25, p. 104385. DOI: 10.1016/j.isci.2022.104385 (cit. on p. 18).

Coifman, R. R. and M. V. Wickerhauser (1992). "Entropy-based algorithms for best basis selection". In: *IEEE Transactions on Information Theory* 38.2, pp. 713–718. DOI: 10.1109/18.119732 (cit. on pp. 79, 177, 178, 206).

Cotilla-Sanchez, E., P. D. Hines, and C. M. Danforth (2012). "Predicting critical transitions from time series synchrophasor data". In: *IEEE Transactions on Smart Grid* 3.4, pp. 1832–1840. DOI: 10.1109/TSG.2012.2213848 (cit. on p. 21).

Courbariaux, M. et al. (2022). "A flexible smoother adapted to censored data with outliers and its application to {SARS-CoV-2} monitoring in wastewater". In: *Frontiers in Applied Mathematics and Statistics* 8. DOI: https://doi.org/10.3389/fams.2022.836349 (cit. on p. 153).

Crawford, J. D. (1991). "Introduction to bifurcation theory". In: *Reviews of Modern Physics* 63.4, pp. 991–1037. DOI: 10.1103/RevModPhys.63.991 (cit. on p. 185).

Cuevas, E. (2020). "An agent-based model to evaluate the COVID-19 transmission risks in facilities". In: *Computers in biology and medicine* 121, p. 103827. DOI: 10.1016/j.compbiomed.2020.103827 (cit. on p. 130).

Currie, C. S. M. et al. (2020). "How simulation modelling can help reduce the impact of COVID-19". In: *Journal of Simulation* 14.2, pp. 83–97. DOI: 10.1080/17477778.2020.1751570 (cit. on p. 64).

Dablander, F. et al. (2021). "Overlapping Time Scales Obscure Early Warning Signals of the Second COVID-19 Wave". In: *medRxiv*. DOI: 10.1101/2021.07.27.21261226 (cit. on pp. 21, 131).

Dablander, F. et al. (2022). "Anticipating Critical Transitions in Psychological Systems using Early Warning Signals : Theoretical and Practical Considerations Theory of Critical Slowing Down". In: *Psychological Methods*. DOI: 10.31234/osf.io/5wc28 (cit. on p. 80).

Dai, L. et al. (2012). "Generic indicators for loss of resilience before a tipping point leading to population collapse". In: *Science* 336.6085, pp. 1175–1177. DOI: 10.1126/science.1219805 (cit. on pp. 44, 88, 99, 100, 109, 124).

Dai, L., K. S. Korolev, and J. Gore (2013). "Slower recovery in space before collapse of connected populations". In: *Nature* 496.7445, pp. 355–358. DOI: 10.1038/nature12071 (cit. on pp. 21, 124).

Dai, L. et al. (2015). "Relation between stability and resilience determines the performance of early warning signals under different environmental drivers". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.32, pp. 10056–10061. DOI: 10.1073/pnas.1418415112 (cit. on pp. 20, 21, 28, 38, 111, 112, 124–126, 175).

Dakos, V. et al. (2008). "Slowing down as an early warning signal for abrupt climate change". In: *Proceedings of the National Academy of Sciences* 105.38, pp. 14308–14312. DOI: 10.1353/elh.2014.0029 (cit. on pp. 21, 35, 45, 76).

Dakos, V. et al. (2012a). "Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data". In: *PloS one* 7.7, e41010. DOI: 10.1371/journal.pone.0041010 (cit. on pp. 21, 51, 79, 132).

Dakos, V. et al. (2012b). "Robustness of variance and autocorrelation as indicators of critical slowing down". In: *Ecology* 93.2, pp. 264–271. DOI: 10.1890/11-0889.1 (cit. on pp. 35, 88).

Dakos, V., E. H. van Nes, and M. Scheffer (2013). "Flickering as an early warning signal". In: *Theoretical Ecology* 6.3, pp. 309–317. DOI: 10.1007/s12080-013-0186-4 (cit. on pp. 19, 21).

Dakos, V. et al. (2015). "Resilience indicators: Prospects and limitations for early warnings of regime shifts". In: *Philosophical Transactions of the Royal Society B* 370.1659, pp. 1–10. DOI: 10.1098/rstb.2013.0263 (cit. on pp. 21, 33, 51, 54, 60).

D'Aoust, P. M. et al. (2021). "Catching a resurgence: Increase in SARS-CoV-2 viral RNA identified in wastewater 48 h before COVID-19 clinical tests and 96 h before hospitalizations". In: *Sci. Total Environ.* 770, p. 145319. DOI: https://doi.org/10.1016/j.scitotenv.2021.145319 (cit. on pp. 153, 168).

Daughton, C. G. (2020). "Wastewater surveillance for population-wide {COVID-19}: the present and future". In: *Sci. Total Environ.* 736, p. 139631. DOI: https://doi.org/10.1016/j.scitotenv.2020.139631 (cit. on p. 152).

Davis, M and S Leo (2013). "Black–Litterman in continuous time: the case for filtering". In: *Quant. Financ. Lett.* 1.1, pp. 30–35. DOI: https://doi.org/10.1080/21649502.2013.803794 (cit. on p. 69).

De-Camino-Beck, T. (2020). "A modified SEIR Model with Confinement and Lockdown of COVID-19 for Costa Rica". In: *medRxiv* (cit. on p. 64).

De Mot, L. et al. (2016). "Cell fate specification based on tristability in the inner cell mass of mouse blastocysts". In: *Biophysical journal* 110.3, pp. 710–722. DOI: 10.1016/j.bpj.2015.12.020 (cit. on p. 61).

Del Vecchio, D., A. J. Dy, and Y. Qian (2016). "Control theory meets synthetic biology". In: *Journal of The Royal Society Interface* 13.120, p. 20160380. DOI: 0.1098/rsif.2016.0380 (cit. on pp. 89, 91).

Demenocal, P. et al. (2000). "Abrupt onset and termination of the African Humid Period:: rapid climate responses to gradual insolation forcing". In: *Quaternary science reviews* 19.1-5, pp. 347–361. DOI: 10.1016/S0277-3791(99)00081-5 (cit. on p. 17).

Deris, J. B. et al. (2013). "The innate growth bistability and fitness landscapes of antibiotic-resistant bacteria". In: *Science* 342.6162, p. 1237435. DOI: 10.1126/science.1237435 (cit. on pp. 112, 119).

Desbarats, A. J. and R Dimitrakopoulos (2000). "Geostatistical simulation of regionalized pore-size distributions using min/max autocorrelation factors". In: *Mathematical Geology* 32.8, pp. 919–942. DOI: 10.1023/A:1007570402430 (cit. on p. 83).

Dessavre, A. G. et al. (2019). "The problem of detrending when analysing potential indicators of disease elimination". In: *Journal of theoretical biology* 481, pp. 183–193. DOI: 10.1016/j.jtbi.2019.04.011 (cit. on pp. 54, 65, 76, 77, 80, 139, 143, 150, 210).

Dethier, J. et al. (2015). "A positive feedback at the cellular level promotes robustness and modulation at the circuit level". In: *American Journal of Physiology-Heart and Circulatory Physiology* 114.4, pp. 2472–2484. DOI: 10.1152/jn.00471.2015 (cit. on p. 39).

Dhooge, A, W Govaerts, and Y. A. Kuznetsov (2003). "MATCONT: A MATLAB Package for Numerical Bifurcation Analysis of ODEs". In: *ACM Trans. Math. Softw.* 29.2, pp. 141–164. DOI: 10.1145/779359.779362 (cit. on p. 66).

Di Santo, S. et al. (2016). "Self-Organized Bistability Associated with First-Order Phase Transitions". In: *Physical Review Letters* 116.24, pp. 1–5. DOI: 10.1103/PhysRevLett.116.240601 (cit. on p. 39).

Diamond, J., J. A. Robinson, and Others (2010). *Natural experiments of history*. Harvard University Press (cit. on p. 131).

Dickman, R. et al. (2000). "Paths to Self-Organized Criticality". In: *Brazilian Journal of Physics* 30.1, pp. 27–41. DOI: 10.1590/S0103-97332000000100004 (cit. on pp. 39, 56).

Dickman, R., A. Vespignani, and S. Zapperi (2003). "Self-organized criticality as a phase transition". In: *Physical Review E* 57.5, pp. 5095–5105. DOI: 10.1103/PhysRevE.57.5095 (cit. on p. 39).

Diekmann, O, J. A. P. Heesterbeek, and M. G. Roberts (2010). "The construction of next-generation matrices for compartmental epidemic models". In: *Journal Royal Society Interface* 7.47, pp. 873–885. DOI: 10.1098/rsif.2009.0386 (cit. on p. 65).

Diks, C., C. Hommes, and J. Wang (2019). "Critical slowing down as an early warning signal for financial crises?" In: *Empirical Economics* 57.4, pp. 1201–1228. DOI: 10.1007/s00181-018-1527-3 (cit. on pp. 17, 21).

Ditlevsen, P. D., K. K. Andersen, and A. Svensson (2007). "The DO-climate events are probably noise induced: statistical investigation of the claimed 1470 years cycle". In: *Climate of the Past* 3.1, pp. 129–134. DOI: 10.5194/cp-3-129-2007 (cit. on p. 57).

Ditlevsen, P. D. and S. J. Johnsen (2010). "Tipping points: Early warning and wishful thinking". In: *Geophysical Research Letters* 37.19, pp. 2–5. DOI: 10.1029/2010GL044486 (cit. on pp. 33, 51, 58, 135, 176, 177).

Dmitriev, A. et al. (2017). "The Application of Stochastic Bifurcation Theory to the Early Detection of Economic Bubbles". In: *Procedia Computer Science* 122, pp. 354–361. DOI: 10.1016/j.procs.2017.11.380 (cit. on p. 17).

Dobrushkin, V. A. (2014). *Applied differential equations: the primary course*. Vol. 18. CRC Press (cit. on p. 181).

Dobson, I. et al. (2007). "Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization". In: *Chaos* 17.2, 026103 (2007); DOI: 10.1063/1.2737822 (cit. on p. 18).

Dominguez, L. G. (2005). "Enhanced Synchrony in Epileptiform Activity? Local versus Distant Phase Synchronization in Generalized Seizures". In: *Journal of Neuroscience* 25.35, pp. 8077–8084. DOI: 10.1523/jneurosci.1046-05.2005 (cit. on p. 57).

Dong, E., H. Du, and L. Gardner (2020). "An interactive web-based dashboard to track COVID-19 in real time". In: *The Lancet infectious diseases* 20.5, pp. 533–534. DOI: 10.1016/S1473-3099(20)30120-1 (cit. on pp. 136, 137).

Drake, J. M. (2013). "Early warning signals of stochastic switching". In: *Proceedings of the Royal Society B: Biological Sciences* 280.1766, pp. 15–18. DOI: 10.1098/rspb.2013.0686 (cit. on pp. 21, 81).

Drake, J. M. and B. D. Griffen (2010). "Early warning signals of extinction in deteriorating environments". In: *Nature* 467.7314, pp. 456–459. DOI: 10.1038/nature09389 (cit. on pp. 21, 44).

Drijfhout, S. et al. (2015). "Catalogue of abrupt shifts in Intergovernmental Panel on Climate Change climate models". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.43, E5777–E5786. DOI: 10.1073/pnas.1511451112 (cit. on p. 17).

Drion, G. et al. (2012). "A novel phase portrait for neuronal excitability". In: *PloS one* 7.8, e41806. DOI: 10.1371/journal.pone.0041806 (cit. on p. 39).

Drion, G. et al. (2015a). "Dynamic input conductances shape neuronal spiking". In: *eneuro* 2.1. DOI: 10.1523/ENEURO.0031-14.2015 (cit. on p. 39).

Drion, G. et al. (2015b). "Neuronal behaviors: A control perspective". In: *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, pp. 1923–1944 (cit. on pp. 31, 39).

Drion, G., A. Franci, and R. Sepulchre (2018). "Cellular switches orchestrate rhythmic circuits". In: *Biological cybernetics* 113.1, pp. 71–82. DOI: 10.1007/s00422-018-0778-6 (cit. on p. 39).

Du, Z. et al. (2020). "Serial interval of COVID-19 among publicly reported confirmed cases". In: *Emerging infectious diseases* 26.6, p. 1341. DOI: 10.3201/eid2606.200357 (cit. on p. 138).

Duff, C. et al. (2012). "Mathematical modelling of stem cell differentiation: The PU.1-GATA-1 interaction". In: *Journal of Mathematical Biology* 64.3, pp. 449–468. DOI: 10.1007/s00285-011-0419-3 (cit. on p. 40).

Dunlop, M. J. et al. (2008). "Regulatory activity revealed by dynamic correlations in gene expression noise". In: *Nature Genetics* 40.12, pp. 1493–1498. DOI: 10.1038/ng.281 (cit. on p. 106).

Dutta, P. S., Y. Sharma, and K. C. Abbott (2018). "Robustness of early warning signals for catastrophic and non-catastrophic transitions". In: *Oikos* 127.9, pp. 1251–1263. DOI: 10.1111/oik.05172 (cit. on pp. 33, 52, 108, 191).

Ebert, M. S. and P. A. Sharp (2012). "Roles for microRNAs in conferring robustness to biological processes". In: *Cell* 149.3, pp. 515–524. DOI: 10.1016/j.cell.2012.04.005 (cit. on p. 112).

Ebrahimzadeh, E. et al. (2018). "Prediction of paroxysmal Atrial Fibrillation: A machine learning based approach using combined feature vector and mixture of expert classification on HRV signal". In: *Computer methods and programs in biomedicine* 165, pp. 53–67. DOI: 10.1016/j.cmpb.2018.07.014 (cit. on p. 178).

Eckmann, J.-P. et al. (1995). "Recurrence plots of dynamical systems". In: *World Scientific Series on Nonlinear Science Series A* 16, pp. 441–446 (cit. on p. 75).

Eelbode, T. et al. (2020). "Optimization for medical image segmentation: theory and practice when evaluating with Dice score or Jaccard index". In: *IEEE Transactions on Medical Imaging* 39.11, pp. 3679–3690. DOI: 10.1109/TMI.2020.3002417 (cit. on p. 45).

Eissing, T. et al. (2011). "A computational systems biology software platform for multiscale modeling and simulation: integrating whole-body physiology, disease biology, and molecular reaction networks". In: *Frontiers in physiology* 2, p. 4 (cit. on p. 10).

Ermentrout, G. B. and D. Kleinfeld (2001). "Traveling electrical waves in cortex: insights from phase dynamics and speculation on a computational role". In: *Neuron* 29.1, pp. 33–44 (cit. on p. 57).

Eugenio, M. et al. (2014). "Bifurcation analysis of single-cell gene expression data reveals epigenetic landscape". In: *Proceedings of the National Academy of Sciences of the United States of America* 111.52, E5643–E5650. DOI: 10.1073/pnas.1408993111 (cit. on pp. 40, 70–72, 177).

Fan, J et al. (2020). *The University of Maryland Social Data Science Center Global COVID-19 Trends and Impact Survey, in partnership with Facebook*. https://covidmap.umd.edu/api.html (cit. on p. 179).

Farkas, K. et al. (2020). "Wastewater and public health: the potential of wastewater surveillance for monitoring {COVID-19}". In: *Curr. Opin. Envir. Sci. Heal.* 17, pp. 14–20. DOI: https://doi.org/10.1016/j.coesh.2020.06.001 (cit. on p. 152).

Fawcett, T. (2006). "An introduction to ROC analysis". In: *Pattern recognition letters* 27.8, pp. 861–874 (cit. on p. 82).

Feng, S. et al. (2016). "Core signalling motif displaying multistability through multi-state enzymes". In: *Journal of the Royal Society Interface* 13.123. DOI: 10.1098/rsif.2016.0524 (cit. on p. 36).

Ferguson, N. M. et al. (2020). "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand". In: *Imperial College, London* March, p. 20. DOI: 10.25561/77482 (cit. on p. 65).

Fernandez-Cassi, X. et al. (2021). "Wastewater monitoring outperforms case numbers as a tool to track {COVID-19} incidence dynamics when test positivity rates are high". In: *Water research* 200, p. 117252. DOI: https://doi.org/10.1016/j.watres.2021.117252 (cit. on p. 167).

Ferrell, J. E. (2012). "Bistability, bifurcations, and Waddington's epigenetic landscape". In: *Current Biology* 22.11, R458–R466. DOI: 10.1016/j.cub.2012.03.045 (cit. on p. 40).

Feudel, U., A. N. Pisarchik, and K. Showalter (2018). "Multistability and tipping: From mathematics and physics to climate and brain - Minireview and preface to the focus issue". In: *Chaos* 28.3, p. 033501. DOI: 10.1063/1.5027718 (cit. on p. 17).

Fiorentino, J., M.-E. Torres-Padilla, and A. Scialdone (2020). "Measuring and Modeling Single-Cell Heterogeneity and Fate Decision in Mouse Embryos". In: *Annual Review of Genetics* 54, pp. 167–187. DOI: 10.1146/annurev-genet-021920-110200 (cit. on p. 61).

Fisher, G. H. (1967). "Measuring ambiguity". In: *The American journal of psychology* 80.4, pp. 541–557. DOI: 10.2307/1421187 (cit. on p. 39).

Fisher, J. and T. A. Henzinger (2006). "Executable Biology". In: *Proceedings of the 2006 Winter Simulation Conference*, pp. 1675–1682. DOI: 10.1109/WSC.2006.322942 (cit. on p. 5).

Fisher, L. (2015). "More than 70 ways to show resilience". In: *Nature* 518.7537, p. 35 (cit. on p. 20).

Flaxman, S. et al. (2020). "Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe". In: *Nature* 584.7820, pp. 257–261. DOI: 10.1038/s41586-020-2405-7 (cit. on p. 133).

Foo, J. C. et al. (2017). "Dynamical state transitions into addictive behaviour and their early-warning signals". In: *Proceedings of the Royal Society B: Biological Sciences* 284.1860, p. 20170882. DOI: 10.1098/rspb.2017.0882 (cit. on p. 17).

Fraccascia, L., I. Giannoccaro, and V. Albino (2018). "Resilience of complex systems: state of the art and directions for future research". In: *Complexity* 2018. DOI: 10.1155/2018/3421529 (cit. on pp. 17, 20).

Franci, A., G. Drion, and R. Sepulchre (2012). "An organizing center in a planar model of neuronal excitability". In: *SIAM Journal on Applied Dynamical Systems* 11.4, pp. 1698–1722. DOI: 10.1137/120875016 (cit. on p. 39).

Franci, A. et al. (2013). "A balance equation determines a switch in neuronal excitability". In: *PLoS computational biology* 9.5, e1003040. DOI: 10.1371/journal.pcbi.1003040 (cit. on p. 39).

Franci, A., G. Drion, and R. Sepulchre (2014). "Modeling the modulation of neuronal bursting: a singularity theory approach". In: *SIAM Journal on Applied Dynamical Systems* 13.2, pp. 798–829. DOI: 10.1137/13092263X (cit. on p. 39).

Franci, A. and R. Sepulchre (2016). "A three-scale model of spatio-temporal bursting". In: *SIAM Journal on Applied Dynamical Systems* 15.4, pp. 2143–2175. DOI: 10.1137/15M1046101 (cit. on p. 39).

Franci, A., G. Drion, and R. Sepulchre (2018). "Robust and tunable bursting requires slow positive feedback". In: *Journal of neurophysiology* 119.3, pp. 1222–1234. DOI: 10.1152/jn.00804.2017 (cit. on pp. 38, 39).

Frank, T. (2004). "Complete description of a generalised Ornstein-Uhlenbeck process related to the nonextensive Gaussian entropy". In: *Physica A: Statistical Mechanics and its Applications* 340.1-3, pp. 251–256. DOI: 10.1016/j.physa.2004.04.014 (cit. on p. 206).

Frank, T. D., A. Daffertshofer, and P. J. Beek (2002). "Impacts of statistical feedback on the flexibility-accuracy trade-off in biological systems". In: *Journal of Biological Physics* 28.1, pp. 39–54. DOI: 10.1023/A:1016256613673 (cit. on pp. 38, 112).

Fraser, C. et al. (2009). "Pandemic potential of a strain of influenza A (H1N1): early findings". In: *Science* 324.5934, pp. 1557–1561. DOI: 10.1126/science.1176062 (cit. on p. 64).

Frei, M. G. and E. al (2010). "Controversies in epilepsy: debates held during the Fourth International Workshop on Seizure Prediction". In: *Epilepsy & Behavior* 19.1, pp. 4–16. DOI: 10.1016/j.yebeh.2010.06.009.CONTROVERSIES (cit. on p. 57).

Freidlin, M. I. and A. D. Wentzell (1998). "Random perturbations of dynamical systems". In: Springer (cit. on pp. 29, 196, 205).

Freidlin, M. and A. Wentzell (1979). *Random Perturbations of Dynamical Systems*. Springer (cit. on pp. 19, 72, 213).

Frias-Martinez, E., G. Williamson, and V. Frias-Martinez (2011). "An agent-based model of epidemic spread using human mobility and social network information". In: *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*. IEEE, pp. 57–64 (cit. on p. 63).

Friedman, N., L. Cai, and X. S. Xie (2006). "Linking stochastic dynamics to population distribution: an analytical framework of gene expression". In: *Phys. Rev. Lett.* 97.16, p. 168302. DOI: P10.1103/PhysRevLett.97.168302 (cit. on p. 114).

Frigola, D. et al. (2012). "Asymmetric stochastic switching driven by intrinsic molecular noise". In: *PloS one* 7.2, e31407. DOI: 10.1371/journal.pone.0031407 (cit. on pp. 26, 61, 63, 112).

G. Butler, R. and G. R. Butler (2014). *Exploratory vs Confirmatory Research*. https://bit.ly/3roPRj1 (cit. on p. 7).

Gal, A. et al. (2010). "Dynamics of Excitability over Extended Timescales in Cultured Cortical Neurons". In: *Journal of Neuroscience* 30.48, pp. 16332–16342. DOI: 10.1523/JNEUROSCI.4859-10.2010 (cit. on pp. 39, 57).

Gal, A. and S. Marom (2013). "Self-organized criticality in single-neuron excitability". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 88.6, pp. 1–6. DOI: 10.1103/PhysRevE.88.062717 (cit. on pp. 39, 57).

Gallo, L. et al. (2021). "Lack of practical identifiability may hamper reliable predictions in COVID-19 epidemic models". In: *Science advances* 8.3, eabg5234. DOI: 10.1126/sciadv.abg5234 (cit. on p. 68).

Gammaitoni, L. et al. (1998). "Stochastic resonance". In: *Reviews of Modern Physics* 70.1, pp. 223–287. DOI: 10.1103/RevModPhys.70.223 (cit. on p. 33).

Gao, J., B. Barzel, and A. L. Barabási (2016). "Universal resilience patterns in complex networks". In: *Nature* 530.7590, pp. 307–312. DOI: 10.1038/nature16948 (cit. on pp. 31, 84, 116, 178).

Gardiner, C. W. (1985). *Handbook of Stochastic Methods*. Boca Raton: Springer (cit. on pp. 29, 63, 67, 72, 91, 93, 94, 99, 102, 117, 196, 203, 204, 206, 211, 213).

Gatto, M. et al. (2020). "Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures". In: *Proceedings of the National Academy of Sciences* 117.19, pp. 10484–10491. DOI: https://doi.org/10.1073/pnas.2004978117 (cit. on p. 64).

Gavidia, M. (2022). "Real-time Monitoring and Early Warning Of Atrial Fibrillation". In: *OrbiLu* (cit. on p. 178).

Gelman, A. and D. B. Rubin (1992). "Inference from Iterative Simulation Using Multiple Sequences". In: *Statist. Sci.* 7.4, pp. 457–472. DOI: 10.1214/ss/1177011136 (cit. on p. 69).

George, S. V., S. Kachhara, and G Ambika (2021). "Early warning signals for critical transitions in complex systems". In: *arXiv preprint arXiv:2107.01210* (cit. on p. 47).

Gertsev, V. I. and V. V. Gertseva (2004). "Classification of mathematical models in ecology". In: *Ecological Modelling* 178.3-4, pp. 329–334. DOI: 10.1016/j.ecolmodel.2004.03.009 (cit. on p. 5).

Ghaffarizadeh, A., N. S. Flann, and G. J. Podgorski (2014a). "Multistable switches and their role in cellular differentiation networks". In: *BMC bioinformatics* 15.7, pp. 1–13. DOI: 10.1186/1471-2105-15-S7-S7 (cit. on p. 18).

Ghaffarizadeh, A., N. S. Flann, and G. J. Podgorski (2014b). "Multistable switches and their role in cellular differentiation networks". In: *BMC Bioinformatics* 15.Suppl 7, S7. DOI: 10.1186/1471-2105-15-S7-S7 (cit. on p. 40).

Gidea, M. and Y. Katz (2018). "Topological data analysis of financial time series: Landscapes of crashes". In: *Physica A: Statistical Mechanics and its Applications* 491, pp. 820–834. DOI: 10.1016/j.physa.2017.09.028 (cit. on p. 177).

Gidea, M. et al. (2020). "Topological recognition of critical transitions in time series of cryptocurrencies". In: *Physica A: Statistical Mechanics and its Applications* 548, p. 123843. DOI: 10.1016/j.physa.2019.123843 (cit. on pp. 177, 179).

Giere, R. (2004). "How models are used to represent reality". In: *Philosophy of science* 71.5, pp. 742–752 (cit. on p. 4).

Gillespie, D (2000a). "The chemical Langevin equation". In: *J. Chem. Phys.* 113.1, pp. 297–306. DOI: 10.1063/1.481811 (cit. on pp. 90, 156).

Gillespie, D. T. (2000b). "Chemical Langevin equation". In: *Journal of Chemical Physics* 113.1, pp. 297–306. DOI: 10.1063/1.481811 (cit. on pp. 63, 91, 197).

Giordano, G. et al. (2020). "Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy". In: *Nature Medicine* 26.6, pp. 855–860. DOI: 10.1038/s41591-020-0883-7 (cit. on pp. 64, 65, 136).

Goldberg, Y. et al. (2021). "Waning immunity after the {BNT162b2} vaccine in {I}srael". In: *New England Journal of Medicine* 385.24, e85. DOI: https://doi.org/10.1056/NEJMoa2114228 (cit. on p. 156).

Gollas, F. and R. Tetzlaff (2005). "Modeling brain electrical activity in epilepsy by reaction-diffusion cellular neural networks". In: *Bioengineered and Bioinspired Systems II* 5839, p. 219. DOI: 10.1117/12.608643 (cit. on p. 57).

Golubitsky, M., I. Stewart, and D. G. Schaeffer (2012). *Singularities and Groups in Bifurcation Theory-II*. Springer Science \& Business Media (cit. on pp. 26, 181, 186).

Gondelfeld, N. (1992). *Lectures on phase transitions and the renormalization group*. Taylor & Francis (cit. on p. 38).

Gopalakrishnan, E. A. et al. (2016). "Early warning signals for critical transitions in a thermoacoustic system". In: *Scientific Reports* 6.September, pp. 1–10. DOI: 10.1038/srep35310 (cit. on p. 21).

Gowda, K. and C. Kuehn (2015). "Early-warning signs for pattern-formation in stochastic partial differential equations". In: *Communications in Nonlinear Science and Numerical Simulation* 22.1-3, pp. 55–69. DOI: 10.1016/j.cnsns.2014.09.019 (cit. on pp. 21, 31, 52, 179).

Gross, J. L. and J. Yellen (2005). *Graph theory and its applications*. CRC press (cit. on pp. 48, 198).

Guckenheimer, J., P. Holmes, and M. Slemrod (2013). *Nonlinear Oscillations Dynamical Systems, and Bifurcations of Vector Fields*. Springer Science \& Business Media. DOI: 10.1115/1.3167759 (cit. on pp. 25, 185, 186).

Guinn, M. T. et al. (2020). "Observation and control of gene expression noise: barrier crossing analogies between drug resistance and metastasis". In: *Frontiers in Genetics* 11, p. 1292. DOI: 10.3389/fgene.2020.586726 (cit. on p. 61).

Gunawardena, J. (2014). "Time-scale separation–Michaelis and Menten's old idea, still bearing fruit". In: *The FEBS journal* 281.2, pp. 473–488. DOI: 10.1111/febs.12532 (cit. on p. 89).

Gundy, P. M., C. P. Gerba, and I. L. Pepper (2009). "Survival of coronaviruses in water and wastewater". In: *Food Environ. Virol.* 1.1, pp. 10–14. DOI: https://doi.org/10.1007/s12560-008-9001-6 (cit. on p. 156).

Gutierrez, P. S., D. Monteoliva, and L. Diambra (2009). "Role of cooperative binding on noise expression". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 80.1, pp. 1–7. DOI: 10.1103/PhysRevE.80.011914 (cit. on p. 116).

Guttal, V. and C. Jayaprakash (2008). "Changing skewness: An early warning signal of regime shifts in ecosystems". In: *Ecology Letters* 11.5, pp. 450–460. DOI: 10.1111/j.1461-0248.2008.01160.x (cit. on pp. 21, 99, 100, 109, 144, 147).

Haario, H. et al. (2006). "DRAM: efficient adaptive MCMC". In: *Statistics and computing* 16.4, pp. 339–354 (cit. on p. 69).

Haragus, M. and G. Iooss (2010). *Local Bifurcation, Center Manifolds and Normal Forms in Infinte-Dimensional Dynamical Systems*. Springer Science \& Business Media (cit. on p. 185).

Haragus, M. et al. (2018). "Probability of noise and rate-induced tipping". In: *Chaos* 475.3, pp. 1–15. DOI: 10.1115/1.3167759 (cit. on pp. 21, 34).

Harris, M. J., S. I. Hay, and J. M. Drake (2020). "Early warning signals of malaria resurgence in Kericho, Kenya". In: *Biology Letters* 16.3, pp. 1–6. DOI: 10.1098/rsbl.2019.0713 (cit. on p. 131).

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*. Cambridge university press (cit. on p. 69).

Hasan, S. W. et al. (2021). "Detection and quantification of {SARS-CoV-2 RNA} in wastewater and treated effluents: Surveillance of {COVID-19} epidemic in the {United Arab Emirates}". In: *Sci. Total Environ.* 764, p. 142929. DOI: https://doi.org/10.1016/j.scitotenv.2020.142929 (cit. on p. 152).

Hasty, J. et al. (2000). "Noise-based switches and amplifiers forgene expression". In: *Proceedings of the National Academy of Sciences of the United States of America* 97.5, pp. 2075–2080. DOI: 10.1073/pnas.040411297 (cit. on pp. 18, 29, 61, 62, 90, 91, 105, 111, 113, 119, 121, 126).

He, S., Y. Peng, and K. Sun (2020). "SEIR modeling of the COVID-19 and its dynamics". In: *Nonlinear dynamics* 101.3, pp. 1667–1680. DOI: 10.1007/s11071-020-05743-y (cit. on p. 154).

Heagy, J. F. and S. M. Hammel (1994). "The birth of strange nonchaotic attractors". In: *Physica D: Nonlinear Phenomena* 70.1-2, pp. 140–153 (cit. on p. 31).

Heino, M. T. J. et al. (2021). "Studying behaviour change mechanisms under complexity". In: *Behavioral Sciences* 11.5, p. 77. DOI: 10.3390/bs11050077 (cit. on p. 75).

Heino, M. T. J. et al. (2022). "Attractor landscapes: A unifying conceptual model for understanding behaviour change across scales of observation". In: DOI: 10.31234/osf.io/3rxyd (cit. on pp. 14, 21).

Heinrich, M. et al. (2010). "Symmetry-breaking transitions in networks of nonlinear circuit elements". In: *New Journal of Physics* 12, p. 113030. DOI: 10.1088/1367-2630/12/11/113030 (cit. on p. 39).

Heller, P. (1967). "Experimental investigations of critical phenomena". In: *Reports on Progress in Physics* 30.2, p. 731. DOI: 10.1088/0034-4885/30/2/307 (cit. on p. 43).

Hepburn, B. and H. Andersen (2021). "Scientific Method". In: *The {Stanford} Encyclopedia of Philosophy*. Ed. by E. N. Zalta. {S}ummer 2. Metaphysics Research Lab, Stanford University (cit. on pp. 1, 7).

Herrera-Valdez, M. A., M. Cruz-Aponte, and C. Castillo-Chavez (2011). "Multiple outbreaks for the same pandemic: Local transportation and social distancing explain the different "Waves" of A-H1N1PDM

cases observed in México during 2009". In: *Mathematical Biosciences and Engineering* 8.1, pp. 21–48. DOI: 10.3934/mbe.2011.8.21 (cit. on p. 65).

Heseltine, J. and E. J. Kim (2016). "Novel mapping in non-equilibrium stochastic processes". In: *Journal of Physics A: Mathematical and Theoretical* 49.17, p. 175002. DOI: 10.1088/1751-8113/49/17/175002 (cit. on p. 206).

— (2019). "Comparing information metrics for a coupled Ornstein-Uhlenbeck process". In: *Entropy* 21.8, pp. 1–16. DOI: 10.3390/e21080775 (cit. on p. 206).

Hesse, J. and T. Gross (2014). "Self-organized criticality as a fundamental property of neural systems". In: *Frontiers in systems neuroscience* 8, p. 166. DOI: 10.3389/fnsys.2014.00166 (cit. on p. 56).

Hill, N. R. et al. (2019). "Predicting atrial fibrillation in primary care using machine learning". In: *PLoS one* 14.11, e0224582. DOI: 10.1371/journal.pone.0224582 (cit. on p. 178).

Hillmer, S. C. and G. C. Tiao (1982). "An ARIMA-model-based approach to seasonal adjustment". In: *Journal of the American Statistical Association* 77.377, pp. 63–70. DOI: 10.1080/01621459.1982.10477767 (cit. on p. 78).

Hirota, M. et al. (2011). "Global Resilience of Tropical Forest". In: *Science* 334.October, pp. 232–235. DOI: 10.1126/science.1210657 (cit. on p. 17).

Hodges, W. (2020). "Model Theory". In: *The {Stanford} Encyclopedia of Philosophy*. Ed. by E. N. Zalta. {W}inter 2. Metaphysics Research Lab, Stanford University (cit. on p. 4).

Holling, C. S. (1996). "Engineering resilience versus ecological resilience". In: *Engineering within ecological constraints* 31.1996, p. 32 (cit. on pp. 20, 91).

Hooker, C., ed. (2011). *Philosophy of complex systems*. Elsevier B.V. (cit. on pp. 8, 10).

Hoppensteadt, F. C. and E. M. Izhikevich (1998). "Thalamo-cortical interactions modeled by weakly connected oscillators: Could the brain use FM radio principles?" In: *BioSystems* 48.1-3, pp. 85–94. DOI: 10.1016/S0303-2647(98)00053-7 (cit. on p. 57).

Horstmeyer, L., C. Kuehn, and S. Thurner (2018). "Network topology near criticality in adaptive epidemics". In: *Physical Review E* 98.4, p. 42313. DOI: 10.1103/PhysRevE.98.042313 (cit. on pp. 18, 65, 130).

Hsiang, S. et al. (2020). "The effect of large-scale anti-contagion policies on the COVID-19 pandemic". In: *Nature* 584.7820, pp. 262–267. DOI: 10.1038/s41586-020-2404-8 (cit. on p. 130).

Huang, S. et al. (2007). "Bifurcation dynamics in lineage-commitment in bipotent progenitor cells". In: *Developmental biology* 305.2, pp. 695–713. DOI: 10.1016/j.ydbio.2007.02.036 (cit. on pp. 40, 61).

Huang, Y. and R. Gottardo (2013). "Comparability and reproducibility of biomedical data". In: *Briefings in bioinformatics* 14.4, pp. 391–401. DOI: 10.1093/bib/bbs078 (cit. on pp. 2, 3).

Huisman, J. S. et al. (2021). "Wastewater-based estimation of the effective reproductive number of {SARS-CoV-2}". In: *medRxiv*, p. 2021.04.29.21255961. DOI: https://doi.org/10.1101/2021.04.29.21255961 (cit. on pp. 152, 165, 170).

Humpherys, J., P. Redd, and J. West (2012). "A Fresh Look at the Kalman Filter". In: *SIAM Review* 54.4, pp. 801–823. DOI: 10.1137/100799666 (cit. on p. 69).

Hunt, C. A. et al. (2008). "Dichotomies between computational and mathematical models". In: *Nature Biotechnology* 26.7, pp. 737–738. DOI: 10.1038/nbt0708-737 (cit. on p. 5).

Iglesias, P. A. and B. P. Ingalls (2010). *Control theory and systems biology*. MIT Press (cit. on pp. 8, 31, 39).

Izhikevich, E. M. (1998). "Multiple cusp bifurcations". In: *Neural Networks* 11.3, pp. 495–508. DOI: 10.1016/S0893-6080(97)00117-2 (cit. on p. 39).

— (2007). *The Geometry of Excitability and Bursting Eugene*, p. 111. DOI: 10.1017/S0143385704000173 (cit. on p. 28).

Jeffries, C. and K. Wiesenfeld (1985). "Observation of noisy precursors of dynamical instabilities". In: *Physical Review A* 31.2, pp. 1077–1085. DOI: 10.1103/PhysRevA.31.1077 (cit. on p. 39).

Jensen, H. J. (1998). *Self-Organized Criticality - Emergent Complex Behavior in Physical and Biological Systems*. Cambridge: Cambridge University Press (cit. on pp. 4, 39).

Jiang, J. et al. (2018). "Predicting tipping points in mutualistic networks through dimension reduction". In: *Proceedings of the National Academy of Sciences of the United States of America* 115.4, E639–E647. DOI: 10.1073/pnas.1714958115 (cit. on p. 85).

Jiang, J., A. Hastings, and Y.-C. Lai (2019). "Harnessing tipping points in complex ecological networks". In: *Journal of the Royal Society Interface* 16.158, p. 20190345. DOI: 10.1098/rsif.2019.0345 (cit. on p. 17).

Jirsa, V. e al. (2014). "On the nature of seizure dynamics". In: *Brain* 53.8-9, pp. 992–995. DOI: 10.1093/brain/awu133 (cit. on pp. 18, 30, 57).

Jones, C. R. and M. Jahanshahi (2014). "Contributions of the Basal Ganglia to Temporal Processing: Evidence from Parkinson's Disease". In: *Timing and Time Perception* 2.1, pp. 87–127. DOI: 10.1163/22134468-00002009 (cit. on p. 50).

Jones, D. S., M. Plank, and B. D. Sleeman (2009). *Differential equations and mathematical biology*. CRC press (cit. on p. 5).

Jurczyk, J. et al. (2017). "Measuring critical transitions in financial markets". In: *Scientific Reports* 7.1, pp. 1–6. DOI: 10.1038/s41598-017-11854-1 (cit. on p. 17).

Kaern, M. et al. (2005). "Stochasticity in gene expression: from theories to phenotypes". In: *Nature Reviews Genetics* 6.6, pp. 451–464. DOI: 10.1038/nrg1615 (cit. on p. 61).

Kalman, R (1960). "A new approach to linear filtering and prediction problems". In: *J. Basic Eng.* 82.82, pp. 35–45 (cit. on pp. 69, 79).

Kanamaru, T., H. Fujii, and K. Aihara (2013). "Deformation of Attractor Landscape via Cholinergic Presynaptic Modulations: A Computational Study Using a Phase Neuron Model". In: *PLoS ONE* 8.1, e53854. DOI: 10.1371/journal.pone.0053854 (cit. on p. 40).

Kapo, K. E. et al. (2017). "Estimation of {US} sewer residence time distributions for national-scale risk assessment of down-the-drain chemicals". In: *Sci. Total Environ.* 603, pp. 445–452. DOI: https://doi.org/10.1016/j.scitotenv.2017.06.075 (cit. on p. 157).

Katok, A. and B. Hasselblatt (1997). *Introduction to the modern theory of dynamical systems*. 54. Cambridge university press (cit. on p. 19).

Kaur, T. et al. (2020). "Anticipating the Novel Coronavirus Disease (COVID-19) Pandemic". In: *Frontiers in Public Health* 8.September, pp. 1–12. DOI: 10.3389/fpubh.2020.569669 (cit. on p. 52).

Kéfi, S. et al. (2013). "Early warning signals also precede non-catastrophic transitions". In: *Oikos* 122.5, pp. 641–648. DOI: 10.1111/j.1600-0706.2012.20838.x (cit. on pp. 52, 191).

Kéfi, S., M. Holmgren, and M. Scheffer (2016). "When can positive interactions cause alternative stable states in ecosystems?" In: *Functional Ecology* 30.1, pp. 88–97. DOI: 10.1111/1365-2435.12601 (cit. on pp. 38, 47, 49).

Kemp, F. F. et al. (2021). "Modelling COVID-19 dynamics and potential for herd immunity by vaccination in Austria, Luxembourg and Sweden". In: *Journal of Theoretical Biology* 530, p. 110874. DOI: https://doi.org/10.1016/j.jtbi.2021.110874 (cit. on pp. 15, 64, 65, 69, 129, 130, 136, 156, 160, 162, 164, 168, 170).

Kendall, M. G. (1938). "A new measure of rank correlation". In: *Biometrika* 30.1/2, pp. 81–93 (cit. on p. 80).

Kepler, T. B. and T. C. Elston (2001). "Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations". In: *Biophysical journal* 81.6, pp. 3116–3136. DOI: 10.1016/S0006-3495(01)75949-8 (cit. on p. 88).

Kermack, W. O. and A. G. McKendrick (1927). "A contribution to the mathematical theory of epidemics". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 115.772, pp. 700–721. DOI: 10.1098/rspa.1927.0118 (cit. on pp. 53, 63).

Kerr, C. C. et al. (2021). "Covasim: an agent-based model of COVID-19 dynamics and interventions". In: *PLOS Computational Biology* 17.7, e1009149. DOI: 10.1371/journal.pcbi.1009149 (cit. on p. 130).

Kerr, N. L. (1998). "HARKing: Hypothesizing after the results are known". In: *Personality and social psychology review* 2.3, pp. 196–217. DOI: 10.1207/s15327957pspr0203_4 (cit. on p. 7).

Kerr, R., S. Jabbari, and I. G. Johnston (2019). "Intracellular Energy Variability Modulates Cellular Decision-Making Capacity". In: *Scientific Reports* 9.1, pp. 1–12. DOI: 10.1038/s41598-019-56587-5 (cit. on p. 40).

Khas' minskii, R. Z. (1966). "A limit theorem for the solutions of differential equations with random right-hand sides". In: *Theory of Probability & Its Applications* 11.3, pp. 390–406 (cit. on pp. 198, 204).

Kheir Gouda, M., M. Manhart, and G. Balázsi (2019). "Evolutionary regain of lost gene circuit function". In: *Proceedings of the National Academy of Sciences* 116.50, pp. 25162–25171. DOI: 10.1073/pnas.1912257116 (cit. on p. 112).

Kiers, C. and C. K. Jones (2019). "On Conditions for Rate-induced Tipping in Multi-dimensional Dynamical Systems". In: *Journal of Dynamics and Differential Equations* 3, pp. 1–15. DOI: 10.1007/s10884-019-09730-9 (cit. on pp. 31, 178).

Kirby, J. F. (2005). "Which wavelet best reproduces the Fourier power spectrum?" In: *Computers and Geosciences* 31.7, pp. 846–864. DOI: 10.1016/j.cageo.2005.01.014 (cit. on pp. 79, 178, 206).

Kitano, H. (2002). "Systems biology: a brief overview". In: *science* 295.5560, pp. 1662–1664. DOI: 10.1126/science.1069492 (cit. on pp. 5, 22).

— (2004). "Biological robustness". In: *Nature Reviews Genetics* 5.11, pp. 826–837. DOI: 10.1038/nrg1471 (cit. on p. 20).

Kleinen, T., H. Held, and G. Petschel-Held (2003). "The potential role of spectral properties in detecting thresholds in the Earth system: Application to the thermohaline circulation". In: *Ocean Dynamics* 53.2, pp. 53–63. DOI: 10.1007/s10236-002-0023-6 (cit. on p. 21).

Klose, A. K. et al. (2020). "Emergence of cascading dynamics in interacting tipping elements of ecology and climate". In: *Royal Society open science* 7.6, p. 200599. DOI: 10.1098/rsos.200599 (cit. on p. 18).

Klosin, A. et al. (2020). "Phase separation provides a mechanism to reduce noise in cells". In: *Science* 367.6476, pp. 464–468. DOI: 10.1126/science.aav6691 (cit. on p. 112).

Kolata, G. B. (1977). "Catastrophe theory: the emperor has no clothes". In: *Science* 196.4287, pp. 287–351 (cit. on pp. 39, 222).

Kollepara, P. K., A. F. Siegenfeld, and Y. Bar-Yam (2021). "Modeling complex systems: A case study of compartmental models in epidemiology". In: *arXiv*, arXiv:2110.02947 (cit. on p. 170).

Kolodziejczyk, A. A. et al. (2015). "The technology and biology of single-cell RNA sequencing". In: *Mol. Cell* 58.4, pp. 610–620. DOI: 10.1016/j.molcel.2015.04.005 (cit. on p. 80).

Komin, N. and A. Skupin (2017). "How to address cellular heterogeneity by distribution biology". In: *Current Opinion in Systems Biology* 3, pp. 154–160. DOI: 10.1016/j.coisb.2017.05.010 (cit. on p. 61).

Kopell, N. and G. Ermentrout (1990). "Phase Transitions and Other Phenomena in Chains of Coupled Oscillators". In: *SIAM Journal on Applied Mathematics* 50.4, pp. 1014–1052. DOI: 10.1137/0150062 (cit. on p. 38).

Korda, M. and I. Mezić (2018). "Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control". In: *Automatica* 93, pp. 149–160. DOI: 10.1016/j.automatica.2018.03.046 (cit. on p. 30).

Korolev, K. S., J. Xavier, and J. Gore (2014). "Turning ecology and evolution against cancer". In: *Nature Reviews Cancer*, pp. 1–10. DOI: 10.1038/nrc3712 (cit. on pp. 41, 88).

Kostanjcar, Z et al. (2016). "Estimating Tipping Points in Feedback-Driven Financial Networks". In: *IEEE Journal of Selected Topics in Signal Processing* 10.6, pp. 1040–1052. DOI: 10.1109/JSTSP.2016.2593099 (cit. on p. 17).

Kramer, M. A. et al. (2012). "Human seizures self-terminate across spatial scales via a critical transition". In: *Proceedings of the National Academy of Sciences* 109.51, pp. 21116–21121. DOI: 10.1073/pnas.1210047110 (cit. on p. 57).

Kruüsemann, J. L. "Modelling the behaviour of a synthetic gene circuit for controlled gene expression in Chlamydomonas reinhardtii". Master thesis. ETH Zurich, University of Cambridge (cit. on p. 7).

Kuehn, C. (2011). "A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics". In: *Physica D: Nonlinear Phenomena* 240.12, pp. 1020–1035. DOI: 10.1016/j.physd.2011.02.012 (cit. on pp. 18, 27, 32, 135, 185, 192, 195).

— (2013). "A mathematical framework for critical transitions: Normal forms, variance and applications". In: *Journal of Nonlinear Science* 23.3, pp. 457–510. DOI: 10.1007/s00332-012-9158-x (cit. on pp. 28, 52, 54, 76, 77, 141, 144, 202, 209, 220, 221).

Kuehn, C. and C. Meisel (2011). "On spatial and temporal multilevel dynamics and scaling effects in epileptic seizures". In: *arXiv:1103.5934*, pp. 1–24 (cit. on pp. 28, 29, 32, 46, 52–54, 65, 116, 121, 191, 205).

Kuehn, C., G. Zschaler, and T. Gross (2015). "Early warning signs for saddle-escape transitions in complex networks". In: *Scientific Reports* 5, p. 13190. DOI: 10.1038/srep13190 (cit. on pp. 21, 178).

Kuehn, C. and F. Romano (2019). "Scaling laws and warning signs for bifurcations of SPDEs". In: *European Journal of Applied Mathematics* 30.5, pp. 853–868. DOI: 10.1017/S0956792518000438 (cit. on pp. 21, 179).

Kuehn, C. and C. Bick (2021). "A universal route to explosive phenomena". In: *Science Advances* 7.16, pp. 1–7. DOI: 10.1126/sciadv.abe3824 (cit. on pp. 25, 26, 31, 89, 91, 116, 186).

Kuehn, C., K. Lux, and A. Neamtu (2022). "Warning Signs for Non-Markovian Bifurcations: Color Blindness and Scaling Laws". In: *Proceedings of the Royal Society A* 478.2259, p. 20210740. DOI: 10.1098/rspa.2021.0740 (cit. on pp. 33, 52, 88).

Kuhn, T. (1962). *The structure of scientific revolutions*. Princeton University Press (cit. on p. 2).

Kumar, M. et al. (2020). "First proof of the capability of wastewater surveillance for COVID-19 in India through detection of genetic material of SARS-CoV-2". In: *Sci. Total Environ.* 746, p. 141326. DOI: https://doi.org/10.1016/j.scitotenv.2020.141326 (cit. on p. 152).

Kumar, M. et al. (2021). "Unravelling the early warning capability of wastewater surveillance for COVID-19: A temporal study on SARS-CoV-2 RNA detection and need for the escalation". In: *Environ. Res.* 196, p. 110946. DOI: https://doi.org/10.1016/j.envres.2021.110946 (cit. on pp. 153, 168).

Kumar, N., T. Platini, and R. V. Kulkarni (2014). "Exact distributions for stochastic gene expression models with bursting and feedback". In: *Phys. Rev. Lett.* 113.26, p. 268105. DOI: 10.1103/PhysRevLett.113.268105 (cit. on p. 114).

Kuramoto, Y. (1984). *Chemical Oscillations, Waves, and Turbulence*. Vol. 19. DOI: 10.1007/978-3-642-69689-3 (cit. on p. 57).

Kuznetsov, Y. A. (2013). *Elements of applied bifurcation theory*. Vol. 112. Springer Science & Business Media, p. 591. DOI: 10.1007/b98848 (cit. on pp. 26, 181, 185–187, 201).

La Salle, J. P. (1976). *The stability of dynamical systems*. SIAM (cit. on p. 19).

Lade, S. J. and T. Gross (2012). "Early warning signals for critical transitions: A generalized modeling approach". In: *PLoS Computational Biology* 8.2. DOI: 10.1371/journal.pcbi.1002360 (cit. on pp. 21, 45, 51).

Lade, S. J. et al. (2013). "Regime shifts in a social-ecological system". In: *Theoretical Ecology* 6.3, pp. 359–372. DOI: 10.1007/s12080-013-0187-3 (cit. on p. 17).

Ladyman, J., J. Lambert, and K. Wiesner (2013). "What is a complex system?" In: *European Journal for Philosophy of Science* 3.1, pp. 33–67. DOI: 10.1007/s13194-012-0056-8 (cit. on p. 8).

Lahrich, S et al. (2021). "Review on the contamination of wastewater by {COVID-19} virus: Impact and treatment". In: *Science of The Total Environment* 751, p. 142325. DOI: https://doi.org/10.1016/j.scitotenv.2020.142325 (cit. on p. 152).

Lai, S. et al. (2020). "Effect of non-pharmaceutical interventions to contain COVID-19 in China". In: *Nature* 585.2, pp. 410–413. DOI: 10.1038/s41586-020-2293-x (cit. on pp. 64, 156).

Larsen, D. A. and K. R. Wigginton (2020). "Tracking COVID-19 with wastewater". In: *Nat. Biotechnol.* 38.10, pp. 1151–1153. DOI: https://doi.org/10.1038/s41587-020-0690-1 (cit. on p. 152).

Laurence, E. et al. (2019). "Spectral Dimension Reduction of Complex Dynamical Networks". In: *Physical Review X* 9.1, pp. 1–17. DOI: 10.1103/PhysRevX.9.011042 (cit. on p. 85).

Lee, D. and H. Park (2019). *Permutation entropies of short-term interest rates as an early-warning signal*. DOI: 10.1111/infi.12348 (cit. on p. 21).

Legrand, J. et al. (2007). "Understanding the dynamics of Ebola epidemics". In: *Epidemiology & Infection* 135.4, pp. 610–621. DOI: 10.1017/S0950268806007217 (cit. on pp. 64, 130).

Lenton, T. M. et al. (2012). "Early warning of climate tipping points from critical slowing down: comparing methods to improve robustness". In: *Philosophical Transactions of the Royal Society A* 370.1962, pp. 1185–1204. DOI: 10.1098/rsta.2011.0304 (cit. on pp. 17, 21, 45, 57, 76, 139).

Lestas, I. et al. (2008). "Noise in gene regulatory networks". In: *IEEE Transactions on Automatic Control* 53.Special Issue, pp. 189–200. DOI: 10.1109/TAC.2007.911347 (cit. on pp. 62, 88).

Lestas, I., G. Vinnicombe, and J. Paulsson (2010). "Fundamental limits on the suppression of molecular fluctuations". In: *Nature* 467.7312, pp. 174–178. DOI: 10.1038/nature09333 (cit. on pp. 62, 88, 112).

Lever, J. J. et al. (2020). "Foreseeing the future of mutualistic communities beyond collapse". In: *Ecology Letters* 23.1, pp. 2–15. DOI: 10.1111/ele.13401 (cit. on p. 83).

Li, C. and J. Wang (2015). "Quantifying the landscape for development and cancer from a core cancer stem cell circuit". In: *Cancer research* 75.13, pp. 2607–2618 (cit. on p. 177).

Li, M. et al. (2014). "Detecting tissue-specific early warning signals for complex diseases based on dynamical network biomarkers: study of type 2 diabetes by cross-tissue analysis". In: *Briefings in bioinformatics* 15.2, pp. 229–243. DOI: 10.1093/bib/bbt027 (cit. on p. 88).

Li, X. et al. (2021a). "Data-driven estimation of {COVID-19} community prevalence through wastewater-based epidemiology". In: *Sci. Total Environ.* P. 147947. DOI: https://doi.org/10.1016/j.scitotenv.2021.147947 (cit. on pp. 152, 157).

Li, Y. and S.-W. Zhang (2021b). "Resilience function uncovers the critical transitions in cancer initiation". In: *Briefings in Bioinformatics* 22.6, bbab175 (cit. on p. 88).

Liu, Q.-H. et al. (2018). "Measurability of the epidemic reproduction number in data-driven contact networks". In: *Proceedings of the National Academy of Sciences* 115.50, pp. 12680–12685 (cit. on p. 138).

Liu, R. et al. (2021). "Predicting local COVID-19 outbreaks and infectious disease epidemics based on landscape network entropy". In: *Science Bulletin*. DOI: 10.1016/j.scib.2021.03.022 (cit. on pp. 17, 131).

Liu, X. et al. (2020a). "Network resilience". In: *arXiv*, pp. 125–154. DOI: 10.2495/978-1-84564-562-5/09 (cit. on pp. 20, 133, 179).

Liu, X. M. et al. (2009). "Effect of multiplicative and additive noise on genetic transcriptional regulatory mechanism". In: *Physica A: Statistical Mechanics and its Applications* 388.4, pp. 392–398. DOI: 10.1016/j.physa.2008.10.030 (cit. on p. 106).

Liu, Y. et al. (2020b). "The reproductive number of COVID-19 is higher compared to SARS coronavirus". In: *Journal of Travel Medicine* 27.2. DOI: 10.1093/jtm/taaa021 (cit. on p. 66).

Livina, V. N., F. Kwasniok, and T. M. Lenton (2010). "Potential analysis reveals changing number of climate states during the last 60 kyr". In: *Climate of the Past* 6.1, pp. 77–82. DOI: 10.5194/cp-6-77-2010 (cit. on pp. 17, 45, 57, 71, 72).

Loppini, A., S. Filippi, and H. E. Stanley (2019). "Critical transitions in heterogeneous networks: Loss of low-degree nodes as an early warning signal". In: *Physical Review E* 99.4, pp. 1–5. DOI: 10.1103/PhysRevE.99.040301 (cit. on pp. 21, 178).

Lu, M. et al. (2013). "Tristability in cancer-associated microRNA-TF chimera toggle switch". In: *Journal of Physical Chemistry B* 117.42, pp. 13164–13174. DOI: 10.1021/jp403156m (cit. on p. 36).

Ma, H. et al. (2014). "Predicting time series from short-term high-dimensional data". In: *International Journal of Bifurcation and Chaos* 24.12, p. 1430033 (cit. on pp. 74, 179).

Mackey, M. C. and J. Milton (1987). "Dynamical diseases". In: *Neuro—ophthalmology* 21, p. 24. DOI: 10.1111/j.1749-6632.1987.tb48723.x (cit. on p. 18).

Macy, M. W. et al. (2021). "Polarization and Tipping Points". In: *PNAS Conference on Polarization* 118.50, e2102144118. DOI: 10.1073/pnas.2102144118/-/DCSupplemental.y (cit. on p. 17).

Maienschein-Cline, M., A. Warmflash, and A. R. Dinner (2010). "Defining cooperativity in gene regulation locally through intrinsic noise". In: *IET Systems Biology* 4.6, pp. 379–392. DOI: 10.1049/iet-syb.2009.0070 (cit. on p. 113).

Mandelj, S., I. Grabec, and E. Govekar (2002). "Statistical Approach To Modeling of Spatiotemporal Dynamics". In: *International Journal of Bifurcation and Chaos* 11.11, pp. 2731–2738. DOI: 10.1142/s0218127401003802 (cit. on p. 57).

Marchesseau, S. et al. (2013). "Personalization of a cardiac electromechanical model using reduced order unscented Kalman filtering from regional volumes". In: *Med. Im. An.* 17.7, pp. 816–829. DOI: 10.1016/j.media.2013.04.012 (cit. on p. 69).

Marinazzo, D. et al. (2019). "Synergy as a warning sign of transitions: The case of the two-dimensional Ising model". In: *Physical Review E* 99.4, p. 040101. DOI: 10.1103/PhysRevE.99.040101 (cit. on p. 38).

Markdahl, J., D. Proverbio, and J. Goncalves (2020). "Robust synchronization of heterogeneous robot swarms on the sphere". In: *Proceedings of the IEEE Conference on Decision and Control* 2020-Decem.Cdc, pp. 5798–5803. DOI: 10.1109/CDC42340.2020.9304268 (cit. on p. 15).

Markdahl, J. et al. (2021). "Almost global convergence to practical synchronization in the generalized Kuramoto model on networks over the n-sphere". In: *Communications Physics* 4.1, pp. 1–9. DOI: 10.1038/s42005-021-00689-y (cit. on p. 16).

Marković, D. and C. Gros (2014). "Power laws and self-organized criticality in theory and nature". In: *Physics Reports* 536.2, pp. 41–74. DOI: 10.1016/j.physrep.2013.11.002 (cit. on p. 38).

Matsumori, T., H. Sakai, and K. Aihara (2019). "Early-warning signals using dynamical network markers selected by covariance". In: *Physical Review E* 100.5, pp. 1–9. DOI: 10.1103/PhysRevE.100.052303 (cit. on pp. 21, 178).

Maturana, M. I. et al. (2020). "Critical slowing down as a biomarker for seizure susceptibility". In: *Nature Communications* 11.1, pp. 1–12. DOI: 10.1038/s41467-020-15908-3 (cit. on p. 57).

Max Roser Hannah Ritchie, E. O.-O. and J. Hasell (2020). "Coronavirus Pandemic (COVID-19)". In: *Our World in Data* (cit. on pp. 136, 137, 165).

May, R. (1977). "Thresholds and breakpoints in ecosystems with a multiplicity of stable states". In: *Nature* 269, pp. 471–477. DOI: 10.1038/267585a0 (cit. on pp. 17, 44).

Mayer, G. et al. (2021). "Implementing FAIR data management within the German Network for Bioinformatics Infrastructure (de. NBI) exemplified by selected use cases". In: *Brief. Bioinform.* DOI: 10.1093/bib/bbab010 (cit. on p. 7).

Mazzocchi, F. (2019). "Scientific research across and beyond disciplines". In: *EMBO reports* 20.6, pp. 1–6. DOI: 10.15252/embr.201947682 (cit. on p. 2).

McMahan, C. S. et al. (2021). "COVID-19 wastewater epidemiology: a model to estimate infected populations". In: *The Lancet Planetary Health* 5.12, e874–e881. DOI: https://doi.org/10.1016/S2542-5196(21)00230-8 (cit. on p. 152).

McNamee, R. (2003). "Confounding and confounders". In: *Occupational and environmental medicine* 60.3, pp. 227–234. DOI: 10.1136/oem.60.3.227 (cit. on p. 87).

Meijer, H. (2014). "Matcont Tutorial : ODE GUI version". In: *Change*, pp. 1–7. DOI: 10.1097/COH.0b013e328349592a (cit. on pp. 35, 66).

Meisel, C. et al. (2012a). "Failure of adaptive self-organized criticality during epileptic seizure attacks". In: *PLoS Computational Biology* 8.1. DOI: 10.1371/journal.pcbi.1002312 (cit. on p. 57).

Meisel, C. and C. Kuehn (2012b). "Scaling effects and spatio-temporal multilevel dynamics in epileptic seizures". In: *PLoS ONE* 7.2. DOI: 10.1371/journal.pone.0030371 (cit. on pp. 18, 57, 79, 205).

Meisel, C. et al. (2015). "Critical Slowing Down Governs the Transition to Neuron Spiking". In: *PLoS Computational Biology* 11.2, pp. 1–21. DOI: 10.1371/journal.pcbi.1004097 (cit. on pp. 44, 52, 57, 95).

Meng, Y. et al. (2020). "Tipping point and noise-induced transients in ecological networks". In: *Journal of the Royal Society Interface* 17.171, p. 20200645. DOI: 10.1098/rsif.2020.0645 (cit. on pp. 31, 178).

Miles, P. R. (2019). "pymcmcstat: A python package for bayesian inference using delayed rejection adaptive metropolis". In: *Journal of Open Source Software* 4.38, p. 1417. DOI: 10.21105/joss.01417 (cit. on p. 69).

Miller, P. B. et al. (2017). "Forecasting infectious disease emergence subject to seasonal forcing". In: *Theoretical Biology and Medical Modelling* 14.1, pp. 1–14 (cit. on p. 131).

Mills, E. (2004). *Regime Shifts in Lake Ecosystems: Pattern and Variation* (cit. on p. 17).

Milo, R. et al. (2002). "Network motifs: simple building blocks of complex networks". In: *Science* 298.5594, pp. 824–827 (cit. on pp. 8, 62).

Mitchell, M. (2009). *Complexity - a guided tour*. Oxford University Press (cit. on p. 8).

Miura, F., M. Kitajima, and R. Omori (2021). "Duration of {SARS-CoV-2} viral shedding in faeces as a parameter for wastewater-based epidemiology: Re-analysis of patient data using a shedding dynamics model". In: *Sci. Total Environ.* 769, p. 144549. DOI: https://doi.org/10.1016/j.scitotenv.2020.144549 (cit. on p. 157).

Mojtahedi, M. et al. (2016). "Cell Fate Decision as High-Dimensional Critical State Transition". In: *PLoS Biology* 14.12, pp. 1–28. DOI: 10.1371/journal.pbio.2000640 (cit. on pp. 18, 21, 41, 45, 46, 70, 76, 88, 109, 112, 127).

Mombaerts, L. "Dynamical modeling techniques for biological time series data". PhD thesis. University of Luxembourg (cit. on pp. 53, 58, 75, 177).

Montanari, A. N. et al. (2022a). "Functional observability and subspace reconstruction in nonlinear systems". In: *OrbiLu* (cit. on pp. 14, 177, 179).

Montanari, A. N. et al. (2022b). "Functional observability and target state estimation in large-scale networks". In: *Proceedings of the National Academy of Sciences* 119.1. DOI: 10.1073/pnas.2113750119 (cit. on p. 180).

Mora, T. and W. Bialek (2011). "Are biological systems poised at criticality?" In: *Journal of Statistical Physics* 144.2, pp. 268–302. DOI: 10.1007/s10955-011-0229-4 (cit. on p. 112).

Morales, I. O. et al. (2015). "Behavior of early warnings near the critical temperature in the two-dimensional Ising model". In: *PloS one* 10.6, e0130751. DOI: 10.1371/journal.pone.0130751 (cit. on p. 38).

Moris, N., C. Pina, and A. M. Arias (2016). "Transition states and cell fate decisions in epigenetic landscapes". In: *Nature Reviews Genetics* 17.11, pp. 693–703. DOI: 10.1038/nrg.2016.98 (cit. on pp. 41, 90, 177).

Mormann, F. et al. (2000). "Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients". In: *Physica D* 144.3-4, pp. 358–369. DOI: 10.1016/S0167-2789(00)00087-7 (cit. on p. 57).

Morone, F., G. Del Ferraro, and H. A. Makse (2019). "The k-core as a predictor of structural collapse in mutualistic ecosystems". In: *Nature Physics* 15.1, pp. 95–102. DOI: 10.1038/s41567-018-0304-8 (cit. on p. 178).

Moutsinas, G. and W. Guo (2020). "Node-Level Resilience Loss in Dynamic Complex Networks". In: *Scientific Reports* 10.1, pp. 1–12. DOI: 10.1038/s41598-020-60501-9 (cit. on p. 178).

Nagy, P. and P. Tasnádi (2013). "Zeeman catastrophe machines as a toolkit for teaching chaos". In: *European Journal of Physics* 35.1, p. 15018 (cit. on pp. 215, 216, 222).

Namachchivaya, N. S. and G. Leng (1990). "Equivalence of stochastic averaging and stochastic normal forms". In: *Journal of Applied Mechanics* 57.4, pp. 1011–1017. DOI: 10.1115/1.2897619 (cit. on pp. 29, 198, 204).

Naughton, C. C. et al. (2021). "Show us the data: Global {COVID-19} wastewater monitoring efforts, equity, and gaps". In: *medRxiv*, p. 2021.03.14.21253564 (cit. on pp. 152, 153).

Navarro, V. et al. (2002). "Seizure anticipation in human neocortical partial epilepsy". In: *Brain* 125.3, pp. 640–655. DOI: 10.1093/brain/awf048 (cit. on p. 57).

Navid Moghadam, N. et al. (2020). "Studying the performance of critical slowing down indicators in a biological system with a period-doubling route to chaos". In: *Physica A: Statistical Mechanics and its Applications* 544, p. 123396. DOI: 10.1016/j.physa.2019.123396 (cit. on p. 21).

Néant, N. et al. (2021). "Modeling {SARS-CoV-2} viral kinetics and association with mortality in hospitalized patients from the {French COVID} cohort". In: *Proc. Natl. Acad. Sci. USA* 118.8. DOI: https://doi.org/10.1073/pnas.2017962118 (cit. on pp. 156, 157).

Nemudryi, A. et al. (2020). "Temporal detection and phylogenetic assessment of {SARS-CoV-2} in municipal wastewater". In: *Cell Rep. Med.* 1.6, p. 100098. DOI: https://doi.org/10.1016/j.xcrm.2020.100098 (cit. on p. 152).

Nené, N. R., J. Garca-Ojalvo, and A. Zaikin (2012). "Speed-dependent cellular decision making in nonequilibrium genetic circuits". In: *PLoS ONE* 7.3, pp. 1–7. DOI: 10.1371/journal.pone.0032779 (cit. on p. 34).

Nishiura, H., N. M. Linton, and A. R. Akhmetzhanov (2020). "Serial interval of novel coronavirus (COVID-19) infections". In: *International Journal of Infectious Diseases* 93, pp. 284–286. DOI: 0.1016/j.ijid.2020.02.060 (cit. on p. 139).

Noakes, L. (1991). "The Takens embedding theorem". In: *International Journal of Bifurcation and Chaos* 1.04, pp. 867–872 (cit. on p. 76).

Nolting, B. and K. Abbott (2016). "Balls, cups, and quasi-potentials: quantifying stability in stochastic systems". In: *Ecology* 97.4, pp. 850–864. DOI: 10.1890/15-1047.1 (cit. on pp. 71, 205).

Norman, T. M. et al. (2015). "Stochastic switching of cell fate in microbes". In: *Annual review of microbiology* 69, pp. 381–403. DOI: 10.1146/annurev-micro-091213-112852 (cit. on pp. 62, 88, 91).

Nyman, K., P. Ashwin, and P. Ditlevsen (2020). "Bifurcation of critical sets and relaxation oscillations in singular fast-slow systems". In: *Nonlinearity* 33.6, p. 2853. DOI: 10.1088/1361-6544/ab7292 (cit. on p. 27).

O'Brien, D. A. and C. F. Clements (2021). "Early warning signals predict emergence of COVID-19 waves". In: *medRxiv* (cit. on p. 131).

O'Connor, T. and Y. H. Wong (2015). "Emergent Properties". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Summer 201. Metaphysics Research Lab, Stanford University (cit. on p. 10).

O'Dea, E. B. and J. M. Drake (2019). "Disentangling reporting and disease transmission". In: *Theoretical Ecology* 12.1, pp. 89–98 (cit. on pp. 54, 131, 147).

O'Keeffe, P. E. and S. Wieczorek (2020). "Tipping Phenomena and Points of No Return in Ecosystems: Beyond Classical Bifurcations". In: *SIAM Journal on Applied Dynamical Systems* 19.4, pp. 2371–2402. DOI: 10.1137/19M1242884 (cit. on p. 28).

Olde Rikkert, M. G. et al. (2016). "Slowing Down of Recovery as Generic Risk Marker for Acute Severity Transitions in Chronic Diseases". In: *Critical Care Medicine* 44.3, pp. 601–606. DOI: 10.1097/CCM.0000000000001564 (cit. on p. 21).

Olier, I. et al. (2021). "How machine learning is impacting research in atrial fibrillation: implications for risk prediction and future management". In: *Cardiovascular Research* 117.7, pp. 1700–1717. DOI: 10.1093/cvr/cvab169 (cit. on p. 178).

Olthof, M. et al. (2020). "Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders". In: *Clinical Psychological Science* 8.1, pp. 25–35. DOI: 10.1177/2167702619865969 (cit. on pp. 17, 21, 33).

Omidvarnia, A. et al. (2018). "Range entropy: A bridge between signal complexity and self-similarity." In: *arXiv Information Theory (cs.IT)*, pp. 1–12. DOI: arXiv:1809.06500v1 (cit. on p. 57).

O'Regan, S. M. et al. (2013). "Theory of early warning signals of disease emergenceand leading indicators of elimination". In: *Theoretical Ecology* 6.3, pp. 333–357. DOI: 10.1007/s12080-013-0185-5 (cit. on pp. 18, 53, 65, 73, 131–135).

O'Regan, S. M., J. W. Lillie, and J. M. Drake (2016). "Leading indicators of mosquito-borne disease elimination". In: *Theoretical Ecology* 9.3, pp. 269–286. DOI: 10.1007/s12080-015-0285-5 (cit. on pp. 18, 65).

O'Regan, S. M. and D. L. Burton (2018). "How stochasticity influences leading indicators of critical transitions". In: *Bulletin of mathematical biology* 80.6, pp. 1630–1654. DOI: 10.1007/s11538-018-0429-z (cit. on pp. 29, 52, 88–90, 105, 135, 144, 210).

Oreskes, N., K. Shrader-Frechette, and K. Belitz (1994). "Verification, validation, and confirmation of numerical models in the earth sciences". In: *Science* 263.5147, pp. 641–646. DOI: 10.1126/science.263.5147.641 (cit. on pp. 3, 5).

Orzack, S. H. (2012). "The philosophy of modelling or does the philosophy of biology have any use?" In: *Philosophical Transactions of the Royal Society B* 367.1586, pp. 170–180 (cit. on p. 2).

Osorio, I. et al. (2009). "Pharmaco-resistant seizures: Self-triggering capacity, scale-free properties and predictability?" In: *European Journal of Neuroscience* 30.8, pp. 1554–1558. DOI: 10.1111/j.1460-9568.2009.06923.x (cit. on p. 57).

Osorio, I. et al. (2010). "Epileptic seizures: Quakes of the brain?" In: *Physical Review E* 82.2, pp. 1–13. DOI: 10.1103/PhysRevE.82.021919 (cit. on pp. 56, 57).

Ott, E. (2002). *Chaos in dynamical systems*. Cambridge university press (cit. on pp. 9, 29, 75).

Ozbudak, E. M. et al. (2004). "Multistability in the lactose utilization network of Escherichia coli". In: *Nature* 427.6976, pp. 737–740. DOI: 10.1038/nature02298 (cit. on p. 62).

Paine, R. T. (1966). "Food web complexity and species diversity". In: *The American Naturalist* 100.910, pp. 65–75 (cit. on p. 17).

Pal, M., S. Ghosh, and I. Bose (2015). "Non-genetic heterogeneity, criticality and cell differentiation". In: *Physical Biology* 12.1, p. 016001. DOI: 10.1088/1478-3975/12/1/016001 (cit. on p. 40).

Pan, L. et al. (2020). "Phase diagrams of interacting spreading dynamics in complex networks". In: *Physical Review Research* 2.2, pp. 1–16. DOI: 10.1103/PhysRevResearch.2.023233 (cit. on p. 85).

Pan, W. et al. (2010). "Robust H∞ feedback control for uncertain stochastic delayed genetic regulatory networks with additive and multiplicative noise". In: *Internatiional Journal of Robust and Nonlinear Control* 18.5, pp. 557–569. DOI: 10.1002/rnc (cit. on pp. 106, 126).

Papoulis, A. and S. U. Pillai (2002). *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education (cit. on pp. 29, 196, 203, 204, 206, 208, 210).

Parish, L. M. et al. (2004). "Long-range temporal correlations in epileptogenic and non-epileptogenic human hippocampus". In: *Neuroscience* 125.4, pp. 1069–1076. DOI: 10.1016/j.neuroscience.2004.03.002 (cit. on p. 57).

Park, M. et al. (2020). "A systematic review of COVID-19 epidemiology based on current evidence". In: *Journal of clinical medicine* 9.4, p. 967. DOI: 10.3390/jcm9040967 (cit. on p. 138).

Pathria, R. and P. Beale (2011). *Statistical Mechanics*. 3rd. Elsevier Ltd (cit. on pp. 17, 22, 30, 38, 192, 199).

Pavithran, I. and R. I. Sujith (2021). "Effect of rate of change of parameter on early warning signals for critical transitions". In: *Chaos* 31.1. DOI: 10.1063/5.0025533 (cit. on p. 21).

Peccia, J. et al. (2020). "Measurement of {SARS-CoV}-2 {RNA} in wastewater tracks community infection dynamics". In: *Nat. Biotechnol.* 38.10, pp. 1164–1167. DOI: https://doi.org/10.1038/s41587-020-0684-z (cit. on p. 152).

Pedersen, S. F., Y.-C. Ho, and Others (2020). "SARS-CoV-2: a storm is raging". In: *The Journal of clinical investigation* 130.5 (cit. on p. 130).

Pedraza, J. M. and J. Paulsson (2008). "Effects of molecular memory and bursting on fluctuations in gene expression". In: *Science* 319.5861, pp. 339–343. DOI: 10.1126/science.1144331 (cit. on p. 106).

Peng, L. et al. (2020). "Epidemic analysis of COVID-19 in China by dynamical modeling". In: *arXiv* preprint a, pp. 1–18. DOI: 10.1101/2020.02.16.20023465 (cit. on p. 132).

Perryman, C. and S. Wieczorek (2014). "Adapting to a changing environment : non-obvious thresholds in multi-scale systems". In: *Proceedings of the Royal Society A* 470.2170, p. 20140226. DOI: 10.1098/rspa.2014.0226 (cit. on p. 34).

Petersen, S. V., D. P. Schrag, and P. U. Clark (2013). "A new mechanism for Dansgaard-Oeschger cycles". In: *Paleoceanography* 28.1, pp. 24–30. DOI: 10.1029/2012PA002364 (cit. on p. 57).

Petraitis, P. S. and S. R. Dudgeon (2004). "Detection of alternative stable states in marine communities". In: *Journal of Experimental Marine Biology and Ecology* 300.1-2, pp. 343–371. DOI: 10.1016/j.jembe.2003.12.026 (cit. on p. 17).

Petropoulos, F. and S. Makridakis (2020). "Forecasting the novel coronavirus {COVID-19}". In: *PloS one* 15.3, e0231236. DOI: 10.1371/journal.pone.0231236 (cit. on p. 167).

Pfeuty, B. and K. Kaneko (2014). "Reliable binary cell-fate decisions based on oscillations". In: *Physical Review E* 89.2, pp. 1–6. DOI: 10.1103/PhysRevE.89.022707 (cit. on p. 41).

Phillips, B., M. Anand, and C. T. Bauch (2020). "Spatial early warning signals of social and epidemiological tipping points in a coupled behaviour-disease network". In: *Scientific reports* 10.1, pp. 1–12 (cit. on p. 131).

Pikovsky, A., M. Rosenblum, and J. Kurths (2001). *Synchronization - a universal concept in nonlinear science*. Cambridge university press (cit. on p. 57).

Pisarchik, A. N. and U. Feudel (2014). "Control of multistability". In: *Physics Reports* 540.4, pp. 167–218 (cit. on p. 61).

Pollán, M. et al. (2020). "Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study". In: *The Lancet* 396.10250, pp. 535–544. DOI: 10.1016/S0140-6736(20)31483-5 (cit. on p. 153).

Pomerening, J. R. (2008). "Uncovering mechanisms of bistability in biological systems". In: *Current opinion in biotechnology* 19.4, pp. 381–388 (cit. on p. 177).

Popper, K. (2014). *Conjectures and refutations: The growth of scientific knowledge*. Routledge (cit. on p. 3).

Poston, T. and A. E. R. Woodcock (1973). "Zeeman's catastrophe machine". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 74. 2. Cambridge University Press, pp. 217–226 (cit. on p. 215).

Poston, T. and I. Stewart (1979). *Catastrophe theory and its applications*. Courier Corporation (cit. on pp. 215, 216, 218).

Pötzsche, C. and M. Rasmussen (2006). "Taylor approximation of integral manifolds". In: *Journal of Dynamics and Differential Equations* 18.2, pp. 427–460. DOI: 10.1007/s10884-006-9011-8 (cit. on p. 185).

Proverbio, D. "A Multi Agent System approach to complex micro-biological systems - In silico simulation of Dictyostelium Discoideum colonies". Master thesis. University of Turin (cit. on p. 6).

Proverbio, D. et al. (2020). "Assessing the robustness of decentralized gathering: a multi-agent approach on micro-biological systems". In: *Swarm Intelligence* 14.4, pp. 313–331. DOI: 10.1007/s11721-020-00186-y (cit. on pp. 6, 13).

Proverbio, D. et al. (2021). "Dynamical SPQEIR model assesses the effectiveness of non-pharmaceutical interventions against COVID-19 epidemic outbreaks". In: *PLoS ONE* 16.5, e0252019. DOI: 10.1371/journal.pone.0252019 (cit. on pp. 15, 64–66, 69, 129–133, 154).

Proverbio, D. et al. (2022a). "Buffering variability in cell regulation motifs close to criticality". In: *Physical Review E* 106 (3), p. L032402. DOI: https://doi.org/10.1103/PhysRevE.106.L032402 (cit. on pp. 13, 111).

Proverbio, D. et al. (2022b). "Model-based assessment of COVID-19 epidemic dynamics by wastewater analysis". In: *Science of The Total Environment* 827, p. 154235. DOI: 10.1016/j.scitotenv.2022.154235 (cit. on pp. 14, 69, 151, 154, 165, 166, 169).

Proverbio, D. et al. (2022c). "Performance of early warning signals for disease re-emergence: A case study on COVID-19 data". In: *PLOS Computational Biology* 18.3, e1009958. DOI: 10.1371/journal.pcbi.1009958 (cit. on pp. 14, 76, 88, 122, 129, 136, 139, 140, 144, 210).

Qian, H. (2012). "Cooperativity in cellular biochemical processes: Noise-enhanced sensitivity, fluctuating enzyme, bistability with nonlinear feedback, and other mechanisms for sigmoidal responses". In: *Annual Review of Biophysics* 41.1, pp. 179–204. DOI: 10.1146/annurev-biophys-050511-102240 (cit. on p. 112).

Quail, T., A. Shrier, and L. Glass (2015). "Predicting the onset of period-doubling bifurcations in noisy cardiac systems". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.30, pp. 9358–9363. DOI: 10.1073/pnas.1424320112 (cit. on p. 18).

Quilliam, R. S. et al. (2020). "COVID-19: The environmental implications of shedding SARS-CoV-2 in human faeces". In: *Environ. Int.* 140, p. 105790. DOI: https://doi.org/10.1016/j.envint.2020.105790 (cit. on p. 152).

Radchuk, V. et al. (2019). "The dimensionality of stability depends on disturbance type". In: *Ecology Letters* 22.4, pp. 674–684. DOI: 10.1111/ele.13226 (cit. on p. 111).

Randazzo, W. et al. (2020). "Metropolitan wastewater analysis for {COVID-19} epidemiological surveillance". In: *Int. J. Hyg. Envir. Heal.* 230, p. 113621. DOI: https://doi.org/10.1016/j.ijheh.2020.113621 (cit. on p. 152).

Rasheed, K. et al. (2020). "Machine learning for predicting epileptic seizures using EEG signals: A review". In: *IEEE Reviews in Biomedical Engineering* 14, pp. 139–155. DOI: 10.1109/RBME.2020.3008792 (cit. on p. 178).

Ravnås, E. (2008). "Continuation and Bifurcation software in MATLAB". In: *57* August (cit. on p. 66).

Reeves, K. et al. (2021). "High-resolution within-sewer SARS-CoV-2 surveillance facilitates informed intervention". In: *Water Res.* 204, p. 117613. DOI: 10.1016/j.watres.2021.117613 (cit. on p. 152).

Reiner, R. C. et al. (2020). "Modeling COVID-19 scenarios for the United States". In: *Nature Medicine* 27, pp. 94–105. DOI: 10.1038/s41591-020-1132-9 (cit. on p. 136).

Ren, H. and D. Watts (2015). "Early warning signals for critical transitions in power systems". In: *Electric Power Systems Research* 124, pp. 173–180. DOI: 10.1016/j.epsr.2015.03.009 (cit. on p. 21).

Reno, C. et al. (2020). "Forecasting COVID-19-Associated Hospitalizations under Different Levels of Social Distancing in Lombardy and Emilia-Romagna, Northern Italy: Results from an Extended SEIR Compartmental Model". In: *Journal of clinical medicine* 9.5, p. 1492. DOI: 10.3390/jcm9051492 (cit. on p. 136).

Reyes, B. C., I. Otero-Muras, and V. A. Petyuk (2020). "A general technique for the detection of switch-like bistability in chemical reaction networks governed by mass action kinetics with conservation laws". In: *bioRxiv*, p. 2020.11.06.372235 (cit. on p. 44).

Richard, A. et al. (2016). "Single-Cell-Based Analysis Highlights a Surge in Cell-to-Cell Molecular Variability Preceding Irreversible Commitment in a Differentiation Process". In: *PLoS Biology* 14.12, pp. 1–35. DOI: 10.1371/journal.pbio.1002585 (cit. on pp. 40, 88).

Ricon-Becker, I. et al. (2020). "A seven-day cycle in COVID-19 infection, hospitalization, and mortality rates: Do weekend social interactions kill susceptible people?" In: *medRxiv* 10.1101/20 (cit. on p. 139).

Risken, H. and T. K. Caugheyz (1996). "The Fokker-Planck Equation: Methods of Solution and Application, 2nd ed." In: *Journal of Applied Mechanics* 58.3, pp. 860–860. DOI: 10.1115/1.2897281 (cit. on pp. 198, 203, 211).

Ritchie, P. and J. Sieber (2016). "Early-warning indicators for rate-induced tipping". In: *Chaos* 26.9, pp. 1–15. DOI: 10.1063/1.4963012 (cit. on p. 34).

Ritchie, P., Ö. Karabacak, and J. Sieber (2019). "Inverse-square law between time and amplitude for crossing tipping thresholds". In: *Proceedings of the Royal Society A* 475.2222, p. 20180504. DOI: 10.1098/rspa.2018.0504 (cit. on p. 17).

Ritchie Jr, P. D. L. (2016). "Early-warning indicators for tipping points" (cit. on p. 211).

Rivieccio, B. A. et al. (2020). "Heterogeneity of COVID-19 outbreak in Italy". In: *Acta Bio Medica: Atenei Parmensis* 91.2, p. 31. DOI: 10.23750/abm.v91i2.9579 (cit. on p. 136).

Rocha, J. C. et al. (2018). "Cascading regime shifts within and across scales". In: *Science* 362.6421, pp. 1379–1383. DOI: 10.1101/364620 (cit. on pp. 18, 179).

Roda, W. C. et al. (2020). "Why is it difficult to accurately predict the COVID-19 epidemic?" In: *Infectious Disease Modelling* 5, pp. 271–281. DOI: https://doi.org/10.1016/j.idm.2020.03.001 (cit. on p. 156).

Rodrigues, A. C. et al. (2014). "Brain network dynamics characterization in epileptic seizures". In: *The European Physical Journal Special Topics* 223.13, pp. 2933–2941. DOI: 10.1140/epjst/e2014-02306-8 (cit. on p. 57).

Rodríguez-Méndez, V. et al. (2016). "Percolation-based precursors of transitions in extended systems". In: *Scientific Reports* 6, pp. 1–11. DOI: 10.1038/srep29552 (cit. on pp. 21, 31, 178).

Rosen, A. et al. (2010). "Experimental observation of critical phenomena in a laser light system". In: *Physical review letters* 105.1, p. 013905. DOI: 10.1103/PhysRevLett.105.013905 (cit. on p. 43).

Rosso, O. A. et al. (2006). "EEG analysis using wavelet-based information tools". In: *Journal of Neuroscience Methods* 153.2, pp. 163–182. DOI: 10.1016/j.jneumeth.2005.10.009 (cit. on p. 79).

Rosso, O. A. et al. (2007). "Distinguishing noise from chaos". In: *Physical Review Letters* 99.15, pp. 1–4. DOI: 10.1103/PhysRevLett.99.154102 (cit. on p. 57).

Rubino, D., K. A. Robbins, and N. G. Hatsopoulos (2006). "Propagating waves mediate information transfer in the motor cortex". In: *Nature neuroscience* 9.12, p. 1549. DOI: 10.1038/nn1802 (cit. on p. 57).

Rundle, J. B. et al. (2002). "Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neurobiology, and forecasting". In: *Proceedings of the National Academy of Sciences* 99.Supplement 1, pp. 2514–2521. DOI: 10.1073/pnas.012581899 (cit. on p. 39).

Ryan, A. J. (2007). "Emergence is coupled to scope, not level". In: *Complexity* 13.2, pp. 67–77 (cit. on p. 10).

Saaid, F., D. Nur, and R. King (2012). "Change points detection of vector autoregressive model using SDVAR algorithm". In: *Proceedings of the Fifth Annual ASEARC Conference - Looking to the future* (cit. on p. 73).

Saetzler, K., C. Sonnenschein, and M. A. Soto (2011). "Systems biology beyond networks: generating order from disorder through self-organization". In: *Seminars in cancer biology*. Vol. 21. 3. Elsevier, pp. 165–174 (cit. on pp. 2, 3).

Sala-Comorera, L. et al. (2021). "Decay of infectious {SARS-CoV-2} and surrogates in aquatic environments". In: *Water Res.* P. 117090. DOI: 10.1016/j.watres.2021.117090 (cit. on p. 156).

Santillán, M. (2008). "On the use of the Hill functions in mathematical models of gene regulatory networks". In: *Mathematical Modelling of Natural Phenomena* 3.2, pp. 85–97 (cit. on pp. 62, 112, 113).

Santosh, K. C. (2020). "COVID-19 prediction models and unexploited data". In: *Journal of medical systems* 44.9, pp. 1–4. DOI: https://doi.org/10.1007/s10916-020-01645-z (cit. on p. 168).

Sardanyés, J. and T. Alarcón (2018). "Noise-induced bistability in the fate of cancer phenotypic quasispecies: A bit-strings approach". In: *Scientific Reports* 8.1, pp. 1–10. DOI: 10.1038/s41598-018-19552-2 (cit. on p. 40).

Sarkar, S. et al. (2019). "Anticipating critical transitions in epithelial-hybrid-mesenchymal cell-fate determination". In: *Proceedings of the National Academy of Sciences of the United States of America* 116.52, pp. 26343–26352. DOI: 10.1073/pnas.1913773116 (cit. on p. 36).

Saunders, P. T. (2018). "The organism as a dynamical system". In: *Thinking about biology*. CRC Press, pp. 41–63 (cit. on p. 28).

Scheffer, M et al. (2012). "Anticipating Critical Transitions". In: *Science* 338.6105, pp. 344–348. DOI: 10.1126/science.1225244 (cit. on p. 51).

Scheffer, M. (2009). *Critical transitions in nature and society*. Princeton University Press (cit. on pp. 17, 19, 26, 32, 35, 45, 47, 49, 51, 90, 112, 187, 203).

Scheffer, M. and S. R. Carpenter (2003). "Catastrophic regime shifts in ecosystems: linking theory to observation". In: *Trends in ecology & evolution* 18.12, pp. 648–656 (cit. on p. 51).

Scheffer, M. et al. (2009). "Early-warning signals for critical transitions". In: *Nature* 461.7260, pp. 53–59. DOI: 10.1038/nature08227 (cit. on pp. 21, 44, 112, 117, 121, 209).

Scheffer, M. et al. (2015). "Generic Indicators of Ecological Resilience: Inferring the Chance of a Critical Transition". In: *Annual Review of Ecology, Evolution, and Systematics* 46.1, pp. 145–167. DOI: 10.1146/annurev-ecolsys-112414-054242 (cit. on pp. 21, 60).

Schiff, S. J. et al. (1994). "Controlling chaos in the brain". In: *Nature* 370, pp. 615–620. DOI: 10.1038/370615a0 (cit. on p. 57).

Schittkowski, K. (2002). *Numerical data fitting in dynamical systems: a practical introduction with applications and software*. Vol. 77. Springer Science & Business Media (cit. on p. 68).

Scholz, J. P., J. A. S. Kelso, and G Schöner (1987). "Nonequilibrium phase transitions in coordinated biological motion: critical slowing down and switching time". In: *Physics Letters A* 123.8, pp. 390–394. DOI: 10.1016/0375-9601(87)90038-7 (cit. on pp. 21, 112).

Scott, J. F., M Sargent III, and C. D. Cantrell (1975). "Laser-phase transition analogy: Application to first-order transitions". In: *Optics Communications* 15.1, pp. 13–16. DOI: 10.1016/0030-4018(75)90171-6 (cit. on p. 38).

Segel, L. A. (1988). "On the validity of the steady state assumption of enzyme kinetics". In: *Bulletin of mathematical biology* 50.6, pp. 579–593. DOI: 10.1007/BF02460092 (cit. on p. 89).

Sharma, Y., P. S. Dutta, and A. K. Gupta (2016a). "Anticipating regime shifts in gene expression: The case of an autoactivating positive feedback loop". In: *Physical Review E* 93.3, pp. 1–13. DOI: 10.1103/PhysRevE.93.032404 (cit. on pp. 33, 52, 62, 63, 94, 99, 104, 105, 112, 113, 211).

Sharma, Y., P. S. Dutta, and A. K. Gupta (2016b). "Anticipating regime shifts in gene expression: The case of an autoactivating positive feedback loop". In: *Physical Review E* 93.3, p. 32404 (cit. on p. 126).

Sharma, Y. and P. S. Dutta (2017). "Regime shifts driven by dynamic correlations in gene expression noise". In: *Physical Review E* 96.2, pp. 1–14. DOI: 10.1103/PhysRevE.96.022409 (cit. on pp. 18, 33).

Sheffer, M. et al. (2001). "Catastrophic shifts in ecosystems". In: *Nature* 413.6856, pp. 591–596. DOI: 10.1038/35098000 (cit. on p. 17).

Sherry, D. (2006). "Mathematical reasoning: Induction, deduction and beyond". In: *Studies in History and Philosophy of Science Part A* 37.3, pp. 489–504. DOI: 10.1016/j.shpsa.2005.06.012 (cit. on pp. 2, 5).

Shi, J., T. Li, and L. Chen (2016). "Towards a critical transition theory under different temporal scales and noise strengths". In: *Physical Review E* 93.3, pp. 1–13. DOI: 10.1103/PhysRevE.93.032137 (cit. on pp. 19, 21, 29, 30, 36, 70).

Shi, J., K. Aihara, and L. Chen (2021). "Dynamics-based data science in biology". In: *National Science Review* 8.5, nwab029 (cit. on p. 74).

Siciliano, V. et al. (2013). "MiRNAs confer phenotypic robustness to gene networks by suppressing biological noise". In: *Nature communications* 4.1, pp. 1–7 (cit. on p. 112).

Sidney, R. C., M. Dunlop, and M. B. Elowitz (2010). "A synthetic three-color reporter framework for monitoring genetic regulation and noise". In: *Journal of Biological Engineering* 4.10, pp. 1–12. DOI: 10.1186/1754-1611-4-10 (cit. on p. 106).

Sima, A., A. Paul, and M. Schulz (2004). "The Younger Dryas—an intrinsic feature of late Pleistocene climate change at millennial timescales". In: *Earth and Planetary Science Letters* 222.3-4, pp. 741–750. DOI: 10.1016/j.epsl.2004.03.026 (cit. on p. 57).

Simon, M. A. (1996). "Beyond inductive and deductive reasoning: The search for a sense of knowing". In: *Educational Studies in mathematics* 30.2, pp. 197–210 (cit. on pp. 2, 3).

Sjoedin, H. et al. (2020). "COVID-19 healthcare demand and mortality in Sweden in response to non-pharmaceutical mitigation and suppression scenarios". In: *International Journal of Epidemiology*. DOI: 10.1093/ije/dyaa121 (cit. on pp. 64, 65).

Skogestad, S. and I. Postlethwaite (2005). *Multivariable Feedback Control: Analysis and Design, 2nd Edition*. Wiley and Sons. DOI: 978-0-470-01167-6 (cit. on p. 39).

Smits, W. K., O. P. Kuipers, and J.-W. Veening (2006). "Phenotypic variation in bacteria: the role of feedback regulation". In: *Nature Reviews Microbiology* 4.4, pp. 259–271. DOI: 10.1038/nrmicro1381 (cit. on p. 61).

Smolen, P., D. A. Baxter, and J. H. Byrne (1998). "Frequency selectivity, multistability, and oscillations emerge from models of genetic regulatory systems". In: *American Journal of Physiology-Cell Physiology* 274.2, pp. C531–C542 (cit. on p. 112).

Snijder, B. and L. Pelkmans (2011). "Origins of regulated cell-to-cell variability". In: *Nature reviews Molecular cell biology* 12.2, pp. 119–125. DOI: 10.1038/nrm3044 (cit. on p. 127).

Snoeck, C. J. et al. (2020). "Prevalence of SARS-CoV-2 infection in the Luxembourgish population – the CON-VINCE study". In: *medRxiv*, p. 2020.05.11.20092916. DOI: 10.1101/2020.05.11.20092916 (cit. on pp. 153, 160).

Solé, R. V. (2003). "Phase transitions in unstable cancer cell populations". In: *European Physical Journal B* 35.1, pp. 117–123. DOI: 10.1140/epjb/e2003-00262-8 (cit. on pp. 18, 38).

Sornette, D. (2013). *Critical phenomena in natural sciences*. Vol. 53. 9, pp. 1689–1699. DOI: 10.1017/CBO9781107415324.004 (cit. on pp. 17, 39, 56).

Southall, E., M. J. Tildesley, and L. Dyson (2020). "Prospects for detecting early warning signals in discrete event sequence data: Application to epidemiological incidence data." In: *PLoS Comput Biol* 16.9, e1007836. DOI: 10.1371/journal.pcbi.1007836 (cit. on pp. 54, 65, 80, 89, 98, 99, 109, 131, 133, 136, 144, 146, 147).

Southall, E. et al. (2021). "Early warning signals of infectious disease transitions: A review". In: *J. of the Royal Society Interface* 18.182, p. 20210555. DOI: 10.1098/rsif.2021.0555 (cit. on pp. 21, 53, 54, 131).

Stamovlasis, D. (2014). "Bifurcation and hysteresis effects in student performance: The signature of complexity and chaos in educational research". In: *Complicity: An International Journal of Complexity and Education* 11.2 (cit. on p. 194).

Stanoev, A., C. Schröter, and A. Koseska (2021). "Robustness and timing of cellular differentiation through population-based symmetry breaking". In: *Development* 148.3, dev197608. DOI: 10.1242/dev.197608 (cit. on p. 40).

Stapelberg, N. J. C. et al. (2018). "From Feedback Loop Transitions to Biomarkers in the Psycho-Immune-Neuroendocrine Network: Detecting the Critical Transition from Health to Major Depression". In: *Neuroscience & Biobehavioral Reviews* 90, pp. 1–15. DOI: 10.1016/j.neubiorev.2018.03.005 (cit. on p. 17).

Stefanescu, R. A., R. G. Shivakeshavan, and S. S. Talathi (2012). "Computational models of epilepsy". In: *Seizure* 21.10, pp. 748–759. DOI: 10.1016/j.seizure.2012.08.012 (cit. on p. 57).

Storch, L. and S. Day (2019a). "Towards the prediction of critical transitions in spatially extended populations with cubical homology". In: 92, pp. 31–48. DOI: 10.1090/conm/736/14846 (cit. on p. 179).

Storch, L. S. and S. L. Day (2019b). "Towards the prediction of critical transitions in spatially extended populations with cubical homology". In: *arXiv preprint arXiv:1912.01021* 92, pp. 31–48. DOI: 10.1090/conm/736/14846 (cit. on p. 45).

Stovner, B. N. et al. (2018). "Attitude estimation by multiplicative exogenous Kalman filter". In: *Automatica* 95, pp. 347–355. DOI: https://doi.org/10.1016/j.automatica.2018.05.038 (cit. on p. 69).

Strogatz, S. H. (2018). *Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering*. CRC press (cit. on pp. 24, 26, 27, 62, 63, 112, 181, 183, 186, 187, 192, 205).

Su, Y. et al. (2019). "Phenotypic heterogeneity and evolution of melanoma cells associated with targeted therapy resistance". In: *PLoS Computational Biology* 15.6, pp. 1–22. DOI: 10.1371/journal.pcbi.1007034 (cit. on pp. 29, 211).

Sugihara, G. (1994). "Nonlinear forecasting for the classification of natural time series". In: *Philosophical Transactions of the Royal Society of London. Series A* 348.1688, pp. 477–495. DOI: 10.1098/rsta.1994.0106 (cit. on p. 74).

Sugihara, G. et al. (1999). "Residual delay maps unveil global patterns of atmospheric nonlinearity and produce improved local forecasts". In: *Proceedings of the National Academy of Sciences* 96.25, pp. 14210–14215. DOI: 10.1073/pnas.96.25.14210 (cit. on p. 74).

Suhandynata, R. T. et al. (2021). "SARS-CoV-2 serology status detected by commercialized platforms distinguishes previous infection and vaccination adaptive immune responses". In: *The Journal of Applied Laboratory Medicine* 6.5, pp. 1109–1122. DOI: 10.1093/jalm/jfab080 (cit. on p. 171).

Sutherland, J. P. (1974). "Multiple stable points in natural communities". In: *The American Naturalist* 108.964, pp. 859–873 (cit. on p. 17).

Syafruddin, S. and M. S. M. Noorani (2012). "SEIR model for transmission of dengue fever in Selangor Malaysia". In: *IJMPS* 9, pp. 380–389. DOI: 10.1142/S2010194512005454 (cit. on p. 64).

Systrom, K., T. Vladek, and M. Krieger (2020). *Rt.live*. \url{https://github.com/rtcovidlive/covid-model} (cit. on p. 138).

Taha, A. A. and A. Hanbury (2015). "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool". In: *BMC medical imaging* 15.1, pp. 1–28. DOI: 10.1186/s12880-015-0068-x (cit. on p. 179).

Takens, F. (1981). "Detecting strange attractors in turbulence". In: *Dynamical systems and turbulence, Warwick 1980*. Springer, pp. 366–381 (cit. on p. 74).

Tan, J., L. Zhao, and H. Chen (2019). "A meta-analysis of the effectiveness of gradual versus abrupt smoking cessation". In: *Tobacco Induced Diseases* 17. DOI: 10.18332/tid/100557 (cit. on p. 17).

Taylor, J. R. (1997). *An Introduction to Error Analysis*. Mill Valley, California: University Science Books (cit. on pp. 68, 96, 118, 139).

Tejedor, A. et al. (2017). "Network robustness assessed within a dual connectivity framework: joint dynamics of the Active and Idle Networks". In: *Scientific reports* 7.1, pp. 1–10. DOI: 10.1038/s41598-017-08714-3 (cit. on p. 20).

Terry, J. R., O. Benjamin, and M. P. Richardson (2012). "Seizure generation: the role of nodes and networks". In: *Epilepsia* 53.9, e166–e169. DOI: 10.1111/j.1528-1167.2012.03560.x (cit. on p. 57).

Thom, R. et al. (1977). "Structural Stability and Morphogenesis". In: *Science* 196.5631, pp. 1270–1272. DOI: 10.1038/269759a0 (cit. on pp. 38, 186, 193, 216, 218, 222).

Thomas, P., N. Popović, and R. Grima (2014). "Phenotypic switching in gene regulatory networks". In: *Proceedings of the National Academy of Sciences* 111.19, pp. 6994–6999 (cit. on p. 61).

Thompson, J. M. T. and J. Sieber (2011). "Predicting climate tipping as a noisy bifurcation: a review". In: *International Journal of Bifurcation and Chaos* 21.02, pp. 399–423. DOI: 10.1142/S0218127411028519 (cit. on pp. 17, 19, 29, 32, 51, 181, 202, 209).

Thompson, J. and S. Wattam (2021). "Estimating the impact of interventions against COVID-19: from lockdown to vaccination". In: *PloS one* 16.12, e0261330. DOI: 10.1371/journal.pone.0261330 (cit. on p. 130).

Tiwari, S. B. et al. (2021). "Surveillance of Wastewater for Early Epidemic Prediction (SWEEP): Environmental and health security perspectives in the post COVID-19 Anthropocene". In: *Environ. Res.* 195, p. 110831. DOI: https://doi.org/10.1016/j.envres.2021.110831 (cit. on p. 152).

Touboul, J. and A. Destexhe (2010). "Can power-law scaling and neuronal avalanches arise from stochastic dynamics?" In: *PLoS ONE* 5.2. DOI: 10.1371/journal.pone.0008982 (cit. on p. 56).

— (2017). "Power-law statistics and universal scaling in the absence of criticality". In: *Physical Review E* 95, p. 012413. DOI: 10.1103/PhysRevE.95.012413 (cit. on p. 56).

Trefois, C. (2014). "Detection and characterization of critical transitions in mitochondrial activity via high content screening". PhD thesis. University of Luxembourg (cit. on pp. 70, 93, 98, 107, 108).

Trefois, C. et al. (2015). "Critical transitions in chronic disease: Transferring concepts from ecology to systems medicine". In: *Curr. Opin. Biotech.* 34, pp. 48–55. DOI: 10.1016/j.copbio.2014.11.020 (cit. on pp. 18, 88, 117, 121).

Truong, C., L. Oudre, and N. Vayatis (2020). "Selective review of offline change point detection methods". In: *Signal Processing* 167, p. 107299 (cit. on p. 73).

Tsimring, L. S. (2014). "Noise in biology". In: *Reports on Progress in Physics* 77.2.2, p. 026601. DOI: 10.1088/0034-4885/77/2/026601.Noise (cit. on pp. 20, 29).

Tu, C., P. D'Odorico, and S. Suweis (2021). "Dimensionality reduction of complex dynamical systems". In: *iScience* 24.1, p. 101912. DOI: 10.1016/j.isci.2020.101912 (cit. on p. 85).

Tyson, J. J., K. C. Chen, and B. Novak (2003). "Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell". In: *Current opinion in cell biology* 15.2, pp. 221–231 (cit. on p. 61).

Usman, S. M., M. Usman, and S. Fong (2017). "Epileptic seizures prediction using machine learning methods". In: *Computational and mathematical methods in medicine* 2017. DOI: 10.1155/2017/9074759 (cit. on p. 178).

Vallejo, J. A. et al. (2021). "Modeling the number of people infected with {SARS-COV-2} from wastewater viral load in Northwest {S}pain". In: *Science of The Total Environment*, p. 152334. DOI: 10.1016/j.scitotenv.2021.152334 (cit. on pp. 152, 157, 161, 165).

Van De Leemput, I. A. et al. (2014). "Critical slowing down as early warning for the onset and termination of depression". In: *Proceedings of the National Academy of Sciences of the United States of America* 111.1, pp. 87–92. DOI: 10.1073/pnas.1312114110 (cit. on pp. 17, 35, 45).

van den Driessche P. (2017). "Reproduction numbers of infectious disease models". In: *Infectious Disease Modelling* 29;2.3, pp. 288–303. DOI: 10.1016/j.idm.2017.06.002 (cit. on p. 65).

Van Kampen, N. G. (1992). *Stochastic processes in physics and chemistry*. Vol. 1. Elsevier (cit. on pp. 90, 134, 211).

Van Nes, E. H. and M. Scheffer (2004). "Large species shifts triggered by small forces". In: *American Naturalist* 164.2, pp. 255–266. DOI: 10.1086/422204 (cit. on p. 17).

Veening, J.-W., W. K. Smits, and O. P. Kuipers (2008). "Bistability, epigenetics, and bet-hedging in bacteria". In: *Annu. Rev. Microbiol.* 62, pp. 193–210. DOI: 10.1146/annurev.micro.62.081307.163002 (cit. on p. 61).

Velazquez, J. L. P. et al. (2011). "Experimental observation of increased fluctuations in an order parameter before epochs of extended brain synchronization". In: *Journal of Biological Physics* 37.1, pp. 141–152. DOI: 10.1007/s10867-010-9205-5 (cit. on p. 57).

Venkadesan, M., J. Guckenheimer, and F. F. J. Valero-Cuevas (2007). "Manipulating the edge of instability". In: *Journal of biomechanics* 40.8, pp. 1653–1661. DOI: 10.1038/jid.2014.371 (cit. on p. 202).

Veraart, A. A. J. et al. (2012). "Recovery rates reflect distance to a tipping point in a living system". In: *Nature* 481.7381, pp. 357–359. DOI: 10.1038/nature10723 (cit. on pp. 21, 45).

Waddington, C. H. (1957). *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. Allen & Unwin (cit. on p. 40).

Wald, A. (1945). "Sequential tests of statistical hypotheses". In: *The annals of mathematical statistics* 16.2, pp. 117–186. DOI: 10.1214/aoms/1177731118 (cit. on p. 73).

Wambui, G. D., G. A. Waititu, and A. Wanjoya (2015). "The power of the pruned exact linear time (PELT) test in multiple changepoint detection". In: *American Journal of Theoretical and Applied Statistics* 4.6, p. 581 (cit. on p. 73).

Wang, C. J. and K. L. Yang (2016). "Correlated noise-based switches and stochastic resonance in a bistable genetic regulation system". In: *European Physical Journal B* 89.8. DOI: 10.1140/epjb/e2016-70224-2 (cit. on p. 106).

Wang, C., J. Bi, and M. G. Olde Rikkert (2018). "Early warning signals for critical transitions in cardiopulmonary health, related to air pollution in an urban Chinese population". In: *Environment International* 121.163, pp. 240–249. DOI: 10.1016/j.envint.2018.09.007 (cit. on pp. 45, 80).

Wang, J., C. Li, and E. Wang (2010). "Potential and flux landscapes quantify the stability and robustness of budding yeast cell cycle network". In: *Proceedings of the National Academy of Sciences of the United States of America* 107.18, pp. 8195–8200. DOI: 10.1073/pnas.0910331107 (cit. on p. 71).

Wang, J. et al. (2011). "Quantifying the Waddington landscape and biological paths for development and differentiation". In: *Proceedings of the National Academy of Sciences* 108.20, pp. 8257–8262. DOI: 10.1073/pnas.1017017108 (cit. on p. 40).

Wang, P., J. Lu, and X. Yu (2015). "Colored noise induced bistable switch in the genetic toggle switch systems". In: *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 12.3, pp. 579–589. DOI: 10.1109/TCBB.2014.2368982 (cit. on p. 106).

Wang, R. et al. (2012a). "Flickering gives early warning signals of a critical transition to a eutrophic lake state". In: *Nature* 492.7429, pp. 419–422. DOI: 10.1038/nature11655 (cit. on pp. 17, 21, 44, 53, 70, 73–75).

Wang, X. et al. (2012b). "Construction of gene regulatory networks with colored noise". In: *Neural Computing and Applications* 21.8, pp. 1883–1891. DOI: 10.1007/s00521-011-0584-8 (cit. on p. 106).

Weber, M. and J. Buceta (2013). "Stochastic stabilization of phenotypic states: the genetic bistable switch as a case study". In: *PloS one* 8.9, e73487. DOI: 10.1371/journal.pone.0073487 (cit. on pp. 61–63, 112, 114).

Weidhaas, J. et al. (2021). "Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds". In: *Sci. Total Environ.* 775, p. 145790. DOI: https://doi.org/10.1016/j.scitotenv.2021.145790 (cit. on p. 152).

Weinans, E. et al. (2019). "Finding the direction of lowest resilience in multivariate complex systems". In: *Journal of the Royal Society Interface* 16.159, p. 20190629. DOI: 10.1098/rsif.2019.0629 (cit. on pp. 47, 52, 76, 83).

Weinans, E. et al. (2021). "Evaluating the performance of multivariate indicators of resilience loss". In: *Scientific Reports* 11.1, pp. 1–11. DOI: 10.1038/s41598-021-87839-y (cit. on pp. 47, 83).

Wernberg, T. et al. (2016). "Climate-driven regime shift of a temperate marine ecosystem". In: *Science* 353.6295, pp. 169–172. DOI: 10.1126/science.aad8745 (cit. on p. 17).

West, R. and T. Sohal (2006). ""Catastrophic" pathways to smoking cessation: findings from national survey". In: *Bmj* 332.7539, pp. 458–460. DOI: 10.1136/bmj.38723.573866.AE (cit. on p. 17).

Wichers, M. et al. (2016). "Critical slowing down as a personalized early warning signal for depression". In: *Psychotherapy and psychosomatics* 85.2, pp. 114–116 (cit. on p. 21).

Widder, A. and C. Kuehn (2016). "Heterogeneous population dynamics and scaling laws near epidemic outbreaks". In: *Mathematical Biosciences and Engineering* 13.5, pp. 1093–1118. DOI: 10.3934/mbe.2016032 (cit. on p. 65).

Wieczorek, S. et al. (2011). "Excitability in ramped systems: The compost-bomb instability". In: *Proceedings of the Royal Society A* 467.2129, pp. 1243–1269. DOI: 10.1098/rspa.2010.0485 (cit. on pp. 17, 18, 34).

Wiggins, S. et al. (1991). *Introduction to applied nonlinear dynamical systems and chaos*. Vol. 33. 1, p. 81. DOI: 10.1016/0378-4754(91)90060-g (cit. on p. 19).

Wilkat, T., T. Rings, and K. Lehnertz (2019). "No evidence for critical slowing down prior to human epileptic seizures". In: *Chaos* 29.9, pp. 2–7. DOI: 10.1063/1.5122759 (cit. on p. 57).

Williamson, M. S., S. Bathiany, and T. M Lenton (2016). "Early warning signals of tipping points in periodically forced systems". In: *Earth System Dynamics* 7.2, pp. 313–326. DOI: 10.5194/esd-7-313-2016 (cit. on pp. 21, 34).

Wilmes, P. et al. (2020). "Mass Screening for SARS-CoV-2 Uncovers Significant Transmission Risk from Asymptomatic Carriers". In: *SSRN Electronic Journal*, pp. 1–45. DOI: 10.2139/ssrn.3738086 (cit. on p. 130).

Wilmes, P. et al. (2021). "SARS-CoV-2 transmission risk from asymptomatic carriers: results from a mass screening programme in Luxembourg". In: *Lancet Reg. Heal.-Europe* 4, p. 100056. DOI: 10.1016/j.lanepe.2021.100056 (cit. on pp. 136, 160).

Winfree, A. T. and S. H. Strogatz (1984). "Organizing centres for three-dimensional chemical waves". In: *Nature* 311.5987, p. 611. DOI: 10.1038/311611a0 (cit. on p. 57).

Wittenberg, R. W. and P. Holmes (1999). "Scale and space localization in the Kuramoto-Sivashinsky equation". In: *Chaos* 9.2, pp. 452–465. DOI: 10.1063/1.166419 (cit. on p. 79).

Wölfel, R. et al. (2020). "Virological assessment of hospitalized patients with {COVID-2019}". In: *Nature* 581.7809, pp. 465–469. DOI: https://doi.org/10.1038/s41586-020-2196-x (cit. on p. 157).

Woodcock, A. E. R. and T. Poston (1976). "A higher catastrophe machine". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 79. 2. Cambridge University Press, pp. 343–350 (cit. on p. 215).

Wu, J. T., K. Leung, and G. M. Leung (2020). "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study". In: *The Lancet* 395.10225, pp. 689–697. DOI: https://doi.org/10.1016/S0140-6736(20)30260-9 (cit. on pp. 66, 132, 150, 168, 199).

Wurtzer, S. et al. (2020). "Evaluation of lockdown effect on SARS-CoV-2 dynamics through viral genome quantification in waste water, Greater Paris, France, 5 March to 23 April 2020". In: *Eurosurveillance* 25.50, p. 2000776. DOI: 10.2807/1560-7917.ES.2020.25.50.2000776 (cit. on p. 152).

Xie, Q. et al. (2018). "Predicting fluctuations-caused regime shifts in a time delayed dynamics of an invading species". In: *Physica A* 493.November, pp. 69–83. DOI: 10.1016/j.physa.2017.10.036 (cit. on p. 21).

Xu, S. and Y. Li (2020). "Beware of the second wave of COVID-19". In: *The Lancet* 395.10233, pp. 1321–1322 (cit. on p. 131).

Yan, J. et al. (2021). "Identifying Critical States of Complex Diseases by Single-Sample Jensen-Shannon Divergence". In: *Frontiers in oncology* 11, p. 1824. DOI: 10.3389/fonc.2021.684781 (cit. on pp. 45, 76).

Yan, P. and S. Liu (2006). "SEIR epidemic model with delay". In: *The ANZIAM Journal* 48.1, pp. 119–134. DOI: https://doi.org/10.1017/S144618110000345X (cit. on p. 64).

Yang, B. et al. (2018). "Dynamic network biomarker indicates pulmonary metastasis at the tipping point of hepatocellular carcinoma". In: *Nature Communications* 9.1, pp. 1–14. DOI: 10.1038/s41467-018-03024-2 (cit. on pp. 21, 88, 127).

Yang, H. (2010). "Multiscale recurrence quantification analysis of spatial cardiac vectorcardiogram signals". In: *IEEE Transactions on Biomedical Engineering* 58.2, pp. 339–347. DOI: 10.1109/TBME.2010.2063704 (cit. on p. 58).

Yasemi, M. and M. Jolicoeur (2021). "Modelling cell metabolism: a review on constraint-based steady-state and kinetic approaches". In: *Processes* 9.2, p. 322. DOI: 10.3390/pr9020322 (cit. on p. 89).

Zahler, R. S. and H. J. Sussmann (1977). "Claims and accomplishments of applied catastrophe theory". In: *Nature* 269.5631, pp. 759–763. DOI: 10.1038/269759a0 (cit. on pp. 38, 39, 222).

Zahri, M. (2014). "Multidimensional Milstein scheme for solving a stochastic model for prebiotic evolution". In: *Journal of Taibah University for Science* 8.2, pp. 186–198. DOI: https://doi.org/10.1016/j.jtusci.2013.12.002 (cit. on p. 67).

Zaikin, A and J Kurths (2001). "Additive noise in noise-induced nonequilibrium transitions". In: *Chaos* 11.3, pp. 570–580. DOI: 10.1063/1.1380369 (cit. on p. 205).

Zarei, F., S. Moghimi-Araghi, and F. Ghanbarnejad (2019). "Exact solution of generalized cooperative susceptible-infected-removed (SIR) dynamics". In: *Physical Review E* 100.1, p. 012307. DOI: 10.1103/PhysRevE.100.012307 (cit. on p. 65).

Zeeman, E. C. (1979). "Catastrophe theory". In: *Structural Stability in Physics*. Springer, pp. 12–22 (cit. on pp. 17, 215).

Zhang, H., Y. Chen, and Y. Chen (2012). "Noise Propagation in Gene Regulation Networks Involving Interlinked Positive and Negative Feedback Loops". In: *PLoS ONE* 7.12, pp. 1–8. DOI: 10.1371/journal.pone.0051840 (cit. on p. 106).

Zhang, W., L. M. Wahl, and P. Yu (2016). "Backward bifurcations, turning points and rich dynamics in simple disease models". In: *Journal of Mathematical Biology* 73.4, pp. 947–976. DOI: 10.1007/s00285-016-0976-6 (cit. on pp. 45, 65, 81).

Zhang, X., C. Kuehn, and S. Hallerberg (2015). "Predictability of critical transitions". In: *Physical Review E* 92.5, pp. 1–15. DOI: 10.1103/PhysRevE.92.052905 (cit. on pp. 21, 33, 81).

Zheng, X. D., X. Q. Yang, and Y. Tao (2011). "Bistability, probability transition rate and first-passage time in an autoactivating positive-feedback loop". In: *PLoS ONE* 6.3, pp. 9–11. DOI: 10.1371/journal.pone.0017104 (cit. on p. 114).

Zhong, Y. D. and N. E. Leonard (2019). "A Continuous Threshold Model of Cascade Dynamics". In: *2019 IEEE 58th Conference on Decision and Control (CDC)*. DOI: 10.1109/CDC40024.2019.9029844 (cit. on p. 179).

Zhou, J. X. et al. (2012). "Quasi-potential landscape in complex multi-stable systems". In: *Journal of the Royal Society Interface* 9.77, pp. 3539–3553. DOI: 10.1098/rsif.2012.0434 (cit. on pp. 71, 72, 205).

Zhou, J. X. and S. Huang (2013). *Theoretical Considerations for Reprogramming Multicellular Systems*. First Edit. Elsevier Inc., pp. 81–99. DOI: 10.1016/B978-0-12-394430-6.00005-4 (cit. on p. 40).

Zhu, Y. et al. (2021). "Early warning of {COVID-19} via wastewater-based epidemiology: potential and bottlenecks". In: *Sci. Total Environ.* P. 145124. DOI: https://doi.org/10.1016/j.scitotenv.2021.145124 (cit. on p. 152).

Zomorodian, A. and G. Carlsson (2005). "Computing persistent homology". In: *Discrete & Computational Geometry* 33.2, pp. 249–274. DOI: 10.1007/s00454-004-1146-y (cit. on p. 179).