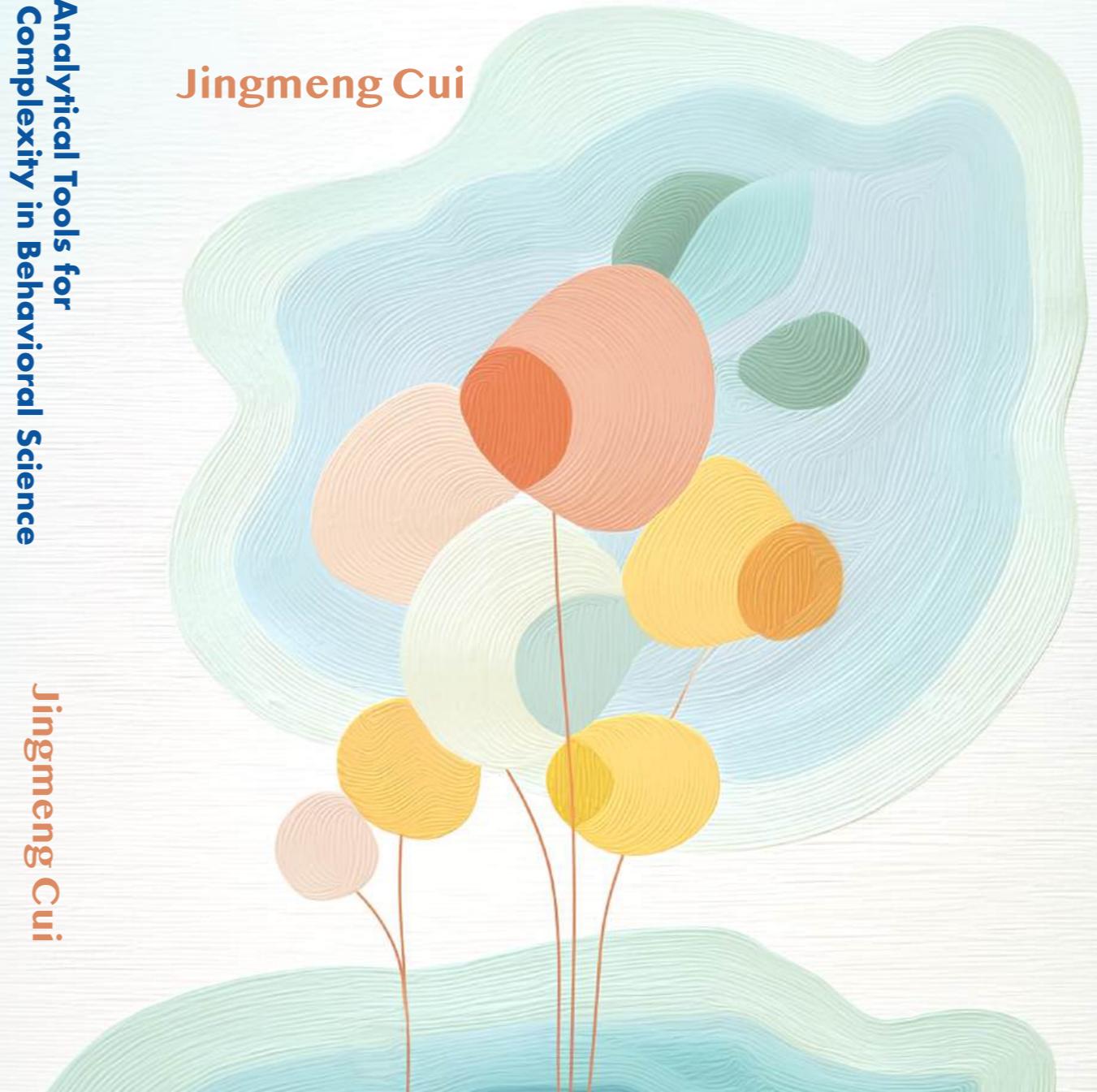


Advancing the analytical tools for understanding complexity in behavioral science

Jingmeng Cui

Advancing the Analytical Tools for
Understanding Complexity in Behavioral Science

Jingmeng Cui



Advancing the analytical tools for understanding complexity in behavioral science

Jingmeng Cui

Cover design: studio shu | shu.artdesign@gmail.com

Layout: Yuerui Network & Jingmeng Cui

Printing: Ridderprint | www.ridderprint.nl

Copyright © Jingmeng Cui, 2025

Copyright for the articles rests with the respective journals or the author.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without prior written permission from the author or the copyright-owning journal.



/ university of
groningen

Advancing the analytical tools for understanding complexity in behavioral science

PhD thesis

to obtain the degree of PhD of the
University of Groningen
on the authority of the
Rector Magnificus Prof. J.M.A. Scherpen
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Thursday 20 November 2025 at 12.45 hours

by

Jingmeng Cui

born on 12 February 1997

Supervisor

Prof. dr. A. Lichtwarck-Aschoff

Co-supervisor

Dr. F.W. Hasselman

Assessment Committee

Prof. dr. R.F.A. Cox

Prof. dr. D. Borsboom

Prof. dr. E. Ceulemans

Table of Contents

Chapter 1 General Introduction 7

PART 1. STABILITY LANDSCAPE METHODS

Chapter 2 From Metaphor to Computation: Constructing the Potential Landscape for Multivariate Psychological Formal Models 21

Chapter 3 simlandr: Simulation-Based Landscape Construction for Dynamical Systems 51

Chapter 4 Unlocking Nonlinear Dynamics and Multistability from Intensive Longitudinal Data: A Novel Method 75

Chapter 5 Quantifying the Stability Landscapes of Psychological Networks 97

PART 2. EARLY WARNING SIGNALS AND TRANSITIONS

Chapter 6 Examining the Research Methods of Early Warning Signals in Clinical Psychology through a Theoretical Lens 119

Chapter 7 Understanding Types of Transitions in Clinical Change: An Introduction from the Complex Dynamic Systems Perspective 147

PART 3. NONLINEAR DYNAMICS

Chapter 8 Examining the Feasibility of Nonlinear Vector Autoregressions for Psychological Intensive Longitudinal Data 169

Chapter 9 Analyzing Formal Dynamic Models in Psychology: A Tutorial Using Graphical Tools 195

Chapter 10 General Discussion 221

Appendix A. Supplementary Materials for Chapter 2 237

Appendix B. Supplementary Materials for Chapter 4 243

Appendix C. Supplementary Materials for Chapter 5 247

Appendix D. Supplementary Materials for Chapter 6 249

Appendix E. Supplementary Materials for Chapter 7 265

Appendix F. Supplementary Materials for Chapter 8 268

Appendix G. Supplementary Materials for Chapter 9 273

References 278

Dutch Summary (Nederlandse Samenvatting) 316

List of Publications 320

About the Author 322

Acknowledgments 323

Chapter 1

General Introduction



Prologue: How Did All These Things Start?

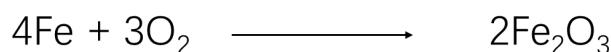
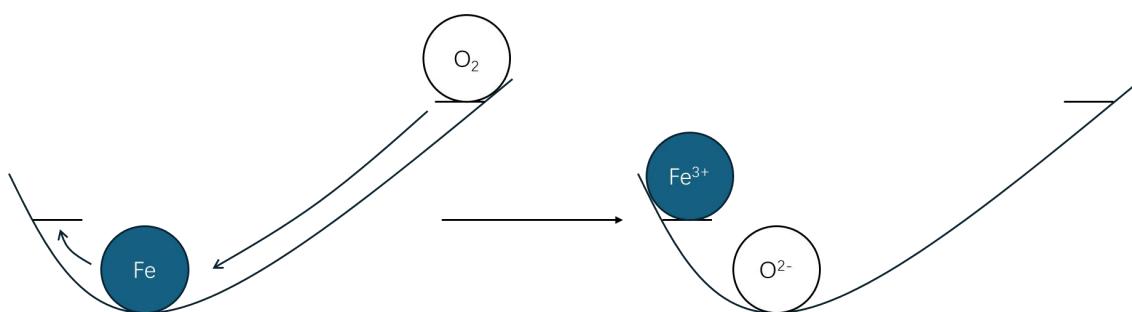
During the past few years, after I explained what my PhD is about, I often got this question: “What’s your background?” So, I feel maybe it is a good idea to start this thesis with some anecdotes.

My Story about Chemistry

In high school, I had a keen interest in chemistry. Each element has its unique properties, and together, they can have various reactions and make up all kinds of substances. I found this fascinating. After this initial excitement, I soon wanted to know the rules behind those phenomena. One of the first rules I learned was about reduction-oxidation reactions in which an oxidant receives electrons from a reducer. An everyday example is that when iron rusts, the iron atoms (Fe) give electrons to oxygen molecules (O_2). Consequently, iron atoms become ferric ions (Fe^{3+}), and oxygen molecules become oxide ions (O^{2-}). The most basic rule about this type of reaction is that a strong oxidant and a strong reducer will react, yielding a weak oxidant and a weak reducer. Therefore, the reaction has a specific direction, making the reversed reaction (iron rusts become iron and oxygen) impossible to happen automatically.

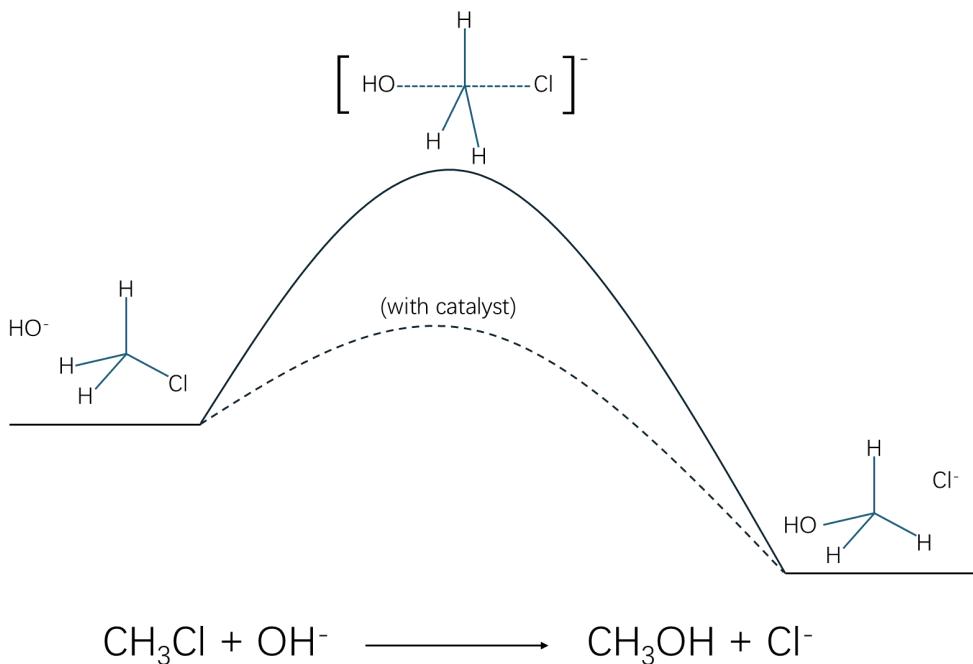
This rule made sense to me but still sounded abstract. To gain more intuition, I came up with the idea to represent such reactions with balls and landscapes (Figure 1). In this idea, the strength of the oxidants is represented by the height of the ball, and the reaction is represented as a collision of two balls. A ball from a higher position can fall down and knock another ball to a lower position, but not the other way around. Similarly, although the oxidation product (Fe^{3+}) now has a weak ability to act as an oxidant and receive electrons from others, it cannot get electrons from oxide ions (O^{2-}).

Figure 1. Potential landscape illustration of a simple reduction-oxidation reaction (iron rusts).



My chemistry teacher, Mr. Wang, quite liked this idea at that time. As my interest in chemistry continued to grow, I followed him in an extracurricular training for the Chinese Chemistry Olympiad (CChO) and learned more about chemical reactions and substances. In organic chemistry, I saw my old friends, landscapes, again, which were used to represent the detailed change process from reactants to products. During an organic reaction, the reactants first need to overcome an energy barrier (Figure 2). The lower the barrier, the faster the reaction is. Many organic reactions are reversible, meaning that the products may revert to the reactants. In the long run, the extent of the reaction depends on the relative stability of the reactants and the products. If the reactants have high energy (are less stable), the reaction tends to be complete, meaning that most of the reactants will become products. Once again, the idea of potential landscapes stood out to me as a powerful way to think about stability and transitions. *Many things in the world work like balls on a landscape. They have energy-like properties that represent the stability of their states. Understanding this can be quite valuable in making sense of how systems shift from one state to another.* Although I didn't realize it at the time, this idea would later become important in my journey in psychology.

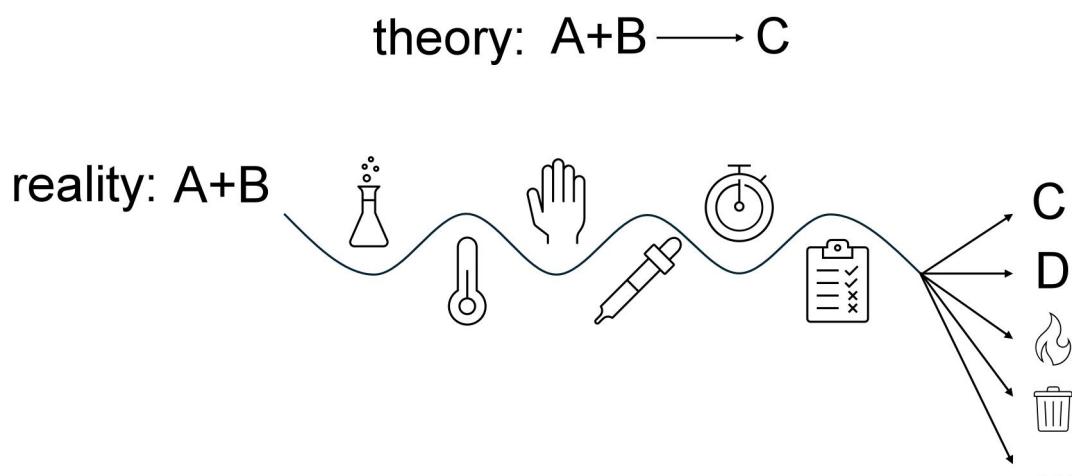
Figure 2. Illustration for the reaction potential energy surface of the hydrolysis of methyl chloride. Some catalysts can lower the barrier of the reaction, thus increasing the reaction speed.



The result of the CChO, by the way, was that I won a national gold medal and was admitted by my dream school, Peking University, for a bachelor's program in chemistry. In the

university, I gained more knowledge about how chemistry works in real life. Researchers in social sciences may think exact sciences¹ like chemistry have “cleaner” theories and phenomena. Well, that might be true in textbooks but is not actually the case in research labs. To produce a reaction is much more difficult than writing its formula down on paper. The flasks and tubes need to be arranged smartly to allow adding reagents halfway and to prevent leaking; the reaction temperature needs to be carefully controlled to make sure the reaction does not stop but also not set up fire (I did once set up a small fire); the process needs to be monitored in various ways to ensure the reaction completes properly, not too soon nor too late... From there, I learned an important lesson about the gap between theory and reality. *Scientific theories often appear clear and concise because they describe idealized situations. In practice, however, things are far more complex, with many factors playing a role. It is crucial to carefully assess whether the conditions underlying a theory still hold in real life.* This insight, as you will see, also applies to other fields.

Figure 3. Illustration of chemistry in theory and real-life labs.



Perhaps unsurprisingly, chemistry students don’t just take chemistry courses, but also mathematics, physics, and programming. Through those courses, I was introduced to the basic concepts of dynamic systems and their formal analysis methods and gained coding skills. One key difference between these fields and chemistry is their bottom-up approach. They often start with simple, idealized models that may not be entirely realistic but can still capture the essential aspects of complex phenomena. For example, a gas is composed of billions of molecules. It is infeasible to precisely describe the interactions of each pair of molecules. However, if we assume all the molecules only have elastic collisions, we can reach a very simple

¹ Or “beta subjects” as known by Dutch people.

description that is accurate enough for everyday applications. When needed, we can introduce adaptations to reach higher precision, but these adaptations are still based on the original simple form. This taught me that *while it is important to describe a complex system realistically, sometimes starting with simpler building blocks and gradually refining the model from there is also valuable*. This idea would also come back later for me in my PhD.

My Story about Psychology

Let me now return back to my high school days and tell you the other side of the story, about my interest in psychology. During high school, perhaps fueled by a mix of puberty and the extra adventure (or pressure) of the Olympiad competition, I also became deeply curious about the workings of the mind, both my own and that of others. Our competition program was quite intense, as we spent a large amount of time alongside the normal curriculum, characterized by uncertainty, planning, information searching, and frequent interaction and competition with other teammates. At the same time, like many teenage boys, I began to feel things I didn't quite have words for. Many questions came to my mind at the time. I wanted to know who I was and what I wanted to be, why I felt nervous sometimes or had strange feelings towards friends, and why people around me had different beliefs, motivations, and ways of acting. Psychology wasn't a subject taught in Chinese high schools. Nevertheless, in search of answers to those questions, I explored many introductory and folk psychology books during my free time.

After I started university, I enrolled in a double bachelor's program in psychology, thinking it would remain just a serious hobby. However, in my second year of chemistry, I realized that while I loved learning chemistry, I lacked the curiosity to push its boundaries through research. No topic truly sparked my desire to know beyond the textbooks. Two other things also happened at the same time. First, I found several friends and classmates of mine were struggling with mental issues. I tried to help them with the knowledge I had, but I often felt a deep sense of powerlessness, not only because I didn't know much about psychopathology at the time, but also because even the experts they turned to couldn't always offer much help, either. Second, in a general education course, we were required to read a book from a classic book list, and I chose *The Art of Loving* by Erich Fromm. This book was really inspiring to me, so I also continued reading several other books about the history of psychology, psychopathology, and psychotherapy. My interest in psychology grew so large that I started to feel that maybe I wanted to be a psychologist more than a chemist.

My journey took a new turn when I first visited the research group of Prof. Yanjie Su. Her work spans a wide range of topics in developmental psychology, and I immediately knew I

wanted to do research like that. I immersed myself in more psychology courses and started an undergraduate research project in Prof. Su's lab, investigating a question that stemmed from my own observations and curiosity. I loved the project and knew then that I wanted to pursue research in psychology. I decided to go abroad for a master's in psychology and received an offer from Radboud University, Nijmegen. Although I completed bachelor's degrees in both chemistry and psychology, I initially thought that my chemistry knowledge wouldn't be helpful to me any further. After all, it is not seen as useful in 99% of psychology labs.

I say 99%, because I happened to come across the field of complexity sciences in the last year of my bachelor's degree and found that there was a complex systems group in the Netherlands, the place I was already planning to go. Everything started to come together when I joined the Complexity in Behavioural Science (CiBS) group and met my master's and PhD supervisors, Prof. Anna Lichtenwarck-Aschoff and Dr. Fred Hasselman. There, many ideas from mathematics and natural sciences about how complex systems with many interacting elements work were applied to understanding psychological phenomena. As I immersed myself in this approach, I not only found it intuitive but also started generating quantitative ideas naturally. I saw ways to move beyond abstract and metaphorical concepts and actually calculate things that I felt were critical to the field. These ideas felt very concrete, thanks to the strong foundation I had built during my earlier studies in psychology, statistics, mathematics, physics, chemistry, and coding, essentially everything I needed. Moreover, many lessons I learned from my chemistry studies also echoed again, shaping some of the key ideas in my PhD project.

In the next section, I will introduce you to some specific topics that I have been working on.

Introduction to This Thesis

It is difficult to understand the human mind and why some people suffer from mental disorders. A major reason is the inherent complexity of humans. This complexity lies on different levels. On the biological level, millions of neurons in the brain coordinate in a way that is still far from being understood, producing all kinds of fascinating phenomena. On the level of groups and society, individuals interact with others, forming diverse relationships and collective behaviors. Even on the individual level, which is the primary focus of psychology, thousands of psychological constructs and processes unfold over time (Borsboom et al., 2022; Elson et al., 2023; Hasselman, 2023b). All these layers of complexity contribute to the challenges in understanding the mind and mental disorders scientifically.

Nonetheless, social scientists are not alone in this adventure. Scientists in other fields, such as mathematics, physics, and biology, also face the challenge of understanding complex phenomena and have made significant efforts to address them. Many researchers across fields have recognized that properly handling complexity is the key to understanding many important systems. Complexity science, therefore, emerged as an intrinsically multidisciplinary field that offers many conceptual and analytical tools, to study complex systems. Conceptual tools are the fundamental ideas and frameworks within complexity science. Some examples include self-organization, emergence, critical transitions, and chaotic dynamics (Borsboom et al., 2022; Bringmann et al., 2023; Cramer et al., 2016; Hasselman, 2023; Hayes & Andrews, 2020; Olthof et al., 2023; Thelen & Smith, 1998; van der Maas, 2024; also see Table 1). These ideas are typically explored in philosophical or narrative terms, providing qualitative insights without requiring technical or mathematical formulation. Analytical tools, in contrast, involve more rigorous approaches, such as mathematical derivations, computer simulations, and statistical modeling. These tools aim to quantify, model, or test the phenomena in complex systems, offering a formal bridge between abstract concepts and concrete observations. Some examples that fall into this category include complexity measures, stability analysis of equilibrium points, and non-equilibrium thermodynamics (Di Vita, 2022; Lloyd, 2001; Sayama, 2015; also see Table 1).

Of course, there is no clear boundary between the two types of tools, as most conceptual frameworks are grounded in earlier analytical work, and analytical tools are often guided by or require conceptual ideas. Yet, for psychologists, especially those who work with *behavioral data* rather than *neurophysiological data*, it seems that the conceptual part of complexity science has been more readily applied than formal analytical tools (Wagenmakers et al., 2012; but also see Stephen & Van Orden, 2012). This comes from at least two reasons. First, the measurement issue in psychology has long been an important challenge (Fried & Flake, 2018; Hasselman, 2023a; Lakens, 2025). Psychological constructs, especially those related to mental disorders and well-being, are often assessed by self-report instruments. In a typical example, a participant is asked to report their happiness on a 1-7 Likert scale. This measurement procedure can, at most, distinguish seven levels across the full spectrum of happiness, and for many participants, the difference between choosing a 2 or a 3 may be negligible. This limited resolution imposes a ceiling on measurement precision. If it is impossible to measure mental states with sufficient precision, applying sophisticated analytical tools would be infeasible, leaving their adoption lagging behind. Second, the research culture and training of social scientists place greater emphasis on concepts, ideas, or heuristics than formal analytical

methods (Borsboom, 2006). As a result, neither the researcher nor their target phenomena are fully prepared to embrace the formal side of complexity science.

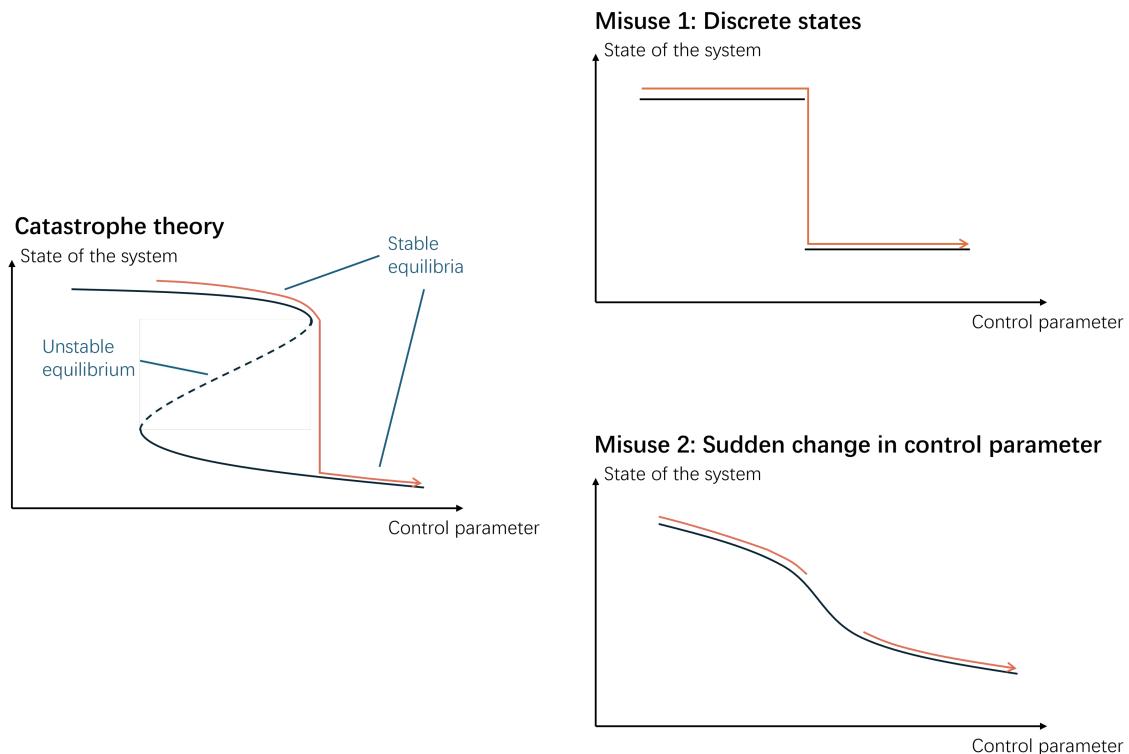
Table 1. Some key terms in complexity sciences and their definitions (Scott, 2005).

Classification	Terms	Definition
Conceptual tools		
	Self-organization	Spontaneous formation of spatiotemporal or functional structures in complex open systems.
	(Modest)	The whole has features that are different in kind from those of its parts (or alternatively that could be had by its parts).
	Emergence	
	Critical transition	A change from one system phase to another.
	Chaos	Systems that evolve in time according to a deterministic rule and demonstrate capricious and seemingly unpredictable behavior.
Analytical tools		
	(Algorithmic)	The length of the shortest algorithm, or the most parsimonious recipe that can generate the studied entity.
	Complexity	
	Stability of equilibrium in dynamic systems	An equilibrium point is stable if all local trajectories flow into it and is unstable if at least one local trajectory moves out of the neighborhood. Can be determined by calculating the Lyapunov exponent.
	Non-equilibrium statistical mechanics	A branch of statistical mechanics that aims to describe the behavior of large systems of particles that are moved away from the thermodynamic equilibrium, using the properties of each individual element and their interactions specified by physical laws. Such systems produce irreversible behaviors and form thermodynamic equilibrium states in isolated situations but may form a stationary nonequilibrium state under constraints.

The strong conceptual focus, sometimes at the expense of technical precision, can cause problems, as critical details may be lost during explanation and interpretation. A notable example is the hype and subsequent misuse of catastrophe theory. Catastrophe theory showed

an interesting mathematical phenomenon that the stable equilibria of dynamic systems may suddenly appear or disappear in response to continuous changes in control parameters (Thom, 1975; Zeeman, 1976). However, later studies often overlook the mathematical underpinnings and use the theory to explain various discontinuous phenomena, many of which stem from very different causes, such as the possible states of the system are inherently discrete, or that the sudden change in the state of the system comes from the sudden change of the control parameter (Sussmann & Zahler, 1978; Zahler & Sussmann, 1977; see Figure 4). Such misunderstandings do not only assign incorrect explanations to observations but also encourage misguided applications, reducing the effectiveness of the tools and the insights they could offer.

Figure 4. Illustration of the catastrophe theory and two misused cases.



Things are changing, though. In recent years, various methods have been proposed to bring formal analytical techniques into the field of psychology. Several notable examples include network psychometrics models (Borsboom et al., 2021), formal dynamic models (Haslbeck et al., 2022), drift-diffusion models (Loossens et al., 2020; Oravecz et al., 2009), early warning signal research (Helmich et al., 2024), and recurrence-based analysis tools (Hasselman & Bosman, 2020). These methods provide important tools with different emphases. Network psychometric models and formal dynamic models focus on the structure

and form of *dynamic interactions* among multiple variables, with the former being data-driven and the latter theory-driven. Their central goal is to investigate the complex dynamics of important elements in a system and figure out how one element influences another. Drift-diffusion models, early warning signal research, and recurrence-based analysis, on the other hand, focus more on the *general stability* of a system's macro states. They aim to identify the stable states of the system, describe how much the system fluctuates around those stable states, and predict when such change is likely to happen.

However, this progress is still not enough. There are still many concepts remaining vague, many conditions remaining implicit, and many mechanisms remaining unexplained. This is, of course, a broad issue that could be discussed endlessly. To make the discussion more concrete, I focus on three concrete questions in the subsequent sections, which form the core guiding my PhD work. First, what is the meaning of stability of states in a complex psychological system, and how, if possible, should we calculate it? Second, what are the possible mechanisms underlying change in psychological phenomena, and how, if possible, should we predict when change will occur? Third, how do nonlinear interactions give rise to complex behaviors in psychological systems, and how, if possible, should we model and analyze them?

Potential Landscape: Representing Stability in Psychological Systems

Although the concept of potential landscape originates from physics, the idea of using landscapes to represent stability is already built into our intuition. The famous visual cliff experiment by Gibson and Walk (1960) shows that infants as young as 6 months old already show scare in front of a cliff – they know that if they move forward, their current position will no longer be stable and they will have the tendency to fall down. Therefore, it is not surprising that many fields of science use potential landscapes to represent stability and the tendency to change. In the 1910s, René Marcelli, a physical chemist, already proposed the use of potential energy surface (PES) to represent the energy of the molecules as a function of the atom coordinates (Lewars, 2003). In biology, Wright (1932) used the fitness landscape to describe the relationship between genotypes and reproductive success, and Waddington (1966) used the epigenetic landscape to represent the process of cell development and differentiation. In ecology, similar analogies were introduced by Williams in 1970 to describe possible ecological states, their relative stability and resilience, and regime shifts (Lamothe et al., 2019).

Psychology has also adopted the landscape metaphor to represent the stability of different psychological states. Haken et al. (1985), for example, used potential landscapes to represent the bistability in finger movement patterns, and van der Maas et al. (2003) used landscapes to represent polarization in attitudes. Many researchers in psychopathology also

started to use the landscape metaphor recently (Hayes & Andrews, 2020; Olthof et al., 2023; van de Leemput et al., 2014; Wichers et al., 2019). In those applications, there are often two basins in the landscape, one representing a healthy phase, and the other representing an unhealthy phase (e.g., depression). The relative depths of the basins and the barrier between them are then used to represent the stability of the two phases and the difficulty for the system to transition from one phase to the other. Although the landscape metaphor conveys ideas of multistability and transitions, it cannot be directly linked to empirical observations or quantitative models. As a result, the relative stability of different psychological phases, as well as changes in the stability landscape, remain speculative, with no way to evaluate whether these guesses are accurate. Therefore, to go a step further, we need a way to quantitatively calculate the potential landscapes.

The calculations of potential landscape and dynamics can go in two main directions: calculating the potential energy directly, or calculating the potential landscape from dynamic equations. In some fields, the calculation of potential energy is straightforward. The PES in chemical reactions is one example of such. There, using some quantum mechanics methods, it is possible to calculate the energy of the reactants, the products, and the transition state (i.e., the saddle point of the potential energy surface, which is the state with the highest energy along the reaction process). The speed of the chemical reaction is then exponentially related to the energy difference between the transition state and the reactants. In this example, the potential energy is known and we use it to infer the dynamic properties of the system. For other systems, however, having a dynamic function is more straightforward. This is the case in biochemical systems with multiple species (e.g., Li & Wang, 2013; Li & Ye, 2019; Wang et al., 2008). For example, there are multiple genes in a cell with complex relationships, some genes activating others while some inhibiting others. Overall, they create several gene expression patterns, corresponding to different types of cells (e.g., stem cells, neurons, blood cells, and muscle cells). The relationships among those cells can be expressed in certain biochemical formulas so that we know how the expression level of genes changes over time. However, there is no well-defined “energy” for each possible state of gene expression. It is the dynamic interactions among the genes that make it *as if* the cell can fall into various basins on a landscape, representing different cell types. For this kind of system, we know the dynamic functions in the system and use those to calculate the potential landscape.

For psychological systems, both cases are possible, but mostly the dynamics come first, corresponding to the latter approach discussed above. A way to derive the dynamic equations of psychological systems is to use *formal dynamic models*. Those models translate verbal

psychological theories into mathematical equations. Once we have such a mathematical model, we can then adapt the methods from biochemical systems to calculate the potential landscape of psychological systems. I show a concrete example in **Chapter 2** using a formal model of panic disorder, and I discuss the technical details of this method in **Chapter 3**. The dynamics of the system can also be estimated from empirical data. In **Chapter 4**, I describe a method to estimate the nonlinear dynamics from intensive longitudinal data and calculate the potential landscape from there.

In some cases, however, the potential energy of a psychological system is available, making the direct approach possible. In **Chapter 5**, I describe a method to estimate the potential landscape for Ising network models. The Ising network is a kind of network model that allows bistability (van Borkulo et al., 2014). This model has an intrinsic potential energy measure. Therefore, instead of relying on the dynamics of the system, we try to summarize this intrinsic energy measure to provide more intuitive metrics.

Transitions and Early Warning Signals: Theory and Reality

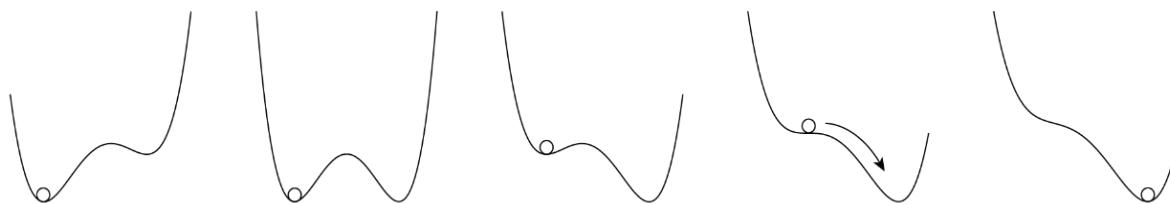
The past few years have seen ups and downs in the research of early warning signals (EWSs). While complex systems research in psychology is often rather abstract, the idea of EWSs appears to bear very practical promise. This promise builds on their ability to predict sudden changes in many types of systems without requiring detailed knowledge of the underlying mechanisms. Indeed, EWSs have been observed in a wide range of complex systems, including various examples in natural sciences (Scheffer et al., 2009). Therefore, many researchers also hope to use EWSs to predict sudden changes in psychopathology. Those sudden changes often have critical implications. For example, if we know a patient is about to have a sudden deterioration, we can apply timely intervention to prevent it.

The theoretical basis for classical EWSs² is often explained through ball-and-landscape illustrations, shown in Figure 5. The ball is in a deep basin at the beginning, but it gradually loses stability, represented by the basin becoming shallower. During this process, the ball becomes more prone to noise and starts to have larger fluctuations. If we record the system's state, we can observe its variance increasing. Moreover, it also takes longer for the ball to return to the local equilibrium, leading to higher autocorrelation. In the end, the basin becomes completely unstable, and the ball shifts to another basin, which represents a sudden transition in psychological states. Therefore, this transition can be forecasted by the increasing

² Some other statistical indicators may exist before transitions in very different kinds of systems than the original EWSs, yet sometimes are also called early warning signals (e.g., Evers et al., 2024). I see this more as a terminology issue instead of a substantial discrepancy in theories and mostly focus on the classical EWSs in this dissertation.

variance and autocorrelations in the record of the system beforehand. These statistical indicators, along with several others, are referred to as EWSs because they can help predict critical transitions (Dablander et al., 2023; Scheffer et al., 2009, 2012).

Figure 5. Illustration of a bifurcation-induced transition.



EWSs are universal to some extent because they do not require a specific substantial nature of the system — it can be physical, biological, financial, or psychological, and they do not require a specific type of interaction among system elements. However, they require that the system undergoes a specific *type* of transition similar to what is represented in Figure 5 and that researchers calculate EWSs correctly. If these conditions are not met, the observed statistical indicators may not reflect the reduced stability of the system and fail to predict a transition. For example, if a system shifts to an alternative basin because of a random fluctuation rather than a change in stability, no EWSs will appear; if the calculation of EWSs includes not only the period before the transition but also during or after the transition, the resulting statistical indicator may be confounded. Therefore, careful examination of the research methods for EWSs, as well as a better understanding of possible transition types in dynamic systems and their relationships to clinical phenomena, is required.

In **Chapter 6**, I will elaborate on the conditions under which EWSs are valid and discuss how research methods of EWS studies can be better aligned with the underlying theoretical assumptions. After that, in **Chapter 7**, I will introduce possible types of transitions in complex dynamic systems and explore how these transition types can be conceptualized within the context of psychopathology.

Complex Dynamic Interactions: Building from Basics

It is important not to miss the forest for the trees, but it is also unwise to ignore the trees once we have seen the forest. In previous sections, I introduced some tools to describe the broader picture of complex psychological systems, mapping their global stability, attractors, and transitions. However, after mapping out these overarching dynamics, it is now the time to dive deeper into the finer details, namely how the elements in a complex system interact with each

other. The specific forms of these interactions play an important role in shaping the global features of the system. For example, panic disorder is characterized by repeated panic attacks, during which patients feel an extreme level of fear. Those panic attacks may come from the vicious circle between physical arousal and perceived threat: if a patient's level of physical arousal (e.g., heart rate and breathing rate) increases, the person may take it as an indicator of danger, which further increases the physical arousal level (Clark, 1986). Yet, this can only lead to panic attacks when the relationship between physical arousal and perceived threat is nonlinear; otherwise, the system will either always stay in the healthy phase or always be in a panic phase, which is unrealistic (Robinaugh et al., 2021).

Most of the traditional statistical methods focus on linear interactions, which are easier to estimate, suffer less from overfitting issues, and require fewer data points. However, they are often not designed for estimating nonlinear dynamic interactions. Methods specialized in recovering nonlinear relationships from data exist in other fields (e.g., Brunton et al., 2016), but they are often data-hungry, requiring time series of lengths that are not feasible for self-report data collection. In **Chapter 8**, I describe a new method that extends the traditional linear vector autoregressive model by incorporating nonlinear terms. We show both its advantages and limitations, and we recommend it mainly be applied for exploratory purposes. Nevertheless, such an approach provides valuable insight into how far we can go in estimating nonlinear dynamic relationships among psychological variables using the concurrent data collection approach.

The data-driven approach is not the only possibility to investigate nonlinear interactions. Nonlinear dynamic models can also start from theory, using mathematical modeling and simulations to examine the relationship between the form of nonlinear interactions and the global behavior of the system. The development of formal dynamic models is a relatively new direction in psychology. Previous researchers mainly focused on simulations, searching for specific dynamic features from simulated time series. Although this approach is intuitive, it relies on specific simulation instances and parameter settings. It is unclear whether the observed dynamic features only emerge under these particular conditions, and whether changing the models will change the system's overall behavior as well. In **Chapter 9**, I introduce some graphical techniques from dynamic systems analysis and biological modeling. These techniques can help researchers better understand how the form of dynamic equations determines the system's global behavior.

Okay, that's it for now! I hope you enjoy the empirical chapters, and this narrative voice of mine will meet you again in **Chapter 10**, the general discussion :)

Chapter 2

From Metaphor to Computation: Constructing the Potential Landscape for Multivariate Psychological Formal Models

This chapter is based on:

Cui, J., Lichtwarck-Aschoff, A., Olthof, M., Li, T., & Hasselman, F. (2023). From metaphor to computation: Constructing the potential landscape for multivariate psychological formal models. *Multivariate Behavioral Research*, 58(4), 743–761.
<https://doi.org/10.1080/00273171.2022.2119927>



Abstract

For psychological formal models, the stability of different phases is an important property for understanding individual differences and change processes. Many researchers use landscapes as a metaphor to illustrate the concept of stability, but so far there is no method to quantify the stability of a system's phases. We here propose a method to construct the potential landscape for multivariate psychological models. This method is based on the generalized potential function defined by Wang et al. (2008) and Monte Carlo simulation. Based on potential landscapes we define three different types of stability for psychological phases: absolute stability, relative stability, and geometric stability. The panic disorder model by Robinaugh et al. (2024) is used as an example, to demonstrate how the method can be used to quantify the stability of states and phases, illustrate the influence of model parameters, and guide model modifications. An R package, simlandr, was developed to provide an implementation of the method.

Introduction

The past decades have seen a rapid growth of models and theories in the field of psychology, and more specifically in the field of psychopathology. These models, however, are not without critiques. Many verbal theories acknowledge the complex and dynamic nature of mental disorders, but they are not always able to make precise and falsifiable predictions (Borsboom et al., 2021; Robinaugh et al., 2021). Statistical models can provide quantitative estimations, but traditional and even cutting-edge statistical methods are largely based on linear, static, and homogeneous assumptions, and usually fail to draw correct conclusions about the nature of the underlying process (Granic & Hollenstein, 2003; Haslbeck et al., 2019; Olthof, Hasselman, & Lichtwarck-Aschoff, 2020). In recent years, formal models are gaining momentum in the field of psychology, aimed to address the abovementioned problems and provide a quantitative foundation for theoretical inferences (e.g., Burger et al., 2020; Cramer et al., 2016; Robinaugh et al., 2019; Schiepek et al., 2014).¹

In those formal models, the elements of psychological systems and their interactions are described mathematically. Based on these specifications, one can simulate how the model evolves over time, observe the characteristics of the model, and investigate how the model output corresponds to real-life phenomena. In complex systems, higher-order, macroscopic psychological phases² – distinct patterns of psychological systems – can emerge from the self-organization of these microscopic elements and their interactions (Goldstein, 1999; Olthof et al., 2023). Different psychological phases can sometimes be assigned to differences in mental health: a system can for instance be in an anxious, panicky phase or in a calm and relaxed phase (Robinaugh et al., 2024).

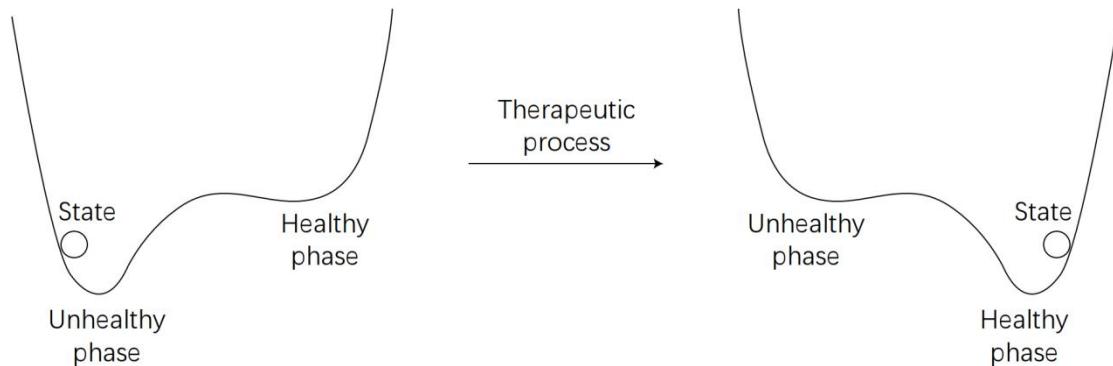
For formal models, the stability of different states and phases is an important quantity that relates to individual differences and change processes of the system. Previous research has often used the landscape metaphor to illustrate this idea: the state of the system is like a ball on the landscape. If the ball is in “a deep valley”, the system is stable; if the ball is on “a hill” or in “a shallow valley”, the system is unstable, and it tends to “fall down” to a more stable place. The “valleys” or “attractors” correspond to possible phases of the system (Lamothe et al., 2019; Olthof, Hasselman, Oude Maatman, et al., 2020; Wichers et al., 2019; see Figure 1 for a diagram). If the valley of an unhealthy phase of some individuals is deeper, their mental

¹ In this paper, we use the term “system” for real-life or modelled systems that contains interactive elements; “theory” for a set of ideas that explain how the system work, and “model” for the tool that researchers use to give a simplified description of the system mechanism.

² The term “phase” is sometimes used interchangeably with “state”, as in the “mania state/phase” of bipolar disorder and “liquid state/phase” of matter. To avoid confusion, we use the term “state” in this paper for the more specific conditions of the system which are defined by the values of the system variables, and “phase” for the higher-level patterns of the system that consist of a group of states.

systems are more likely to be trapped there, and they will be more vulnerable to mental disorders. Also, the process of treating mental disorders can be seen as changing the landscape of the system in a way that the stability of the unhealthy phase is decreased while at the same time strengthening the stability of healthy phases (Hayes et al., 2015).

Figure 1. Diagram of concepts in the landscape metaphor of psychopathology.



While the metaphorical use of the landscape is certainly a good way of illustration, the concept of stability does not yet have a formal, quantitative representation, which hinders further investigation. This calls for a new line of methods: quantitatively computing the potential landscape from formal dynamic models.

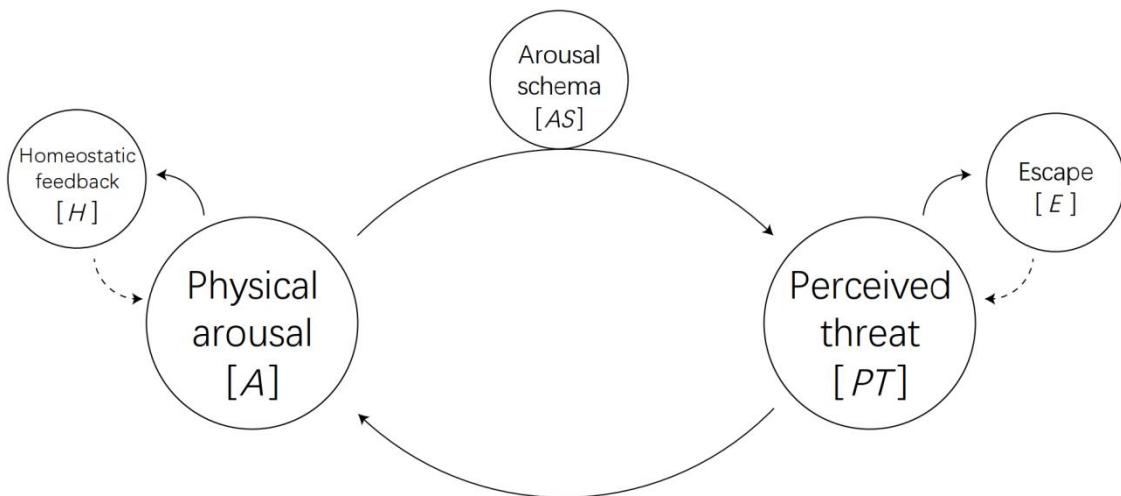
Formal Models and Case Simulations

As its name suggests, a psychological formal model of psychological phenomena mathematically defines how variables evolve over time and how they interact with each other. Often the evolution of such systems can be described by a set of (stochastic) differential equations. These equations specify the forms and strengths of these interactions among variables and the magnitude of noise in mathematical form. With this precise description, one can determine how the system evolves over time from a starting point and gain knowledge about the theory-implied behavior of the system (Robinaugh et al., 2021).

A well-known psychological formal model is the panic disorder model by Robinaugh et al. (2019; also see Borsboom et al., 2021, and Haslbeck et al., 2019, for discussions on this model). This model is well constructed and contains many typical features that are common in psychological models (e.g., nonlinear relationships, feedback loops, the dependence of system behavior on its unique history, adaption to the environment, and a rather large number of variables). Therefore, we use this model as an example to explain our ideas. Here we briefly introduce the model specification and the main variables and parameters. The relationships of

the most important variables that we use in the current paper are shown in Figure 2, and the full description of the model can be found in Robinaugh et al. (2024). The variables and parameters that will be investigated in the current paper are marked in bold font on their first occurrence.

Figure 2. Simplified causal diagram of the panic disorder model (adapted from Robinaugh et al., 2019). Each circle represents a variable of the model. The solid lines represent positive influences on the changing rate of the target variable, and the dashed lines represent negative influences on the changing rate of the target variable. The circle on the path from arousal to perceived threat represents the moderating effect of arousal schema on this relationship.



In this model, a panic disorder is considered to emerge from mutually interacting system variables. The core variables of this system are physical arousal (**A**, the level of arousal-related sensations, e.g., heart rate) and perceived threat (**PT**, the cognitive perception that the situation is threatening). The changing rate of physical arousal (**A**), dA/dt , is influenced by its own value, perceived threat (**PT**), and homeostatic feedback (**H**, the strength of the homeostatic processes that counteract the unsustainably elevated physical arousal),

$$\frac{dA}{dt} = r_A(s_{PT,A}PT - A - s_{H,A}H), \quad (1)$$

which represents that physical arousal (**A**) tends to decrease when itself and homeostatic feedback (**H**) is high and tends to increase when perceived threat (**PT**) is high. The parameters r_A , $s_{PT,A}$, and $s_{H,A}$ represent the strength of these influences. The changing rate of homeostatic feedback (**H**), dH/dt , is influenced by its own value and physical arousal (**A**),

$$\frac{dH}{dt} = r_H \left(\frac{A^{p_{A,H}}}{A^{p_{A,H}} + h_{A,H}^{p_{A,H}}} - H \right), \quad (2)$$

which represents that homeostatic feedback (H) tends to decrease when itself is high and tends to increase when physical arousal (A) is high. The parameters r_H , $p_{A,H}$, and $h_{A,H}$ represent the strength of these influences. The changing rate of perceived threat (PT), dPT/dt , is influenced by its own value, physical arousal (A), and escape behavior (E),

$$\frac{dPT}{dt} = r_{PT} \left(\frac{A^{p_{A,PT}}}{A^{p_{A,PT}} + h_{A,PT}^{p_{A,PT}}} - PT - s_{E,PT}E \right), \quad (3)$$

which represents that perceived threat (PT) tends to decrease when itself and escape behavior (E) is high and tends to increase when physical arousal (A) is high. r_{PT} , $p_{A,PT}$, $h_{A,PT}$, $s_{E,PT}$ are parameters representing the strength of these influences. Here, the influence of A on the changing rate of PT is not linear, but in the form of an S-shaped sigmoid function. Robinaugh et al. (2021) showed that this sigmoid function is necessary for the formal model in order to generate theory-implied behavior as observed in real life. The changing rate of escape behavior (E), dE/dt , is influenced by its own value and perceived threat,

$$\frac{dE}{dt} = r_E \left(\frac{PT^{p_{PT,E}}}{PT^{p_{PT,E}} + h_{PT,E}^{p_{PT,E}}} - E \right), \quad (4)$$

which represents that escape behavior (E) tends to decrease when it is high itself and tends to increase when perceived threat (PT) is high.

People with a panic disorder over-interpret their physical arousal as an indication of danger. When this over-interpretation, termed as *arousal schema (AS)* in the model, is high, an increase in physical arousal can result in a larger increase in perceived threat. This effect of AS is represented as its influence on the parameter $h_{A,PT}$,

$$h_{A,PT} = 1 - \frac{AS}{AS + h_{AS,APT}} - s_{AS,APT}c. \quad (5)$$

When AS is higher, $h_{A,PT}$ is lower, which makes the influence of A more dominant in Equation 3. Therefore, it is easier for a relatively high level of physical arousal (A) to lead to a large increase in perceived threat (PT). The increase of perceived threat (PT) can, in turn, amplify physical arousal (A). This generates a vicious circle between the two and finally lead to a panic attack. A collective variable **fear** is defined as the geometric mean of A and PT ,

$$fear = \sqrt{A \times PT}, \quad (6)$$

and is used to represent the general symptom severity. Panic attacks, therefore, manifest as a sudden increase in the level of **fear**. If panic attacks happen relatively often, the person can be said to have a panic disorder. Arousal schema (AS) is also influenced by a learning mechanism.

Its changing rate, dAS/dt , depends on its own value and the previous history of perceived threat (PT), *fear*, and escape behavior (E),

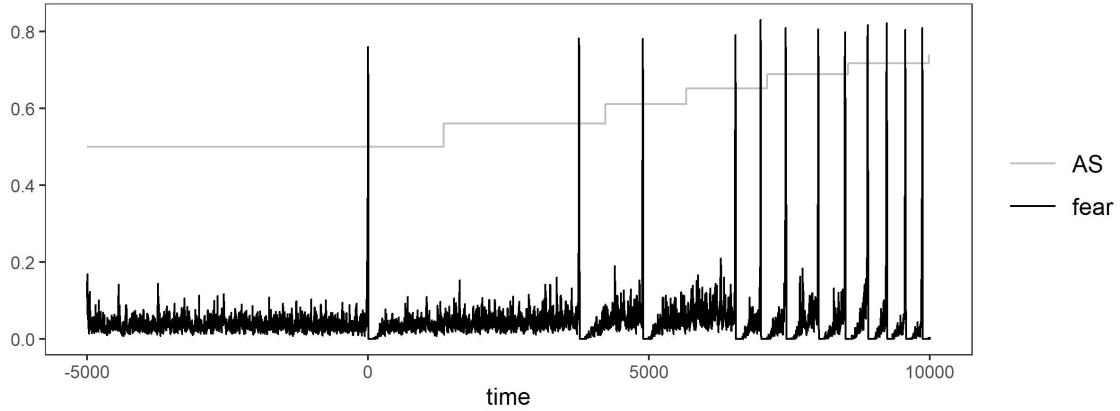
$$\frac{dAS}{dt} = \begin{cases} 0, & \text{if } \max(fear_{t-\Omega}, \dots, fear_t) < cr_{fear,AS} \\ r_{AS,a}(\max(PT_{t-\Omega}, \dots, PT_t) - AS), & \text{if } \max(fear_{t-\Omega}, \dots, fear_t) \geq cr_{fear,AS} \text{ and } \max(E_{t-\Omega}, \dots, E_t) > cr_{E,AS} \\ -r_{AS,e}AS, & \text{if } \max(fear_{t-\Omega}, \dots, fear_t) \geq cr_{fear,AS} \text{ and } \max(E_{t-\Omega}, \dots, E_t) \leq cr_{E,AS} \end{cases}, \quad (7)$$

which represents three different learning conditions. When the maximum value of *fear* in the previous Ω time points is lower than a critical threshold $cr_{fear,AS}$, no learning processes happened; when the maximum value of *fear* in the previous Ω time points is higher than the critical threshold $cr_{fear,AS}$, the direction of the learning process depends on whether the individual's escape behavior (E) in the previous Ω time points is higher than another threshold $cr_{E,AS}$. If the previous escape behavior (E) is high, the individual does not know how threatening the actual situation is, so he or she will learn to update the arousal schema (AS) according to the highest perceived threat (PT) during this time period with an acquiring rate parameter $r_{AS,a}$; if the previous escape behavior (E) is low, the individual will find that the actual situation is not so threatening, so his or her arousal schema (AS) will decrease with an extinguishing rate parameter $r_{AS,e}$.

The model is intended to produce two qualitatively different phases of the (patient) system: a healthy phase and a panic phase. The simulation results of this model (with the parameter values in Robinaugh et al., 2019; Figure 3) indeed show clear sudden increases in *fear*, which represent panic attacks. In these panic attacks, the system moves from the healthy phase to the panic phase, and quickly transitions back to the healthy phase. When the arousal schema (AS) becomes higher, the panic attacks become more frequent, which qualitatively shows the stability of the panic phase increasing, and the stability of the healthy phase diminishing. These simulation results align well with the theoretical foundation of the model.

In this approach, the performance of the model is evaluated by case simulation result: the output of directly simulating the dynamic model. Case simulation is an important way of model evaluation and deduction because it shows how the system evolves if the model correctly represents the system. A mismatch of the simulation result and real-life observations indicates something must be wrong in the model. Case simulation, however, cannot directly provide information about the stability of states: what it shows is how the state changes over time (i.e., state as a function of time), not the stability of different states (i.e., stability as a function of states). Therefore, we need to find a way to define and calculate the stability of the states in a psychological system.

Figure 3. The simulation results using the panic disorder model by Robinaugh et al. (2024). The first panic attack of the system appears probabilistically. To make simulation results comparable, *time* was set as zero at the first panic attack (the time when the peak value of *fear* is reached).



Potential Landscape for a Dynamical System

Defining or representing stability is not a totally new subject. In physics, the quantity that is used to represent the stability of a state is the *potential function*. Take the gravitational potential energy as an example: for a given object, if it is at a higher position, its gravitational potential energy is higher. This means that the object is more unstable, and it is more likely to “fall down” to more stable states in lower places. Mathematically, if a potential function can be (strictly) defined for a system, its velocity (i.e., how the system state changes over time) should be proportional to the gradient of the potential function. Intuitively, this means that the system always tends to move to the place with lower potential energy, just like a ball on a hill tends to fall along the most convenient way into the valley.

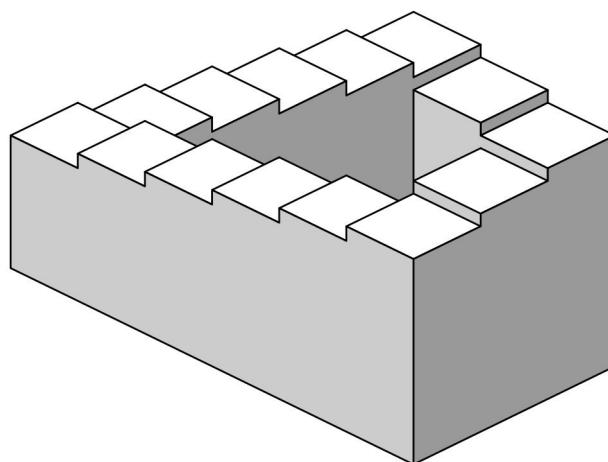
For a unidimensional deterministic system, the potential function can be easily obtained from taking the integral of the dynamic function. This method has already been used for representing the stability of states in unidimensional psychological systems (e.g., Dablander et al., 2020; Robinaugh et al., 2019). However, most psychological models contain many variables, thus are multidimensional. Analyzing the stability for multidimensional systems is more challenging because the multivariate dynamic functions are often not integrable. One way to understand this issue is by looking at the Penrose impossible stairs (Rodríguez-Sánchez et al., 2020; see Figure 4): it is possible that a system keeps whirling around, but then it cannot be represented as always going downstairs in a real 3D space. Therefore, the potential landscape cannot be directly obtained. Mathematicians have developed several generalized methods to construct generalized potential functions (P. Zhou & Li, 2016). These methods relax the requirements of integrability, but the resulting functions can also be used to represent the

stability of system states. Among different ways of generalization, we found the potential function by Wang et al. (2008; also see Li & Wang, 2013, and Li & Ye, 2019, for examples of usage in complex biochemical systems) the most suitable for psychological formal models because it can be estimated with the Monte Carlo method, thus does not have a strict requirement concerning the properties of the dynamic functions. Other generalizations often require the dynamic functions to be continuous, derivable, and independent of the history, properties that psychological models often do not meet (e.g., Equation 7 of the panic disorder model is dependent on the model history and thus not derivable). Wang's definition of potential function is based on the steady-state distribution of the system (denoted as P_{ss}), which refers to the distribution of states that holds constant over time. If we have a single system evolving over time according to a set of (stochastic) differential equations, its state is likely to change every now and then. However, if we have an infinite bunch of systems with the steady-state distribution and let them all evolve together, although each system's state still changes, their distribution can be invariant.³ Following a generalization of Boltzmann distribution, the potential (U) of a state \mathbf{X} is then given by

$$U(\mathbf{X}) = -\ln P_{ss}(\mathbf{X}) \quad (8)$$

which means that the potential is equal to the logarithm of the steady-state distribution. If the probability density for the steady-state distribution is lower, then its potential is higher.

Figure 4. The Penrose impossible stairs (Sakurambo, 2005).



³ A concrete illustration: imagine an infinite number of copies of an athlete running on a 400-meter track, and imagine they are independent (so they do not run into each other). Even if they are copies of the same athletes, due to random noises, they will gradually distance themselves from each other. After an infinitely long time, these people will be very evenly distributed in the track. Now even if each person is still running, the population distribution on the track will not change anymore. If each running person is a stochastic dynamic system, then this population distribution is the steady-state distribution for the system.

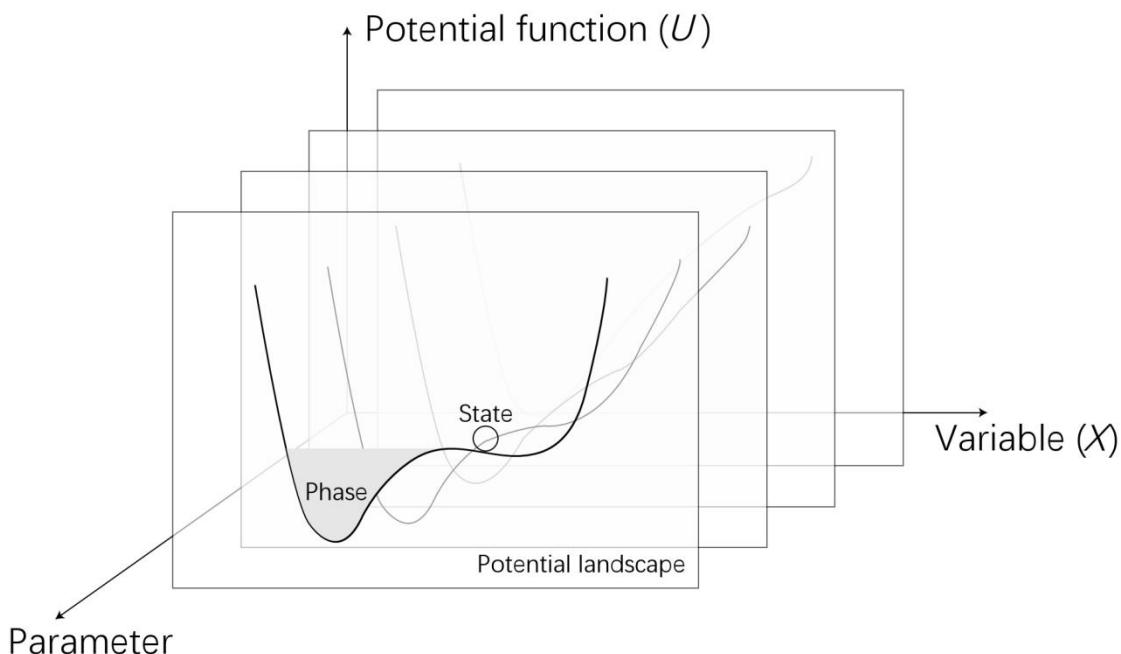
This potential function is also related to some other properties expected from strictly defined potential functions. Here we also explain the usefulness of generalized potential functions from the force decomposition perspective (P. Zhou & Li, 2016). As mentioned above, the main issue in constructing potential landscapes for multidimensional systems is that they often show whirling behaviors that cannot be represented with a potential landscape. These whirling behaviors are not the whole picture because the total forces in the system also contain gradient parts that represent the general tendency for the system to move to some specific regions. The idea of force decomposition means to decompose those complex forces of a system into a curl part and a gradient part. The stability information of the system is mainly contained in the gradient part, and the gradient part is integrable. Therefore, the generalized potential landscape can be obtained by integrating the gradient part of the forces. The curl part does not contain direct information about the stability because if the system just has the tendency to oscillate between two states without a preference, it is not meaningful to say that one state is more stable than the other. Wang's landscape is originally defined from the steady-state distribution, but it can also be proved that it is equivalent to a possible way of force decomposition (P. Zhou & Li, 2016). Therefore, the potential landscape constructed with Wang's method can be seen as a representation of the gradient part of the system dynamics.

From either perspective, the potential landscape of the system shows the tendency that the system resides or leaves a specific state. If a state has a lower density in the steady-state distribution, the system is less likely to be around this state in the long run; and if there are a bunch of systems starting with a uniform distribution in the state space, it is more likely that the systems starting around this state will move to other states. If a system is in a higher position on the gradient part of its dynamics, it means that the system tends to fall down to a lower position if not affected by the curl forces and random noises. Therefore, the potential landscape can efficiently represent the stability of psychological systems on the state level.

Having the potential for specific states, we can now describe the stability of the phases. How to clearly define psychological phases is a complex issue on its own. One may propose that a collection of qualitatively similar states constitutes a phase (e.g., the mental states when a patient with depression has a high overall symptom severity), in which case the phase may be seen as a *point attractor*. In contrast, it can also be the case that certain kind of trajectories constitutes a phase (e.g., a mental trajectory in which a patient with cyclothymic disorder switches between states with high depressive symptoms and states with elevated moods), in which case the phase may be seen as an *oscillating attractor* (Barton, 1994). Sometimes those two definitions can also be interchangeable. For example, if we use the mood variability over a

period instead of the valence and intensity of mood as the key variable to describe the cyclothymic disorder, then the patient's mood variability is always high, and may be described as a point attractor. Fully investigating this issue is beyond the scope of the current paper. For clarity and simplicity, we will only look into the phases consisting of similar states and bounded by barriers in the potential landscape (i.e., point attractor-like phases). In other words, we use the term *state* for a *point* in the state space (e.g., the state with $A = 0.1$ and $PT = 0.1$) and the term *phase* for a larger *region* that contains many states within it (e.g., the healthy phase). This distinction, although not commonly made in psychological literature, is important for the method we are introducing. The relationships among the potential function, parameters, variables, state, and phase we used in the current paper are shown in Figure 5.

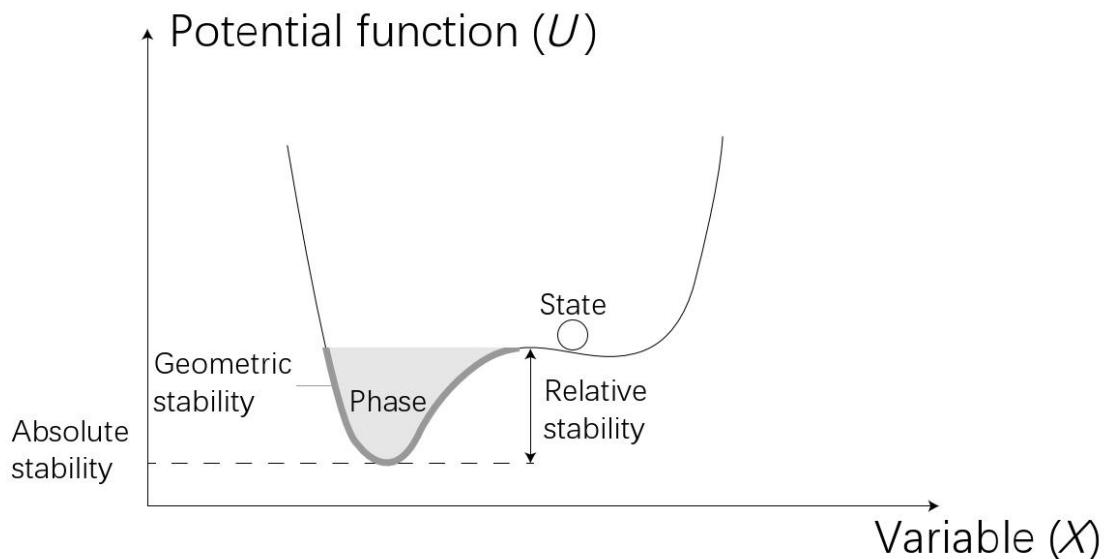
Figure 5. The relationships among the potential function, parameters, variables, state, and phase in the framework of our method.



Based on the potential function of states (defined in Equation 8), we can go one step further to define the stability of phases, which is often more important for psychological systems. From the potential landscape perspective, the stability of phases can be characterized in three ways (see Figure 6). The first one is the local minimum of the potential function within a given phase. Although the phase contains a collection of many states, the local minimum is the most stable state in the phase, hence can provide a quantitative representation. We refer to the potential of the local minimum as the *absolute stability* of the phase. Second, the potential difference between the local minimum and the barrier of the phase is directly related to the

difficulty for the system to move out of that phase. We refer to it as the *relative stability*⁴ of the phase. The difference between the two is that the former represents if given infinite time, how probable it is for the system to be in a given phase (e.g., the healthy phase in the panic disorder model); the latter represents, if putting the system in a given phase, how difficult it is for the system to escape that phase. Finally, the shape of the potential landscape within and around a certain phase represents how probable the system vibrates within or leaves a phase in a certain direction. We refer to this as the *geometric stability*. Although the geometric stability, as the shape of the landscape, is described qualitatively in the current work, we should note that this qualitative information is based on the quantitative information of the stability of the states within and around a phase of the system.

Figure 6. The relationships among the absolute stability, relative stability, and geometric stability of a psychological phase.



Besides quantifying the stability, potential landscapes can also be used for other purposes. One advantage of the potential landscape over the case simulation is that it can summarize the stability information concisely and be directly compared across different parameter settings, therefore enables systematic investigation of the influence of various parameters on the model. For example, Robinaugh et al. (2024) only provided one set of values for all 22 parameters in the model, and these parameter values were chosen based on their ability to produce reasonable output. The plausible range of each parameter, however, was not investigated. This

⁴ This is related to the term “resilience” used in some papers (e.g., Dablander et al., 2020).

is a common practice for performing case simulations for formal models but leaves the robustness of a specific parameter setting questionable. Those parameters also are related to trait-like psychological properties of the system, which are of theoretical interest. It is understandable that most modelers do not show the simulation results with all different parameter settings because it does not show the influence of parameters in an informative way. With the help of potential landscapes, these problems can be addressed more clearly.

The potential landscape can also provide guidance on model modification. It is nearly impossible to have a model that successfully explains every real-life phenomenon. Modification is often needed to continuously improve the model. Case simulations do not always provide enough guidance on how to modify the model. For example, there is a clinical phenomenon that the panic attack model could not explain: some people only have some (non-clinical) panic attacks but do not develop a panic disorder (Robinaugh et al., 2024). In the original model, however, if the panic attack happens once, the system would always develop into a panic disorder. The authors provided a way of mending as an example of model improvement, namely adding another parameter called “escape schema” (S_E). When S_E is higher, the parameter $h_{PT,E}$ in Equation 4 is higher. This parameter represents the extent to which the individual believes escaping could help to cope with the perceived threat. After adding this parameter, individuals with a low S_E would have panic attacks but would not develop a panic disorder. The rationale behind this solution is mainly based on a known theoretical mechanism that is translated into a model parameter, but not based on the model output. In other words, the simulation results only provide information about *whether* there is a problem, but not *how* the problem arises. The potential landscape method in contrast can provide insight into the problem – at least problems can be systematically analyzed from the perspective of stability.

Aim of the Current Research

The current research aims to provide a method to compute the potential landscape for psychological formal models and examine its usefulness in understanding psychological systems. First, we develop a set of tools to compute the potential landscape and related stability indicators from psychological dynamic models. Then, we illustrate how the procedure works by using the panic disorder model by Robinaugh et al. (2024). We will show how to (1) analyze the stability of states and phases from the potential landscape perspective, (2) systematically investigate the influence of various parameters from the potential landscape, and (3) use the potential landscape to guide model modification.

Methods

The method section is divided into three parts. In the first part, we show the adjustments we made to make the panic disorder model more suitable for landscape construction. In the second part, we explain the preparatory analyses we did that ensure the validity of the landscape results. In the third part, we introduce the method that we used in the main analysis. The simulations and analyses in this study were performed in R 4.1.2 (R Core Team, 2021). The replicable R scripts used for this study, as well as the animation or interactive version of the landscapes, can be found at the OSF repository of this project (<https://osf.io/ke3xb/>). An R package, *simlandr*⁵, was developed to organize the methods we used and to facilitate future applications. We try to involve minimal yet sufficient mathematics in the current article. For readers seeking more rigorous technical details of this method, we refer them to A3. *Practical information on programming issues* in Appendix A and Cui et al. (2021).

Model Adjustments

In order to construct the landscape for the system, some modifications are needed. First, we adjusted some ways of computational implementations to make the simulation more effective. The new implementation produces, in principle⁶, the same outputs as the implementation by Robinaugh et al. (2024). We refer to this model as the *original model*.

Then, as we use long-term simulations to estimate the steady-state distribution of the model, we need the model to have global ergodicity. In other words, the model should travel through its entire realistic phase space in a sufficiently long time. Complex systems often display local ergodicity breaking, which occurs when the system gets trapped in a local minimum, or, attractor state. To ensure global ergodicity, we want the model to have large enough noise terms so that the system can escape local minima. The original model does not meet this requirement because there is only one noise term on the changing rate of physical arousal (*A*). This term is not enough to ensure the ergodicity of the whole system. Therefore, we added several Gaussian noise terms on all model variables to replace the single noise term in the original model. We also calibrated the standard deviation of the noise term to make sure the panic phase still exists. Second, in the original model, there is a short-long term difference: some variables (e.g., arousal schema, *AS*) are updated every “day”, while other variables are updated per “minute”. However, in a Monte Carlo simulation, the purpose is to estimate the steady-state distribution, not to represent the actual time scale of real-life changes. Keeping

⁵ The package and its vignettes are available at <https://cran.r-project.org/package=simlandr>.

⁶ The simulation function is implemented using Rcpp (Eddelbuettel & François, 2011) instead of the implementation in R by Robinaugh et al. (2024), and random numbers are generated using RcppZiggurat (Eddelbuettel, 2020) in this implementation. Some differences in pseudo-random number generation may lead to very minor differences in the model output. These differences, however, do not influence the results meaningfully.

the short-long term difference in the model for Monte Carlo simulation is not meaningful and will reduce the effective sample size for those slowly updated variables. Therefore, we deleted the short-long term difference by updating all the variables in the same frequency. We refer to this model as the *simplified model*.

In the original model, *AS* is an important variable that influences the stability of the system, and it changes much more slowly compared to other variables. In the simplified model, we added noise to it and made it change faster. These modifications make it difficult to see how the value of *AS* changes the stability of the system. Therefore, we also made a model that is based on the simplified model but holding *AS* constant. In this model, *AS* is a parameter instead of a variable, and it is useful for investigating the influence of *AS*. We refer to it as the *constant AS model*.

Finally, we also extracted the deterministic part of the constant *AS* model (i.e., all noise terms were deleted). This is because a model of deterministic ordinary differential equations (ODEs) enables mathematical analyses of the stable points of the model. We refer to this model as the *deterministic model*.

Preparatory Analyses

The Number and Stability of Equilibrium Points. Using the deterministic model, it is possible to analytically tell how many equilibrium states there are in the system. If the system is in its equilibrium points, all the time derivatives of the models should be zero. Moreover, if the eigenvalues of the Jacobian matrix of the system all have negative real parts, the equilibrium point is stable; if some of the eigenvalues have positive real parts, the equilibrium point is unstable (see Sayama, 2015, for explanations of this method). Therefore, the stabilities of the equilibrium points are determined by the dominant eigenvalue (λ_d), which means the eigenvalue with the largest real part. To know how many equilibrium points there are in the system, we hold every time derivative zero except for dPT/dt , and then calculate how it changes over *PT*. The result is shown in Figure A1 of Appendix A. Based on these results, we can find that when *AS* is low, there is only one zero point at $PT = 0$. However, as *AS* increases, two new equilibrium points appear.

The real part of these λ_d s is shown in Table A1. In the one-equilibrium-point cases, that point is stable; in the three-equilibrium-point cases, the first and the third points are stable, but the second one is unstable. The first and the third point here correspond to the healthy and the panic phase, respectively. The third equilibrium point only emerges when *AS* is sufficiently large. The second (unstable) equilibrium point corresponds to the saddle point of the system.

Ideally, the system can be in equilibrium there, but with a small disturbance, the system will move to either of those two stable phases. Based on the results of stability analysis, we can confirm the heuristic that the landscape we construct should have one or two phases, depending on parameter values.

Checking Convergence and Determining the Simulation Length. Monte Carlo estimation of the steady-state distribution is only valid if the simulation *converges*, which means that there are enough data points sampled that the joint distribution of the variables will not change even when the simulation length is extended. We checked this by comparing the distribution of key variables in the initial, middle, and final stages of the simulation. With a simulation length of 10^7 timesteps, the distributions in different stages are sufficiently stable (Figure A2). Therefore, we use 10^7 as the simulation length for constructing landscapes.

Main Analysis

The Stability of States and Phases. For constructing the potential landscape function for each possible state, we first estimate the steady-state distribution (PSS) of the model with Monte Carlo simulation. The raw potential landscape function is defined in a high dimensional space, where the dimension equals the number of variables in the model. To make this function understandable, we need to perform a dimensionality reduction. This was done with a simple but widely used approach, which is obtaining the marginal distributions (e.g., Li & Wang, 2013; Zhang et al., 2019). This method is capable of visualizing up to three selected variables for the model. Kernel smooth methods were used to calculate smooth distribution density with a reduced set of variables each time, and Equation 8 was used to calculate the potential landscape.

For the stability of phases, we first calculated the absolute stability by finding the local minima of the potential function within each phase. Then, we looked for the *minimum energy path* (MEP) and the *saddle point* between the two phases. The minimal energy path is the path that the system would be most likely to travel from one local minimum to another if the system was purely gradient. The point with the highest potential in the MEP is the saddle point. It can be proved that this path should first go along the steepest ascending path from the starting point and then go along the steepest descending path to the end point (E & Vanden-Eijnden, 2010). Its geometric form, from which a Dijkstra algorithm (Dijkstra, 1959) can be derived, was used to find the MEP (Heymann & Vanden-Eijnden, 2008). After that, the relative stability defined by the barrier height was calculated as the potential difference between the saddle point and the local minima. The geometric stability of the phases is described qualitatively.

The Effect of Parameters. To investigate the influence of parameters on the stability of the phases, multiple simulations with different parameter values were performed. For a single parameter, its value was sampled evenly within a parameter space. The range of this parameter space is roughly centered around the original parameter values used by Robinaugh et al. (2024) with a plausible width that is large enough to clearly show the influence of the parameter on the potential landscape. The potential landscape was constructed separately for each parameter value, and the barrier height was computed respectively. Later, the barrier heights were compared across parameter values to show the influence of parameters on the difficulty for the system to escape a certain phase and transition into another phase. For the joint influence of two parameters, a sample grid was made for the combination of parameter values, and the potential landscape was calculated for each condition.

Model Modification. Based on the information provided by the potential landscape, we propose the following general strategy for model modification. In the first step, the problem of the model is identified. It is usually some differences in stability between model outcome and real-life phenomena. In the second step, the reason for this inconsistency is analyzed from the potential landscape perspective. It can be that the stabilities of different phases are not suitable, the barrier height between phases is too high or too low, or the landscape has more or fewer phases than it should have. After that, the model is adjusted accordingly. For example, adding or removing time derivative terms can tilt the landscape and stabilize the states in a certain direction. Finally, both case simulation outputs and the potential landscape of the modified model are checked to test if the problem has been solved. Using this strategy, we analyzed a problem of the panic disorder model, provided a way of modification, and evaluated the modified model.

Results

Stability of States and Phases

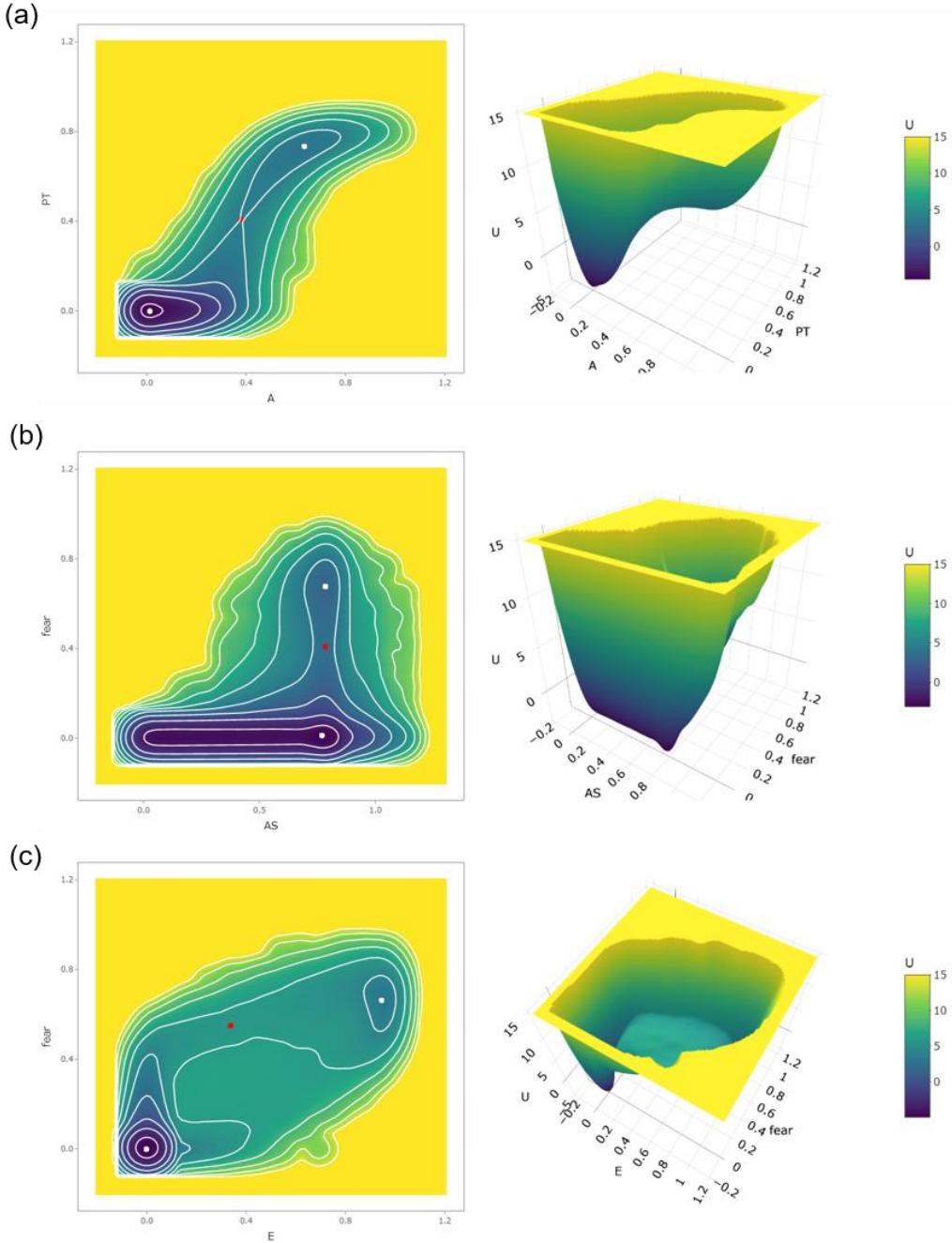
Among all the model variables, physical arousal (A), perceived threat (PT), and *fear* are the core variables representing the symptoms of panic disorder. A higher value of these variables represents higher symptom severity. Besides that, arousal schema (AS) represents the key control variable for the disorder. Therefore, we first constructed potential landscapes for these variables with the simplified model. In the potential landscape of A and PT (Figure 7a), we can find two local minima. The position of the first one is at $A = 0.01$, $PT = 0.00$ ($U = -4.47$), and the second one is at $A = 0.63$, $PT = 0.73$ ($U = 3.01$). The saddle point is at $A = 0.38$, $PT = 0.41$ ($U = 3.99$). The barrier heights are $\Delta U = 8.46$ and $\Delta U = 0.98$ for the two phases, respectively. The first phase has a lower symptom severity, hence corresponds to the healthy phase of the

system. The symptom severity of the second one is higher, hence corresponds to the panic phase of the system. The potential of the local minimum within the healthy phase is lower than that of the panic phase, and the barrier height of the healthy phase is higher than that of the panic phase, indicating that the healthy phase has a higher absolute and relative stability. Both phases show a regular circle-like shape, which means that the system state tends to vibrate around the local minimum symmetrically in both A and PT directions. There is a single pathway connecting them, which is the path that the system is likely to take when transitioning from one to the other. For example, if the system is going from the healthy phase to the panic phase, A and PT will increase together until the system reaches the region of the panic phase.

The landscape of AS and $fear$ is shown in Figure 7b. Similarly, the phase with a lower fear level is the healthy phase, and the phase with a higher fear level is the panic phase. According to the landscape, the panic phase only appears when AS is high enough. The local minimum with lower fear is at $AS = 0.77, fear = 0.01 (U = -3.08)$, and the local minimum with higher fear is at $AS = 0.78, fear = 0.68 (U = 2.83)$. The saddle point is at $AS = 0.78, fear = 0.41 (U = 3.54)$. The barrier heights are $\Delta U = 6.62$ and $\Delta U = 0.71$ for the two phases, respectively. Again, both absolute and relative stability indices support that the healthy phase is more stable than the panic phase. While the panic phase shows a circle-like shape, the healthy phase shows a slender shape, indicating that the potential landscape is rather flat along the AS direction. This shows that when AS is lower than the local minimum point (0.78), the higher and lower AS ranges do not differ much in stability, which means there is little resistance for AS to increase or decrease within the healthy phase. However, when AS is higher than the local minimum point (0.78), there is a strong tendency for AS to decrease. This indicates that there are some mechanisms in the system preventing AS from rising too much.

Besides the most central variables, the panic disorder model also has many other elements. Here, we calculated the landscape of escape (E) and $fear$ to show the stability-related properties of E , which is an important behavioral mechanism in panic disorder. This potential landscape is shown in Figure 7c. As the previous potential landscapes, this potential landscape also shows two phases, with the healthy one being more stable than the panic phase (local minimum in the healthy phase: $E = 0.00, fear = 0.00, U = -4.97$; local minimum in the panic phase: $E = 0.95, fear = 0.66, U = 3.10$; saddle point: $E = 0.34, fear = 0.55, U = 5.65$; barrier heights: $\Delta U = 10.63$ and $\Delta U = 2.55$). E is in a lower range when the system is in the healthy phase and in a higher range when the system is in the panic phase, which indicates that the panic phase is related to a higher tendency of escaping.

Figure 7. Bivariate potential landscapes of (a) A and PT , (b) AS and *fear*, and (c) E and *fear*. All three landscapes were constructed using the simplified model. The first plot in each row is the 2D heatmap with contours, and the second plot is the 3D surface plot. The white and red dots on the 2D heatmaps represent the local minima of the phases and the saddle points, respectively.



While the potential landscape (Figure 7c) around two local minima shows a regular shape, the landscape between them shows a unique phenomenon: between the healthy phase and the panic phase, there are two pathways instead of one. Between these two paths, there is also a

small “hill” inside the “valley”. This indicates that there are two possible paths of transitioning from one phase to the other. The path taken by the system from the healthy phase to the panic phase is different from the path taken by the system from the panic phase back to the healthy phase. For systems where two variables have asymmetric relationships (e.g., the Lotka–Volterra predator-prey model), this type of behavior is not uncommon. It indicates that maybe the process of a panic attack is a one-way street: if a panic attack has started, it may not be possible to stop it before the *fear* level reaches its peak. Further, because the path taken by the system from the healthy phase to the panic phase is different to the path from the panic phase to the healthy phase, maybe it is possible to tell the transition direction from the state of the system.⁷

Influence of Parameters

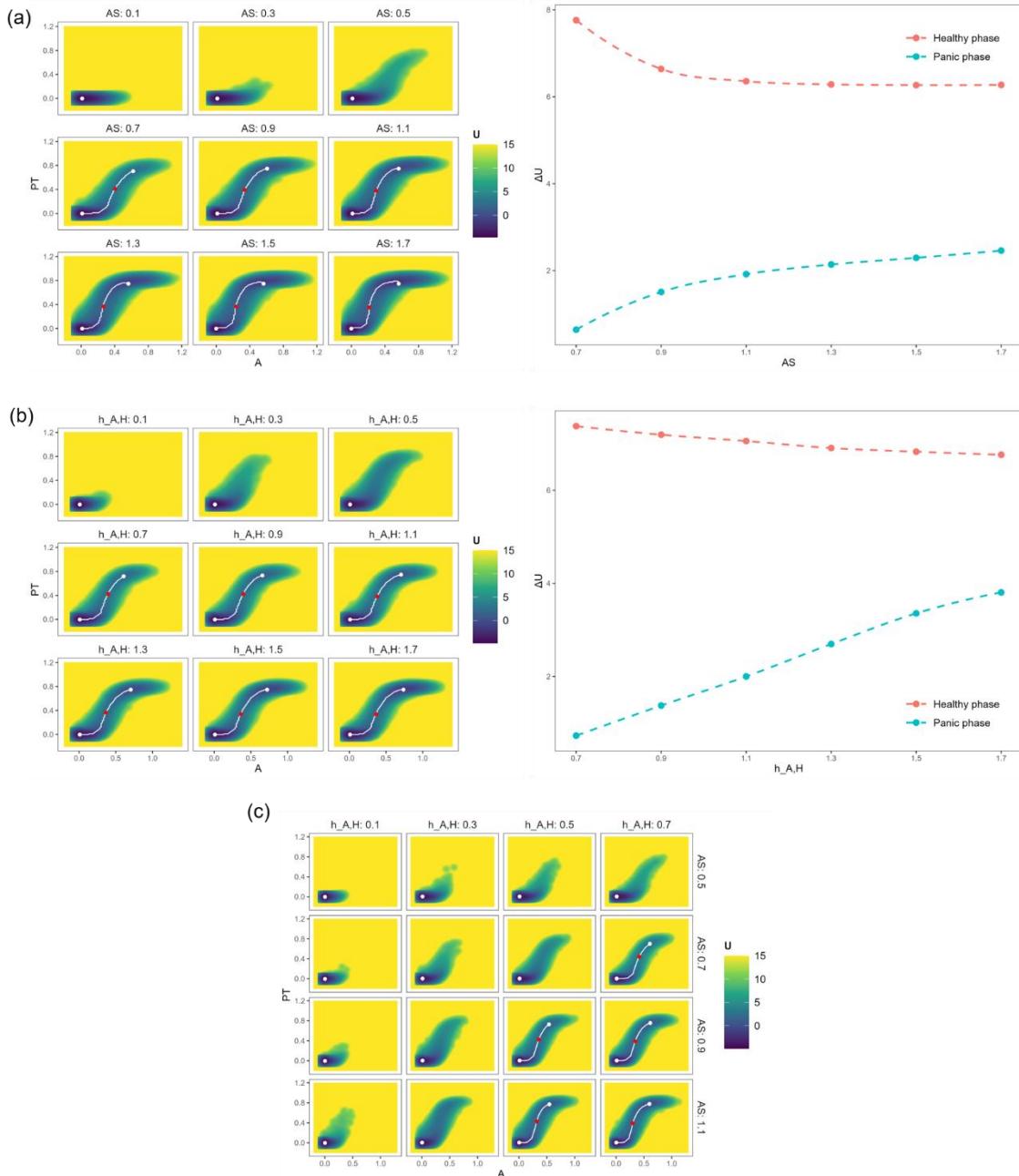
As mentioned earlier, *AS* is an important variable in the model that controls the stability of the phases. In the constant *AS* model, *AS* is a parameter instead of a variable. Therefore, we can first investigate *AS* again using the method for parameters. We constructed a series of landscapes from the constant *AS* model (Figure 8a). Comparing those landscapes, we can find that when *AS* is 0.3 or lower, there is only one phase in the system, namely the healthy phase. The panic phase only appears when *AS* is 0.7 or higher. For the critical condition that *AS* = 0.5, the system can go to the region of panic phase, but there is not a local minimum in that region (i.e., the potential landscape increases monotonically in the direction towards higher *A* and *E*). We call it a *quasi-stable phase*. Because there is no local minimum for a quasi-stable phase, its potential value and barrier height cannot be calculated.

The barrier heights for the two phases are also shown in Figure 8a. As *AS* becomes larger, the healthy phase gradually becomes more unstable, while the panic phase becomes more stable. Nevertheless, the healthy phase is always more stable than the panic phase. These results again confirm the findings above.

While *AS* is related to the relationship between *A* and *PT*, which is at the core of the system, there are also some other variables whose effects are less obvious. Here we investigate another parameter, $h_{A,H}$, as an example. This parameter represents the relationship between physical arousal (*A*) and homeostatic feedback of arousal (*H*): the physical arousal at a later point in time tends to reduce more if it is higher in the previous time point. *A* has a positive influence on *H* and *H* has a negative influence on *A*. Therefore, the state of *A* is “stored” in *H* and will be influenced by *H* later. When $h_{A,H}$ is higher, the influence of *A* on *H* is weaker.

⁷ The landscape for *AS*, *E*, and *fear* also confirms the results above. In this landscape plot, three variables are represented with x-, y-, and z-axis, and the potential value is represented by color. See the OSF repository (<https://osf.io/ke3xb/>) for the visualization of this landscape.

Figure 8. Bivariate potential landscapes of A and PT with different (a) AS , (b) $h_{A,H}$, and (c) both AS and $h_{A,H}$. The landscapes were constructed using the constant AS model. The first plot in each row is the potential landscapes represented in 2D heatmaps. The white dots represent the local minima of the phases in the landscape, the white lines connecting two white dots represent the minimum energy path between two local minima, and the red dots represent the saddle points in the paths. The second plot in (a) and (b) shows the barrier heights (ΔU) of both phases in the potential landscapes for different parameter values. For cases where there is only one phase in the system, it is not possible for the system to transition to an alternative phase, hence the barrier height cannot be defined. Therefore, barrier heights were not calculated for those cases.



The landscapes and barrier heights are shown in Figure 8b. Comparing the landscapes with different $h_{A,H}$ values (Figure 8b) and those with different AS values (Figure 8a), we can find that both parameters have a similar role in controlling the landscape of A and PT . When arousal has weaker homeostatic feedback of arousal (i.e., when the delayed negative feedback gets weaker), the panic phase gradually appears and stabilizes. If changing both parameters together, their roles are similar and independent: increasing either will stabilize the healthy phase of the system (see Figure 8c). This suggests that the same phenomenon of the system (e.g., panic disorder) can stem from different underlying mechanisms.

Guidance for Model Modification

Potential landscapes can also guide model modification, as we show here with the panic disorder model as an example. As mentioned by Robinaugh et al. (2024), the problem of the original model is that the system should have non-clinal panic attacks, but the model failed to generate them. From the landscape with AS and *fear* (Figure 7b), we can see that the panic phase only appears when AS is high enough. However, the landscape in the AS direction is rather flat, which means that AS does not have a strong tendency to increase or decrease within its plausible range. This is related to the problem mentioned above because whenever AS happens to be in the high range, it does not go back. Hence, an increase in AS can easily lead to a panic disorder.

To solve this problem, what we should do is tilt the landscape to make the low- AS region more stable. The way to tilt the landscape is straightforward. As the gradient of the landscape corresponds to the changing rate of the variables, we can simply add a negative term to the dynamic function to tilt it towards zero. Here we added a small negative term to the time derivative of AS :

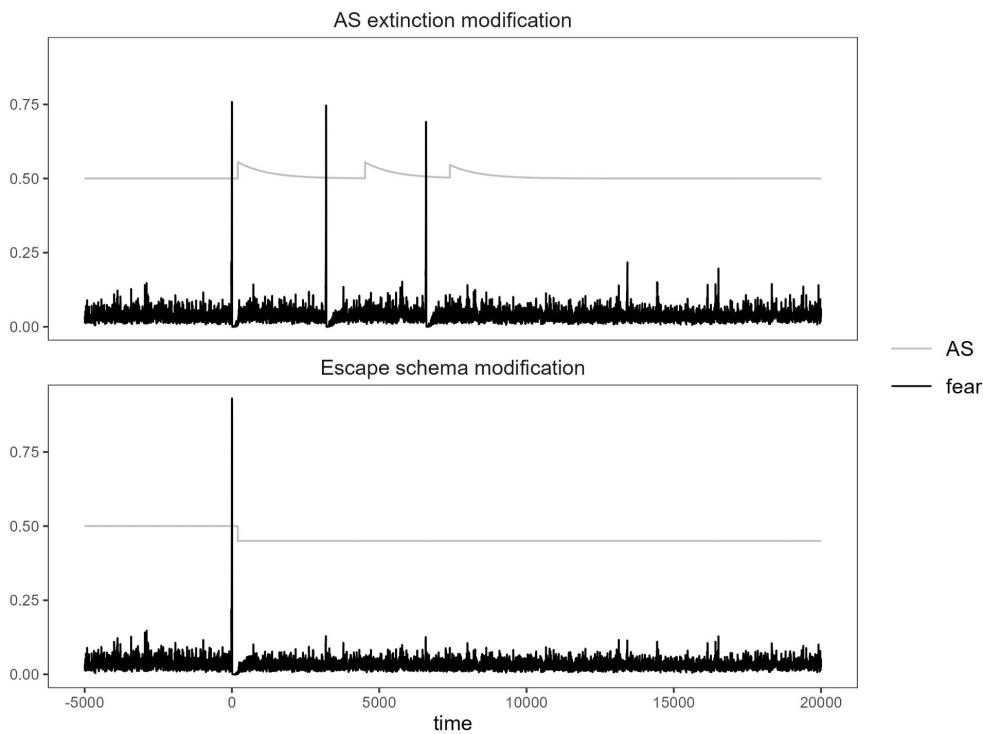
$$\frac{dAS}{dt} = g(X) + \begin{cases} -r_{\text{extinction}}(AS - AS_{\text{baseline}}), & AS > AS_{\text{baseline}} \\ 0, & AS \leq AS_{\text{baseline}} \end{cases} \quad (9)$$

Where $g(X)$ represents the terms in the original functions, and the term to the right of the brace is the added term: when AS is larger than a given baseline value (AS_{baseline}), it declines exponentially. This setting can add a small tendency towards lower AS around its critical range while not letting it decline to zero. To distinguish this way of modification and the way in the original paper by Robinaugh et al. (2019, which used a higher S_E to increase the parameter $h_{PT,E}$ as a constant; see Equation 4), we refer to the modification in Equation 9 as the *AS extinction modification* and the modification by Robinaugh et al. (2024) as the *escape schema modification*.

Both modifications show a good effect on solving the problem in the original model. As shown in Figure 9, for both modifications, the system can have a limited number of panic

attacks without later going into a full-blown panic disorder. After the first panic attack, the *AS* value of the original model increases and stays in a higher value until the next panic attack, which further increases *AS* (Figure 3), whereas the *AS* value decreases after the first panic attack for both modified models. The difference in the direction of change for *AS* shows the effects of both modification mechanisms. The potential landscapes of both modifications are shown in Figure 10. Comparing them with the landscape of the original model (Figure 7b), the landscapes of those modified models show the expected tendency that the lower-*AS* region of the healthy phase is more stable.

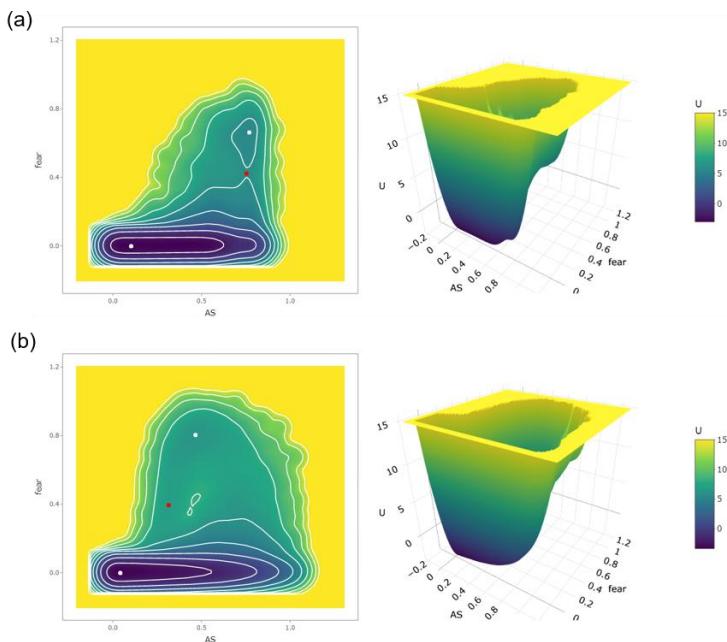
Figure 9. Model simulation results of *fear* and *AS* for the two modifications. The first panic attack of the system appears probabilistically. To make simulation results comparable, *time* was set as zero at the first panic attack (the time when the peak value of *fear* is reached). In the *AS* extinction modification, $r_{\text{extinction}} = 0.001$, $AS_{\text{baseline}} = 0.5$; in the escape schema modification, $S_E = 0$.



However, there are several important differences between the two modifications. For the escape schema modification, the *fear* value reaches a higher peak level and the *AS* value directly decreases after the panic attack; in the *AS* extinction model, the peak *fear* value is similar to the case in the original model, and the *AS* value first increases before it declines back to a lower value (Figure 9). Why is this the case? The landscapes of the models can provide

further information. The general shape of the landscape of the AS extinction modification (Figure 10a) is similar to the original model (Figure 7b), but only the stability of the lower AS range of the healthy phase is increased. The shape of the landscape of the escape schema modification (Figure 10b), however, is much different from the original model, especially around the panic phase. There is a small island in the middle between the two phases. As explained earlier (for the landscape of E and *fear* in Figure 7c), this indicates that the pathway the system takes from the healthy phase to the panic phase is different than the other way around, so that the system does not recover in the same way. In this model, this means that the system goes to the panic phase when AS is high, and then the AS level declines *during* the same panic attack. When the system goes back to the healthy phase, the AS value is already at a lower level. In the original model and the AS extinction modification, AS decreases at a slower time scale, so that the system recovers to the healthy phase with a similarly high AS value. These differences enable further theoretical and empirical examinations of these two modifications.

Figure 10. Bivariate potential landscapes of AS and *fear* for (a) the AS extinction modification and (b) the escape schema modification. The first plot in each row is the 2D heatmap with contours, and the second plot is the 3D surface plot. The white and red dots on the 2D heatmaps represent the local minima of the phases and the saddle points, respectively.



Discussion

We here introduced a new method to construct potential landscapes for multivariate psychological formal models. Based on the steady-state distribution, the stability of any state

of the system can be quantified with the potential function. We illustrated the method with the panic disorder model by Robinaugh et al. (2024). After several adjustments to the original model, we constructed potential landscapes for the system and analyzed the absolute, relative, and geometric stability of the healthy phase and the panic phase. Then, we examined the influence of two model parameters on the potential landscape of the system. Finally, based on the information from the potential landscape, we came up with a new way of model modification and compared the simulation output and the potential landscape of it with the modification suggested by Robinaugh et al. (2024).

The results of the potential landscape showed that there are one or two phases in the system, depending on the parameter settings. The healthy phase is always present, while the panic phase only appears under certain conditions (e.g., with high *AS*) and is always less stable than the healthy phase. Increasing *AS* can thus stabilize the panic phase and destabilize the healthy phase. These results are well aligned with the conclusions from the case simulations in the original paper. It is important to note that we do not think that the potential landscape method can or should replace case simulations. On the contrary, we claim that both methods provide important information about the nature of psychological formal models. The advantage of our method is that it can present the concept of stability in a clear, explicit way. By filtering out the time-related information in the model output, the stability-related information that does not change over time emerges clearly on the potential landscapes. Instead of relying on the heuristics from observation, the stability of states is now specified as positions in a potential landscape. Specifically, the stability of psychological phases can be described accurately based on three aspects: absolute stability, relative stability, and geometric stability. This makes the stability information more apparent for researchers, hence facilitating understanding and communication of the models.

Moreover, the potential landscape method also enables easy investigation of variables and parameters. A common problem for psychological formal models is that the number of variables and parameters used for constructing the model is much higher than the critical variables that are examined to evaluate the model. In this paper, we showed how to construct the potential landscape for *A*, *PT*, *AS*, *E*, and *fear* and how the potential landscape changes with *AS* and $h_{A,H}$. With the tools we provided in the simlandr package, the same landscape construction method can be easily extended to all other variables and parameters in the model. Our method allows researchers to evaluate each of their variables and parameters systematically and see if the resulting potential landscape is consistent with theory and empirical findings, which parameter range produces the expected behavior, and how the

parameters influence the stability. This not only makes model evaluation more effective and comprehensive but also helps to clarify the scope and boundaries of a model or theory.

Based on the information from the potential landscape, we provided a new modification of the original model. The advantage of using a stability measure to guide model modification is that, in some cases, the discrepancy between model output and real-life phenomena is more closely related to the stability of states or phases rather than the simulated trajectories. In the case of the panic disorder model, for example, the problem of non-clinical panic attacks can be directly attributed to the geometric stability of the healthy phase (i.e., its landscape on the AS direction is too flat). Also, because of the close link between the potential landscape and the dynamic functions, changing the landscape of the system is usually not difficult, making it straightforward to solve the identified problems. In this work, we proposed a way of modification for the panic disorder model based on its potential landscape. There is also a possible theoretical explanation for this modification, namely fear extinction. After the association (arousal schema, AS) between physical arousal (*A*) and perceived threat (*PT*) is learned, even when there are no new fear-inducing events, the strength of this association still decreases (Mattera et al., 2020; Milad & Quirk, 2012). It is important to emphasize that the decision to add a negative value to the AS time derivative in the AS extinction modification was inspired by the potential landscape of the model, not based on a theoretical analysis. Nevertheless, it later helps to point out a direction in which researchers could search for relevant theories.

Limitations and Possible Pitfalls

Despite the advantages of the method discussed above, we also want to point out several limitations and possible pitfalls of our method. First, the potential landscape is calculated from a generalized potential function. This means that it does not contain all the dynamic information of the model. In other words, psychological systems are not totally the same as a ball on a landscape. Some additional non-gradient forces also influence the system, which can drive it in a different direction instead of the direct path towards the most stable state. For example, in the landscape of *E* and *fear*, there are two paths connecting two phases of the system. We suppose these paths are different in directions (i.e., the system state goes through one path to the healthy phase and the other path to the panic phase). However, the choice of which path to take is influenced by some non-gradient forces, which are not shown in the potential landscape. In psychology, non-gradient forces may be related to emotional or behavioral inertia, which means that some emotional or behavioral variables in the system may be more resistant to change (e.g., *E* and *fear* in the panic disorder model; Alós-Ferrer et al.,

2016; Kuppens et al., 2010). These non-gradient forces are canceled out when calculating the stability of states. Nevertheless, researchers should be aware of their existence and investigate their influences if the dynamic properties of the system are also a concern.

Second, to ensure the ergodicity in the Monte Carlo simulation, we added some noise terms to the original model. We suggest that researchers do the same if their simulation is difficult to converge. However, we want to point out that for some systems these noise terms can affect the stability of phases. For example, Van den Broeck et al. (1994) showed that for a specific kind of dynamic system, an ordered phase only exists if the noise is in a certain range. The conditions under which noise terms can affect the stability of phases are not yet clear. Further investigations on this issue are needed.

Third, although psychological models usually have a large number of variables, the potential landscape can only be visualized in a lower (up to three) dimensional space. This means one phase or transition path on the landscape can actually be several ones in the high-dimensional space. For example, there is only one pathway between the healthy phase and the panic phase in the landscape of *A* and *PT*, but two pathways in the landscape of *fear* and *E*. Therefore, we suggest researchers look into different combinations of variables and investigate how the phases and paths in those low dimension landscapes correspond.

Fourth, the potential landscape can only show the stability of individual states. For a system that has periodic or chaotic phases (e.g., Schiepek et al., 2017), the potential landscape may not be able to clearly represent its stability. In these cases, some preprocessing of the original variables may be needed. For example, if two phases of a system have similar mean values but differ in their variation, then the moving-window standard deviation can be used to construct the landscape instead of the original variable values. The suitable way of transformation depends on the exact system of interest.

Future Directions

Incorporating the potential landscape method opens up new avenues for psychological (modeling) research. Here we want to point out several possible directions. First, there is still much room to further develop the potential landscape method for psychological models. Apart from the generalized potential function used in our work, there are also several other potential functions that have different definitions and calculation methods (see P. Zhou & Li, 2016, for a review). These methods may have a higher requirement for the form of dynamic functions, but they also have strengths in representing other aspects of stability (e.g., the quasi-potential

landscape by J. X. Zhou et al., 2012, emphasizes more on the transition path between states). We encourage future research to test the use of those methods for psychological models.

Second, we encourage modelers to further explore the usage of this method and apply this method to different kinds of models. In this paper, we showed several applications of the potential landscape methods, namely representing stability, investigating the influence of parameters, and guiding modifications, and the formal model we used is an emotion-cognition-behavior model of a mental disorder. However, we are confident that the usage of this method is not limited to this range. For example, it may be possible to draw a phase diagram to show the parameter ranges where different phases exist and use the potential landscape as a way to choose parameter values systematically. Some cognitive models also have multistability (e.g., Kogo et al., 2011), which can possibly be analyzed with this method. We look forward to future researchers exploring those possibilities.

Third, potential landscapes can also be helpful in clinical practice. Understanding individual differences of psychological phases and clinical change processes is critical to establish more effective therapies. From the potential landscape perspective, the difference in people's parameter values makes them differ in their vulnerability to mental disorders, and successful therapies are related to changes in parameter values. This idea can be better understood if we can construct landscapes with different parameters to represent different individuals or different therapeutic stages. We illustrated how systems with different AS and $h_{A,H}$ values differ in the stability of the healthy phase and the panic phase. These AS and $h_{A,H}$ values may correspond to different individuals and/or stages. Apart from panic disorder, this approach can also be readily applied for conceptualizing other mental disorders. Constructing landscapes from empirical data, however, requires future methodological development. Previous research found that recovering the psychological dynamics from experience sampling data is difficult, while the distribution of psychological variables is more assessable (Haslbeck & Ryan, 2021). The generalized potential landscape is defined from the steady-state distribution of the system, which can be possibly estimated from the observed variable distribution. Therefore, maybe fitting the potential landscape directly from empirical data is a possible alternative to fitting the dynamic functions.

Finally, we want to note that, in the abstract sense, the form of psychological formal models does not differ much from many dynamic models in biology, chemistry, and other natural science fields. For example, we coincidentally found out that the panic disorder model also has an equivalent form represented by chemical reactions (see A4. Equivalent chemical representations in Appendix A). While formalizing psychological theory is a new trend in

psychology, the same effort has been undertaken in many fields, with various analytical methods readily available. With the current work as an example, we would encourage further interdisciplinary cooperation in the psychological modeling field. We hope the insights from expertise in all fields of science can somehow come together, helping us to better understand human psychology, a complex yet fascinating subject.

Conclusion

The stability of states and phases is an important property for psychological formal models, yet not concretely addressed with the common case simulation method. By incorporating the generalized potential function by Wang et al. (2008) and Monte Carlo simulation, we developed a method to construct the potential landscape for multivariate psychological dynamic models. This method can contribute to a better understanding of the stability concept, the influence of model parameters, and the way to modify a model. We hope this method can help researchers to better evaluate and develop their models and ultimately help to guide clinical practice in the future.

Acknowledgments

The authors would like to thank Freek Oude Maatman who provided valuable inputs for the conceptual basis of this work. They would like to thank Wendy Post, Jill de Ron, and Niklas D. Neumann for their feedback on previous versions of this manuscript. They also thank Alberto Maydeu-Olivares, Emilio Ferrer, and the anonymous reviewers who gave helpful suggestions during the editorial process.

Chapter 3

simlandr: Simulation-Based Landscape Construction for Dynamical Systems

This chapter is based on:

Cui, J., Olthof, M., Lichtwarck-Aschoff, A., Li, T., & Hasselman, F. (in press). simlandr: Simulation-based landscape construction for dynamical systems. *The R Journal*. Preprint available at PsyArXiv: <https://doi.org/10.31234/osf.io/pzva3>



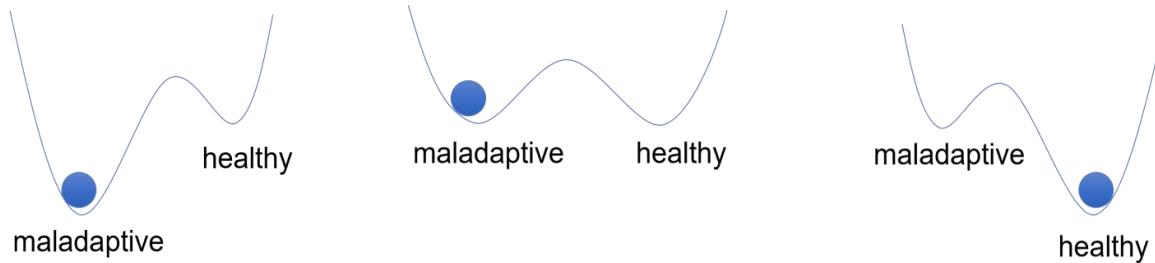
Abstract

We present the simlandr package for R, which provides a set of tools for constructing potential landscapes for dynamical systems using Monte Carlo simulation. Potential landscapes can be used to quantify the stability of system states. While the canonical form of a potential function is defined for gradient systems, generalized potential functions can also be defined for non-gradient dynamical systems. Our method is based on the potential landscape definition from the steady-state distribution, and can be used for a large variety of models. To facilitate simulation and computation, we introduce several novel features, including data structures optimized for batch simulations under varying conditions, an out-of-memory computation tool with integrated hash-based file-saving systems, and an algorithm for efficiently searching the minimum energy path. Using a multitable cell differentiation model as examples, we illustrate how simlandr can be used for model simulation, landscape construction, and barrier height calculation. The simlandr package is available at <https://CRAN.R-project.org/package=simlandr>, under GPL-3 license.

Introduction

To better understand a dynamical system, it is often important to know the stability of different states. The metaphor of a potential landscape consisting of hills and valleys has been used to illustrate differences in stability in many fields, including genetics (Waddington, 1966; Wang et al., 2011), ecology (Lamothe et al., 2019), and psychology (Olthof et al., 2023). In such a landscape, the stable states of the system correspond to the lowest points (minima) in the valleys of the landscape. Just like a ball that is thrown in such a landscape will eventually gravitate towards such a minimum, the dynamical system is conceptually more likely to visit its stable states in which the system is also more resilient to noise. For example, in the landscape metaphor of psychopathology (Figure 1), the valleys represent different mental health states, their relative depth represents the relative stability of the states, and the barriers between valleys represent the difficulty of transitioning between these states Hayes and Andrews (2020). When the healthy state is more stable, the person is more likely to stay mentally healthy, whereas when the maladaptive state is more stable, the person is more likely to suffer from mental disorders.

Figure 1. Illustration of the ball-and-landscape metaphor commonly used in the field of psychopathology.



Yet, formally quantifying the stability of states is a nontrivial question. Here we present an R package, *simlandr*, that can quantify the stability of various kinds of systems without many mathematical restrictions.

Dynamical systems are usually modeled by stochastic differential equations, which may depend on the past history (i.e., may be non-Markovian, Stumpf et al., 2021). They take the general form of

$$dX_t = b(X_t, H_t)dt + \sigma(X_t, H_t)dW, \quad (1)$$

where X_t is the random variable representing the current state of the system and H_t represent

the past history of the system $\mathbf{H}_t = \{\mathbf{X}_s | s \in [0, t)\}$ ¹. The first term on the right-hand side of Equation 1 represents the deterministic part of the dynamics, which is a function of the system's current state $\mathbf{b}(\mathbf{X}_t, \mathbf{H}_t)$. The second term represents the stochastic part, which is standard white noise $d\mathbf{W}$ multiplied by the noise strength $\sigma(\mathbf{X}_t, \mathbf{H}_t)$.

If the dynamical equation (Equation 1) can be written in the following form

$$dX = -\nabla U dt + \sqrt{2dW}, \quad (2)$$

then U is the potential function of the system.² However, this is not possible for general dynamical systems. The trajectory of such system may contain loops which are not possible to be represented by a gradient system (this issue was compared to Escher's stairs by Rodríguez-Sánchez et al., 2020). In this case, further generalization is needed. The theoretical background of simlandr is the generalized potential landscape by Wang et al. (2008), which is based on the Boltzmann distribution and the steady-state distribution of the system. The Boltzmann distribution is a distribution law in physics, which states the distribution of classical particles depends on the energy level they occupy. When the energy is higher, the particle is exponentially less likely to be in such states

$$P(x) \propto \exp(-U). \quad (3)$$

This is then linked to dynamical systems by the steady-state distribution. The steady-state distribution of stochastic differential equations is the distribution that does not change over time, denoted by P_{SS} which satisfies

$$\frac{\partial P_{SS}(x, t)}{\partial t} = 0. \quad (4)$$

The steady-state distribution is important because it extracts time-invariant information from a set of stochastic differential equations. Substituting the steady-state distribution to Equation 3 gives Wang's generalized potential landscape function (Wang et al., 2008)

$$U(x) = -\ln P_{SS}(x). \quad (5)$$

If the system has ergodicity (i.e., after sufficient time it can travel to all possible states in the state space), the long-term sample distribution can be used to estimate the steady-state distribution, and the generalized potential function can be calculated.

Our approach is not the only possible way for constructing potential landscapes. Many other theoretical approaches are available, including the SDE decomposition method by Ao (2004) and the quasi-potential by Freidlin and Wentzell (2012). Various strategies to numerically compute these landscapes have been proposed (see Zhou & Li, 2016, for a review).

¹ The corresponding variable representing positions in the state space is not a random variable, so we use lowercase x for it. This convention will be followed throughout this article.

² Under zero inertia approximation.

However, available realizations are still scarce. To our knowledge, besides simlandr, there are two existing packages specifically for computing potential landscapes: the waydown package (Rodríguez-Sánchez et al., 2020) and the QPot package (Dahiya & Cameron, 2018; Moore et al., 2016). The waydown package uses the skew-symmetric decomposition of the Jacobian, which theoretically produces landscapes that is similar to Wang et al. (2008; but see Cui et al. (2023) for a potential technical issue with this package.) The QPot package uses a path integral method that produces quasi-landscapes following the definition by Freidlin and Wentzell (2012). Because of the analytical methods used by waydown and QPot, they both require the dynamic function to be Markovian and differentiable in the whole state space. Moreover, they can only be used for systems of up to two dimensions. simlandr, in contrast, is based on Monte Carlo simulation and the steady-state distribution. It does not have specific requirements for the model. Even for the models that are not globally differentiable, have history-dependence, and are defined in a high-dimensional space, simlandr is still applicable (e.g., Cui, Lichtwarck-Aschoff, Olthof, et al., 2023). Therefore, simlandr can be applied to a much wider range of dynamical systems, illustrate a big picture of dominated attractors, and investigate how the stability of different attractors may be influenced by model parameters. As a trade-off, simlandr is not designed for rare events sampling in which the noise strength $\diamond (X)$ is extremely small, nor for the precise calculation of the tail probability and transition paths. Instead, it is better to view simlandr as a semi-quantitative tool that provides a broad overview of key attractors in dynamic systems, allowing for comparisons of their relative stability and the investigation of system parameter influences. We will show some typical use cases of in later sections. Some key terms used in this article are summarized in Table 1.

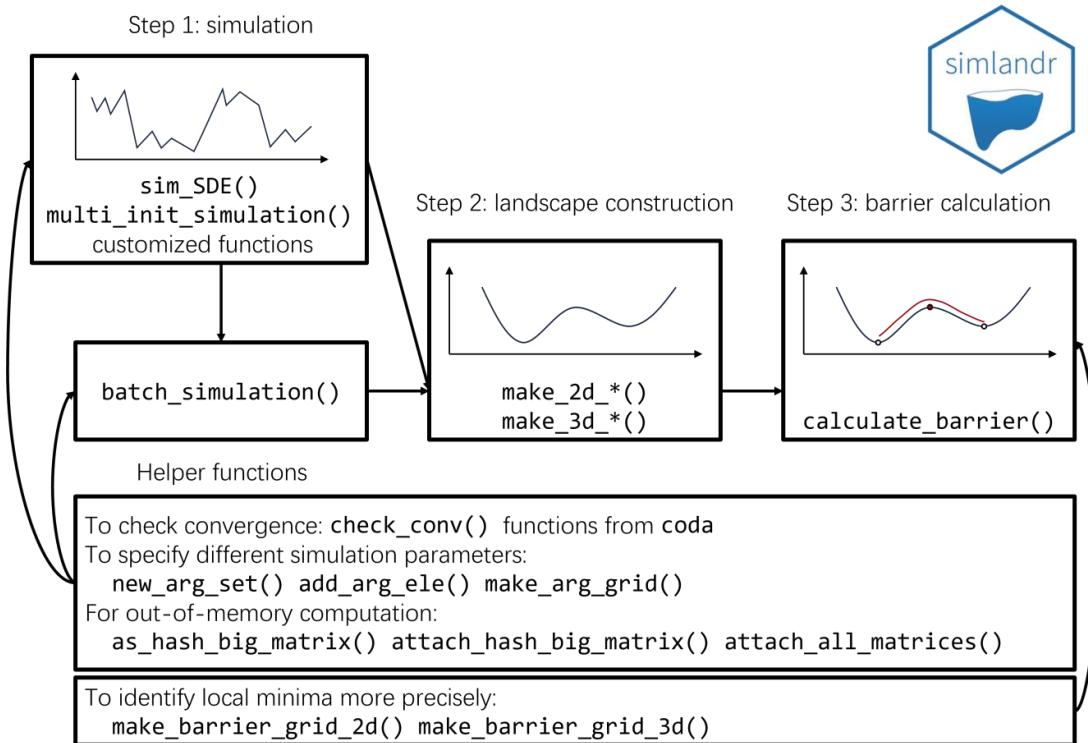
Table 1. Summary of key terms used in this article.

Term	Explanation
Potential landscape metaphor	A conceptual metaphor representing the stability of a complex dynamic system as an uneven landscape, with a ball on it representing the system's state. This can be quantitatively realized in various ways (Zhou and Li 2016).
Gradient system	A system whose deterministic motion can be described solely by the gradient of a potential function.
Non-Markovian system	A system whose future evolution depends not only on its current state but also on its past history.
Steady-state distribution	The probability distribution of a dynamic system that remains unchanged over time.
Ergodic system	A dynamic system that, given enough time, will eventually pass through all possible states.
Minimum energy path (MEP)	A transition path linking two local minima and passing through a saddle point (Wan et al. 2024). It is always parallel to the gradient of the energy landscape, representing an efficient transition route.

Design and Implementation

The general workflow of simlandr involves three steps: model simulation, landscape construction, and barrier height computation. See Figure 2 for a summary.

Figure 2. The structure and workflow of simlandr.



Step 1: Model Simulation

For the first step, a simulation function should be created by the user. This function should be able to simulate how the dynamical system of interest evolves over time and record its state in every time step. This can often be done with the Euler-Maruyama method. If the SDEs are up to three dimensions and Markovian, a helper function from simlandr, `sim_SDE()`, can also be used. This function is based on the simulation utilities from the Sim.DiffProc package (Guidoum & Boukhetala, 2020) and the output can be directly used for later steps. Moreover, the `multi_init_simulation()` function can be used to simulate trajectories from various starting points, thus reduce the possibility for system to be trapped in a local minimum. The `multi_init_simulation()` function also supports parallel simulation based on the future framework to improve the time efficiency (Bengtsson, 2021).

For Monte Carlo methods, it is important that the simulation *converges*, which means the distribution of the system is roughly stable. The precision of the steady-state distribution

estimation, according to Equation 5, determines the precision of the distribution estimation. simlandr provides a visual tool to compare the sample distributions in different stages (`check_conv()`), whereas the coda package (Plummer et al., 2006) can be used for more advanced diagnostics. The output of the `sim_SDE()` and `multi_init_simulation()` functions also used the classes from the coda package to enable easy convergence diagnosis. To achieve ergodicity in reasonable time, sometimes stronger noises need to be added to the system.

A simulation function is sufficient if the user is only interested in a single model setting. If the model is parameterized and the user want to investigate the influence of parameters on the stability of the system, then multiple simulation need to be run with different parameter settings. simlandr provides functions to perform batch simulations and store the outputs for landscape construction as one object. This can later be used to compare the stability under different parameter settings or produce animations to show how a model parameter influences the stability of the model.

In many cases, the output of the simulation is so large that it cannot be properly stored in the memory. simlandr provides a `hash_big_matrix` class, which is a modification of the `big.matrix` class from the `bigmemory` package (Kane et al., 2013), that can perform out-of-memory computation and organize the data files in the disk. In an out-of-memory computation, the majority of the data is not loaded into the memory, but only the small subset of data that is used for the current computation step. Therefore, the memory occupation is dramatically reduced. The `big.matrix` class in the `bigmemory` package provides a powerful tool for out-of memory computation. It, however, requires an explicit file name for each matrix, which can be cumbersome if there are many matrices to be handled, and this is likely to be the case in a batch simulation. The `hash_big_matrix` class automatically generates the file names using the md5 values of the matrices with the `digest` package (Eddelbuettel et al., 2021) and store it within the object. Therefore, the file links can also be restored automatically.

Step 2: Landscape Construction

Simlandr provides a set of tools to construct 2D, 3D, and 4D landscapes from single or multiple simulation results. The steady-state distribution for selected variables of the system is first estimated using the kernel density estimates (KDE). The `density` function in R is used for 2D landscapes, whereas the `ks` package (Duong, 2021) is used by default for 3D and 4D landscapes because of its higher efficiency. Then the potential function U is calculated from Equation 5. The landscape plots without a z-axis are created with `ggplot2` (Wickham, 2016), and those with a z-axis are created with `plotly` (Sievert, 2020). These plots can be further refined using

the standard `ggplot2` or `plotly` methods. See Table 2 for an overview of the family of landscape functions.

Table 2. Overview of various landscape functions provided by `simlandr`. Dimensions in bold represent the potential U calculated by the function. Dimensions in italic represent model parameters. Dimensions in parentheses are optional.

Type of Input	Function	Dimensions
Single simulation data	<code>make_2d_static()</code>	x, y
	<code>make_3d_static()</code>	(1) x, y, z+color ; (2) x, y, color
	<code>make_4d_static()</code>	x, y, z, color
Multiple simulation data	<code>make_2d_matrix()</code>	x, y, cols, (rows)
	<code>make_3d_matrix()</code>	x, y, z+color, cols, (rows)
	<code>make_3d_animation()</code>	(1) x, y, z+color, fr ; (2) x, y, color, fr ; (3) x, y, z+color, cols

Step 3: Barrier Height Calculation

An important property of the states in a landscape is their stability, which can be indicated by the barrier height that the system has to overcome when it transitions between one stable state to another adjacent state (see Cui, Lichtwarck-Aschoff, Olthof, et al., 2023, for further discussions about different stability indicators). The barrier height is also related to the escape time that the system transition from one valley to another, which can be tested empirically (Wang et al., 2008). `simlandr` provides tools to calculate the barrier heights from landscapes. These functions look for the local minima in given regions and try to find the saddle point between two local minima. The potential differences between the saddle point and local minima are calculated as barrier heights.

In 2D cases, there is only one possible path connecting two local minima. The point on the path with the highest U is identified as the saddle point. For 3D landscapes, there are multiple paths between two local minima. If we treat the system *as if* it is a gradient system with Brownian noise, then the most probable path (termed as the *minimum energy path*, MEP) that the system transitions is that it first goes along the steepest *ascent* path from the starting point, and then goes along the steepest *descent* path to the end point (E & Vanden-Eijnden, 2010). We find this path by minimizing the following action using the Dijkstra (1959) algorithm (Heymann & Vanden-Eijnden, 2008)

$$\varphi_{\text{MEP}} = \underset{\varphi}{\operatorname{argmin}} \int_A^B |\nabla U| |\mathrm{d}\varphi| \left(\approx \underset{\varphi}{\operatorname{argmin}} \sum_i |\nabla U_i| |\Delta\varphi_i| \right), \quad (6)$$

where A and B are the starting and end points and φ is the path starting at A and ending in B . After that, the point with the maximum potential value on the MEP is identified as the saddle point. Note that while the barrier height still indicates the stability of local minima, the MEP may not be the true most probable path for a nongradient system to transition between stable states.

Examples

We use two dynamical systems to illustrate the usage of the simlandr package. The first one is a two-dimensional stochastic non-gradient gene expression model, which was used by Wang et al. (2011) to represent cell development and differentiation. The second example is a dynamic model of panic disorder Robinaugh et al. (2024) which contains much more variables and parameters, non-Markovian property, and non-differentiable formulas. We mainly use the first example to show the agreement of the results from simlandr with previous analytic results, and the second example as a typical use case of a complex dynamic model which is not treatable with other methods (also see Cui, Lichtwarck-Aschoff, Olthof, et al., 2023, for more substantive discussions). Note that, either system includes more than two variables, making it impossible to perform the landscape analysis with other available R packages.

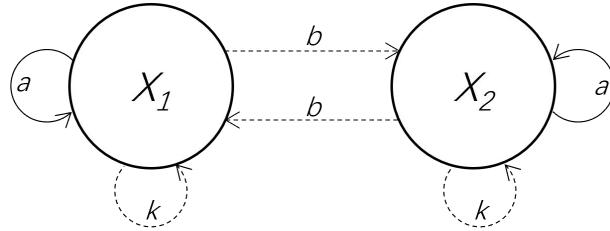
Example 1: The Gene Expression Model

This model is built on the mutual regulations of the expressions of two genes, in which X_1 and X_2 represent the expression levels of two genes which activate themselves and inhibit each other. A graphical illustration is shown in Figure 3 (adapted from Wang et al., 2011). Their dynamic functions can be written as

$$\begin{aligned} \frac{dX_1}{dt} &= \frac{ax_1^n}{S^n + x_1^n} + \frac{bS^n}{S^n + x_2^n} - kx_1 + \sigma_1 \frac{dW_1}{dt}, \\ \frac{dX_2}{dt} &= \frac{ax_2^n}{S^n + x_2^n} + \frac{bS^n}{S^n + x_1^n} - kx_2 + \sigma_2 \frac{dW_2}{dt}, \\ \frac{da}{dt} &= -\lambda a + \sigma_3 \frac{dW_3}{dt}, \end{aligned} \quad (7)$$

where a represents the strength of self-activation, b represents the strength of mutual-inhibition, and k represents the speed of degradation. The development of an organism is modeled as a decreasing at a certain speed λ . In the beginning, there is only one possible state for the cell. After a certain milestone, the cell differentiates into one of the two possible states.

Figure 3. A graphical illustration of the relationship between the activation levels of the two genes. Solid arrows represent positive relationships (i.e., activation) and dashed arrows represent negative relationships (i.e., inhibition).



This model can be simulated using the `sim_SDE()` function in `simlandr`, with the default parameter setting $b = 1, k = 1, S = 0.5, n = 4, \lambda = 0.01$, and $\sigma_1 = \sigma_2 = \sigma_3 = 0.2$.

```
# Load the package.
library(simlandr)

# Specify the simulation function.
b <- 1
k <- 1
S <- 0.5
n <- 4
lambda <- 0.01

drift_gene <- c(rlang::expr(z * x^(!!n)/((!!S)^ (!!n) + x^(!!n)) + (!!b) *
  (!!S)^ (!!n)/((!!S)^ (!!n) + y^(!!n)) - (!!k) * x), rlang::expr(z * y^(!!n)/((!!S)
^ (!!n) +
  y^(!!n)) + (!!b) * (!!S)^ (!!n)/((!!S)^ (!!n) + x^(!!n)) - (!!k) * y),
  rlang::expr(-(!!lambda) * z)) |>
  as.expression()

diffusion_gene <- expression(0.2, 0.2, 0.2)
# Perform a simulation and save the output.
set.seed(1614)
single_output_gene <- sim_SDE(drift = drift_gene, diffusion = diffusion_gene,
  N = 1e+06, M = 10, Dt = 0.1, x0 = c(0, 0, 1), keep_full = FALSE)
```

After the simulation, we perform some basic data wrangling to produce a dataset that can be used for further analysis. We create a new variable `delta_x` as the difference between `X1` (`X`) and `X2` (`Y`), and we rename the variable `Z` as `a`.

```
single_output_gene2 <- do.call(rbind, single_output_gene)
single_output_gene2 <- cbind(single_output_gene2[, "X"] - single_output_gene2[,
```

```
"Y"], single_output_gene2[, "Z"])
colnames(single_output_gene2) <- c("delta_x", "a")
```

We then perform the convergence check on the simulation result. First, we convert the simulation output to the format for the coda package, and thin the output to speed up the convergence check.

```
single_output_gene_mcmc_thin <- as.mcmc.list(lapply(single_output_gene,
  function(x) x[seq(1, nrow(x), by = 100), ]))
```

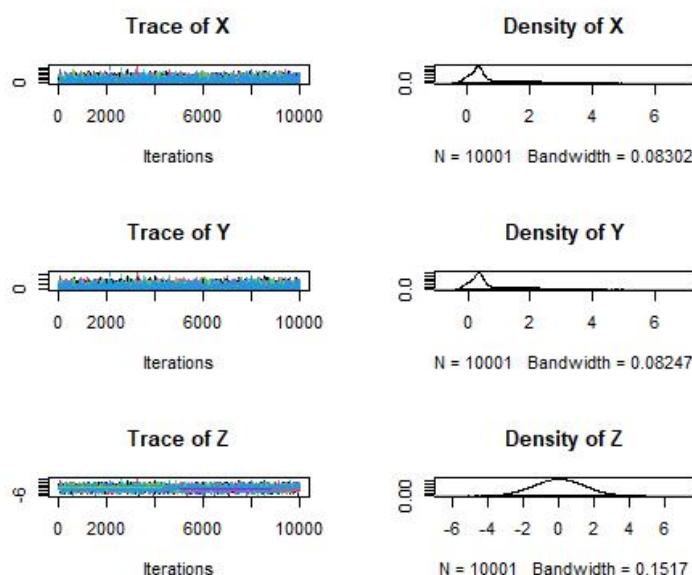
We then shown the convergence diagnosis plot to check the convergence of the simulation in Figure 4. The distribution of the two key variables in different simulation stages are converging, indicating that the simulation is long enough to provide reliable estimation of the steady-state distribution. Other convergence checks can also be readily performed using the coda package.

```
plot(single_output_gene_mcmc_thin)
```

We generated the 3D landscape for this model with `make_3d_single()`. Here, we use `x, y` to specify the variables of interest, and use `lims` to specify the limits of the `x` and `y` axes for the landscape. The `lims` argument can be left blank, then the limits will be automatically calculated.

```
l_single_gene_3d <- make_3d_single(single_output_gene2, x = "delta_x",
  y = "a", lims = c(-1.5, 1.5, 0, 1.5), Umax = 8)
```

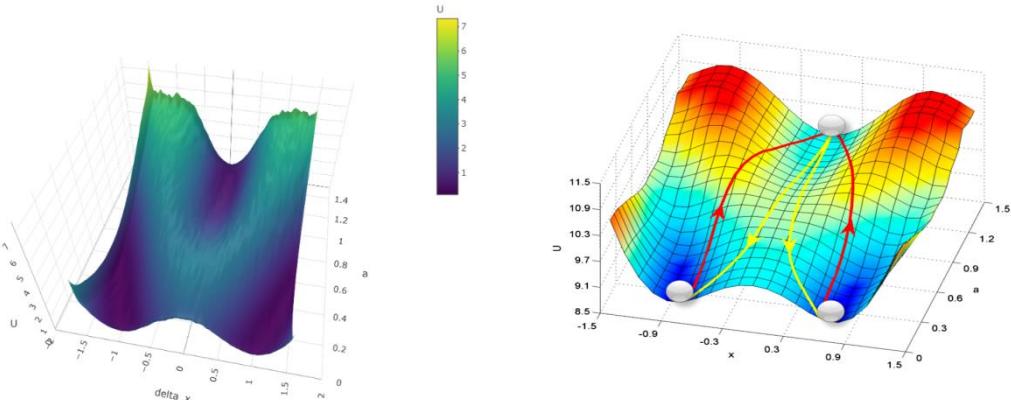
Figure 4. The convergence check result for the simulation of the gene expression model. The variables in different simulation stages did not show distributional differences, indicating that the simulation is long enough to provide a reliable estimation of the steady-state distribution.



The resulted landscape is shown in the left panel of Figure 5. In this plot, the x-axis represents $\Delta x (= x_1 - x_2)$, and the y-axis represents a . To compare with, the potential landscape obtained analytically by Wang et al. (2011) is shown in the right panel of Figure 5. The result of simlandr appears to be very close to the result based on the analytical derivation. Note that because different normalization methods were used, the U values of the two landscapes are not directly comparable. Here, we are mainly interested in their relative shape.

```
plot(l_single_non_grad_3d)
```

Figure 5. The 3D landscape (potential value as z-axis) for the gene expression model. The left panel is the plot produced by simlandr; the right panel is the potential landscape obtained analytically by Wang et al. (2008), reproduced with the permission of the authors and in accordance with the journal policy.



We then calculate the barrier for the landscape using `calculate_barrier()`. The barrier is calculated by specifying the start and end locations, and the radii of the start and end locations. The height of the barrier from two sides can be calculated with `get_barrier_height()`.

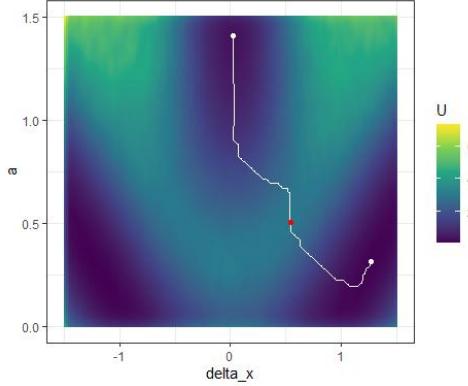
```
b_single_gene_3d <- calculate_barrier(l_single_gene_3d, start_location_value = c(0,
  1.2), end_location_value = c(1, 0.2), start_r = 0.3, end_r = 0.3)

get_barrier_height(b_single_gene_3d)
## delta_U_start    delta_U_end
##      2.572507     2.829729
```

The local minima, the saddle point, and the MEP can be added to the landscape with `autolayer()`, shown in Figure 6.

```
plot(l_single_gene_3d, 2) + autolayer(b_single_gene_3d)
```

Figure 6. The landscape for the gene expression model. The local minima are marked as white dots, the saddle points are marked as red dots, and the MEPs are marked as white lines.



Next, we use multiple simulations to investigate the influence of two parameters, k and b , on the stability of the system. As explained above, the parameter b represents the strength of mutual-inhibition between the two genes. Therefore, as b increases, we expect the differentiation is more extreme, that is, the cell is more likely to develop into one of the two cell types with very different gene expression levels. The valleys in the landscape representing the two types will become further apart and the barrier becomes clearer. The parameter k represents the speed of degradation of the gene products. As k increases, the gene products degrade faster, and this effect is more pronounced when the gene products are at high levels. Therefore, the dominant gene will express in a less extreme level, and we expect that the two valleys become closer, and the barrier will become less clear as k increases.

We use the batch simulation functions of simlandr. First, we create the argument set for the batch simulation. This specifies the parameters to be varied. We examined three b values, 0.5, 1, 1.5, and three k values, 0.5, 1, 1.5, which form 9 possible combinations.

```
batch_arg_set_gene <- new_arg_set()
batch_arg_set_gene <- batch_arg_set_gene |>
  add_arg_ele(arg_name = "parameter", ele_name = "b", start = 0.5, end = 1.5,
              by = 0.5) |>
  add_arg_ele(arg_name = "parameter", ele_name = "k", start = 0.5, end = 1.5,
              by = 0.5)
batch_grid_gene <- make_arg_grid(batch_arg_set_gene)
```

We then perform the batch simulation with the `batch_simulation()` function. Here, we specify the simulation function to be used, which is similar to single simulation we showed above, together with the data wrangling procedure. The simulation function is defined with

the `sim_fun` argument. We also use `bigmemory = TRUE` to store the simulation results in the `hash_big_matrix` format, which is more memory-efficient.

```
batch_output_gene <- batch_simulation(batch_grid_gene, sim_fun = function(parameter)
{
  b <- parameter[["b"]]
  k <- parameter[["k"]]
  drift_gene <- c(rlang::expr(z * x^(!!n)/ (!!S)^(!!n) + x^(!!n)) + (!!b) *
    (!!S)^(!!n)/ (!!S)^(!!n) + y^(!!n)) - (!!k) * x, rlang::expr(z *
    y^(!!n)/ (!!S)^(!!n) + y^(!!n)) + (!!b) * (!!S)^(!!n)/ (!!S)^(!!n) +
    x^(!!n)) - (!!k) * y), rlang::expr(- (!!lambda) * z)) |>
  as.expression()
  set.seed(1614)
  single_output_gene <- sim_SDE(drift = drift_gene, diffusion = diffusion_gene,
    N = 1e+06, M = 10, Dt = 0.1, x0 = c(0, 0, 1), keep_full = FALSE)
  single_output_gene2 <- do.call(rbind, single_output_gene)
  single_output_gene2 <- cbind(single_output_gene2[, "X"] - single_output_gene2[, "Y"],
    single_output_gene2[, "Z"])
  colnames(single_output_gene2) <- c("delta_x", "a")
  single_output_gene2
}, bigmemory = TRUE)
```

If the output is saved in an RDS file, upon next use, it can be read as follows.

```
saveRDS(batch_output_gene, "batch_output_gene.RDS")
batch_output_gene <- readRDS("batch_output_gene.RDS") |>
  attach_all_matrices()
```

We then make the 3D matrix for the batch output, using `make_3d_matrix()`.

```
l_batch_gene_3d <- make_3d_matrix(batch_output_gene, x = "delta_x", y = "a",
  cols = "b", rows = "k", lims = c(-5, 5, -0.5, 2), h = 0.005, Umax = 8,
  kde_fun = "ks", individual_landscape = TRUE)
```

For the barrier calculation step, the start and end points of the barrier may be different for each landscape plot. The following code shows how to create a barrier grid for each landscape plot. First, we create a barrier grid template using the function `make_barrier_grid_3d()`. Next, we modify the barrier grid template to create a barrier grid for the landscape plot.

```
make_barrier_grid_3d(batch_grid_gene, start_location_value = c(0, 1.5),
  end_location_value = c(1, -0.5), start_r = 1, end_r = 1, print_template = TRUE)
```

```

## structure(list(start_location_value = list(c(0, 1.5), c(0, 1.5),
## ), c(0, 1.5), c(0, 1.5), c(0, 1.5), c(0, 1.5), c(0, 1.5), c(0,
## 1.5), c(0, 1.5)), start_r = list(c(1, 1), c(1, 1), c(1, 1), c(1,
## 1), c(1, 1), c(1, 1), c(1, 1), c(1, 1)), end_location_value = list(
##     c(1, -0.5), c(1, -0.5), c(1, -0.5), c(1, -0.5), c(1, -0.5),
##     c(1, -0.5), c(1, -0.5), c(1, -0.5)), end_r = list(
##     c(1, 1), c(1, 1), c(1, 1), c(1, 1), c(1, 1), c(1, 1), c(1,
##     1), c(1, 1))), row.names = c(NA, -9L), class = c("arg_grid",
## "data.frame"))

##   ele_list b k start_location_value start_r end_location_value end_r
## 1 0.5, 0.5 0.5 0.5      0.0, 1.5    1, 1       1.0, -0.5 1, 1
## 2 1.0, 0.5 1.0 0.5      0.0, 1.5    1, 1       1.0, -0.5 1, 1
## 3 1.5, 0.5 1.5 0.5      0.0, 1.5    1, 1       1.0, -0.5 1, 1
## 4 0.5, 1.0 0.5 1.0      0.0, 1.5    1, 1       1.0, -0.5 1, 1
## 5 1, 1 1.0 1.0      0.0, 1.5    1, 1       1.0, -0.5 1, 1
## 6 1.5, 1.0 1.5 1.0      0.0, 1.5    1, 1       1.0, -0.5 1, 1
## 7 0.5, 1.5 0.5 1.5      0.0, 1.5    1, 1       1.0, -0.5 1, 1
## 8 1.0, 1.5 1.0 1.5      0.0, 1.5    1, 1       1.0, -0.5 1, 1
## 9 1.5, 1.5 1.5 1.5      0.0, 1.5    1, 1       1.0, -0.5 1, 1

bg_gene <- make_barrier_grid_3d(batch_grid_gene, df = structure(list(start_location_value = list(c(0,
1.5), c(0, 1.5), c(0, 1.5), c(0, 1.5), c(0, 1.5), c(0, 1.5),
c(0, 1.5), c(0, 1.5)), start_r = list(c(0.2, 1), c(0.2, 1),
c(0.2, 1), c(0.2, 0.5), c(0.2, 0.5), c(0.2, 0.3), c(0.2, 0.3),
c(0.2, 0.3)), end_location_value = list(c(2, 0), c(2, 0),
c(2, 0), c(1, 0), c(1, 0), c(1, 0), c(1, 0), c(1, 0)), end_r = list(c(1,
1), c(1, 1), c(1, 1), c(1, 1), c(1, 1), c(1, 1), c(1, 1),
c(1, 1))), row.names = c(NA, -9L), class = c("arg_grid", "data.frame")))

```

3

With the barrier grid template, we can calculate the barrier for each landscape plot.

```
b_batch_gene_3d <- calculate_barrier(l_batch_gene_3d, bg = bg_gene)
```

If the barrier grid was not needed, the following code can be used to calculate the barrier.

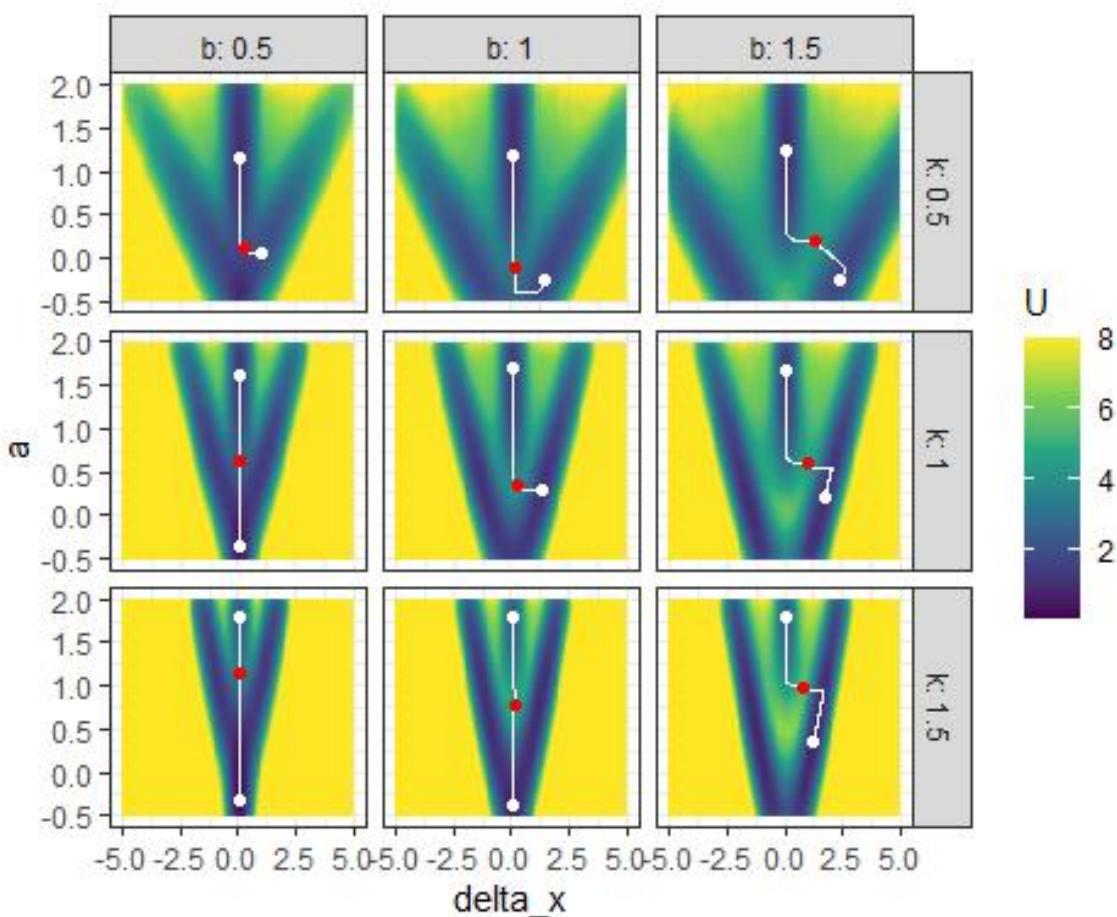
```
b_batch_gene_3d <- calculate_barrier(l_batch_gene_3d, start_location_value = c(0,
1.5), end_location_value = c(1, 0), start_r = 1, end_r = 1)
```

The resulted landscapes and the MEPs between states are shown in Figure 7.

```
plot(l_batch_gene_3d) + autolayer(b_batch_gene_3d)
```

From the landscapes, it is clear that increasing b , which represents higher strength of gene mutual prohibition, makes the two differentiated states further apart from each other. Increasing k , which represents faster degradation, makes the undifferentiated state disappear earlier, thus makes the differentiation earlier. When b is low enough and k is high enough, there is no differentiation anymore because the two differentiated states merge together and form a more stable state at $\Delta x = 0$. In this case, there is no actual differentiation in the system, but only a one-to-one conversion of cell types. Only when b is high enough and k is low enough is it possible for the cell to differentiate into two types.

Figure 7. The landscape for the gene expression model of different b and k values. The local minima are marked as white dots, the saddle points are marked as red dots, and the MEPs are marked as white lines.

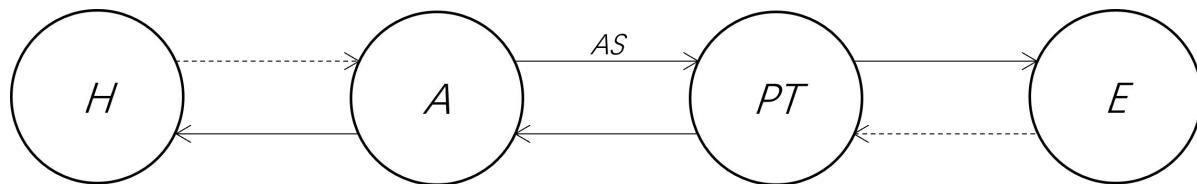


Example 2: Panic Disorder Model

The second example we use is the panic disorder model proposed by Robinaugh et al., (2024). The model is implemented in the PanicModel package (<https://github.com/jmbh/PanicModel>). It contains 12 variables and 33 parameters and also involves history dependency and non-diffe

rentiable formula (such as if-else conditions) to model the complex interplay of individual and environmental elements in different time scales. The most important variables of the model are the physical arousal (A) of a person (e.g., heart beat, muscle tension, sweating), the person's perceived threat (PT , how dangerous the person cognitively evaluate the environment), and the person's tendency to escape from the situation (E). The core theoretical idea of the model is that physical arousal and perceived threat of a person may strengthen each other in certain circumstances, leading to sudden increases in both variables, manifesting as panic attacks. The tendency that a person tends to use physical arousal as cognitive evidence of threat is represented by another variable, arousal schema AS . A comprehensive introduction of the model is beyond the scope of the current article, and we would like to refer interested readers to Robinaugh et al. (2024). Here, to simplify the context, we assume that AS does not change over time, and no psychotherapy is being administered. We focus on the influence of AS on the system's stability represented by A and PT . A graphical illustration of several core variables of this model is shown in Figure 8 (adapted from Cui, Lichtwarck-Aschoff, Olthof, et al., 2023).

Figure 8. A graphical illustration of the relationships between several important psychological variables in the panic disorder model. Solid arrows represent positive relationships and dashed arrows represent negative relationships.



To construct the potential landscapes for this model, we first create a function that performs a simulation using the `simPanic()` function from `PanicModel`. This is required as we need to modify some default options for illustration.

```

library(PanicModel)

sim_fun_panic <- function(x0, par) {

  # Change several default parameters
  pars <- pars_default

  # Increase the noise strength to improve sampling efficiency
  pars$lambda_N <- 200

  # Make S constant through the simulation
  pars$TS$r_S_a <- 0
}
  
```

```

pars$TS$r_S_e <- 0

# Specify the initial values of A and PT according to the format
# requirement by `multi_init_simulation()`, while the other
# variables use the default initial values.
initial <- initial_default
initial$A <- x0[1]
initial$PT <- x0[2]

# Specify the value of S according to the format requirement by
# `batch_simulation()`.

initial$S <- par$S

# Extract the simulation output from the result by simPanic().
# Only keep the core variables.
return(as.matrix(simPanic(1:5000, initial = initial, parameters = pars)$outmat[, 
  c("A", "PT", "E")]))
}

```

We then perform a single simulation from multiple starting points. To speed up the simulation, we use parallel computing.

```

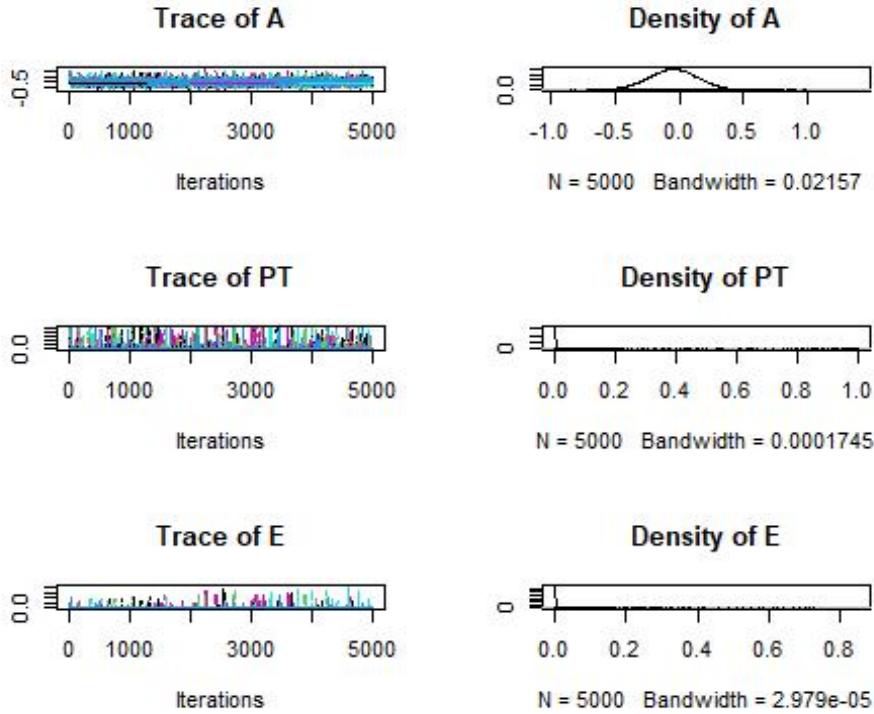
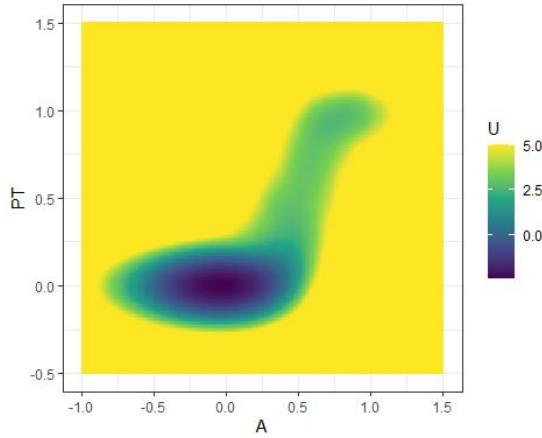
future::plan("multisession")
set.seed(1614, kind = "L'Ecuyer-CMRG")
single_output_panic <- multi_init_simulation(sim_fun = sim_fun_panic, range_x0 = c
(0, 1, 0, 1), R = 4, par = list(S = 0.5))

```

The convergence check results of the simulation, shown in Figure 9, indicate that the time series of the first 100 data points are strongly influenced by the choice of initial value. Therefore, we remove the first 100 data points in the following analysis.

```
plot(single_output_panic)
```

We then create the 3D landscape for the panic disorder model, shown in Figure 10. The landscape shows that the system has two stable states, which are represented by the valleys in the landscape. The system can be trapped in these valleys, leading to different levels of physical arousal and perceived threat. The valley with higher potential value represents a state with higher physical arousal and perceived threat, which corresponds to a panic attack. In contrast, the valley with lower potential value represents a state with lower physical arousal and perceived threat, which corresponds to a healthy state.

Figure 9. The convergence check result for the simulation of the panic disorder model.**Figure 10.** The 3D landscape (potential value as color) for the panic disorder model

We now investigate the effect of the parameter S on the potential landscape. This parameter represents the tendency that a person consider physical arousal as a sign of danger. Therefore, we expect that higher S will stabilize the panic state and destabilize the healthy state.

We perform a batch simulation with varying S values to construct the potential landscapes for different S values. This, again, starts with the creation of a grid of parameter values.

```
batch_arg_grid_panic <- new_arg_set() |>
  add_arg_ele(arg_name = "par", ele_name = "S", start = 0, end = 1, by = 0.5) |>
  make_arg_grid()
```

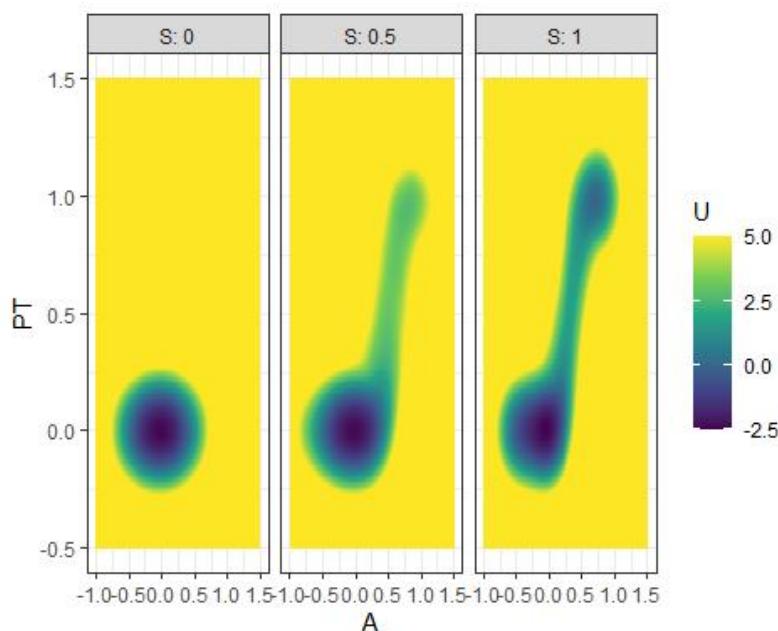
We then perform the batch simulation using parallel computing.

```
future::plan("multisession")
set.seed(1614, kind = "L'Ecuyer-CMRG")
batch_output_panic <- batch_simulation(batch_arg_grid_panic, sim_fun = function(par)
{
  multi_init_simulation(sim_fun_panic, range_x0 = c(0, 1, 0, 1), R = 4,
    par = par) |>
  window(start = 100)
})
```

The 3D landscapes for different S values are shown in Figure 11. The landscapes show that the system only has one stable state when S is low, but has two stable states when S is high. The stability of the panic state also increases when S is higher. This indicates that a higher S value corresponds to a higher risk of panic attacks.

```
l_batch_panic_3d <- make_3d_matrix(batch_output_panic, x = "A", y = "PT",
  cols = "S", h = 0.005, lims = c(-1, 1.5, -0.5, 1.5))
plot(l_batch_panic_3d)
```

Figure 11. The landscape for the panic disorder model of different S values. Two landscapes are shown for different variable combinations, A and PT, or A and E.



Discussion

Potential landscapes can show the stability of states for a dynamical system in an intuitive and quantitative way. They are especially informative for multistable systems. In this article, we illustrated how to construct potential landscapes using simlandr. The potential landscapes generated by simlandr are based on the steady-state distribution of the system, which is in turn estimated using Monte Carlo simulation. Compared to analytic methods, Monte Carlo estimation is more flexible and thus more applicable for complex models. The flexibility comes together with a higher demand for time and storage, which is necessary to make the estimation precise enough. The `hash_big_matrix` class partly solved this problem by dumping the memory storage to hard disk space. Also, it is important that the simulation function itself is efficient enough. The functions `sim_SDE()` and `multi_init_simulation()` make use of the efficient simulations provided by Sim.DiffProc (Guidoum & Boukhetala, 2020) and the parallel computing with the future framework (Bengtsson, 2021). For customized simulation functions, there are also multiple approaches that can be used to improve the performance, for which we refer interested readers to Wickham (2019). In Online Supplementary Materials A (available at <https://osf.io/9kgx7/>), we provide a benchmark of the typical time and memory usage of the procedures in simlandr. From there, we can see that time and memory usage are acceptable in most cases on a personal computer. When the transition between attractors is rare, the `multi_init_simulation()` function may help to speed up the convergence, and more advanced sampling methods like importance sampling or rare event sampling may be needed in more complex situations. The detailed ways to implement such methods are highly dependent on the specific model and are beyond the scope of this package. We direct interested readers to Rubino and Tuffin (2009) and Kloek and van Dijk (1978) for a comprehensive review of rare event simulation methods. Nevertheless, the landscape construction functions in simlandr allow users to provide weights for the simulation results, which can be used to adjust the sampling distribution.

In addition, the length of the simulation and the choice of noise strength may also have an important influence on the results. If the length is too short, the density estimation will be inaccurate, resulting in rugged landscapes. If the length is too long, the simulation part would require more computational resources, which is not always realistic. If the noise is too weak, the system may not be able to converge in a reasonable time, resulting in problems in convergence checks, overly noisy landscapes, or failure to show valleys that are theoretically present. If the noise is too strong, the simulation may be unstable and the boundaries between valleys may be blurry. In Online Supplementary Materials B (available at <https://osf.io/9kgx7/>),

we showed the influence of simulation length and noise strength on the landscape output. With some theoretical expectation of the system's behavior, it is not difficult to spot that the simulation is too short, or the noise level might be unsuitable. In that case, some adjustments are required before the landscape can be well constructed.

All landscape construction and barrier calculation functions in simlandr contain both visual aids and numerical data that can be used for further processing. The html plots based on plotly are more suitable for interactive illustrations, while it is also possible to export them to static plots using `plotly::orca()`. The `ggplot2` plots are readily usable for flat printing.

We also want to note some limitations of the potential landscape generated by simlandr. First, the generalized potential landscape is not a complete description of all dynamics in a system. It emphasizes the stability of different states by filtering out other dynamical information. Some behaviors are not possible in gradient systems (e.g., oscillations and loops), thus cannot be shown in a potential landscape (Zhou & Li, 2016). Second, since the steady-state distribution is estimated using a kernel smoothing method, which depends on stochastic simulations, the resulting potential function may not be highly accurate. Its accuracy is further affected by the choice of kernel bandwidth and noise strength. This issue is particularly pronounced at valley edges, where fewer samples are available for estimation. Similar limitations apply to MEP calculations, as they are derived from the generalized landscape rather than the original dynamics. Therefore, we do not recommend directly interpreting the potential function or barrier height results for applications requiring high precision. Instead, the potential landscape is best used as a semi-quantitative tool to gain insights into the system's overall behavior, guide further analysis, and compare system behavior under different parameter settings, provided the same simulation and kernel estimation conditions are used. The examples in this article illustrated some typical use cases we recommend.

Availability and Future Directions

This package is publicly available from the Comprehensive R Archive Network (CRAN) at <https://CRAN.R-project.org/package=simlandr>, under GPL-3 license. The results in the current article were generated with simlandr 0.4.0 version. R script to replicate all the results in this article can be found at <https://osf.io/9kgx7/>.

The barrier height data calculated by simlandr can also be further analyzed and visualized. For example, sometimes it is helpful to look into how the barrier height changes with varying parameters (e.g., Cui et al. (2023)). We encourage users to explore other ways of analyzing and visualizing the various results provided by simlandr.

The method we chose for simlandr is not the only possible one. The generalized landscape by Wang et al. (2008), which we implemented, is more flexible and emphasizes the possibility that the system is in a specific state, while other methods may have other strengths (e.g., the method by Rodríguez-Sánchez et al., 2020, emphasizes the gradient part of the vector field, and the method by Moore et al., 2016, emphasizes the possibility of transition processes under small noise). We look forward to future theoretical and methodological developments in this direction.

Chapter 4

Unlocking Nonlinear Dynamics and Multistability from Intensive Longitudinal Data: A Novel Method

This chapter is based on:

Cui, J., Hasselman, F., & Lichtwarck-Aschoff, A. (2023). Unlocking nonlinear dynamics and multistability from intensive longitudinal data: A novel method. *Psychological Methods*. <https://doi.org/10.1037/met0000623>



Abstract

The availability of smart devices has made it possible to collect intensive longitudinal data (ILD) from individuals, providing a unique opportunity to study the complex dynamics of psychological systems. Existing time-series methods often have limitations, such as assuming linear interactions or having restricted forms, leading to difficulties in capturing the complex nature of these systems. To address this issue, we introduce fitlandr, a method with implementation as an R package that integrates nonparametric estimation of the drift-diffusion function and stability landscape. The drift-diffusion function is estimated using the Multivariate Kernel Estimator (MVKE, Bandi & Moloche, 2018), and the stability landscape is estimated through Monte-Carlo estimation of the steady-state distribution (Cui et al., 2021, 2022). Using a simulated emotional system, we demonstrate that fitlandr can effectively recover bistable dynamics from data, even in the presence of moderate noise, and that it primarily relies on dynamic information from the system instead of distributional information. We then apply the method to two empirical single-participant ESM datasets and compared the results with the simulation datasets. Whereas both datasets show a bimodal distribution, fitlandr only revealed bistability in one of them, indicating that bimodality in ILD does not necessarily imply the existence of bistability in the underlying system. These results demonstrate the potential of fitlandr as a tool for uncovering the rich, nonlinear dynamics of psychological systems from ILD.

Introduction

Psychologists have a long-standing interest in understanding changes in mental states over time. With the rise of smart devices, intensive longitudinal data (ILD) from individuals are now readily available, providing new opportunities to study the evolution of psychological systems (Hamaker et al., 2015; Myin-Germeys & Kuppens, 2021; Trull & Ebner-Priemer, 2009). The new type of data not only brings about new information but also bears challenges that call for innovative analytical methods. Psychological systems, like other complex systems in nature, consist of a large number of elements interacting continuously and nonlinearly (Olthof et al., 2020, 2023). As a result, bimodality and multimodality, which means that the data distribution has two or more modes instead of only one mode in a Gaussian distribution, are the rule rather than the exception (Delignières et al., 2004; Haslbeck et al., 2023; Haslbeck & Ryan, 2022). Also, change processes are characterized by sudden changes or regime shifts, indicating the nonstationarity of the data (Helmich et al., 2020; Olthof et al., 2020). Most existing models assume constant linear interactions of variables discretely and are therefore less suited to handle data that show multimodality and nonstationarity.

Additionally, previous methods primarily focus on the *dynamics* of a complex psychological system (i.e., indicating the most probable direction of the system at the next time point based on its current state.) But next to these dynamics, it is also essential to gain a deeper understanding of the system's *stability*. The stability of a system describes the number of qualitatively different phases in the system, the range of these phases, and in which phase the system is most likely to reside. Understanding the stability of psychological systems has important theoretical and practical implications. For example, there is a growing interest in sudden changes in experienced severity of symptoms (Cramer et al., 2016; Cui et al., 2023; Haslbeck et al., 2022; Olthof et al., 2023; Wichers et al., 2019). This line of work conceptualizes mental disorders as a stable emergent phase¹ of a person-specific mental system embedded in a specific context. The ball-and-landscape metaphor is often used to illustrate this concept, where the state of a person's mental system is like a ball on a landscape that fluctuates around a local minimum but may transition to another minimum under certain conditions, resulting in a change to a different phase (e.g., from the healthy phase to the depressive phase or the anxious phase). In this metaphor, the stability of the system is represented by the altitude of the landscape. A higher position indicates less stability and a tendency to fall, and a lower

¹ The term “attractor” is sometimes employed to convey the same idea as the “phase” we used here. However, “attractor” can be used to refer to a particular point at which the system is most stable, akin to a local minimum on the landscape rather than a basin of attraction encompassing the local minimum (Sayama, 2015). To prevent any potential confusion, we adhere to the term “phase” in this article, in line with our previous work (Cui et al., 2023).

position indicates more stability and a tendency to remain in the same area. If multiple basins exist on the landscape, indicating there are multiple phases that the system can reside in, the system can be characterized as exhibiting multistability. It is important to distinguish multistability from multimodality: multimodality is only a feature of the data distribution, whereas multistability signifies the existence of distinct phases that the system tends to *stay* in a short time, but may occasionally transition between them. Although the multimodality of data may suggest the presence of multistability in the system, it is not always the case, as we will elaborate in subsequent sections. Whereas the ball-and-landscape metaphor is conceptually well recognized and accepted, quantifying the potential landscape function remains a challenge. Although some recent studies have proposed methods to estimate the stability landscapes for psychological formal models (Cui et al., 2023), there are currently no methods available to estimate stability landscapes from multivariate continuous data, limiting further progress in this field.

In the current article, we propose a novel nonparametrically method to estimate the dynamic functions and the generalized potential function of psychological systems from ILD. The dynamics can be represented with vector fields, and the generalized potential function can be visualized by potential landscapes. By this method, we aim to provide a quantitative and rather assumption-free description of the dynamics and the stability of psychological systems. An R package, *fitlandr*, was developed to provide an implementation of the method and can be accessed at the Comprehensive R Archive Network (<https://CRAN.R-project.org/package=fitlandr>). The name *fitlandr* will be used throughout the current article for the whole method that we are proposing. In the following sections, we will first introduce the currently available methods for psychological ILD, and describe the similarities and differences with our method. After that, we describe our method and briefly introduce related algorithms. We will then demonstrate the method using simulation data and empirical data. With the simulation data, we examine if our method can recover known multistability from short time series, how much it is robust under noise, and how much it relies on the dynamical information (i.e., the information regarding how the current state of the system influences its state at the subsequent time point) from the data rather than the distributional information. With the empirical data, we examine if our method can recover multistability suggested by psychological theories and whether it provides insights regarding the nature of psychological systems. Finally, we discuss the advantages, disadvantages, and potential applications of this method. The necessary code for replicating the findings presented in this article can be obtained from the following URL: <https://osf.io/s7dq4/>.

A Survey of Current Methods and Models

In this section, we present a survey of commonly used methods and models for analyzing psychological ILD. We summarize their limitations in capturing complex dynamics and stability, and highlight the added value of the method we are proposing.

The vector autoregression (VAR) model is the most widely used approach for multivariate time series analysis and serves as a foundation for longitudinal network analysis (Bringmann et al., 2013). In a VAR model, the state of the system from the previous time point, represented as a vector, is used to predict the state of the system at the next time point through a regression model. Despite its simplicity, VAR provides a powerful way to gain insights from ILD. Yet, VAR models assume constant linear interactions of the variables discretely, thus fail to capture complex features such as nonstationarity and multistability (Haslbeck et al., 2022; Haslbeck & Ryan, 2022; Olthof et al., 2020).

To overcome this issue, many variants of the original VAR model have been developed, seeking to capture more complex dynamics in the system. For example, the time-varying VAR (TV-VAR) model allows for smooth change in the parameters of the VAR model, accounting for nonstationary behavior in the data (Bringmann et al., 2017; Haslbeck et al., 2021). Threshold-VAR (De Haan-Rietdijk et al., 2016; Tong & Lim, 1980) and Markov-Switching-VAR (Hamilton, 1989; Haslbeck & Ryan, 2022) are models that apply different regression equations based on a threshold variable or a latent variable that indicates which phase the system is in. Although these variants improve upon the VAR model, they do not explicitly model nonlinear dynamics. More concretely, even though these models allow the linear interaction coefficient to change nonlinearly over time, they still do not permit the function form to be nonlinear (i.e., the influence of any variable on itself or other variables is always linear). As a result, the system at each time point is a linear system, which means that at any given moment, when a higher x corresponds to a higher y , it follows that an even greater x will consistently result in a proportionally higher y , without any exceptions. This limits the ability to capture nonlinear interactions among variables. Moreover, they represent nonstationary and multistable behaviors with one or more variables (such as the time-varying coefficients in the TV-VAR model, one or more assessed variables in the Threshold-VAR model, and latent states in the Markov-Switching-VAR model). These variables can undergo discrete or smooth changes, making the model undergo transformations, but they are not explanans per se. As a result, the model does not inherently account for the reasons behind their changes. Last, if the transition between different phases of the system is the primary focus rather than the specific dynamics, a hidden Markov model (HMM) can also be used, which assumes that the data come from

multiple probability distributions (Haslbeck & Ryan, 2022; Visser, 2011). However, in HMM, only the specific transition probabilities between hidden states are defined, whereas within each hidden state, the time points are considered independent from each other, lacking any temporal dependencies. Consequently, HMM primarily focuses on capturing the transition possibilities between hidden states rather than providing a detailed description of the specific dynamics.

Another branch of methods for analyzing ILD is based on the drift-diffusion process, in which the moving trajectory of a system is governed by two types of forces: the drift part, which represents the deterministic forces, and the diffusion part, which represents the stochastic forces. The drift-diffusion process is modeled using stochastic differential equations instead of linear regressions, thus the models based on the drift-diffusion process are inherently continuous. The Fokker-Planck Equation Model (FPEM) by Tschacher and Haken (2020) is an example of a simple drift-diffusion model that uses the binned average of the differences between sample points as the drift term and the variance as the diffusion term. The method calculates a potential landscape function by integrating the drift term. An advantage of this model is that the local dynamics and the global multistability are uniformly represented in one dynamical equation. The multimodality is the property of the system itself, not the consequence of switching between different systems. This method, however, requires enough data points in each bin to estimate the mean and the variance reliably, and the integral can only be calculated in a single dimension, thus the method is only applicable for one-dimensional systems. Another method based on the drift-diffusion process is the Ornstein-Uhlenbeck Model (OUM, also known as the DynAffect model, Kuppens et al., 2010; Oravecz et al., 2011). This model assumes that the drift is proportional to the distance between the current state of the system and a single point (the “homebase” of the system). The direction of the drift term always points towards this point. The diffusion term is set as a multidimensional Gaussian noise. This method can be applied to two or more dimensional systems, but it does not allow multistability due to its linear form in the drift part. Another related method is the Continuous-Time VAR (CT-VAR) model (Ryan et al., 2018; Ryan & Hamaker, 2022), which is a continuous-time extension to VAR models. It assumes linear continuous interactions among variables, thus the dynamics can also be described by a drift-diffusion process similar to the OUM, whereas the CT-VAR model is intended to be used for datasets with more variables and can be represented in a network. However, like the OUM, it assumes linear dynamics, thus does not show multistability. To explain the typical skewed and bimodal distribution in emotional ILD, Loossens et al. (2020) proposed the Affective Ising Model (AIM), which

assumes a specific type of the drift function inspired by the Ising model in physics. This model assumes two subsystems: the positive emotion system and the negative emotion system. Each element strengthens other elements in the same system, and inhibits the elements in the other system. The AIM can well represent a specific type of data distribution, namely the V-shaped distribution in the state space of positive and negative emotions. The model, however, is limited to a specific set of assumptions about positive and negative emotion systems and does not generalize to the ILD of other psychological processes.

We summarize the key features of the above-mentioned methods in Table 1. Inspecting the table, we can see a clear trade-off between the flexibility of the model and the feasibility of the model estimation. Taking the representation of multistability as an example: models that specify the number of phases beforehand (such as Threshold-VAR, Markov-Switching-VAR, HMM, and AIM) can handle more variables and require fewer data points, whereas models that let the number of phases emerge from the estimation process (such as FPEM) require more data points and can handle fewer variables. This mostly comes from the methodological challenges posed by psychological ILD (van de Leemput et al., 2014): Psychological ILD often include a few dozens to one hundred data points from self-reported questionnaires. Although this length is large in the field of psychology, it falls short in comparison to other fields that typically have thousands of data points, and data collected through questionnaires are often more prone to measurement noise. For example, the animal movement data in ecology used by Brillinger (2007) has 1,571 data points measured by GPS with a precision in the meter level and range of kilometers, and the long-term geographical data used by Livina et al. (2010) has 3,000 data points of isotope concentration which could also be measured precisely comparing with the range of fluctuation. To compare with, ESM datasets often contains around a hundred data points, with the measurement noise comprising about a quarter of the total variation (Dejonckheere et al., 2022).

Table 1. A comparison of the commonly used IID methods and the *fitlandr* method presented in the current article.

Method	Multidimensional ¹	Nonlinear dynamics	Continuous dynamics	Number of stable states	Transitions	Selected references
VAR-based						
VAR	Yes (high)	No	No	1	N/A	(Bringmann et al., 2013; Hamilton, 1994)
TV-VAR	Yes (high)	No	No	1 (smooth changes allowed)	N/A	(Bringmann et al., 2017; Haslbeck et al., 2021)
Threshold-VAR	Yes (high)	No	No	≥ 2 (prespecified)	Change models	(De Haan-Rietdijk et al., 2016; Tong & Lim, 1980)
Markov-Switching-VAR	Yes (high)	No	No	≥ 2 (prespecified)	Change models	(Hamilton, 1989; Haslbeck & Ryan, 2022)
Distribution-based						
HMM	Yes (high)	N/A	N/A	≥ 2 (prespecified)	Change models	(Haslbeck & Ryan, 2022; Visser, 2011)
VAR- and drift-diffusion-based						
CT-VAR	Yes (high)	No	Yes	1	N/A	(Ryan et al., 2018; Ryan & Hamaker, 2022)
Drift-diffusion-based						
FPEM	No	Yes	Yes	Any (emerged)	Within model	(Tschacher & Haken, 2020)
OU M	Yes (low)	No	Yes	1	N/A	(Kuppens et al., 2010; Oravecz et al., 2011)
AIM	Yes (2)	(specific type)	Yes	2 (specific type)	Within model	(Loossens et al., 2020)
<i>fitlandr</i>	Yes (low)	Yes	Yes	Any (emerged)	Within model	<i>The current article</i>

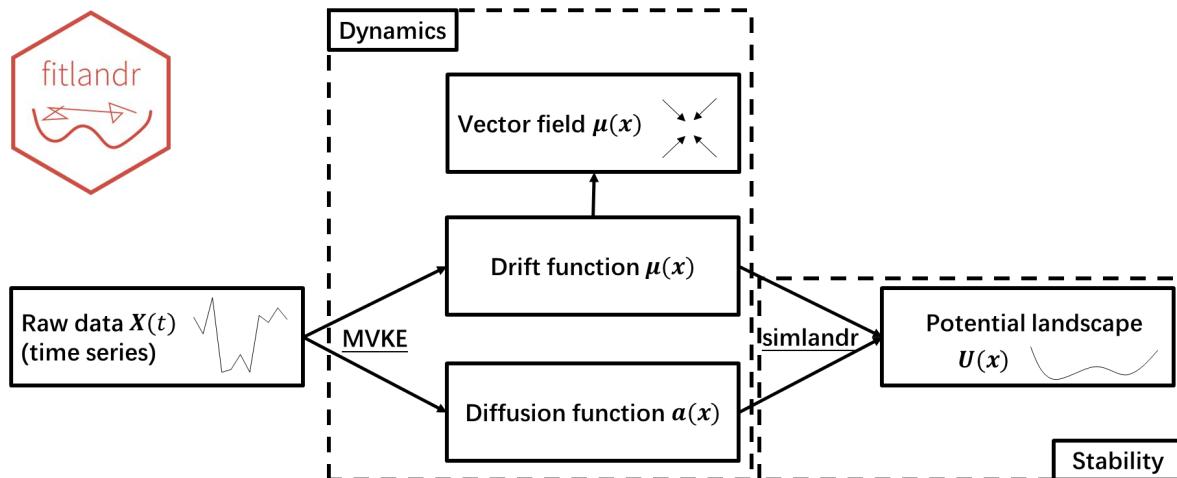
Notes: ¹VAR-based models can be easily fitted for dozens of variables, and the resulted model can be meaningfully represented in a network structure, thus practically they more often used for rather high-dimensional data. Distribution- and drift-diffusion-based models sometimes, in principle, can also be applied for high dimensional data, but the required amount of data is often larger, and the resulted model is often presented as distribution plots, vector fields, or potential

Nevertheless, in the middle of this trade-off, there is a gap in the existing range of methods: none of them can effectively represent all types of multistability in a multidimensional space for psychological ILD. This is where fitlandr comes in. The specifics of fitlandr will be discussed in depth in the subsequent sections.

The fitlandr Method

The fitlandr method draws inspiration from the dynamo project (Qiu et al., 2022), which aims to describe single cell RNA dynamics and stability with vector fields and potential landscapes. Whereas fitlandr uses similar steps as dynamo, different algorithms and procedures were implemented to accommodate the characteristics of psychological ILD, which are often shorter, more prone to noise, have fewer dimensions, and display more homeostatic behavior than single cell RNA data. The workflow of fitlandr is illustrated in Figure 1, with further explanations provided in the following subsections. The main characteristics of fitlandr, in comparison with previous methods, are shown in Table 1.

Figure 1. The workflow of fitlandr.



Step 1: Estimating System Dynamics

The first step of the method is to estimate the dynamics of the system from raw data. In fitlandr, we employ the general drift-diffusion form of stochastic differential equations to describe the dynamics of the system, written as follows¹:

¹ Note that, although sharing the same mathematical origin, the model used in this context is not directly related to the drift-diffusion model (DDM) used in cognitive decision modeling (Ratcliff et al., 2016; Ratcliff & McKoon, 2008). DDM assumes a very specific process of cognitive decision making. In contrast, the method proposed in the current article is not intended to model any decision-making behavior, nor does it have strong assumptions regarding a specific mental process.

$$d\mathbf{X} = \boldsymbol{\mu}(\mathbf{X})dt + \boldsymbol{\sigma}(\mathbf{X})d\mathbf{W}, \quad (1)$$

in which the first term in the right-hand side, $\boldsymbol{\mu}(\mathbf{X})$, represent the deterministic force, and the second term represents the noise of the system (the diffusion matrix, $\mathbf{a} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T$, is more commonly used). In fitlandr, we use a flexible kernel algorithm, namely the Multivariate Kernel Estimator (MVKE, Bandi & Moloche, 2018) to estimate the drift-diffusion equation. Assuming \mathbf{x}_i ($i = 1, 2, 3, \dots, n$) are n observations of \mathbf{X}_t in $t = 1, 2, 3, \dots, n$, the kernel estimators for $\boldsymbol{\mu}(\mathbf{x})$ and $\mathbf{a}(\mathbf{x})$ are given by (Bandi & Moloche, 2018):

$$\hat{\boldsymbol{\mu}}(\mathbf{x}) = \frac{\sum_{i=1}^{n-1} K(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_{i+1} - \mathbf{x}_i)}{\sum_{i=1}^n K(\mathbf{x}_i - \mathbf{x})}, \quad (2)$$

$$\hat{\mathbf{a}}(\mathbf{x}) = \frac{\sum_{i=1}^{n-1} K(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_{i+1} - \mathbf{x}_i)(\mathbf{x}_{i+1} - \mathbf{x}_i)^T}{\sum_{i=1}^n K(\mathbf{x}_i - \mathbf{x})}, \quad (3)$$

in which $K(\mathbf{x})$ is a product kernel function, as follows in our implementation:

$$K(\mathbf{x}) = \frac{1}{h^D} \prod_{j=1}^D e^{-\frac{x_j}{h}}, \quad (4)$$

where x_j is the j -th element of \mathbf{x} , and h is a user-specified parameter controlling the width of the kernel estimator. Intuitively, MVKE is a weighted average of the dynamics from the data. The weight of a data point depends on its distance with the evaluation point, $K(\mathbf{x}_i - \mathbf{x})$. The closer the evaluation point is to a data point, the more similar the dynamics is.

Due to the use of a kernel algorithm, the output cannot be represented as parameter estimates. Instead, the force of the system in different directions is shown through a vector field. A vector field plot displays multiple arrows in a grid, representing the direction and tendency of how the system evolves from a given point. However, it is not as straightforward to see how variables influence one another. In a VAR model, a positive coefficient of variable X_1 on variable X_2 , $c_{1,2}$ indicates that a higher value of X_1 at time t predicts a higher value of X_2 at time $t + 1$. In a vector field, however, the influence of X_1 to X_2 may not be monotonous. A higher value of X_1 may predict a higher value of X_2 at the next time point in a certain range, but the relationship may be reversed in another range, and the relationship may also depend on the value of X_2 . It is exactly this flexibility that we see as the key advantage of the current method. Whereas previous time-series methods such as VAR and OUM can also be depicted with a vector field, they limit the shape of the vector field (see Appendix B1). In the proposed method, we do not impose a specific form on the drift part. Instead, we use kernel methods to estimate it in a flexible manner.

Step 2: Estimating Stability Landscape

After obtaining the dynamic function of the system, we can go one step further to analyze the stability by constructing the system's potential landscape. The formal representation of stability is based on the concept of a potential function, which assumes that the deterministic force of the system can be represented as the gradient of the potential. In this framework, a lower potential indicates greater stability, as the system naturally moves towards states with lower potential. The steeper the slope of the potential, the stronger the system's tendency to move towards lower potential states. However, because of the mathematical properties of multivariate functions (as explained in Cui et al., 2021, 2022), constructing a well-defined potential function from an arbitrary vector field is often impossible. This is because not all the forces in the vector field can be represented as the gradient of a stability function. Therefore, we use a generalized potential function defined by Wang et al. (2008). The generalized potential landscape provides an informative summary of the stability of the states, although it may not encompass all details of the system's dynamics. This function is calculated as the negative logarithm of the steady-state distribution of the system,

$$U(\mathbf{x}) = -\ln P_{SS}(\mathbf{x}), \quad (5)$$

in which U is the generalized potential function, the higher the value of U , the more unstable the state is; and P_{SS} is the steady-state distribution of the system and estimated through Monte Carlo simulation in the fitlandr package. The drift-diffusion function obtained from the previous step, is used to simulate the system from multiple starting points until the distribution converges. The simlandr package (Cui et al., 2021, 2023) is then used to calculate and visualize the potential landscape function and to find the position of the local minima and the saddle point on the potential landscape.

Recover Nonlinear Dynamics and Stability from Simulation Data

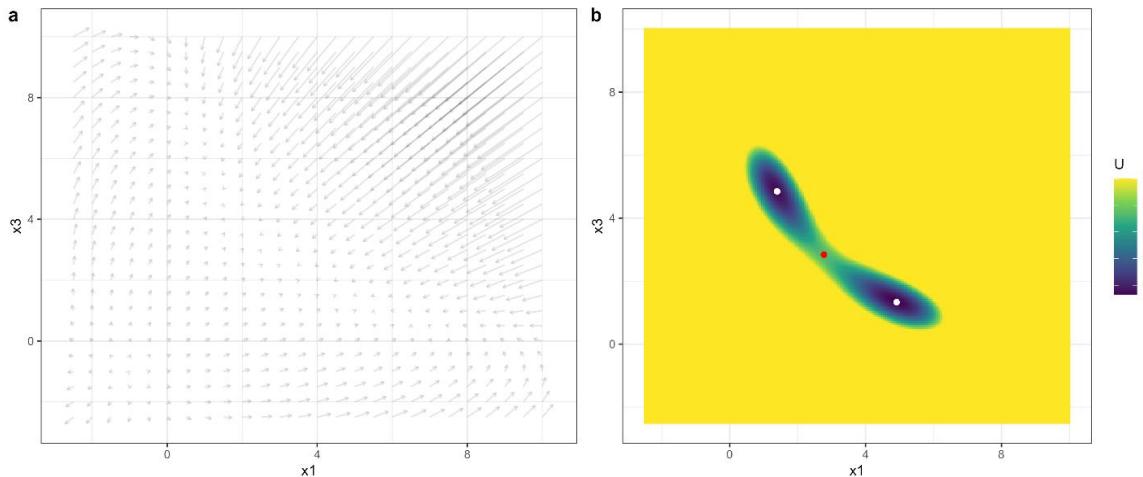
In this section, we evaluate the performance of fitlandr using model simulations. We use the mood system model by van de Leemput et al., (2014, also see Haslbeck & Ryan, 2021) for illustration. This model is based on a Generalized Lotka-Volterra model with four variables: x_1 and x_2 represent positive emotions and x_3 and x_4 represent negative emotions. Interactions between these variables are captured by the quadric terms $C_{ij}x_i x_j$ ($i, j = 1, 2, 3, 4$; corresponding to the four types of emotions), where positive emotions strengthen each other and weaken negative emotions, and vice versa. These variables also have self-reinforcing effects represented by $r_i x_i$ ($i = 1, 2, 3, 4$). The model also includes a white noise term, $\frac{\sigma_i dW}{dt}$ ($i = 1, 2, 3, 4$), for each variable. Written together:

$$dx_i = \left(1.6 + r_i x_i + \sum_{j=1}^4 C_{ij} x_i x_j \right) dt + \sigma_i dW \quad (i = 1, 2, 3, 4). \quad (6)$$

We simulate the dataset from this model using the same parameter settings as in

Haslbeck and Ryan (2021), namely $C = [C_{ij}] = \begin{bmatrix} -0.2 & 0.04 & -0.2 & -0.2 \\ 0.04 & -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.04 \\ -0.2 & -0.2 & 0.04 & -0.2 \end{bmatrix}$, and $\sigma_i = 4.5$ ($i = 1, 2, 3, 4$). This is a four-variable model, for which the full vector field and the potential landscape in the original state space can only be shown in a four-dimensional state space. However, with the aforementioned parameter settings, the dynamic equations for x_1 and x_2 are exactly symmetric, and the same holds for x_3 and x_4 . As a result, the majority of the behavior of the system can be understood by examining x_1 and x_3 . Therefore, we will only investigate x_1 and x_3 in the remaining part of this section. The true vector field and potential landscape of the system are shown in Figure 2.

Figure 2. The true (a) vector field and (b) potential landscape for the model.



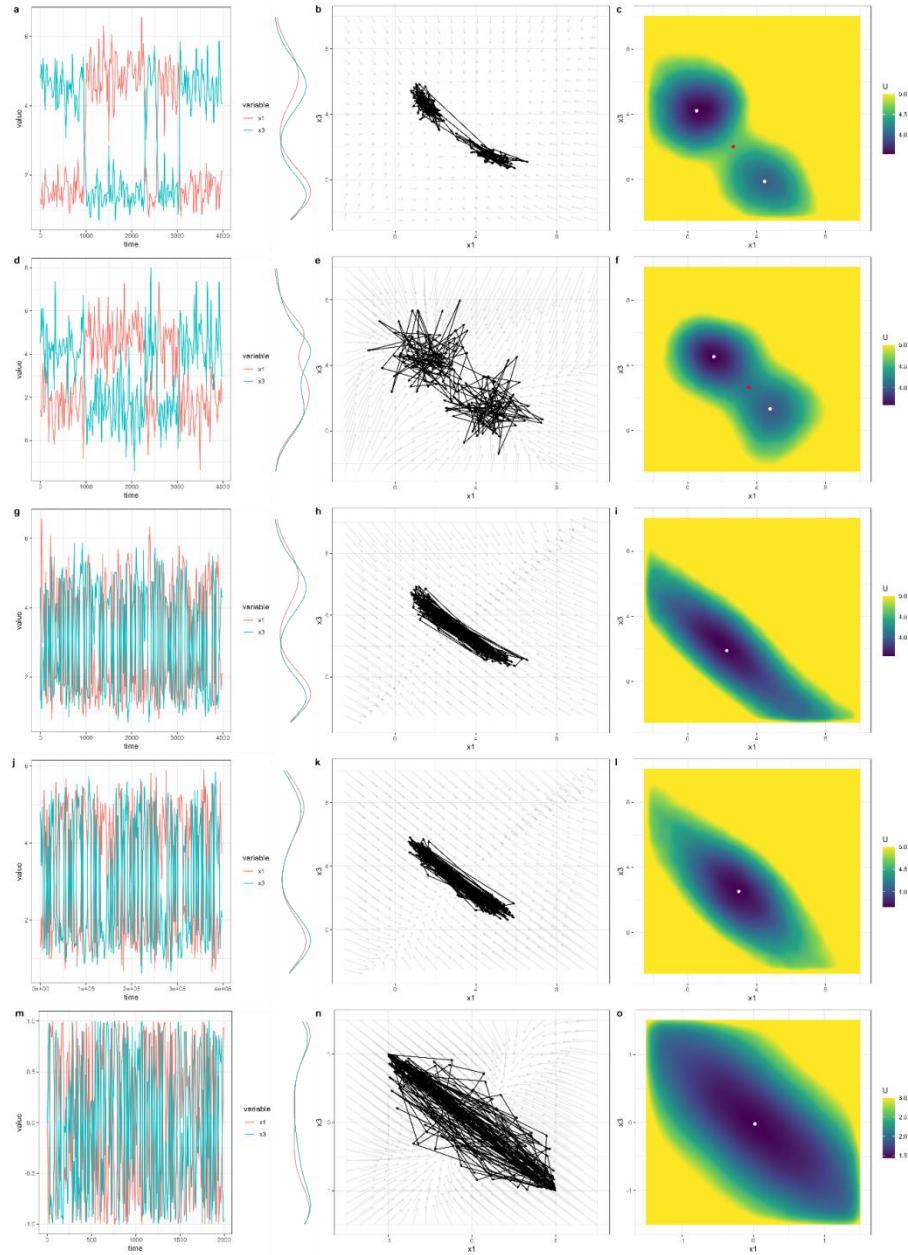
The simulation was performed under five conditions to evaluate whether and when the bistability of the system can be recovered. The five conditions are: the baseline condition, the noisy condition, the permutation condition, the long interval condition, and the polarized interpretation condition. We first focus on explaining the first two conditions. In the baseline condition, the output time series was down-sampled to every 20 time units, and only the first 200 time points after down-sampling were retained (i.e., the variable values at the $t = 0, 20, 40, \dots, 4000$ were used.) This was to ensure that the sampling frequency and the length of the data are comparable to psychological ILD. The resulting time series is shown in Figure 3a. In

the noisy condition, we added white noise drawn from $N(0, 1)$ to the baseline data to evaluate the robustness of the method against noise. We assume this noise comes from the measurement but not the system itself. Therefore, we first run the simulation as in the baseline condition, and then add a white noise to the simulation result. The resulting time series is shown in Figure 3d.

We estimated the vector fields and the potential landscapes for both conditions. The results are shown in Figure 3b-c and Figure 3e-f, respectively. In the baseline condition, the estimated vector field (Figure 3b) was largely consistent with the true vector field (Figure 2a) in terms of direction, but not in magnitude. All the arrows in the vector fields point towards one of the two regions that the system is stable in. This can be attributed to the limited information present in the short time series, which did not provide enough detail to perfectly estimate the movement tendency of the system in the unstable regions. Despite this, the potential landscape (Figure 3c) accurately captured the bistable nature of the system. There are clearly two basins on the potential landscape, and the positions of the local minima (denoted by the white dots) were close to the true positions (Figure 2b). The phases have similar stability, as indicated by the similar depth of the basins, which aligns with the characteristics of the true potential landscape as well. In the noisy condition, the addition of measurement noise disrupted the clear separation between the two phases of the system (Figure 3d). But the estimated vector field and potential (Figure 3e-f) still retained the general trends and stability features, despite becoming fuzzier. This indicates fitlandr is robust against moderate levels of noise. The results for smaller and larger noise levels can be found in Appendix B2. The results demonstrate when the noise level is too large, fitlandr is not able to recover the bistability of the system anymore.

Figure 3. The simulated time series, their vector field estimation results, and their landscape construction results. The first row is the baseline condition; the second row is the noisy condition; the third row is the permutation condition; the fourth row is the long interval condition; the fifth row is the polarized interpretation condition. The three columns are the time series with density plot aside, the vector field estimation result, and the landscape construction result. In the vector field plots, gray arrows represent the estimated vector field; the black arrows represent the vector samples from the datasets used for estimation. In the potential landscape plots, the color at the purple side represents higher stability; the color at the yellow side represents lower stability; the white dots represent local minimums of the

phases; when there are two phases on the landscapes, the red dots represent the saddle point on the path connecting the two local minimum.



At this point, it is tempting to suspect that `fitlandr` simply reflects the density distribution of the time series, given that the landscape results are similar to the density distribution of the raw data, and the density distribution of a system is known to be more robust than its dynamic features (Haslbeck & Ryan, 2022). We tested this with three other simulation conditions. In the permutation condition, the data from the baseline condition was randomly shuffled. This manipulation retains the distribution of the data points but removed the dynamical information from the data, and we use this condition to test if the method extracts the dynamical information from data or relies solely on the distributional information. The

resulting time series is shown in Figure 3g. In the long interval condition, the time series was down-sampled to every 2,000 time points instead of 20 time points in the baseline condition, and the length of the data remained the same (i.e., the variable values at the $t = 0, 2000, 4000, \dots, 400000$ were used.) This condition was used to assess the impact of having a much lower sampling frequency compared to the time scale of the changing process of interest. The resulting time series is shown in Figure 3j. In addition, the polarized interpretation condition was not simulated using the bistable Generalized Lotka-Volterra model, but a monostable VAR model specified by:

$$\begin{bmatrix} x'_{1,t+1} \\ x'_{3,t+1} \end{bmatrix} = \begin{bmatrix} 0.2 & 0.2 \\ 0.2 & 0.2 \end{bmatrix} \begin{bmatrix} x'_{1,t} \\ x'_{3,t} \end{bmatrix} + \varepsilon, \quad \varepsilon \sim N(\mathbf{0}, \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 0.5 \end{bmatrix}), \quad (7)$$

and the results were transformed with a hyperbolic tangent function,

$$\begin{aligned} x_{1,t} &= \tanh(x'_{1,t}), \\ x_{3,t} &= \tanh(x'_{3,t}), \end{aligned} \quad (8)$$

to create a bimodal sample distribution. This transformation makes x_1 and x_3 have a bimodal distribution (Figure 3m), so that it mimics a scenario where the psychological process is not inherently bistable, but the participants tend to use both extremes of the scale instead of the middle part, which creates bimodality in the data. In empirical studies, this can be induced, for example, when a 0-100 slider scale is used and the slider is initialized at 50. In that case, participants may be more likely to move the slider away from 50, thereby inducing bimodality in the data (Haslbeck et al., 2023).

For the permutation condition, the estimated vector field and potential function are shown in Figure 3h-i. It is clear that bistability cannot be recovered in this case. Both the vector field and the landscape showed a clear tendency for the system to move to the middle of the two distribution modes. With the time order randomized, it is equally likely for the system to remain in a stable phase or transition to another stable phase, causing the algorithm to identify the middle as the only stable phase of the system. We can also observe similar behaviors in the long interval condition and the polarized interpretation condition, in which the system also quickly transitioned between two distribution modes and the middle was recognized by the algorithm as the only stable phase of the system (Figure 3k-l, Figure 3n-o). These similar behaviors have different origins. In the permutation condition, the time order was shuffled, removing the time dependency from the data. In the long interval condition, the system is actually bistable, but due to the large measurement interval relative to the dynamics of interest, the dataset hardly retains dynamic information about the bistability of the system. In the polarized interpretation condition, the system is actually monostable, but the hyperbolic

tangent transformation made the data points bimodal. In this case, the monostable results from fitlandr correctly recovered the monostable nature of the system.

To summarize, our results with simulated data demonstrate that fitlandr can effectively capture the bistable nature of a system, displaying robustness in the presence of moderate noise. The results indicate that the method primarily extracts dynamic information from the data rather than simply reflecting its distributional characteristics. If the dataset lacks sufficient dynamic information about the bistability of the system, fitlandr is unlikely to produce bistable outputs, even if the data have a bimodal distribution.

Describe Nonlinear Dynamics and Stability from Empirical Data

After investigations with simulation models in the previous section, we now apply the fitlandr method to two empirical datasets. It is important to note that these empirical applications serve only as illustrations of the method. The selection of participants and variables was arbitrary and not pre-registered. Hence, results shown in this section should not be regarded as empirical claims about the behavior of these clinical groups.

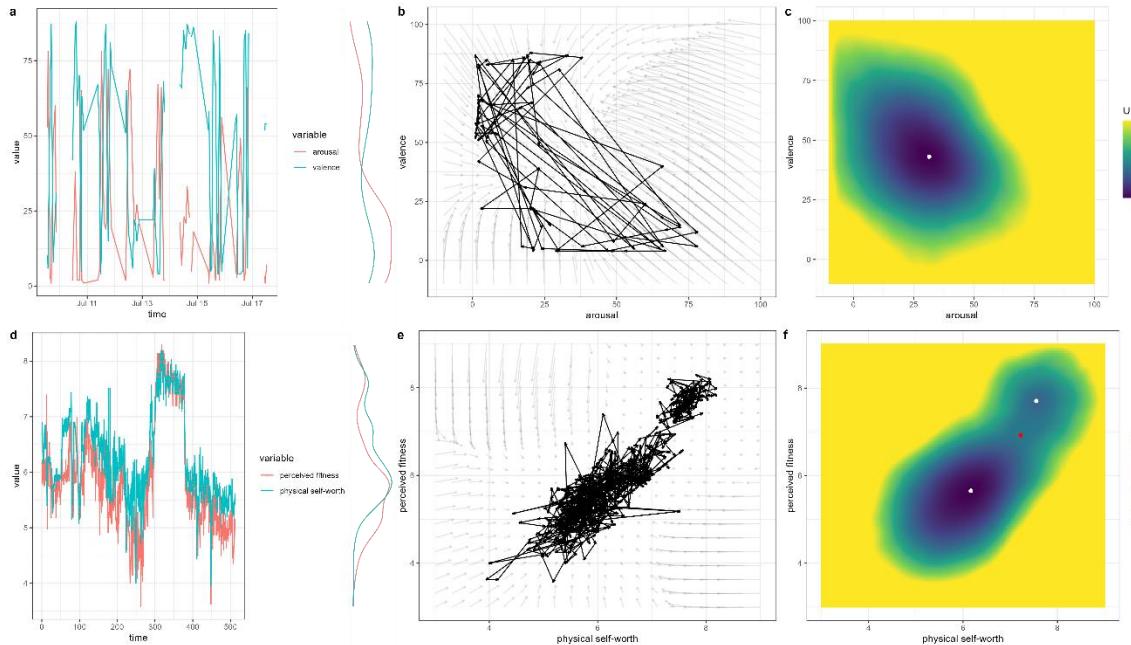
The first dataset is from the Leuven BPD study (Houben et al., 2016) and was obtained from the EMOTE database (Kalokerinos et al., in preparation) with the data request ID 32YDI3R54M. This dataset involves participants with a diagnosis of bipolar personality disorder (BPD) and a healthy control group reporting their emotional states 10 times a day over eight consecutive days. The details of the data collection procedure are described in Houben et al. (2016). We chose this dataset because patients with BPD have “unstable emotional experiences and frequent mood changes” (DSM-5TM, American Psychiatric Association, 2022). Therefore, we expect that their mood system may be multistable. In the current study, the time series data from the participant with a BPD diagnosis and with the longest record was used to estimate the vector field and potential landscape. The time series of this participant (P1) contained 4 missing values and 75 valid points. Two variables, the self-rated emotional arousal and valence, were used in this study. Both variables were measured together from an emotional grid scale and took integer values between 0 and 91. The mean values of emotional arousal and variance of this participant were 25.09 ($SD = 23.05$) and 44.81 ($SD = 33.86$), respectively. Both variables are bimodally distributed, as shown in Figure 4a.

The second dataset used in the current study is from Delignières et al. (2004) and is freely available online.² In this study, four participants filled in the Physical-Self Inventory (PSI-6, Ninot et al., 2001) twice a day for 512 consecutive days, yielding 1024 time points without

² The dataset is available at <https://didierdelignieresblog.wordpress.com/recherche/databank/>.

missing values. We used the data from the second participant (P2) due to its clear bistable patterns (as shown in Figure 4d). Two items from the scale, perceived fitness and physical self-worth, were used in this study. The mean values of perceived fitness and physical self-worth for this participant were 5.88 ($SD = 0.94$) and 6.28 ($SD = 0.79$), respectively.

Figure 4. The empirical time series, their vector field estimation results, and their landscape construction results. The first row corresponds to the data from P1 (Houben et al., 2016); the second row corresponds to the data from P2 (Delignières et al., 2004). The three columns are the time series with density plot aside, the vector field estimation result, and the landscape construction result. The meaning of the symbols in the vector fields and the potential landscapes is the same as in Figure 3.



We estimated the vector fields and the potential landscapes for both participants, and the results are presented in Figure 4. Although both variables of P1 have a bimodal distribution (Figure 4a), the vector field and the potential landscape (Figure 4b-c) both show that the system has the tendency to move to the only local minimum (denoted by the white dot on the landscape) from any starting point. Thus, only one phase was identified for the system, with a moderate level of emotional valance and arousal, which means that the dataset does not contain enough bistability information about the system. As shown in the simulation study, it is possible that the system is monostable but has a bimodal data distribution due to the reaction tendency of the participant or that the system is bistable but the measurement method did not capture enough information to reveal the bistability (e.g., due to inadequate

sampling frequency or measurement precision). Notably, the vector field of P1 showed a rotation tendency. The arrows from the low arousal, low valance region, and the high arousal, high valance region both tend to first move to the low arousal, high valance region before moving to the local minimum. This paves the way for future studies to investigate whether this holds true for the emotion regulation of this particular participant. From the time series of P2, both the vector field and the potential landscape indicate that there are two phases in the system (Figure 4e-f), with positions close to the two modes of the data distribution. This means that the participant's self-esteem regarding physical strength is likely to have bistability. The centers of the phases, which are the local minima, are denoted by the white dots on the landscape. One phase is more stable, characterized by a low level of physical self-worth and perceived fitness, and the other phase is more unstable, with a higher level of physical self-worth and perceived fitness. Because the length of the two time series differs notably, we also estimated vector fields and landscapes with a small section of the time series from P2 as well as a down-sampled time series which is more sparse than the original one. The resulting vector field and potential landscape (Appendix B3) showed similar features as for the original time series (Figure 4e-f).

Discussion

In the current article, we introduce *fitlandr*, a novel method for estimating the dynamic function and the generalized potential function of psychological systems using ILD. The method comprises two steps. The first step involves estimating the drift-diffusion using the multivariate kernel estimation algorithm (MVKE, Bandi & Moloche, 2018). Compared with previous methods, MVKE is a non-parametric, rather assumption-free algorithm, which offers adequate flexibility for recovering complex dynamics from ILD. The second step of the method entails estimating the generalized potential function from the drift-diffusion function, based on the definition by Wang et al. (2008). This definition is grounded on the steady-state distribution of the system, which can be estimated through Monte-Carlo simulation of the drift-diffusion function. We evaluated the performance of *fitlandr* using both simulation and empirical data. Results from simulation data showed that *fitlandr* is capable of recovering bistability of the system from relatively short time series (200 sample points) and is robust against moderate levels of noise. We also showed three conditions in which the datasets had a bimodal distribution but did not contain information of bistable dynamics. In those cases, *fitlandr* produces monostable results. Our analysis of two empirical datasets demonstrated that *fitlandr* generates meaningful results about the stability of the system. Although the two

datasets displayed bimodality in their variable distributions, only one of the outputs from fitlandr showed bistability.

The results of fitlandr highlight the distinction between bimodality and bistability. Whereas bimodality describes the data distribution, bistability refers to the stability of the system: if a system is bistable, and if we know the system is currently in its first phase, then it is likely that the system will still stay around the same phase in the near future. In the simulation study, we showed three cases where the bimodality in the data did not correspond to the bistability of the system. In the permutation condition and the long interval condition, although the data points were generated from a bistable system, there was almost no dynamical information in the data that showed the bistability. As a result, the algorithm was unable to determine whether the data are from a genuine bistable system, or *randomly* drawn from a bimodal distribution. When the data points are shuffled (in the permutation condition) or sampling frequency is much slower than the average time that the system transitions between two stable phases, it is equally likely that the system stays around its previous phase or moves to a new phase. Therefore, it is not possible to tell if bistability exists in the system. This can also be found in the last condition, the polarized interpretation condition. In this condition, the system is actually monostable, and the bimodal data distribution was a result of a transformation. When the underlying data generation process is unknown, those three conditions are indistinguishable to the algorithm, and hence, the algorithm produces similar monostable results. Here we also note an important difference between the characteristics of the underlying system and the information contained in the data. If the data quality is not good enough, in the sense that it does not contain enough information about the true characteristics of the system, no matter how good an algorithm is, it will always lead to faulty outputs.

The same also applies to the empirical analysis, in which we used two $N = 1$ datasets to estimate the dynamical function and the potential landscape. The formal theories about which psychological processes involve bistability is sparse (Fried, 2020; Robinaugh et al., 2021). Based on verbal theories, one may expect the emotional system of a patient with BPD (P1) to be bistable, whereas there may not be an explicit theory suggesting that the physical self-concept of a person (P2) is bistable. However, we found the opposite in the results by fitlandr. There is only one phase for P1, but two phases for P2. The dataset of P1 did show a bimodal distribution, but we did not recover bistability of the system. As shown in the simulation study, multiple reasons can lead to this result. First, the person may have two qualitatively different phases, one with high arousal and low valence emotion and one with low arousal and high valence

emotion, but the sample intervals were much larger than the average time of switching between these two phases (e.g., the average time of switching between the two affective phases may be twenty minutes, but the assessment is conducted every two hours). As a result, the data obtained were too coarse-grained to capture the switches between these phases effectively. Second, it may be that the person actually has only one psychological phase, but the person's responses are influenced by a polarized tendency, which means that the person tends to use the extremes of the scale instead of the middle of the scale. Without additional information, it is impossible to distinguish between these two explanations. Moreover, for other participants in the same group of those studies, we did not find similar results as P1 or P2 (Appendix B4), which suggests that the number and position of stable phases are highly idiographic.

The kernel method used in fitlandr enables the description of various nonlinear dynamics of psychological systems. Although the fitting procedure is rather assumption-free, there are underlying assumptions in the drift-diffusion function. The drift part and the diffusion part are only the functions of the state of the system, x . This implies that during the data collection process, only the state of the system changes, but the dynamic function of the system remains unchanged. This assumption may not be correct when for instance a clinical intervention is administrated during data collection, as this could alter the way that the variables interact with each other (e.g., a cognitive restructuring may change how physical arousal influence perceived threat, Robinaugh et al., 2019). Future development of this method may involve a moving window approach to allow the dynamics of the system to slowly change over time, as in the TV-VAR model (Bringmann et al., 2017; Haslbeck et al., 2021). Another assumption of the drift-diffusion function is the Markov property, which means that the evolution of the system only depends on the current state of the system and not on historical states. As a result, it cannot directly predict periodic changes, such as those in cyclothymic disorder (Akiskal & Pinto, 1999; American Psychiatric Association, 2022a). Including previous time points in the dynamic function may help to remedy this issue, as in phase space reconstruction methods (Marwan et al., 2007; Takens, 1981). MVKE also assumes equal distance between observations, which is often not true. Future statistical development of the method extends the method to handle different time intervals, as in the CT-VAR model. However, it is unclear if the physical time interval should be used in the modeling. Physically, 8 hours from 11 PM to 7 AM and 8 hours from 9 AM to 5 PM may be equal, but the psychological changes during those two time slots are not likely to be the same. Further research is needed to determine the appropriate time scale for measuring psychological variables.

We used around 200 time points in the current article, which is feasible for an ESM study, but is not always easy to conduct. Here, we provide three general principles to guide the data collection for using the fitlandr method. First, the period of data collection should be long enough to observe different phases of the system. For a patient with a bipolar disorder, for example, data points from both the depressive phase and the manic phase should be collected to describe the possible bistability of the system. Second, as fitlandr assumes that the dynamics and stability of the system does not change over time, the data collection period should not be too long that it contains significant developmental changes or treatment changes. Alternatively, if the time series does contain significant changes of the system dynamics, it may be better to use the fitlandr method in moving windows instead of applying it to the entire time series. Third, the measurement frequency should match the change rate of the phenomena, so that there is sufficient information about the stability of the system in the dataset. It is difficult to provide guidance about exactly how many data points are sufficient to produce stable results, because the results shown in the current study may not be applicable to all real-life scenarios, and the dynamic properties of different people and different variables can differ in so many ways. Future work may investigate different specific scenarios and provide estimations of the number of data points required for this method.

Currently, fitlandr can only be applied to a single participant dataset. It might be appealing to develop multilevel extensions to fitlandr, as in the Dynamic Structural Equation Models (DSEM, Hamaker et al., 2018). It is technically possible to build a multilevel model for nonparametric methods (e.g., Li et al., 2006). However, as the dynamic functions and potential landscapes can be very different within the same group of people (Appendix B4), we doubt if aggregating the landscapes will yield meaningful results. It is good to have a method that helps to generalize the results to a larger group, but this is only meaningful to the extent that the result is generalizable. Therefore, we highly suggest the current method to be used as an $N = 1$ method unless there is sufficient evidence suggesting the dynamic functions or stability landscapes at the group level are similar to the landscapes of the individuals that make up the group (Fisher et al., 2018; Hamaker, 2023; Hekler et al., 2019; Molenaar, 2004; Olthof et al., 2023).

Conclusion

We introduce the fitlandr method in the current article to flexibly estimate the dynamical drift-diffusion function and generalized potential landscape function for psychological systems. This approach uses MVKE, a nonlinear kernel method, to estimate the drift-diffusion function, employs Monte Carlo simulation to determine the steady-state distribution, and uses

the simlandr package to calculate the generalized potential landscape function. Our method is effective in detecting bistability from both simulation and empirical data, is robust against noise, and relies on the dynamic information from the data instead of the distributional information.

Chapter 5

Quantifying the Stability Landscapes of Psychological Networks

This chapter is based on:

Cui, J.*, Lunansky, G.*¹, Lichtwarck-Aschoff, A., Mendoza, N., & Hasselman, F. (2023).
Quantifying the Stability Landscapes of Psychological Networks. PsyArXiv.
<https://doi.org/10.31234/osf.io/nd8zc>



Abstract

The network theory of psychopathology proposes that mental disorders can be represented as networks of interacting psychiatric symptoms. These direct symptom-symptom interactions can create a vicious cycle of symptom activation, pushing the network to a self-sustaining dysfunctional phase of psychopathology: a mental disorder. Symptom network models can be estimated from empirical data through statistical models. Although simulation studies have established a relation between the structure of these symptom network models and the probability they end up in a self-sustaining dysfunctional phase, the general stability of the system is left implicit. The general stability includes both the stability of the dysfunctional phase and the stability of the healthy phase. In this paper, we present a novel method to quantify the stability landscapes of network models through stability landscapes. Our method is based on the Hamiltonian of the microstates of Ising models and can be used to show the stability of estimated Ising network models. Compared to simulation-based methods, our approach is computationally more efficient and quantifies the stability of all possible system states. Furthermore, we propose a set of stability metrics to quantify the stability of the healthy and dysfunctional phases and a bootstrapping method for range estimation of the stability metrics. To demonstrate the method's utility, we apply it to an empirical data set and show how it can be used to compare the stability of phases between groups. The presented method is implemented in a freely available R package, *Isinglandr*.

Introduction

Over the past decades, there has been a growing interest in understanding psychopathology through the lens of network theory. This approach conceptualizes mental disorders not as isolated symptoms with a common underlying cause but as complex systems where symptoms interact with and influence one another (Cramer et al., 2016; Olthof, Hasselman, Oude Maatman, et al., 2023). In a network model of psychopathology, the analysis takes place at the level of interacting symptoms, where each symptom is represented as a node, and the connections (called edges) between nodes represent the relationships between symptoms (Borsboom, 2017). The interactions between symptoms are crucial for understanding the overall behavior of the system. For example, sleeping problems can lead to worrying thoughts, which in turn can increase fatigue and depressed mood (Cramer et al., 2016).

Individuals whose symptom networks have strong connections between symptoms may be more vulnerable to developing a mental disorder, as the activation of one symptom can trigger a cascade of others. Conversely, those with weaker connections may be more resilient. Often the system is organized such that certain phases (i.e., qualitatively different patterns of behavior) emerge from the system, such as the healthy phase and the depressive phase (Cui et al., 2023). The stability of these phases can vary, explaining why some individuals show resilience while others experience persistent mental disorders (Cramer et al., 2016; Olthof et al., 2020).

One way to advance the understanding of psychopathology from a network perspective is to estimate statistical network models from empirical data. These models reveal the conditional associations between symptoms, controlling for all other variables (Borsboom et al., 2021; Bringmann et al., 2022; Epskamp et al., 2018). However, while these models provide insights into the relationships between symptoms, they do not directly indicate the stability of different system phases – meaning we cannot easily determine whether a group of people is more likely to be in a healthy or depressive phase based solely on their network structure. Knowing the stability of different phases is crucial both for clinical practice and for better theorizing mental disorders from the network theory, whereas the methods to depict the stability of phases from network models are still sparse.

To address this gap, researchers have used simulation-based approaches to predict the likely phase of a symptom network. For instance, studies have examined whether networks with strong associations among symptoms are more likely to result in a depressive phase (Cramer et al., 2016; Lunansky et al., 2024). However, these simulations often provide only a

rough estimate and do not capture the overall stability of the entire system across different starting points. Moreover, the accuracy of these predictions can be influenced by random noise and other factors, making it challenging to draw definitive conclusions.

A useful way to conceptualize the stability of a symptom network is through the metaphor of a stability landscape. Imagine a ball rolling on an uneven landscape, where lower positions represent more stable states. The shape of the landscape determines where the ball (or system) is likely to settle. In this analogy, different basins on the landscape correspond to different psychological phases, such as the healthy phase and the depressed phase (Hayes & Andrews, 2020). The stability of these phases determines the direction in which the system is likely to evolve. For example, individuals characterized by a stability landscape with a deep and steep basin for the depressive phase are more likely to develop and remain in a depression, while those with a shallow basin for the healthy phase may be more resilient.

While the stability landscape metaphor has been frequently used, it has not been fully computed from estimated psychological networks. Recently, a method was proposed to calculate the stability landscape of formal dynamical models (Cui et al., 2023). In this paper, we will use those insights to introduce a novel method to compute the full stability landscape from networks estimated from cross-sectional data. This method allows us to infer the population-level stability of network states. Most network studies to date use cross-sectional data (Bringmann et al., 2022; Epskamp et al., 2018), however, efforts to estimate person-specific models based on multivariate timeseries data are pursued by many authors (Bringmann, 2021; Hulsmans et al., 2024; Mansueto et al., 2023; Olthof, Hasselman, Aas, et al., 2023; A. G. C. Wright & Woods, 2020). To draw inferences on the individual level we would need to compute stability landscapes from longitudinal individual data. However, the methods to estimate non-linear networks from such data are still undeveloped, making it currently infeasible to compute individual landscapes. Thus in this paper, as the first step, we focus on cross-sectional data and stability landscapes for the group level.

The paper is structured as follows: First, we introduce the concept of formal stability landscapes. Next, we explain how stability landscapes can be calculated from cross-sectional Ising network model, which is particularly well-suited for demonstrating non-linear dynamics and bistability (van Borkulo et al., 2014). We also propose a set of metrics to quantify phase stability, made available through the R-package *Isinglandr*, which can be applied to any psychological Ising network. Finally, we provide an empirical illustration of how stability landscapes can be used to compare different groups and discuss the main contributions and limitations of our method.

Formalizing Stability Landscapes for Psychological Networks

Currently, there are numerous network psychometric models that can be applied to psychological systems to analyze the relationships among variables. The choice of model depends on the type of data being used. For instance, the Gaussian Graphical Model (GGM) is used for cross-sectional, continuous data, the Vector Autoregression model (VAR) is used for longitudinal, continuous data, and the Ising network model is used for cross-sectional, binary data. GGM and VAR are both linear models and cannot represent bistable systems (Haslbeck et al., 2022). However, bistability is often assumed in theoretical formulations of psychopathology (Hayes & Andrews, 2020; Olthof, Hasselman, Oude Maatman, et al., 2023; Scheffer et al., 2018) and observed in clinical settings (Helmich et al., 2020; Miller, 2004; Tang & DeRubeis, 1999). The Ising network model, although not as widely used, is capable of showing the bistability of a system. Therefore, in this article, we use the Ising network model for illustration purposes, as in Cramer et al. (2016) and Lunansky et al. (2024).

In an Ising network, the system state is represented by the activation state of each node, denoted as $\mathbf{a} = [a_1, a_2, \dots, a_n]^T$, where a_i represents the activation state of the i -th node, with 1 indicating an active state and 0 indicating an inactive state. In the context of psychological networks, active and inactive states correspond to the presence or absence of symptoms, respectively. The Ising network model consists of two types of parameters (van Borkulo et al., 2014): node thresholds and edge weights. The threshold of a node reflects its inherent tendency for activation. When all other nodes are in the same state, increasing the threshold of a node makes it less likely to be activated. Edge weights represent the strength of the association between two nodes. The meaning of the edge weight varies depending on the representation of the node states (Haslbeck, Epskamp, et al., 2021). In the current article, we adopt the standard representation used in most psychological literature, where a higher edge weight $w_{i,j}$ between node i and j indicates a higher probability that both nodes will be activated ($a_i = a_j = 1$) relative to all other possible node states (i.e., only one node is activated or neither node is activated). Therefore, decreasing the thresholds of nodes and increasing the edge weight of nodes makes it more likely that more nodes are activated.

The probability that a node is active, given the active states of all other nodes and the network parameters, can be calculated using the following formula (van Borkulo et al., 2014):

$$P(a_i = 1 | a_j, j = 1, 2, \dots, i - 1, i + 1, \dots, N) = \frac{1}{1 + e^{-\beta(m_i + \sum_{j=1, j \neq i}^N w_{i,j} a_j)}}, \quad (1)$$

where m_i is the threshold for the i -th node, $w_{i,j}$ is the edge weight between the i -th and j -th nodes ($w_{i,i} = 0$), and β represents the randomness of the system (typically set to 1). The thresholds for all nodes, denoted by $\mathbf{m} = [m_1, m_2, \dots, m_N]^T$, and the edge weights for each

pair of nodes, denoted by $\mathbf{W} = \begin{bmatrix} w_{1,1} & \cdots & w_{1,N} \\ \vdots & \ddots & \vdots \\ w_{N,1} & \cdots & w_{N,N} \end{bmatrix}$, can be estimated using logistic regressions

(van Borkulo et al., 2014). Together, these parameters fully describe an Ising network.

To formalize the stability landscape for psychological networks, draw upon the concept of the stability landscape from physics, which represents the potential energy in each possible state, and the movement tendency of the system is along the gradient of the landscape¹:

$$\frac{dx(t)}{dt} = -\frac{dU(x)}{dx}, \quad (2)$$

in which x is the state of the system, dx/dt represent the speed of movement for the system, $U(x)$ is the potential energy for a given state x , and $dU(x)/dx$ is how the potential energy increase or decrease along x , representing the steepness of the landscape. Intuitively speaking, this means the state of the system tends to go down to a lower place of the landscape, and when the landscape is steeper, it also moves faster. This is exactly what we mean by the ball-and-landscape analogy. Therefore, to search for a landscape definition for networks, we need to have a variable x to summarize the state of the system, as well as a function $U(x)$ to represent the stability of the system in a given state.

For Ising networks, there is already an intrinsic energy measure for each possible activation state, namely, the Hamiltonian energy (Ising, 1925; also see Brusco et al., 2021, for theoretical discussions in the psychological literature). For a state of the network, \mathbf{a} , the Hamiltonian, H , is given by:

$$H(\mathbf{a}) = -\sum_{i=1}^{n-1} \sum_{j=i+1}^n w_{i,j} a_i a_j - \sum_i a_i m_i = -\frac{1}{2} \mathbf{a}^T \mathbf{W} \mathbf{a} - \mathbf{a}^T \mathbf{m}, \quad (3)$$

The probability that the system is in a state \mathbf{a} is given by:

$$P(\mathbf{a}) \propto e^{-\beta H(\mathbf{a})}. \quad (4)$$

Can we use $H(\mathbf{a})$ to construct the stability landscape of the system? While theoretically possible, it is not practical due to the vast number of possible states (2^N), making it difficult to comprehensively list the Hamiltonian energy of each. However, in the case of psychopathology, the number of symptoms is often used as an indicator of the severity of the psychopathology,

¹ Under zero-inertia condition; the constant coefficient is omitted.

especially for diagnostic purposes (American Psychiatric Association, 2022). Therefore, as a solution to the vast number of possible states in the network model we opt to use the number of active nodes,

$$n = \sum_{i=1} a_i, \quad (5)$$

instead of the vector \mathbf{a} as the representation of the system state. To clarify, in the remaining part of this article, we only use the term state for x (sometimes also referred to as *macrostate* in the literature), in contrast to the states represented by \mathbf{a} as *microstates* (Dalege et al., 2018). Note that the state of the system represented by n is a meaningful abstraction of the microstates only when the nodes are similar enough so that it is reasonable to calculate the number of active nodes. If the nodes have diverse meanings, such as having both positive and negative values, then the summation of the nodes may not be appropriate.

One state of the network may be generated by many potential microstate configurations. For example, if there are five nodes in a network, then there are $C_5^2 = 10$ possible microstates belonging to the state that two nodes are active (see Figure 1 for several examples)². To define the potential function of a state, we need to effectively summarize the information from its microstates. Here we introduce the generalized potential function by Wang et al. (2008; see also Cui et al., 2023). Wang's landscape is defined by the steady-state distribution of the system, P_{SS} , which is a distribution of systems that does not change over time. If a large number of copies of an Ising network evolve for a long period of time then ultimately, the distribution of the number of active nodes n will converge to $P_{SS}(n)$. Wang's landscape function is defined as:

$$U(x) = -\ln P_{SS}(x). \quad (6)$$

At first glance, this definition of stability landscapes appears circular and equivalent to Equation 4. However, Equation 6 is more general because it can be used for other variables that represent the state of the system (here, the state n instead of the microstate of the system \mathbf{a}) while still providing meaningful results. For Ising networks, the steady-state distribution is the same as defined by Equation 4³. Hence, we can use Equation 4 to calculate the steady-state

² The number of microstates belonging to a state is termed as the multiplicity of the state.

³ Strictly speaking, it is the other way around. The Ising network model itself does not specify how the system evolves over time. Therefore, the steady-state distribution of common simulation algorithms for Ising systems all converge to the possibility distribution defined by Equation 4 (e.g., the Glauber dynamics, Glauber, 1963).

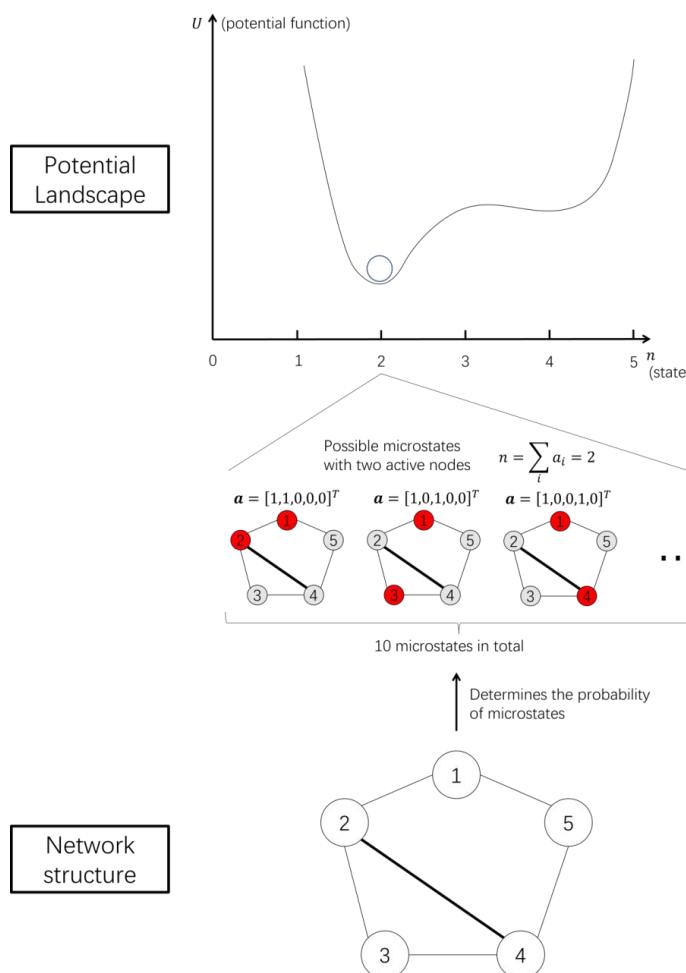
distribution of the system analytically, without running simulations. Combining Equations 4-6, we define the generalized stability landscape for the state variable n as follows⁴:

$$U(n) = -\ln P_{SS}(n) = -\ln \sum_{a_i, \sum a_i = n} P(a_i) = -\ln \sum_{a_i, \sum a_i = n} e^{-\beta H(a_i)}, \quad (7)$$

where $H(\mathbf{a})$ is defined by Equation 3.

We developed `Isinglandr`, an R package that implements the methods presented in this article and produces the results reported here. The package is available from the Comprehensive R Archive Network (CRAN) at <https://cran.r-project.org/package=Isinglandr>, and the code used to generate the results is available on the Open Science Framework (OSF) repository at <https://osf.io/y3kju/>.

Figure 1. Diagram of the relationship between the network structure, the microstate and the state of the system, and the stability landscape. The network structure is arbitrarily chosen for illustration. Red nodes represent active nodes and gray nodes represent inactive nodes. The thickness of the edges represents the strength of the connection. See Equation 3.



⁴ Constants are omitted for brevity.

Calculating the Stability Landscape of Depression

In this section, we show how variations in the symptom network structure lead to different stability landscapes. This illustrates the relationship between the symptom network structures of the cross-sectional Ising network and the stability landscapes. Building on previous studies, we systematically adjust the parameters of an estimated cross-sectional baseline Ising network to demonstrate how the stability landscape changes with different network structures (Lunansky et al., 2024).

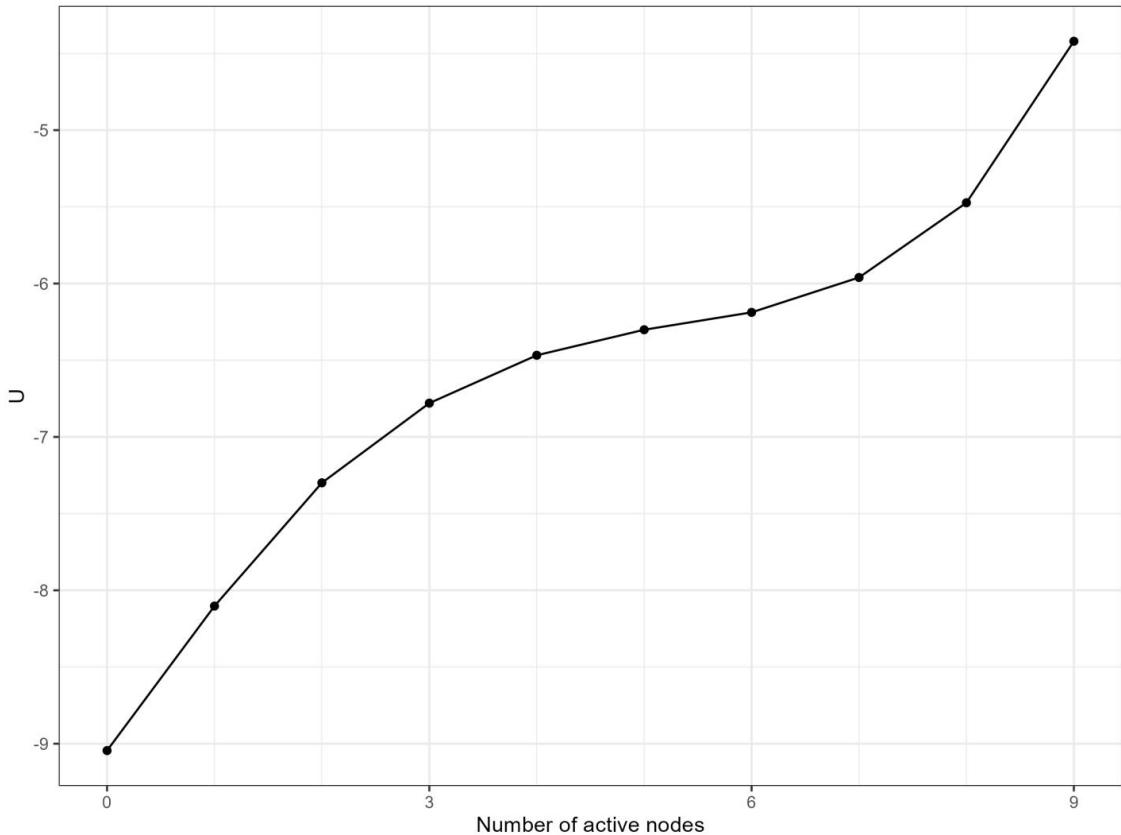
Estimating the Baseline Model

The baseline network is estimated from data from the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD; Kendler, Kenneth S. & Prescott, Carol A., 1999). The data contain binary data on the presence/absence of nine depression symptoms from 8973 twins from the Mid-Atlantic Twin Registry (see Appendix C for a description of all nine depression symptoms). We received the estimated Ising network parameters from these data, which were estimated with the *IsingFit* package in R (van Borkulo et al., 2014; van Borkulo & Epskamp, 2023)

The stability landscape of the baseline networks is shown in Figure 2. The x-axis represents the number of active nodes (i.e., the system's state), and the y-axis represents the generalized potential function of each state. Higher potential values indicate less stability for that state, on average within the sample. The potential function has two local minima at $n = 0$ and $n = 6$, indicating two phases: the *healthy phase* (fewer active symptoms, $n = 0$ to $n = 5$) and the *depressive phase* (more active symptoms, $n = 5$ to $n = 9$). The lower potential at $n = 0$ suggests that the healthy phase is more stable than the depressive phase, meaning that, although the system can reside in the depressive phase, it is likely to transition back to the healthy phase.

Stability Landscapes of Varying Symptom Network Structures

Alternative network structures were created by multiplying the baseline model's connectivity and threshold parameters by constants (0.8, 1, or 1.2), allowing us to simulate increases, decreases, or no changes to these values. Figure 3 shows the resulting stability landscapes for these different parameter configurations. The center panel shows the stability landscape when both the threshold parameters and overall connectivity parameters are unchanged (multiplied by 1), matching the baseline stability landscape in Figure 2.

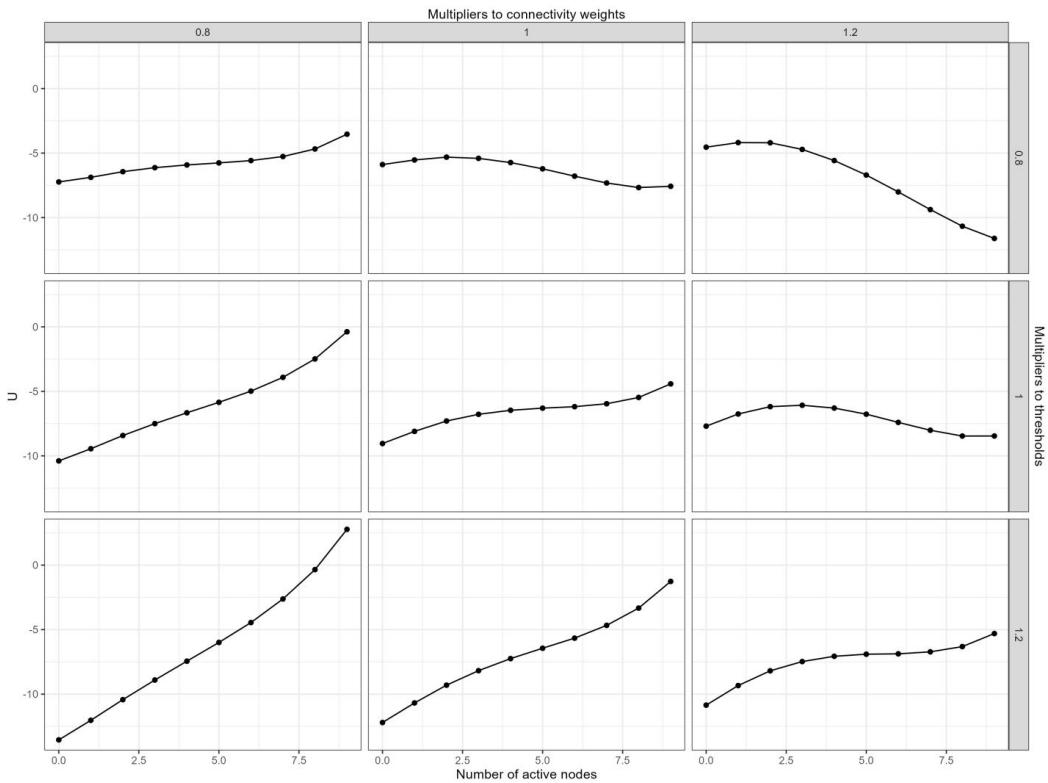
Figure 2. The landscape of the baseline network.

Overall, we observe that weaker connectivity (multiplied by 0.8) and stronger thresholds (multiplied by 1.2) lead to fewer activated nodes and a more stable healthy phase. This is evident in the lower-left panel, where the stability landscape shows only one phase with a local minimum at $n = 0$, indicating that the corresponding network tends toward having fewer active symptoms. Conversely, increased connectivity (multiplied by 1.2) and weaker thresholds (multiplied by 0.8) result in more activated symptoms, enhancing the stability of the depressive phase. This scenario is illustrated in the upper-right panel, where the depressive phase becomes more stable, with a local minimum at $n = 9$.

When changes in connectivity and threshold values offset each other, the effects on the stability landscape may cancel out, as shown in the upper-left and lower-right panels, where the stability landscapes remain similar to the baseline.

For readers interested in further exploring how parameter changes influence system stability, the *Isinglandr* package's Shiny app (`Isinglandr::shiny_Isingland_MDD()`) allows for interactive parameter adjustments and visualizations.

Figure 3. The landscapes of variations of the baseline network structure. Overall connectivity and threshold parameters of the baseline network are systematically multiplied with 0.8, 1, or 1.2, and the panels show the resulting stability landscapes.



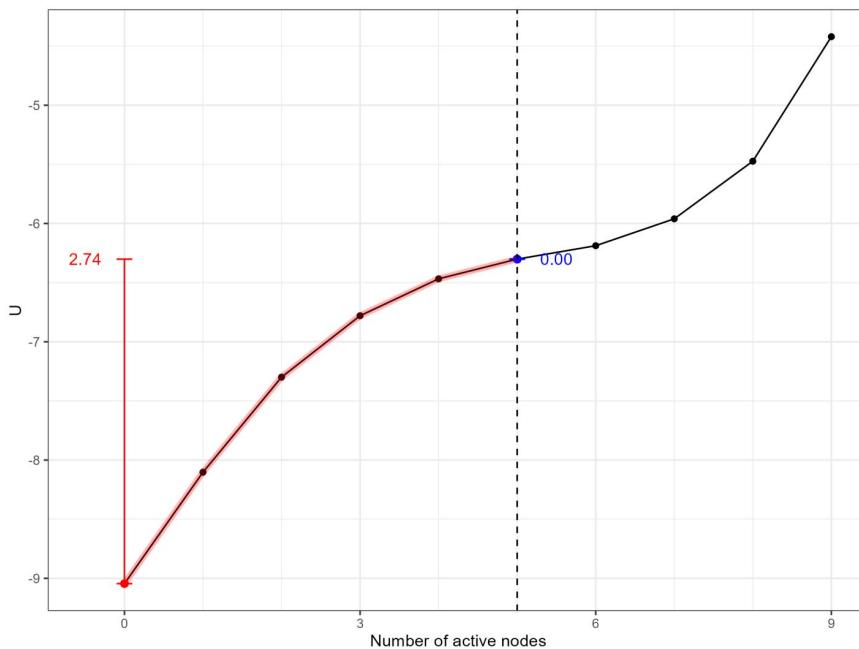
Stability Indicator Metrics

To better interpret the stability landscapes, we propose a set of metrics to quantify the stability of the healthy and depressive phases for different network structures. The stability landscape provides a quantitative representation of the stability of the network states, serving as a basis for describing the relative stability of these phases. Previous studies have suggested using barrier height as a relative stability measure of psychological phases (Cui et al., 2023), but this metric is undefined when there is only one stable phase in the landscape, and it does not account for the clinical significance of the number of symptoms.

To address these limitations, we use the clinical cutoff of depression (in this example, five active symptoms out of nine nodes as in DSM-5-tr, American Psychiatric Association, 2022) to demarcate the healthy and depressive phases on the stability landscape. Specifically, the stability of the healthy phase is calculated based on the left portion of the potential function (states with fewer than five active symptoms), and the stability of the depressive phase is calculated from the right portion (five or more active symptoms).

Within the left portion, the difference between the leftmost local minimum and the maximum value to the right of that local minimum is used as the stability of the healthy phase, and the stability of the depressive phase is defined symmetrically. This approach ensures consistency with clinical definitions of depressive diagnosis and severity. The resulting stability metrics are shown in Figure 4 and Figure 5. The figures illustrate how different parameterizations affect the stability of each phase, with the healthy phase becoming less stable and the depressive phase more stable as network connectivity increases and thresholds decrease.

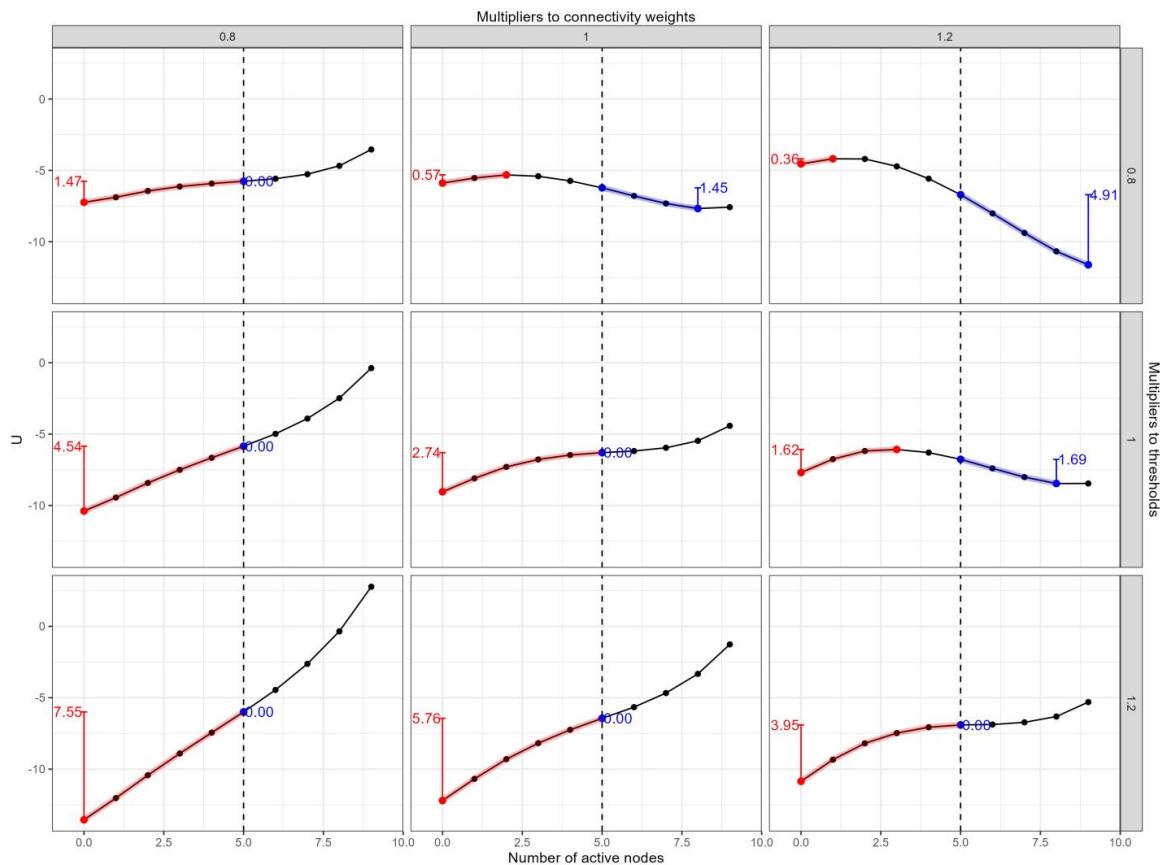
Figure 4. Stability metrics for the baseline network. The red lines indicate the range of the landscape function used for calculating the stability of the healthy phase, and the blue lines indicate the range of the landscape function used for calculating the stability of the depressive phase. The vertical lines represent the difference in the potential function U between two reference points. The red and blue numbers are the values of the stability metrics for the healthy and depressive phases, respectively.



To capture the overall stability of both phases, we calculate the stability difference by subtracting the stability metric for the depressive phase from that of the healthy phase. A positive stability difference indicates that the healthy phase is more stable than the depressive phase, with a larger absolute value reflecting a greater relative stability of the healthy phase. In contrast, a negative stability difference suggests that the depressive phase is more stable, with larger values indicating a stronger tendency in the sample towards the depressive phase. The

stability differences for all the landscapes of variations of the baseline network structure are shown in Table 1.

Figure 5. Stability metrics for the landscapes of variations of the baseline network structure. The red lines indicate the range of the landscape function used for calculating the stability of the healthy phase, and the blue lines indicate the range of the landscape function used for calculating the stability of the depressive phase. The vertical lines represent the difference in the potential function U between two reference points. The red and blue numbers are the values of the stability metrics for the healthy and depressive phases, respectively.



Bootstrapping for Stability Metric Uncertainty. For empirical datasets, we can also calculate the uncertainty of the stability metrics with bootstrapping. The bootstrapping method has been widely used in various psychometric contexts for estimating the uncertainty of parameters (e.g., Epskamp et al., 2018; Mallinckrodt et al., 2006; D. B. Wright et al., 2011). It resamples participants with replacement from the original dataset many times, estimates the parameters from the resampled datasets, and use those parameter values to obtain the range estimation of the parameter. Compared with parametric methods, the bootstrapping method

does not require specific knowledge about the parameter distribution, thus is suitable for the stability metrics. The resampling process should be operated from the original dataset instead of the network parameters. Therefore, we will only explain the method here, and leave concrete examples for the next section, where the original dataset used for network estimation is available.

Table 1. Stability differences for the landscapes of variations of the baseline network structure. Higher values indicate that the healthy phase is more stable compared with the depressive phase.

Multipliers to thresholds	Multipliers to connectivity weights	Stability difference
0.80	0.80	1.31
0.80	1.00	-0.88
0.80	1.20	-4.56
1.00	0.80	3.73
1.00	1.00	2.58
1.00	1.20	-0.08
1.20	0.80	6.10
1.20	1.00	4.95
1.20	1.20	3.79

In the Isinglandr package, we use the *boot* (Davison & Hinkley, 1997) and *boot.pval* (Thulin, 2023) packages to estimate the standard error, confidence interval, *p*-values, as well as the significance of stability metrics. For the stability difference between the two phases and the group comparison of the stability differences, we apply the bias-corrected and accelerated (BCa) method as the bootstrap distribution of the stability difference is often skewed. Previous research has shown that the BCa method performs the best in such cases (Puth et al., 2015). The stability metrics of a single phase are often highly zero-inflated, which can lead to errors or unrealistic estimations. Therefore, we use the percentile method for the stability of a single phase. Note that the percentile method is less accurate than the BCa method (Puth et al., 2015). Thus, the range estimates for the stability of a single phase should be interpreted with caution.

Using Stability Landscapes to Compare Groups

Network analysis is a powerful tool that can provide an overview of the specific interactions between variables in a population of interest, making it highly suitable for comparisons between groups (e.g., males and females; Burger et al., 2023). Network comparison offers a comprehensive way to gain insight into group differences, such as whether specific edges

between nodes vary across populations (e.g., gender differences in psychopathology networks, Kendler et al., 2022, or differences in stress disorder networks between young adults and adolescents, Sun & Zhou, 2023). In the context of psychopathology, group-level network comparisons have been used to understand whether different groups, based on, for example, age or gender, may need different clinical treatments (Lee & Hu, 2022). In this section, we demonstrate how stability landscapes can be used for group-level network comparisons with an empirical illustration, and we explain how this approach differs from a widely used *Network Comparison Test* (NCT, van Borkulo et al., 2022).

The NCT is a statistical test that compares network structures between groups using resampling-based permutation (van Borkulo et al., 2022). The test checks for significant differences in the (1) overall network structure, (2) global connectivity of the networks, and, if global differences are detected, (3) differences in strength between specific edges. The test is implemented in the software package *NetworkComparisonTest* (van Borkulo et al., 2022) within the R-environment. The test can be used for various network models, including the Ising network model.

Empirical Illustration

Here, we give an empirical illustration in which we compare psychopathology symptom networks between individuals with low and normal to high resilience. We show how to compare the stability landscapes and compute the stability difference between the groups, in addition to applying the Network Comparison Test (van Borkulo et al., 2022).

Methods

Data. The data contains assessments of coronavirus anxiety symptoms using the Coronavirus Pandemic Anxiety Scale (CPAS-11, Bernardo et al., 2022) and the Brief Resilience Scale (BRS, Smith et al., 2008) in 2436 participants. The data was collected online during the outbreak of the Delta variant of SARS-CoV-2 in the Philippines (see Dizon et al., 2023; Mendoza et al., 2022 for more details on the data collection procedure).

Networks. We split the group based on their BRS scores according to the cut-off values as reported by Smith et al. (2013): low resilience ($BRS < 3$, $n = 836$) versus normal to high resilience ($BRS \geq 3$, $n = 1551$). When comparing groups, it is best that sample sizes are equal to ensure that differences in the estimated networks are not due to differences in power (van Borkulo et al., 2022). Therefore, we randomly sampled 836 participants from the normal to high resilience group. In this way, both networks are estimated from the same number of participants.

We then estimated the Ising network models of the coronavirus anxiety symptoms (CPAS-11) for these groups using the IsingFit package in R (van Borkulo et al., 2014; van Borkulo & Epskamp, 2023). To do so, we first binarized the CPAS-11 scores as the Ising network model is estimated from binary data (van Borkulo et al., 2014). The CPAS-11 items range from 0 (no symptom presence at all) to 3 (symptom presence nearly every day). We chose to binarize on symptom absence versus some symptom presence. This means that all CPAS-11 values ≥ 1 are recoded into 1. The cutoff value of CPAS-11 is 15 for the 0-3 scale (Bernardo et al., 2022). This corresponds to 8 active nodes for the binarized scale (see Appendix C for details), and we will use this value for calculating the stability metrics. This means that the healthy phase consists of a sum score of actively present symptoms below 8, and the anxious phase starts at 8 actively present symptoms or more. Note that we did not include the resilience factors (BRS) in the networks, as these variables were used to split the groups. Including variables into the networks that were initially used to split the groups could potentially introduce bias (de Ron et al., 2021; Haslbeck, Ryan, et al., 2021).

Comparison of the Two Networks. From these networks we (1) compute the stability landscapes as described in the previous section, (2) compute the stability difference, and (3) apply the NCT to test for significant differences in global connectivity and overall structure between the two symptom networks of resilience groups (van Borkulo et al., 2022).

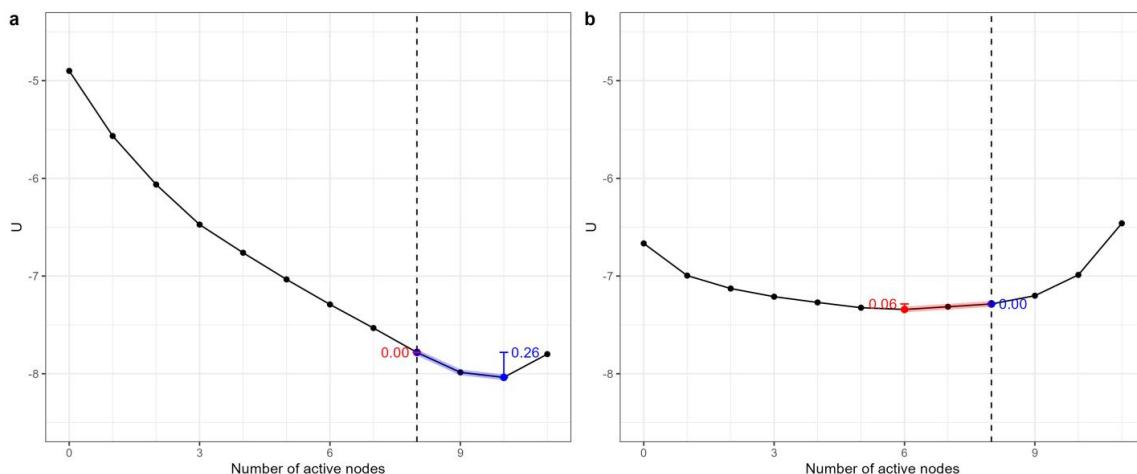
Results

Figure 6 shows the results for the stability landscapes computed from the symptom network of low resilience participants (left; panel a) and normal to high resilience participants (right; panel b). The stability landscape of the low resilience group goes in a steep line towards a stable point of 10 actively present symptoms. This is above the cut-off value of 8 symptoms or higher, meaning that the anxious phase is stable for the low resilience group. The stability metric for the healthy phase is 0.00 ($SE = 0.01$, 95% CI [0.00, 0.00], $p < .001^5$), and the stability metric for the anxious phase is 0.26 ($SE = 0.11$, 95% CI [0.00, 0.41], $p = .204$). However, the landscape for the normal to high resilience group shows a much wider basin with stability around 5 actively present symptoms, which falls below the cut-off value of the anxious phase. Thus, the healthy phase is more stable for the normal to high resilience group. The stability metric for the healthy phase is 0.06 ($SE = 0.11$, 95% CI [0.00, 0.38], $p = .422$), and the stability metric for the anxious phase is 0.00 ($SE = 0.01$, 95% CI [0.00, 0.03], $p < .001$). The stability difference for the low resilience group is -0.26 ($SE = 0.11$, 95% CI [-0.66, -0.13], $p = .001$) and for

⁵ As the variance of the bootstrapping samples is extremely low, the p-value is not very meaningful here.

the normal to high resilience group is $.06$ ($SE = 0.11$, 95% CI $[-0.05, 0.33]$, $p = .573$). The difference between the two groups is 0.31 ($SE = 0.16$, 95% CI $[0.11, 0.99]$, $p = .002$).

Figure 6. The landscapes of two groups. The cut-off value for the anxious phase is 8 symptoms or more (the dotted vertical line). Panel (a) shows the stability landscape for the low resilience group. The stable point is on 10 actively present symptoms, which falls within the anxious phase. Panel (b) shows the stability landscape for the normal to high resilience group. The stability landscape shows a much wider basin with a stable point that falls within the healthy phase



The NCT reports no significant differences in overall structure ($M = 0.51$, $p = .98$) nor global connectivity ($S = 0.85$, $p = .69$). As the test found no global differences, we did not look further into potential local differences between specific edges.

Conclusion and Discussion

In this section, we demonstrated the use of stability landscapes of networks to compare groups and provided an empirical illustration that can serve as an example for future research. We found that the anxious phase was stable for the low resilience group, while the healthy phase was stable for the normal to high resilience group. However, the NCT did not reveal differences between the two estimated networks. Thus, comparing networks of two groups based on their underlying stability landscape may yield different results than the NCT. This is because the stability landscapes are computed based on Ising networks, which consist of both connectivity and threshold parameters. The NCT can be used on Ising networks, but it only compares the networks on their structure (i.e., edge weights). There may be differences between groups that are captured within the threshold parameters but not considered by the NCT. Contrarily, the stability landscapes for Ising network model networks are calculated from both the edge

weight parameters and threshold parameters. As a sensitivity analysis, we computed the stability landscapes from the two networks while holding the thresholds constant over the networks. In other words, we computed the stability landscapes using the two networks with different connectivity parameters but using the threshold parameters from the low resilience group in both models. In this case, the landscapes were much more similar (see Appendix C). This explains potential differences between the NCT and stability landscapes. We recommend using both to complement each other.

General Discussion

In this article, we proposed a novel method to quantify the stability of Ising networks in psychology using the generalized stability landscape function. The generalized stability landscape function represents the stability of the network's different states (i.e., the sum of actively present nodes). The lower the potential, the more likely that the network will end up in that state – just like a ball on a landscape that tends to roll down to a lower place. The main purpose of the method is to provide a quantitative, intuitive representation of the stability landscape of psychological networks. Instead of only studying the stability of the current phases of the system using simulations (Lunansky et al., 2024), the method presented in this paper computes the full stability landscape of the system as a whole. This allows us to understand the system's phases and their relative stability, as well as compare the stability of the system across groups.

We connected the computation of stability landscapes with the network theory of psychopathology (Borsboom, 2017) and demonstrated how the potential function can be used to quantify the stability landscape of psychopathology networks estimated from cross-sectional data that consist of Major Depressive Disorder (MDD) symptoms. As an illustration, we showed how variations of the depression symptoms network lead to different stability landscapes. Furthermore, we proposed a set of stability metrics from the landscape function.

Finally, we used an empirical example to demonstrate how different stability landscapes can be empirically estimated from cross-sectional data. We estimated networks of PTSD symptoms in two groups with different levels of resilience, as measured by the Brief Resilience Scale. Afterward, we estimated the corresponding stability landscapes and showed how group-level differences in resilience indeed indicated differences in stability landscapes. By also considering the threshold parameters of the networks, comparing groups by their stability landscapes provides an added value next to existing network comparison methods, such as the

Network Comparison Test (van Borkulo et al., 2022) As such, the proposed method is a meaningful complement to the current network analysis toolbox.

The current method holds several assumptions that we would like to clarify. First, the method uses sum scores as a representation of the network. For example, using the number of depression symptoms as an indicator of depression severity. As such, it does not differentiate between specific active symptoms. However, specific symptoms can have different thresholds for activation, for example, depressed mood may be more easily activated than suicidal ideation. But, once active, both nodes contribute equally to the symptom severity. We believe this approach is an acceptable simplification for symptom networks of several specific disorders. However, this assumption may not necessarily hold for networks with more differentiated symptoms. For example, networks consisting of combinations of depression symptoms and anxiety symptoms, bipolar disorder symptoms, or networks with nodes that extend beyond mere symptoms, such as risk or protective factors.

It is also important to note that we used the Ising network models estimated from cross-sectional group data to represent psychological systems and construct landscapes. This means that conclusions can only be made at the group level. If one is interested in interpretations on the individual level, cross-sectional models are only meaningful when the group is homogeneous enough, or in other words, with high ergodicity (Fisher et al., 2018; Molenaar, 2004). Therefore, all the limitations of cross-sectional models for individual inferences will also apply to the landscapes constructed therefrom. One possibility to overcome this limitation would be the development of idiographic methods to estimate Ising models from longitudinal, individual data, which are currently lacking. A second possibility would be to compute stability landscapes from other types of network models which are estimated from individual data, such as Vector Autoregression (VAR) models. However, these models are linear and cannot account for multistability, which is only possible in nonlinear models (Haslbeck et al., 2022; Haslbeck & Ryan, 2022). But even if methods to compute stability landscapes from idiographic network models were developed in future research, it would not be straightforward to collect the necessary data from an individual to compute their stability landscape. The data collection period would need to span a long period of time with enough variability to assess not only the individual's current phase but also potential alternative phases that are captured by the landscape. An alternative way forward would be to combine cross-sectional and idiographic approaches. Cross-sectional methods, such as the proposed approach in this paper, could identify typical networks and stability landscapes that are characteristic of averaged healthy or unhealthy phases. Idiographic data collection would then be used to identify the type of

network and stability landscape that best describes specific individuals. Currently, it remains an open question how this could be done exactly, but we believe it would make a highly interesting novel research line to investigate.

We see several topics for future research that could further the presented approach. First, instead of defining a clinical cut-off value between a healthy or dysfunctional state beforehand, one could, in theory, use the *shape* of the stability landscape to this aim. For example, if one has data from a general population and the stability landscape shows bistability, the area that separates the two stable states could contain a meaningful cut-off value that separates the healthy sample from the clinical sample. The simulations with adjusted empirical networks in the current paper did not find such bistability, but future research could focus on finding an example of a dataset that leads to a bistable stability landscape and determine the optimal method to use the shape of this landscape for diagnostic purposes.

Second, the empirical validation of the presented method is of course a pressing topic for future research. An interesting topic would be further investigating the relationship between the stability metrics and specific network structures. In the current study, we found the landscape difference is mainly associated with node thresholds instead of the strength of network connectivity. This may be different in another research context. Researchers may also look into the relationship between the stability metrics and other variables, for example, treatment outcome. In this way, we can gain more understanding of the meaning of the system's stability and its practical implications.

Finally, we used symptom networks as illustrations, but the use of this method is not restricted to symptom networks. In principle, the same method can also be used for any other Ising networks for which the aforementioned two assumptions hold. For example, to understand how attitudes (feelings, beliefs, and behaviors) evolve. The Causal Attitude Network (CAN) model uses the Ising network model to explain how people move from negative attitudes about something or someone towards a more positive attitude, or vice versa (Dalege et al., 2017). It could be interesting to compute the stability landscapes of attitude networks to understand how stable positive or negative attitudes about, for example, politicians are (Dalege et al., 2017). Similarly, while we used IsingFit (van Borkulo et al., 2014; van Borkulo & Epskamp, 2023) to estimate Ising network models throughout the current article, the landscape construction method we propose is not exclusive to this estimation technique. Alternative approaches for estimating Ising network models have been suggested recently, including nonregularized and multivariate estimations (Brusco et al., 2023), as well as methods for correcting selection bias (Boot et al., 2023). Despite the differences in

estimation methods, as long as the resulting network model remains an Ising network, the structure of the Hamiltonian remains consistent. Consequently, Ising networks derived from other estimation techniques can be employed for landscape construction using the same methodology outlined in this article. With this proposed method, we are one step closer to understanding the complexity and dynamics of psychological systems.

Chapter 6

Examining the Research Methods of Early Warning Signals in Clinical Psychology through a Theoretical Lens

This chapter is based on:

Cui, J., Olthof, M., Hasselman, F., & Lichtwarck-Aschoff, A. (2025). Examining the research methods of early warning signals in clinical psychology through a theoretical lens. *BMC Psychiatry*, 25, 261. <https://doi.org/10.1186/s12888-025-06688-5>



Abstract

The past few years have seen a rapid growth in research on early warning signals (EWSs) in the psychopathology domain. Whereas early studies found EWSs to be associated with sudden changes in clinical change trajectories, later findings showed that EWSs may not be general across variables and cases and have low predictive power. These mixed results may be explained by the diverse methods employed in clinical EWS studies, with some of these approaches and practices potentially misaligned with the underlying theory of EWSs. This article employs a variety of methods, such as a narrative review, mathematical derivations, simulations, and visual illustrations, to support our claims, explain specific assumptions, and guide future empirical research. This multitude of methods serves our aim to provide theoretical as well as methodological contributions to the field. We identify the following key assumptions for EWS validation studies: the system departs from a point attractor, EWSs appear before the critical transition, and EWS variables align with system destabilization. The literature review shows that the common research practices in the field are often not in line with those assumptions, and we provide specific suggestions corresponding to each of the assumptions. More rigorous empirical evidence is needed to better validate the existence of EWSs in clinical sudden changes and fully realize their clinical potential. As theory-based prediction tools, EWSs require stronger alignment between theory and practice to enhance both theoretical understanding and predictive accuracy.

Introduction

Clinical change is not always linear and gradual. In fact, about half of the clients undergo one or more sudden changes in their symptom trajectory over the course of treatment (Helmich et al., 2020; Shalom & Aderka, 2020; Tang & DeRubeis, 1999). Those sudden changes are difficult to predict, yet have important clinical implications because they are often closely linked to treatment outcomes (Helmich et al., 2020; Shalom & Aderka, 2020). Recently, researchers have proposed to use early warning signals (EWSs) to predict sudden changes in clinical trajectories (Dakos et al., 2012; Lenton, 2011; Scheffer et al., 2009, 2012). EWSs are a group of statistical indicators depicting the instability of a complex system, that are generally observable before a system changes to a new stable state. Intuitively, the existence of EWSs can be understood by the following idea. Different *phases*¹ may emerge from the interactions among interdependent biological, psychological, and sociocultural factors within a complex person-environment system. In psychopathology, those phases may represent mental suffering (e.g. depressed mood, anxiety, panic), but also a healthy, well-functioning phase (Cui et al., 2023; Granic et al., 2018; Hasselman, 2022; Hayes & Andrews, 2020; Hayes & Yasinski, 2015; Olthof, Hasselman, Oude Maatman, et al., 2023). Often, such phases are relatively stable, which means that the system may be stuck in a certain phase, having difficulties disengaging or escaping from it. For example, people with depression may be stuck in a phase characterized by depressed mood, reduced sleep quality, and difficulties in concentration (Hayes & Andrews, 2020; Holtzheimer & Mayberg, 2011). Yet the stability of phases may decrease over time, making it easier for the system to escape. The destabilization of the unhealthy phase, which corresponds to the loosening of undesirable cognitive, affective, or behavioral patterns (Hayes & Andrews, 2020), then functions as a mechanism for the sudden gain toward a healthier phase. Likewise, the destabilization of a healthy phase may function as a mechanism for the sudden loss toward the unhealthy phase. When a certain phase is destabilized, the system state is more likely to fluctuate, and after perturbation, the system takes longer to recover. As a result, various early-warning indicators, such as increasing variance (associated with stronger fluctuations) and autoregressive coefficients (associated with slower recovery), can be observed in the data.²

¹ Here we use the term “phase” to represent a collection of system states that are related to a system attractor and have similar qualities. Other literature may use different terms for a similar concept, whereas we aim to make clear the differences between a single state and a collection of states and stay consistent with our previous work (Cui et al., 2023, 2025).

² A related, well-studied concept is emotional inertia, which is also operationalized as high autocorrelations (specifically in affective states). Conceptually, emotional inertia refers to the persistence of feelings across time and contexts, or the inability of affective states to be adaptively regulated to baseline after perturbations, and a large body of empirical findings has suggested that high emotional inertia is related to psychological maladjustments and various mental disorders (Koval et al., 2015; Koval & Kuppens, 2024; Kuppens et al., 2010). Yet recently, Koval and Kuppens (2024) proposed to understand emotional inertia as EWSs for sudden loss based on both conceptual and empirical evidence. Both emotional inertia and EWSs can show the inability to return to the stable attractor (also known as the adaptive home base in some literature) after perturbations and can be assessed through autocorrelation in individual time series. Moreover, most consistent findings regarding emotional inertia were found by comparing healthy individuals with different vulnerability levels instead of

Therefore, we may use the EWSs evidenced in measurements of the mental state of a client to predict whether the client is likely to have a sudden change soon.

Several studies have found empirical evidence that EWSs may exist for clinical changes (Olthof et al., 2020; Wichers et al., 2016, 2020) and hypothesized that EWSs can potentially be used for detecting vulnerable individuals, determining the timing of interventions, and predicting the direction of change (Helmich et al., 2021; Schreuder et al., 2022). Later studies, however, found the predictive power of EWSs to be generally weak, EWSs only occurring in some variables different for each client, and overall, not showing a clear consistent pattern (F. M. Bos et al., 2022; Curtiss et al., 2023; Dablander et al., 2023; Helmich et al., 2022; Schreuder et al., 2020). Those mixed findings may arise from individual differences or measurement techniques, but it is also likely that they stem from methodological issues (Helmich et al., 2024). Considering the relative novelty of the field, various research methodologies and paradigms to study EWS coexist. Some studies monitor symptom levels with frequently repeated measures (Olthof et al., 2020; Wichers et al., 2016), and some studies only use pre- and post-assessments to evaluate whether a change has occurred (Curtiss et al., 2023; van de Leemput et al., 2014); some studies use emotion items to calculate EWSs (F. M. Bos et al., 2022; Curtiss et al., 2023; Wichers et al., 2016), and others use treatment process measures, for example, therapy progress and insight (Olthof et al., 2020). Among this variety of methods, some might be more suitable than others, yielding more reliable results. Assessing these methods requires a thoughtful examination of their alignment with the foundational theory of EWSs. While some studies and reviews have discussed several potential issues of research methods in this field (Dablander et al., 2023; Helmich et al., 2024), we identified several important facets that have received limited attention. To address these points, a rigorous evaluation is needed to establish a direct link between the mathematical theory of EWSs and specific research practices in psychopathology. Therefore, in this paper, we aim to formulate recommendations for improved research methodology in clinical EWS studies based on a mathematical derivation of EWSs in multivariate dynamic systems.

comparing people with and without mental disorders, or people with different severity levels of mental disorders. Therefore, high emotional inertia is more likely to be a warning sign for healthy individuals instead of a general indicator of mental illness. Taken together, emotional inertia may be a special case of EWSs indicative of maladaptive transitions. In other words, emotional inertia can be used as an indicator of vulnerability in healthy individuals, but in general, emotional inertia may be better conceived of as a signal of change instead of a maladaptive characteristic. We would like to refer readers interested in the relationship between emotional inertia and EWSs to Koval and Kuppens (2024), for more comprehensive discussion. Although there is a close link between the two concepts, most studies under the term emotional inertia are based on interindividual comparisons instead of studying whether emotional inertia increases within an individual. Consequently, those studies aim to detect who is likely to have mental disorders instead of when a person is likely to have a sudden change. In the current article, we focus on the studies under the term EWSs, whose primary focus is to predict whether a sudden change is likely to happen in the near future within an individual.

Methods

The current article is a theoretical and methodological contribution. We apply multiple methods to support our claims, evaluate the common research practice in the field, and provide practical suggestions for empirical researchers.

We first provide a mathematical, theoretical derivation of EWSs in general multivariate dynamic systems, which clarifies the original EWS theory that psychopathology borrowed from complexity science. We show the mathematical derivations in Appendix D1, and in the main text, we explain the gist of the derivations with verbal language, visual illustrations, and simulations (with the simulation details in Appendix D2). From the derivation, we identify three key assumptions that should be considered when designing EWS studies: the system departs from a point attractor, EWSs appear before the critical transition, and EWS variables align with system destabilization.

After that, we elaborate on each assumption and discuss their implications for research in the psychopathology domain. For each assumption, we investigate whether the research practices in previous empirical studies were in line with those assumptions and provide suggestions for future studies. We used narrative reviews considering the field of EWS studies in psychopathology is still in its initial stage and the number of empirical studies is rather limited. Some other methods were also used to better explain each specific assumption. For the first assumption, we introduce a new visual check method, namely the distance plot, to assist researchers in examining if this assumption is met. We also include brief simulations and empirical examples for this method. For the third assumption, we provide a brief literature review of the EWS investigations in other scientific fields to illustrate the difference in research methods between psychopathology and other disciplines.

Finally, we provide an integrative summary and discussion of our findings.

Results

The Formal Theory of EWSs

Unlike other verbal theories in psychopathology (Robinaugh et al., 2021), the theory of EWSs has a formal background rooted in mathematical derivations, enabling the analysis of its key assumptions. The formal background of EWSs is based on bifurcation theory, which explains how a gradual change in a system parameter may lead to a qualitative change in the functioning of a system (Gilmore, 1993; Haken, 2011; Scheffer et al., 2009; Thom, 1975; Zeeman, 1976). We first use the case of a cusp bifurcation as a conceptual explanation, which is a rather simple scenario from bifurcation theory. Figure 1a (first row) shows a ball on a landscape with

two local basins in different configurations (columns 1-7). The position of the ball represents the *state* of the system, the basins represent the *phases* of the system³, and the color represents the altitude of the landscape, showing the *stability* of the system⁴. In clinical cases, for example, the x -axis may represent mood, and the y -axis may represent sleep quality. The right basin could correspond to the healthy phase (i.e., high mood and high sleep quality), and the left basin to the depressive phase of a client (i.e., low mood and low sleep quality). When the basin is deeper (represented by darker blue colors), the phase is more stable; when the basin is shallower (represented by lighter blue and green colors), the landscape is shallower, and the phase is more unstable.

The shape of the stability landscape of the system is often determined by one or more control parameters. By adjusting the control parameter (which can be, e.g., alleviating financial stressors, improving social contacts, or making progress in psychotherapy), the landscape changes smoothly from the first column to the last column in Figure 1a, resulting in a destabilization of the left, depressive basin and a stabilization of the right, healthy basin. At a certain point (termed the bifurcation point), the left basin no longer exists, causing the ball to abruptly move to the right basin, this occurs in the second to last column in Figure 1a. This represents a qualitative transition in the state of the system, for example, a sudden gain in treatment. Before the transition, although the state of the system (i.e., the position of the ball) does not change much, the phase's stability (i.e., the depth of the basin) decreases. As a result, when the system is under small perturbations, it is easier for the system to move to another position (i.e., the ball moves further away from the equilibrium point), and it is harder for the system to recover (i.e., the ball returns more slowly to the equilibrium point). For example, after experiencing positive events, the client becomes more joyful than before and does not return to the depressed state as quickly, which represents the destabilization of the depressed phase.

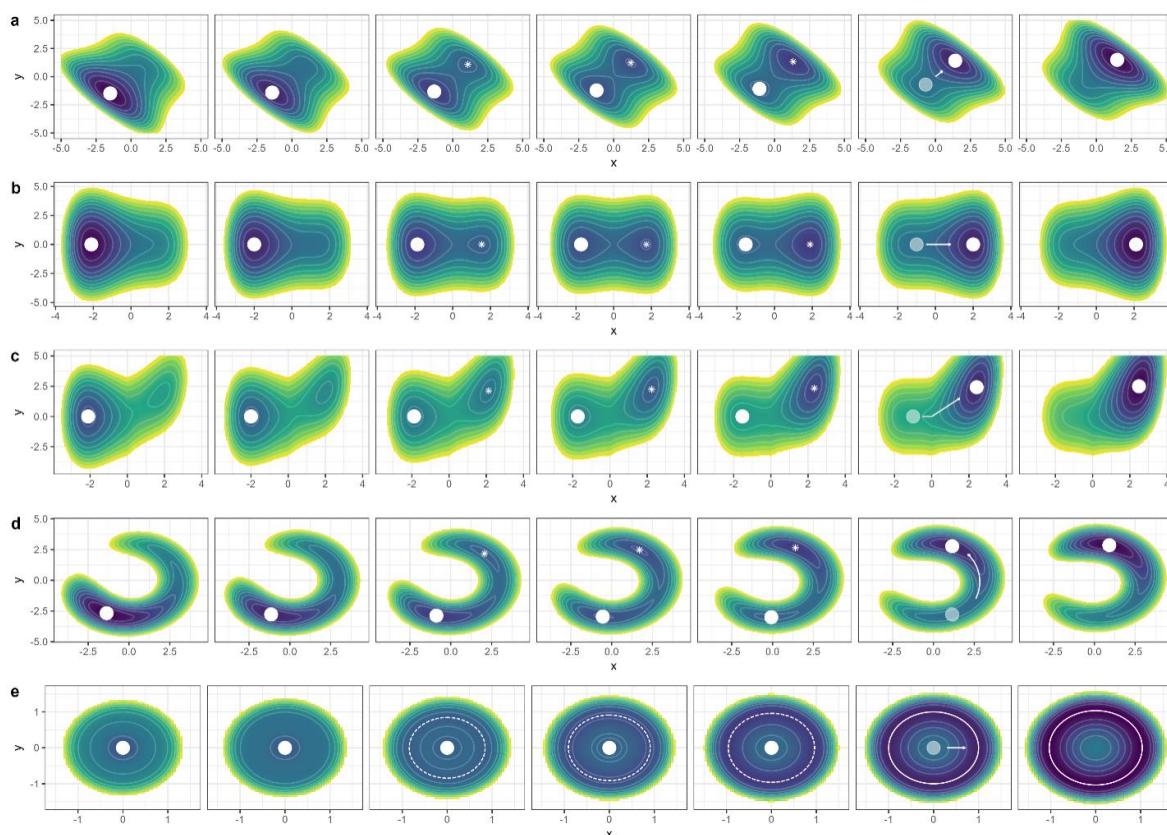
For real-life systems, it is often hard to tell how stable a phase exactly is (i.e., the altitude of the ball's position is not easy to measure). However, if we observe that the state of the system (i.e., the position of the ball) exhibits increasing variance and increasing autocorrelation, we can infer that the stability of the phase is decreasing, and a critical transition may happen in the near future. Take a client with depression for example, it is difficult to exactly determine the client's stability of the current depressive phase, yet we can

³ To avoid confusion, the term “state” is used in this article for the specific condition defined by the values of the system variables, and “phase” is used for the higher-level patterns of the system, which consist of many states that are qualitatively similar (Cui et al., 2023).

⁴ Note that the definition of “stability” in this article may not be the same as some other works in this field (e.g., the time that a system spends to go back to its local minimum after a small perturbation, Dablander et al., 2023). See Cui et al. (2023) for the relationships between different stability measures based on a stability landscape.

assess how the mood level and sleep quality of this person change over time and infer the stability of the person's depressive phase based on such statistical information.

Figure 1. Different examples for bifurcation-induced transitions. Colors that are darker blue represent deeper regions of the stability landscape, and therefore stabler system states. Points and solid circles represent the actual state of the system, and asterisks and dashed circles represent the attractors that are not occupied by the system. For each example, the tipping point is at the second column from the right, represented by arrows. (a) The direction of stability loss and the transition involve both the x - and the y -axes. (b) The direction of stability loss and the transition involve only the x -axis. (c) The direction of stability loss only involves the x -axis; the transition starts along the x -axis but later involves the y -axis. (d) The direction of stability loss involves both the x - and the y -axis; the transition starts involving both the x - and the y -axis, but the x -value does not change much after the transition. (e) The direction of stability loss involves both the x - and the y -axes; the new attractor is not a point attractor but a limit circle, and both the x and the y values overlap with the previous point attractor.



EWSs do not only exist for the simple cusp bifurcation. Many real-life systems are similar to the cusp bifurcation in the sense that one basin of the system disappears at the tipping point

(Scheffer et al., 2012), which means that there is at least one direction in which the system can tip over to a new phase.⁵ Mathematically, approaching the bifurcation point is asymptotic behavior, that is the autocorrelation of the system state will approach 1, and the variance of the system state will approach infinity (assuming small random perturbations). In real-life observations, it is of course not possible to observe this asymptotic behavior (i.e., we cannot use the local stability characteristics of the system to infer the dynamic properties when the system is *infinitely* close to the bifurcation point, and we cannot observe a system that is *infinitely* close to the bifurcation point while keeping the system from transitioning to another phase). Hence practically, the aforementioned conclusion implies that when the system is close *enough* to the bifurcation point, the primary factor that drives the changes in variance and autocorrelation is the destabilization of the current phase (i.e., the flattening of that valley in the potential landscape). The destabilization thus causes the increase in variation and autocorrelation, overshadowing the influence of other factors. This means that if we observe an *unusual* increase in variation and autocorrelation in a client's symptom severity ratings (for instance calculated in a moving window), it is more likely that this increase is driven by the destabilization of the client's landscape and that a sudden change is about to happen.

Previous researchers have provided mathematical proof for the presence of EWSs in one-dimensional systems (Scheffer et al., 2009) and multidimensional systems that can be sufficiently described by the tendency of descending along a landscape without involving curling forces (Dablander et al., 2023). Yet, real-life psychological systems are likely to be general multidimensional systems, for which the simplified assumptions above may not hold. In Appendix D1, we conduct a mathematical derivation for multivariate systems in general, which is more realistic for real-life systems in the psychopathology domain and enables us to draw implications for multivariate research. As the details of mathematical derivation will not be informative for all empirical EWS researchers, we try to explain the gist of the theory in verbal form in the main text, with several specific scenarios shown with ball-and-landscape illustrations (the rows in Figure 1). Those illustrations are all variants of the cusp model we depicted in the first row, Figure 1a, with details explained in Appendix D2. Using verbal descriptions and specific examples instead of mathematical derivations for general cases inevitably trades rigor for readability. Therefore, if a step is logically not strict enough in the verbal explanation or the generalizability of the specific scenarios is questioned, we refer the interested reader to the mathematical proof in Appendix D1.

⁵ In mathematical language, this means the dominant eigenvalue at that equilibrium becomes zero.

We start by introducing the assumed theoretical object of study as a multivariate stochastic dynamic system. This implies that there are multiple (cognitive, emotional, biological, social, etc.) variables that describe the mental state of an individual, and those variables have deterministic influences on each other, yet all those variables are also perturbed by random noise (e.g. everyday events). We also assume that for an individual diagnosed with a mental disorder, the dynamic interactions between the variables create a stable, system-wide attractor state that can be labeled as pathological (e.g., a depressive state) and that the system's behavior remains in this stable state, even after some perturbations (Hayes & Andrews, 2020). Consequently, it is difficult for the individual's mental system to move far away from this pathological state. To simplify the subsequent discussion, we use the phrase the *strength of attraction* to refer to the strength of the pull exerted on the mental system of the individual to remain in the stable pathological state and the difficulty with which the system moves away from this stable state under perturbations.

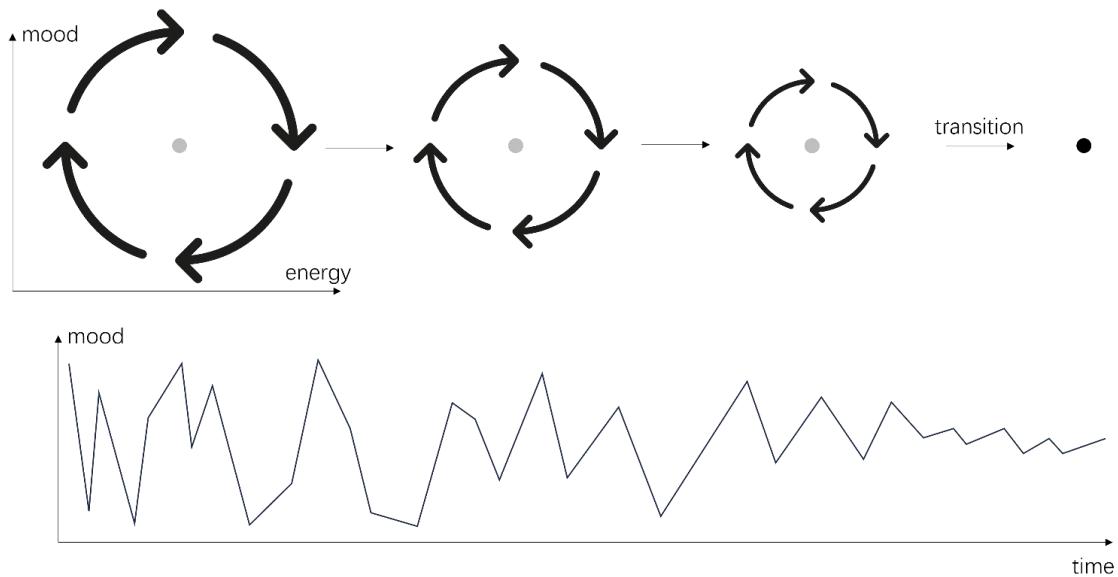
If the strength of attraction remains high, the system is unlikely to escape from the pathological state. This must hold for all directions, which means that a small perturbation in *any* variable or any combination of variables cannot drive the system far away from this stable state. However, when the strength of attraction gradually decreases and approaches zero, a small perturbation can take the system away from this pathological state and make it transition to another state, which might be a healthier one. Just prior to the transition, the system still has the tendency to go back to the old stable state, but this tendency is weaker, so the speed at which the system returns to the old stable state is slower, which leads to an increase in the variance and autocorrelation of the system, known as the EWSs.

Here we emphasize the first important assumption of EWSs in multivariate systems: *the system should experience bifurcation-induced tipping, in which the system is attracted by a point attractor that loses its stability after the transition*. This means that before the transition, the system has the tendency to move to the single most stable *point*, but this tendency becomes weaker and weaker before the transition. Only then do variance and autocorrelation show the instability of the system. In a clinical scenario, a point attractor may correspond to a certain stable level of depressive symptomatology of a client. Sometimes the client may feel a little better or a little worse, due to all kinds of everyday events, yet the client always quickly returns to the same baseline level of depressive symptoms. Over the course of treatment, the depressed phase gets destabilized (Hayes & Andrews, 2020), which makes the client's depressive symptoms fluctuate more. Also, little moments of feeling better may last longer and longer (i.e.,

the return time to the depressed state increases). Such increased fluctuations and return times are then EWS indicating that a sudden change will happen soon.

In all the ball-and-landscape scenarios in the rows of Figure 1, the system starts from a point attractor (such as a stable level of depression). Note that in Figure 1e, the system ends in a cyclic attractor, which is not a problem because what matters is that the system *starts* in a point attractor. After the transition, the system may end in a different type of attractor. This requirement is not to be taken lightly: if the system does not start in a point attractor, the variance and autocorrelation *will not* be indicators of the stability of the system. In Figure 2 we depict a particular situation in which the system starts from a cyclic attractor. In this situation, the system tends to cycle between various states and the variance or autocorrelation of the variables represents the strength of the fluctuation instead of the instability of the fluctuating state. This type of situation may happen when cyclical fluctuations are an intrinsic property of the disorder, such as for rapid switching forms of bipolar disorder or borderline personality disorder (MacKinnon & Pies, 2006). In those clients, according to the model by Kraepelin (MacKinnon & Pies, 2006), mood states are unstable and change rapidly, as well as energy levels (see Figure 2, in which the horizontal and vertical dimensions represent the mood and energy of a client). In the beginning, a client may be in a high energy and manic state, then the energy of the person runs out, leading the client to a low energy and manic state. After that, the client goes to a low energy and depressed state, which leads to the cumulation of energy, and the client goes back to the high energy and manic state. The recovery of those clients manifests as the stabilization of the mood state and energy level, which means their mood and energy states become less extreme (i.e., the radius in the cyclic attractor of Figure 2 becomes smaller). Thus, one would expect to find decreases in variance and autocorrelation of mood and energy measures in the case of successful treatment. Hence, the decrease of variance and autocorrelation does not mean that the client's phase of bipolar disorder is becoming more stable because the attractor state of a (rapid switching) bipolar disorder, in terms of mood and energy variables, is not a point attractor. The system does not have the tendency to return to a single stable point but quickly alternates between different states. Therefore, it is not suitable to calculate EWSs for those clients based on mood and energy measures to use them for the prediction of sudden changes in their psychopathology, at least not from the methodological framework presented in the current paper.

Figure 2. An example scenario without EWSs. For clients with rapid-switching forms of bipolar disorder or borderline personality disorder, the system transitions from a cyclic attractor to a point attractor. Thus, the variation and autocorrelation in mood measures represent the strength of fluctuation of the cyclic attractor instead of the stability of a point attractor.



The second assumption is that *the increase in variance and autocorrelation should appear closely before the transition*. This is already implied in the general concept of EWSs as the heralds of imminent change. The mathematical derivation that the variance will approach infinity and autocorrelation will approach 1 holds *only as* the system approaches the bifurcation point. As explained above, for real-life observation, this means EWSs overshadow other factors when the system is sufficiently close to the tipping point. If we calculate the variance or autocorrelation far before the transition, they are not able to pick up on rising instability among all other causes of variance and autocorrelation. If we calculate the variance or autocorrelation of the system during the transition, or after it has already transitioned to the new basin, the variance and autocorrelation of the variables are not, or not exclusively, related to the stability of the *previous* phase and may actually be relatively high as a consequence of the shift itself, instead of the rising instability. Taking the example of sudden gains in depression again, if we want to infer the stability of the depressive phase, we should not include observations of the healthy phase in calculations because anything calculated therefrom would become a representation of both phases instead of only the destabilization of the depressive phase. Therefore, it is only meaningful to use EWSs to predict a transition that is *about* to happen.

Note that we frame this assumption mainly for empirical research testing whether and when EWSs exist for psychological sudden changes. Practically, it would not make sense to calculate EWS after knowing that the transition has happened. But using EWS as a predictor in clinical practice is only possible after the phenomenon of EWS is sufficiently understood for clinical changes (i.e., after knowing when we can use what EWS to predict which kind of sudden changes for whom), which, according to our evaluation, is not fully established at the moment. It could be possible that, with sufficient empirical evidence, in the future, we know that a certain amplitude of EWSs is very likely to predict a clinical sudden change, which can then be used for practical purposes. Yet, this kind of empirical evidence needs to first be cumulated through studies with sound methodology, which is the core argument of the current article.

Besides the two assumptions mentioned above, there are some additional considerations for multivariate systems. For those systems, a key point is that a state is only stable if it is stable in *all* directions, but it is unstable if it is unstable in *any* direction. We again use the ball-and-landscape metaphor to describe this idea. Note that we can only show two variables, yet for real-life systems, there might be a much larger number of variables creating a multidimensional landscape. For a certain real-life transition, there are some directions⁶ in which the system loses its stability, and those directions may involve one or more variables. In Figure 1a, this direction is to the upper right corner and involves both x and y , whereas, in Figure 1b, this direction is parallel to the x -axis and only involves the variable x . In clinical cases, those directions may correspond to the symptoms that first start to appear or alleviate (we will provide more detailed examples later). The important point here is that even if it involves only one variable it still induces a leak in the basin, making the system transition out. Thus, our third assumption is that the direction in which the basin becomes flat is also the direction in which the system leaves the basin. Therefore, the direction of EWSs is also the direction of the start of the sudden change. In other words, *the EWSs and the start of the transition involve the same set of variables*⁷. Empirically, this simply implies that if a transition is evidenced in a (set of) variables, then the EWSs should be studied in the same variables, not in other simultaneously observed variables that do not show a transition. Note that the direction in which a system *starts a transition* does not necessarily map onto the straight line from the previous attractor to the new attractor. We illustrate this with Figure 1c, in which the transition is along the x -axis at the beginning, but later also involves another variable y . Nevertheless, if one variable is

⁶ Mathematically, those directions are pointed by the leading eigenvectors.

⁷ Mathematically, this implies that only variables with nonzero loadings in the eigenvector will exhibit variance approaching infinity and autocorrelation approaching 1 as the system approaches the bifurcation point. Detailed illustrations can be found in Appendix D1

involved in the destabilization of the system, then it must be involved in the transition process as well.⁸

Take two clients with depression for example. Assume the first one only has depressed mood as the major symptom (represented by the *x*-axis in Figure 1a), and the second one has both depressed mood and sleeping problems (represented by the *x*- and *y*-axes in Figure 1b or Figure 1c). During treatment, both clients had a sudden gain and recovered. If we would monitor mood and sleeping problems across treatment, we may see a sudden change in mood for the first person, but sudden changes in both sleep quality and mood for the second person. In this case, we would expect the first person to have EWSs in mood before the transition, but not in sleep quality because the first person did not have sleep problems to begin with. For the second person, EWSs may occur in mood or sleep quality or both, because the person's depressive phase may first destabilize along the sleep quality axis or the mood axis. If the person's depressive phase first destabilizes along the sleep quality axis, (i.e., the person first improves in sleep which then also positively affects the mood), EWSs should be found in sleep quality but not mood, because changes in mood follow the initial destabilization in sleep. If the person's depressive phase destabilizes along both axes simultaneously, EWSs should occur in both variables. If it is unknown on which axis the depressive phase first destabilizes, monitoring both variables for EWSs is best. There is no need to monitor other variables that are not a part of this person's change process because these other variables (e.g. appetite) are not involved in the transition.

In sum, our mathematical derivation illustrates three important general assumptions for EWS research: (1) the system starts from a point attractor, and this attractor becomes unstable after the transition; (2) EWSs appear right before the transition, not after the transition or far before; (3) EWSs occur only in the variables in which the (start of the) transition occurs. These three assumptions are vital prerequisites for investigating whether EWSs exist or not as precursors to sudden changes in mental disorders. In the following sections, we further discuss the assumptions in relation to empirical studies and provide recommendations for future research.

⁸ A special case is that the new state may have a similar value on an axis even if the system destabilized along this axis before the transition (e.g., the *x*-axis in Figure 1d). Therefore, if we only observe the system before and entirely after the transition, we may not be able to observe the change in this axis during the transition. This may happen in real life, for example, if the *x*-axis represents the external behavior of a child, and the *y*-axis represents the conflict resolution of parents. If the conflict resolution of parents is inferior, the child may not dare to exhibit externalizing behavior to prevent aggravating the conflicts; when the conflict resolution of parents improves, the child may start to exhibit externalizing behavior; when the conflict resolution of parents improves even more, the externalizing behavior of the child may reduce again to a level similar to the beginning, but now the whole family system is much healthier. In this case, the child's externalizing behavior may show EWSs before the transition even if the externalizing behavior's beginning and end levels are similar. Nevertheless, if we observe the whole trajectory, it would be obvious that the child's externalizing behavior also plays a role and changes through the transition, so our conclusion still holds in this case.

The First Assumption: The System Departs from a Point Attractor

As shown in our derivation, an important assumption of the underlying change mechanism is that the system before the transition is in a point attractor, but after the transition, that point attractor diminishes. Previous studies used different methods for detecting whether a sudden transition has happened, yet they do not always establish whether the transition departs from a point attractor (Table 1). Some studies used the difference in symptom severity scores before and after an assessment period (two measurement occasions) to indicate whether a transition had taken place. Using two measurement points does not provide enough information to investigate which type of attractor (point, cycle, or some other type) the system was previously in. Some other studies used several repeated assessments and examined if the change between successive assessment points exceeded a certain clinical threshold. The problem here is that if an assessment touches the threshold and then comes back to its previous value, it is still counted as a transition (F. M. Bos et al., 2022), even though the system did not enduringly leave the original phase. Helmich et al. (Helmich et al., 2022) used a similar method, but with an additional requirement that the mean level difference before and after the identified transition had to be large enough, making the results more robust. Other researchers (Olthof et al., 2020; Wichers et al., 2020) used change point analysis, which is a group of statistical methods seeking to find transition points in the time series. In general, those methods try to split the whole time series into parts to make the data points in each part relatively stable around their own means, but the mean value may differ significantly across different parts (Cabrieto et al., 2017; Helmich et al., 2021). The advantage of change point analysis is that it takes the whole time series into account, therefore includes more information, and it performs better in ruling out false positives because it only identifies a transition if the new phase is different enough from the previous one and also relatively stable. Change point analysis is powerful in detecting a point-to-point transition because it tries to make the data points stay close to the mean level, corresponding to the point attractor, before and after the transition. However, it may not work well for more complex transition types, for example, the point-to-circle transition shown in Figure 1e. Traditional change point analysis used in most previous studies also only works for single-variate data, although multivariate extensions are available (Matteson & James, 2014).

Table 1. An overview of the transition detection methods in previous studies.

Transition detection method	Examples of empirical studies	Method details
Difference in two assessments (continuous; no criteria)	(van de Leemput et al., 2014) (Curtiss et al., 2019) (Schreuder et al., 2020)	Difference of the HDRS-17 and SCL-90-R scores in the baseline and follow-up assessments
Difference in subsequent assessments above a threshold	(F.M. Bos et al., 2022)	Difference of the IDS-C score in the baseline and follow-up assessments
Change point detection	(Helmlich et al., 2022)	Difference of the QIDS-SR score in the baseline and follow-up assessments An increase of ≥ 6 points in ASRM or QIDS-SR, without such increases in the previous two weeks. (The scores may decrease immediately after the transition.)
	(Wichers et al., 2020)	A change that is more than a threshold determined by DaRCI (Helmlich, 2024) in the SCL-90-R depression subscale, with a stability check that there is a large change in the mean level before and after the transition.
	(Olthof et al., 2020)	Change point analysis (Matteson & James, 2014) of the SCL-90-R depression subscale
		Change point analysis (Lewis & Stevens, 1991) of the item of problem intensity in TPQ

Note. Abbreviations: HDRS-17, Hamilton Depression Rating Scale; SCL-90-R, Symptom Checklist 90 Revised; IDS-C, Inventory of Depressive Symptomatology – Clinician Rating; QIDS-SR: Quick Inventory of Depression Symptomatology-Self Report; ASRM, Altman Self Rating Mania Scale; TPQ, Therapy Process Questionnaire

In summary, if it can be assumed that both attractors before and after the transition are point attractors, we suggest future researchers use change point analysis as a main approach for identifying critical transitions; if a threshold-based method is used, it is important to check if the system is stably residing in the new phase to rule out the influence of a single outlier. We also suggest a more general, descriptive method, namely the distance plot, to accompany the methods described above as a way to visually check whether there is a transition from a point attractor. A similar method based on recurrence plots has been proposed by Viol et al. (2022), yet our method is more specific to check this specific assumption of EWSs. The distance plot is a two-dimensional plot that shows the Euclidian distance between system states at different points in time (Iwanski & Bradley, 1998; Marwan et al., 2007).

$$D_{t_i,t_j} = \|\mathbf{x}_i - \mathbf{x}_j\|, \quad (1)$$

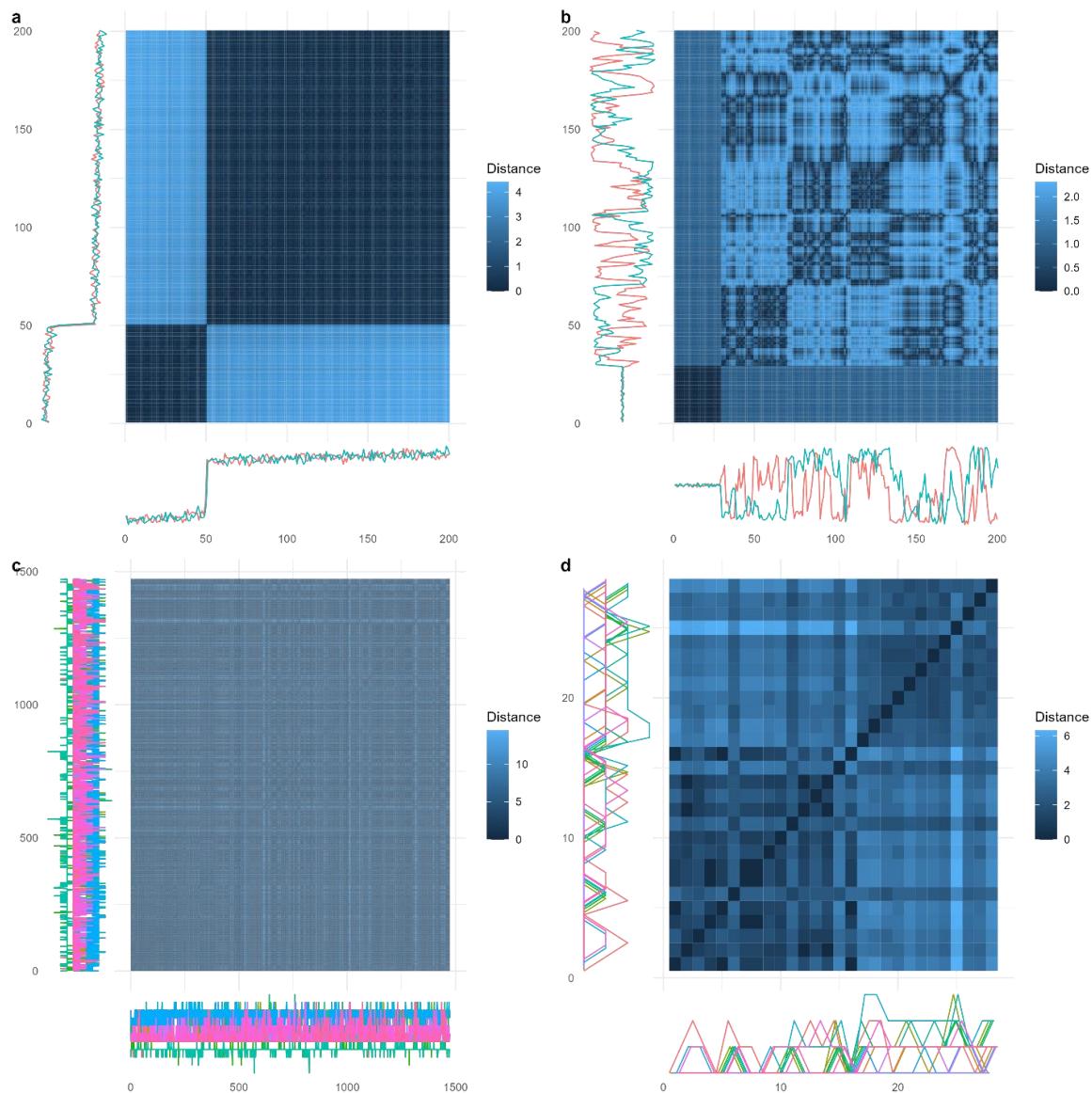
in which the Euclidian distance for two vectors means the square root of the sum of squares of the differences in variable values. For example, if there are four variables assessed over time, and at time 1 the variables take the values of $(1, 2, 3, 4)$, and at time 2 the variables take the values of $(2, 3, 2, 3)$, then the Euclidian distance of the two time points is calculated by:

$$D_{1,2} = \|(1, 2, 3, 4) - (2, 3, 2, 3)\| = \sqrt{(1-2)^2 + (2-3)^2 + (3-2)^2 + (4-3)^2} = 2.$$

If the system is attracted by a point attractor, under a small perturbation, the system will fluctuate closely around this point. Therefore, the distance between each pair of observations should be relatively small. However, after the transition, the system no longer stably moves around this point attractor. Its state is either far from the point attractor or only transiently crosses the point attractor. In a distance plot, both the horizontal axis and the vertical axis represent time, and each pairwise Euclidian distance of points is shown as the color of a pixel. The pixel color in column 1 and row 2 of the matrix, for example, represents the magnitude of the Euclidian distance between the state of the system observed at time 1 compared to the state at time 2. Since the distance between the two points is the same regardless of the temporal order (i.e. time 2 compared to time 1), a distance plot is always symmetric around the diagonal line. If there is a dark region on the plot, it would indicate the time points along it have similar values. In contrast, if there is a light region, it would indicate the time points along it have very different values. Therefore, if the system leaves a point attractor, the distance plot will show a square of dark region before the transition, and a light rectangle of rather far distance next to the square. In Figure 3a-b, we show the distance plots for the bivariate simulated data from examples of point-to-point transition and point-to-circle transition in Figure 1a and 1e (with the raw time series shown along the axes; a description of the simulation procedure can be found in Appendix D2). From the results, it is clear that the method can detect the system

leaving a point attractor, no matter the shape of the new attractor. In Figure 3b, the new phase does not show a square, but different blocks, which indicates a non-point attractor. This is not a problem because what is important in the theoretical assumption is about the attractor *before* the transition, not *after* the transition.

Figure 3. Distance plots for (a) simulated system based on the example in Figure 1a, (b) simulated system based on the example in Figure 1e, (c) ESM affect items, and (d) item scores of the depression subscale of SCL-90-R from Kossakowski et al. (2017).



We also apply this method to an empirical dataset described and made open access by Kossakowski et al. (2017). This dataset is from a client with major depressive disorder (MDD) who completed daily ESM measures up to 10 times per day for 239 consecutive days. Besides

that, the participant also completed the depression subscale of the Symptom Checklist-Revised (SCL-90-R). For illustration, we use the time series of 13 ESM affect items¹ used in Wichers et al. (2016) and the item scores from the depression subscale of SCL-90-R (28 time points; $M = 1.57$, $SD = 0.44$). The distance plots for both datasets are shown in Figure 3c-d. From the results, it becomes clear that the symptom severity of the system assessed by the depression subscale of SCL-90-R may have experienced a transition out of a point attractor, but the ESM affect items did not. Therefore, the critical transition is more evident in the symptom severity scale instead of the ESM affect measures. We will further elaborate on this point when discussing the third assumption. Note that, although differences in time intervals between observations are an important issue to solve for dynamic statistical models, it is of minor relevance for the inferences we can draw from the distance plot as long as the total observation time is sufficiently long to capture a potential transition. This is mainly because the distance plot is used to detect the critical change that is evident from the whole time series instead of dynamic relationships between consecutive time points.

The Second Assumption: EWSs Appear before the Critical Transition

As shown previously, EWSs should appear before the sudden transition and predict a forthcoming transition. This is also a conceptual and practical requirement if we want to identify *early* warning signals to predict future sudden transitions. For empirical studies validating the existence of EWSs, it is also crucial to make sure EWSs are calculated before the transition because the transition itself can also lead to an increase in variance and autocorrelation which is not an *early* warning signal (see Figure 4 for an example and see Appendix D2 for a simulation study). The influence of the transition itself on EWSs persists even after detrending the time series, a standard procedure in EWS calculation (Wichers et al., 2016; the time series in Figure 4 was also detrended before EWS calculation). Although detrending can remove the effect of gradual trends, it is less effective for a sudden change. Looking into previous literature, we found that although some studies explicitly make sure EWSs were calculated before the transitions, it is not always the case for other studies (Table 2).

In order to examine possible EWSs strictly before the transition, the first prerequisite is to rigorously pinpoint the moment of the transition in empirical studies. Some studies calculate EWSs for the whole assessment period, without identifying the exact moment of the transition. For example, the study by van de Leemput et al. (2014) used the difference in depression score

¹ Those items include irritated, content, lonely, anxious, enthusiastic, cheerful, guilty, indecisive, strong, restless, agitated, suspicious, and worries.

before and after treatment (for depressed clients) or observation period (for the general population) to represent whether a transition had taken place or not. Results of this study showed that participants with a greater change in their depression scores also had a larger correlation between emotion scores. However, because depression scores were not measured repeatedly throughout the study period, it was not possible to distinguish between a critical transition and a gradual change (E. H. Bos & De Jonge, 2014), and even if we assume that a sudden change took place, it is impossible to identify the exact time point of that transition. Consequently, if a transition had taken place, the EWS calculation period could have included the transition.

Figure 4. A simulation example of a transition with EWSs. The variance and ACF increase before the transition, which are the true EWSs. However, the transition itself creates an even higher peak in variance and ACF, which are not EWSs. Those increases after the transition may be mistakenly taken as evidence of EWSs if the time window of calculation is not strictly before the transition. See Appendix D2 for details of the simulation setup.

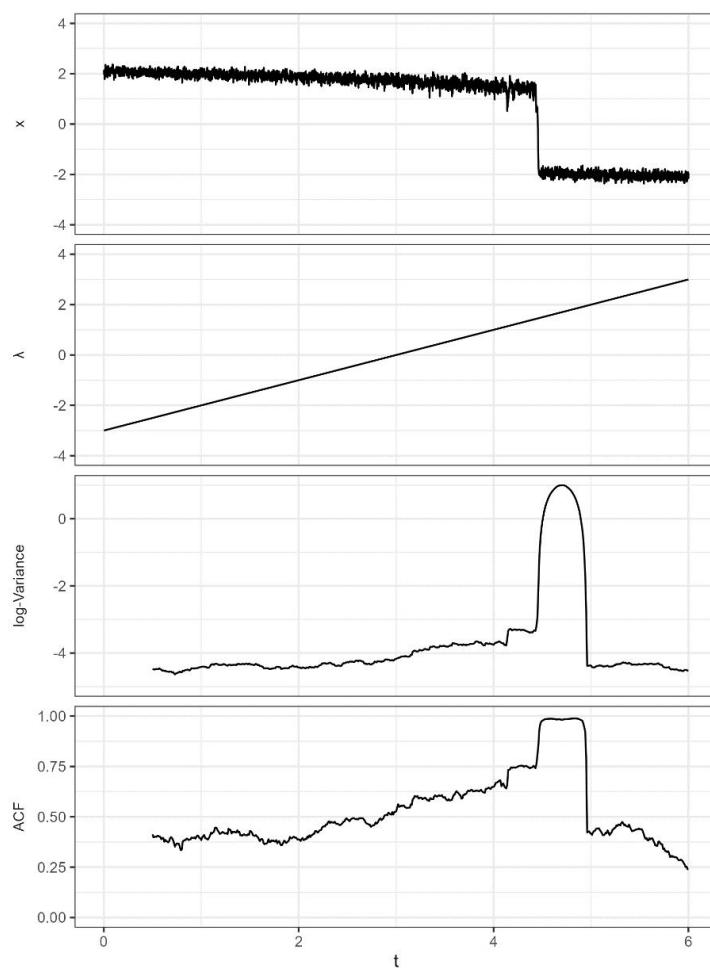


Table 2. An overview of assessment periods of early warning signals in previous studies.

EWS calculation period	Examples of empirical studies	Details
Around the transition	(Curtiss et al., 2019, 2023; van de Leemput et al., 2014)	During the whole assessment; between-participant comparison
Roughly before the transition	(F. M. Bos et al., 2022)	Before the time that the transition was detected, including the week before the sudden increase
Strictly before the transition	(Helmich et al., 2022; Schreuder et al., 2020; Wickers et al., 2016, 2020) (Olthof et al., 2020)	Before the time that the transition was detected, excluding the week before the sudden increase Before the time the transition was detected (the variables used for transition detection were assessed daily)

For studies in which the timing of the transition is identified, it is still possible that the EWS calculation period involves the transition. For example, in the study by F. M. Bos et al., (2022), the variables used for detecting transitions were measured weekly, but the data used for calculating EWSs was measured daily up until the transition point. Therefore, the time window for calculating EWSs may also include the transition point because the transition may have happened anytime during that week. In that case, the transition itself would lead to an increase in variance and autocorrelation. The EWSs detected may then not be true EWSs, but rather, a statistical byproduct of the transition. For many other studies using the same assessment frequency (F. M. Bos et al., 2022; Helmich et al., 2022; Schreuder et al., 2020), the week before the detected transition was excluded from EWS calculations. In those cases, you move further away from the transition point and lose information by excluding the prior week but at least it is certain that the EWSs were calculated strictly before the transition. For the studies using the same assessment frequency for all variables (Olthof et al., 2020), there is no such pitfall, so it is relatively straightforward to make sure EWSs are calculated before the transition.

For future research that aims to *validate* the existence of EWSs, our suggestion is that EWSs should always be calculated in a period preceding the transition. At the same time, we do not recommend discarding data before and close to the transition or assessing EWSs too early before the transition because EWSs are the most salient right before the transition. Therefore, we suggest the best approach to take is trying to detect the transition with high time precision, preferably as frequently as all other measures in the study. When that is not

possible, researchers should try to make sure that the EWS assessment period does not include the transition, even if it is further apart from the transition.

We would also like to note that the implication of this assumption for EWS *validation* studies is naturally different for the research that aims to *apply* EWSs in clinical practice. The application of EWSs in clinical practice is ideally undertaken after the predictive power of EWSs is validated to a certain degree. To our knowledge, most of the studies in the past decade are validation-type studies, and we suggest given the current status of the field, further validation studies are still needed before real-life applications. Therefore, our investigations in this section are mainly intended for validation studies. In validation studies, often the whole observation period is recorded and analyzed afterward which means that the transition point is known and can be pinpointed in the time series. Once EWSs have been validated we can move to the application scenario in clinical practice. Here, the patient's mental state is monitored in real time not knowing when a transition is going to happen. When EWSs are detected (e.g., alleviated variance and autocorrelation of the time series exceeds a threshold), the therapist can take certain measures to prevent an undesirable transition or to promote a desirable transition. This action should be taken, theoretically, before the transition, thus what is observed is that the patient is still in the previous phase but showing EWSs of an upcoming transition. In this case, EWSs are naturally calculated before a possible transition to meet this assumption.

The Third Assumption: EWS Variables Align with System Destabilization

As shown in the previous derivation, the variables that show EWSs are the same ones that point to the direction of destabilization. In other words, it is the same set of variables that show EWSs and that are involved in the start of the transition. This does not mean that the variables that show sudden transitions *always* have preceding EWSs (see Figure 1c and Boerlijst et al., 2013), because EWSs only show the direction in which the system loses its stability at the *beginning*. After leaving the previous stability basin, it is possible that the system changes its direction and involves more variables. Nevertheless, as there is not enough evidence, to our knowledge, to specify which variables are likely to only show transitions without EWSs, it is more reasonable to calculate EWSs and detect transitions for all variables of interest. In other words, there is no reason to *exclude* the variables that have sudden transitions from the calculation of EWSs, while there are good theoretical reasons to include them (see Appendix D2 for a simulation study showing the possible consequences of excluding variables with sudden transitions from EWS calculations).

In previous studies on clinical EWSs, however, we found that it is very common to exclude the variables for sudden change detection from the EWS calculation (Table 3). Actually, no studies that we found in the field of clinical psychology have any shared variables for detecting EWSs and sudden changes. Some of them used different scales for the two sets of variables. For example, in the study by Wichers et al. (2016), SCL-90-R was used for detecting sudden changes, whereas the ESM affect measures were used to detect EWSs. As shown previously, the ESM affect measures of this dataset did not show a clear transition during the assessment period, which means that the sudden change was not evident enough from the ESM measures of the client. The EWSs calculated for those variables are therefore possibly unrelated to the destabilization before the transition in symptom severity.² Some studies used a single scale for both assessments but excluded the variables that were used for detecting sudden changes for calculating EWSs (Olthof et al., 2020). This kind of exclusion is unnecessary and makes it difficult to tell if the variables that undergo the transition also show EWSs. In Table 3, we show several examples in other fields (e.g., physics, ecology, and movement science). Actually, we did not find any studies in other fields, from which EWS studies in clinical study drew much inspiration, that separate the variables into two sets, one for transition detection and another for EWS calculation. Therefore, we suggest future researchers consistently use the same variable(s) to detect transitions and EWSs.³

² Wichers et al. (2016) conducted two sets of assessments of the client, one is ESM measures of momentary affects, and the other is weekly, retrospective measures of symptom severity. Whereas the symptom severity time series showed a clear sudden change, there was no clear sudden change in the ESM measures. Wichers et al. (2016) claimed that they found evidence of EWSs in ESM measures, but based on our argument, it is unclear whether it is related to the sudden change in symptom severity. As the number of assessments of symptom severity is too few in this dataset (15 data points before the transition) and typically considered unsuitable for EWS calculations, we could not examine whether EWSs were present in the weekly symptom measures.

³ There are several methods designed specifically for multivariate EWSs (Weinans et al., 2021) and some of them (e.g., PCA-based approach, Olthof, Hasselman, Aas, et al., 2023; Schreuder et al., 2022) have been used for psychological data. Those methods also have the potential to be incorporated into the theoretical framework we proposed. Although providing derivations for those methods is out of the scope of the current article, we encourage future research to take this further step.

Table 3. An overview of the variables used in EWS studies in clinical psychology and other fields of complex systems.

Same variable(s)	Examples of empirical studies	Variable(s) of sudden change	EWS
No			
	(van de Leemput et al., 2014)	HDRS-17 or SCL-90-R	Variance, ACF, cross-correlations of ESM measures of four emotions
	(Wichers et al., 2016)	SCL-90-R	Variation, ACF, cross-correlations of ESM measures of affect
	(Olthof et al., 2020)	The item of problem intensity in TPQ	The dynamic complexity of all other variables in the TPQ
	(Curtiss et al., 2023)	QIDS-SR	Variance, ACF, cross-correlations of 10 items from the PANAS
	(F. M. Bos et al., 2022)	ASRM and QIDS-SR	ACF of 17 items of momentary mood and symptoms
Yes (examples from other fields)	(Kramer & Ross, 1985)	Light absorption	Variance of light absorption of different experiment trials
	(Tredicce et al., 2004)	Laser intensity	The time that laser intensity reaches the second phase
	(Dakos et al., 2008)	Temperature and deuterium concentration	ACF of temperature and deuterium concentration
	(Meisel & Kuehn, 2012)	EEG channel activity	Variance of EEG channel activity
	(Kelso et al., 1986)	Phase modulus of finger movements	Variance of phase modulus of finger movements

Note. We did not find empirical studies in clinical psychology that use the same set of variables for sudden change detection and EWS calculation. Therefore, we included several studies in other fields, selected from Scheffer et al. (2012), as examples. Abbreviations: HDRS-17, Hamilton Depression Rating Scale; SCL-90-R, Symptom Checklist 90 Revised; TPQ, Therapy Process Questionnaire; QIDS-SR, Quick Inventory of Depression Symptomatology-Self Report; ASRM: Altman Self-Rating Mania Scale; PANAS, Positive and Negative Affect Schedule.

Determining which variables to measure and how to measure them is a non-trivial question, especially in the field of ESM (Fried et al., 2022). Take the affect measures used in Wichers et al. (2016), items such as “I feel down” are in the literature used as a symptom, emotion, affective state, or mood measure. This blurry distinction in the field of ESM studies makes it difficult to compare studies and look for converging evidence. A general strategy for choosing the “right” variable cannot be given but depends on the research question and the phenomenon of interest. If one is interested in transitions in symptom severity, then symptom severity should be tracked over time to determine the transition and look for EWS. If the focus lies on transitions in mood, then mood is what needs to be measured. Once that more conceptual decision has been made the assessment procedure should be adjusted accordingly, that is the measurement frequency should match the phenomenon of interest and the phrasing of the question should be explicit about the type of process the item is tapping into. More concretely, if the researcher is interested in momentary affect transitions, the measurement frequency should be dense enough to be able to pick up on transitions at this rather fast time scale and the question asked should make clear that respondents are asked to reflect on their momentary affective state (e.g. “I feel down *now*” rather than “I feel down”). Another complicating factor in the field of psychology is that different people may interpret certain questions differently and may vary in their response tendencies (across people and within people over time). For this issue, an idiographic approach, where patients collaborate with their therapists to identify personally meaningful items in their lives, represents a promising strategy (Olthof, Hasselman, Aas, et al., 2023).

Discussion

This article aims to examine the current research methodology in clinical EWSs from a theoretical perspective. The motivation for this examination comes from the variety of methods and mixed research findings in this field and also from the nature of theory-based prediction of EWS studies. In order to find out which study methods and designs are the most suitable for studying clinical EWSs, we first investigated the theory of EWS for multivariate systems using a mathematical derivation. From the theoretical investigation, we identified three key assumptions used during the derivation: (1) the system starts from a point attractor that disappears after the critical transition, (2) early warning signals appear strictly before the critical transition, and (3) the same set of variables are involved both in EWSs sudden changes. Based on those assumptions, we evaluated common practices in recent literature on EWSs in clinical psychology and found that those assumptions were not always met or examined in empirical studies. Finally, we provided suggestions for future empirical studies.

The strength of our approach lies in the close alignment of methods and theory. All the evaluations and suggestions proposed are grounded in the mathematical derivation of early warning signals in general multivariate dynamic systems, drawing from the foundation of bifurcation theory. This imparts a robust framework to guide future investigations. However, we acknowledge the essential need for a critical examination of our theory's real-world applicability. In complex, real-life systems, various forms of changes may exist (Ashwin et al., 2012; Proverbio et al., 2023; Shi et al., 2016). These changes may give rise to different statistical indicators and require different investigation methods, yet they have remained relatively understudied within the field of clinical psychology (Cui et al., 2025). We advocate for a more dedicated exploration of the essence of these transitional phenomena in future research. Only by shedding light on less-explored facets can we ensure that our statistical advancements yield practical utility and relevance.

The three key assumptions presented here are certainly not sufficient to guarantee the detection of EWSs. We highlighted them in this article as we consider them the most salient ones in the current methodological context of the field and thus warrant specific attention. All three assumptions are important and do not have trade-off relationships, thus the best scenario would be to ensure that all three assumptions are met. Practically, however, we can see that researchers may not always be completely certain, especially for the first and the second assumptions which are usually bound to the accuracy of visual checks and statistical inferences. In contrast, the third assumption can be met with certainty by methodological choice. Therefore, we encourage researchers to always check whether the third assumption is met, while for the other two, transparent reporting about related procedures and communicating how likely it is that those assumptions are met might suffice in uncertain situations.

Besides the conditions discussed in this article, several recent studies have pointed out other important conditions for EWSs, such that the noise in the system should be Markovian white noise (Kuehn et al., 2022; Morr & Boers, 2024), that the variables of interest should be sufficiently sampled with adequate accuracy (Dablander et al., 2023; Helmich et al., 2024), that both variance and autocorrelation should increase instead one of them (Ditlevsen & Johnsen, 2010), and so on. While a detailed summary of all the important methodological considerations is out of the scope of the current article, we encourage methodological researchers to continue the investigation of the applicability of those conditions in the psychopathology field and search for ways to improve the detection of EWSs, and we

encourage empirical researchers to actively follow the methodological advances in the field and apply available assumption checking methods in their data.

In this article, we examined whether the current research methodology in clinical EWSs is aligned with the theory, and we provided suggestions according to the theory. EWSs are essentially theory-based techniques, and we would like to emphasize the meaning and value of theory-based prediction. There is no guarantee that a theory-based approach can always predict sudden changes in clinical trajectories successfully or even have better predictive power than methods that are not well aligned with theory. It is possible that another prediction method, for example, a machine-learning-based approach (Wang et al., 2024) or outlier detection methods (Smit & Snippe, 2023), yields a better prediction precision than EWSs. Still, the merit of theory-based predictions lies in several other aspects. First, we can gain a better understanding of the transition process itself and use the results to enhance theory formation. In EWS studies, for example, we start with the theoretical hypothesis that sudden gains and sudden losses in clinical psychology can be conceptualized as bifurcations. From there, we deduct the inference that EWSs should exist with certain assumptions. Therefore, rigorous examination of EWSs within a strong theoretical framework can provide evidence consistent or inconsistent with the hypothesis that clinical transitions are driven by bifurcation tipping. However, because EWSs may also arise in other types of changes, their presence or absence should be interpreted in combination with additional theoretical and empirical considerations (see Cui et al., 2025, for detailed discussion), rather than being taken as definitive evidence. If bifurcation as the change mechanism is supported, some control parameters may exist for such bifurcation-induced changes. Those control parameters have important clinical implications because they are the key reason why a client becomes stuck in a maladaptive phase and, therefore, the key intervention targets. At the same time, they may not be as salient as the symptoms themselves and thus can be overlooked if the treatment only focuses on symptoms. Further research may be conducted to determine the underlying control parameters that determine the stability landscape of the system, which might be related to stressors or vulnerability in general or more specific factors for an individual or their social, economic, or cultural environment (Olthof, Hasselman, Oude Maatman, et al., 2023), and find ways to intervene on the control parameters, thereby benefiting intervention science. Only knowing that a statistical indicator can predict a forthcoming transition accurately without a theoretical reason, on the other hand, does not increase understanding of the clinical change processes. Second, with theory-based predictions, we can understand in which case the prediction method is generalizable, and how reliable it is. A theory states under which

conditions a conclusion can be made, whereas statistical predictors do not specify those. In fact, without a theory, it is questionable if any conclusions based on the sample can be generalized to a new situation (Cartwright, 2009).

In summary, the investigation of EWSs in clinical settings represents a promising research direction that has the potential to significantly enhance our understanding of change mechanisms in mental health and psychopathology and ultimately inform intervention science. While empirical evidence is crucial, it is equally important to prioritize theoretical refinement and methodological evaluation within the field. Such efforts will not only drive theoretical advances but also lay the foundation for a more accurate and responsible use of EWSs in real-life practice, which is the ultimate goal that researchers in this field aspire to achieve in the future. By fostering a synergistic relationship between theory, methodology, and empirical work, we can strive for comprehensive progress and advance the field of clinical EWSs toward improved well-being for individuals.

Conclusion

EWSs are theory-based prediction tools and require a good alignment of the methodology and the underlying theory. Based on our investigation of EWS theories and research practices, we found that the popular research designs in the field do not always align with the theory, leaving room for improvement. We suggest future research in clinical EWSs should be designed so that the transition is visually checked with distance plots, the EWS assessment period is strictly before the transition, and a consistent set of variables is used.

Acknowledgments

We thank Freek Oude Maatman for his valuable input on the conceptual framework of this work, we thank Daniele Proverbio for his helpful suggestions on the theoretical arguments and derivations in this work, and we thank Arjen van Assen for providing valuable insights to help the manuscript establish a better connection with clinical research. We thank the Center for Information Technology of the University of Groningen for their support and for providing access to the Peregrine high-performance computing cluster.

Chapter 7

Understanding Types of Transitions in Clinical Change: An Introduction from the Complex Dynamic Systems Perspective

This chapter is based on:

Cui, J., Hasselman, F., Olthof, M., & Lichtwarck-Aschoff, A. (2025). Understanding types of transitions in clinical change: An introduction from the complex dynamic systems perspective. *Journal of Psychopathology and Clinical Science*, 134(4), 469–482.
<https://doi.org/10.1037/abn0000991>



Abstract

Sudden changes are common in clinical trajectories. While theoretical work in complex dynamic systems has provided mathematical theories for various types and mechanisms of change, a concrete application for the field of psychopathology is still lacking. We aim to bridge this gap by outlining an applied theoretical framework using theoretical concepts of the natural sciences for the field of clinical psychopathology, also devoting attention to issues and providing recommendations that are specific to the psychopathology domain. First, the mechanisms and features of four distinct types of transitions are introduced: bifurcation-induced tipping (B-tipping), noise-induced tipping (N-tipping), rate-induced tipping (R-tipping), and noise-induced diffusion (N-diffusion). Those types of transitions differ in the main cause of the change and data characteristics. To illustrate their application to clinical phenomena, we present two real-life scenarios using simulated time series. These examples demonstrate how theoretical types of change may connect to clinical phenomena and highlight how different types of transitions can co-occur in various subsystems. In the first example, we show that the mood system and the momentary affect system of a patient with sudden loss may show B-tipping and N-diffusion at the same time; in the second example, we show that increasing the stimulus strengthening speed in exposure therapy may lead to R-tipping, while the therapeutic decision in this context may be caused by N-tipping. Finally, we lay out possible pathways for determining the appropriate type of transition for future empirical research, highlighting methods both from dynamic system research and special opportunities for research in clinical psychology.

Introduction

Changes in psychopathology are often not a smooth, gradual process. In psychotherapy, around one-third of patients experience a sudden gain (i.e., abrupt alleviation in their symptom severity; Hayes et al., 2007; Helmich et al., 2020; Olthof, Hasselman, & Lichtwarck-Aschoff, 2020; Tang & DeRubeis, 1999; see Shalom & Aderka, 2020 for a meta-analysis), and patients with sudden gains tend to have better treatment outcomes (Helmich et al., 2020; Lutz et al., 2013; Shalom & Aderka, 2020). Studies on sudden losses (i.e., abrupt deterioration in symptom severity) are rare, possibly because people who are about to have an onset or relapse are seldom included in monitoring studies. Nevertheless, several studies observed sudden losses in the general population (Eckes & Nestler, 2023; Olthof et al., 2024), in patients undergoing treatment (Helmich et al., 2020; Lutz et al., 2013), or tapering (Wichers et al., 2016), suggesting that sudden changes may also be a prominent phenomenon for symptom worsening.

Despite the high prevalence and clinical relevance of sudden symptom changes, the mechanisms behind them are still unclear. Empirical studies found no significant predictors for sudden changes even with large samples and machine learning methods, evidencing that they cannot be explained by demographical or symptom (severity) measures (Aderka et al., 2021; Aderka & Shalom, 2021; Zilcha-Mano et al., 2019). One possible approach to explain such sudden changes comes from the perspective of complex dynamic systems, which conceptualize sudden changes as critical transitions (Hayes et al., 2007; Olthof, Hasselman, Oude Maatman, et al., 2023; Schiepek & Tschacher, 1992; Thelen & Smith, 1998). The interactions of multiple social, behavioral, and biological elements relevant to mental disorders lead to the emergence of separate attractors or patterns. Those attractors may be labeled as healthy or pathological, depending on the person's subjective feeling and level of functioning in such attractors. Each attractor is locally stable, which means that it is hard to transition from a healthy state to a pathological state, and vice versa. Under certain conditions, however, the system may switch from one attractor to another, leading to a sudden change in the state of the system.

While the notion of a transition from one attractor to another can explain *how* clinical change occurs, it does not explain *why* it occurs. In complex systems, causal explanations of change are quite different from linear understandings. Rather than searching for isolatable and single causal factors explaining sudden changes, complex causality acknowledges the large amounts of interdependent elements that are involved in the transition (Lichtwarck-Aschoff & van Geert, 2004; van Geert, 2019). Consider the idiom, the straw that broke the camel's back.

Even if we can identify this last straw, we cannot say it is the only cause of the breakdown as each straw has its own contribution. What we can do, however, is to investigate the mechanisms of change on a more abstract level, looking for a way to conceptualize the essence of the change. Go back to the camel: although we cannot point to a single straw as the causal force, we know that the camel's back broke under the gradually increasing load, whether that comes from straws, clothes, or whatever, and it is easy to tell that this situation is intrinsically different from a camel tripped by a rope or pulled by a fast-racing car. Thus, we can still conceptualize the underlying change mechanism at a more qualitative, abstract level, despite the large number of causal elements involved at the more detailed level.

In this article, we aim to provide a theoretical introduction to the types of changes in complex dynamic systems. We first introduce a classification system of critical transitions, namely B-tipping, N-tipping, R-tipping, and N-diffusion. Some authors have provided systematic criteria to distinguish transition types with detailed mathematical underpinnings (e.g., Proverbio et al., 2023; Shi et al., 2016; Thompson & Sieber, 2011a). We have chosen a more intuitive way, by using ball-and-landscape illustrations (which have been used in many previous studies in natural sciences, e.g., Lamothe et al., 2019; Proverbio et al., 2023; Scheffer et al., 2009; Waddington, 1966, as well as in psychology, Cui et al., 2023; Heino et al., 2022; Olthof, Hasselman, Oude Maatman, et al., 2023), together with simulations, to explain the different transitions types. We focus on the conceptual differences of the transitions in the main text, leaving simulation details to Appendix E1 and the reproducible code (available at <https://osf.io/4jaqk/>). After that, we use several clinical scenarios to illustrate why different types of transitions may happen in the same system depending on the variables and levels of interest. We will also explain why setting clear boundaries for a system of interest is necessary for clarifying the research question about transitions and end with a roadmap and several directions for future research designs.

Types of Transitions

In this section, we introduce the basic types of critical transitions, summarized in Table 2. Readers who are not familiar with dynamic systems theory may refer to Table 1 for the definitions of some key terminologies.

Table 1. Definition of key concepts in the complex dynamic systems theory.

Concept	Definition
Complex dynamic system	A system that consists of a large number of elements that interact in a non-simple way and evolve over time, such that even if we know the properties of each individual part and interaction, it is still non-trivial to infer the property of the whole (Simon, 1962).
State	The property of a system at a certain time point, characterized by relevant variable values.
Parameters	The values or conditions that define the exact form of the dynamic interactions of a system.
Noise	The internal or external forces that affect a system's state in an essentially random way, in contrast to deterministic forces that are predictable given the system's state (Ashwin et al., 2012; Forgoston & Moore, 2018).
Attractor	A state or a collection of states that the system tends to evolve towards from many initial starting points (Milnor, 2006).
Phase ¹ (or basin)	A region in the state space that is close enough to an attractor. If the system is in it, the system has the tendency to go back to the attractor (Cui, Lichtwarck-Aschoff, et al., 2023; Milnor, 2006).
Critical transition (or tipping)	The phenomena that a dynamic system may abruptly transition from one phase to another in a relatively short period (Ashwin et al., 2012; Scheffer et al., 2009).
Ball-and-landscape metaphor	A simplification of certain complex dynamic systems to represent the stability of different states with a low-dimensional, generalized potential function (Cui, Lichtwarck-Aschoff, et al., 2023). In such a metaphor, the position of the ball represents the state of the system, the altitude of the ball represents the stability ² of the system, and the landscape represents the deterministic force of the system.

Note. ¹ Alternative meanings of the term “phase” exist (e.g., in “phase space”). In this article, we follow the definition provided here and use the word “phase” to refer to a set of regions. ² Here by “stability”, we mean thermodynamic stability instead of kinetic metastability. The latter is better represented by the barrier height on the landscape (see Moore et al., 2006).

Table 2. Summary of the features of three transition types.

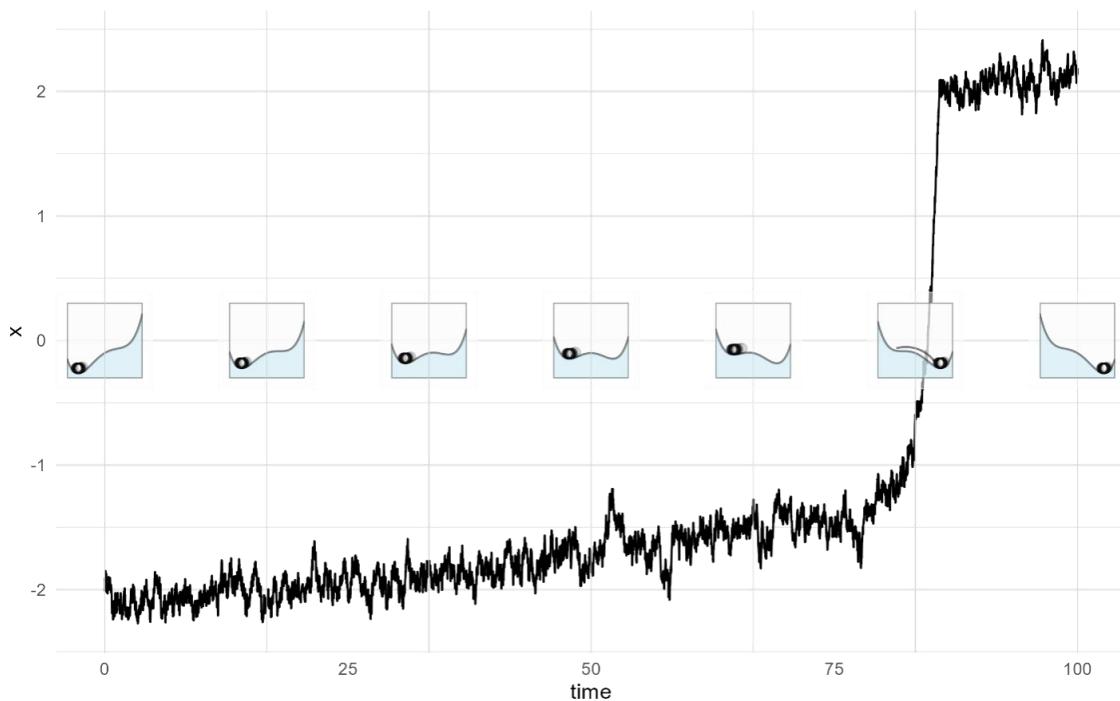
Transition type	Main cause	Predictability
B-tipping	Destabilization of the current attractor	Yes (under conditions)
N-tipping	Sudden, large perturbation	No
R-tipping	Moving rate of the attractor exceeds the ability of the system to follow	Yes (under conditions)
N-diffusion	Constant, large perturbations	N/A (always happening) ¹

Note. ¹ In N-diffusion, the system constantly transitions between alternative basins. Therefore, it is not meaningful to predict whether a transition is about to happen.

The first possible mechanism for sudden changes is destabilization, which is often called bifurcation-induced tipping, or B-tipping (Ashwin et al., 2012, 2017; Boettiger & Batt, 2020; Kuehn, 2011; Thompson & Sieber, 2011b). In Figure 1 we illustrate B-tipping induced by a fold bifurcation, where the system has two basins at the beginning but one loses stability later, making the system transition to the alternative basin. We show the time series together with ball-and-landscape plots that represent the stability of the system. In the ball-and-landscape plots, the two basins on the landscape represent the two alternative phases of the system (healthy versus pathological). The deeper the basin, the more stable the system is. The shape of the landscape is determined by one or more parameters that are called *control parameters*. When the control parameters change gradually, the depth of the basin decreases over time, up until a certain point at which the basin does not exist anymore. This point is called the *bifurcation point* at which the ball abruptly transitions to the other basin, representing a sudden transition. Note that in real-life situations, there is always (small) noise present in the system, causing the system's state to slightly fluctuate around the local minimum. Nevertheless, the noise is weak enough in B-tipping, hence the bifurcation is the main reason for the transition.

B-tipping has prompted the line of studies of early warning signals in psychopathology (e.g., Bos et al., 2022; Helmich et al., 2021, 2022; Olthof, Hasselman, Strunk, et al., 2020; van de Leemput et al., 2014; Wichers et al., 2016). Early warning signals (EWSs) refer to a set of statistical parameters, such as increasing variance and autocorrelation, that appear during the destabilization process and can be used to predict a sudden change (Cui et al., 2022; Gilmore, 1993; Scheffer et al., 2009, 2012; Thom, 1975; Zeeman, 1976). Because the previous attractor of the system becomes less stable, disturbances of the system by random noise will move the

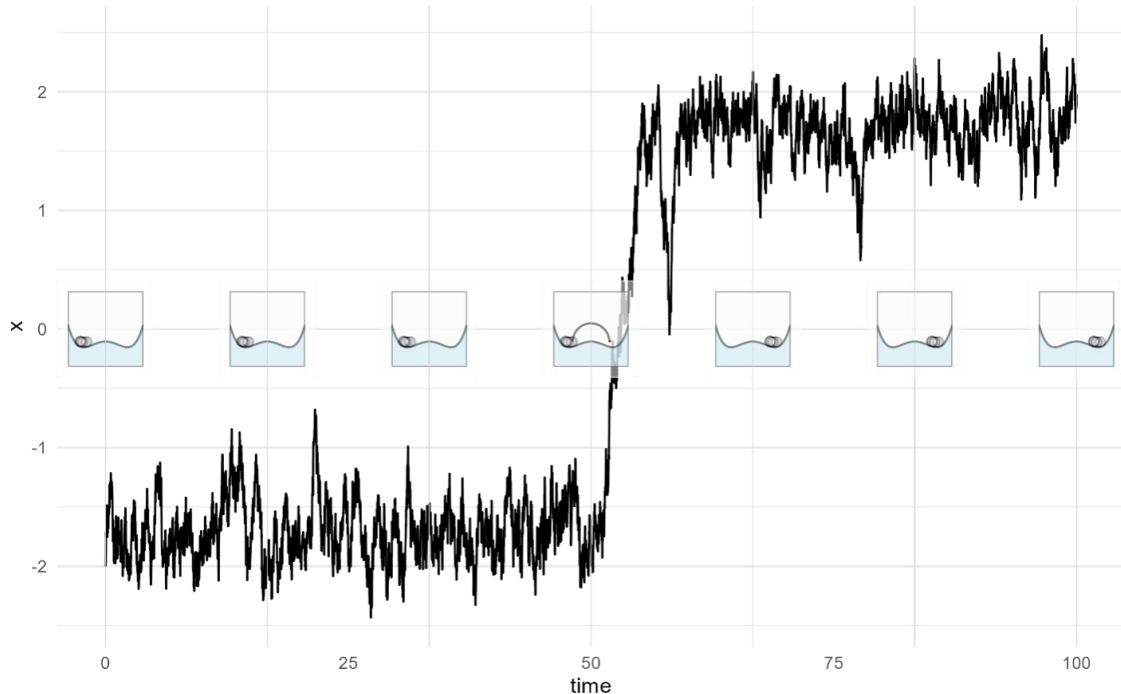
system further away from the attractor point (increasing variance), although still in the basin, and it will take longer to return (increasing autocorrelation). EWSs bear important clinical promises because if a sudden change can be predicted precisely, practitioners may be able to perform just-in-time interventions to facilitate desired and prevent undesired changes (Granic, 2005). The initial enthusiasm is somewhat damped because of the many mixed results found in empirical studies (Helmich et al., 2024). Several explanations may account for this, such that the research methodology has potential issues, that the type of bifurcation does not involve EWSs, or that the clinical change process corresponds to another type of transition (Boettiger et al., 2013; Boettiger & Hastings, 2012; Cui et al., 2022; Dablander et al., 2022; Dakos et al., 2015; Ditlevsen & Johnsen, 2010; Evers et al., 2024; Kuehn et al., 2022; Morr & Boers, 2024; Proverbio et al., 2022). Simulation examples of other bifurcation types are provided in Appendix E2, and we will introduce other types of transitions in the next paragraphs.



The second type of change is noise-induced tipping, or N-tipping (Ashwin et al., 2012; Boettiger & Batt, 2020; Forgoston & Moore, 2018; Ritchie & Sieber, 2017; Van den Broeck et al., 1994). It represents the case where the stability of the system does not change, but the state of

the system moves from one attractor to another due to external or internal noise. A simulation and ball-and-landscape illustration of N-tipping is shown in Figure 2. It is difficult to predict N-tipping before it happens because there is no landscape change before the transition, hence no EWSs can be detected. In other words, there was no destabilization of the attractor state prior to the event but the impact of the sudden event is large enough to let the system jump out of the original basin and transition into an alternative basin. Theoretically, this alternative phase must already exist for N-tipping to occur, which is also hard to see from the time series before the tipping. But the system will visit a region that has rarely been visited before during the sudden transition, which can be taken as a signal that a sudden change is taking place (Shi et al., 2016).

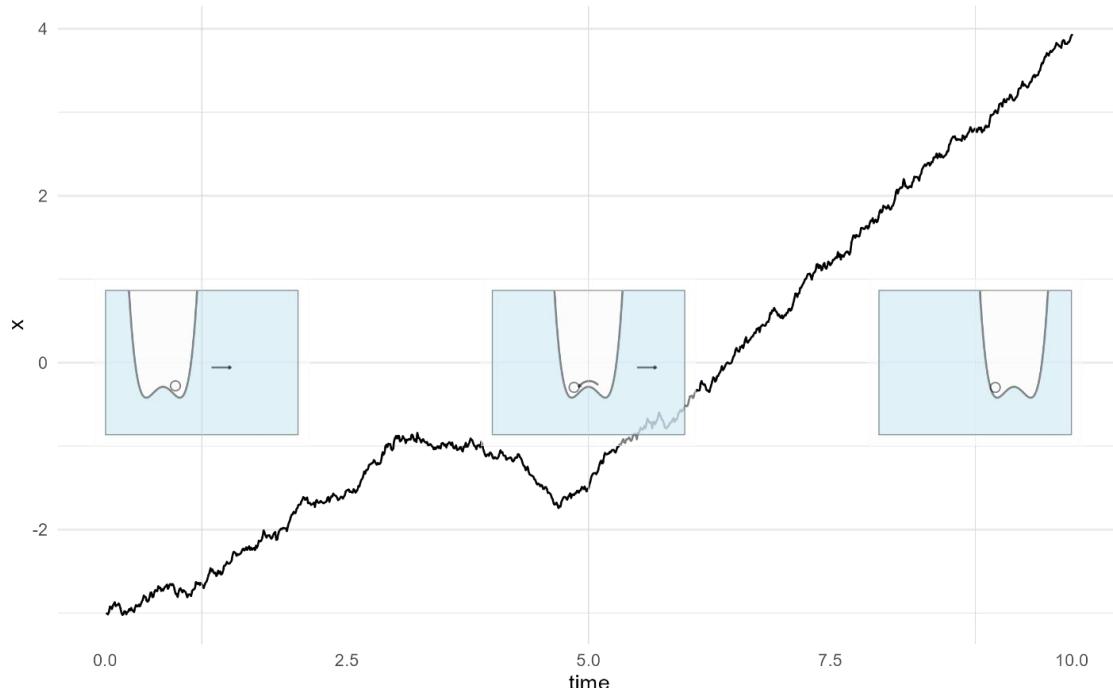
Figure 2. A simulated time series and ball-and-landscape illustrations for N-tipping. The ball-and-landscape illustrations represent the stability of the system at certain time points. Various balls are shown on each landscape to represent the variability of the system's state under noise.



The third type of tipping is the rate-induced tipping, or R-tipping (Ashwin et al., 2012, 2017; Ritchie et al., 2023; Ritchie & Sieber, 2017). It represents the case in which the stability of the system does not change, but the position of the attractor changes so quickly that the state of the system cannot catch up with the (previous) attractor. A ball-and-landscape illustration of R-tipping is shown in Figure 3. R-tipping also has EWSs, such as increasing variance and autocorrelation before the tipping, but they do not occur as early as EWSs for B-tipping and have different implications compared to B-tipping (Ritchie & Sieber, 2016). In B-tipping, the

EWSs before the transition represent the decreasing stability of the former phase; in R-tipping, the increasing variance and autocorrelation show that the system is trying to catch up with the former phase before it transitions to the new phase, leaving the stability of the former phase unchanged.

Figure 3. A simulated time series and ball-and-landscape illustrations for R-tipping. The ball-and-landscape illustrations represent the stability of the system at certain time points. The ball on each landscape represents the real-time state of the system.

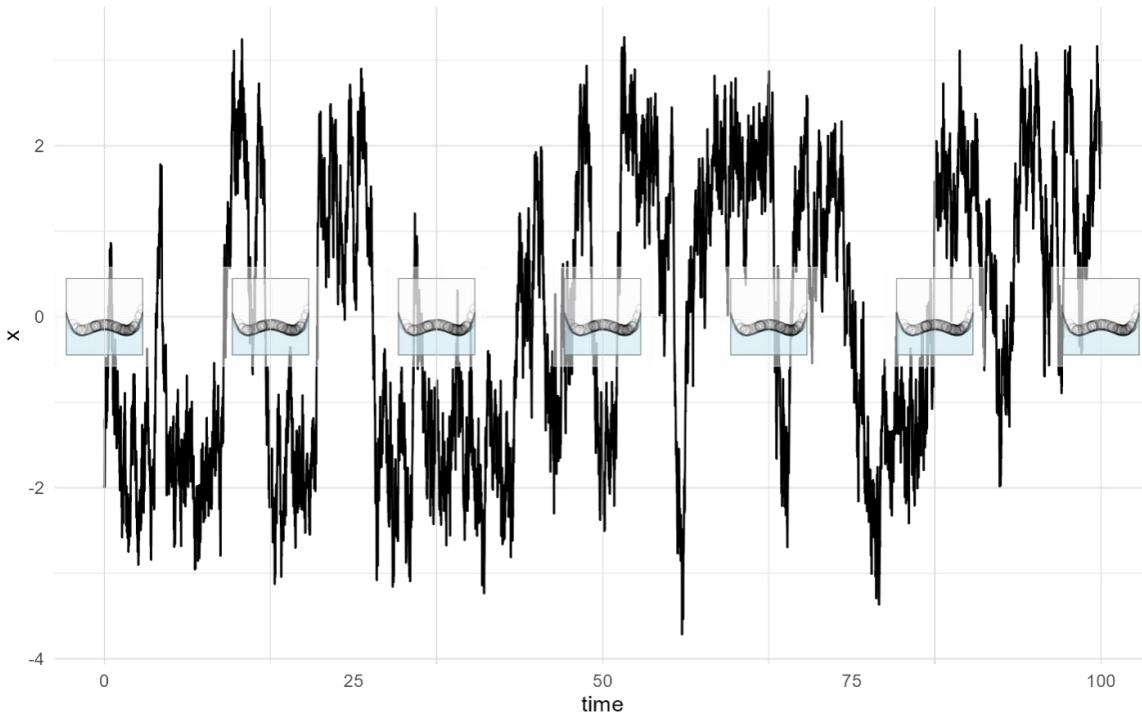


Finally, there is a type of situation that does not involve a single transition from one attractor to another, but frequent transitions throughout the whole state space due to a high level of noise. The stability of states then manifests as the probability density: the more stable the state, the higher the probability density. We call this type of situation *noise-induced diffusion*, or N-diffusion, as there is no single identifiable tipping point.¹ A ball-and-landscape illustration of N-diffusion is shown in Figure 4. Because there is no single transition in N-diffusion, there are no sudden transitions to be predicted. Nevertheless, in a system with high levels of noise, the property of the whole state space can be retrieved from the time series, allowing for the estimation of the landscape and the number of phases for the system (Cui,

¹ The phenomenon we are discussing is referred to as distribution transition or ergodic behavior in previous literature (Proverbio et al., 2023; Shi et al., 2016). In the current article, we have opted to use the term “N-diffusion” to emphasize that it does not involve a single shift in the time series, and to maintain clarity and avoid terminology that might be less familiar to our target audience.

Lichtwarck-Aschoff, et al., 2023; Livina et al., 2010). If a new attractor appears, the system can also quickly move to the new basin and travel back and forth, a phenomenon sometimes referred to as “flickering” (Dakos et al., 2013; Scheffer et al., 2012).

Figure 4. A simulated time series and ball-and-landscape illustrations for N-diffusion. The ball-and-landscape illustrations represent the stability of the system at certain time points. Various balls are shown on each landscape to represent the variability of the system’s state under noise.



In sum, these four types of transitions correspond to three main conditions under which a dynamic system transitions to an alternative phase: (1) decreasing stability of the original phase, (2) noise, and (3) positional change of the original phase. In idealized cases, only one condition is present, being responsible for the transition. Thus, in our idealized cases above, we kept the landscapes and dynamic functions equal and only varied the parameters controlling the speed with which the landscape changes and the strength of noise. These simulations demonstrate that the type of transition is determined by the parameters that are relatively more dominant (see Table E1 in Appendix E1). It is important to emphasize, however, that in real-life scenarios, those conditions may often co-occur, leading to transitions that may exhibit mixed characteristics of different transition types. For example, if the stability of a phase has

already decreased but has not become entirely unstable yet, strong perturbations (i.e. noise) can more easily tip the system to another phase. In this case, the system may experience a mixture of B- and N-tipping. If the stability of the system slowly changes, while the noise in the system is strong enough to cause an N-diffusion, we may observe the system's distribution changes accordingly (Shi et al., 2016). Another complicating factor is that the transition types cannot be differentiated purely on empirical grounds but what can be observed in the data strongly depends on the research focus and interest (i.e., how the system and its boundaries are conceptualized). A point that we will discuss in the next section.

Transition Types in Clinical Psychology

Conceptually, it is not difficult to differentiate between the types of transitions in a psychological context: B-tipping may happen when a slowly changing process leads to a sudden change in the symptom level; N-tipping is triggered by an unforeseeable external event; R-tipping arises when a developmental or environmental process requires a rapid adjustment of the system to maintain stability, but the change occurs too rapidly for the system to adapt; N-diffusion is likely to happen when the system is highly variable and frequently switches between different states. However, difficulties may arise when we link those metaphorical illustrations to real-life scenarios because there are often multiple ways to set the boundary of the system. We explain this with two concrete clinical scenarios, and after that, reflect on the complexity of classifying real-life scenarios at a more abstract level.

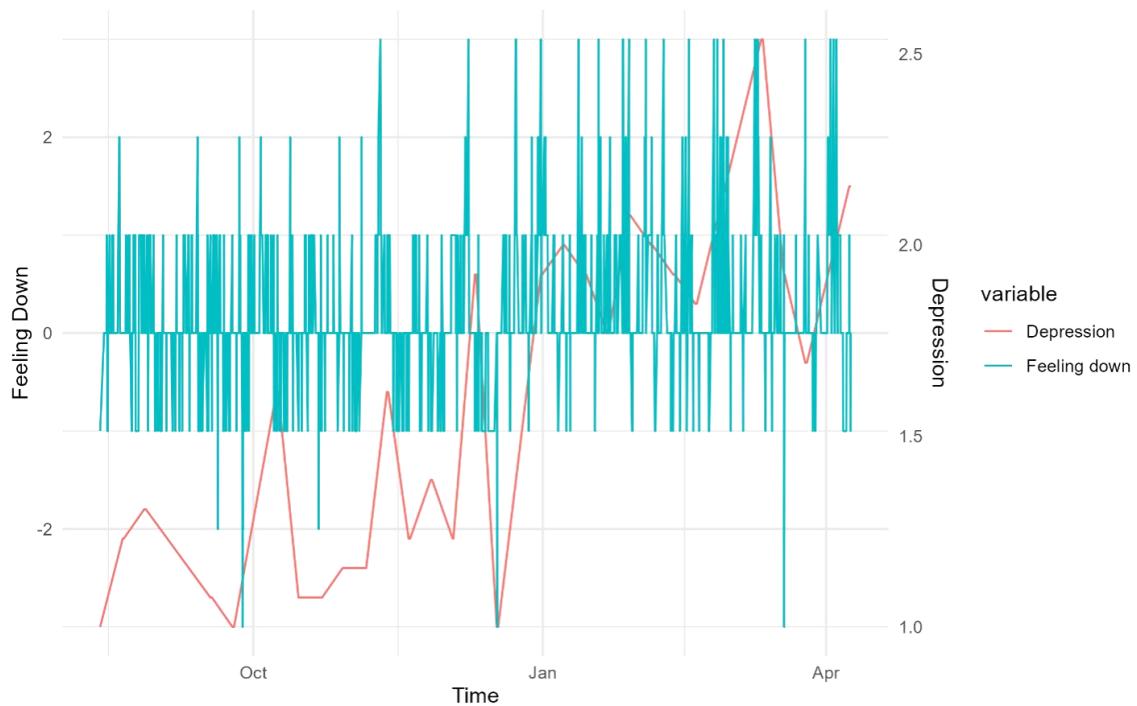
First Scenario: Mood and Affect in Depression Relapse

Mood and affect are closely related, yet different. Whereas mood is more stable, reflecting a more sustained emotional climate, affect is characterized by fluctuating changes, responding to daily events and variations (Alpert & Rosen, 1990). Major depressive disorder is characterized by persistent low mood (American Psychiatric Association, 2022) not necessarily the absence of positive affect. But individuals with depression do differ in the distributional properties of their affect compared to healthy individuals, sometimes showing bimodality (Hosenfeld et al., 2015; Loossens et al., 2020; Rottenberg, 2005; Watson et al., 1988). Due to the differences between the mood and affective system, it is possible that the affective system of a person undergoes N-diffusion, while the same person's mood system may experience B-tipping.

Take the specific example where a patient in the tapering procedure for antidepressants had a sudden loss in the severity of symptoms (Kossakowski et al., 2017; Wichers et al., 2016; Figure 5). Assuming that the sudden transition in depressive symptoms (mood system) is led

by B-tipping, a possible explanation of the mechanism can be the slowly changing neurotransmitter concentration during the tapering procedure, destabilizing the healthy phase and strengthening the depressive phase. At a certain point, the healthy phase is not stable enough anymore, which leads to a relapse, a sudden transition to the depressive phase. However, if we look at the affective system, which is more variable than the mood system, we may reach a different conclusion. In Figure 5, we can clearly see that even before the relapse, the person feels down occasionally. Thus, although the distribution of the momentary affect changed slightly before and after the sudden loss, it did not show a clear transition, resembling the characteristics of N-diffusion.

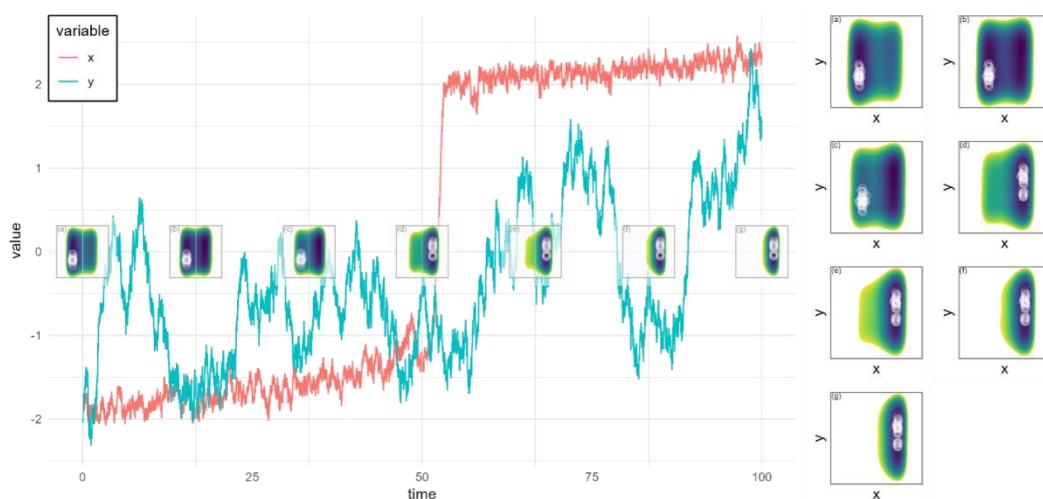
Figure 5. The general mood (depression) and momentary affect (feeling down) of a patient who had a sudden loss in the tapering process. Data are retrieved from Kossakowski et al. (2017).



We simulated the time series according to this scenario, where we used x to represent the depressive mood and y to represent the negative affect. We set the noise level in the variable x weaker than in the variable y and let the stability of the system gradually change along the x -axis so that x undergoes B-tipping and y undergoes N-diffusion (Figure 6). The distribution of y can be influenced by the value of x , that when x is higher, y is more likely to be high. The detailed simulation settings are available in Appendix E1. We consider this to be a realistic relationship between general mood and momentary affect. Comparing the simulated data

with the real-life dataset (Figure 5), we see clear similarities. The variable depression is closer to x and contains a clear transition, although it contains more fluctuations than the simulated variables. The variable “feeling down” is closer to y with much stronger fluctuations. Therefore, we consider the simulated transition types one of the possible explanations for the observed change. The identified type of transition then depends on the relative variability of the variables that researchers are focusing on. If researchers are interested in the slower changing processes (e.g., general mood), they will observe B-tipping; if researchers are interested in the faster changing processes (e.g., momentary affect), they will observe N-diffusion combined with a stability change.²

Figure 6. Simulated time series, where x experiences B-tipping and y experiences N-diffusion. The ball-and-landscape illustrations represent the stability of the system at certain time points. To illustrate the landscape with two variables, we use color to represent the altitude of the landscape, where dark blue represents deeper, more stable positions on the landscape, and light yellow represents higher, less stable positions on the landscape. Various balls are shown in each landscape to represent the variability of the system’s state under noise. The landscapes are also presented as zoomed-in subfigures to enhance readability. Each inset panel (labeled (a)–(g)) in the main figure directly corresponds to its enlarged version on the right.



² This theoretical analysis also has important implications for EWS research, calling into question the previous conclusions on EWSs in this specific dataset, where the authors used the variance and autocorrelation of the affective variables as EWSs (Wichers et al., 2016). As the data of the affective variables have the fingerprint of N-diffusion rather than B-tipping, their (change in) variance and autocorrelations should not be taken as EWSs but may be due to the distribution change associated with the N-diffusion (e.g., when the depth of the two basins becomes closer, the variance and autocorrelation of the N-diffusion variable can also increase). A detailed discussion of EWS research methodology is out of the scope of the current article, and we refer interested readers to Cui et al. (2022).

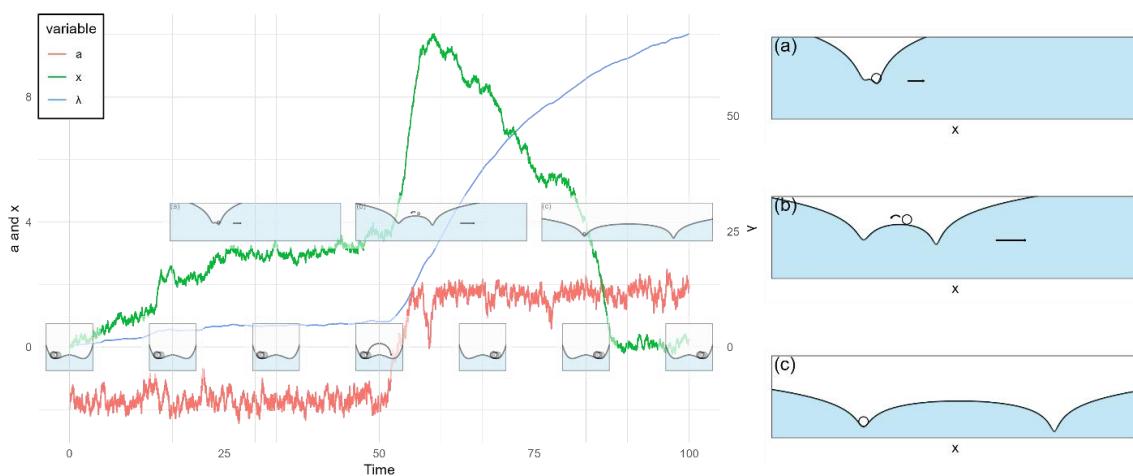
Second Scenario: Gradual and Intensive Exposure Therapy

Exposure therapy is frequently used in the treatment of phobias and anxiety disorders (Craske et al., 2008; Foa & Kozak, 1986; Foa & McLean, 2016; Powers et al., 2010). Multiple types of exposure therapy are available, such as gradual exposure (also known as systematic desensitization) where the strength of the stimuli slowly increases, and intensive exposure (also known as flooding) where the strength of the stimuli rapidly increases to the maximum (Schumacher et al., 2015). The findings about the relative effectiveness of the two approaches are mixed (Kazdin & Wilcoxon, 1976), with some studies reporting a clear superiority of the intensive approach (e.g., Boulogeuris et al., 1971), and others reporting a much higher relapse rate of the intensive approach (De Moor, 1970). Relatedly, a recent study found that gradual exposure in small steps is more effective than gradual exposure in large steps, although patients in the large-step group reached higher levels in the fear hierarchy (de Jong et al., 2023).

To illustrate the possibility of R-tipping, we simulated a scenario where a therapist initially uses gradual exposure, and later changes to intensive exposure, which exceeds the patient's adjustment ability and later leads to a relapse (Figure 7). In the simulation, λ represents the current strength of the stimulus, a represents the speed of increase in the strength of the stimulus, and x represents the patient's tolerance to a specific fear-evoking stimulus. At the beginning, a is small, and λ increases slowly, representing the gradual exposure. At a certain point, the therapist switches to intensive exposure. The detailed simulation settings of this scenario are also available in Appendix E1. At first, the person's tolerance level is only stable at a low value, representing high levels of anxiety. During the treatment, a new form of reaction to the stimulus emerges (Foa & McLean, 2016) and gets strengthened, which means that the patient can increasingly tolerate the threatening object. This is shown as a new basin on the landscape that gradually moves towards high tolerance. The state of the patient (i.e., the actual level of tolerance of the patient) moves along the new basin in the beginning. After the therapist switches to an intensive approach, however, the patient's tolerance level, although initially showing improvement, cannot keep up with the moving speed of the new basin, which leads to a return to the previous basin, manifesting as a relapse (De Moor, 1970). Here, the high speed of the basin's movement is the main cause of the transition, which means x experiences R-tipping. Different from the previous example, the position of the previous basin does not change, but the position of the new basin with higher x moves quickly. Nevertheless, it is still R-tipping as the fast-moving speed of the new basin is the main reason for the transition.

Similar to the previous scenario, if we look closer into this one, we can see that another variable of the system may also experience another type of change because the therapist switching to an intensive approach is also a transition in itself. Several potential causes may drive the therapist's decision to switch, including the patient's factors and the therapist's experience and evaluation (van Minnen et al., 2010). In some cases, the transition in exposure techniques may be a predictable B-tipping, such as when a therapist adjusts their technique based on the patient's recovery speed, and over time, realizes that the progress is slower than expected and decides to switch to an intensive approach. But it is also possible that the switching is caused by some random event (i.e. noise), such as when a therapist unexpectedly meets a colleague who advocates for the intensive approach and persuades them to adopt it. In this case, we would observe the therapist's switch as N-tipping. Here, we show an example of N-tipping in the simulated time series (a in Figure 7). Taking together, in this clinical scenario, the system exhibits R-tipping when the focus is on the patient's tolerance level, and N-tipping if the focus shifts to the therapist's decision. This illustrates that different levels of the system may show different types of transitions at the same time.

Figure 7. Simulated time series, where a experiences N-tipping and x experiences R-tipping. The ball-and-landscape illustrations represent the stability of the system at certain time points. Two sets of landscapes are shown in the figure, representing the stability of a (below) and x (above), respectively. The landscapes of x are also presented as zoomed-in subfigures to enhance readability. Each inset panel (labeled (a)–(c)) in the main figure directly corresponds to its enlarged version on the right.



Reflections on Linking Transition Types to Clinical Phenomena

The two examples above showed that the type of transition is not bound to a specific real-life phenomenon, but more to a specific subsystem in this phenomenon. This is mainly because psychological phenomena are nested across levels and consequently the boundary of the system can be defined in multiple ways (Hasselman, 2023; Noble, 2011). Let us take a closer look at the ball-and-landscape metaphor. This metaphor is useful not only because of its intuitive appeal, but also because it delineates the difference between variable, force, and noise (see Table 1), which is crucial in understanding the types of transitions. The ball represents the state of the system, and the landscape represents the deterministic dynamic function of the system – it tells us in which direction the system is likely to move. Noise is all other forces that influence the state of the system. Thus, if we use one variable x to represent the state of the system or the position of the “ball”, then the deterministic part of its dynamic function, for example, $\frac{dx}{dt} = -x$, shows the deterministic force on the system which can be represented by the landscape³. The ball may change its position (i.e., value) without changes in the dynamic function. In our example, x can be 0, 1, -1, under the same dynamic function, which means that the position of the ball changes, but the landscape does not. When the dynamic function changes, however, it means that the landscape changes. In our example, if now we have $\frac{dx}{dt} = -2x$, we have a steeper valley on the landscape because there is a stronger force pulling the system to the original point. We can also add a noise term to the function, $\frac{dx}{dt} = -x + \varepsilon$. Here, the noise term, ε , cannot be inferred from the landscape. Its force on the state of the system can be along the descending direction of the landscape, or it can be the opposite.

So far, it seems that there is a clear difference between the dynamic function and the value of a variable. However, changing the dynamic function of x can almost always be expressed with the value of another variable, which is called *parameterizing* a function. We can write $\frac{dx}{dt} = -ax$, and change the value of a to change the dynamic function of x . If the aim is to model a real-life system, allowing the dynamic function to change is important because the deterministic dynamic function (i.e., the stability landscape of the system) may change due to changes in the environment of the system. Therefore, the parameter value (i.e., the shape of the landscape) also represents the environmental influences that are external to the system. Further complications will appear if we let a also have its own dynamics and let it be

³ Strictly speaking, the dynamic function has to specify a gradient system to make sure the force can be represented as a landscape. In a one-dimensional system, this always hold, but it may not hold for multivariate systems. In those cases, generalized potential landscapes are available. A detailed analysis of this issue is out of the scope of the current article. We refer the interested readers to Cui, Hasselman, et al., (2023) and Cui, Lichtwarck-Aschoff, et al. (2023).

dependent on x . This makes the boundary between variable and parameter much fuzzier. Unfortunately, this complication is almost always present when we seek to analyze (sudden) changes in clinical psychology (Eronen, 2019). In the first example scenario above, the patient's momentary affect can be taken as the variable x , and the patient's mood as the parameter a . The patient's general mood influences the affective dynamics, but at the same time, the mood of the patient is also influenced by the neurotransmitter concentration, altering during the tapering. Affect and mood could therefore be taken together, constituting the state space of the system, and let their joint dynamics be affected by some other factor (i.e. neurotransmitter levels) outside this system. Similarly, in the second scenario, if we take the patient's tolerance to stimuli as the system, it is influenced by the switching in exposure technique; if we take both the patient's state and the type of exposure together as the system, it can also be influenced by external factors, such as colleagues' recommendations. The nesting does not stop here, because patients and therapists in the region may also form a larger system, that can be collectively influenced by higher-level factors, such as the local policy.

Importantly, we argue that different ways of setting the boundary of the system are simultaneously correct because the difference in the state of the system and the dynamic function of the system depends on the definition of the object system, which can be arbitrarily chosen depending on the research question. Once the boundary of the system is determined, there will be a clear difference between the state of the system and its environment. Note that this arbitrariness does not stem from a particular ambiguity in psychological science, it is a common problem also in natural sciences⁴. Differentiating the system and environment in no way denies the interdependence between the two. On the contrary, it is well accepted that psychological systems are open systems that constantly exchange energy and information with their environment (Bronfenbrenner, 1979). Clarifying the boundaries is necessary to actually link the theoretical types of transitions to specific psychological phenomena. Furthermore, analyzing a phenomenon at different levels does not yield contradictory conclusions – they will yield conclusions at different levels that complement each other.

Roadmap for Future Research

Studies of clinical sudden changes started with clinician's observations and routine outcome monitoring across treatment sessions (Aderka et al., 2012; Hardy et al., 2005; Lutz et al., 2013;

⁴ In physics, the object system is also arbitrarily chosen depending on the research question. Imagine a person standing within a big box, trying to push one side of the box to make it move. Taking the person as the object system, the hands of the person experience an equal and opposite force from the box, according to Newton's Third Law. If, however, we take the box and the person together as the object system, the system does not experience an external horizontal force; thus, its center of mass does not move. There is no right or wrong for those two choices, and both choices lead to valid conclusions. What differs is the focus of the question and the problem to be solved.

Shalom & Aderka, 2020; Stiles et al., 2003; Tang & DeRubeis, 1999). Since then many researchers have investigated sudden changes using intensive longitudinal assessment methods (Helmich et al., 2020; Hulsmans et al., 2024; Olthof et al., 2024; Olthof, Hasselman, Aas, et al., 2023; Olthof, Hasselman, & Lichtwarck-Aschoff, 2020; Wichers et al., 2016, 2020). Yet, studies specifically designed to identify the type of transition are still sparse. In this section, we took inspiration from other fields to provide empirical guidance to identify the type of transitions in clinical psychology. A summary of the possible directions is presented in Table 3.

Table 3. Summary of the possible directions in determining the type of transition in clinical psychology.

Direction	Advantage	Disadvantage
Repeated controlled experiments (in vivo or in silico)	Provide strong evidence of the type of transition and rule out alternative explanations.	Difficult to implement experiments or validate computational models.
Detecting EWSs	More feasible, only the variables need to be monitored, and can be examined in a more natural setting.	Difficult to rule out alternative explanations, strong theoretical or supplementary empirical evidence needed.
(Semi-) qualitative approach	Make better use of the participants' own reasoning and can acquire information from time scales that are unfeasible for real-time monitoring.	Retrospective bias is likely to present, the participants may not be able to understand or answer some difficult questions.

The first general approach is to track changes in control parameters and state variables simultaneously to investigate the influence of parameter changes on the sudden change of the state variables, preferably using experimental manipulations. Many previous studies applied this approach using numeric simulations, which although not empirical, can still shed some light on possible empirical manipulations. For example, Shi et al. (2016) showed that when continuously changing a control parameter, B-tipping only happens at a late time point where the bifurcation point is almost reached, N-tipping happens probabilistically during a long time, and for N-diffusion, the system reaches the alternative basin as long as the new basin appears.

Similarly, if the transition only happens when the parameter changing speed is above a threshold, not bounded to a specific parameter value, then the system is likely to experience R-tipping (Ashwin et al., 2012; Ritchie et al., 2023).

The difficulty underlying this approach is that it requires the researcher to be able to manipulate the control parameter of the system, making sure that it changes in a certain direction and speed, and be able to replicate the experiments multiple times to rule out chance possibilities. In natural sciences, this approach is often applied to highly modellable and controllable small-scale systems. For example, Kramer and Ross (1985) described a chemical system that is bistable given a certain range of incident laser power. The authors recorded the chemical concentration along the process of increasing or decreasing laser power. Because they observed a clear transition in chemical concentration at a later stage of parameter change, which was preceded by critical slowing down, they concluded that the system experiences B-tipping. In psychology, Haken et al., (1985) have conducted similar work for bi-manual coordination, where changing the movement speed can lead to transitions between in-phase and counter-phase movement patterns. Such controlled experiments are rather difficult in a clinical context. The tapering of medication (Wichers et al., 2016) and exposure therapy (Foa & Kozak, 1986) may represent possibilities, but those experiments cannot be exactly replicated, which is a classical problem in psychological measurements (Lord & Novick, 1968). Lab experiments, which are common in the field of computational psychiatry, may provide conditions for control and replication, but whether lab experiments translate to clinically relevant transitions is questionable (Hitchcock et al., 2022; Seriès, 2020).

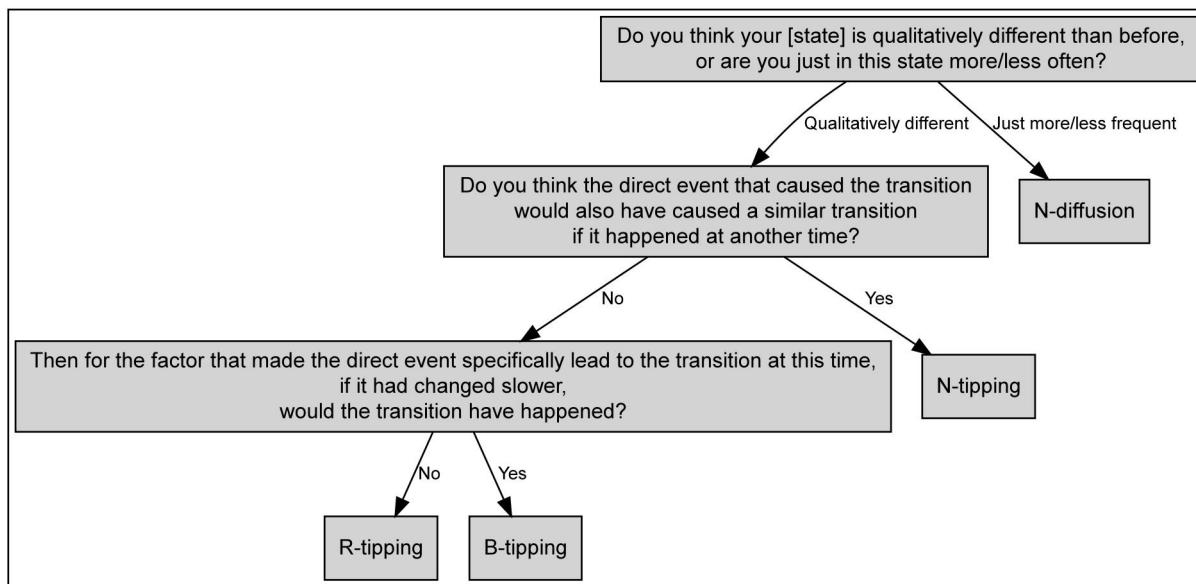
Another possibility involving experimental manipulation is the development of formal dynamic models (Haslbeck et al., 2022; Robinaugh et al., 2021). Just as in ecological models (e.g., Ashwin et al., 2012; Boerlijst et al., 2013; Kuehn, 2011; Weinans et al., 2019), if we have enough evidence to believe that a formal dynamic model is a sufficient representation of actual systems in real life, we can analyze the model through simulations or analytically to understand the type of transition. A substantial difficulty in this direction is that for the formal dynamic models in psychology, the bottom-up building blocks (i.e., the basic function forms and relationships among model elements) are not as theory-informed and empirically validated as in natural sciences. Formal dynamic models are often evaluated by how they resemble real-life transitions, which may induce circular arguments. Therefore, to proceed in this direction, we require stronger support for the validity of formal dynamic models.

The second general approach relies on EWSs as a way to distinguish between B-tipping and N-tipping. For example, early work used the existence of EWSs as evidence that B-tipping

is widely prevalent (Dakos et al., 2008; Scheffer et al., 2009, 2012). Ditlevsen and Johnsen (2010) found no EWSs before the Dansgaard-Oeschger events (a type of rapid climate change), thus concluded those events are probably noise-driven. Similarly, Guttal et al. (2016) found only rising variance but no rising autocorrelation before financial crushes, and suggested the crushes were due to N-tipping led by increasing noise. This approach is clearly more feasible in empirical studies, but an often raised issue is that the relationship between EWSs and B-tipping is not one-to-one, as EWSs may appear without bifurcations, and bifurcations may appear without EWSs (Boettiger et al., 2013; Dablander et al., 2022). One possible solution for this issue is to back up the empirical tests of EWSs with a clear theoretical foundation and descriptive analysis to exclude other types of transitions. For example, if we know that a system departs from a point attractor, the noise is present in the direction of the change, and the stability of the system has monotonic changes, then the system is likely to show EWSs if the transition is B-tipping (Cui et al., 2022). On the other hand, if the noise in the system seems to be so strong that N-diffusion is likely to happen, or if there are clear theoretical arguments to assume that the system departs from a circular attractor (e.g., in the case of bipolar disorder), or if R-tipping is a possible explanation of the transition, EWSs should not be taken as evidence of B-tipping.

The third general approach is that we also see specific opportunities unique to the field of psychology. A key difference between psychology and other disciplines is that our research objects, human participants, have the ability to reason and conceptualize their own lives (Rozin, 2001). Therefore, it may be possible to ask participants counterfactual questions and gain information from there (Pearl & Mackenzie, 2018). For example, after a transition, we may conduct interviews guided by the questions in the decision tree shown in Figure 8 or embed those questions in an ESM assessment routine. This (semi-)qualitative approach is also advantageous as it can assess the processes that are too fast or too slow for traditional ESM designs. For example, Bossenbroek et al. (in preparation) interviewed youngsters in residential youth care and identified various significant transitions in their lives. Some said that the impact of certain stimuli heavily depended on their current state, i.e., the impact was different in other periods of their lives. Information on such a time scale cannot be assessed with here-and-now ESM designs. Of course, such an approach also comes with limitations, such as retrospective bias, researchers' bias, and participants' limited ability of counterfactual reasoning and to remember and evaluate past events (Blome & Augustin, 2015; Morrow, 2005). Nevertheless, the (semi-)qualitative approach still represents a valuable direction.

Figure 8. A decision tree to hypothesize about the type of transition using a semi-qualitative approach. Exact questions may be modified according to the specific research question.



Conclusion

Critical transitions in dynamic systems can be classified into four main categories: B-tipping, N-tipping, R-tipping, and N-diffusion, and possible combinations of those. Different types of transitions are triggered by different causal mechanisms and result in different dynamic patterns in time-series data. Clinical changes in real life involve a large number of variables on multiple levels, with different variables undergoing different types of transitions. The type of transition can therefore not be defined for the change as a whole but is specific to the variables of investigation. To empirically examine types of transition, researchers may apply one of the following directions: repeated controlled experiments (*in vivo* or *in silico*), detecting EWSs, or (semi-) qualitative approaches. Those directions each have advantages and disadvantages, with some being more accurate and rigorous, and others being practically more feasible. We encourage future researchers to continue on this avenue enhancing our understanding of clinical change processes.

Acknowledgments

We would like to thank Freek Oude Maatman for his helpful suggestions on the earlier version of this manuscript, and we would like to thank Yannick Wuppermann for conducting the literature review about sudden losses.

Chapter 8

Examining the Feasibility of Nonlinear Vector Autoregressions for Psychological Intensive Longitudinal Data

This chapter is based on:

Cui, J., Lichtwarck-Aschoff, A., & Hasselman, F. (2024). *Examining the feasibility of nonlinear vector autoregressions for psychological intensive longitudinal data*. PsyArXiv.
<https://doi.org/10.31234/osf.io/5s8jc>



Abstract

Nonlinear relationships among variables play an important role in psychological modeling and understanding changes over time from intensive longitudinal data (ILD). Most methods focus on linear models, with a few exceptions developed for specific nonlinear interactions or general system dynamics. Methods considering multiple possible nonlinear relationships among all variables remain sparse, hindered by challenges like overfitting and interpretation difficulties. This article examines the feasibility of applying such a method to psychological ILD, using the Regularization Algorithm under Marginality Principle (RAMP) alongside a local linearization method for interpretation. We evaluated its performance with simulated and empirical datasets using classification metrics, information criteria, and cross-validation. Results show this method often requires long series for satisfactory performance. While information criteria favor the nonlinear method in empirical datasets, cross-validation favors simpler AR models. Nonetheless, these challenges are comparable to those in linear idiographic modeling. Clear evidence of nonlinear relationships among variables supports the value of this method for exploratory studies. We developed an R package, quadVAR, as an implementation of this method.

Introduction

Dynamics, how variables change over time, is at the heart of many psychological inquiries, such as development, affective, and treatment processes (Bogartz, 1994; Boker & Moulder, 2023; Granic, 2005; Hamaker et al., 2015; Myin-Germeys & Kuppens, 2021). Empirical studies have observed various complex behaviors of the system, such as multistability, sensitive dependence on the initial condition, and sudden changes (Granic et al., 2018; Hayes et al., 2015; Loossens et al., 2020; Moulder et al., 2022; Olthof et al., 2020, 2023; Richardson et al., 2017; Tang & DeRubeis, 1999). Those behaviors are not possible to be generated from linear dynamics, as linear interactions among variables can only create one equilibrium point (Cui et al., 2023; Haslbeck et al., 2022; Haslbeck & Ryan, 2022; Hoekstra et al., 2024; Sayama, 2015). Therefore, to gain a better understanding of psychological dynamics, it is essential to move beyond linear statistics and start investigating the nonlinear part of dynamic interactions.

In the cross-sectional context, it is not a new topic to go beyond linear models and investigate nonlinear relationships among variables (Cohen et al., 2003; Failenschmid et al., 2025; Guastello, 1982; Kroc & Olvera Astivia, 2023; Rigdon et al., 1998). Although linear methods (e.g., Pearson correlation, linear regression) are common to perform and provide straightforward interpretations, they assume a stable relationship of variables throughout the variables' range. This assumption is often unrealistic, as many relationships in psychology are known to be nonlinear, for example, the floor and ceiling effects (Liu & Wang, 2021; L. Wang et al., 2008), or the inverted-U curve for arousal and performance (Yerkes & Dodson, 1908). In those cases, it is better to use a nonlinear technique instead for more realistic inferences (Kroc & Olvera Astivia, 2023). Such nonlinear techniques may include, for example, polynomial regression, which adds nonlinear terms to linear regression, and nonparametric methods, such as splines and local regressions.

A challenge in dynamic modeling, however, is the balance between the number of variables, the large number of nonlinear forms they may constitute, and the relatively small number of data points usually available in psychological time series. The dynamics of psychological variables are often investigated using intensive longitudinal data (ILD), assessed with experience sampling methods (ESM, also known as ecological momentary assessment, EMA, or ambulatory assessment, AA; Kalokerinos et al., *in preparation*; Myin-Germeys & Kuppens, 2021; Robinaugh et al., 2020; A. A. Stone et al., 2023). Those data rely on the self-reports of experiences, which can be evaluations of internal mental or emotional states, or evaluations of the experience of external events. The number of assessments per participant in ESM studies is mostly around a hundred (Vachon et al., 2019). Estimating complex models

from limited data points can lead to severe overfitting issues. Previous studies have shown that even linear VAR models may be overfitting and do not provide more accurate predictions than single-variate AR models in typical psychological settings (Bulteel et al., 2018; Dablander et al., 2020; Mansueto et al., 2023). Adding a large number of nonlinear terms without proper regularization will only make the overfitting problem more severe.

Whereas the majority of ILD modeling work focuses on linear models (e.g., Bringmann et al., 2013, 2015; Epskamp et al., 2018; Hamaker et al., 2018; Jordan et al., 2020; Robinaugh et al., 2020), there are several nonlinear methods available. Several methods focus on nonlinear trends over time (Bringmann et al., 2017; Failenschmid et al., 2025), which is an important pathway to tackle nonlinear phenomena in the time series, but not directly relevant for our aim as they do not explicitly model the relationships among variables. Some methods aim to describe more complex forms of nonlinear interactions from data. The output of those methods is often the attractors in the original or reconstructed state space. The Hankel alternative view of Koopman analysis (HAVOK), for example, can mimic the behavior of dynamic systems with complex mathematical forms (Moulder et al., 2022); the fitlandr method can describe the nonlinear dynamics of a system with vector fields and stability landscapes (Cui et al., 2023); recurrence quantification analysis (RQA) and recurrence networks can describe when and how a system revisits similar states (Hasselman & Bosman, 2020; Marwan et al., 2007; Zou et al., 2019). However, those methods do not provide a direct interpretation of the *effects* among variables (i.e., whether a higher value in X can predict a higher or lower value of Y), which is often of interest in psychological research. Some models are closer to the traditional linear vector autoregression framework but add a few nonlinear terms to it. For example, the moderated VAR (M-VAR) model (Adolf et al., 2017; Bringmann et al., 2024) allows adding an external moderator that can influence the VAR parameters. Adding more variables is challenging as the number of possible terms can increase exponentially as the possible nonlinear forms and the possible variables involved expand. The sparse identification of nonlinear dynamics (SINDy) algorithm, for example, applies a penalized regression for a large amount of nonlinear dynamic forms and aims to recover the dynamic equations of a system from time series (Brunton et al., 2016). However, this method has a high requirement on the sampling frequency, length, and precision, thus also unsuitable for psychological ILD.

So, is it possible to develop a method that can estimate the coefficients of some nonlinear terms in the multidimensional dynamic equations from empirical time series? What we are looking for is a method with more balanced characteristics in interpretability, feasibility, and ability to represent nonlinear dynamics than currently available methods. Therefore, in this

article, we introduce a new technique, quadratic vector autoregressive (quadVAR, accompanied by an R package of the same name) models, which provides a method for fitting quadratic relationships among many variables in psychological ILD. Here, by quadratic we mean a quadratic polynomial with the quadratic terms of a single variable (e.g., x_1^2) as well as all the possible second-order interaction terms (e.g., x_1x_2). We chose quadratic models that are similar to Haslbeck, Borsboom, et al. (2021) and Bringmann et al. (2024) instead of more flexible curves, such as the generalized additive model, GAM, and the kernel smoothing method (Haslbeck et al., 2021; Wood, 2017), because the number of possible nonlinear terms is already very large in the case of interactions of quadratic polynomials. Consider an assessment of 10 items, the linear terms (x_1, x_2, \dots, x_{10}) and the quadratic terms ($x_1^2, x_1x_2, x_1x_3, \dots, x_2^2, x_2x_3, \dots, x_{10}^2$) already have 65 degrees of freedom. Fitting this model for each outcome variable would increase the total degrees of freedom to 650. Considering the amount of data available in psychological studies, the possible terms in the model are already numerous. Haslbeck, Bringmann, et al. (2021) only used time as the moderator, which reduces the number of effects to be estimated and more degrees of freedom can be used to refine the shape of the curves. For the same reason, we did not include higher-order terms and non-polynomial terms, unlike Brunton et al. (2016). We will revisit the pros and cons of this choice in the discussion section.

After choosing the general strategy of quadratic regression, two further issues remain to be addressed. First, how can the structure of the whole model be constrained? For models with interaction terms, it is desirable that the model structure has a hierarchy guided by the *marginality principle*, which means corresponding linear terms (e.g., x_1) enter the model before interaction terms (e.g., x_1x_2 , Hao & Zhang, 2017; Nelder, 1977). Otherwise, an arbitrary change of the zero point of the scale can lead to a change in the number of terms in the resulting model, thus influencing the estimation of model complexity in the penalized regression procedure. For example, if we obtain a model where $y = x_1x_2$, there is only one term chosen for the predictor side, but if we arbitrarily choose another scale, $x'_1 = x_1 + 10$, then we have, $y = (x'_1 - 10)x_2 = x'_1x_2 - 10x_2$, where the term x_2 appears again. As most of psychological scales are interval scales, we do not have reason to believe one particular zero point is superior to another. This makes it important to keep the aforementioned hierarchy in the regression structure. In quadVAR, we deal with this issue by using the Regularization Algorithm under Marginality Principle (RAMP, Hao et al., 2018) algorithm, which we will introduce later in more detail.

Second, even if a nonlinear model can be estimated, it can be difficult to interpret all the coefficients and draw meaningful conclusions. The widely used network representation uses the width of edges to represent the linear relationships among variables. If the relationship is nonlinear, a large number of hyperedges (i.e., edges connecting three or more nodes) may be needed to describe the full dynamics, which makes inference and visualization more complicated (Bringmann et al., 2024; Haslbeck et al., 2021). To deal with this issue, we adapted the proposal by Kroc and Olvera Astivia (2023) and propose using linearized networks for interpretations. We will also introduce this method in more detail in later sections.

The remainder of this article is structured as follows. First, we introduce the quadVAR method, which includes the RAMP algorithm for quadratic regressions and the local linearized networks for interpretation. After that, we use simulated datasets and empirical datasets to examine the performance of quadVAR. With simulated datasets, we examine the performance based on information criterion, cross-validation, and inference based on the true model. With the empirical datasets, we show how the method performs in real life. Finally, we discuss the pros and cons of using nonlinear autoregressive models for psychological ILD and possible future developments in nonlinear time series modeling.

The quadVAR Method

In this section, we introduce the quadVAR method, which includes an estimation part using RAMP, and an interpretation part using local linearized networks. We will explain the motivation, specification, and advantages of the two parts separately.

RAMP: Quadratic Regularized Regression with Hierarchical Structure

The core algorithm of model estimation in quadVAR is based on the RAMP algorithm, (Hao et al., 2018). High dimensional regression with nonlinear terms is a complex statistical question by itself (Hao & Zhang, 2017). Many statistical methods have been developed in the past decade, mainly in the context of bioinformatics. Among those methods, RAMP is a recent one that retains a hierarchical structure, and previous research has shown that it performs similarly or better than other comparable models (e.g., hierNet, Bien et al., 2013, and iFOR, Hao & Zhang, 2014) with high computational efficiency and an implementation in R (Hao et al., 2018; Luo & Chen, 2021; C. Wang et al., 2021). We briefly introduce the algorithm as follows. A linear vector autoregressive model with p variables can be formulated as

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{B}\mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (1)$$

where \mathbf{Y}_t is the p -dimensional vector representing the variables at time t , \mathbf{c} is a vector of the same length representing the constant term, \mathbf{B} is a $p \times p$ coefficient matrix, and $\boldsymbol{\varepsilon}_t$ is a

p -dimensional vector of error terms at time t . To simplify the notifications, we only consider the regressive model for the i -th outcome variable. All the estimations are independent of each other, so we can always first get the regression model for each variable as the outcome variable, and then combine them together to constitute the full VAR model.

$$Y_{t,i} = c_i + \boldsymbol{\beta}_i Y_{t-1} + \varepsilon_{t,i}. \quad (2)$$

The standard ordinary least square (OLS) estimation for this regression model with T observations is then given by

$$(\hat{c}_i, \hat{\boldsymbol{\beta}}_i) = \underset{c_i, \boldsymbol{\beta}_i}{\operatorname{argmin}} \sum_{t=2}^T (y_{t,i} - c_i - \boldsymbol{\beta}_i Y_{t-1})^2. \quad (3)$$

As shown above, there are $p + 1$ parameters to be estimated for a single outcome variable, thus $p(p + 1)$ parameters for the whole VAR model, which is often a large number compared to the number of observations. Therefore, to increase the parsimony of the model and prevent overfitting, a shrinkage method, namely the least absolute shrinkage and selection operator (LASSO) is often used (e.g., Epskamp & Fried, 2018; Haslbeck et al., 2021). This method has a penalty term for the coefficients,

$$(\hat{c}_i, \hat{\boldsymbol{\beta}}_i) = \underset{c_i, \boldsymbol{\beta}_i}{\operatorname{argmin}} \sum_{t=2}^T (y_{t,i} - c - \boldsymbol{\beta}_i Y_{t-1})^2 + \lambda_i \|\boldsymbol{\beta}_i\|_1, \quad (4)$$

which also does a variable selection, that the coefficients of unimportant terms will become zero, making the model more parsimonious. If we add quadratic terms, the regressive model becomes:

$$Y_{t,i} = c_i + \boldsymbol{\beta}_{Mi} Y_{t-1,i} + \boldsymbol{\beta}_{li} Y_{t-1} \circ Y_t + \varepsilon_{t,i}, \quad (5)$$

where we follow the notation by Hao et al. (2018) and define $\mathbf{Y}_t \circ \mathbf{Y}_t$ as all the pairwise products of \mathbf{Y}_t , namely $(Y_{t,1}Y_{t,1}, Y_{t,1}Y_{t,2}, \dots, Y_{t,p}Y_{t,p})^T$, which is a vector with $p(p + 1)/2$ dimensions, and $\boldsymbol{\beta}_{li}$ is a $p(p + 1)/2$ coefficient vector of all quadratic terms. Taken together, there are $p(p + 1) + p^2(p + 1)/2$ coefficients to be estimated for the whole quadVAR model, and the possible coefficients for quadratic terms are much more than the coefficients of the main effects. If we use LASSO for all the linear and quadratic terms altogether, two issues will arise. First, as the number of quadratic terms is much larger than the number of linear terms, quadratic terms are more likely to be selected in the final model, which has a negative effect on model parsimony because we would only want to add nonlinear terms and complicate the model when the quadratic terms are necessary. Second, if we add quadratic terms without corresponding linear effects, the structure of the model is subject to change with a simple coding transformation, as the marginality principle we explained above.

The RAMP algorithm solves those problems with special treatment for interaction terms and hierarchical structure between the main effects and interactions. Specifically, the model is built with a sequence of penalty parameters, in descending order ($\lambda_{i,1} > \lambda_{i,2} > \dots > \lambda_{i,K}$). In each step k , all the main effects that have been involved in the interactions in step $k - 1$ will not be panelized, and the other main effects that have not been involved in the interactions in step $k - 1$, together with the permitted interactions based on the main effects that are already in the equation in step $k - 1$, may enter the equation subject to the penalty parameter of the current step. Written in statistical representation:

$$\begin{aligned} & (\widehat{c}_{i,k}, \widehat{\boldsymbol{\beta}}_{Mi,k}, \widehat{\boldsymbol{\beta}}_{Ii,k}) \\ &= \operatorname{argmin}_{c_{i,k}, \boldsymbol{\beta}_{Mi,k}, \boldsymbol{\beta}_{Ii,k}} \sum_{t=2}^T \left(Y_{t,i} - c_i - \boldsymbol{\beta}_{Mi,k} \mathbf{Y}_{t-1,i} - \mathbf{PI}_{\widehat{\boldsymbol{\beta}}_{Mi,k-1}}(\boldsymbol{\beta}_{Ii,k}) \mathbf{Y}_{t-1} \circ \mathbf{Y}_{t-1} \right)^2 \\ &+ \lambda_{i,k} \left(\left\| \mathbf{OM}_{\widehat{\boldsymbol{\beta}}_{Ii,k-1}}(\boldsymbol{\beta}_{Mi,k}) \right\|_1 + \left\| \boldsymbol{\beta}_{Ii,k} \right\|_1 \right), \end{aligned} \quad (6)$$

where $\mathbf{PI}_{\widehat{\boldsymbol{\beta}}_{Mi,k-1}}(\boldsymbol{\beta}_{Ii,k})$ represents choosing permitted interaction terms from $\boldsymbol{\beta}_{Ii,k}$ based on $\widehat{\boldsymbol{\beta}}_{Mi,k-1}$, and $\mathbf{OM}_{\widehat{\boldsymbol{\beta}}_{Ii,k-1}}(\boldsymbol{\beta}_{Mi,k})$ represents choosing other main effects from $\boldsymbol{\beta}_{Mi,k}$ that did not participate in $\widehat{\boldsymbol{\beta}}_{Ii,k-1}$.

With different $\lambda_{i,k}$ values, a set of candidate models is generated for the i -th outcome variable. All the models are then evaluated using extended Bayesian information criteria (EBIC; Chen & Chen, 2008) to find the model that performs the best. EBIC is a variation of BIC with the model searching space considered. The original BIC only considers the number of parameters estimated in the penalty term. This is problematic when a large number of models are estimated and compared because when more parameters are estimated, there are often a large number of possible models, and it is likely that some of them by chance have a better fit. By penalizing the model searching space, EBIC can better ensure the model parsimony and have a better estimate of the model performance in new datasets. The EBIC metric has also been used in many popular linear, binary, and mixed network estimation methods (Epskamp et al., 2012; Epskamp & Fried, 2018; Haslbeck & Waldorp, 2020; van Borkulo et al., 2014).

Local Linearized Networks

After obtaining the estimation for the quadratic regression model, we still need to know what inferences we can draw from the estimation result. Drawing inference from a nonlinear model can be tricky because, as in analyzing interactions for ANOVA, the size of an effect may be dependent on the value of another variable, or the variable itself. Here, we take the proposal by Kroc and Olvera Astivia (2023) to use a linearization method. Technically, it takes the partial

derivative of the regression as the size of an effect, which means that given all other variables held constant, how much the expected change of the outcome variable is given one unit increase of the input variable. All the pairwise effects can be visualized as a network, with the weight of edges representing the effects, the same as linear longitudinal networks (Bringmann et al., 2013). For linear VAR models, the size of the effect is always the same, independent of the variable values, and equal to the corresponding coefficient. In contrast, for quadVAR models, the effects may change with the variable values, which means the variable values will change the structure of the autoregressive networks. Examples will be given in subsequent sections.

Performance of quadVAR: A Simulated System

To give a general illustration and examine the performance of quadVAR, we first show its performance based on a simulated system. We use a nonlinear, bistable emotional model adapted from van de Leemput et al. (2014), which is an extension of the Lotka–Volterra equation. The basic assumption of the model is that there are two positive emotions and two negative emotions in an emotional system. The positive emotions strengthen each other, weaken the negative emotions, and vice versa. Many previous studies have used this model to illustrate nonlinear dynamics and bimodality in psychological ILD (e.g., Cui et al., 2023; Haslbeck & Ryan, 2022). The original model is described continuously, yet in the current work, we rewrite it as a discrete version because the quadVAR model we introduce is a discrete model. If we estimate a discrete model from a continuous system, the coefficients do not have a one-on-one relationship, and it is difficult to obtain the true coefficients (Ryan & Hamaker, 2022).

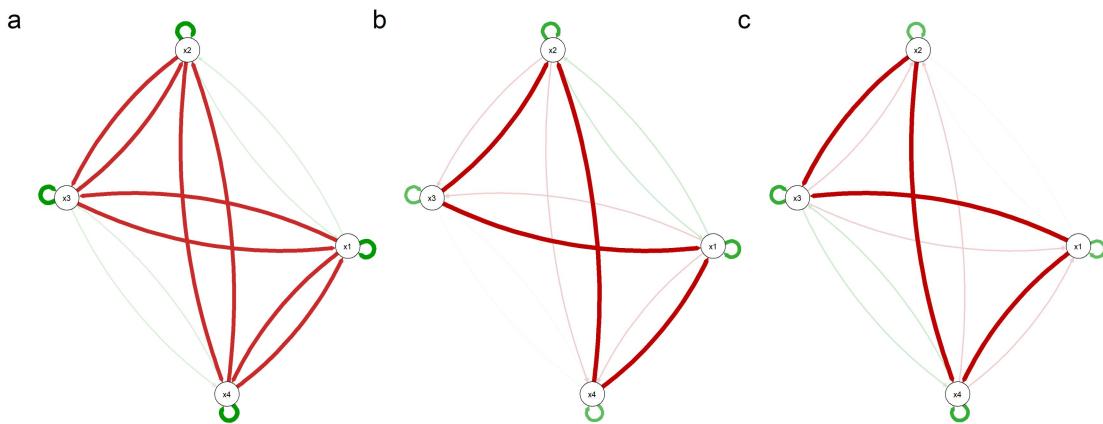
The model we use is specified as follows:

$$\begin{aligned}x_{1,t+1} &= 0.8 + 1.5x_{1,t} - 0.1x_{1,t}^2 + 0.02x_{1,t}x_{2,t} - 0.1x_{1,t}x_{3,t} - 0.1x_{1,t}x_{4,t} + \varepsilon_1, \\x_{2,t+1} &= 0.8 + 1.5x_{2,t} - 0.1x_{2,t}^2 + 0.02x_{1,t}x_{2,t} - 0.1x_{2,t}x_{3,t} - 0.1x_{2,t}x_{4,t} + \varepsilon_2, \\x_{3,t+1} &= 0.8 + 1.5x_{3,t} - 0.1x_{3,t}^2 + 0.02x_{3,t}x_{4,t} - 0.1x_{1,t}x_{3,t} - 0.1x_{2,t}x_{3,t} + \varepsilon_3, \\x_{4,t+1} &= 0.8 + 1.5x_{4,t} - 0.1x_{4,t}^2 + 0.02x_{3,t}x_{4,t} - 0.1x_{1,t}x_{4,t} - 0.1x_{2,t}x_{4,t} + \varepsilon_4.\end{aligned}\tag{7}$$

In this system, x_1 and x_2 represent two positive emotions, for example, cheerful and content, and x_3 and x_4 represent two negative emotions, for example, sad and anxious (van de Leemput et al., 2014) and ε_i ($i = 1, 2, 3, 4$) are Gaussian noise with $SD = 1$. We first show the networks for the true model in Figure 1. Since the model is nonlinear, the network structure differs per variable value. We can see that in the neutral situation (Figure 1a), the influence of positive emotions on negative emotions is as strong as the effect of negative emotions on positive emotions, which is expected because the model has a symmetric structure. Yet, when the system is dominated by positive emotions (Figure 1b), the strength of the edges becomes

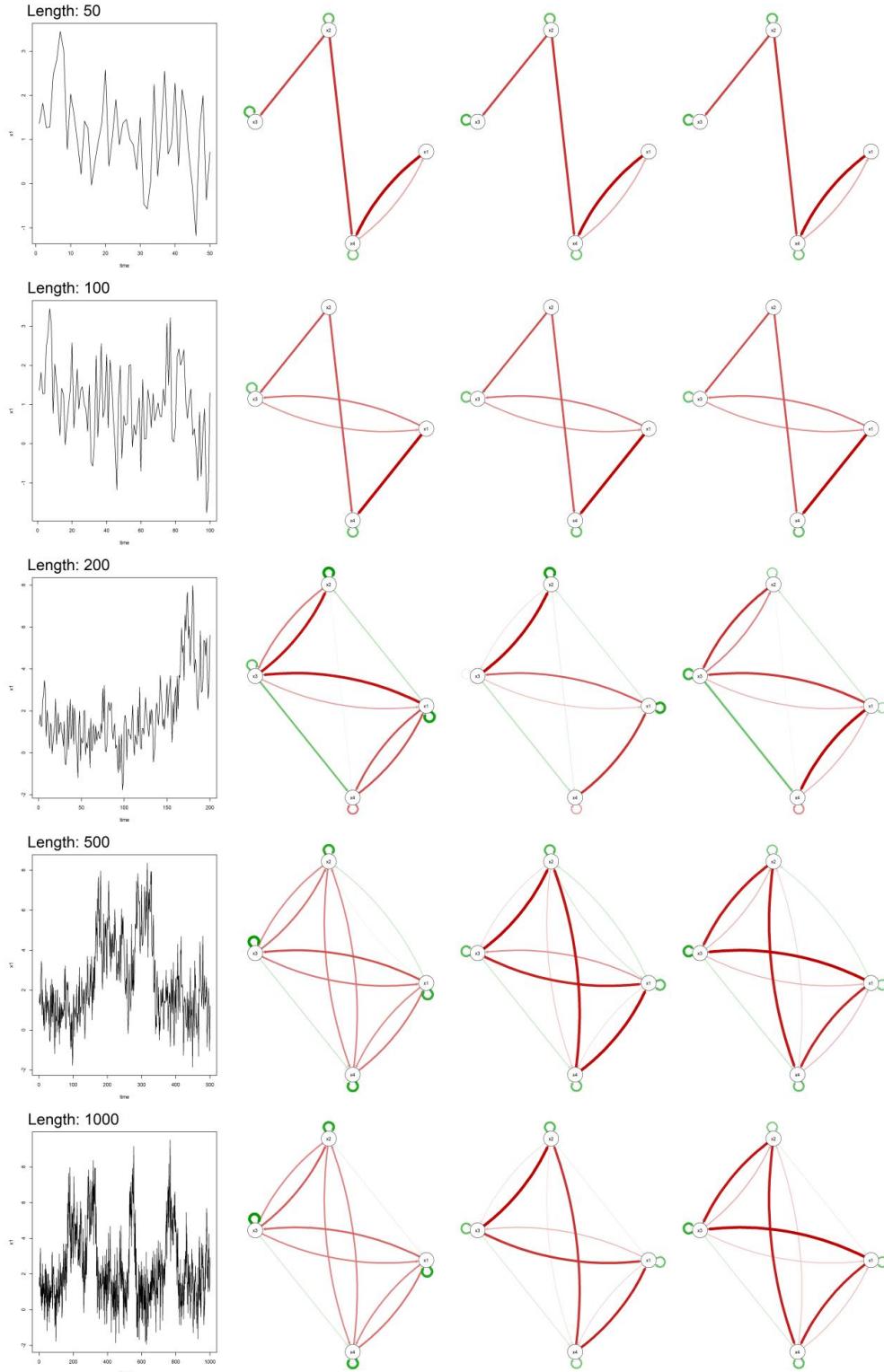
asymmetric. The effect of positive emotions on negative emotions is not strong anymore, but negative emotions have a stronger influence on positive emotions. The opposite pattern appears when the system is dominated by negative emotions (Figure 1c).

Figure 1. Linearized networks for the true model for different system states: (a) neutral state, which is the saddle point between two phases: $x_1 = x_2 = x_3 = x_4 = 2.80$; (b) positive state, which is the local minimum of the phase dominated by positive affects, $x_1 = x_2 = 4.89$, $x_3 = x_4 = 1.36$; (c) negative state, which is the local minimum of the phase dominated by negative affects, $x_1 = x_2 = 1.36$, $x_3 = x_4 = 4.89$.



We then simulate the models for different lengths, namely 50, 100, 200, 500, and 1,000 timesteps, and estimate quadVAR models from the simulated datasets. Those data lengths correspond to different scenarios in psychological ILD collection: collecting 50 data points is an entry-level for ILD studies, and collecting 1,000 data points corresponds to the maximum possibility in ILD collection in previous studies (e.g., Delignières et al., 2004; Kossakowski et al., 2017). We first show an example for each simulation length and the networks estimated thereafter are shown in Figure 2. From the simulation sample, we can observe that for 50 data points, the bistability of the system is only vaguely observable, whereas for 200 or more data points, the bistability of the system is clearer. Similarly, the estimated networks also become close to the true networks for longer time series. During the simulation trials, we also found an interesting phenomenon that is specific to nonlinear multistable systems, decreasing noise does not always lead to better model estimates; on the contrary, it may make the system stay longer in one phase and reduce information on alternative phases, therefore decreasing model performance. We show an example of this finding in Appendix F1.

Figure 2. Linearized networks for the simulation examples with different lengths. In each row, the first plot shows the raw time series of x_1 , the second to the fourth plots show the linearized networks estimated from simulation data for the neutral state ($x_1 = x_2 = x_3 = x_4 = 2.80$), the positive state ($x_1 = x_2 = 4.89, x_3 = x_4 = 1.36$), and the negative state ($x_1 = x_2 = 1.36, x_3 = x_4 = 4.89$).

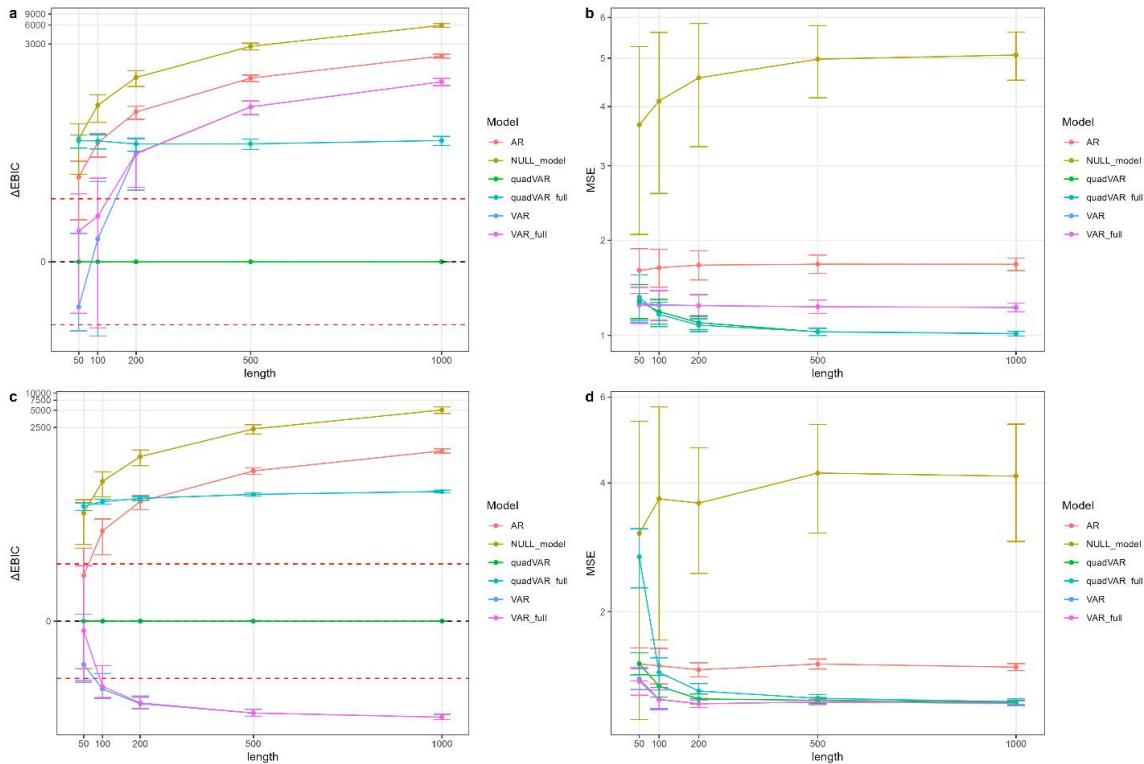


We then run simulations for each length 1,000 times and evaluate the performance of the quadVAR models in two parts. First, we compare the performance of quadVAR with the NULL models (i.e., the intercept-only models that simply use the mean as the prediction), the AR models, the unregularized VAR models, the regularized VAR models (with the `ncvreg` package in R, Breheny & Huang, 2011), and the unregularized quadVAR models, based on the EBIC metrics and cross-validation. To consider the scenario that the true system is linear, we also simulate a linear system based on a regularized VAR estimation (coefficients reported in Appendix F2) of the nonlinear model described above. This makes sure the linear model is comparable with the nonlinear model. To evaluate the model performance in predicting unseen new data points, we perform 10-fold block cross-validation based on the suggestions by Bergmeir and Benítez (2012) and Bulteel et al. (2018) and take the mean squared error (MSE) as the performance metric. Although theoretical work has shown that the performance measures based on information criteria will be consistent with the measures based on cross-validation asymptotically (Shao, 1997; M. Stone, 1977), for real-life situations with limited observations, cross-validation may still give different results than information criteria. Most previous studies only used the cross-validation method for evaluating model performances (e.g., Bulteel et al., 2018; Dablander et al., 2020). In the current study, we use both to compare their differences. Second, we calculate the correlation of the estimated nonlinear coefficients from quadVAR and the true nonlinear coefficients, as well as several classification metrics, including correlation, precision, recall, specificity, sensitivity, and accuracy (see Appendix F3 for the definition of those metrics; Mansueto et al., 2022). Compared with information criteria and cross-validation, which focus more on the overall predictive power, the correlation of coefficients and classification metrics focus more on statistical inferences of individual relationships among variables. In addition, we also examine the classification metrics for the key nonlinear effect described above, namely the effects of positive emotions on negative emotions diminish when positive emotions are dominant (see Figure 1). Specifically, we define a true positive when the effect of cheerful on sad in a positive state ($x_1 = x_2 = 4.89, x_3 = x_4 = 1.36$) is higher than the neutral state ($x_1 = x_2 = x_3 = x_4 = 2.80$), and is a false negative otherwise. We also compare the same effect at the positive state ($x_1 = x_2 = 4.89, x_3 = x_4 = 1.36$) and the positive-neutral state ($x_1 = x_2 = 4.89, x_3 = x_4 = 2.80$), which should be the same for the true system. Therefore, we define a true negative when the effect of cheerful on sad at the positive state and the positive-neutral state is the same, and a false positive otherwise.

Assessing Model Performance with EBIC and Cross-Validations

As the raw EBIC values are not comparable for different datasets, we report Δ EBIC values instead, which are the raw EBIC values subtracted by the EBIC values of the quadVAR model. The results for the original nonlinear system are shown in Figure 3a, and the results for the comparable linear system are shown in Figure 3c. According to the guidance by Burnham et al. (2002), an EBIC difference of 10 or larger is considered essential. We draw this criterion as a red dashed line in Figure 3a and 4c. From the results, it is clear that quadVAR performs best for the nonlinear simulation model when the input dataset contains 100 or more data points, whereas the VAR model performs best when the actual system is linear. For short time series with a length of 50, the VAR model outperforms the quadVAR model. This is in agreement with the results by Bulteel et al. (2018) because more complex models need more data points to be precisely estimated. If the available data points are not enough, the estimation results of the model may be inaccurate, making the performance worse than a simpler, yet mis-specified model. The AR model never performs well, which is different from previous studies (e.g., Bulteel et al., 2018; Dablander et al., 2020), and the unregularized full VAR model performs almost as well as regularized VAR models. This may be because the simulation model we used contains dense mutual influences among variables, and this is also the case for the comparable linear system, which only contains one absent effect. We will revisit how realistic the simulated systems are in later discussions. We show the MSE values from cross-validations with datasets of different lengths in Figure 3b and 4d. In general, cross-validation MSE shows a similar pattern as the information criteria (Figure 3a and 4c). One difference is that, when the true system is nonlinear, the unregularized quadVAR model yields similar or even better predictions than the regularized quadVAR model yields. This indicates that the regularization procedure in quadVAR is more conservative and tends to produce models with higher parsimony, but this does not necessarily lead to a better prediction performance. Overall, both EBIC and cross-validation results indicate that, as the dataset is long enough, the model that is closer to the true dynamics of the system tends to perform better, and the general performance of quadVAR is satisfactory when modeling sufficient data from nonlinear systems.

Figure 3. EBIC and cross-validation MSE comparison for repeated simulations. The first row is the results from the original nonlinear system, and the second row is the results from the comparable linear system. The suffix “_full” represents the unregularized models (i.e., none of the parameters shrink to zero). The error bars represent standard deviation. Note that to represent large ΔEBIC values, we used pseudo-logarithm scales for the y-axis, which means the scale is logarithmic for large values but smoothly transforms to a linear scale around 0.

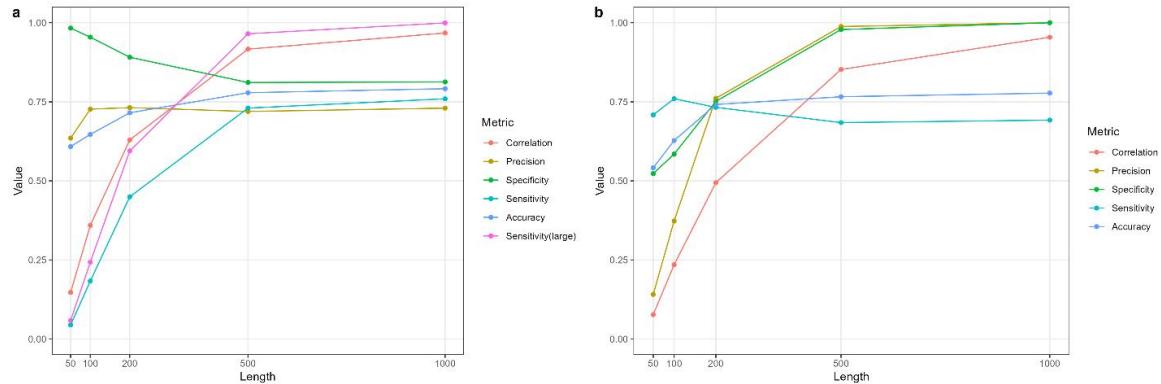


Assessing Model Performance with Classification Metrics

The classification metrics for the coefficient estimations and the nonlinear effects are summarized in Figure 4. We first look into Figure 4a, which is the classification metrics for nonlinear parameters. In general, many indicators seem suboptimal for short time series, especially the sensitivity. Because large effects are more important than small effects in real-life scenarios, we also checked the sensitivity for coefficients with an absolute true value larger than or the same as 0.10. As expected, the sensitivity for those coefficients is higher than that for all coefficients, but still rather low for small sample sizes. Precision and specificity are better for small sample sizes, yet the specificity declines as the sample size grows. The accuracy and correlation also increase steadily as the sample size grows. Taking together, although the

overall performance of the method increases as the sample size increases, it seems to be more on the liberal side for larger sample sizes at the coefficient level.

Figure 4. Classification metrics for the models estimated from simulation data.



In contrast, the classification metrics for nonlinear effects (Figure 4a) show a different pattern. The sensitivity is rather high for smaller sample sizes and slightly declines as the sample size grows, yet the precision starts low and quickly grows as the sample size increases. The reason for this difference could be that multiple coefficients can lead to similar nonlinear effects. In the true system, all the variables have symmetric forms of dynamic functions. The values of x_1 and x_2 are always similar, so as x_3 and x_4 . If in the coefficient estimation misattributed an effect from x_1 to x_2 , the nonlinear effect may still be correct. Therefore, as the sample size grows, the estimation of the nonlinear effects also improves but leans toward the conservative side.

Discussion for the Simulation Study

In this section, we tested quadVAR on a simulated nonlinear system. We found that, overall, quadVAR has an adequate performance. For nonlinear systems, it performs better than linear models both in EBIC and cross-validation when the sample size is not too small. For linear systems, using quadVAR can result in lower EBIC than linear models, as expected, but the cross-validation result is not much worse than the linear models, which indicates that the higher flexibility of quadVAR did not lead to excessive overfitting. For statistical inferences, the overall performance of quadVAR is not entirely satisfactory but is comparable to the individual linear models reported in previous studies (Mansueto et al., 2023), which also indicates that the overfitting issue of quadVAR is well controlled. Although some classification metrics are not very high for usual sample sizes (especially sensitivity for coefficients and precision for nonlinear effects), we would like to emphasize that, as far as we know, quadVAR is the only

method to date that allows estimating nonlinear dynamics from individual ILD, which means that if we also draw the classification metrics of other methods in Figure 4, they will all be a straight line of zero because they will assume all nonlinear coefficients to be zero. In this sense, we suggest that quadVAR is a suitable method to estimate nonlinear dynamics for exploration purposes, yet the inferences should be drawn with caution. We also note that all the results we reported are only for one simulated system, which as we show later, is quite different from the dynamic structure estimated from real-life data. We will revisit this issue after showing the results from empirical data.

Performance of quadVAR: Empirical Datasets

In this section, we assess the performance of quadVAR using empirical datasets. As the true system is unknown, we cannot use classification metrics to evaluate the model performance. Therefore, we focus on cross-validation and EBIC results. Apart from the general performance of the models, we also show several examples to draw nonlinear inferences from quadVAR estimations. We note that the following results are mainly used for examining the method instead of drawing empirical conclusions, and the analysis procedures were not preregistered.

We used two datasets from previous studies. The datasets were chosen with three criteria. First, the datasets should contain enough data points for each individual, preferably around or more than 200 data points. This is based on our simulation results, that the performance of quadVAR is only adequate when the data length is around this amount. Second, the assessment should be done on a visual analog scale (VAS) instead of a Likert scale, because we consider a continuous scale is needed to adequately estimate nonlinear influences, yet a Likert scale would be too discrete and restricted to allow nonlinear features to present. Third, the datasets should be openly available (or easily available upon request) so that the results can be easily reproduced.

We found two datasets that suffice for the above criteria. The first one is the dataset from Rowland and Wenzel (2020). This dataset was especially interesting because that Haslbeck et al. (2023) reported this dataset contains strong bimodality that cannot be recovered using the VAR model. Therefore, by using this dataset, we can examine whether involving nonlinear terms would help to describe datasets with nonlinear features. The dataset contains 125 participants, and each of them answered 8 questions about their current affective states¹ 6 times per day in a 0-100 VAS for 40 consecutive days. On average, the participants had 172.56 valid responses (ranging from 76 to 226, $SD = 35.13$). To avoid the influence of different time

¹ Those affective states include happy, excited, relaxed, satisfied, angry, anxious, depressed, and sad.

intervals, all the overnight lags and lags with missing data in between were removed before the modeling procedure. The effective data length of the participants then had a mean of 114.06 (ranges from 30 to 180, $SD = 36.83$).

The second one is the dataset from Kuppens et al. (2010), requested from the EMOTE database (Kalokerinos et al., in preparation) under request number 3UUCEGX6EU. In this study, each participant answers 8 questions about their affective state and emotional regulation strategies. On average, the participants had 173.6 valid responses (ranging from 46 to 211, $SD = 26.54$). As the two datasets showed similar patterns, we focus on the results from the first dataset in the main text and report the results from the second dataset in Appendix F4.

For cross-validation, we winsorized the prediction within the possible range of answering to account for the restriction of the scale range (e.g., if a model prediction is 150 but the original data was collected using a scale from 0 to 100, we change the prediction to 100). All other settings for the cross-validation are the same as in the simulation study reported in the previous section.

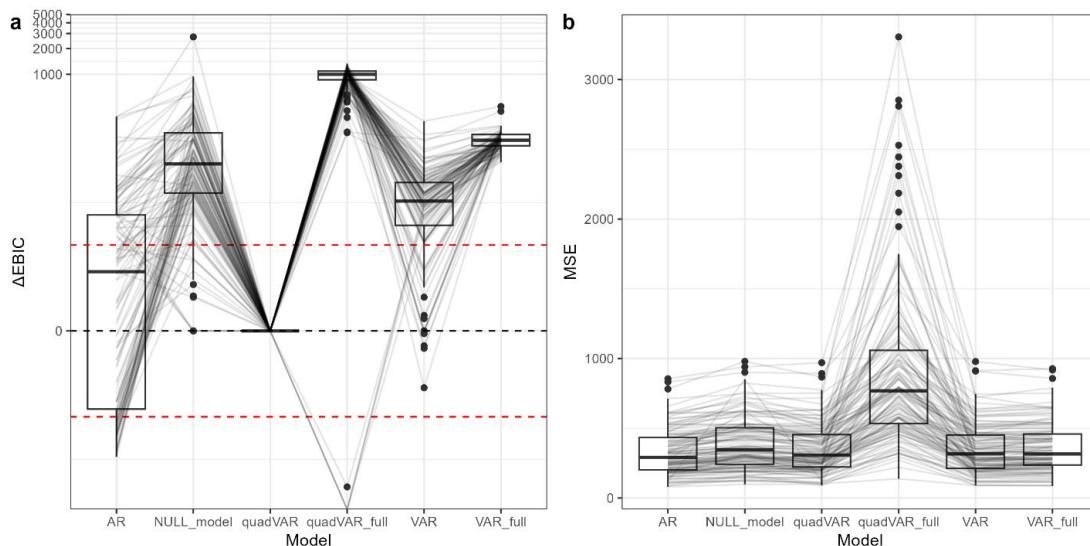
General Model Performance by EBIC and Cross-Validation

The EBIC and cross-validation results for the dataset from Rowland and Wenzel (2020) are summarized in Figure 5. For empirical data, the model selection results by EBIC and cross-validation have very different patterns. For the dataset from Rowland and Wenzel (2020), if we use EBIC for model selection, the quadVAR model is the best one for 69 out of 125 participants (55.20%), followed by the AR model (48 participants, 38.40%), the VAR model (5 participants, 4.00%), the NULL model (2 participants, 1.60%), and the unregularized quadVAR model (1 participant, 0.80%)². In general, the AR model and the quadVAR models perform the best, and their relative model performances are related to the effective data length. The correlation for the relative EBIC of the AR model is 0.19 ($p = 0.03$, 95% CI [0.02, 0.36]), indicating that quadVAR performs better with longer datasets. However, if we use cross-validation for model selection, the AR model is the best for 105 participants (84.00%), followed by the unregularized VAR model (9 participants, 7.20%), the VAR model (5 participants, 4.00%), the NULL model (5 participants, 4.00%), and the quadVAR model (2 participants, 1.60%). This strong discrepancy between EBIC and cross-validation is not expected from either the theory (Shao, 1997; M. Stone, 1977) or the simulation study we performed in the previous section. The performance of the quadVAR model is well judged by EBIC, but not as much judged by cross-validation. The performance of the AR model is much better with cross-

² Those numbers do not sum up to 125 participants because for some participants multiple procedures had the same model selected and they are tied for the best.

validation than with EBIC. The NULL model also performs better in cross-validation than EBIC. Interestingly, even the best-performing AR model is not much superior to the NULL model, which indicates that time series modeling in general does not provide strong improvement over using the mean for prediction. It is also worth noting that the quadVAR models do not have more degrees of freedom compared to VAR models. The quadVAR models, on average, have 10.07 degrees of freedom ($SD = 8.40$), and the VAR models have 11.78 degrees of freedom ($SD = 9.43$) for the dataset from Rowland and Wenzel (2020). To compare with, the AR models have 8 degrees of freedom, and there are 576 candidate terms in total for quadVAR models to select from. Therefore, using quadVAR models does not necessarily lead to an increase in the complexity of the outcome model.

Figure 5. EBIC and cross-validation MSE for participants in Rowland and Wenzel (2020). Gray lines represent individual participants. In Figure 5a, there are 4 participants for whom the unregularized quadVAR models were overspecified. The EBIC values for the unregularized quadVAR models were thus negative infinity and not shown in the figure.



Examples of Nonlinear Features in Empirical Data

To illustrate the usage of quadVAR in discovering nonlinear relationships for idiographic data, we show several examples from the two datasets. The examples are arbitrarily chosen from the participants for whom the quadVAR model is the best model selected with EBIC and has at least one nonzero nonlinear term estimated. The first one is from the 6th participant from Rowland and Wenzel (2020). The nonzero parameters estimated using quadVAR for this participant are shown in Appendix F5. Here, we focus on the nonlinear effect from *excited:angry*

to *angry*. This indicates that different values of *excited* can result in different effects of *angry* on itself for the next time point. This can be seen more clearly from the linearized networks shown in Figure 6, that with a higher level of *excited*, the self-reinforce effect of *angry* becomes stronger. However, after a closer look at the data by plotting the residuals of *angry* controlled on all other effects of *angry* at the previous time point, we found that this effect is likely driven by several influential cases (Figure 7). There was once that the participant had an exceptionally high level of *angry*. The level of *excited* is low at this point, but high at the previous time point. This single time point contributed strongly to a significant estimate of the interactions between *excited* and *angry*.

Figure 6. Linearized networks for the 6th participant in Rowland and Wenzel (2020) with different levels of *excited*: (a) *Mean – SD*; (b) *Mean*; (c) *Mean + SD*. The following abbreviations are used: "hpp" represents happy, "exc" represents excited, "rlx" represents relaxed, "sts" represents satisfied, "ang" represents angry, "anx" represents anxious, "dpr" represents depressed, "sad" represents sad.

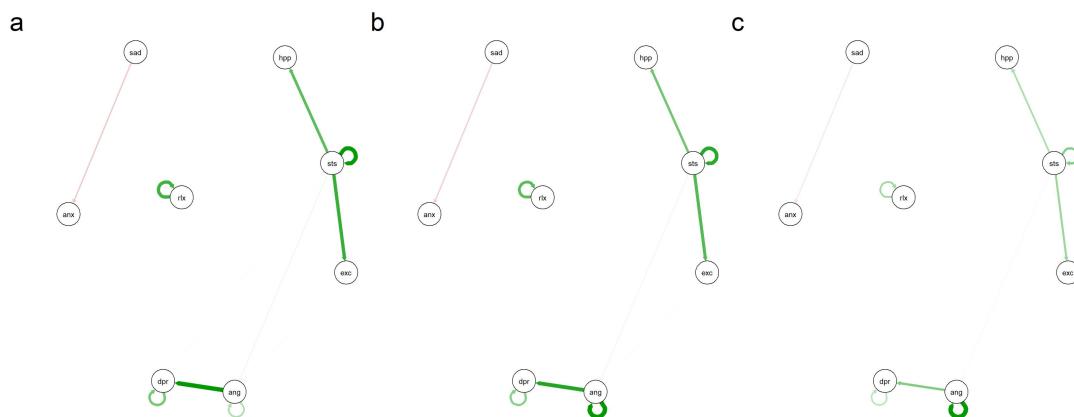


Figure 7. Partial plot of the residual of *angry*, controlled for all other variables excluding *angry* at the previous time point, on *angry* at the previous time point. Different colors represent high or low values of the moderator, *excited*, at the previous time point, divided by the median.

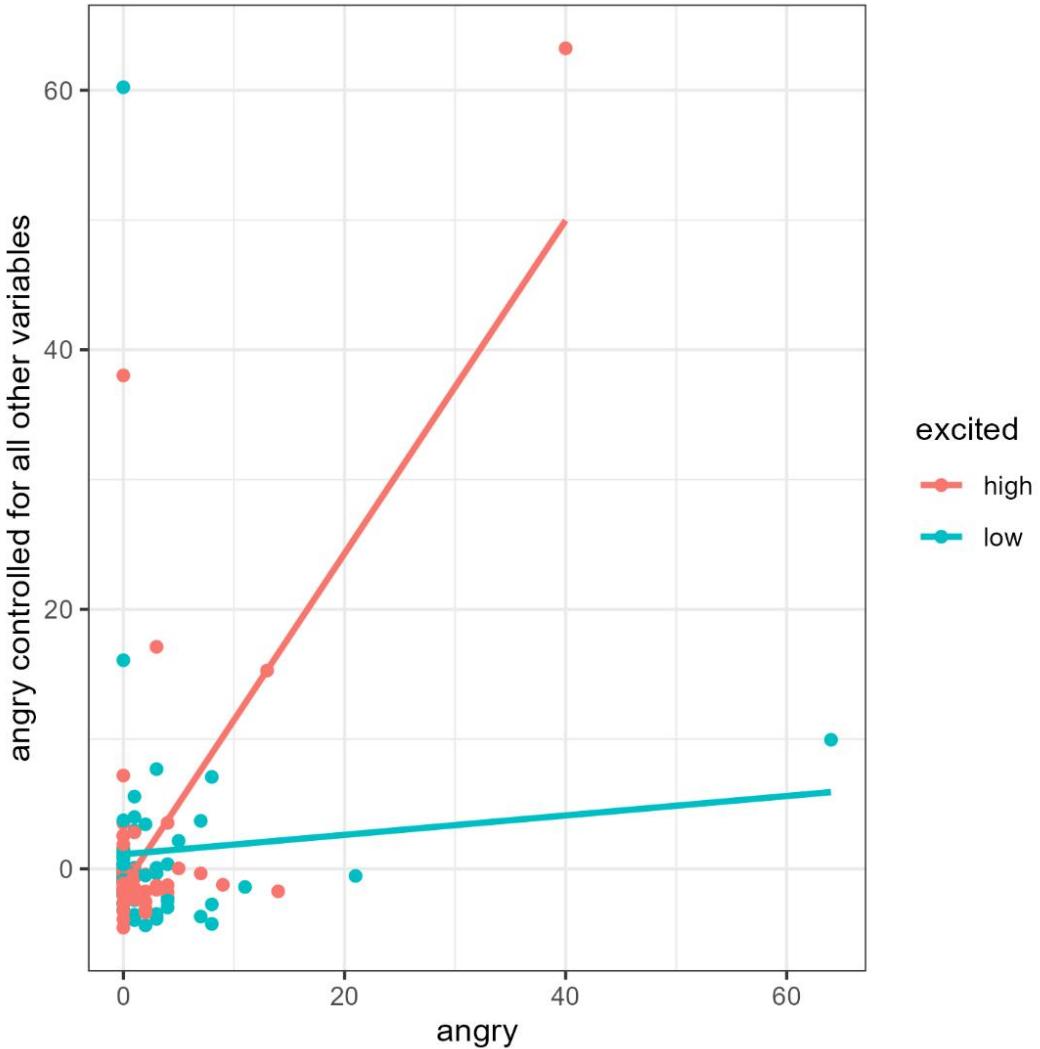


Figure 8. Linearized networks for the 8th participant in Rowland and Wenzel (2020) with different levels of *satisfied*: (a) *Mean – SD*; (b) *Mean*; (c) *Mean + SD*. The following abbreviations are used: "hpp" represents happy, "exc" represents excited, "rlx" represents relaxed, "sts" represents satisfied, "ang" represents angry, "anx" represents anxious, "dpr" represents depressed, "sad" represents sad.

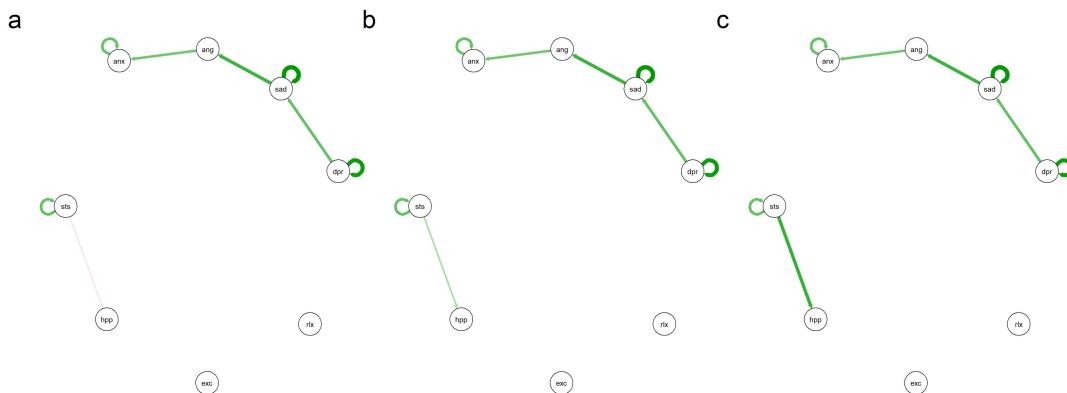
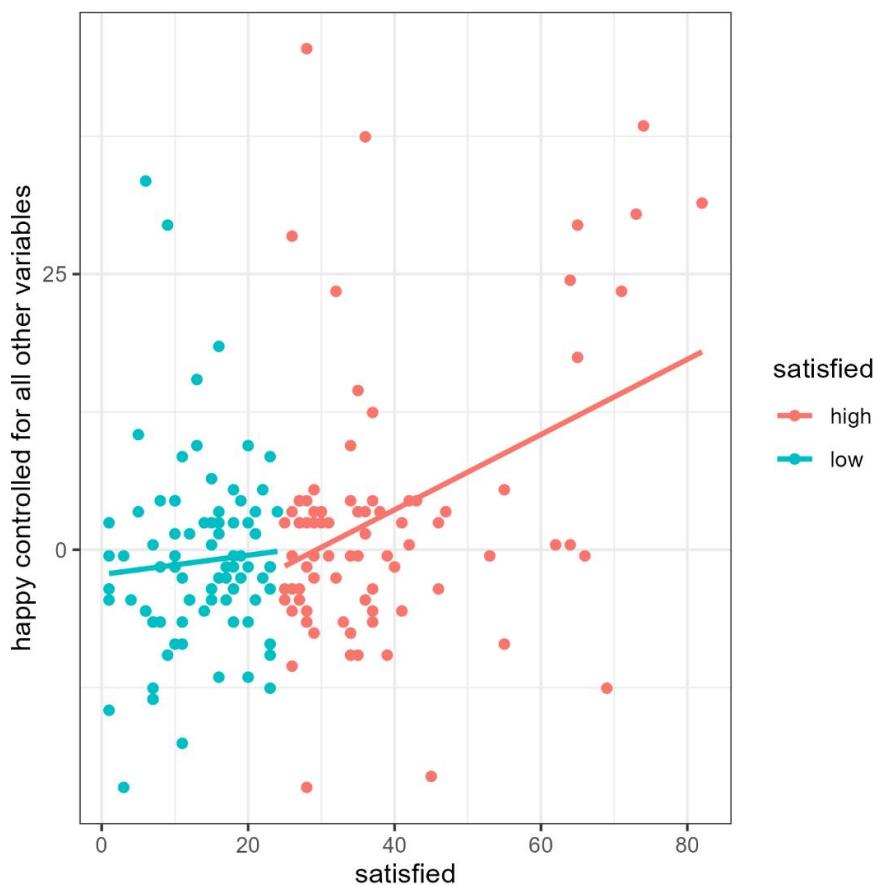


Figure 9. Partial plot of the residual of *happy*, controlled for all other variables excluding *satisfied* at the previous time point, on *satisfied* at the previous time point. Different colors represent high or low values of the moderator, *satisfied*, at the previous time point, divided by the median.



Discussion for the Results from Empirical Datasets

Unlike the simulation datasets, for empirical datasets, the model performance was judged by information criteria (EBIC in our study) and the cross-validation results showed strong discrepancies. If we look at the results from cross-validation, we will reach a similar conclusion as Bulteel et al. (2018), that the AR model performs the best in predicting unseen data. However, if we look at EBIC, which is also intended to be an estimate of the predictive power for unseen data (Burnham et al., 2002; Chen & Chen, 2008; Wagenmakers & Farrell, 2004), quadVAR performs better for more than half of the participants. The most likely explanation for the better performance of the AR model compared to the simulation data is the differences in the data generation process. Compared to the networks for simulation data (Figure 1), the networks estimated with quadVAR from empirical data contain more self-reinforced loops and much fewer edges, which indicates that the empirical data generation process is intrinsically more similar to an AR process than the simulated system. However, this does not explain the difference between EBIC results and cross-validation results. To our knowledge, the studies regarding the difference between the two performance indicators are sparse, especially in the context of psychology. We suspect that the influential cases and multistability in the datasets may contribute to this, as some cross-validation folds may not include those influential cases, thus having very different parameter estimates than using the whole dataset. Future studies are still needed to fully resolve this issue. At the same time, we would not suggest removing those influential data points before analysis, as those observations are from the same individual, and the influential data points may represent rare yet important state of the system. Therefore, we recommend always checking the influential cases before drawing conclusions and treating them with caution.

The example we presented showed nonlinear patterns that are theoretically interesting. From the 8th participant from Rowland and Wenzel (2020), for example, we found a nonlinear relationship between *satisfied* and *happy* (Figure 9), and the relationship seems explainable because when the level of satisfaction is low, it may not be significant enough to influence the feeling of happiness, but when the level of satisfaction is high, it should be able to contribute to happiness notably. However, we also saw a few overlaps of nonlinear effects between the two participants shown here. As the quadVAR model is currently an idiographic model, it does not consider any possible between-person similarities. Therefore, whether the nonlinear relationships we found are only applicable to a single participant, or can be generalized to the group level, is still a question to be investigated.

General Discussion

In this article, we introduced the quadVAR method and the accompanying R package, that enables testing non-linear effects in vector autoregressive models for psychological ILD. Our method implements a state-of-the-art quadratic regression algorithm, RAMP (Hao et al., 2018), that takes the hierarchy of regression structure and a large amount of possible nonlinear terms into account. The linearization method was used to facilitate the interpretation of the nonlinear regression output (Kroc & Olvera Astivia, 2023). The performance of the method with both simulated and empirical datasets was examined. In simulation datasets, we found that quadVAR outperforms linear methods for a nonlinear system, as long as the sample size is large enough. When we look at the performance for individual parameters and nonlinear effects, quadVAR has a similar problem as linear models in that the general accuracy is not very high and requires a large sample size to reach satisfactory performance. This indicates that quadVAR is more suitable for exploratory purposes and requires cautious interpretation and confirmatory repetitions before reaching a solid conclusion regarding specific nonlinear characteristics of a process. With empirical datasets, we found that quadVAR performs the best for more than half of the participants judged by EBIC, yet cross-validation results favor the simpler AR model. Further detailed investigations into individual participants showed that, although sometimes the nonlinear effects were driven by several influential data points, there were cases where a robust nonlinear relationship between psychological variables seemed to exist, which evidences the usefulness of investigating nonlinear relationships for psychological ILD but requires future investigations.

To our knowledge, the current work is the first in the field to explicitly model continuous, intrinsic, nonlinear dynamics from psychological ILD. Including nonlinear terms in autoregression models clearly induced more challenges. For the method development part, we need special treatments for both the estimation and interpretation of the results. A challenge in the simulation part is that nonlinear systems can easily become unstable (i.e., grow into infinity), which makes it difficult to systematically vary the parameters to test their influences on model performance. It also prohibits us from using a more realistic simulation example using parameters estimated with empirical datasets. The noise level is another challenge because it does not have a monotonic relationship with the model's performance due to the double role noise plays. On the one hand, stronger noise decreases the signal-to-noise ratio, making it more difficult to retrieve dynamic information. On the other hand, driven by noise the system travels through the state space more easily, adding more information to the data. Therefore, it is difficult to systematically evaluate the influence of noise on the model's

performance. Challenges also arise for the performance benchmarks. In nonlinear models, the linear coefficients are not independent of nonlinear coefficients, making it difficult to tell how well the linear part is estimated within the nonlinear framework. We also saw the EBIC and the cross-validation results favor different models, making it more difficult for model selection, and this may also be related to the nonlinearity of the quadVAR method.

Nevertheless, we also see promising signs of introducing such a nonlinear method into the world of idiographic dynamic analysis and network modeling. First, counterintuitively, we found that introducing nonlinear terms does not necessarily lead to a stronger overfitting issue. The degree of freedom of the quadVAR models is even slightly lower than the VAR models for the empirical datasets we used. This indicates that the nonlinear estimation algorithm we implemented performed well in constraining the overall model complexity and choosing a small number of nonlinear terms from a large reservoir. Theoretically, if some dynamic characteristics of a system can be modeled with one nonlinear term, forcing the system to be linear only makes it necessary to have more terms, which actually increases model complexity. Therefore, using nonlinear terms does not necessarily lead to more severe overfitting problems. We also see from the partial plots that quadVAR has the potential to discover nonlinear dynamical relationships from multidimensional time series. Taking all the evidence together, we suggest that exploring nonlinear dynamical relationships for psychological ILD is a promising direction, and when a linear method is used, the existence of nonlinearity and outliers in the datasets should be checked.

Several directions of development are worth future consideration. First, the order of the quadratic terms may be extended further. As shown by Kroc and Olvera Astivia (2023), including cubic terms is beneficial to model many nonlinear relationships in psychology, including the ceiling and flooring effects. Using higher-order terms inevitably makes the model selection process more difficult. Although it may not be feasible to estimate all possible cubic and even higher-order interaction terms for psychological ILD, adding cubic terms for single variables may be possible. Future studies may also consider using robust estimations in quadVAR and idiographic modeling in general, as we saw influential data points may be present in individual ILD that may change the parameter estimations significantly. Yet no matter such robust method can be developed or not, we would emphasize again the importance of checking influential cases in dynamic modeling studies in general, which is hardly done in current empirical studies, as most of the shrinkage methods and multilevel methods based on least squares estimates are actually sensitive to outliers (Dedrick et al., 2009; H. Wang et al., 2007).

Second, it may be possible to develop a multilevel extension for the quadVAR method. Previous research has shown that multilevel models perform better in cross-validation because they can also use information from other participants for estimation, therefore partly overcoming the constraints of limited data points per participant (Bulteel et al., 2018). However, multilevel models assume the coefficients of each participant vary around the group mean, which means that the within-person dynamics should be similar enough across people. This may not hold, as we see great variability in network structures among individuals. Alternatively, more idiographic frameworks, like the group iterative multiple model estimation (GIMME, Lane & Gates, 2017; Wright et al., 2019) may also be considered to estimate the generalizability of the relationship at the group level. A major difficulty in developing such group-level methods is that the RAMP algorithm and some alternative quadratic regression algorithms (e.g., hierNet, Bien et al., 2013) are all based on the LASSO penalized regression framework, which shrinks the parameter estimates to zero instead of the group mean. Recent developments have shown the possibility of combining multilevel modeling and LASSO in a single estimation procedure (e.g., glmmLasso, Schelldorfer et al., 2014). Combining this idea with the quadVAR method may be a productive approach.

Third, data collection and study design may also be improved to provide datasets that are more suitable for discovering nonlinear dynamics. In the empirical examples, we found the estimated networks are rather sparse, with very few connections across variables. This may partly be because the variables assessed were aimed at providing a general description of the affective state of participants, instead of examining a specific relationship between some important variables. With the nonlinear estimation available, future research may be specifically designed to recover nonlinear dynamics that are theoretically predicted. A well-designed algorithm may help to discover potential effects and examine certain hypotheses, but it can never do the conceptual job for researchers. As written by Burnham et al. (2002, p.147), “‘Let the computer find out’ is a poor strategy for researchers who do not bother to think clearly about the problem of interest and its scientific setting.” Therefore, what should be fed into the method should always be carefully considered.

Acknowledgments

We thank the Center for Information Technology of the University of Groningen, and especially Pedro Santos Neves therefrom, for their support and for providing access to the Hábrók high performance computing cluster. We thank Laura Bringmann and Yong Zhang for their helpful feedback and suggestions on the earlier version of this manuscript.

Chapter 9

Analyzing Formal Dynamic Models in Psychology: A Tutorial Using Graphical Tools

This chapter is based on:

Cui, J., Wagenmakers, D., Doorn, G. S. van, Hasselman, F., & Lichtwarck-Aschoff, A. (2025). *Analyzing formal dynamic models in psychology: A tutorial using graphical tools*. PsyArXiv. https://doi.org/10.31234/osf.io/fhstw_v1



Abstract

Formal theories translate verbal theories into a mathematical representation, such as a coupled differential equation or other dynamical systems, intending to strengthen the deductive power of (clinical) theories and to formulate testable and novel hypotheses. Work in clinical formal theories mainly relies on simulations, which is an intuitive method to evaluate overall model performance but may fall short in establishing a precise link between the mathematical properties of the model and the dynamic property of the model outcome. Moreover, when the model outcome contradicts clinical observations, it is unclear where the discrepancy comes from, and how to improve the model. In this current article, we introduce formal mathematical techniques for graphical model analysis, including phase plane analysis, which allows identifying a system's stable and unstable equilibria, and bifurcation analysis, a framework to delineate parameter regimes corresponding to qualitatively different dynamical outcomes for a model. Using two formal dynamic models in psychology (one for panic disorder and one for suicidal ideation), we illustrate those methods through an easy-to-use R package, *deBif*, with a graphical user interface. With these examples, we show the importance of using graphical tools to investigate the hypothesized mechanisms of psychological systems.

Introduction

Psychology has a long history of theory development, yet the progress of integrating and examining theories still faces strong obstacles (cf. Meehl, 1990a, 1990b). Many researchers have used the term “theory crisis” to describe the fact that many theories appear to coexist in many fields of psychology even though they often seek to explain the same empirical phenomena. They almost all lack the ability to generate predictions that would allow experimentalists to distinguish between the veracity of competing theories (i.e. by comparing their empirical accuracy), and they lack the required level of formal description that would allow theorists to determine whether they should be amended, integrated with competing theories, or have to be abandoned entirely (Eronen & Bringmann, 2021; Fried, 2020; Oberauer & Lewandowsky, 2019). One possible avenue to advance theory construction that has been proposed by many researchers is the use of formal dynamic models as a way to clarify psychological theories which should allow for the deduction of more precise testable predictions (Borgstede & Eggert, 2023; Borsboom et al., 2021; Haslbeck et al., 2022; Robinaugh et al., 2021, 2024, but also see Eronen & Bringmann, 2021; Oude Maatman, 2021, for opposite opinions). By using formal models, researchers can translate verbal theories about psychological phenomena into mathematical representations of how psychological processes in an individual system interact and evolve. Some early formal models in psychology include the classical conditioning model by Rescorla and Wagner (1972), the hand movement model by Haken et al. (1985), and the cognitive development model by van Geert (1991).

One type of formal model that was recently proposed is the formal dynamic model defined using ordinary differential equations (ODEs) or stochastic differential equations (SDEs). A distinct feature of these models is their ability to be simulated, allowing researchers to explore the dynamic characteristics of the system in different contexts and compare it to the real-life psychological phenomena that inspired the theory (Borsboom et al., 2021; Robinaugh et al., 2021). Many formal dynamic models have been proposed in the past years (Burger et al., 2020; Schöller et al., 2018; van Dongen et al., 2025; Wang et al., 2023), and some researchers have also advocated using formal dynamic models to integrate interpersonal and intrapersonal theories (Borsboom & Haslbeck, 2024) and test the effectiveness of psychotherapies *in silico* (Ryan et al., 2025).

Most of the available studies on formal dynamic models solely rely on simulations for generating predictions and model evaluations. Usually, simulation results are used to either qualitatively compare it to the phenomena observed in real life (Burger et al., 2020; Robinaugh et al., 2024; Schöller et al., 2018) or make comparisons between some statistical indicators of

the simulated data and real-life data (Haslbeck et al., 2022). An ultimate validation of the formal model would be the case in which an as-yet-unknown phenomenon is predicted, which is subsequently corroborated to exist in an empirical study (e.g., Simmering et al., 2008; Spencer et al., 2001). Simulation-based approaches are intuitive and easily built on the existing toolkits of psychologists, yet they also have considerable shortcomings. One problem is that simulations do not provide direct insight into what the critical components and interactions are that give rise to a certain phenomenon (Cui et al., 2023). When the simulation output is in accordance with real-life observations, we do not know whether this only holds for a very specific set of function forms and parameter values, or whether it also generalizes to a wider range of (similar) conditions. When the simulation output is not in line with the real-life phenomena, we do not know what the problem might be, nor do we know what we can do to solve it. Those shortcomings may limit progress in studying the potential of formal dynamic models in advancing psychological theories.

Several other fields of science, such as biology and physics, have a much longer history of applying dynamic models for advancing theories, and they have developed concrete mathematical methods for analyzing dynamic models that we can also apply in psychology. One helpful technique, for example, is the *phase plane analysis* (Kuznetsov, 2023). This technique uses both mathematical calculations and graphical representations to find out the equilibria of the system and studies the evolution of a system from a given starting place. Another technique is the *bifurcation analysis* (Kuznetsov, 2023), which focuses on the parameter space instead of the variable space. By using this technique, we can systematically investigate how the parameter setting changes the stability features of the system and explore the qualitatively different patterns of the system. To our knowledge, only a few psychological studies in formal dynamic models have applied a simple version of phase plane analysis (Cui et al., 2023; Cui, Olthof, et al., 2025; Dablander et al., 2023; Robinaugh et al., 2024), whereas bifurcation analysis has not been used in the field of psychology so far.

Therefore, in this tutorial, we aim to provide a comprehensive introduction to both techniques and illustrate their usage with concrete examples. The graphical tools that we introduce here, should not be seen as replacements for the current simulation-based approach. Rather, we see those different approaches as complementary, the graphical tools are important for understanding the role of elements and interactions in a deterministic setting, whereas a simulation-based approach can better accompany situations with noise and when there is a rather large number of interacting elements involved. To improve the readability of the manuscript, we limit the amount of mathematical derivations and provide graphical tools to

reduce potential barriers to applying them to psychological systems. The tutorial is structured as follows. First, we introduce the basic ideas of formal dynamic models in psychology and outline two example models. Second, we explain the two techniques, namely the phase plane analysis and the bifurcation analysis by analyzing these two models. Third, we illustrate the usage of those two techniques by applying them to the two example systems. Finally, we discuss the role of the analyses in modeling and reflect on the purpose of modeling and the limitations of such approaches.

Introduction of Formal Dynamic Models in Psychology

A formal dynamic model describes how the state of a system evolves over time, using formal, mathematical language. The most common form of a formal model in psychology is based on the stochastic differential equation (SDE), which takes the general form of $dx/dt = f(x)dt + g(x)dW$, where x represents all relevant variables, $f(x)$ represents the deterministic part of the changing rate of the system, and $g(x)$ represents the stochastic part of the system. Analyzing SDEs is relatively difficult but analyzing the deterministic part of them $dx = f(x)dt$, called ordinary differential equation (ODE), is often much easier. Therefore, we first focus on the ODE skeletons of the models.

In the psychological context, variables may be emotions, perceptions, physiological states, or any other quantities that are deemed important for a certain research question (but also see discussions by Kalis & Borsboom, 2020; Oude Maatman, 2020). In the panic disorder model by Robinaugh et al. (2024), for example, the core variables in the model are perceived threat (PT) and physical arousal (A). According to previous theories by Clark (1986), the two variables can strengthen each other, perceived threat increases when physical arousal is higher, and vice versa, which forms a vicious circle, potentially leading to panic attacks. The theory does not specify the exact form of the function. Therefore, various types of functions may apply. Here we first show the function form by Robinaugh et al. (2024),

$$\frac{dA}{dt} = r_A(S_{PT,A}PT - A - H), \quad (1)$$

$$\frac{dPT}{dt} = r_{PT} \left(\frac{1}{1 + \exp(-k_{PT}(A - h_{A,PT}))} - PT \right), \quad (2)$$

$$\frac{dH}{dt} = r_H \left(\frac{1}{1 + \exp(-k_H(A - h_{A,H}))} - 0.5H \right). \quad (3)$$

The functions between the two variables have different forms, one is a straight line, and the other is an S-shaped curve. In later work by the same authors (Robinaugh et al., 2021), various function forms were tried, and they found that the combination shown in Equations 1-3 was the only one that produced the expected panic attacks. We will perform a detailed analysis of this conclusion in subsequent sections. The original model by Robinaugh et al. (2024) also involved many other variables. We will introduce some of them sequentially in later sections, along with our analysis, to emphasize their specific roles in the model performance. For now, we focus on these three variables. A simulation example using default parameters (Table 1) is shown in Figure 1a. The spikes shown in this figure represent panic attacks, characterized by short periods of high physical arousal (A) and perceived threat (T).

Table 1. The default parameter values for the panic disorder model.

Parameter	Value
r_A	0.5
$S_{PT,A}$	1
r_{PT}	1
S	1
k_{PT}	$20 - 10 * 0.1^S$
$h_{A,PT}$	0.25^S
r_H	0.05
k_H	20
$h_{A,H}$	0.4

Note. All model parameters are taken from the PanicModel package associated with Robinaugh et al. (2024). The package can be accessed from <https://github.com/jmbh/PanicModel/>.

The panic disorder model has been used for illustration in many previous articles (e.g., Cui et al., 2023; Haslbeck et al., 2022; Robinaugh et al., 2021). To show the broad applicability of the tools introduced here, we will also analyze another recently published formal model, namely the model of suicidal ideation by Wang et al. (2023). In this model, the authors aim to explain the observation that suicidal thoughts often have rapid onsets, short durations, and

are close to zero most of the time, although the level of external stressors and aversive internal states may fluctuate a lot. The core variables in the model are aversive internal states (A), urge to escape (U), and suicidal thoughts (T). Their relationships are mathematically expressed as follows:

$$\frac{dA}{dt} = b_2A(K - A) + a_2S - d_2T, \quad (4)$$

$$\frac{dU}{dt} = -c_3U + b_3A, \quad (5)$$

$$\frac{dT}{dt} = -d_4T + \frac{1}{1 + \exp[-c_{41}(U - c_{42})]}, \quad (6)$$

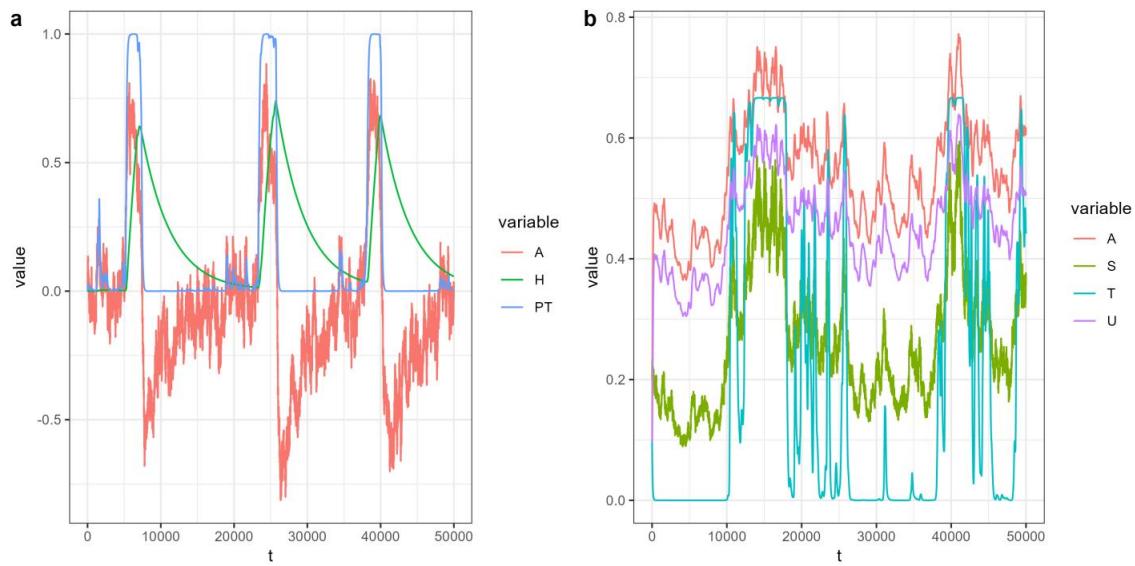
Where S represents stressors, and other letters represent model parameters. The default parameter values of this model are summarized in Table 2, and a simulation example of the variables using the default parameters is shown in Figure 1b.

Table 2. The default parameter values for the suicidal ideation model.

Parameter	Value
a_2	1.5
b_2	1.5
d_2	0.1
K	0.1
b_3	2.5
c_3	3
d_4	1.5
c_{41}	90
c_{42}	0.5

Note. All model parameters are taken from the code associated with Wang et al. (2023). The code can be accessed from https://github.com/ShirleyBWang/math_model_suicide.

Figure 1. Simulation examples for (a) the panic disorder model and (b) the suicidal ideation model. We used the default parameter values shown in Table 1 and Table 2 for the simulations. Random noise was added to specific variables according to the original specifications. For the panic disorder model, Gaussian noise with $sd = 0.01$ was added to A. This is a simpler form of noise than the one used by Robinaugh et al. (2024) as the form of noise is not the central topic of investigation of the current work, and we found using this simpler form of noise in this model to produce similar outputs. In addition, we removed the range restriction in the original model to avoid artificial influence on the mathematical property of the dynamic system. For the suicidal ideation model, we used the same form of geometric Brownian noise as used by Wang et al. (2023).

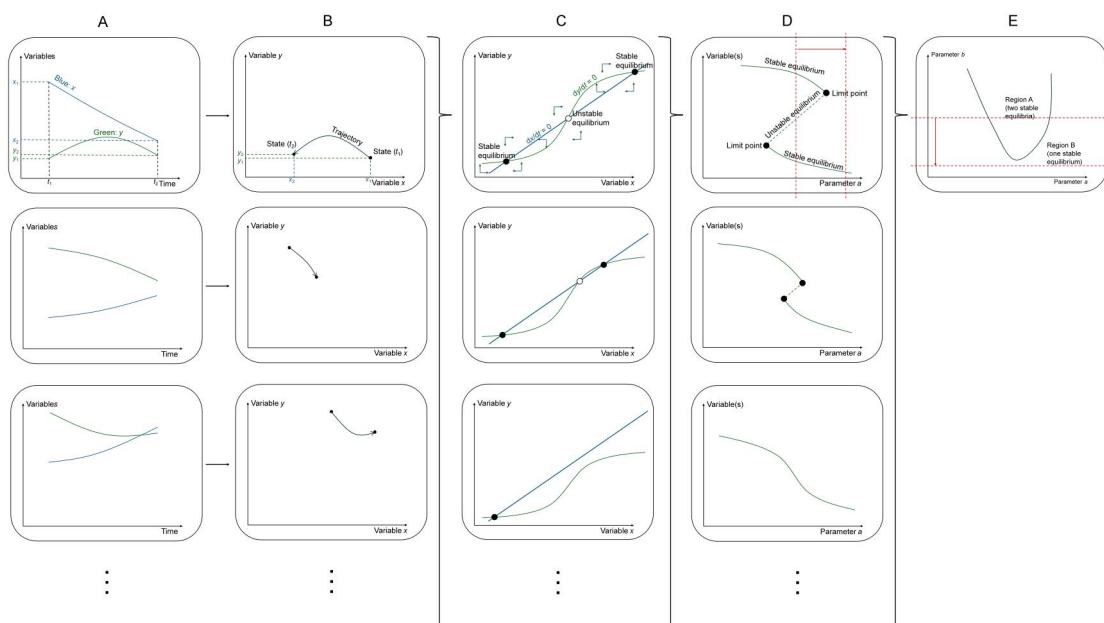


Now we will move to the next section, in which we introduce the basic concepts of the two graphical tools.

Two Graphical Tools

In this section, we introduce important concepts used in two graphical tools for the analysis of the dynamics of formal models. We introduce an easy-to-use R package with a graphical interface, which makes it possible to analyze more complex systems automatically through software. A brief illustration of the tools is shown in Figure 2. As we can already see there, the graphical tools are shown in 2D spaces, which means that we can only analyze two variables or parameters at the same time. This does not mean that those methods can only be applied to dynamic systems with very few elements. As we will show later in concrete examples, we can use those graphical tools to analyze various parts of the system step-by-step and eventually reach a more comprehensive understanding of the dynamic features of the system as a whole.

Figure 2. Illustration of several steps of the phase plane analysis and the bifurcation analysis. Column A: simulated time series of variables x and y . Column B: various trajectories on the phase plane of variables x and y . Column C: phase plane analysis with nullclines and equilibrium points. Analysis results with various parameter values of a are shown. Column D: one-parameter bifurcation analysis of the parameter a . Only the values of variable y are shown for simplicity, while in actual analysis the values of multiple variables can be shown together. Analysis results with various parameter values of b are shown. Column E: two-parameter bifurcation analysis of both parameters a and b .



The first graphical method is *phase plane analysis*. Here the term phase plane can be understood as a two-dimensional plot¹, in which each axis represents the value of a variable. Each point on the plot reflects a specific state of the system, characterized by the combined values of two variables, and a trajectory of the system can be shown as a line on the phase space (Figure 2, Column B). Compared to the time series plots (Column A), the phase plane plots represent information about variables while omitting time. Phase plane analysis aims to understand the direction of change, or how the state of the system evolves over time. For example, we can investigate in which range the variable of a system will increase, and when it will decrease (Column C). This is important for many psychological questions. If we know how to increase a person's positive emotion, we can understand how to help the person escape the depressed state; if we know how to make the perceived threat decrease, we can understand when a person having a panic attack can gradually recover to the normal state. In order to

¹ Strictly speaking, this is the definition of 2D phase space, which is rather easy to interpret and informative. Therefore, in the remaining part of this article, we use the term phase space to represent 2D phase space by default.

know when a variable increases or decreases, we first need to find out the *nullcline*, the line on which the variable keeps the same value. The nullclines also set the boundaries of the increase and decrease region. For example, in Column C, the blue and the green curves are nullclines for the two variables, x and y , respectively. For the region above the blue curve, x increases, and for the region below the blue curve, x decreases. Similarly, for the region above the green curve, y decreases, and for the region below the green curve, y increases. At the intersection points, both variables are kept constant. Hence, those points are the *equilibrium points* of the system (Column C). An equilibrium point can be stable or unstable. For a stable equilibrium point, the system can return to it after a small perturbation, but for an unstable equilibrium point, the system will go away from it after a small perturbation. Those two situations are often illustrated with the ball-and-landscape metaphor, in which a stable equilibrium is like a ball in the lowest position of a valley, and an unstable equilibrium is like a ball at the highest position of a hill. Therefore, in the presence of noise, the system is more likely to remain close to stable equilibria and stay away from unstable equilibria.

The equilibria of a dynamic system can be altered by changing the *parameters*. Those parameters often define the strength of influence from one variable to another, or some intrinsic features of some variables' dynamics. In principle, if we want to investigate the influence of parameters on the system's dynamics, we can perform a series of phase space analyses for various parameter values (Figure 2, Column C). However, this approach would be quite cumbersome in practice. A more concise method for summarizing how parameters influence the system's dynamics would be useful. Looking at Column C, we can find that a minor change in the parameters only affects the *position* of equilibria, which is depictive in most cases. Only in some special cases, a change in the parameters can lead to changes in the stability of equilibria, the appearance of a new equilibrium point, or the disappearance of an existing equilibrium point. Those situations, where a small change in the parameters leads to a qualitative change in equilibrium points, are called *bifurcations*, and the method to evaluate whether a bifurcation is happening is called *bifurcation analysis*.

In a one-parameter bifurcation analysis, a plot is drawn to show the system's position of the equilibrium points as a function of the parameter value (Figure 2, Column D). The procedure of drawing such a plot consists of performing the phase plane analysis (Column C) many times, each time with a specific parameter value, and then putting the variable values of the corresponding equilibria states as dots in the one-parameter bifurcation plot on the corresponding parameter value (marked as red dotted lines on the first subplot of Column D). In the opposite direction, we can also infer the position of the equilibrium points in the phase

plane analysis (Column C) as slices from the one-parameter bifurcation analysis (the first subplot Column D). In the first two subplots of Column D, we see two bifurcation points, at which a stable equilibrium and an unstable equilibrium point merged and disappeared. The behavior of the system changes qualitatively when the parameter a crosses those two points. In the range between the two parameter points, the system has two stable equilibria, but out of this range, the system only has one equilibrium point. The bifurcation points may take different forms than the one shown in our example. Much mathematical work has been done in the 20th century to classify the types of bifurcation points. A more comprehensive overview of the types of bifurcation points is available in Thom (1975) or Abraham and Shaw (1992).

A dynamic model often contains multiple parameters. If we want to investigate how two parameters jointly influence the dynamics of the system, we can perform multiple one-parameter bifurcation analyses with various values of the second parameter, as in Column D of Figure 2. However, most of the bifurcation analyses will only have small, quantitative differences, as is the case for the phase plane analyses. To oversee how two parameters influence the behavior of the system, we can perform a two-parameter bifurcation analysis instead. One-parameter bifurcation analysis can then be seen as taking slices of the two-parameter bifurcation, depicting the bifurcation points (compare Column D and Column E along the red dotted lines). The curve in a two-variable bifurcation analysis divides the parameter space into two regions. Each region has different numbers or types of equilibrium points. In the example of Column E, the region above the curve has two stable equilibria, whereas the region below the curve has only one stable equilibrium.

Conducting phase-plane analysis for more variables or bifurcation analysis for more parameters together is theoretically possible, but would require drawing plots in higher dimensions, which are difficult to comprehend (but see de Boer, 2024, for an application for 3D phase plane analysis). We therefore focus on the cases up to two dimensions in this article. To make the analysis readily available for psychologists we introduce the R package *deBif*, developed by de Roos (2025). The *deBif* package provides an easy-to-use graphical interface that can be used to perform phase plane analysis and bifurcation analysis implemented by the functions *phaseplane()* and *bifurcation()*, respectively. The *phaseplane()* function can produce results like Columns A and B in Figure 2, and the *bifurcation()* function can produce results like Columns C and D. Users can run those functions to evoke corresponding Shiny apps (Chang et al., 2024), and from there, perform various analyses interactively. The functions of the package, important parameters, and tabs are summarized in Table 3 and Table 4. We will illustrate the

usage of the functions, the graphical interface, and specific analysis steps together with examples.

Table 3. Key functions of the deBif package, the tabs of the corresponding Shiny apps, and explanations.

Function	Tab	Explanation
phaseplane()	Time plot	Show the simulated time series from the given starting point.
	Nullclines	Show the nullclines.
	Steady states	Show the steady states, and their stability, together with the nullclines.
	Vector field	Show the nullclines, steady states, and the vector field (representing the direction of change) of different regions.
	Trajectories	Show the simulation trajectory (in Time plot) on the phase plane, together with the nullclines and the steady states.
	Portrait	Show multiple trajectories starting from different states, together with the nullclines and the steady states.
bifurcation()	Time series	Show the simulated time series from the given starting point.
	1 parameter bifurcation	Perform one-parameter bifurcation analysis (starting point required) ¹ .
	2 parameter bifurcation	Perform two-parameter bifurcation analysis (starting point required).

Note. ¹ Researchers often need to supply a new starting point for those calculations, preferably from a steady state.

Many formal models contain more than two variables and more than two parameters. For those models, we still recommend that researchers focus on at most two variables and two parameters at the same time to enhance the interpretability of the model results. The treatment of other variables depends on the relative time scale of the system (Cui, Hasselman,

et al., 2025; Hasselman, 2023). Variables evolving on faster timescales than the focus variables will reach their steady state quickly, so we can assume that their time derivative is zero, and therefore solve the fast variables as a function of slower variables to reduce the number of variables in the analysis (Bertram & Rubin, 2017; Okino & Mavrovouniotis, 1998). Variables that evolve much slower than the focus variables can be put as parameters in the model instead of variables, to investigate how changing the values of these slow variables affects the dynamics of the focus variables. We will show further illustrations with concrete examples in subsequent sections.

Table 4. Key parameters for the functions in the deBif package.

Parameter	Instruction
model	An R function that describes the dynamic system. The function should take three parameters, <i>t</i> , <i>state</i> , and <i>parms</i> , where <i>t</i> represents time, <i>state</i> is a named vector of all the variables, and <i>parms</i> is a named vector of all the parameters. The output of the function should be a vector of all the derivatives, in the order of the input variables. This model parameter is rather abstract. Readers may use the code we share as the starting point and modify it from there.
parms	A named vector of all parameters. This parameter can be overwritten during the execution of the Shiny app.
state	A named vector of the initial values of all the variables. This parameter can be overwritten during the execution of the Shiny app.
...	There are other parameters to modify the graphic output in the Shiny app. See the help document of the functions for details.

In what follows, we analyze the two systems introduced earlier. To prioritize clarity and brevity, we focus on interpreting the analysis results in the main text and leave detailed commands and procedures to produce the results in Online Supplementary Materials (available at <https://osf.io/ym9vt/>).

Example 1: Analysis of the Panic Disorder Model

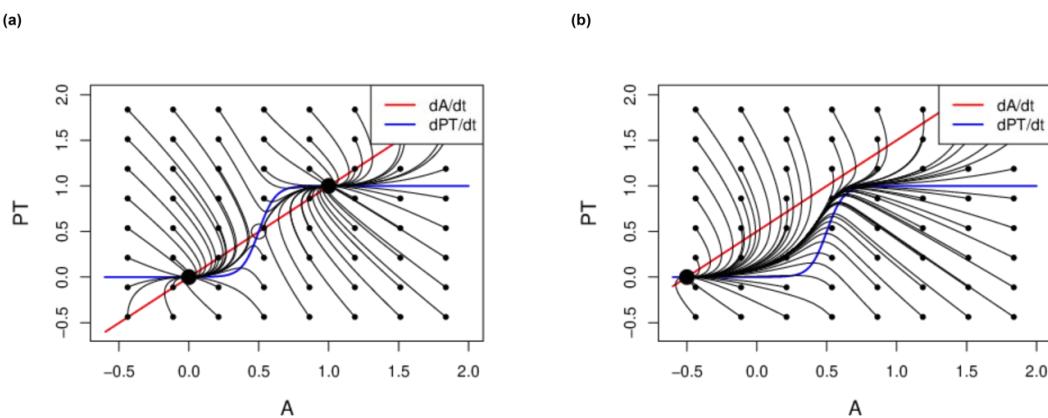
The panic disorder model is specified by Equations (1-3), with the default parameters in Table 1. There are three core variables in the system: physical arousal (*A*), perceived threat (*T*), and homeostatic feedback (*H*).

Analysis with H as a Parameter

From both the default parameter values (Table 1) and the simulations (Figure 1a), we can see that H changes at a much lower rate than the other two variables A and T . Hence, we first look at the faster time scale constituted by A and T and treat H as a parameter.

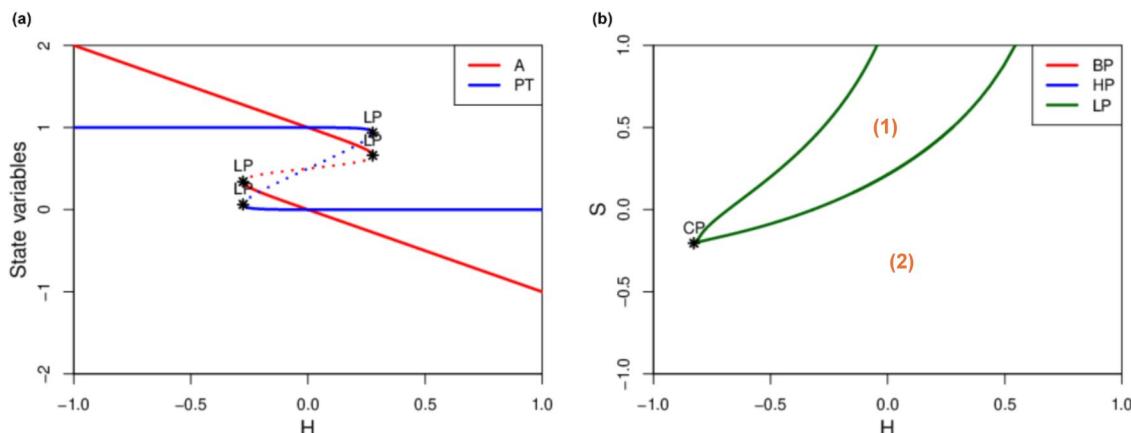
We first investigate the results with the default parameter setting (Figure 3a). The two nullclines represent the states where A and PT remain unchanged. The red line is the nullcline for A . For all the states to the lower right direction of this line, A decreases, and vice versa. The blue line represents the nullcline for PT . For all states to the upper left direction of this line, PT decreases. The two nullclines cross at three points, of which the first and the third one are stable equilibria, and the middle one is an unstable equilibrium point. The two stable steady states represent the healthy and panic states of the system, respectively. From this analysis, it seems that the system can reside in either state, yet from the simulations (Figure 1), we see that the system can only move to the panic state for a short time, and after that, quickly falls back to the healthy state. This needs to be explained by changing the value of H . After increasing H from 0 to 0.5 (and enlarging the display range of the variables accordingly, Figure 3b), the dynamic features of A and PT change. From Figure 3b, we can see that the nullcline of A (the red line) moves to the upper left direction. Now the two nullclines only intersect at one point, which is a healthy state as both A and PT have a rather low value at this point. Therefore, the system can only gravitate towards a healthy state when H is high enough.

Figure 3. Phase plane analyses results of the panic disorder model, with H as a parameter. (a) $H = 0$; (b) $H = 0.5$. The plot shows the nullclines (the blue and the red lines), the steady states, their stability (solid points for stable steady states and hollow points for unstable steady states), and some trajectories.



The one-parameter bifurcation analysis results are shown in Figure 4a. In this analysis, we can see that both variables A and PT have two stable states between two bifurcation points, where $-0.28 < H < 0.28$ (the specific coordinate values can be extracted from the Shiny application). If $H < -0.28$, the system only has the panic stable state, and if $H > 0.28$, the system only has the healthy stable state. As H increases, the value of A for the stable states decreases both for the panic state and the healthy state, whereas the value of PT does not change much. Therefore, the bistability of the system only appears when homeostatic feedback is in a certain range, and the strength of homeostatic feedback also influences the level of physical arousal at the equilibrium states.

Figure 4. Bifurcation analyses results of the panic disorder model. (a) The one-parameter bifurcation analysis result, with H as the parameter. “LP” means limit points. Solid lines represent stable steady states, and dashed lines represent unstable steady states. (b) The two-parameter bifurcation result, with H and S as parameters. “CP” means cusp points, and “BP”, “HP”, and “LP” represent different types of bifurcation points, which we did not elaborate on in this tutorial. “HP” means Hopf bifurcation points (not present for this system), “BP” means branching points (not present for this system), and “LP” means limit points. The points labeled with numbers (e.g., 1, 2) were added by the author for clarity and are marked in orange. These labels are not part of the raw software output.



We can now include a second parameter in the analysis, S (Table 1). This parameter represents an important psychological construct, namely arousal schema. It shows how much a person perceives high physical arousal as a sign of threat. In the original model by Robinaugh et al. (2024), arousal schema is a variable that changes at a much slower rate, and high arousal schema is seen as the key mechanism explaining panic disorder. We omitted the dynamic equation of S in our analysis as it takes an irregular form with if-else conditions and depends

on the history of the system, thus cannot be analyzed with the methods introduced in this tutorial. Nevertheless, we can still use S in the two-parameter bifurcation analysis and see how this variable changes the stability of the steady states. In Figure 4b, we can see that the two green curves together divide the parameter plane into two regions, (1) and (2). The two regions have different numbers of stable steady states. From the results of the one-parameter bifurcation analysis, we can infer that in the small triangle-like area (1), the system has two stable states, but in region (2), it has only one stable state. Specifically, when arousal schema is low, there is always only one stable state, but when arousal schema is high, the strength of the homeostatic feedback determines whether there are one or two stable states in the system and whether the stable state is a healthy state or a panic state. Thus, the stability feature of the system is dependent on the joint influence of both variables, and the panic disorder phenomenon can only be observed when arousal schema is high enough and homeostatic feedback is moderate.

Analysis with H as a Variable

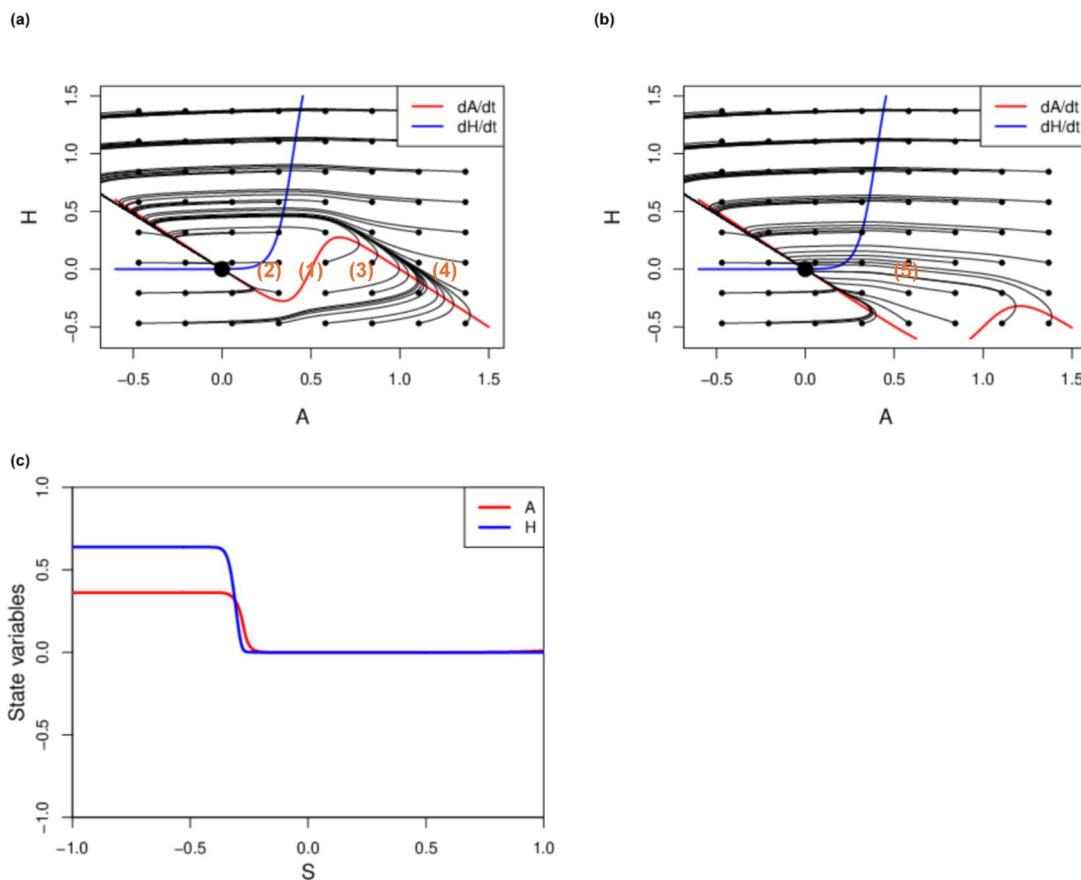
We then investigate the dynamics of H . To do so, we need to eliminate one of the fast variables to make the total number of variables two. This can be done by assuming that one fast variable is always at its equilibrium (i.e., the variable's derivative is zero). The rationale for this treatment is that the fast variables approach their equilibrium much faster than H , hence assuming that they are always at the equilibrium does not affect the results much (Bertram & Rubin, 2017; Okino & Mavrovouniotis, 1998). Here we arbitrarily chose to solve PT (solving A would give similar results) from Equation (2). Solving $dPT/dT = 0$ results in:

$$PT = \frac{1}{1 + \exp(-k_{PT}(A - h_{A,PT}))}. \quad (7)$$

This equation can be used with the panic disorder model as presented in Equations 1-3 to run the phase plane analysis and the bifurcation analysis. The results are depicted in Figure 5. We first look at the phase plane analysis results with arousal schema fixed at a moderate level, $S = 0.5$ (Figure 5a). Although there is only one stable state in the system, represented by the large black dot close to $(0,0)$, the system may move in opposite directions under noise: When A is above a certain threshold (1), it does not directly decrease towards the equilibrium of the system (the black dot). Rather, A first increases until its value reaches its nullcline (the red line) and then turns around and decreases again. This is only possible when the nullcline of A (the red line) has this particular curvilinear shape. If we look along the line of $H = 0$, there is first a region (2) where A decreases to make sure that the steady state is stable, then a region (3) where A increases to make it possible to show spikes, and then a region (4) where A decreases

again to ensure A will not increase too much. Also, the change rate of H is low enough compared to A to make sure that the dynamic trajectories (the black curves in Figure 5a) first almost go along the horizontal direction, so that A can show a clear spike.

Figure 5. Phase plane analyses and bifurcation analyses results of the panic disorder model, with H as a variable. (a) phase plane analysis result using the default parameter ($S = 0.5$); (b) phase plane analysis result with $S = 0$; (c) one-parameter bifurcation analysis result. The points labeled with numbers (e.g., 1, 2) were added by the author for clarity and are marked in orange. These labels are not part of the raw software output.



We then use Figure 5b to see why the system does not have panic attacks when S is fixed at 0. The results show that the nullcline of A (the red line) moves to the lower right direction, and compared to Figure 5a, now the intersection point (1) disappears at (5). Therefore, even if there is noise driving A to higher values around (5), as the state of the system does not cross the nullcline of A , A will always tend to decrease, which means that there will be no panic spikes anymore (compare with the trajectories starting from (3)). From the bifurcation analysis result in Figure 5c, we can see that as long as S is around the range between 0 to 1 (which is always the case in the complete model by Robinaugh et al., 2024), the position of the

stable state hardly changes. Therefore, the level of physical arousal and the perceived threat of the healthy stable state will not change following the changes in S .

Brief Discussion

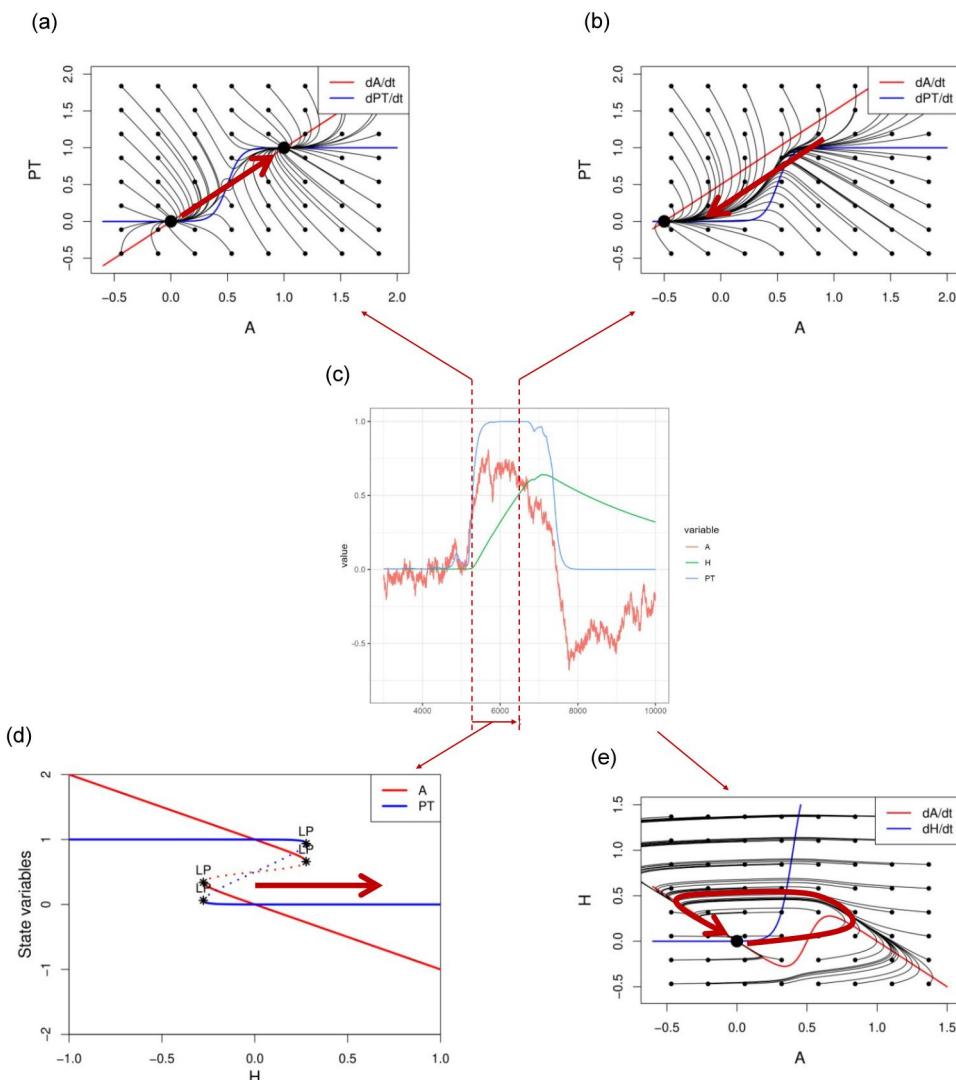
In the analysis example shown above, we investigated the dynamics of the panic disorder model on two different time scales. For the first one, the fast time scale of A and PT , we found that both H and S can influence the stability of the system. When S is low, the system always has only one stable state, the healthy state (Figure 4b). When S is higher, the stability of the system will depend on H (Figure 4b). If H is at its default value (Figure 3a), the system can have two stable states, one being healthy and the other being in panic. Therefore, with some noise, it is possible for the system to move from the healthy state to the panic state. But if H is higher (Figure 3b), the panic state will become unstable again, and the system will fall back to the healthy state. The required H value to make the panic state unstable depends on S . The higher S is, the higher H must be to make the panic state disappear (Figure 4b). The second part of the analysis focused on the slower time scale of H . We found that, because of the shape of the nullclines, if A is higher than a threshold (1), it will first increase before decreasing back to the healthy stable state, making panic spikes appear in the time series of A . However, in the long run, the system always tends to go back to the healthy state as there is only one stable state in the system.

In Figure 6, we provide an illustration to link the graphical analysis results back to the simulated time series. For Figure 6c, we zoom into a segment of the time series shown in Figure 1a which represents a panic attack. The two panels above, which are the phase space analysis results for $H = 0$ (Figure 6a) and $H = 0.5$ (Figure 6b), correspond to the stability characteristics of A and PT with different H values. When $H = 0$, the subsystem of A and PT is bistable, thus noise is possible to drive the system from one stable equilibrium to another. However, later H increases, making the subsystem of A and PT monostable, forcing the system to go back to the only remaining equilibrium. This relationship between H and the stability of A and PT can also be found in the one-parameter bifurcation analysis, shown in Figure 6d. Finally, in the phase space analysis on the slower time scale (Figure 6e), we can see that after A increases from random noise, the system's state has a tendency to go through a long detour before returning to the previous equilibrium, which corresponds to the phenomenon of panic attacks.

In summary, the key dynamic features giving rise to the spike-like behavior of the panic disorder model are the following: (1) The nullclines of A and PT may form one or two stable states depending on H . When H is high, two nullclines detach at the right side, making the stable panic state disappear. (2) The increase of A and/or PT can lead to a slow increase in H .

We may change the specific function forms and still create similar behaviors in the system, as long as the key conditions are met. Robinaugh et al. (2021) tried several alternative forms of the functions of A and PT and suggested that only the combination of a linear form in the dynamic function of A (Equation (1)) and the S-shape form in the dynamic function of PT (Equation (2)) can produce behaviors like panic attacks. Based on the analysis above, however, we can conclude that this statement is likely not entirely correct. We infer that if both equations are S-shaped, the model can also show similar behavior. Indeed, we successfully formulated a model with both S-shaped function forms that matches the phenomenon of panic disorder, detailed in Appendix G. The graphical tools we propose here, thus lead to a clearer explanation of why, or what exact mechanism, can lead to specific behaviors in the simulation results.

Figure 6. The relationship between the graphical analysis outputs and the simulated time series for the model of panic disorder. The bold, red arrows in panels (a), (b), (d), and (e) correspond to various features of the time series in panel (c).



Example 2: Analysis of the Suicidal Ideation Model

To illustrate the generalizability of the proposed methods, we now move to the second example, the suicidal ideation model by Wang et al. (2023). The model contains three core variables: aversive internal states (A), urge to escape (U), and suicidal thoughts (T). The model is specified by Equations 4-6, with the default parameter values shown in Table 2.

The key feature of the model is that, although the aversive internal states and urge to escape have many peaks and fluctuations, suicidal thoughts are highly zero-inflated and only have a few peaks (Wang et al., 2023; also see Figure 1b). We now investigate why the system has this feature. As the system has three variables with similar time scales (Equations 4-6), we arbitrarily solve out the variable U by assuming $dU/dt = 0$. From Equation 5, we have that

$$U = \frac{b_3 A}{c_3}. \quad (8)$$

To investigate how different variables of the system react to random stressors, we use the variable representing the external stressor, S , as the model parameter in our analysis. The results are shown in Figure 7. We can see that the system always has a single stable point but at different locations. As S increases from $S = 0.2$ in Figure 7a to $S = 0.5$ in Figure 7b, the nullcline of A (the red line) moves to the upper right direction, making its intersection with the nullcline of T (the blue line) move from (1) to (2). The nullcline of T is S-shaped so that at the beginning, the intersection does not change much in the T axis (Figure 7a). After a threshold (3), however, the intersection of the nullclines moves to the second half of T 's nullcline, so the T value of the stable point suddenly increases (Figure 7b). We can see this trend more clearly from the bifurcation analysis in Figure 7c. Here, as S increases, A increases smoothly, but T increases abruptly. Therefore, if S randomly fluctuates, we can observe A follows closer with S , whereas T mostly stays around zero, with occasional large spikes. This follows the simulation results shown earlier (Figure 1b). Again, in Figure 8, we show another illustration that emphasizes the relationship between the simulated time series and the results from phase plane analyses and a bifurcation analysis.

In summary, the difference between the dynamics of A and T can mainly be explained by the shape of their nullclines. The nullcline of A is close to a straight line with a large, negative derivative and moves smoothly in the right direction, and the nullcline of T is close to a S-shaped curve that does not change with S . Again, in Appendix G, we formulated an alternative model in which the dynamic function of A takes a linear form instead of a quadratic form, while retaining the relationship between the nullclines. The characteristics of this model closely resemble those of the original model of suicidal ideations, yet the alternative model has

a simpler form. Therefore, using the analysis methods presented in this tutorial, we can see that the quadratic expression in the original model (Wang et al., 2023) is not necessary to produce the phenomenon of interest.

Figure 7. The phase plane analysis and bifurcation analysis screenshots during the analysis of the suicidal ideation model. (a) is the phase plane analysis result with $S = 0.2$; (b) is the phase plane analysis result with $S = 0.5$; (c) is the one-parameter bifurcation analysis result with S as the parameter. The points labeled with numbers (e.g., 1, 2) were added by the author for clarity and are marked in orange. These labels are not part of the raw software output.

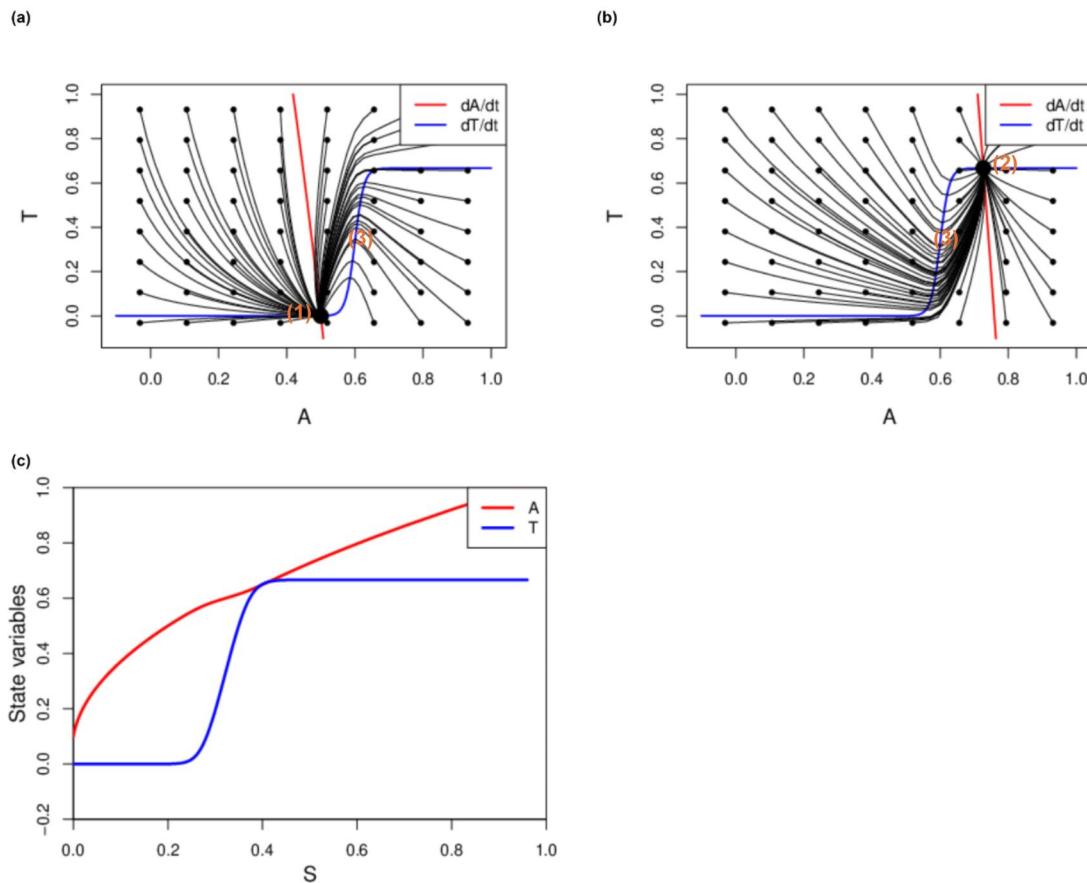
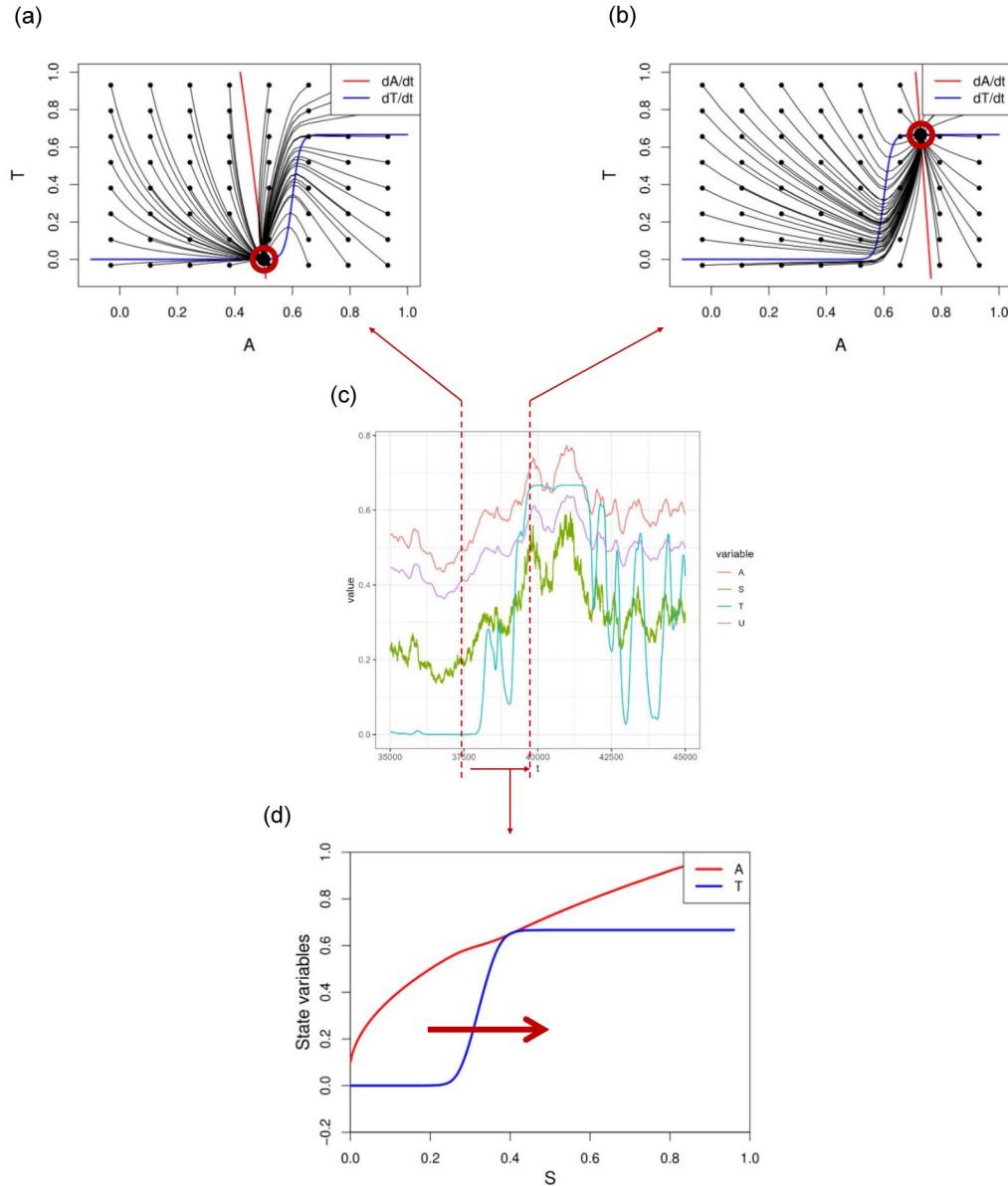


Figure 8. The relationship between the graphical analysis outputs and the simulated time series for the model of suicidal ideations. The bold, red circles and arrows in panels (a), (b), and (d) correspond to various features of the time series in panel (c).



General Discussion

In this tutorial, we introduced two important graphical tools for dynamic systems, namely the phase plane and bifurcation analysis. We explained the meaning of several important plots of the analyses, the related code, and the procedure, and we demonstrated the analysis with two specific formal dynamic models in psychology. For the panic disorder model, we found that the key feature of the system, namely the spikes of panic attacks, come from fast, potentially bistable dynamics between physical arousal (A) and perceived threat (PT), and a slower

variable, homeostatic feedback (H), that controls the dynamics of A and PT . When H hovers around the baseline level, the system of A and PT is bistable, so it may go close to the panic state under noise. However, this will lead to an increase in H , which makes the system of A and PT monostable leaving only the healthy state, so that after a short period of time in the panic state, the system will quickly return to the healthy state. The variable that is even slower, S , controls the required level of H to make A and PT monostable, thus controlling whether the system is likely to show panic attacks. For the suicidal ideation model, we found that the different reactivity of aversive internal states (A) and suicidal thoughts (T) towards external stressors comes from the shape of their nullclines. The nullcline of A is close to a straight line and moves smoothly with increasing external stressors (S), whereas the nullcline of T is an S-shaped curve. When S increases, the intersection point moves along the nullcline of T , leading to a smooth change in A and an abrupt change in T .

For both models, we proposed alternative forms of dynamic equations that are different from the original specifications but still meet the key dynamic features we found. This underscores the importance of the graphical tools we introduced in investigating the underlying dynamics producing a certain psychological phenomenon. Both adapted models showed similar behaviors compared to the original models. Therefore, in contrast to previous arguments by some researchers (e.g., Robinaugh et al., 2021), the alignment between simulation results and real-life observations does not necessarily provide robust support for the exact function form used in formal models. This may render further deductions from those models (e.g., *in silico* experiments of treatments, Ryan et al., 2025) less reliable. Instead, we argue that greater emphasis should be placed on the key dynamic features of the model.

The aim of introducing those graphical methods to the realm of psychology is to gain a better understanding of the mechanisms underlying certain dynamics. In other words, we do not only want to have computational models that behave similarly to real-life systems but we also want to understand which characteristics of the dynamic functions can give rise to a specific phenomenon, hence feature in the time series. Formal dynamic models are often defined with a set of functions and parameters. If the simulation results show the key features of the dynamic model, we can only infer that this specific combination of function forms and parameter values may be a reasonable candidate model for the phenomenon of interest. The amount of information we can gain from a single instance, though, is rather little, as there are many other possibilities of function forms and parameters that may lead to similar results. Optimally, we at least want to have a group of models featured by a set of similar characteristics that may give similar results. Many researchers claim that the benefits of using

formal models are to reduce the ambiguity of verbal descriptions and make the theory more specific (Borsboom et al., 2021; Fried, 2020; Robinaugh et al., 2021), but relying on a single simulation model also has the risk of seeing the trees instead of the forest. By focusing on the characteristics of formal dynamic models instead of specific instances of formal models, we can balance specificity and generalizability and make the results more interpretable.

Analyzing the model at the level of dynamic features also enables model comparison across fields. Several types of models are known in other fields, which may describe very different phenomena but share certain dynamic features. For instance, excitable models in biology, originating from the research on neural activation, describe how a system may leave its equilibrium and go towards another direction, showing a spike in the signal, before returning to the original equilibrium point (Edelstein-Keshet, 2005; Murray, 1989), are very similar to the panic disorder model. Thus, the conditions in excitable models may also apply to formal models of panic disorder. Alternatively, if we want to develop a new model for another type of psychological phenomenon, and we are aware of models in other fields showing similar behaviors, we can take those models as a starting point and adapt them for building psychological models. For example, the Hopf bifurcation, which describes how a dynamic system transitions from a steady state to a limit cycle and exhibits oscillatory behavior, maybe a useful starting point to build a model of bipolar disorder and investigate how a client changes from having a balanced mood to experiencing alternating periods of manic and depressive phases.

The field of dynamic modeling in psychology (and the domain of psychopathology specifically) is still in its infancy. Pioneering work in this field (e.g., the models we used as examples by Robinaugh et al., 2024; Wang et al., 2023; and other models described by Burger et al., 2020; Schöller et al., 2018; van Dongen et al., 2025) often involves a large number of variables or constructs, while the resulting dynamics tend to be relatively simple, with only a few key features of interest. At the same time, it is unclear whether these features arise from the complex interactions among elements or are primarily driven by one or two functions specified by the researcher. Having many elements and complex dynamic functions is not necessarily a problem. However, its level of complexity should be justified by the complexity of the model's outcomes.

Here, it is good to refer to the work by Levins (1966), who has defined three purposes of dynamic modeling: generality, realism, and precision. In reality, those three goals are hard to achieve within the same model. Generality emphasizes the basic understanding of the system's core mechanism. A generalizable mechanism can be transferred to many similar situations but

goes at the cost of ignoring less important elements and the quantitative predictions may not be precise. Realism prioritizes resemblance between model and real-life situations, for example, the number and nature of elements involved, and the forms of interactions. Practically, this may make the system very complex and hence difficult to disentangle the basic mechanisms that give rise to the phenomenon. It will also make the model less generalizable to other, similar situations and the predictive power may actually be weaker compared to simpler models because of the difficulty of estimating all parameters in such a complex model. Last, a model can also aim at achieving high precision in prediction. Machine learning models, for example, would be suitable for this goal, but it is good to be aware that prioritizing prediction precision usually goes at the cost of the generalizability and realism of the model.

The field of psychology currently lacks a basic understanding of the core dynamical mechanisms underlying many psychological phenomena. Striving for high realism by constructing complicated models involving many variables, might therefore be premature. Given that most simulation outputs are only qualitatively compared to empirical data, aiming for high precision also seems less relevant. We therefore advocate prioritizing the generalizability of the underlying mechanisms of the system by building simpler models with fewer core variables, facilitating investigation and interpretation. The introduced methods greatly contribute to this endeavor because they allow for a more systematic investigation of formal models, thereby expanding the potential of formal dynamic models in advancing psychological theories.

Chapter 10

General Discussion



The general aim of this thesis was to advance the analytical tools in the complexity science of psychology, focusing on three main parts: the stability of complex psychological systems, the critical transitions in clinical trajectories and their early warning signals, and the nonlinear dynamic interactions between psychological variables. I will first discuss the findings in these three parts. After that, based on the findings in this thesis, I will reflect on the current status of complex system research in psychology and discuss several challenges and future directions. Finally, this thesis will end with a narrative reflection on the PhD project as a whole.

Discussion of the Findings

Using Potential Landscapes as a Description of Psychological Dynamics

This thesis started with several chapters introducing potential landscape methods. The potential landscape is naturally defined for gradient systems, and its generalized version can be used to describe a wider range of dynamic systems. Although several types of potential landscape definitions are available (Zhou & Li, 2016), in this thesis, I focused on the definition based on the steady-state distribution, because it poses fewer restrictions on the form of the dynamic functions, and therefore, is the easiest to apply for psychological systems. Other methods may be more advantageous for certain needs, as I discussed in Chapters 2-3, but they are less suitable for psychological systems that usually contain a large number of elements with complex interactions.

Describing dynamic systems with landscapes inevitably sacrifices some information. The non-gradient part of the system cannot be described with landscapes. This may correspond to the periodic trends, for example, in bipolar disorder. Also, as landscape visualizations can only be drawn in up to three dimensions, they cannot represent dynamic features in higher dimensions. Nevertheless, the stability of states is often more important for psychological systems, as many psychological theories can be described as the characteristics of stable states (e.g., depressive states, healthy states). Therefore, the potential landscape can still serve as a useful simplification of the overall dynamics.

The landscapes for psychological systems are typically bounded within a certain range. While the system's state may transition from one phase to another, it generally remains within a plausible region of the state space and will not diverge toward infinity. This feature may be bound by measurement techniques, but may also show the intrinsic, adaptive nature of psychological systems: the mental state of a person tends to be relatively stable and often returns to previously visited states rather than moving in a certain direction endlessly. The

adaptive, stable nature of psychological systems is different from many other fields. In RNA dynamics, for example, the molecules often degrade over time (Qiu et al., 2022); in cell differentiation, cells will die out in the end (Waddington, 1966); and in image processing, the vector fields representing lens movement typically do not converge to a stable point (Ma et al., 2013). Methods in these fields typically do not assume the system's state is relatively stable. This is the main reason why we could not apply some methods from other disciplines directly but needed to modify them in Chapter 4.

The landscape from Ising networks is different from others because they are not estimated from time-series data, and therefore, do not have an intrinsic definition of dynamics. Some technical approaches, such as the Glauber dynamics (Glauber, 1963) and the Metropolis algorithm (Binder & Luijten, 2001), can be used to perform simulations from Ising networks, but they are mainly intended to approach the steady-state distribution of the system instead of resembling the actual change process. In Chapter 5, we also discussed the potential issues of using cross-sectional data to infer the stability of a system, and we restrict the inferences of such a method to the group level. Nevertheless, the method described in Chapter 5 still represents a prototypical instance to describe the stability of networks. With some adaptations, similar approaches can be applied to a wide range of network models and help to connect the statistical network models and the theoretical ideas of symptom networks (Hoekstra et al., 2024).

In this thesis, I described two methods to construct potential landscapes for empirical data, one through the vector field estimations using fitlandr (Chapter 4), and the other through network models (Chapter 5). The choice between the two methods depends on an inevitable trade-off between the amount of data available, the number of variables included in the model, and the complexity of the relationships among them. Whereas the vector field in fitlandr supports more flexible relationships between two variables, it cannot generalize to higher-dimensional cases. In contrast, Isinglandr can model multiple variables, but they can only take binary values and have logistic relationships with each other. Moreover, fitlandr is designed for idiographic, longitudinal modeling with limited data length, yet Isinglandr only supports cross-sectional data collected from a large group. Technically, it is possible to construct Ising-like longitudinal models and construct landscapes from there (e.g., Schumacher et al., in preparation), but the resulting model will deviate from the original Ising model, requiring a different estimation procedure.

Rethinking Early Warning Signals and Critical Transitions

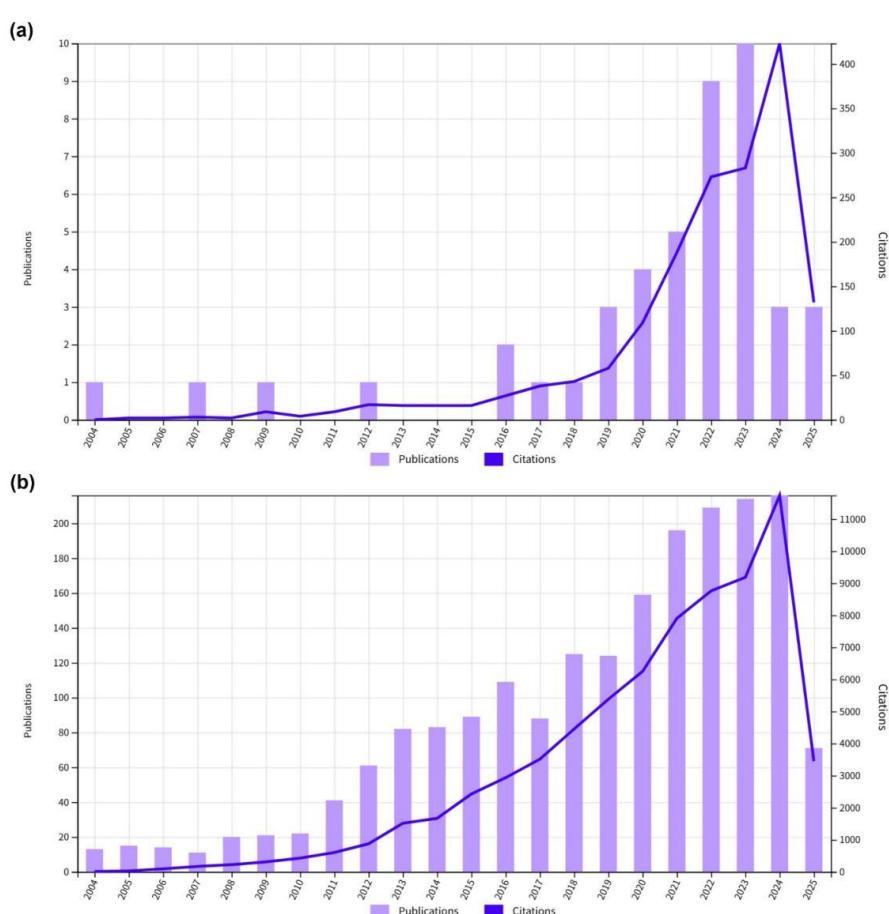
The past few years have witnessed speedy growth in the research of early warning signals in psychology (Figure 1). Recent reviews by Helmich et al. (2021, 2024) highlighted the inconsistency of results and the difficulty in applying EWSs in clinical settings. The inconsistency of EWS results may not necessarily be a problem. As argued by Olthof et al. (2020), the control parameter of the system may not always change monotonically; the system may experience a short period of instability and then fall back to the previous phase. In this scenario, although no transition is observed, EWSs are still valid signals of the instability of the system. Although this argument is valid in theory, it is not easily falsifiable by empirical studies, as the instability of the system is difficult to measure. One possible opportunity is to test whether interventions during unstable periods can yield better outcomes. This approach has been pursued recently by Olthof et al. (2025), although further investigations are still warranted.

Chapters 6 and 7 in the current thesis took a more theoretical and abstract approach to critically assess whether the research practices used in earlier studies, though seemingly intuitive, are appropriate to identify EWSs. The theoretical background of EWSs differs from most psychological theories. In the case of EWSs, specific statistical patterns, such as increasing variance and autocorrelation, can be mathematically derived under certain conditions, whereas predictions in psychology are often based on verbal reasoning and empirical generalizations. The mathematical derivation guarantees that, if all the required conditions are met, EWSs should appear before transitions. The only thing left is how to develop effective assessment methods to detect them. In real-life systems, it is difficult to verify each condition of EWSs. Nevertheless, if certain practices are clearly inconsistent with the conditions required for EWSs, the resulting statistical indicators may not be genuine EWSs. In Chapter 6, I explained several important conditions based on the theoretical deviation of EWSs and pointed out some inconsistencies between theory and common research practices in the field. As a result, many EWSs reported in previous studies may not be valid. While the changes in statistical indicators, such as increases in variance or autocorrelation, may be real, they are not likely to stem from system destabilization.

But how should we understand clinical transitions if the dynamic instability is not the underlying course? I offered a preliminary answer to this question in Chapter 7, where I introduced multiple types of transitions and their underlying mechanisms. In complex dynamic systems, numerous factors may serve as the direct “cause” for a critical transition. Understanding each cause at every level of various systems is clearly not feasible. Therefore, I

limited the discussions in Chapter 7 to a certain level of abstraction. Instead of looking at specific causes, researchers in dynamic systems have summarized several major causes of transitions, including the destabilization of attractors (B-tipping), random fluctuations (N-tipping or N-diffusion), or the rapid movement of the position of the attractor (R-tipping). These types of tipping have distinct theoretical implications and may lead to different observations. Although I identified several major difficulties in distinguishing these transition types, I still see clear benefits to introducing these types to the psychopathology field. With these possible types in mind, we can go broader in ideas and find more plausible explanations for empirical observations. For example, the time series of momentary affects are more likely to have N-diffusion instead of B-tipping. This also echoes the discussion in Chapter 6: if researchers had known the types of transitions prior to designing their studies, would they have still chosen to calculate EWSs from momentary affect?

Figure 1. The frequency of articles including the phrase “early warning signal(s)” in (a) literature in the psychology domain, and (b) in all domains. Data was acquired on April 24th, 2025, through the Web of Science database (Clarivate, 2025).



Approach Nonlinear Dynamic Interactions

The next two chapters of this thesis focused on nonlinear relationships among variables, which are notoriously difficult to detect from empirical data. Compared to the previous two parts, this part of the work addressed a more fine-grained level of dynamic systems.

In Chapter 8, I introduced a new method of quadratic vector autoregression (quadVAR). The method features a systematic variable selection procedure that applies to both the linear terms and the quadratic or interaction terms. The results of the regression can be easily interpreted using linearized networks. The method's general performance was modest, especially in the classification metrics of parameters, yet it also showed promise in discovering nonlinear relationships for exploratory purposes. Given true psychological systems are never linear, should empirical researchers consider using more nonlinear analysis methods? Of course, more flexible methods often come with lower statistical power and more severe overfitting issues. Those issues are especially prominent when the methods are applied to relatively short time series collected by self-reports, as is typically the case in psychological research. With these considerations, I would suggest that nonlinear autoregression methods should always be applied with caution and carefully validated, possibly using information criteria and cross-validation. Yet when linear methods are used, researchers should also routinely check the linearity assumption. This can be done, for example, by using the quadVAR method and looking into any nonlinear relationships found there.

The difficulty in estimating nonlinear relationships from data also motivates the field of formal dynamic models, which use a top-down approach to infer these relationships between variables from theoretical insights and expert knowledge. Early work in formal dynamic models relies on simulations, which intuitively describe how a system changes over time. However, simulation-based methods do not provide a direct explanation of why certain dynamic relationships produce certain global system behaviors. In Chapter 9, I introduced a more analytical approach to examine how the combination of nonlinear dynamic functions gives rise to the system's attractors. The main strength of this approach is the use of mathematical stability analysis, which is not dependent on the specifications of simulations, such as the parameter settings, initial state of the system, and noise strength. In a simulation, it is possible that the system only starts around one attractor and never goes to others, or that two attractors are too close to be visually distinguished. These issues can be solved with the stability analysis methods, as the equilibrium points of the system in the whole state space or parameter space can be found using dynamic equations. Other advantages of the approach in Chapter 9 are better computational efficiency and a clearer link between the function form and

the system's behavior. On the other hand, this method requires a more detailed analysis of the model, including splitting the levels of interest (a topic also discussed in Chapter 7) and adjusting the dynamic functions accordingly. For dynamic models with many different elements and complex forms, this can be challenging.

A Reflection on Complexity Science in Psychology

Till here, the current thesis has explored multiple approaches rooted in complex systems research and applied these to psychological questions. It is often stated that complex systems, although different in nature, show similar behaviors. However, this does not mean that we have a single methodology that can be applied to all complex systems. On the contrary, we must better understand the characteristics of these systems and treat them with more care and nuance.

I first illustrate this point with the large variety of systems that may fall into the term “complex systems” (Figure 2). A complex system (or a model of one) may represent dynamics arising from a single element (e.g., the logistic map), several elements (e.g., the Lorenz model), or billions of elements (e.g., a neural network in the brain). The elements in a system may be exactly the same (e.g., atoms in the Ising model in physics, the molecules in a crystal), somehow similar to each other (e.g., people in a social network, variables in a correlational network, but see Molenaar, 2004, for the ergodicity conditions), or very different from each other (e.g., the servers, cables, personal devices, and humans for the Internet). The form of interactions between elements may be linear (as assumed by VAR networks), slightly nonlinear (the ReLU function used in artificial neural networks), or highly nonlinear (e.g., the influence of psychotherapy on a patient’s mental state). The system may have a specific function designed by humans (e.g., a computer program, the supply chain system), evolved naturally to adapt to the environment (e.g., an animal in its environment), or be purely functionless (e.g., the logistic map or the Lorenz system). The system may be dominated by a rather small number of clearly defined elements and their interactions, or can only be understood as the overall dynamic features (Wallot & Kelty-Stephen, 2018). The key models or methods developed in this thesis also apply to different types of complex systems, summarized in Table 1.

If the system of interest is (assumed to be) simple enough, maybe it is legitimate to use one solution for all. For example, a large amount of empirical work in psychology seems to assume that the influences of different predictors on an outcome variable are linear and additive. Under this assumption, it makes sense to use a single method, namely linear

regression, for all kinds of research questions regardless of the exact content of the variables of interest. That is why we see in many psychology studies that statistical methods seem to be less dependent on the specific research question.

Figure 2. Examples of complex systems or their models. This figure includes images from Wikimedia Commons: logistic bifurcation map ([Macias, CC BY-SA 3.0](#)), Lorenz attractor ([Dschwen, CC BY-SA 3.0](#)), neural network diagram ([KelvinSong, CC BY-SA 3.0](#)), sociogram ([Moreno, Public Domain](#)), fMRI brain scan ([Beao, Public Domain](#)), and Internet structure ([BWenk, CC BY-SA 4.0](#)).

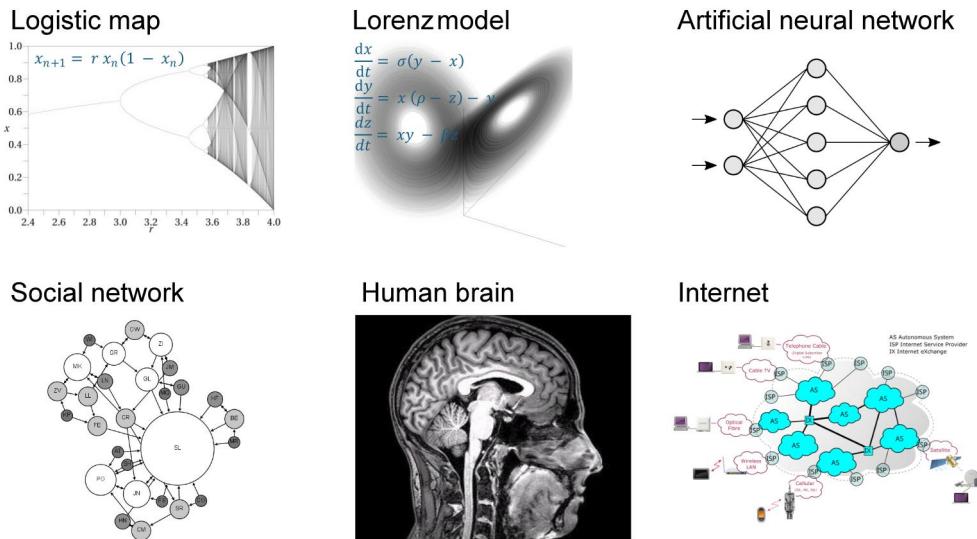


Table 1. Models and topics covered in this thesis, and the type of complex dynamic systems they apply to.

Model/topic	Related chapter(s)	Number of elements	Similarity of elements	Complexity in the form of interactions	Adaptivity	Component-or interaction-dominant
simlandr/formal dynamic models	2-3, 9	Small	Low	High	Low	Component
fitlandr	4	Large	Low	High	Medium	Interaction
Isinglandr	5	Medium	High	Medium	Medium	Both
EWS/transitions	6-7	Large	Low	High	High	Interaction
quadVAR	8	Medium	Medium	Low	Medium	Component

Note. For adaptivity, I label the models or topics that are directly related to the adaptive functions of individuals as “high”, those fitted from individuals’ data as “medium”, and those generated from theories with unknown outcomes as “low”.

For complex systems, some principles can be applied to systems with similar *structures* regardless of the specific *nature* of the elements. For example, the Ising model in physics describes the magnetism of matter, whereas the Ising model in psychology describes the symptoms of mental disorders. Nevertheless, they both show bistability as they have similar forms of interactions among variables. We can also consider the example of early warning signals, which are observed in ecology, physics, and finger movement, as long as a system under a small noise undergoes bifurcation-induced tipping (Scheffer et al., 2012).

However, this generalizability does not mean that *anything* found in many complex systems is automatically applicable to a new complex system, simply knowing that the new system is “complex”. This becomes evident when we examine how key concepts, such as complexity itself, are operationalized differently across disciplines, sometimes even in contradictory ways. For example, there are many measures of the complexity of a system. In the case of the logistic map, its Kolmogorov complexity (the length of the shortest computer code that can produce the output, Korotkikh, 2001) is rather small and does not change for different control parameter values, because the system’s output can always be generated with some simple computer code. Yet, the system’s dynamic complexity (a combined indicator of fluctuation and spreading of a time series) is quite high in the chaotic regime and increases as the control parameter grows. Another related example we found is that, although dynamic complexity sometimes behaves similarly to permutation entropy (Schiepek & Strunk, 2010), they may go in opposite directions before a sudden change (see <https://osf.io/hxqzf> for an example).

Therefore, it is important to realize that borrowing concepts, theories, and methods *within* the complexity science regime actually requires much caution. This is sometimes overlooked in previous literature, especially from social sciences, where theories are often based on verbal reasoning, and researchers are not used to checking the formal conditions of theories (Sussmann & Zahler, 1978; Wagenmakers et al., 2012; but also see Oliva & Capdevielle, 1980). In psychology, this issue has been raised by various researchers for some specific disciplines. For example, the use of centrality measures from physical networks may not be meaningful in correlational networks (Bringmann et al., 2019), and the interpretation of the first EWS results was also questioned at the early stage (Bos & De Jonge, 2014). The complexity of psychological systems suggests that we can only investigate a small aspect of the system in a single study, which requires methods that are specifically suitable for this aspect.

Here, it may be tempting to ask: To which category do psychological systems belong? Knowing this, perhaps we can rule out some methods that are inherently unsuitable for

psychology and focus on the most suitable category. While many researchers have argued that real-life psychological systems are idiographic, interaction-dominant, non-decomposable, and exhibit strong complexity (Hasselman, 2023; Olthof, Hasselman, Oude Maatman, et al., 2023; Oude Maatman & Eronen, 2025), I maintain that the level of these characteristics (Table 1) in a specific research question still requires further empirical examination. Note that the point I want to make is more at the epistemological level rather than the ontological level: Even if a model assumes a much simpler system than the actual one, it may not necessarily be a bad model, as a good simplification that removes unnecessary complexity can often help problem-solving, and a complex model may actually hide simple mistakes. To select a suitable research method for a given question, what is important is determining which model is the most *suitable* for a given empirical question (i.e., it captures the most important facet of the target system for the research question), not which model is the most *similar* to the target system. Of course, once we have gained sufficient knowledge about the target system, we can add more details to the model to make it better resemble real-life systems. However, this must be built on a solid foundation — we cannot skip over Newtonian physics and jump directly to the theory of relativity. Given that complexity research in psychology is still in its early stages and that various types of research questions exist in the field, I believe it is too early to say that any method is generally inappropriate. It is likely that many methods will find their place in different disciplines. More effort should be focused on determining which methods are most suitable for specific purposes.

For the future of complex systems research in psychology, I take a cautiously optimistic stance. On the one hand, it is hard to refute that psychological phenomena are inherently complex, thus, a reductionistic, over-simplified view misses many vital details. On the other hand, adapting concepts and even some methods from the complexity sciences is not likely to produce massive progress in the near future. This is not only due to the technical framework yet to be established, an issue that the current thesis aims to contribute to, but also deeper philosophical questions remain. For instance, what are the psychological constructs, and how can they best be measured? What can we assume about psychological dynamics, and are current methods based on valid assumptions? What can we hypothesize to be the basic principles governing psychological phenomena, and are those principles falsifiable? Compared to incremental research that heavily relies on existing work (some of which ultimately rests on unvalidated assumptions), those fundamental questions are harder to answer, but they need to be carefully treated with more creativity.

Moreover, we also need to investigate the similarities and differences between psychological systems and other complex systems. Although psychological systems are more difficult to measure and perform experiments on, psychologists also have the advantage of being able to talk to participants and make use of participants' memory and reasoning ability (Rozin, 2001; also mentioned in Chapter 7). Therefore, qualitative or mixed-methods studies should also be considered (e.g., Bossenbroek et al., 2025; Hulsmans et al., 2024). Those methods may provide unique opportunities for understanding psychological systems, and compensate for the disadvantages of data length and assessment accuracy.

Future Directions

I suggest the following directions for the further development of complex system methods in psychology. First, we need better integration and an overview of different methods, their applicable conditions or underlying assumptions, and the meaning of their outputs. Many methods have been developed in the field of complex systems, yet their differences are not easy to grasp. As a result, empirical researchers sometimes just start with the statement that psychological systems are complex and simply choose analysis methods that are readily available to them, which often depends on the computational and tutorial resources that they can access, without questioning what the most suitable method for a given research purpose is. In Table 2, I list some methods I am aware of and their key characteristics. From this point onward, a more comprehensive analysis is still needed to further clarify their advantages, disadvantages, and suitable use cases.

Table 2. Several types of complex system methods have been applied to psychological phenomena.

Type of methods	Examples
Wave-or-signal-based	Dynamic complexity (Schiepek & Strunk, 2010), change point analysis (Cabrieto et al., 2017; Lewis & Stevens, 1991; Matteson & James, 2014), statistical process control (Schat et al., 2023)
Recurrence-based	Recurrence quantification analysis (Marwan et al., 2007), recurrence network (Hasselman & Bosman, 2020)
Stability-based	fitlandr (Chapter 4), Ornstein-Uhlenbeck model (Oravecz et al., 2009), Fokker-Planck equation model (Tschacher & Haken, 2020).
Network-based	Network psychometrics (Borsboom et al., 2021), social network analysis (Borgatti et al., 2009)

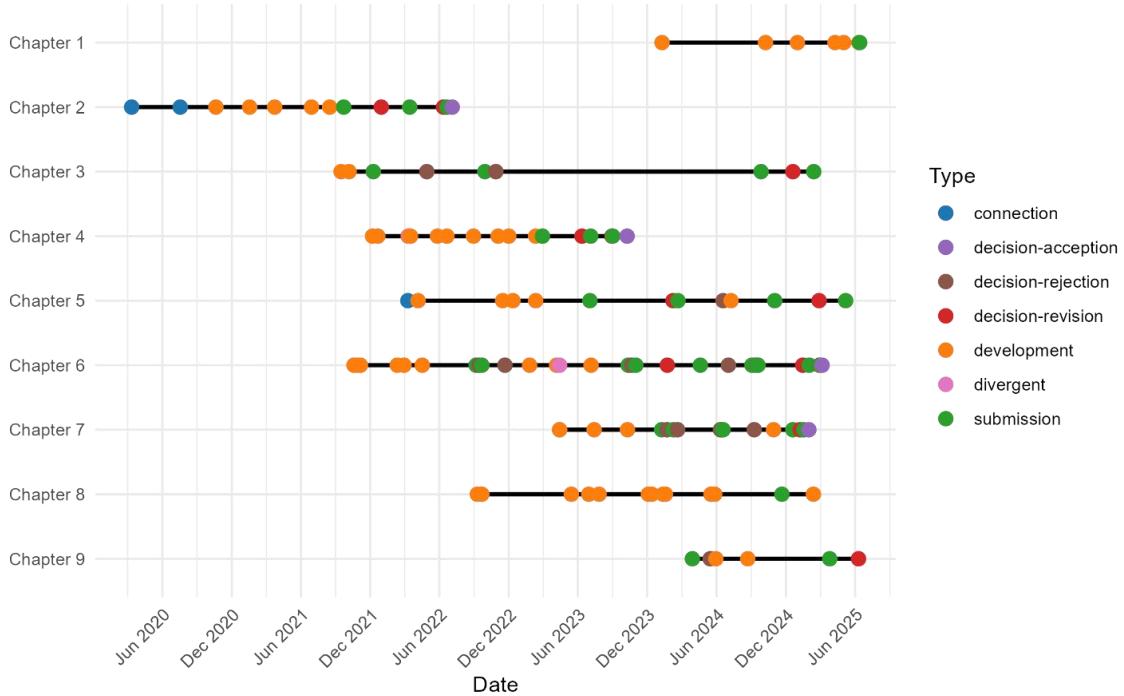
Second, methodological and theoretical developments may also enable new possibilities for empirical study design. Take the correlational network analysis as an example, an important question here is what variables should be included in the assessment. Complex systems may also inspire new assessments. For example, if we care more about the variability of a variable instead of its absolute value, we may ask comparative questions in experience sampling questionnaires (e.g., how much do you think you are *happier* than the previous assessment), which may be more accurate than calculating the variance of the absolute value. If we want to know how far a person is from a stable state, we can also explicitly ask the participant how much more stressed they are compared to a specific situation that the participant can imagine (e.g., a safe place with a relaxing environment). This requires some imagination ability from the participant but may be feasible as similar approaches, such as guided imagery and imagery rescripting, have long been used in psychotherapy (Arntz, 2012; Cumming & Anderson, 2013). To know which questions can best assess the multistability of a person's mental state, we may also ask the participant to come up with customized questions that can best distinguish this person's different (e.g., happy and depressive) modes. The most suitable questions may be highly heterogeneous among people, but can better describe multistability than standard questionnaires (Olthof, Hasselman, Aas, et al., 2023). We may also ask counterfactual questions, as mentioned in Chapter 7: for the events or event series that led to a transition, participants could be asked: Do you think that would also have led to a transition if it had happened at another time?

Third, we should continue adapting methods from other fields. The methodology across different scientific fields is sometimes more similar than it appears to be, although different terminologies may hinder mutual communication. Developments in psychological modeling may stem from previous psychological models but are also likely to come from other fields, such as biology (Chapters 2-5, 8-9), economics (Chapter 5), or ecology (Chapters 6-7). The development of such methods requires more interdisciplinary collaborations, as well as more interdisciplinary knowledge from psychologists to enable such crossing across disciplinary borders. Psychologists should also develop the ability to describe phenomena in an abstract way (e.g., in mathematical formulas or pseudo-codes), a skill that could be facilitated through more interdisciplinary courses, workshops, and educational materials. Another opportunity comes from recent developments in artificial intelligence, which may provide tools to search for relevant information across disciplines or translate useful terminologies from different fields (Jones, 2023).

Epilogue: How Did All These Things Work Out?

By the time I started my PhD project, I had no clue how things would develop throughout the coming four years. A lot of exploration, trials, failures, and progress together shaped it into its current form. Looking back on the past four years, I feel I enjoyed the PhD project a lot, and I also realized my PhD path was quite unique among the colleagues I know. Therefore, I would like to share some experience I gained during this project, which may be interesting for those who are also considering a curiosity-driven PhD project, even without a very mature plan at the beginning. I summarized the timeline of this PhD project in Figure 3, and an interactive version of it can be found at https://jingmeng-cui.netlify.app/assets/files/PhD_timeline.html.

Figure 3. The timeline of this PhD project.



Several characteristics of this project are interesting to discuss. First, learning by trying is a central theme of this project. Compared with empirical work, the methodological and theoretical work presented in this thesis was less straightforward to plan. Sometimes it was clear that there were methods available in other fields, but unclear whether they would work for our situation. Sometimes we did not know if the technique would exist at all. Although the general directions were set at the beginning of the project, concrete plans only emerged along the way; some directions were discovered after explorations, and some were based on the work that was established during my PhD. Many of the methods we use are rather new. The core

idea of the simlandr (Chapters 2-3) is based on the idea by Wang et al. (2008), which was further mathematically validated by Zhou and Li (2016); the Isinglandr project was based on the work by Lunansky et al. (2024) that I only heard halfway through the project; the core idea of the fitlandr (Chapter 5) was based on Qiu et al., (2022), with the MVKE algorithm by Bandi and Moloche (2018); the quadVAR method (Chapter 8) was based on the RAMP algorithm, Hao et al. (2018); the idea of introducing the graphical analysis tools to psychology emerged during the master's thesis project of Gruppen (2023). The chapters about EWSs (Chapters 6-7) also deviated from the initial plan, which was to refine the EWS detection methods to yield better prediction power. More substantive issues were found during the exploration phase, so instead of improving the accuracy, we focused on examining the consistency between theory and method.

Second, information searching played an important role. We not only needed to solve the problem but also needed to define the problem clearly, sometimes using terms from other fields, and find out to what extent the problem could be solved. Also, many efforts were made in implementation, which required coding in R as well as some knowledge of other programming languages, such as C, C++, MATLAB, and Python. When there are multiple methods available, it is also important to find comparison articles and choose the situations that are close to psychological applications. It is even better to find algorithms with available implementations, but I noticed that it is important to be aware that existing algorithms may contain errors, especially when the package is intended for a very specific field without being tested by many users (e.g., the patched version of RAMP, <https://github.com/Sciurus365/RAMP>, and Cui et al., 2023). Therefore, when building on others' work, it is also important to check the output carefully and understand the code before using it.

Third, the project also features a relatively large proportion of examination or correction compared to new developments. After all these years, the application of complex systems in psychology is still experimental¹. Sometimes, new methods attract more attention than work pointing out issues in previous methods. Does that mean correction has less value than creation? I would not agree. The value of efforts in reproducing empirical research has been increasingly recognized in the past decade (Open Science Collaboration, 2015), and many researchers have realized that robust findings are better than fancy findings. A similar idea also needs to be recognized on the methodological side (Borsboom, 2006). Of course, methodologists do not always agree with each other. In this case, more open discussion is certainly helpful in clarifying the viewpoints and outlining the pros and cons of each approach.

¹ Adapted from the statement about the “casnet” R package by Fred Hasselman.

Last, and importantly, this project is inherently interdisciplinary, and my interdisciplinary background certainly played a key role in its development. As mentioned in the prologue, I began my academic journey in chemistry. Although the techniques I learned there were not directly involved in any of the projects, the mathematics, natural sciences, and coding bases I gained there certainly have paved an easier route for me to understand more advanced and specialized techniques. The intuition I gained in natural sciences also motivated several key ideas applied in this project. At the same time, I also completed my bachelor's and master's degrees in psychology, which allowed me to develop both a solid foundation and an understanding of the state-of-the-art in the psychological field. In addition, I was fortunate to work with a research team of collaborators who shared similar interests and brought diverse, specialized expertise. Those aspects all contributed to the completion of this thesis. As shown in a recent study by Xiang et al. (2025), the success of interdisciplinary research does not rely on topic-interdisciplinary (inferred from the abstract of articles), but on knowledge-based interdisciplinarity (inferred from the references of articles). Indeed, the studies included in this thesis would not have been possible without a working knowledge of both psychology and other disciplines. Only by combining solid knowledge from different fields can interdisciplinary research reach its full potential and yield meaningful insights.

Life is certainly a very high-dimensional system with many ups and downs. This is especially prominent if we talk about doing a PhD, a rather lonely and uncertain journey. You own the whole project yourself, although supported by supervisors and colleagues, and you never know whether something you try will work out, and how other researchers will evaluate your work. Everyone has experience with good and bad periods, but for me, it is also interesting to see how my research project gave me more insight into the process of staying in or moving between good and bad periods. When you are in a good or a bad period, you feel you are stable or stuck there, and much of what you do "automatically" may just strengthen the phase that you are in, as a positive feedback loop is required for multistability. For instance, you had a good night of sleep, got up early in the morning, felt fresh, had nice ideas at the office and smoothly got the work done, left the office early and happily, hit the gym and felt energetic, tried a restaurant for the first time and had a very good dinner there, went back home and played some video games together with friends before going to bed early and having a good sleep. Alternatively, you had a bad sleep, got up late, rushed to the office feeling guilty, looked at the work today but didn't want to start anything, just spent some time on social media and got more nervous, headed home only when the building is closing, went to the supermarket but didn't feel like eating anything there, ended up buying something fast and not tasty, after

dinner just wanted to lie on the couch and watch short videos without realizing it's already very late, had a difficult time before being able to sleep.

Both these types of days can happen and did happen to me during my PhD journey. What they illustrate is that each step in the process increases the possibility of a certain next step, forming distinct attractors. Being different from the bi-dimensional case, the thing about high-dimensional attractors is that if you somehow notice it and want to change the trajectory, changing only one thing in the chain of events will still keep you super close to the original attractor. Thus, changing one single step will not push you from the negative loop to the positive one. Let's use math again; if there is only one dimension, changing one variable is enough for you to go from (0) to (1); if there are two dimensions, changing one variable can make you be as close to both attractors (0, 0) and (1, 1); but if there are 10 dimensions, changing one variable only brings you from (0, 0, 0, 0, 0, 0, 0, 0, 0, 0) to (1, 0, 0, 0, 0, 0, 0, 0, 0, 0), which is still far away from (1, 1, 1, 1, 1, 1, 1, 1, 1, 1).) The strength of attraction, or the depth of the basins, then translates to whether the previous step is realized in the chain, and how likely the next step is still in the chain. If you are already in the bad loop and want to go to the good one, simply changing one thing will not be enough. You either need to consciously maintain the good loop for a while, even if you don't want to do the things there, by using your determination or going to another environment where maybe the weather is better, friends are around to support you, the food is more delicious, or work pressure is less. This way you make the good loop more attractive, and easier to start and maintain. Although this idea is not a part of my PhD research, it may be a good way to end this thesis, as it nicely shows how research and researchers grow each other in a complex, dynamic way.

Appendix A. Supplementary Materials for Chapter 2

A1. Stability Analysis

Figure A1. dPT/dt vs PT plots for different AS values. The dash lines represent the zero value of the y-axis.

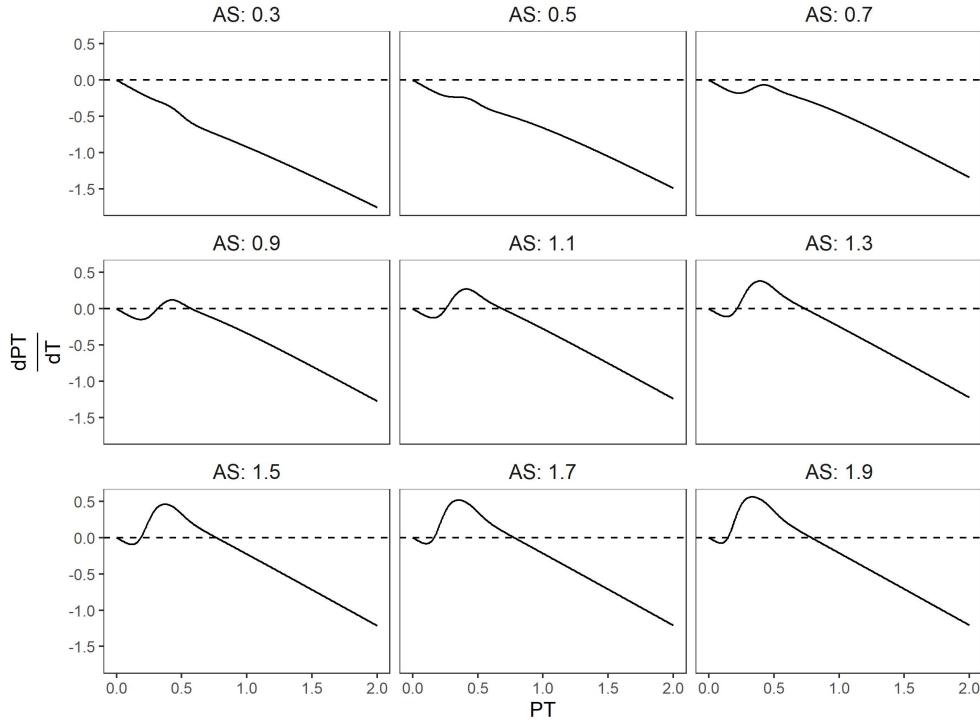
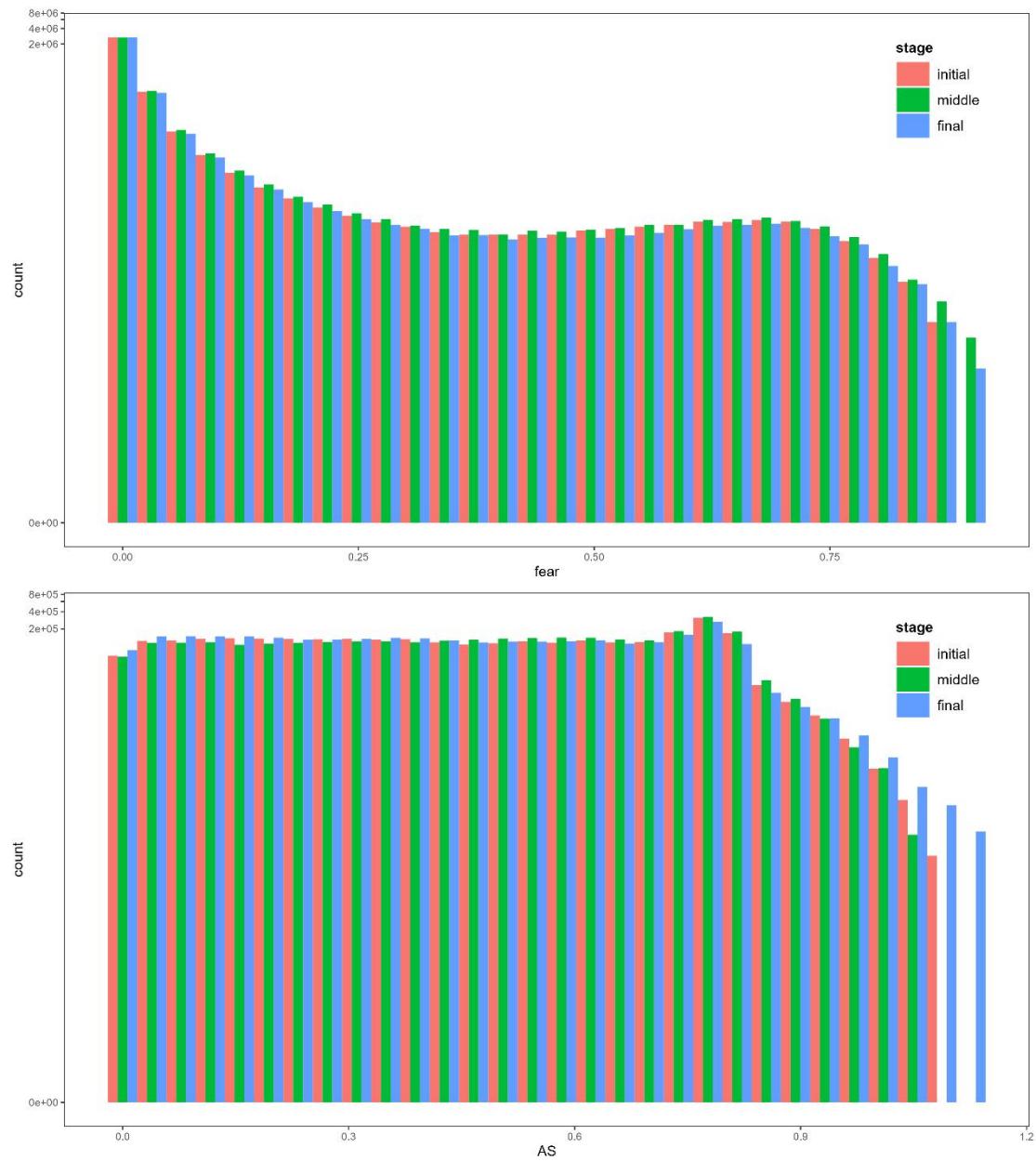


Table A1. The real part of the dominant eigenvalues at three equilibrium points of the system. $Re(\lambda_d)$: the real part of the dominant eigenvalue; EP: equilibrium point. The balance points are numbered in the order of increasing PT .

AS	$Re(\lambda_d)$		
	EP1	EP2	EP3
0.3	-1		
0.5	-1		
0.7	-1		
0.9	-1	0.824	-1
1.1	-1	0.925	-1
1.3	-1	0.979	-1
1.5	-1	1.016	-1
1.7	-1	1.044	-1
1.9	-1	1.066	-1

A2. Checking Convergence

Figure A2. The density distribution of fear and AS in the initial, middle, and final stages of the simulation using the simplified model. The length of the simulation is 107 timesteps. For each stage, 30% of the whole sample was selected. The pseudo-logarithm transformation from the scales package (Wickham & Seidel, 2020) is used for the y-axis to present density values that vary much in magnitude.



A3. Practical Information on Programming Issues

In the main text, we mentioned that because the landscape construction algorithm requires long simulations, the simulation function itself should be fast enough. Improving the performance of codes is a complex issue and the optimal solution differs from case to case (see Wickham, 2019, for an overview). Besides the tutorial paper for the simlandr package (Cui et al., 2021) and our replicable materials (<https://osf.io/ke3xb/>), here we also list several possible directions based on our experience. This section may be helpful if you plan to apply our methods to your own model.

First, profiling the function is a useful way to find out which parts of the function take the most time and memory. The profiling output is often a graph or table which shows how much time or memory the program used to execute each row of codes. It clearly shows the bottleneck of performance. Optimizing these bottleneck steps, if possible, can yield the best improvement. For pure R codes, profiling can be done with the profvis package (Chang et al., 2020) or the profile package (Müller, 2021b). For C++ or Rcpp codes, you may use the pprof tool (<https://github.com/google/pprof>) or the jointprof package (Müller, 2021a). In our case, profiling showed that the most time-consuming step of the original function was not the simulation computing per se, but the maximum value of a constant vector was recalculated in each timestep. Removing that resulted in a dramatic decrease in running time.

Second, for those with more programming experience, you may try to rewrite your R codes in Rcpp (Eddelbuettel & François, 2011), which typically increases the running speed. It is often not possible to divide dynamic simulation functions into several parts, but whenever that is possible, you may try to use parallel computing which computes those parts simultaneously instead of doing one after another. This can be done with the parallel package, which is a part of R (R Core Team, 2021), or the foreach package (Microsoft & Weston, 2020).

Finally, increasing the noise of the system may help to increase the effective sample size, or in other words, make it easier to achieve ergodicity. This is similar to the issue of the variance of proposal distribution in Metropolis sampling (Gelman et al., 1997): when the system is easier to travel to distant regions, a single data point contains more information about the system. However, we should bear in mind that as the noise increases, the property of the original system may be altered to a large extent (as discussed in the *Limitations and possible pitfalls* section of the main text). In our experience, if the noise strength is too high, the model may crash because one or more variables of the system become infinity (thus it is important to check the range of the variables before putting the simulation output into a landscape construction function). This issue is especially common when the step size of the simulation is

large (i.e., the continuous process is approximated by a discrete function that is too rough), and when the dynamic functions are only properly defined in a limited domain. Some trials are often needed to define the optimal magnitude of noise.

A4. Equivalent Chemical Representations

For many psychological formal models, the forms of the functions are similar to those for chemical dynamics. Comparing the two can help to understand the psychological models from another perspective, and potentially enable to borrow the methods from that richer field. Here we use this model of panic disorder as an example to illustrate how to find the equivalent chemical representations. Note that this section mainly serves as a bridge to the chemical dynamics. It does not contain additional technical information by itself, and some background in biochemical modeling may be required for the understanding of the materials.

The dynamic models take a limited set of functions. In the panic disorder model, the basic components are the following (where A and B represent variables, and r , h , and p represent parameters):

$$1. \frac{dA}{dt} = -rA$$

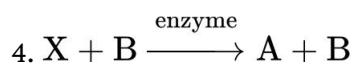
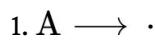
$$2. \frac{dA}{dt} = rB$$

$$3. \frac{dA}{dt} = -rB$$

$$4. \frac{dA}{dt} = r \frac{B^p}{B^p + h^p}$$

$$5. \frac{dA}{dt} = -r \frac{B^p}{B^p + h^p}$$

Correspondingly, there are several kinds of chemical functions for which the changing rate of the species follows the equations above (where X represents an unknown species with constant concentration).



For the readers unfamiliar with chemical formulas, this paragraph provides a short introduction to them. The first three reactions are elementary reactions. According to the law of mass action, the rate of reaction is proportional to the product of the concentrations of the

reactants. Formula 1 is a self-decomposition reaction, in which A decomposes automatically. Formula 2 and 3 are catalyzed reactions. In Formula 2, A is produced from an unspecified species X with a constant concentration, with B as the catalyst. In formula 3, an unspecified species Y is produced in a similar process. Then, Y quickly reacts with A . Because the second step is much faster, the rate that the concentration of A decreases equals the rate that Y is produced, which is proportional to the concentration of B . Formula 4 and 5 are Hill processes. These formulas do not obey the law of mass action because the enzyme has a special property: as more reactants bind with the enzyme, the reaction speed becomes faster. (A well-known large molecule with this behavior is hemoglobin, which binds oxygen molecules faster when there is already one bond molecule.) See Keener and Sneyd (2009) for a comprehensive introduction to this topic.

Using these correspondences, we can rewrite the dynamical equations of the original model in the form of chemical reactions (Table A2).

Table A2. Equivalent chemical reaction representations of the dynamical functions in the model of panic disorder. Xs are unspecified species with a constant concentration; v_{\max} , K_m , and n are the rate parameters for Hill reaction.

Dynamical equations	Chemical reactions
$\frac{dA}{dt} = r_A(s_{PT,A}PT - A - s_{H,A}H)$	$X^1 + PT \xrightarrow{r_A \cdot s_{PT,A}} A + PT$ $A \xrightarrow{r_A} .$ $X^2 + H \xrightarrow{r_A \cdot s_{H,A}} Y^2 + H, Y^2 + A \xrightarrow{\text{fast}} .$
$\frac{dH}{dt} = r_H \left(\frac{A^{p_{A,H}}}{A^{p_{A,H}} + h_{A,H}^{p_{A,H}}} - H \right)$	$A + X^3 \xrightarrow[\substack{v_{\max}=r_H, K_m=h_{A,H}, n=p_{A,H}}]{\text{enzyme}} A + H$ $H \xrightarrow{r_H} .$
$\frac{dPT}{dt} = r_{PT} \left(\frac{A^{p_{A,PT}}}{A^{p_{A,PT}} + h_{A,PT}^{p_{A,PT}}} - PT - S_{E,PT}E \right)$	$A + X^4 \xrightarrow[\substack{v_{\max}=r_{PT}, K_m=h_{A,PT}, n=p_{A,PT}}]{\text{enzyme}} A + PT$ $PT \xrightarrow{r_{PT}} .$ $X^5 + E \xrightarrow{r_{PT} \cdot s_{E,PT}} Y^5 + E, Y^5 + PT \xrightarrow{\text{fast}} .$
$\frac{dE}{dt} = r_E \left(\frac{PT^{p_{PT,E}}}{PT^{p_{PT,E}} + h_{PT,E}^{p_{PT,E}}} - E \right)$	$A + X^4 \xrightarrow[\substack{v_{\max}=r_{PT}, K_m=h_{A,PT}, n=p_{A,PT}}]{\text{enzyme}} A + PT$ $PT \xrightarrow{r_{PT}} .$ $X^5 + E \xrightarrow{r_{PT} \cdot s_{E,PT}} Y^5 + E, Y^5 + PT \xrightarrow{\text{fast}} .$

Appendix B. Supplementary Materials for Chapter 4

B1. Equivalent Vector Fields and Potential Functions for VAR and OUM

In the main text, we present fitlandr, a method to estimate vector fields and potential landscapes for psychological systems. Many previous methods, such as the VAR model and the OUM can also be represented in vector fields and potential landscapes, but in their specific, restricted forms.

In a two-dimensional VAR(1) model, the state of the system is a linear function of the previous state plus a noise term:

$$\mathbf{x}(t+1) = \mathbf{c} + \mathbf{A}\mathbf{x}(t) + \mathbf{e}(t), \quad (1)$$

in which \mathbf{c} is a constant term and \mathbf{A} is the parameter matrix. We can rewrite this equation in terms of the differences between time points:

$$\Delta\mathbf{x} = \mathbf{x}(t+1) - \mathbf{x}(t) = (\mathbf{A} - \mathbf{I})\mathbf{x}(t) + \mathbf{c} + \mathbf{e}(t). \quad (2)$$

Only the first two terms are deterministic. Therefore, $\mathbf{f}(\mathbf{x}) = (\mathbf{A} - \mathbf{I})\mathbf{x}(t) + \mathbf{c}$ is the vector field for the VAR(1) model. As in the difference equation, the vector field function is also restricted to be linear. We draw the vector field of an arbitrary example, where $\mathbf{A} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$, in Figure A1a. Note that because the equation is linear, in usual cases there is only one point where $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, which means there is only one stable point.¹ Therefore, it is not possible to show bistability with a VAR(1) model.

To construct a potential landscape for a VAR(1) model, we estimate the steady-state distribution from simulation, with the noise term drawn from an arbitrary bivariate normal distribution with $\Sigma = \begin{bmatrix} 0.1 & 0.05 \\ 0.05 & 0.1 \end{bmatrix}$. The resulted landscape is shown in Figure A1b.

The OUM assumes the system always tends to go back to a balance point (Kuppens et al., 2010; Oravecz et al., 2011):

$$d\mathbf{x} = -\boldsymbol{\beta}^T(\mathbf{x} - \boldsymbol{\mu})dt + \sigma dB, \quad (3)$$

where $\boldsymbol{\mu}$ is the local minimum (referred to as the home base by Kuppens et al., 2010) of the system, $\boldsymbol{\beta}$ is a vector specifying the stability of the phase (referred to as the strength that the system is attracted to the home base by Kuppens et al., 2010), and σdB is a noise term. The vector field representation of OUM is $\mathbf{f}(\mathbf{x}) = -\boldsymbol{\beta}^T(\mathbf{x} - \boldsymbol{\mu})$, which is again, a linear function.

We draw an arbitrary example where $\boldsymbol{\beta} = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}$, $\boldsymbol{\mu} = \begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix}$ in Figure B1c.

¹ In some special cases there may be infinite number of solutions to $\mathbf{f}(\mathbf{x}) = \mathbf{0}$. For example, if $\mathbf{A} = \mathbf{I}$ and $\mathbf{c} = \mathbf{0}$, then every \mathbf{x} satisfies $\mathbf{f}(\mathbf{x}) = \mathbf{0}$. However, this situation is not likely to happen in a VAR(1) model estimated from empirical data.

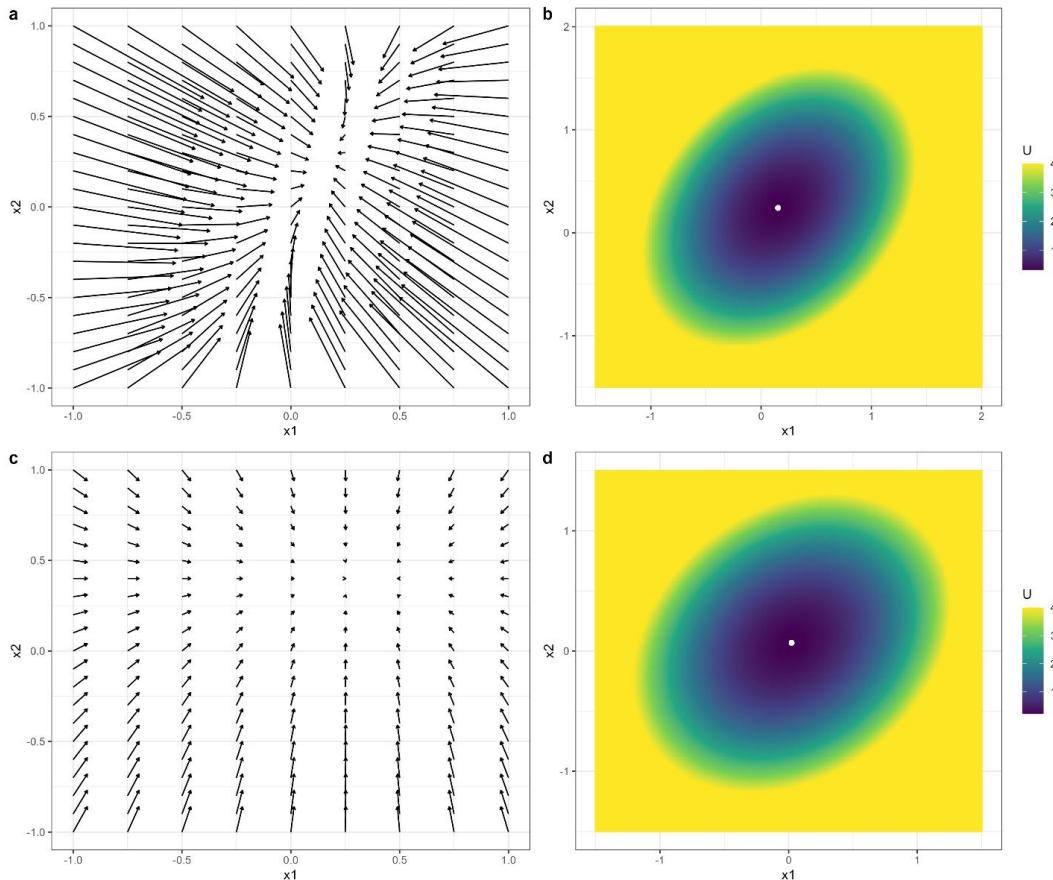
The Jacobian of $\mathbf{f}(\mathbf{x})$ is a diagonal matrix $\text{diag}(\boldsymbol{\beta}) = \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix}$, which is symmetric.

Therefore, it is a gradient model, and its potential function is in a quadratic form:

$$U(\mathbf{x}) = \int \mathbf{f}(\mathbf{x}) d\mathbf{x} = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \text{diag}(\boldsymbol{\beta})(\mathbf{x} - \boldsymbol{\mu}) + U_0. \quad (4)$$

We draw the potential landscape with the noise term drawn from an arbitrary bivariate normal distribution with $\Sigma = \begin{bmatrix} 0.1 & 0.05 \\ 0.05 & 0.1 \end{bmatrix}$. The landscape plot is shown in Figure A1d.

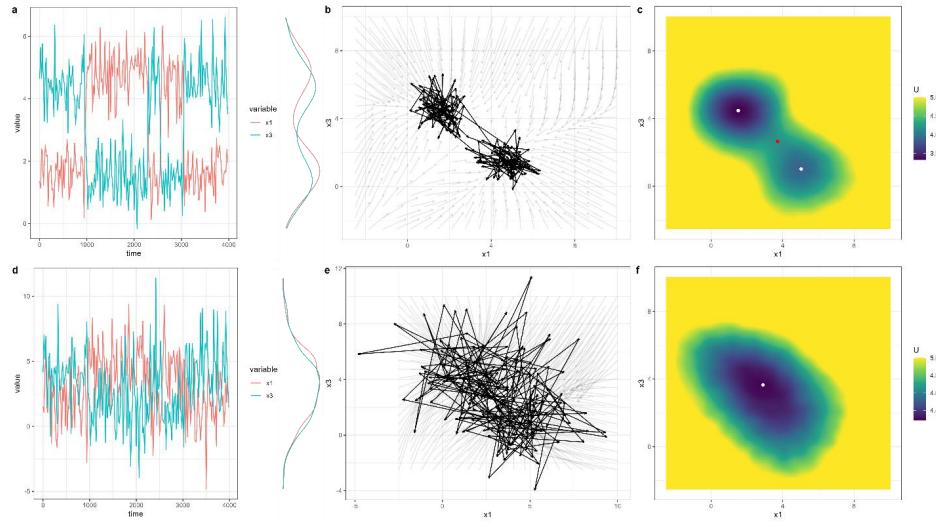
Figure B1. The vector fields and potential landscapes for a two-dimensional VAR(1) model and an OUM.



B2. Performance of fitlandr under Different Levels of Noise

In the main text, we demonstrated that fitlandr is robust under moderate noise ($SD = 1$). When the level of noise is lower ($SD = 0.5$), the results tend to be closer to the baseline condition (Figure B2a-c), but when the noise is too high ($SD = 2$), the bistability cannot be accurately recovered (Figure B2d-f).

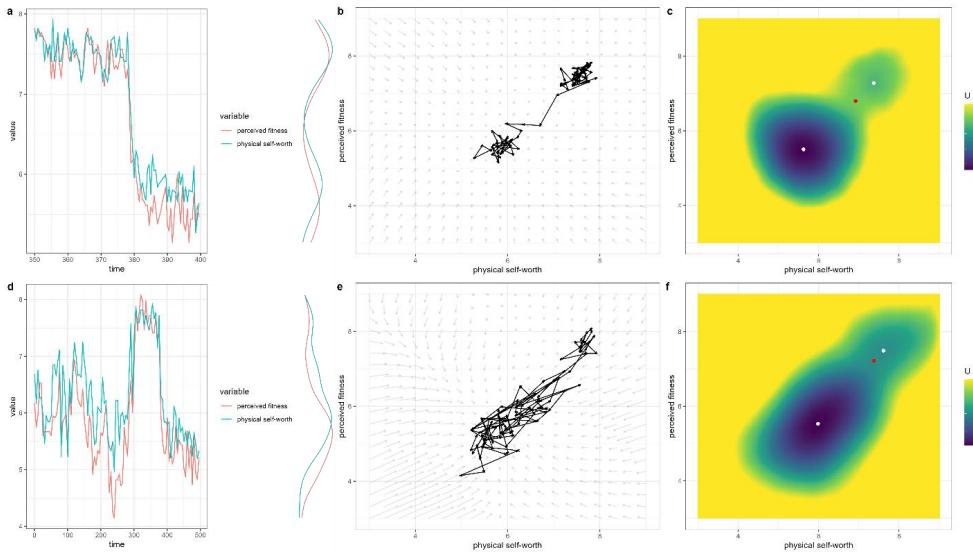
Figure B2. The vector fields and potential landscapes for the model used in the simulation study, with lower and higher levels of noise.



B3. Results from the Shortened Data of P2

The length of the data from P2 is much longer than that of P1. To further demonstrate the robustness of our method, we also present results from shortened data from P2. The data was shortened in two ways: (1) by using only the 700th to 799th data points, and (2) by using every 10th data point from the first data point to the 991st data point. The results are shown in Figure B3a-c and Figure B3d-f, respectively. Despite the reduction in data length, the vector field and the potential landscape generated by fitlandr showed similar features as those based on the original time series.

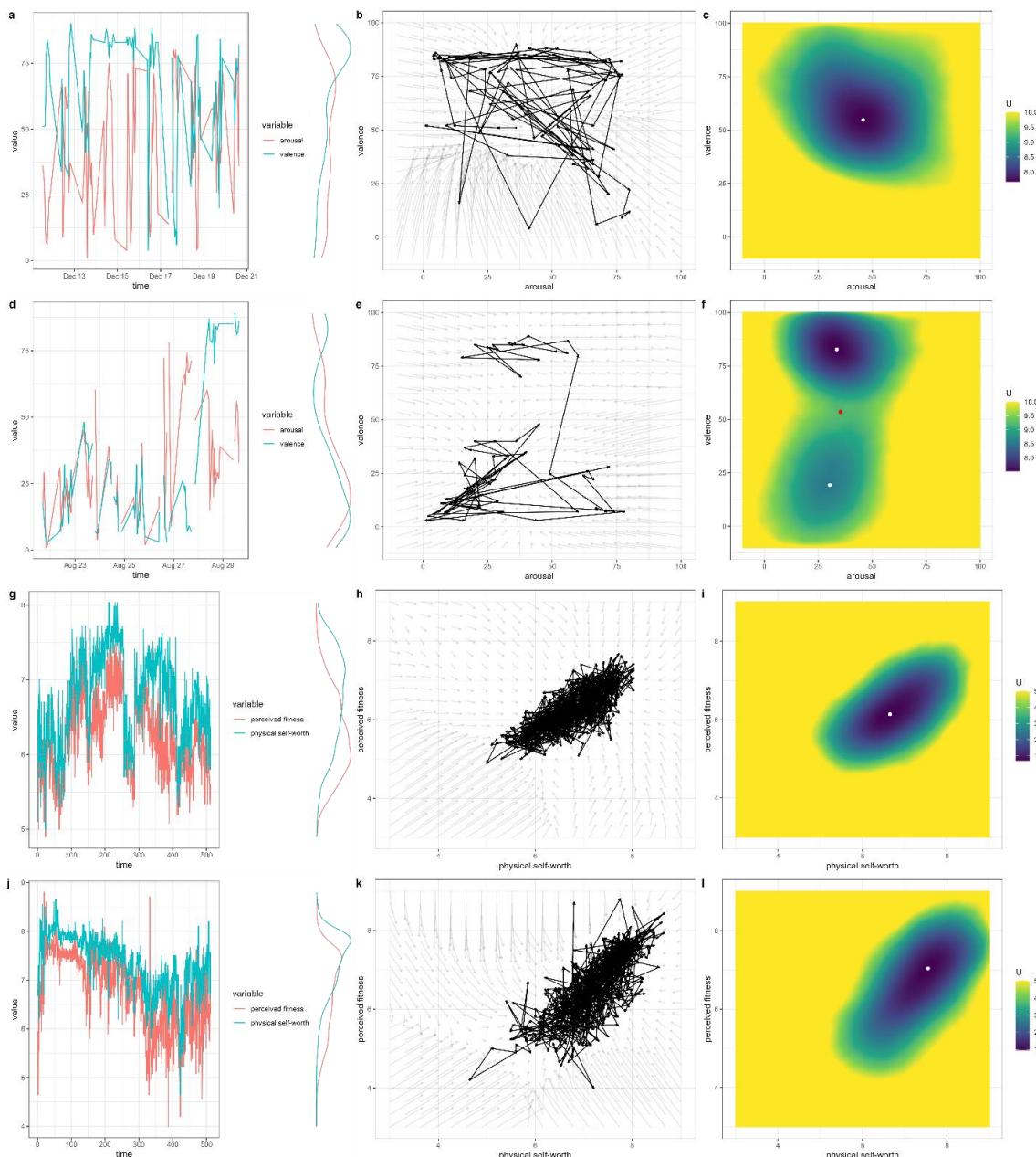
Figure B3. The vector fields and potential landscapes estimated from shortened data from P2.



B4. Vector Fields and Potential Landscapes for Other Participants

In the main text, we showed the results from two participants from two studies as examples. We also applied the same procedure for four additional participants from the same studies. We selected two participants from the Leuven BPD study (Houben et al., 2016), one without a BPD diagnosis (P3) and one with a BPD diagnosis (P4), both with relatively long records. We also selected two participants from the study by Delignières et al. (2004), referred as P5 and P6. The resulted vector fields and potential landscapes are shown in Figure B4. Comparing with the results of P1 and P2, it is clear that each participant has unique dynamics and stability features.

Figure B4. The vector fields and potential landscapes estimated from the data from four additional participants. Each row presents the results for a different participant, P3-P6.



Appendix C. Supplementary Materials for Chapter 5

C1. Description of the Nine Symptom Measures (DSM IV)

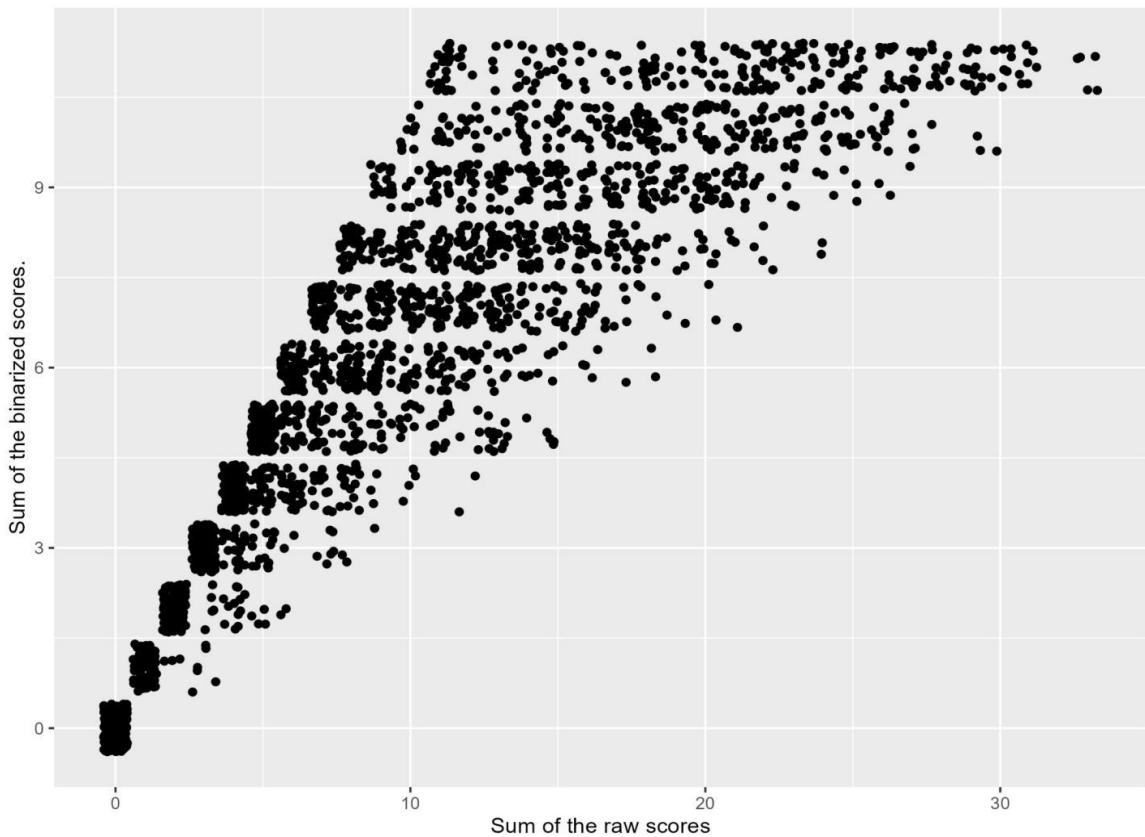
Table C1. MDD Symptoms

DSM Symptom criteria	
(1)	Depressed most of the day, nearly every day
(2)	Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day
(3)	Significant weight loss when not dieting or weight gain (e.g., change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day
(4)	Insomnia or hypersomnia nearly every day
(5)	Psychomotor agitation or retardation nearly every day
(6)	Fatigue or loss of energy nearly every day
(7)	Feelings of worthlessness or excessive or inappropriate guilt nearly every day
(8)	Diminished ability to think or concentrate, or indecisiveness, nearly every day
(9)	Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide

C2. Determining the Cutoff Value for the Binarized CPAS-11 Scale

In order to estimate an Ising network, a crucial step involves binarizing all the responses, whereby any non-zero response is considered "active." For the original CPAS-11 questionnaire, the predefined cutoff value is set at 15. To determine the ideal cutoff value for the binarized data, one that closely aligns with the original data, we employ the algorithm provided by the OptimalCutpoints package (López-Ratón et al., 2014) with Youden's index (Youden, 1950), which maximizes the sum of sensitivity and specificity. The result suggests that the optimal cutoff value for the binarized scale should be 8. See Figure C1 for the distribution relationship between the raw scores and the binarized scores.

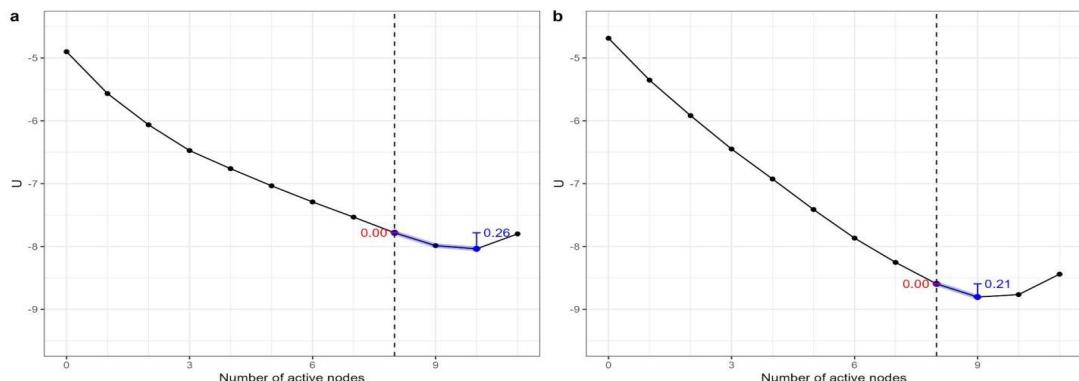
Figure C1. The distribution relationship between the raw scores and the binarized scores.



C3. Landscapes for Networks with the Same Connectivity Matrix

In Figure C2, we show the stability landscapes using the two networks with different connectivity parameters but using the threshold parameters from the low resilience group in both models. In this case, the landscapes were much more similar.

Figure C2. The landscapes for (a) the low resilience group, and (b) the network with the connectivity parameters from the high resilience group and the threshold parameters from the low resilience group.



Appendix D. Supplementary Materials for Chapter 6

D1. Formal Derivation of the EWS Theory for Multivariate Systems

In this part, we aim to provide a formal derivation of the existence of early warning signals for general multivariate systems and systematically examine the assumptions involved in the derivation. The proof for one-dimensional systems and multidimensional systems with symmetric Jacobian has been given in previous work (Dablander et al., 2023; Scheffer et al., 2009).

Assume the dynamic system is governed by the following ordinary differential equation:

$$\frac{dx}{dt} = \mathbf{F}(x), \quad (1)$$

and assume there is an equilibrium point x_0 that $\mathbf{F}(x_0) = \mathbf{0}$. To simplify notations, we assume $x_0 = \mathbf{0}$. When the system is close enough to the equilibrium point, with local linearization, the dynamics can be represented by a linear Jacobian matrix:

$$\frac{dx}{dt} = Jx, \quad (2)$$

where $J = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \dots & \frac{\partial F_n}{\partial x_n} \end{bmatrix}$, n is the dimension of the system, F_i and x_i represents the i -th component of \mathbf{F} and \mathbf{x} . Here, the local linearization means that we assume the force that pulls the system back to (or pushes the system away from) the local minimum is proportional to the distance it is from the local minimum. Holding the same direction, the further the system is from the local minimum, the stronger the force to pull it back (or push it away).

The forces that pull the system back (or push it away) depend on the direction in which the system is away from the local minimum. There are specific directions that characterize the effect of the force, represented by the *eigenvectors* of J . Each eigenvector also has an eigenvalue, representing the strength of the force. The eigenvalue may be complex numbers, but whether the force tends to pull the system back or push the system away only depends on the real part of the eigenvalue. Because the equilibrium point is stable, the system should be pulled back under a small perturbation, no matter in which direction. This means all the eigenvalues of J should only have negative real parts. The eigenvalue with the largest real part (the eigenvalue with the smallest absolute value of the real part) is called the dominant eigenvalue (λ_d). So the equilibrium condition is equivalent to $\text{Re}(\lambda_d) < 0$. We first show the derivations with the

assumption that all eigenvalues and eigenvectors of \mathbf{J} are real values. The situation with complex eigenvalues and eigenvectors has similar conclusions and will be discussed later. When the system is close to the catastrophic bifurcation point λ_d approaches zero, making the system gradually unstable in this direction. This instability manifests when we add noise in the system, as represented by the following multivariate Ornstein-Uhlenbeck process,

$$d\mathbf{X}(t) = \mathbf{J}\mathbf{X}(t)dt + \boldsymbol{\sigma}d\mathbf{W}(t). \quad (3)$$

To simplify the derivation, we first do a transformation of the coordinates with eigenvalue decomposition.

$$d\mathbf{X}'(t) = \boldsymbol{\Lambda}\mathbf{X}'(t)dt + \boldsymbol{\sigma}'d\mathbf{W}(t), \quad (4)$$

where $\mathbf{X}'(t) = \mathbf{Q}^{-1}\mathbf{X}(t)$, $\boldsymbol{\Lambda} = \mathbf{Q}^{-1}\mathbf{J}\mathbf{Q}$, $\boldsymbol{\sigma}' = \mathbf{Q}^{-1}\boldsymbol{\sigma}(t)$. Here \mathbf{Q} is a square matrix whose columns are the eigenvectors of \mathbf{J} , and $\boldsymbol{\Lambda}$ is a diagonal matrix whose diagonal elements are the eigenvalues of \mathbf{J} . The order of the eigenvalues in $\boldsymbol{\Lambda}$ is arbitrary. Here, we assume that the first eigenvalue is the dominant eigenvalue, and it is the only eigenvalue that will approach zero from the negative side ($\lambda_1 \rightarrow 0^-$). Now we investigate the variance and covariance of the process. It can be proved that Ornstein-Uhlenbeck processes always have a multivariate normal distribution, and the variance of the Ornstein-Uhlenbeck process above for a sufficiently long time can be calculated by the following formulas (Meucci, 2009; Vatiwutipong & Phewchean, 2019),

$$\text{vec}(\boldsymbol{\Sigma}'_\infty) = [-(\boldsymbol{\Lambda} \oplus \boldsymbol{\Lambda})]^{-1} \text{vec}(\boldsymbol{\Sigma}'), \quad (5)$$

where $\boldsymbol{\Sigma}' = \boldsymbol{\sigma}'\boldsymbol{\sigma}'^T = \mathbf{Q}^{-1}\boldsymbol{\Sigma}(\mathbf{Q}^{-1})^T$, vec is the vectorization parameter that transforms a matrix to a column vector by stacking the columns of the matrix together and \oplus is the Kronecker sum defined by

$$\begin{aligned} & \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \oplus \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} + b_{11} & \cdots & b_{1n} & \cdots & \cdots & a_{1n} & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ b_{n1} & \cdots & a_{11} + b_{nn} & \ddots & \ddots & 0 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & 0 & \ddots & \ddots & a_{nn} + b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & a_{n1} & \cdots & \cdots & b_{n1} & \cdots & a_{nn} + b_{nn} \end{bmatrix}. \end{aligned} \quad (6)$$

The equations above seem overwhelming. Nevertheless, as $\boldsymbol{\Lambda}$ is a diagonal matrix, $\boldsymbol{\Lambda} \oplus \boldsymbol{\Lambda}$ is also a diagonal matrix, and the first element $\boldsymbol{\Lambda} \oplus \boldsymbol{\Lambda}$ is the only one approaching zero. This

means the term $[-(\Lambda \oplus \Lambda)]^{-1}$ is a diagonal matrix as well, with only the first element approaching infinity. If the first element of Σ' is not zero, only the first element of Σ'_{∞} , which is the variance of the first new axis $\text{var}(X'_1)$, will approach infinity when the system approaches the bifurcation point.

$$\begin{aligned}
 \text{vec}(\Sigma'_{\infty}) &= \begin{bmatrix} \text{var}(X'_1) \\ \text{cov}(X'_1, X'_2) \\ \vdots \end{bmatrix} \\
 &= \begin{bmatrix} -2\lambda_1 & 0 & \cdots \\ 0 & -\lambda_1 - \lambda_2 & \cdots \\ \vdots & \cdots & \cdots \end{bmatrix}^{-1} \begin{bmatrix} \sigma'_{11} \\ \sigma'_{12} \\ \vdots \end{bmatrix} \\
 &= \begin{bmatrix} \frac{1}{-2\lambda_1} & 0 & \cdots \\ 0 & \frac{1}{-\lambda_1 - \lambda_2} & \cdots \\ \vdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} \sigma'_{11} \\ \sigma'_{12} \\ \vdots \end{bmatrix} \\
 &\rightarrow \begin{bmatrix} \infty & 0 & \cdots \\ 0 & \frac{1}{-\lambda_2} & \cdots \\ \vdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} \sigma'_{11} \\ \sigma'_{12} \\ \vdots \end{bmatrix}.
 \end{aligned} \tag{7}$$

The auto-covariance of $\mathbf{X}'(t)$ can be calculated with the following formulas.

$$\lim_{t \rightarrow \infty} \text{cov}(\mathbf{X}'^{(t)}, \mathbf{X}'^{(t+\Delta t)}) = e^{\Lambda \Delta t} \Sigma'_{\infty}. \tag{8}$$

As Λ is a diagonal matrix, its exponential is a matrix with the exponential of diagonal elements:

$$e^{\Lambda \Delta t} = \text{diag}(e^{\lambda_1 \Delta t}, e^{\lambda_2 \Delta t}, \dots, e^{\lambda_n \Delta t}), \tag{9}$$

in which the first element approaches 1, and other elements are between 0 and 1. Therefore, the first element of the auto-covariance matrix, which means the auto-covariance of the first new axis $\text{cov}(X'_1(t), X'_1(t + \Delta t))$ will approach $\text{var}(X'_1)$.

Finally, we derive the situation back to the original coordinates. Because $\Sigma_{\infty} = \mathbf{Q} \Sigma'_{\infty} \mathbf{Q}^T$, the variance and covariance of all the variables with non-zero loading in the first eigenvector will approach infinity, and their correlations will approach 1.

$$\begin{aligned}
 \Sigma_{\infty} &= \begin{bmatrix} a_1 & b_1 & \cdots \\ a_2 & b_2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \text{var}(X'_1) & 0 & \cdots \\ 0 & \text{var}(X'_2) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} a_1 & a_2 & \cdots \\ b_1 & b_2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} a_1^2 \text{var}(X'_1)^2 + b_1^2 \text{var}(X'_2)^2 + \cdots & a_1 a_2 \text{var}(X'_1)^2 + b_1 b_2 \text{var}(X'_2)^2 + \cdots & \cdots \\ a_1 a_2 \text{var}(X'_1)^2 + b_1 b_2 \text{var}(X'_2)^2 + \cdots & a_2^2 \text{var}(X'_1)^2 + b_2^2 \text{var}(X'_2)^2 + \cdots & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \\
 &\rightarrow \begin{bmatrix} \infty & \infty & \cdots \\ \infty & \infty & \cdots \\ \vdots & \vdots & \cdots \end{bmatrix}
 \end{aligned} \tag{10}$$

$$\lim_{t \rightarrow \infty} \text{corr}(X_1, X_2) = \frac{\text{cov}(x_1, x_2)}{\sigma_1 \sigma_2} = \frac{a_1 a_2 \text{var}(X'_1)^2 + b_1 b_2 \text{var}(X'_2)^2 + \cdots}{\sqrt{a_1^2 \text{var}(X'_1)^2 + b_1^2 \text{var}(X'_2)^2 + \cdots + a_2^2 \text{var}(X'_1)^2 + b_2^2 \text{var}(X'_2)^2 + \cdots}} \rightarrow 1. \tag{11}$$

The same also holds for auto-correlation.

The eigenvalues and eigenvectors of \mathbf{J} may contain complex numbers even if we are working with a real function. This is because the local dynamic around the equilibrium point may not only contain shrinking and growing tendencies but also contain rotating. When there are complex eigenvalues and eigenvectors, they always appear in pairs. The imaginary part of the eigenvalues and eigenvectors of \mathbf{J} represents how the system tends to rotate.

If there are complex eigenvalues, we can do a block diagonalization instead of a plain diagonalization (Margalit & Rabinoff, 2019). In a block diagonalization, for each conjugated pair of eigenvalues and eigenvectors, we use the real and the imaginary part of the eigenvectors $[\text{Re}(\nu)\text{Im}(\nu)]$ as the vectors for coordinates transformation, and we use a 2×2 block $\begin{bmatrix} \text{Re}(\lambda) & \text{Im}(\lambda) \\ -\text{Im}(\lambda) & \text{Re}(\lambda) \end{bmatrix}$ to represent the rotating and shrinking or growing tendencies together.

We assume the first pair of complex eigenvalues are the dominant eigenvalues of \mathbf{J} , and only the real parts of them approach zero before the bifurcation point. In this case,

$$\begin{aligned} \text{vec}(\Sigma_{\infty}) &= \begin{bmatrix} \text{var}(X'_1) \\ \text{cov}(X'_1, X'_2) \\ \text{cov}(X'_1, X'_3) \\ \vdots \\ \text{cov}(X'_2, X'_1) \\ \text{var}(X'_2) \\ \text{cov}(X'_2, X'_3) \\ \vdots \end{bmatrix} \\ &= \begin{bmatrix} -2\text{Re}(\lambda_1) & -\text{Im}(\lambda_1) & 0 & \dots & -\text{Im}(\lambda_1) & 0 & 0 & \dots \\ \text{Im}(\lambda_1) & -2\text{Re}(\lambda_1) & 0 & \dots & 0 & -\text{Im}(\lambda_1) & 0 & \dots \\ 0 & 0 & -\text{Re}(\lambda_1) - \lambda_3 & \dots & 0 & 0 & -\text{Im}(\lambda_1) & \dots \\ \dots & \dots \\ \text{Im}(\lambda_1) & 0 & 0 & \dots & -2\text{Re}(\lambda_1) & -\text{Im}(\lambda_1) & 0 & \dots \\ 0 & \text{Im}(\lambda_1) & 0 & \dots & -\text{Im}(\lambda_1) & -2\text{Re}(\lambda_1) & 0 & \dots \\ 0 & 0 & \text{Im}(\lambda_1) & \dots & 0 & 0 & -\text{Re}(\lambda_1) - \lambda_3 & \dots \\ \dots & \dots \end{bmatrix}^{-1} \begin{bmatrix} \sigma'_{11} \\ \sigma'_{12} \\ \sigma'_{13} \\ \vdots \\ \sigma'_{21} \\ \sigma'_{22} \\ \sigma'_{23} \\ \vdots \end{bmatrix}. \end{aligned} \quad (12)$$

To simplify the calculation of the inverse, we change the order of the matrix and vectors to make the values that approach infinity all to the upper-left corner of the matrix. The matrix then becomes a block diagonal matrix that we can obtain the invert of each block separately.

$$\begin{aligned}
& \begin{bmatrix} \text{var}(X'_1) \\ \text{cov}(X'_1, X'_2) \\ \text{cov}(X'_2, X'_1) \\ \text{var}(X'_2) \\ \text{cov}(X'_1, X'_3) \\ \text{cov}(X'_2, X'_3) \\ \vdots \end{bmatrix} \\
&= \begin{bmatrix} -2\text{Re}(\lambda_1) & -\text{Im}(\lambda_1) & -\text{Im}(\lambda_1) & 0 & 0 & 0 & \cdots & \cdots \\ \text{Im}(\lambda_1) & -2\text{Re}(\lambda_1) & 0 & -\text{Im}(\lambda_1) & 0 & 0 & \cdots & \cdots \\ \text{Im}(\lambda_1) & 0 & -2\text{Re}(\lambda_1) & -\text{Im}(\lambda_1) & 0 & 0 & \cdots & \cdots \\ 0 & \text{Im}(\lambda_1) & \text{Im}(\lambda_1) & -2\text{Re}(\lambda_1) & 0 & 0 & \cdots & \cdots \\ 0 & 0 & 0 & 0 & -\text{Re}(\lambda_1) - \lambda_3 & -\text{Im}(\lambda_1) & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \text{Im}(\lambda_1) & -\text{Re}(\lambda_1) - \lambda_3 & \cdots & \cdots \\ \cdots & \cdots \\ \cdots & \cdots \end{bmatrix}^{-1} \begin{bmatrix} \sigma'_{11} \\ \sigma'_{12} \\ \sigma'_{21} \\ \sigma'_{22} \\ \sigma'_{13} \\ \sigma'_{23} \\ \vdots \end{bmatrix}, \quad (13) \\
&= \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{A}_2 & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} \sigma'_{11} \\ \sigma'_{12} \\ \sigma'_{21} \\ \sigma'_{22} \\ \sigma'_{13} \\ \sigma'_{23} \\ \vdots \end{bmatrix}
\end{aligned}$$

in which (symbolic calculation performed with Mathematica, Wolfram Research, Inc., 2022)

$$\begin{aligned}
& \mathbf{A}_1 \\
&= \begin{bmatrix} \frac{-8\text{Re}(\lambda_1)^3 - 4\text{Re}(\lambda_1)\text{Im}(\lambda_1)^2}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{4\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{4\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{4\text{Re}(\lambda_1)\text{Im}(\lambda_1)^2}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} \\ -\frac{4\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{-8\text{Re}(\lambda_1)^3 - 4\text{Re}(\lambda_1)\text{Im}(\lambda_1)^2}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{4\text{Re}(\lambda_1)\text{Im}(\lambda_1)^2}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{4\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} \\ -\frac{4\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{4\text{Re}(\lambda_1)\text{Im}(\lambda_1)^2}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{-8\text{Re}(\lambda_1)^3 - 4\text{Re}(\lambda_1)\text{Im}(\lambda_1)^2}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{4\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} \\ -\frac{4\text{Re}(\lambda_1)\text{Im}(\lambda_1)^2}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{4\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{4\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} & \frac{-8\text{Re}(\lambda_1)^3 - 4\text{Re}(\lambda_1)\text{Im}(\lambda_1)^2}{16\text{Re}(\lambda_1)^4 + 16\text{Re}(\lambda_1)^2\text{Im}(\lambda_1)^2} \end{bmatrix} \\
&\rightarrow \begin{bmatrix} \frac{1}{4\text{Re}(\lambda_1)} & \frac{1}{4\text{Im}(\lambda_1)} & \frac{1}{4\text{Im}(\lambda_1)} & -\frac{1}{4\text{Re}(\lambda_1)} \\ -\frac{1}{4\text{Im}(\lambda_1)} & \frac{1}{4\text{Re}(\lambda_1)} & \frac{1}{4\text{Re}(\lambda_1)} & \frac{1}{4\text{Im}(\lambda_1)} \\ \frac{1}{4\text{Im}(\lambda_1)} & \frac{1}{4\text{Re}(\lambda_1)} & -\frac{1}{4\text{Re}(\lambda_1)} & \frac{1}{4\text{Im}(\lambda_1)} \\ -\frac{1}{4\text{Re}(\lambda_1)} & \frac{1}{4\text{Im}(\lambda_1)} & -\frac{1}{4\text{Im}(\lambda_1)} & -\frac{1}{4\text{Re}(\lambda_1)} \end{bmatrix} \\
& \mathbf{A}_2 \\
&= \begin{bmatrix} \frac{-\text{Re}(\lambda_1) - \lambda_3}{\text{Re}(\lambda_1)^2 + 2\text{Re}(\lambda_1)\lambda_3 + \text{Im}(\lambda_1)^2 + \lambda_3^2} & \frac{\text{Im}(\lambda_1)}{\text{Re}(\lambda_1)^2 + 2\text{Re}(\lambda_1)\lambda_3 + \text{Im}(\lambda_1)^2 + \lambda_3^2} \\ \frac{\text{Im}(\lambda_1)}{\text{Re}(\lambda_1)^2 + 2\text{Re}(\lambda_1)\lambda_3 + \text{Im}(\lambda_1)^2 + \lambda_3^2} & \frac{-\text{Re}(\lambda_1) - \lambda_3}{\text{Re}(\lambda_1)^2 + 2\text{Re}(\lambda_1)\lambda_3 + \text{Im}(\lambda_1)^2 + \lambda_3^2} \end{bmatrix} \rightarrow \begin{pmatrix} -\frac{1}{\lambda_3} & 0 \\ 0 & -\frac{1}{\lambda_3} \end{pmatrix}. \quad (14)
\end{aligned}$$

Because $\sigma'_{12} = \sigma'_{21}$, $\text{cov}(X'_1, X'_2) = \text{cov}(X'_2, X'_1) = 0$. As long as σ'_{11} and σ'_{22} are not both zero, $\text{var}(X'_1)$ and $\text{var}(X'_2)$ will both approach infinity. Because λ_3 does not approach zero, both $\text{cov}(X'_1, X'_3)$ and $\text{cov}(X'_2, X'_3)$ will remain finite.

Following the same procedure as used for the real-valued case, it can be proved that the variance of any variables that have non-zero loading in either the real part or the imaginary part of the first eigenvector will approach infinity. Their correlation and auto-correlation will approach 1 as long as the two variables of consideration both belong to either the real part or the imaginary part of the first eigenvector.

To summarize, with the aforementioned assumptions, we can derive the fact that the variables involved in this critical direction should have the variance, (auto)covariance, and (auto)correlations increasing before the transition.

Examples with the Bivariate Cusp Model

Since the derivation above may be difficult to understand, here we provide an example with the bivariate cusp model (Figure 1A in the main text; also described in Supplementary Materials B). Note that this is a special case of the general derivation above.

The model is specified as follows:

$$\begin{aligned} a &= \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y, \\ b &= \frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y, \\ U &= \frac{1}{4}a^4 - \frac{3}{2}a^2 + \lambda a + b^2 \\ &= \frac{1}{16}(x+y)^4 - \frac{3}{4}(x+y)^2 + \frac{\sqrt{2}\lambda}{2}(x+y) + \frac{1}{2}(x-y)^2. \end{aligned} \tag{15}$$

The partial derivatives of U with respect to x and y are as follows:

$$\begin{aligned} \frac{\partial U}{\partial x} &= \frac{1}{4}(x+y)^3 - \frac{3}{2}(x+y) + \frac{\sqrt{2}\lambda}{2} + (x-y), \\ \frac{\partial U}{\partial y} &= \frac{1}{4}(x+y)^3 - \frac{3}{2}(x+y) + \frac{\sqrt{2}\lambda}{2} - (x-y). \end{aligned} \tag{16}$$

The force on the system is the gradient of U , namely $\mathbf{F} = \begin{bmatrix} -\frac{\partial U}{\partial x} \\ -\frac{\partial U}{\partial y} \end{bmatrix}$. Therefore, the Jacobian

of the system is the following:

$$\begin{aligned}
\mathbf{J} &= \begin{bmatrix} \frac{\partial F_x}{\partial x} & \frac{\partial F_x}{\partial y} \\ \frac{\partial F_y}{\partial x} & \frac{\partial F_y}{\partial y} \end{bmatrix} \\
&= \begin{bmatrix} -\frac{3}{4}(x+y)^2 + \frac{1}{2} & -\frac{3}{4}(x+y)^2 + \frac{5}{2} \\ -\frac{3}{4}(x+y)^2 + \frac{5}{2} & -\frac{3}{4}(x+y)^2 + \frac{1}{2} \end{bmatrix} \\
&= \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & -\frac{3}{2}(x^2 + 2xy + y^2 - 2) \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}
\end{aligned} \tag{17}$$

At the equilibrium points, both partial derivatives should be zero. Comparing the two equations above, it is obvious that $x - y = 0$, $x = y$. Therefore, at the equilibrium point, we have the following condition:

$$\frac{\partial U}{\partial x} = \frac{\partial U}{\partial y} = 2x^3 - 3x + \frac{\sqrt{2}\lambda}{2} = 0 \tag{18}$$

The closed-form solution for this cubic equation with a parameter has a complex form. Interested readers may use numerical software (e.g., Mathematica, Wolfram Research, Inc., 2022) to observe how the solution differs with different λ values. Here, we only provide the conclusion about the asymptotic behavior close to the bifurcation point. As $\lambda \rightarrow 2^-$, the stable equilibrium point at the negative half approaches $(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$ and the Jacobian approaches $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$. The two eigenvalues of the Jacobian then approach 0 and -2 (which can also be seen from the decomposition in Equation 17).

Following Equation 5, the variance-covariance matrix of the system under noise is given by the following (note that, as we have the Jacobian, we calculate the variance-covariance matrix in the original coordinates):

$$\begin{aligned}
\text{vec}(\boldsymbol{\Sigma}_\infty) &= [-(\mathbf{J} \oplus \mathbf{J})]^{-1} \text{vec}(\boldsymbol{\Sigma}) \\
&\rightarrow \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}^{-1} \text{vec}(\boldsymbol{\Sigma}),
\end{aligned} \tag{19}$$

in which the first matrix is singular. To know what exact elements of the matrix are approaching infinity, we do a calculation when λ is a little smaller than 2, that the equilibrium

is a little to the negative side of $(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$. Assume the Jacobian is now $\begin{bmatrix} -1.01 & 1.01 \\ 1.01 & -1.01 \end{bmatrix}$, we have (numerical calculation was conducted by Mathematica, Wolfram Research, Inc., 2022):

$$\begin{aligned}\text{vec}(\Sigma_\infty) &= [-(\mathbf{J} \oplus \mathbf{J})]^{-1} \text{vec}(\Sigma) \\ &= \begin{bmatrix} 2.02 & -1.01 & -1.01 & 0 \\ -1.01 & 2.02 & 0 & -1.01 \\ -1.01 & 0 & 2.02 & -1.01 \\ 0 & -1.01 & -1.01 & 2.02 \end{bmatrix}^{-1} \text{vec}(\Sigma) \\ &= \begin{bmatrix} 6.68851 \times 10^{15} & 6.68851 \times 10^{15} & 6.68851 \times 10^{15} & 6.68851 \times 10^{15} \\ 6.68851 \times 10^{15} & 6.68851 \times 10^{15} & 6.68851 \times 10^{15} & 6.68851 \times 10^{15} \\ 6.68851 \times 10^{15} & 6.68851 \times 10^{15} & 6.68851 \times 10^{15} & 6.68851 \times 10^{15} \\ 6.68851 \times 10^{15} & 6.68851 \times 10^{15} & 6.68851 \times 10^{15} & 6.68851 \times 10^{15} \end{bmatrix} \text{vec}(\Sigma)\end{aligned}\quad (20)$$

Therefore, as long as there is some noise in the system (i.e., $\Sigma \neq \mathbf{0}$), the variance and covariance of x and y will approach infinity.

We also do the same calculation for the second model in the main text (Figure 1B in the main text, also described in Supplementary Materials B). This time the model is specified as:

$$U = \frac{1}{4}x^4 - \frac{3}{2}x^2 + \lambda x + y^2. \quad (21)$$

We then have:

$$\begin{aligned}\frac{\partial U}{\partial x} &= x^3 - 3x + \lambda, \\ \frac{\partial U}{\partial y} &= 2y.\end{aligned}\quad (22)$$

$$\begin{aligned}\mathbf{J} &= \begin{bmatrix} \frac{\partial F_x}{\partial x} & \frac{\partial F_x}{\partial y} \\ \frac{\partial F_y}{\partial x} & \frac{\partial F_y}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} -3x^2 + 3 & 0 \\ 0 & -2 \end{bmatrix}\end{aligned}\quad (23)$$

When $\lambda \rightarrow 2$, the equilibrium point at the negative half approaches $(-1, 0)$, and the Jacobian approaches $\begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}$. Again, we consider the case when λ is a little smaller than 2, and $\mathbf{J} = \begin{bmatrix} -0.01 & 0 \\ 0 & -2 \end{bmatrix}$. Similarly, we have (numerical calculation was conducted by Mathematica, Wolfram Research, Inc., 2022):

$$\begin{aligned}
\text{vec}(\Sigma_{\infty}) &= [-(\mathbf{J} \oplus \mathbf{J})]^{-1} \text{vec}(\Sigma) \\
&= \begin{bmatrix} 0.02 & 0 & 0 & 0 \\ 0 & 2.01 & 0 & 0 \\ 0 & 0 & 2.01 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}^{-1} \text{vec}(\Sigma) \\
&= \begin{bmatrix} 50 & 0 & 0 & 0 \\ 0 & 0.497512 & 0 & 0 \\ 0 & 0 & 0.497512 & 0 \\ 0 & 0 & 0 & 0.25 \end{bmatrix} \text{vec}(\Sigma).
\end{aligned} \tag{24}$$

Here, as long as there is some noise for x (i.e., $\Sigma_{xx} \neq 0$), the variance of x and x only will approach infinity. Other elements of the variance-covariance matrix will remain small.

D2. Landscape Illustrations and Simulations

Multivariate Landscapes and Simulations

In this section, we explain the details of the results shown in Figure 1-3 in the main text. To illustrate different scenarios for landscape changes, we created a set of landscape functions, all of which are based on the general bistable landscape function used by Shi et al. (2016), $U = \frac{1}{4}x^4 - \frac{3}{2}x^2 + \lambda x$. We performed coordinate transformations or added additional components to transform the landscape into certain geometric shapes. Specifications of those landscapes are listed as follows.

System (a) is a coordinate transformation of System (b) (see below) to make both x - and y -axes involved in the transition:

$$a = \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y, \tag{1}$$

$$b = \frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y, \tag{2}$$

$$U = \frac{1}{4}a^4 - \frac{3}{2}a^2 + \lambda a + b^2. \tag{3}$$

System (b) is a multivariate extension of the original one-dimensional landscape function. It adds the y -axis, which takes a quadratic form and does not introduce additional basins:

$$U = \frac{1}{4}x^4 - \frac{3}{2}x^2 + \lambda x + y^2. \tag{4}$$

System (c) added a term, $-2\text{ReLU}(x)y$, to System (b) to show an example in which the starting direction of the transition does not directly point to the endpoint of the transition:

$$U = \frac{1}{4}x^4 - \frac{3}{2}x^2 + \lambda x + y^2 - 2\text{ReLU}(x)y, \text{ in which } \text{ReLU}(x) = \max(x, 0). \quad (5)$$

System (d) is a coordinate transformation of System (b) to make the transition along a curve instead of a straight line:

$$\theta = \arctan \frac{y}{x}, \quad (6)$$

$$a = \frac{4\theta}{\pi} + 0.5, \quad (7)$$

$$b = 3 \left(\sqrt{x^2 + y^2} - 3 \right), \quad (8)$$

$$U = \frac{1}{4}a^4 - \frac{3}{2}a^2 + \lambda a + b^2. \quad (9)$$

System (e) is a coordinate transformation of System (b) to make the new attractor a circle instead of a point:

$$a = 3 \left(\sqrt{x^2 + y^2} \right) - 1, \quad (10)$$

$$U = \frac{1}{4}a^4 - \frac{3}{2}a^2 + \lambda a. \quad (11)$$

We then draw all the landscape functions in Figure 1. All the examples shown in Figure 1 of the main text are cusp bifurcations.

Taking the gradients from the landscape functions, and adding the stochastic term, we can have the stochastic dynamic equations of the system,

$$dx = -\frac{\partial U}{\partial x} dt + \sigma_x dW_x, \quad (12)$$

$$dy = -\frac{\partial U}{\partial y} dt + \sigma_y dW_y, \quad (13)$$

which can be used to simulate the system with the Euler–Maruyama method. In our simulations, we let the control parameter λ change according to the formula $\lambda = (3 - t)/100$, and we used a timestep of 0.01, a simulation length of 700 time units, and a noise level of $\sigma = 0.3$.

EWS Simulation Shown in the Main Text

In this section, we explain the details of the results shown in Figure 4 in the main text.

Again, we use the simple gradient system with noise by Shi et al. (2016) as the model for our simulation. The model contains one state variable, x , and a control parameter, λ . The potential function of the system, U , is specified as follows. Note that this function only differs in a constant coefficient compared with the function introduced in the previous section. This is to make the change rate in the simulations more realistic.

$$U(x, \lambda) = 100 \left(\frac{1}{4}x^4 - \frac{3}{2}x^2 + \lambda x \right). \quad (14)$$

The dynamic functions of the system are then specified as

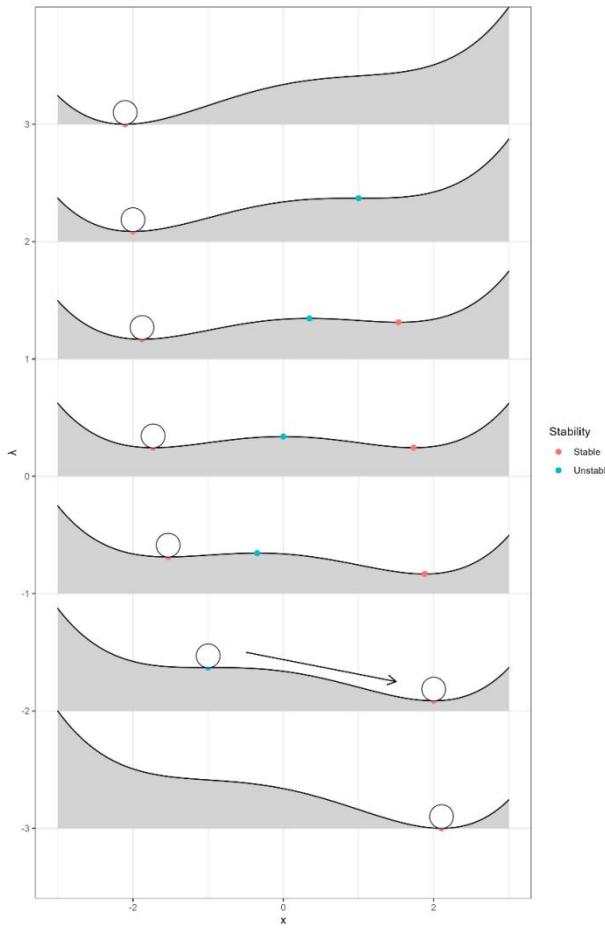
$$\frac{dx}{dt} = -\frac{\partial U(x, \lambda)}{\partial x} + \sqrt{2\sigma}\xi(t), \quad (15)$$

where dx/dt represents the change rate of x , $\partial U(x, \lambda)/\partial x$ represents the gradient of the potential function with respect to x , σ represents the strength of the noise and was set as 10 in this study (as in Shi et al., 2016), and $\xi(t)$ represents standard white noise. The potential landscapes of the system with different λ , as well as the equilibrium points of the system where $\partial U(x, \lambda)/\partial x = 0$, are shown in Figure D1. For simulating the change of the landscape, the initial value of λ is set as -3, and the changing rate of λ is set as $d\lambda/dt = 1$. When the simulation starts, there is only one basin for the system. We refer to this as the *positive phase* because it is in the positive semi-axis of x . As λ increases to -2, the second basin appears, and its stability increases as λ further increases. We refer to this basin as the *negative phase*. When λ increases to 2, the system reaches its bifurcation point. The positive phase of the system disappears and the negative basin becomes the only possible basin. All simulations were numerically performed using the Euler-Maruyama method, with 10^{-4} as the step size and 6 as the total time length. The raw simulation data were subsampled by a factor of 10 to reduce the length of the data. Therefore, the time interval between adjunct time points in the output is 10^{-3} .

Here we also show two early warning indicators that are commonly used in previous empirical studies: increasing variance and increasing autocorrelation function (ACF). All the two parameters were estimated with the overlapping moving window approach. The window size was selected as 200 time points ($\Delta t = 0.2$), and each time the window moved forward for 20 time points ($\Delta t = 0.02$). Here the number of time points in each window is much more than the typical value in empirical studies. We chose this large value because the main purpose of this simulation is to qualitatively show the phenomenon of EWSs, not to provide guidance on the window size for empirical studies. A rather large window size can ensure the stability of

the results. Within each window, the variance and lag-1 ACF were calculated. Specifically, the data were linear-detrended within each window before calculating the autocorrelation coefficient. The right-aligned windows were used, which means that, for example, the variance calculated within the window from $t = 0$ to $t = 0.2$ is regarded as the variance at $t = 0.2$. Thus, no future information is included in the moving windows statistics.

Figure D1. Illustration of a cusp bifurcation specified by Equations 1. The gray areas represent the potential landscapes for different λ values. The red dots represent the stable equilibrium points and the blue points represent the unstable equilibrium points. The white circles represent how the system experiences a sudden change during a cusp bifurcation. At the bifurcation point ($\lambda = 2$), the state of the system became unstable, so it transitions to another state.



Additional Simulation Results

In this section, we provide some additional simulation results related to, but not reported in the main text.

Consequences of Calculating EWSs not Strictly Before the Transition. If the EWS calculation window includes the transition itself, the increase in variance and autocorrelation

may be falsely taken as evidence of EWSs. To illustrate this point, we conducted simulations for another condition, in which the transition is purely driven by a large fluctuation so that there are no EWSs before the transition. This condition is set up the same as described in the *EWS simulation shown in the main text*, except that the parameter λ is held constant at 0, which means that the potential function U does not change through the simulation and that there is a strong noise $\Delta x = -3$ at $t = 3$ that pushes the system to the negative phase. We use the time that the system first crosses the barrier as the time of the transition (t_{trans}), and we use Kendall's τ calculated with the Kendall package (McLeod, 2011) to evaluate the trends of variance and autocorrelation. We investigated three types of ranges in the current research: (1) strictly before the transition, for which τ was calculated in the range from $t_{\text{trans}} - 1.5$ to t_{trans} ; (2) roughly before the transition, for which τ was calculated in the range from $t_{\text{trans}} - 1.5$ to $t_{\text{trans}} + 0.5$; and (3) around the transition, for which τ was calculated in the range from $t_{\text{trans}} - 1.5$ to $t_{\text{trans}} + 1.5$. These conditions were set to mimic different empirical studies where EWSs are calculated strictly before the transition (when the transition indicator is calculated in at least the same frequency as EWSs), roughly before the transition (when the transition indicator is calculated through the whole period but in a lower frequency as EWSs), and in a large range that may contain a transition (when the transition indicator is only calculated before or after the whole study period). The range sizes are set as roughly one order of magnitude larger than the window sizes for moving window statistics, which is often the case in empirical studies. For each condition, the simulation was replicated 10^3 times and the results and statistical indicators were recorded. Examples of the simulated time series are shown in Figure D2, and the distributions of τ in repeated simulations are shown in Figure D3. From the simulation results, it is clear that if the EWS calculation window is not strictly before the transition, the statistical effect of the transition itself may be falsely taken as evidence of EWSs, and true EWSs may not be detected because the decrease in variance and ACF may average out the true EWSs.

Figure D2. Simulation examples for sudden changes caused by (a) the bifurcation of the system, and (2) a large fluctuation

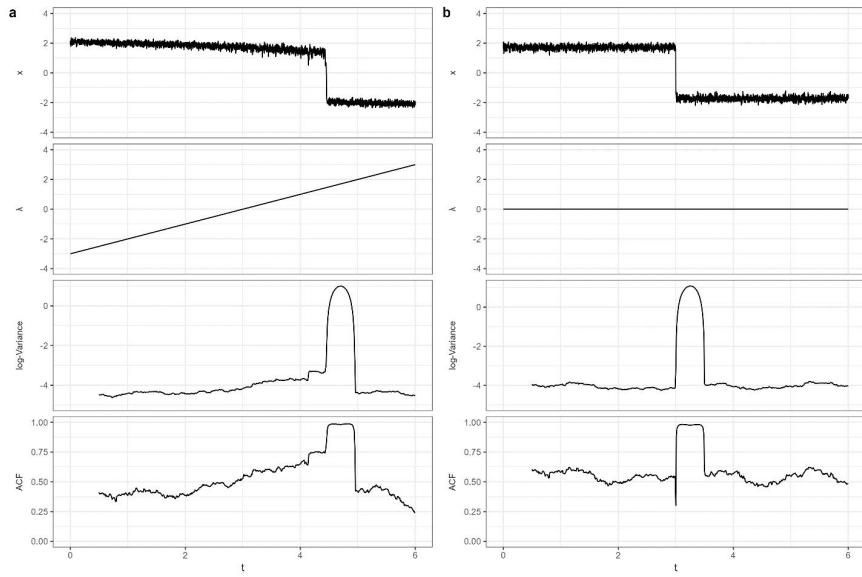
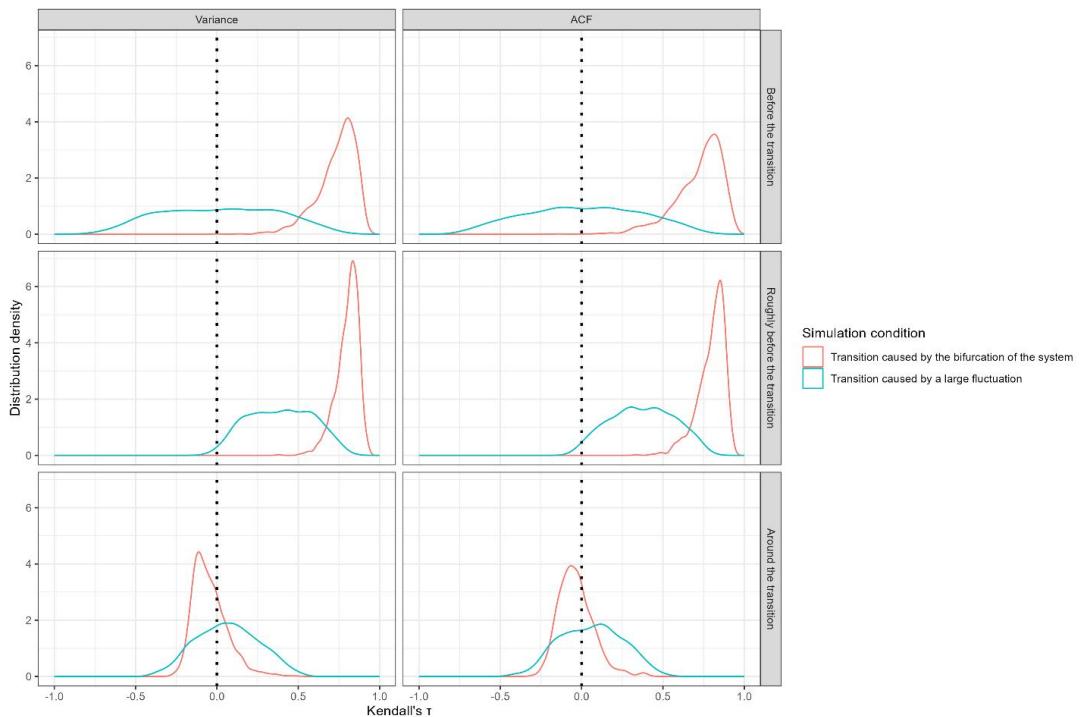


Figure D3. Trends of variance and ACF represented with the distribution of Kendall's τ from 10^3 simulations, for two simulation conditions and different periods.



Consequences of Using Inconsistent Variables for Calculating EWSs and Detecting Sudden Changes. In the main text, we explained why, from theoretical considerations, researchers should use a consistent set of variables to calculate EWSs and detect sudden changes. Here, we use a simplified simulation example to illustrate the possible consequences of using different variables in empirical studies, with some variables *only* used for calculating

EWSs and some *only* used for detecting sudden changes. The model we used here is very similar to the one used in the section *Consequences of calculating EWSs not strictly before the transition*, but there are two sets of equations for two variables x_1 and x_2 :

$$V_1(x_1, \lambda_1) = 100 \left(\frac{1}{4}x_1^4 - \frac{3}{2}x_1^2 + \lambda_1 x_1 \right), \quad (16)$$

$$\frac{dx_1}{dt} = -\frac{\partial V_1(x_1, \lambda_1)}{\partial x_1} + \sqrt{2\sigma_1}\xi_1(t), \quad (17)$$

$$V_2(x_2, \lambda_2) = 100 \left(\frac{1}{4}x_2^4 - \frac{3}{2}x_2^2 + \lambda_2 x_2 \right), \quad (18)$$

$$\frac{dx_2}{dt} = -\frac{\partial V_2(x_2, \lambda_2)}{\partial x_2} + \sqrt{2\sigma_2}\xi_2(t). \quad (19)$$

We set $\sigma_1 = 400$ and $\sigma_2 = 10$ so that x_1 is influenced by strong noise and does not show a single, clear transition, and x_2 has a transition when λ_2 approaches 2 (as in Shi et al., 2016). We further include a weak relationship between x_1 and x_2 by associating λ_1 and λ_2 . The relationship of λ_1 and λ_2 can, in principle, take any form. We illustrate two simple conditions: (a) $\lambda_2 = \lambda_1$, (b) $\lambda_2 = \lambda_1 + 2$. The starting value and changing rate of λ_1 were set the same as in section *Consequences of calculating EWSs not strictly before the transition*: the initial value of λ_1 is -3 and $d\lambda_1/dt = 1$. The simulation methods (e.g., simulation length, step size, etc.) are the same as in the section *Consequences of calculating EWSs not strictly before the transition*. Again, we show simulation examples in Figure D4, and the distribution of Kendall's τ in repeated simulations in Figure D5. From the simulation results, we can see that if EWSs and sudden changes are detected from different variables, the trend of variance or ACF may not be related to the sudden change.

Figure D4. Simulation examples for two conditions.

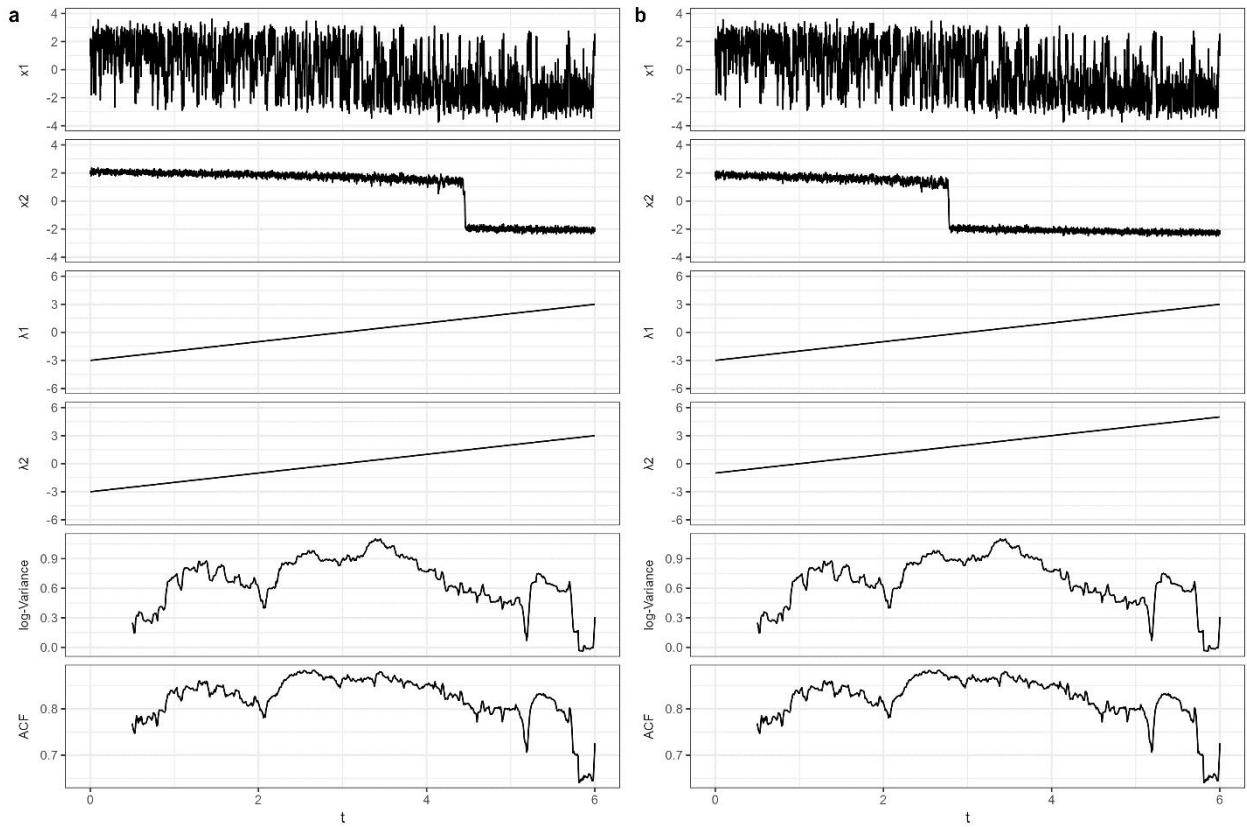
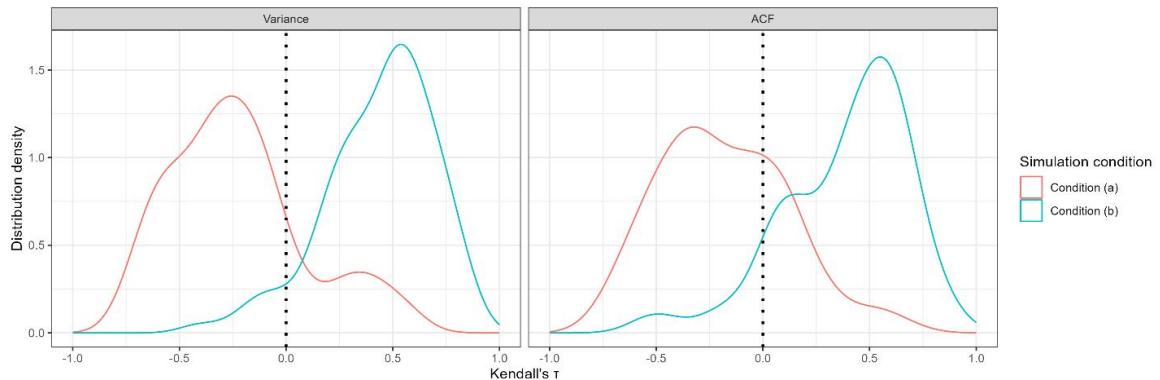


Figure D5. Trends of variance and ACF represented with the distribution of Kendall's τ from 10^3 simulations, for two conditions and different periods.



Code Availability

All the code necessary to replicate the results shown in this file can be found in the OSF repository of this project <https://osf.io/f659u/>.

Appendix E. Supplementary Materials for Chapter 7

E1. Simulation Specifications

In the main text, we presented various simulations to illustrate different types of transitions. We used stochastic differential equations (SDEs) to specify the models and conducted those simulations with the `sde` package (Iacus, 2022) for univariate cases and the `yuima` package (Brouste et al., 2014) for multivariate cases. The R code to replicate all the simulations is available at <https://osf.io/4jaqk/>.

Examples of Various Types of Transitions

We run each situation from $T = 0$ to $T = 100$, with a time step of 0.01, except for the R-tipping case for which we run the simulation from $T = 0$ to $T = 10$.

The landscape function of the B-tipping (fold bifurcation) example shown in the main text is specified as follows (adapted from Shi et al., 2016),

$$U = 0.5 \left(\frac{1}{4}x^4 - \frac{3}{2}x^2 + \lambda x \right), \quad (1)$$

where λ is the control parameter that slowly changes following

$$\lambda = 3 - 0.06t. \quad (2)$$

The dynamics of the variable x is specified with the following SDE,

$$dx = -\frac{\partial U}{\partial x} dt + \sigma dW, \quad (3)$$

where σ represents the strength of noise and takes 0.3 in this example.

For Hopf bifurcation, an additional variable is required to make the vibration behavior possible. Hence, we use the following SDEs,

$$dx = \left((\lambda + x^2 + y^2 - (x^2 + y^2)^2)x - (\omega + b(x^2 + y^2))y \right) dt + \sigma dW_1, \quad (4)$$

$$dy = \left((\lambda + x^2 + y^2 - (x^2 + y^2)^2)y + (\omega + b(x^2 + y^2))x \right) dt + \sigma dW_2, \quad (5)$$

for which the drift function is transformed from the polar coordinates' representation,

$$U = \frac{1}{6}r^6 - \frac{1}{4}r^4 + \frac{1}{2}\lambda r^2, \quad (6)$$

$$\frac{dr}{dt} = -\frac{\partial U}{\partial r}, \quad (7)$$

$$\frac{d\theta}{dt} = \omega + br^2. \quad (8)$$

We use the parameter setting $\omega = 0.5$, $b = 0.1$, $\sigma = 0.1$ in our simulation. Additionally, for the reversed direction, we use $\lambda = -3 + 0.06t$.

For the N-tipping example, we use the same potential function form as the B-tipping example but hold $\lambda = 0$, and we use a stronger noise of $\sigma = 0.65$.

For the R-tipping example, we use the following potential function,

$$U = 0.5 \left(\frac{1}{4}(x - \lambda)^4 - \frac{3}{2}(x - \lambda)^2 \right), \quad (9)$$

and we set $\lambda = -5 + 1.1t$.

For the N-diffusion example, we use the same potential function form as the B-tipping example but hold $\lambda = 0$, and we use a stronger noise of $\sigma = 2$.

The key parameters of those examples are summarized in Table E1

Table E1. Simulation settings in the examples. The parameters corresponding to the main cause of the transitions are marked bold.

Type of transition	Stability change (shape)	Stability change (position)	Variable change	Noise strength
B-tipping	0.06	0.00	0.50	0.30
N-tipping	0.00	0.00	0.50	0.65
R-tipping	0.00	1.20	0.50	0.30
N-diffusion	0.00	0.00	0.50	2.00

Examples of Clinical Scenarios

For the first scenario, we use the following potential function,

$$U = 0.5 \left(\left(\frac{1}{4}x^4 - \frac{3}{2}x^2 + \lambda x \right) + 0.1 \left(\frac{1}{4}y^4 - \frac{1}{2}y^2 - xy \right) \right), \quad (10)$$

the following for the control parameter,

$$\lambda = 1 - 0.06t, \quad (11)$$

and the noise strength is set as $\sigma_x = 0.3$, $\sigma_y = 0.5$.

For the second scenario, we use the same setting as the N-tipping example for x , the following potential function for λ ,

$$U = \ln(1 + (1.5x - \lambda)^2) + \ln(1 + x^2) - 0.18\lambda. \quad (12)$$

For y , we use the same setting as the R-tipping example.

E2. Illustrations of Other Types of Bifurcations

Figure E1. A simulated time series and ball-and-landscape illustrations for Hopf bifurcation.

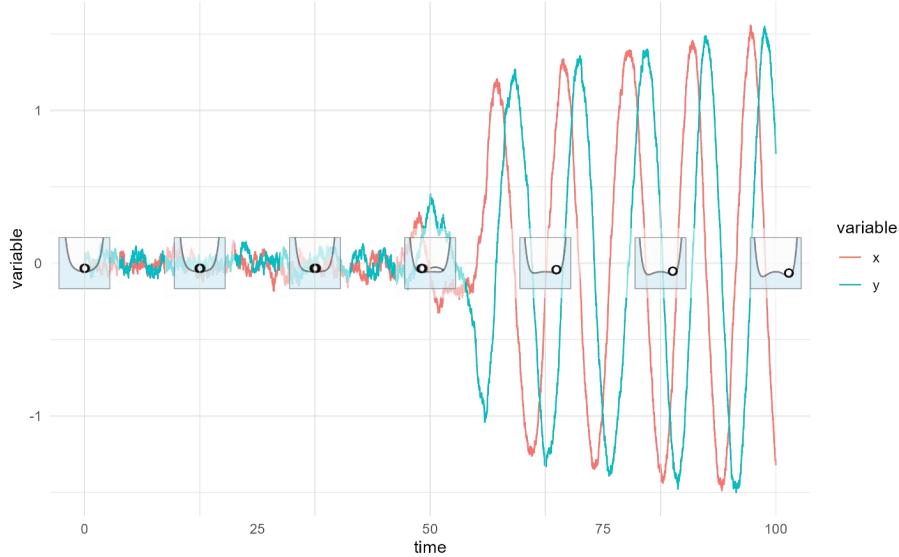
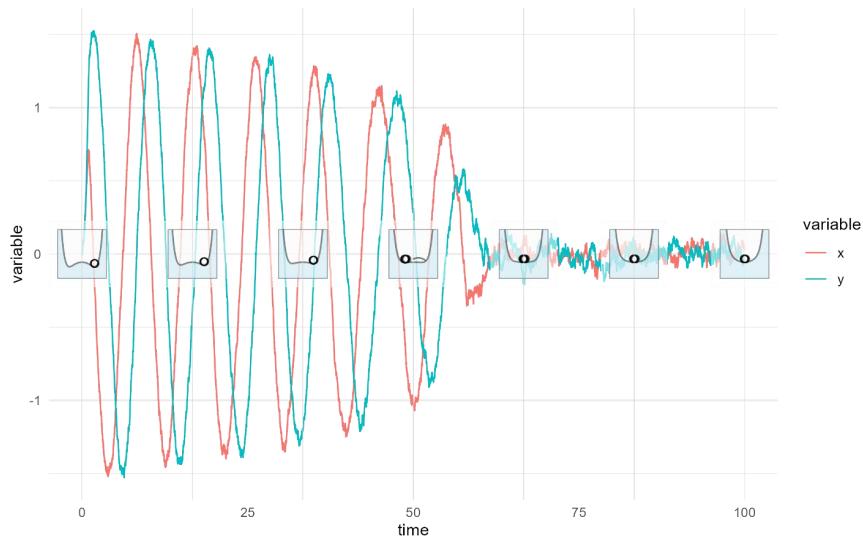


Figure E2. A simulated time series and ball-and-landscape illustrations for Hopf bifurcation (reversed direction).

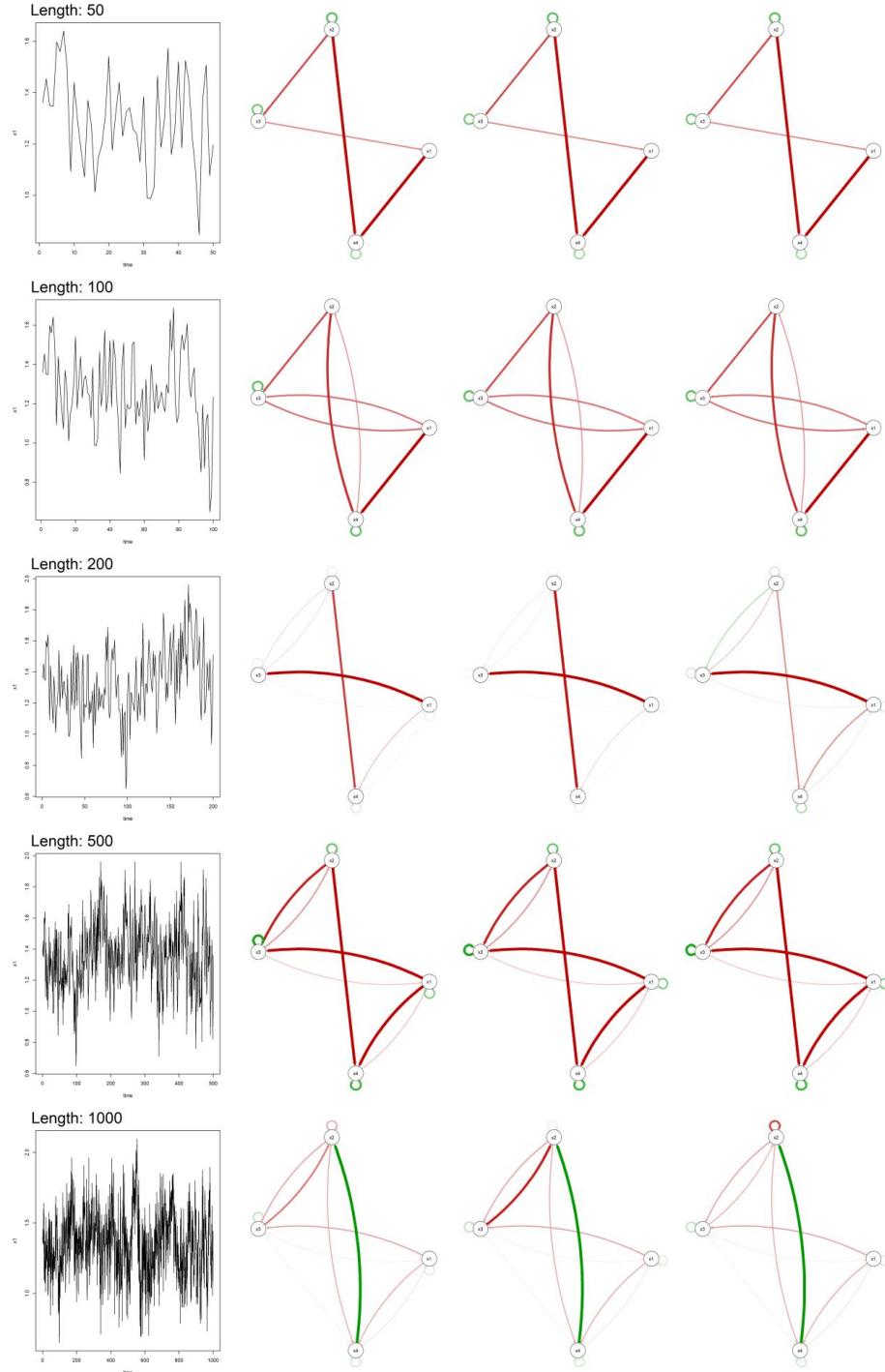


Appendix F. Supplementary Materials for Chapter 8

F1. The Influence of Noise and Transitions on quadVAR Model Estimation

For linear models, the influence of noise on model estimates is rather straightforward, in that larger noise reduces the signal-to-noise ratio, thus making the parameter estimations less accurate (e.g., Mansueto et al., 2022). However, for nonlinear models, the relationship between noise and model estimation is more complicated. We show an example to illustrate this idea, in which we reduce the noise level from $SD = 1$ to $SD = 0.2$. The results are shown in Figure F1. The networks estimated from this example are similar to the linearized networks for the negative state shown in Figure 3. However, no nonlinear effects were estimated in this case, making the networks for different states the same as each other. This is because the noise in the system is too weak, making the system always stay in the negative phase. The simulated data do not contain enough information about the system in the alternative phase, making it difficult to estimate the nonlinear effects.

Figure F1. Linearized networks for the simulation examples with different lengths, with noise $SD = 0.2$. In each row, the first plot shows the raw time series of x_1 , the second to the fourth plots show the linearized networks estimated from simulation data for the neutral state ($x_1 = x_2 = x_3 = x_4 = 2.80$), the positive state ($x_1 = x_2 = 4.89, x_3 = x_4 = 1.36$), and the negative state ($x_1 = x_2 = 1.36, x_3 = x_4 = 4.89$).



F2. The Comparable Linear System

The comparable linear system is specified using parameters of a VAR model estimated from the simulated data from the nonlinear, bistable system. The simulated dataset contains 1000 data points. The coefficients for the VAR model are as follows. The precise value without rounding can also be reproduced with the code online.

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{c} \quad (1)$$

$$\mathbf{A} = \begin{bmatrix} 0.44 & 0.08 & -0.18 & -0.26 \\ 0.14 & 0.41 & -0.18 & -0.23 \\ -0.17 & -0.25 & 0.46 & 0.08 \\ -0.2 & -0.18 & 0.13 & 0.44 \end{bmatrix} \quad (2)$$

$$\mathbf{c} = \begin{bmatrix} 2.76 \\ 2.61 \\ 2.68 \\ 2.46 \end{bmatrix} \quad (3)$$

F3. Definitions of Classification Metrics

For the classification metrics, we do not only consider whether a value is zero but also consider whether the sign of the value is correct according to the true value. Therefore, we are interested in the correctness of the directed inferences of the coefficients. Specifically, we define the following:

- (1) True positive, when the true coefficient is nonzero, and the estimated coefficient is also nonzero and has the same sign as the true coefficient.
- (2) True negative, when the true coefficient is zero, and the estimated coefficient is also zero.
- (3) False positive, when the estimated coefficient is nonzero, but the true coefficient is zero or is nonzero but has the opposite sign as the estimated coefficient.
- (4) False negative, when the estimated coefficient is zero, but the true coefficient is nonzero.

Based on the definitions above, we calculate the following metrics:

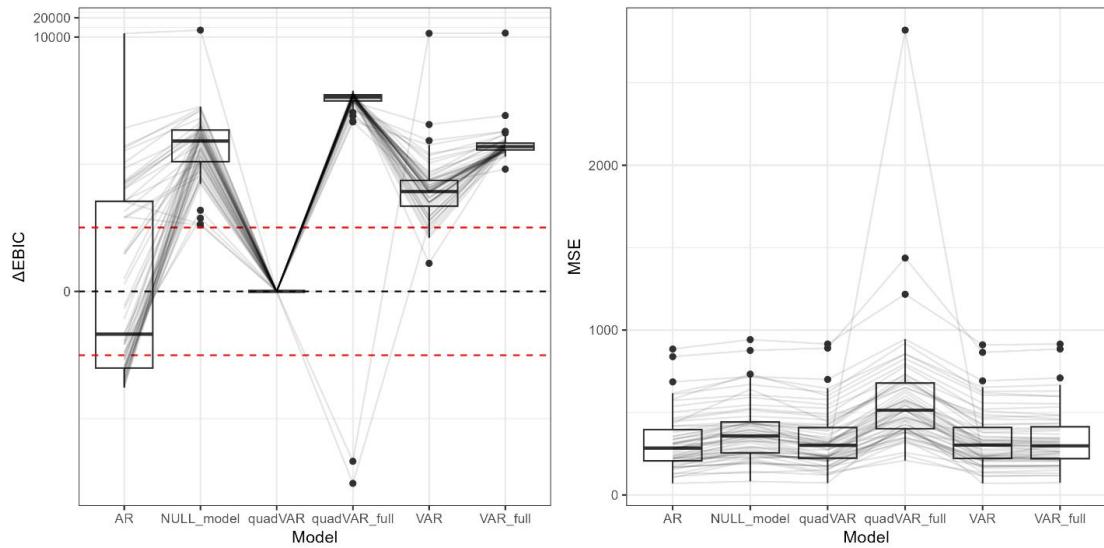
- (1) Accuracy = (true positive + true negative)/(true positive + true negative + false positive + false negative).
- (2) Precision = true positive/(true positive + false positive).
- (3) Sensitivity = true negative/(true negative + false positive).
- (4) Specificity = true positive/(true positive + false negative).

The correlation is defined as the Pearson correlation between the estimated coefficients and the true coefficients.

F4. Results from the Second Empirical Dataset

The results from the dataset from Kuppens et al. (2010) were similar to the results of the dataset Rowland and Wenzel (2020; reported in the main text). The EBIC and cross-validation results for the dataset from Kuppens et al. (2010) are shown in Figure F2.

Figure F2. EBIC and cross-validation MSE for participants in Kuppens et al. (2010). Gray lines represent individual participants.



F5. quadVAR Model Estimations for Selected Participants

Table F1. quadVAR model estimations for the 6th participant in Rowland and Wenzel (2020).

Model	Effect	Estimate
happy	satisfied	0.3551
excited	satisfied	0.4287
relaxed	relaxed	0.4105
satisfied	satisfied	0.5321
angry	excited	-0.0783
angry	satisfied	0.0322
angry	angry	-0.1278
angry	excited:angry	0.0231
anxious	sad	-0.3702
anxious	sad:sad	0.0375
depressed	satisfied	-0.1500
depressed	angry	-1.4642
depressed	depressed	0.2909
depressed	satisfied:angry	0.0391

Table F2. quadVAR model estimations for the 8th participant in Rowland and Wenzel (2020).

Model	Effect	Estimate
happy	satisfied	-0.1669
happy	satisfied:satisfied	0.0056
satisfied	satisfied	0.2585
angry	sad	0.3214
anxious	angry	0.2355
anxious	anxious	0.2537
depressed	depressed	0.4029
sad	depressed	0.2574
sad	sad	0.4248

Appendix G. Supplementary Materials for Chapter 9

G1. Alternative Models

In the main text, we argued that as long as several key dynamic features are preserved, we can alter the function form of the dynamic models and obtain similar features in the dynamic system. Here, we show two examples as alternatives to the original models analyzed in the main text.

The first example is an alternative version of the panic disorder model, in which we use two S-shaped curves for the dynamic equations of A and PT . We chose a different form of the S-shaped curves compared to the original model to ensure that there is always a stable equilibrium point at $A = PT = 0$. The model is specified as follows (the changed parts are marked with gray shading),

$$\frac{dA}{dt} = r_A \left(\frac{PT^{n_A}}{PT^{n_A} + (k_A \exp(\beta_A H))^{n_A}} \right), \quad (1)$$

$$\frac{dPT}{dt} = r_{PT} \left(\frac{A^{n_{PT}}}{A^{n_{PT}} + 1} - PT \right), \quad (2)$$

$$\frac{dH}{dt} = r_H \left(\frac{1}{1 + \exp(-k_H(A - h_{A,H}))} - 0.5H \right). \quad (3)$$

The default parameter set is summarized in Table G1.

Table G1. The default parameter values for the alternative version of the panic disorder model.

Parameter	Value
r_A	1
n_A	6
k_A	0.02
β_A	5
r_{PT}	1
n_{PT}	4
r_H	0.05
k_H	20
$h_{A,H}$	0.4

The simulated output of this model is shown in Figure G1, and the phase plane analyses and bifurcation analysis results are shown in Figure G2. As shown in those results, although we used double S-curved functions for A and PT , because the critical dynamic features remain similar, the model still produces reasonable outcomes for the phenomenon of panic disorder.

Figure G1. Simulated output of the alternative version of the panic disorder model.

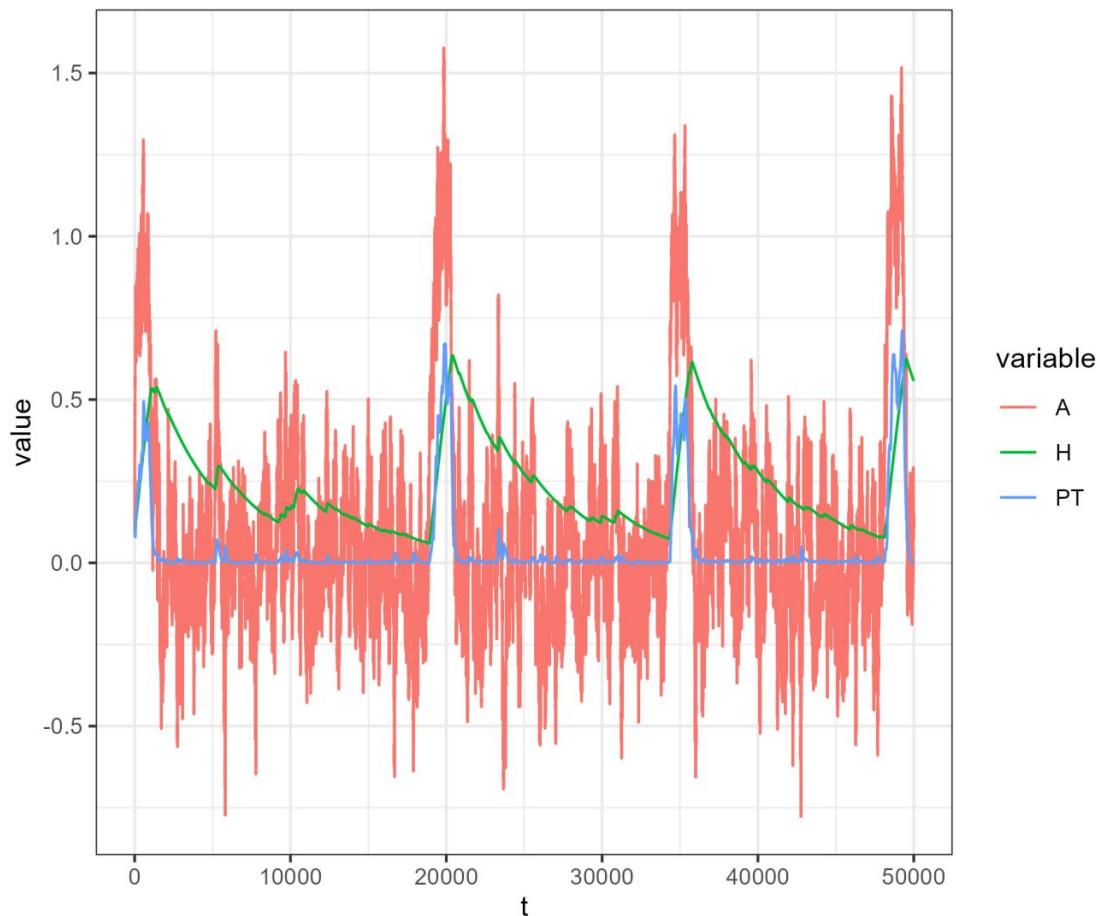
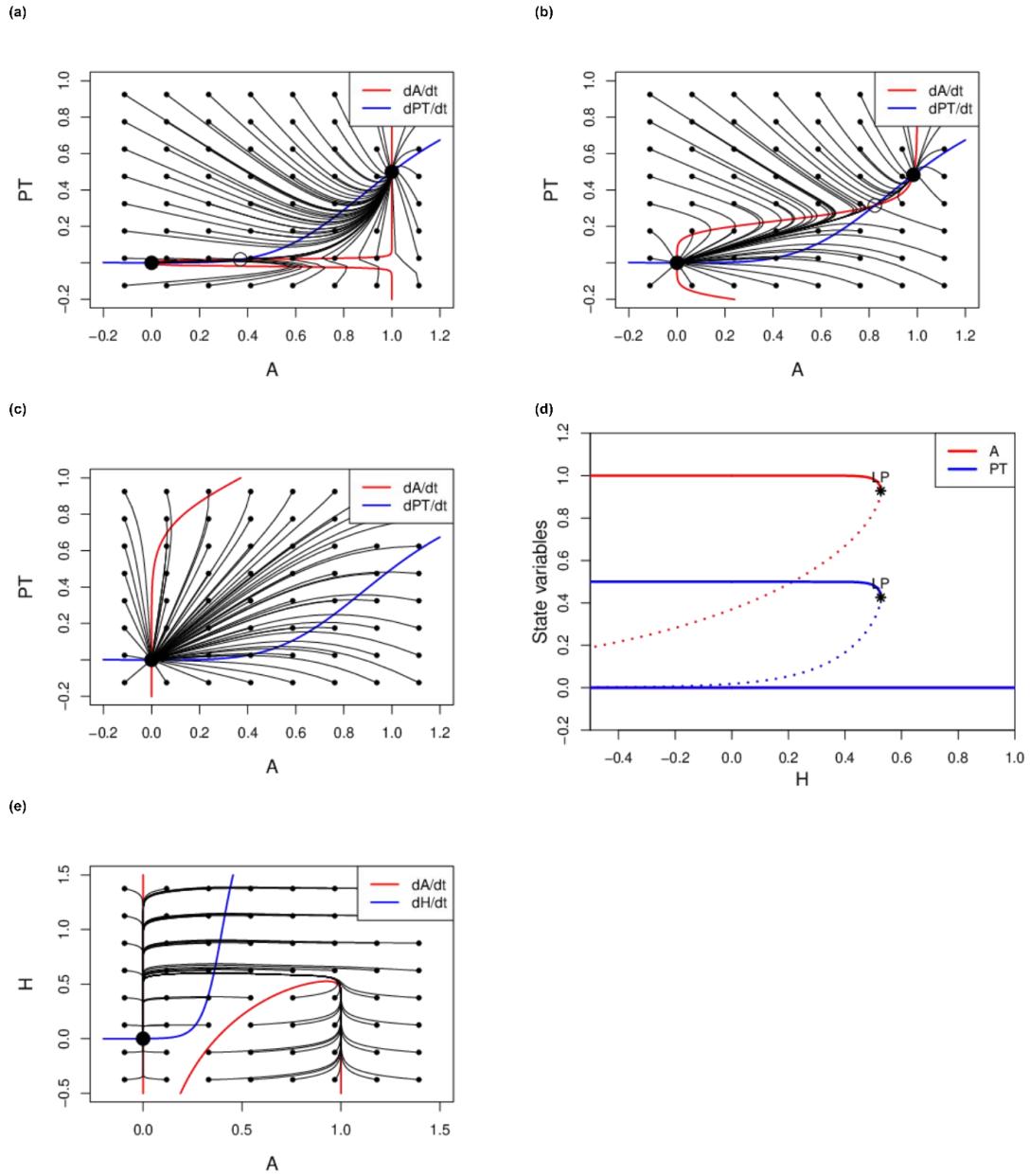


Figure G2. Phase plane analysis and bifurcation analysis results for the alternative model of panic disorder. (a)-(c) Phase plane analyses results with H as a parameter, (a) $H = 0$; (b) $H = 0.5$; (c) $H = 0.8$. (d) The one-parameter bifurcation analysis result with H as the parameter. (e) The phase plane analysis result with H as a variable.



The second example is an alternative version of the suicidal ideation model, in which we changed the form of the dynamic equation for A , making the first term of the right-hand-side from quadratic to linear. The modified model is specified as follows (the changed parts are marked with gray shading),

$$\frac{dA}{dt} = -b_2(A - K) + a_2S - d_2T, \quad (4)$$

$$\frac{dU}{dt} = -c_3U + b_3A, \quad (5)$$

$$\frac{dT}{dt} = -d_4T + \frac{1}{1 + \exp[-c_{41}(U - c_{42})]}. \quad (6)$$

We used the same parameters as the original model, only except that we use $a_2 = 2$ instead of $a_2 = 1.5$ to accommodate the change of the function's form. The simulated output of this model is shown in Figure G3, and the phase plane analyses and bifurcation analyses results are shown in Figure G4. Again, although the function form of A changed from quadratic to linear, because the critical dynamic features remain similar, the model still produces reasonable outcomes for the phenomenon of suicidal ideations.

Figure G3. Simulated output of the alternative version of the suicidal ideation model.

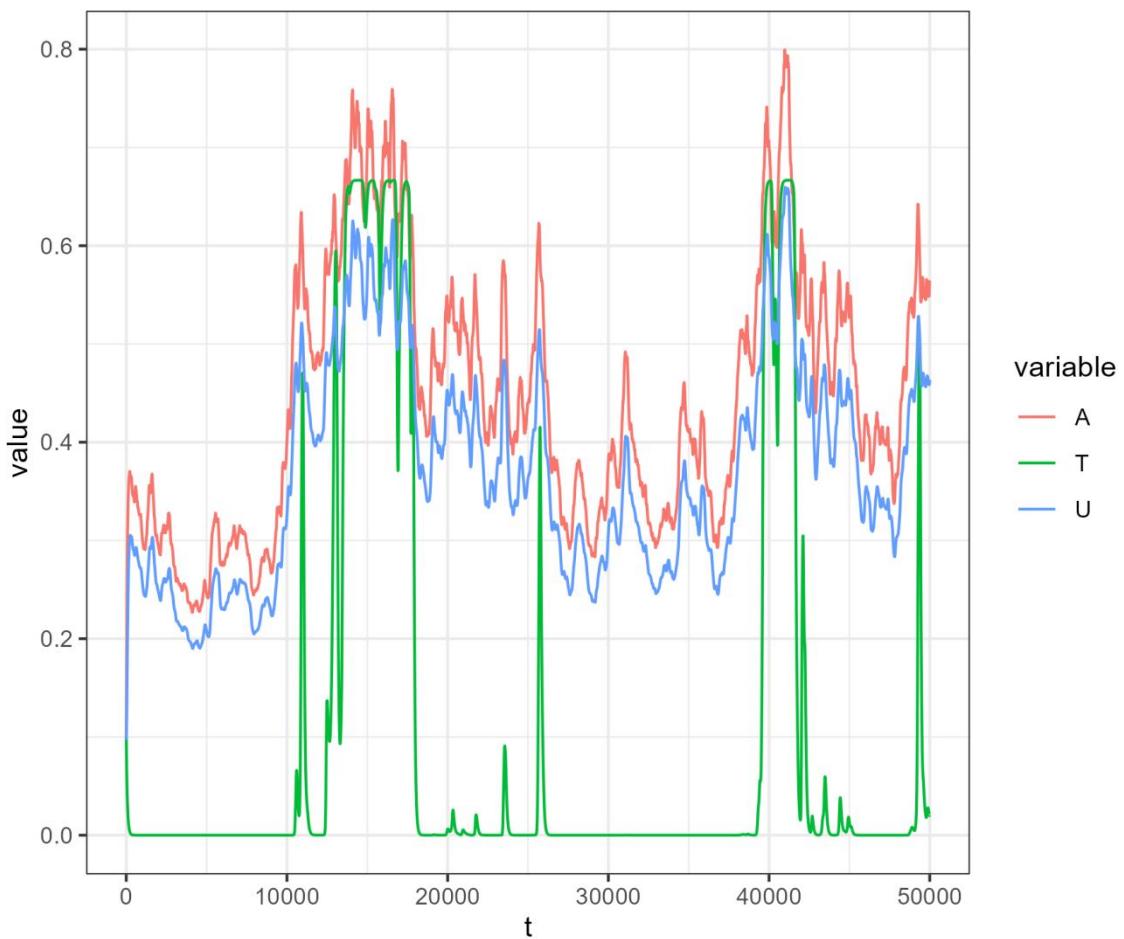
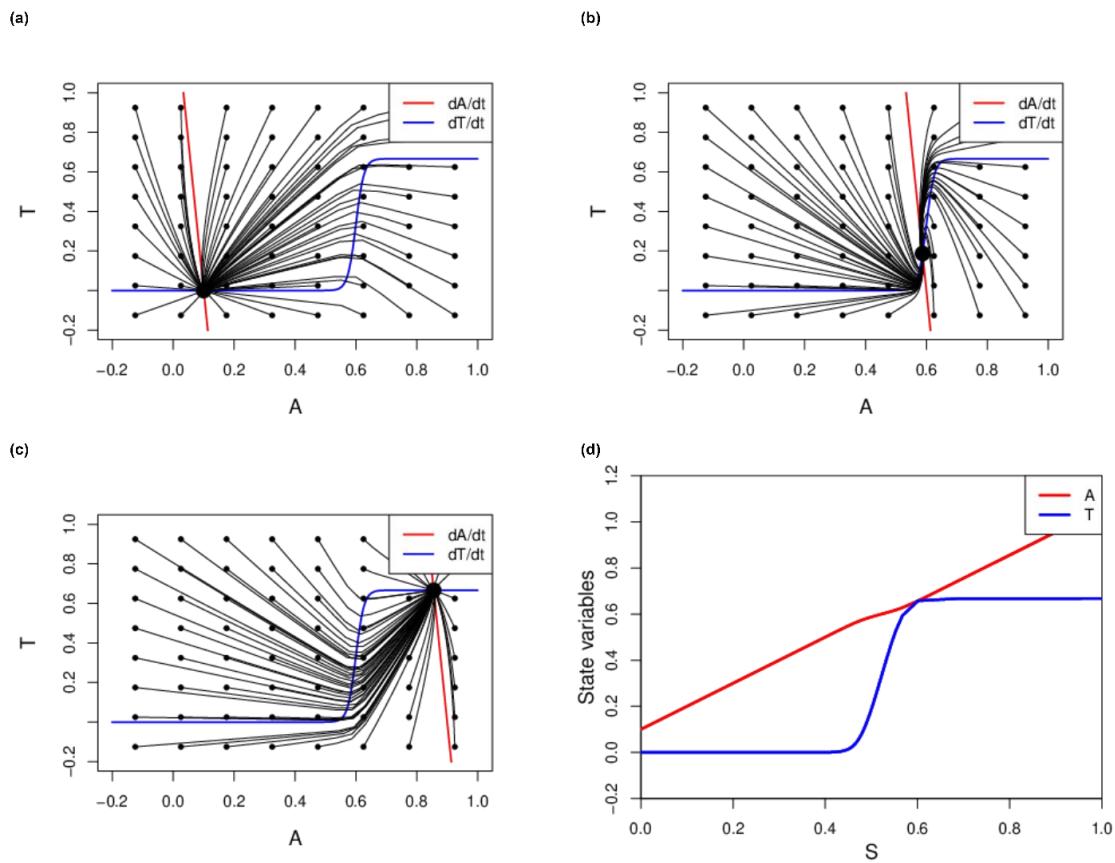


Figure G4. Phase plane analyses and bifurcation analysis results for the alternative model of panic disorder. (a)-(c) Phase plane analyses results for different S values (a) $S = 0$; (b) $S = 0.5$; (c) $S = 0.8$. (d) The one-parameter bifurcation analysis result with S as the parameter.



References

- Abraham, R., & Shaw, C. D. (1992). *Dynamics: The geometry of behavior* (2nd ed.). Addison-Wesley.
- Aderka, I. M., Kauffmann, A., Shalom, J. G., Beard, C., & Björgvinsson, T. (2021). Using machine-learning to predict sudden gains in treatment for major depressive disorder. *Behaviour Research and Therapy*, 144, 103929. <https://doi.org/10.1016/j.brat.2021.103929>
- Aderka, I. M., Nickerson, A., Bøe, H. J., & Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 80(1), 93–101. <https://doi.org/10.1037/a0026455>
- Aderka, I. M., & Shalom, J. G. (2021). A revised theory of sudden gains in psychological treatments. *Behaviour Research and Therapy*, 139, 103830.
<https://doi.org/10.1016/j.brat.2021.103830>
- Adolf, J. K., Voelkle, M. C., Brose, A., & Schmiedek, F. (2017). Capturing context-related change in emotional dynamics via fixed moderated time series analysis. *Multivariate Behavioral Research*, 52(4), 499–531. <https://doi.org/10.1080/00273171.2017.1321978>
- Akiskal, H. S., & Pinto, O. (1999). The evolving bipolar spectrum. *Psychiatric Clinics of North America*, 22(3), 517–534. [https://doi.org/10.1016/S0193-953X\(05\)70093-9](https://doi.org/10.1016/S0193-953X(05)70093-9)
- Alós-Ferrer, C., Hügelschäfer, S., & Li, J. (2016). Inertia and decision making. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00169>
- Alpert, M., & Rosen, A. (1990). A semantic analysis of the various ways that the terms “affect,” “emotion,” and “mood” are used. *Journal of Communication Disorders*, 23(4), 237–246.
[https://doi.org/10.1016/0021-9924\(90\)90002-G](https://doi.org/10.1016/0021-9924(90)90002-G)
- American Psychiatric Association. (2022a). Bipolar and related disorders. In *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). American Psychiatric Association Publishing.
- https://doi.org/10.1176/appi.books.9780890425787.x03_Bipolar_and_Related_Disorders
- American Psychiatric Association. (2022b). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). American Psychiatric Association Publishing.
<https://doi.org/10.1176/appi.books.9780890425787>

- American Psychiatric Association. (2022c). Personality disorders. In *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). American Psychiatric Association Publishing. https://doi.org/10.1176/appi.books.9780890425787.x18_Personality_Disorders
- Ao, P. (2004). Potential in stochastic differential equations: Novel construction. *Journal of Physics A: Mathematical and General*, 37(3), L25–L30. <https://doi.org/10.1088/0305-4470/37/3/L01>
- Arntz, A. (2012). Imagery rescripting as a therapeutic technique: Review of clinical trials, basic studies, and research agenda. *Journal of Experimental Psychopathology*, 3(2), 189–208. <https://doi.org/10.5127/jep.024211>
- Ashwin, P., Perryman, C., & Wieczorek, S. (2017). Parameter shifts for nonautonomous systems in low dimension: Bifurcation- and rate-induced tipping. *Nonlinearity*, 30(6), 2185–2210. <https://doi.org/10.1088/1361-6544/aa675b>
- Ashwin, P., Wieczorek, S., Vitolo, R., & Cox, P. (2012). Tipping points in open systems: Bifurcation, noise-induced and rate-dependent examples in the climate system. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370, 1166–1184. <https://doi.org/10.1098/rsta.2011.0306>
- Bandi, F. M., & Moloche, G. (2018). On the functional estimation of multivariate diffusion processes. *Econometric Theory*, 34(4), 896–946. <https://doi.org/10.1017/S0266466617000305>
- Barton, S. (1994). Chaos, self-organization, and psychology. *American Psychologist*, 49(1), 5–14. <https://doi.org/10.1037/0003-066X.49.1.5>
- Bengtsson, H. (2021). A unifying framework for parallel and distributed processing in R using futures. *The R Journal*, 13(2), 208–227. <https://doi.org/10.32614/RJ-2021-048>
- Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191, 192–213. <https://doi.org/10.1016/j.ins.2011.12.028>
- Bernardo, A. B. I., Mendoza, N. B., Simon, P. D., Cunanan, A. L. P., Dizon, J. I. W. T., Tarroja, M. C. H., Ma. Balajadia-Alcala, A., & Saplala, J. E. G. (2022). Coronavirus Pandemic Anxiety Scale (CPAS-11): Development and initial validation. *Current Psychology*, 41(8), 5703–5711. <https://doi.org/10.1007/s12144-020-01193-2>
- Bertram, R., & Rubin, J. E. (2017). Multi-timescale systems and fast-slow analysis. *Mathematical Biosciences*, 287, 105–121. <https://doi.org/10.1016/j.mbs.2016.07.003>

-
- Bien, J., Taylor, J., & Tibshirani, R. (2013). A lasso for hierarchical interactions. *The Annals of Statistics*, 41(3), 1111–1141. <https://doi.org/10.1214/13-AOS1096>
- Binder, K., & Luijten, E. (2001). Monte Carlo tests of renormalization-group predictions for critical phenomena in Ising models. *Physics Reports*, 344(4–6), 179–253. [https://doi.org/10.1016/S0370-1573\(00\)00127-7](https://doi.org/10.1016/S0370-1573(00)00127-7)
- Blome, C., & Augustin, M. (2015). Measuring change in quality of life: Bias in prospective and retrospective evaluation. *Value in Health*, 18(1), 110–115. <https://doi.org/10.1016/j.jval.2014.10.007>
- Boerlijst, M. C., Oudman, T., & Roos, A. M. de. (2013). Catastrophic collapse can occur without early warning: Examples of silent catastrophes in structured ecological models. *PLOS ONE*, 8(4), e62033. <https://doi.org/10.1371/journal.pone.0062033>
- Boettiger, C., & Batt, R. (2020). Bifurcation or state tipping: Assessing transition type in a model trophic cascade. *Journal of Mathematical Biology*, 80(1), 143–155. <https://doi.org/10.1007/s00285-019-01358-z>
- Boettiger, C., & Hastings, A. (2012). Early warning signals and the prosecutor's fallacy. *Proceedings of the Royal Society B: Biological Sciences*, 279(1748), 4734–4739. <https://doi.org/10.1098/rspb.2012.2085>
- Boettiger, C., Ross, N., & Hastings, A. (2013). Early warning signals: The charted and uncharted territories. *Theoretical Ecology*, 6(3), 255–264. <https://doi.org/10.1007/s12080-013-0192-6>
- Bogartz, R. S. (1994). The future of dynamic systems models in developmental psychology in the light of the past. *Journal of Experimental Child Psychology*, 58(2), 289–319. <https://doi.org/10.1006/jecp.1994.1036>
- Boker, S. M., & Moulder, R. G. (2023). Dynamical systems and differential equation models of change. In *APA handbook of research methods in psychology: Data analysis and research publication*, Vol. 3, 2nd ed (pp. 337–350). American Psychological Association. <https://doi.org/10.1037/0000320-015>
- Boot, J., de Ron, J., Haslbeck, J. M. B., & Epskamp, S. (2023). *Correcting for selection bias after conditioning on a sum-score in the Ising model*. PsyArXiv. <https://doi.org/10.31219/osf.io/xq8ur>
- Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, 323(5916), 892–895. <https://doi.org/10.1126/science.1165821>

- Borgstede, M., & Eggert, F. (2023). Squaring the circle: From latent variables to theory-based measurement. *Theory & Psychology*, 33(1), 118–137.
<https://doi.org/10.1177/09593543221127985>
- Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, 71(3), 425–440.
<https://doi.org/10.1007/s11336-006-1447-6>
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, 16(1), 5–13.
<https://doi.org/10.1002/wps.20375>
- Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, 1(1), 58.
<https://doi.org/10.1038/s43586-021-00055-w>
- Borsboom, D., & Haslbeck, J. (2024). Integrating intra- and interindividual phenomena in psychological theories. *Multivariate Behavioral Research*, 59(6), 1290–1309.
<https://doi.org/10.1080/00273171.2024.2336178>
- Borsboom, D., Haslbeck, J. M. B., & Robinaugh, D. J. (2022). Systems-based approaches to mental disorders are the only game in town. *World Psychiatry*, 21(3), 420–422.
<https://doi.org/10.1002/wps.21004>
- Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory construction methodology: A practical framework for building theories in psychology. *Perspectives on Psychological Science*, 16(4), 756–766.
<https://doi.org/10.1177/1745691620969647>
- Bos, E. H., & De Jonge, P. (2014). “Critical slowing down in depression” is a great idea that still needs empirical proof. *Proceedings of the National Academy of Sciences*, 111(10), E878.
<https://doi.org/10.1073/pnas.1323672111>
- Bos, F. M., Schreuder, M. J., George, S. V., Doornbos, B., Bruggeman, R., van der Krieke, L., Haarman, B. C. M., Wichers, M., & Snippe, E. (2022). Anticipating manic and depressive transitions in patients with bipolar disorder using early warning signals. *International Journal of Bipolar Disorders*, 10(1), 12. <https://doi.org/10.1186/s40345-022-00258-4>
- Bossenbroek, R., Gmeliin, O., Olthof, M., Knot-Dickscheit, J., Hasselman, F., Poelen, E., & Lichtwarck-Aschoff, A. (2025). *Developmental trajectories of young people with a history in*

residential care: A qualitative investigation of lifelines from a complex systems perspective [Manuscript submitted for publication].

- Boulougouris, J. C., Marks, I. M., & Marset, P. (1971). Superiority of flooding (implosion) to desensitisation for reducing pathological fear. *Behaviour Research and Therapy*, 9(1), 7–16. [https://doi.org/10.1016/0005-7967\(71\)90030-1](https://doi.org/10.1016/0005-7967(71)90030-1)
- Breheny, P., & Huang, J. (2011). Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection. *Annals of Applied Statistics*, 5(1), 232–253. <https://doi.org/10.1214/10-AOAS388>
- Brillinger, D. R. (2007). Learning a potential function from a trajectory. *IEEE Signal Processing Letters*, 14(11), 867–870. <https://doi.org/10.1109/LSP.2007.900032>
- Bringmann, L. F. (2021). Person-specific networks in psychopathology: Past, present, and future. *Current Opinion in Psychology*, 41, 59–64. <https://doi.org/10.1016/j.copsyc.2021.03.004>
- Bringmann, L. F., Albers, C., Bockting, C., Borsboom, D., Ceulemans, E., Cramer, A., Epskamp, S., Eronen, M. I., Hamaker, E., Kuppens, P., Lutz, W., McNally, R. J., Molenaar, P., Tio, P., Voelkle, M. C., & Wichers, M. (2022). Psychopathological networks: Theory, methods and practice. *Behaviour Research and Therapy*, 149, 104011. <https://doi.org/10.1016/j.brat.2021.104011>
- Bringmann, L. F., Ariens, S., Ernst, A. F., Snippe, E., & Ceulemans, E. (2024). Changing networks: Moderated idiographic psychological networks. *Advances in Psychology*, 2, e658296. <https://doi.org/10.56296/aip00014>
- Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., Wigman, J. T. W., & Snippe, E. (2019). What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology*, 128(8), 892–903. <https://doi.org/10.1037/abn0000446>
- Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological Methods*, 22(3), 409–425. <https://doi.org/10.1037/met0000085>
- Bringmann, L. F., Lemmens, L. H. J. M., Huibers, M. J. H., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological Medicine*, 45(4), 747–757. <https://doi.org/10.1017/S0033291714001809>
- Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into

- clinical longitudinal data. *PLOS ONE*, 8(4), e60188.
<https://doi.org/10.1371/journal.pone.0060188>
- Bringmann, L., Helmich, M., Eronen, M., & Voelkle, M. (2023). Complex systems approaches to psychopathology. In R. F. Krueger & P. H. Blaney (Eds.), *Oxford Textbook of Psychopathology* (4 edn, pp. 103–122).
<https://doi.org/10.1093/med-psych/9780197542521.003.0005>
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.
- Brouste, A., Fukasawa, M., Hino, H., Iacus, S. M., Kamatani, K., Koike, Y., Masuda, H., Nomura, R., Ogihara, T., Shimuzu, Y., Uchida, M., & Yoshida, N. (2014). The YUIMA project: A computational framework for simulation and inference of stochastic differential equations. *Journal of Statistical Software*, 57(4), 1–51. <https://doi.org/10.18637/jss.v057.i04>
- Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15), 3932–3937. <https://doi.org/10.1073/pnas.1517384113>
- Brusco, M. J., Davis-Stober, C. P., & Steinley, D. (2021). Ising formulations of some graph-theoretic problems in psychological research: Models and methods. *Journal of Mathematical Psychology*, 102, 102536. <https://doi.org/10.1016/j.jmp.2021.102536>
- Brusco, M. J., Steinley, D., & Watts, A. L. (2023). A comparison of logistic regression methods for Ising model estimation. *Behavior Research Methods*, 55(7), 3566–3584.
<https://doi.org/10.3758/s13428-022-01976-4>
- Bulteel, K., Mestdagh, M., Tuerlinckx, F., & Ceulemans, E. (2018). VAR(1) based models do not always outpredict AR(1) models in typical psychological applications. *Psychological Methods*, 23(4), 740–756. <https://doi.org/10.1037/met0000178>
- Burger, J., Isvoranu, A.-M., Lunansky, G., Haslbeck, J. M. B., Epskamp, S., Hoekstra, R. H. A., Fried, E. I., Borsboom, D., & Blanken, T. F. (2023). Reporting standards for psychological network analyses in cross-sectional data. *Psychological Methods*, 28(4), 806–824.
<https://doi.org/10.1037/met0000471>
- Burger, J., van der Veen, D. C., Robinaugh, D. J., Quax, R., Riese, H., Schoevers, R. A., & Epskamp, S. (2020). Bridging the gap between complexity science and clinical practice by formalizing idiographic theories: A computational model of functional analysis. *BMC Medicine*, 18(1), 1–18. <https://doi.org/10.1186/s12916-020-01558-1>

-
- Burnham, K. P., Anderson, D. R., & Burnham, K. P. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed). Springer.
- Cabrieto, J., Tuerlinckx, F., Kuppens, P., Grassmann, M., & Ceulemans, E. (2017). Detecting correlation changes in multivariate time series: A comparison of four non-parametric change point detection methods. *Behavior Research Methods*, 49(3), 988–1005.
<https://doi.org/10.3758/s13428-016-0754-9>
- Cartwright, N. D. (2009). What is this thing called “efficacy”? In C. Mantzavinos (Ed.), *Philosophy of the social sciences: Philosophical theory and scientific practice* (pp. 185–206). Cambridge University Press.
- Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., & Borges, B. (2024). *shiny: Web application framework for R* [Manual].
<https://CRAN.R-project.org/package=shiny>
- Chang, W., Luraschi, J., & Mastny, T. (2020). *profvis: Interactive visualizations for profiling R code* [Manual]. <https://CRAN.R-project.org/package=profvis>
- Chen, J., & Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, 95(3), 759–771. <https://doi.org/10.1093/biomet/asn034>
- Clarivate. (2025). *Web of science*. <https://www.webofscience.com>
- Clark, D. M. (1986). A cognitive approach to panic. *Behaviour Research and Therapy*, 24(4), 461–470. [https://doi.org/10.1016/0005-7967\(86\)90011-2](https://doi.org/10.1016/0005-7967(86)90011-2)
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression correlation analysis for the behavioral sciences* (Third edition). Routledge Taylor & Francis Group.
- Cramer, A. O. J., Borkulo, C. D. van, Giltay, E. J., Maas, H. L. J. van der, Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *PLOS ONE*, 11(12), e0167490. <https://doi.org/10.1371/journal.pone.0167490>
- Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, 46(1), 5–27. <https://doi.org/10.1016/j.brat.2007.10.003>
- Cui, J., Hasselman, F., & Lichtwarck-Aschoff, A. (2023). Unlocking nonlinear dynamics and multistability from intensive longitudinal data: A novel method. *Psychological Methods*.
<https://doi.org/10.1037/met0000623>

- Cui, J., Hasselman, F., Olthof, M., & Lichtwarck-Aschoff, A. (2025). Understanding types of transitions in clinical change: An introduction from the complex dynamic systems perspective. *Journal of Psychopathology and Clinical Science*, 134(4), 469–482.
<https://doi.org/10.1037/abn0000991>
- Cui, J., Lichtwarck-Aschoff, A., & Hasselman, F. (2023). *Comments on “Climbing Escher’s stairs: A way to approximate stability landscapes in multidimensional systems”* (arXiv:2312.09690). arXiv. <https://doi.org/10.48550/arXiv.2312.09690>
- Cui, J., Lichtwarck-Aschoff, A., Olthof, M., Li, T., & Hasselman, F. (2023). From metaphor to computation: Constructing the potential landscape for multivariate psychological formal models. *Multivariate Behavioral Research*, 58(4), 743–761.
<https://doi.org/10.1080/00273171.2022.2119927>
- Cui, J., Olthof, M., Hasselman, F., & Lichtwarck-Aschoff, A. (2025). Examining the research methods of early warning signals in clinical psychology through a theoretical lens. *BMC Psychiatry*, 25, 261. <https://doi.org/10.1186/s12888-025-06688-5>
- Cui, J., Olthof, M., Lichtwarck-Aschoff, A., Li, T., & Hasselman, F. (2021). *simlandr: Simulation-based landscape construction for dynamical systems*. PsyArXiv.
<https://doi.org/10.31234/osf.io/pzva3>
- Cumming, J., & Anderson, G. M. (2013). Guided imagery. In M. D. Gellman & J. R. Turner (Eds.), *Encyclopedia of behavioral medicine* (pp. 881–883). Springer. https://doi.org/10.1007/978-1-4419-1005-9_1341
- Curtiss, J. E., Fulford, D., Hofmann, S. G., & Gershon, A. (2019). Network dynamics of positive and negative affect in bipolar disorder. *Journal of Affective Disorders*, 249, 270–277.
<https://doi.org/10.1016/j.jad.2019.02.017>
- Curtiss, J. E., Mischoulon, D., Fisher, L. B., Cusin, C., Fedor, S., Picard, R. W., & Pedrelli, P. (2023). Rising early warning signals in affect associated with future changes in depression: A dynamical systems approach. *Psychological Medicine*, 53(7), 3124–3132.
<https://doi.org/10.1017/S0033291721005183>
- Dablander, F., Pichler, A., Cika, A., & Bacilieri, A. (2023). Anticipating critical transitions in psychological systems using early warning signals: Theoretical and practical considerations. *Psychological Methods*, 28(4), 765–790. <https://doi.org/10.1037/met0000450>

-
- Dablander, F., Ryan, O., & Haslbeck, J. M. B. (2020). Choosing between AR(1) and VAR(1) models in typical psychological applications. *PLOS ONE*, 15(10), e0240730. <https://doi.org/10.1371/journal.pone.0240730>
- Dahiya, D., & Cameron, M. (2018). Ordered line integral methods for computing the quasi-potential. *Journal of Scientific Computing*, 75(3), 1351–1384. <https://doi.org/10.1007/s10915-017-0590-9>
- Dakos, V., Carpenter, S. R., Brock, W. A., Ellison, A. M., Guttal, V., Ives, A. R., Kéfi, S., Livina, V., Seekell, D. A., Nes, E. H. van, & Scheffer, M. (2012). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLOS ONE*, 7(7), e41010. <https://doi.org/10.1371/journal.pone.0041010>
- Dakos, V., Carpenter, S. R., van Nes, E. H., & Scheffer, M. (2015). Resilience indicators: Prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1659), 20130263. <https://doi.org/10.1098/rstb.2013.0263>
- Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V., & Held, H. (2008). Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences*, 105(38), 14308–14312. <https://doi.org/10.1073/pnas.0802430105>
- Dakos, V., van Nes, E. H., & Scheffer, M. (2013). Flickering as an early warning signal. *Theoretical Ecology*, 6(3), 309–317. <https://doi.org/10.1007/s12080-013-0186-4>
- Dalege, J., Borsboom, D., van Harreveld, F., & van der Maas, H. L. J. (2018). The attitudinal entropy (AE) framework as a general theory of individual attitudes. *Psychological Inquiry*, 29(4), 175–193. <https://doi.org/10.1080/1047840X.2018.1537246>
- Dalege, J., Borsboom, D., van Harreveld, F., Waldorp, L. J., & van der Maas, H. L. J. (2017). Network structure explains the impact of attitudes on voting decisions. *Scientific Reports*, 7(1), Article 1. <https://doi.org/10.1038/s41598-017-05048-y>
- Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap Methods and Their Applications*. Cambridge University Press. <http://statwww.epfl.ch/davison/BMA/>
- de Boer, R. J. (2024). *Grind.R: Phase plane analysis and parameter estimation in R*. <https://tbb.bio.uu.nl/rdb/grindR.html>
- De Haan-Rietdijk, S., Gottman, J. M., Bergeman, C. S., & Hamaker, E. L. (2016). Get over it! A multilevel threshold autoregressive model for state-dependent affect regulation. *Psychometrika*, 81(1), 217–241. <https://doi.org/10.1007/s11336-014-9417-x>

- de Jong, R., Hofs, A., Lommen, M. J. J., van Hout, W. J. P. J., De Jong, P. J., & Nauta, M. H. (2023). Treating specific phobia in youth: A randomized controlled microtrial comparing gradual exposure in large steps to exposure in small steps. *Journal of Anxiety Disorders*, 96, 102712. <https://doi.org/10.1016/j.janxdis.2023.102712>
- De Moor, W. (1970). Systematic desensitization versus prolonged high intensity stimulation (flooding). *Journal of Behavior Therapy and Experimental Psychiatry*, 1(1), 45–52. [https://doi.org/10.1016/0005-7916\(70\)90020-0](https://doi.org/10.1016/0005-7916(70)90020-0)
- de Roos, A. M. (2025). *deBif: Bifurcation analysis of ordinary differential equation systems* [Manual]. <https://CRAN.R-project.org/package=deBif>
- de Ron, J., Fried, E. I., & Epskamp, S. (2021). Psychological networks in clinical populations: Investigating the consequences of Berkson's bias. *Psychological Medicine*, 51(1), 168–176. <https://doi.org/10.1017/S0033291719003209>
- Dedrick, R. F., Ferron, J. M., Hess, M. R., Hogarty, K. Y., Kromrey, J. D., Lang, T. R., Niles, J. D., & Lee, R. S. (2009). Multilevel modeling: A review of methodological issues and applications. *Review of Educational Research*, 79(1), 69–102. <https://doi.org/10.3102/0034654308325581>
- Dejonckheere, E., Demeyer, F., Geusens, B., Piot, M., Tuerlinckx, F., Verdonck, S., & Mestdagh, M. (2022). Assessing the reliability of single-item momentary affective measurements in experience sampling. *Psychological Assessment*, 34(12), 1138–1154. <https://doi.org/10.1037/pas0001178>
- Delignières, D., Fortes, M., & Ninot, G. (2004). The fractal dynamics of self-esteem and physical self. *Nonlinear Dynamics, Psychology, and Life Sciences*, 8(4), 479–510.
- Di Vita, A. (2022). *Non-equilibrium Thermodynamics* (Vol. 1007). Springer International Publishing. <https://doi.org/10.1007/978-3-031-12221-7>
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1), 269–271.
- Ditlevsen, P. D., & Johnsen, S. J. (2010). Tipping points: Early warning and wishful thinking. *Geophysical Research Letters*, 37, L19703. <https://doi.org/10.1029/2010GL044486>
- Dizon, J. I. W. T., Mendoza, N. B., & Nalipay, Ma. J. N. (2023). Anxiety and well-being amidst the COVID-19 outbreak and the moderating role of locus-of-hope: Evidence from a large sample in the Philippines. *Journal of Pacific Rim Psychology*, 17, 18344909231156532. <https://doi.org/10.1177/18344909231156532>

-
- Duong, T. (2021). *ks: Kernel smoothing* [Manual]. <https://CRAN.R-project.org/package=ks>
- E, W., & Vanden-Eijnden, E. (2010). Transition-path theory and path-finding algorithms for the study of rare events. *Annual Review of Physical Chemistry*, 61(1), 391–420. <https://doi.org/10.1146/annurev.physchem.040808.090412>
- Eckes, T., & Nestler, S. (2023). Do I like me now? An analysis of everyday sudden gains and sudden losses in self-esteem and nervousness. *Clinical Psychological Science*, 216770262311656. <https://doi.org/10.1177/21677026231165677>
- Eddelbuettel, D. (2020). *RcppZiggurat: “Rcpp” integration of different ‘Ziggurat’ normal RNG Implementations*. <https://CRAN.R-project.org/package=RcppZiggurat>
- Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8), 1–18. <https://doi.org/10.18637/jss.v040.i08>
- Eddelbuettel, D., Lucas, A., Tuszynski, J., Bengtsson, H., Urbanek, S., Frasca, M., Lewis, B., Stokely, M., Muehleisen, H., Murdoch, D., Hester, J., Wu, W., Kou, Q., Onkelinx, T., Lang, M., Simko, V., Hornik, K., Neal, R., Bell, K., ... Chang, W. (2021). *digest: Create compact hash digests of R objects* [Manual]. <https://CRAN.R-project.org/package=digest>
- Edelstein-Keshet, L. (2005). Limit cycles, oscillations, and excitable systems. In *Mathematical models in biology* (pp. 311–380). Society for Industrial and Applied Mathematics. <https://doi.org/10.1137/1.9780898719147.ch8>
- Elson, M., Hussey, I., Alsalti, T., & Arslan, R. C. (2023). Psychological measures aren't toothbrushes. *Communications Psychology*, 1(1), Article 1. <https://doi.org/10.1038/s44271-023-00026-9>
- Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, 50(1), 195–212. <https://doi.org/10.3758/s13428-017-0862-1>
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48(1), Article 1. <https://doi.org/10.18637/jss.v048.i04>
- Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, 23(4), 617–634. <https://doi.org/10.1037/met0000167>

- Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, 53(4), 453–480. <https://doi.org/10.1080/00273171.2018.1454823>
- Eronen, M. I. (2019). The levels problem in psychopathology. *Psychological Medicine*, 51(6), 927–933. <https://doi.org/10.1017/S0033291719002514>
- Eronen, M. I., & Bringmann, L. F. (2021). The theory crisis in psychology: How to move forward. *Perspectives on Psychological Science*, 16(4), 779–788. <https://doi.org/10.1177/1745691620970586>
- Evers, K., Borsboom, D., Fried, E. I., Hasselman, F., & Waldorp, L. (2024). Early warning signals of complex critical transitions in deterministic dynamics. *Nonlinear Dynamics*. <https://doi.org/10.1007/s11071-024-10023-0>
- Failenschmid, J. I., Vogelsmeier, L. V. D. E., Mulder, J., & Jongerling, J. (2025). Modelling non-linear psychological processes: Reviewing and evaluating non-parametric approaches and their applicability to intensive longitudinal data. *British Journal of Mathematical and Statistical Psychology*. <https://doi.org/10.1111/bmsp.12397>
- Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, 115(27), E6106–E6115. <https://doi.org/10.1073/pnas.1711978115>
- Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, 99(1), 20–35. <https://doi.org/10.1037/0033-2909.99.1.20>
- Foa, E. B., & McLean, C. P. (2016). The efficacy of exposure therapy for anxiety-related disorders and its underlying mechanisms: The case of OCD and PTSD. *Annual Review of Clinical Psychology*, 12, 1–28. <https://doi.org/10.1146/annurev-clinpsy-021815-093533>
- Forgoston, E., & Moore, R. O. (2018). A primer on noise-induced transitions in applied dynamical systems. *SIAM Review*, 60(4), 969–1009. <https://doi.org/10.1137/17M1142028>
- Freidlin, M. I., & Wentzell, A. D. (2012). *Random perturbations of dynamical systems* (3rd ed., Vol. 260). Springer.
- Fried, E. I. (2020). Lack of theory building and testing impedes progress in the factor and network literature. *Psychological Inquiry*, 31(4), 271–288. <https://doi.org/10.1080/1047840X.2020.1853461>

-
- Fried, E. I., Ebrahimi, O. V., Lafit, G., Maciejewski, D., Bringmann, L., Elmer, T., Helmich, M. A., Dejonckheere, E., Hasselman, F., & Roefs, A. (2022, November 30). *MITNB consortium on improving validity of time-series measurement in social sciences*. OSF. <https://doi.org/10.17605/OSF.IO/M2V38>
- Fried, E. I., & Flake, J. K. (2018). Measurement matters. *APS Observer*, 31(3). <https://www.psychologicalscience.org/observer/measurement-matters>
- Gelman, A., Gilks, W. R., & Roberts, G. O. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, 7(1), 110–120. <https://doi.org/10.1214/aoap/1034625254>
- Gibson, E. J., & Walk, R. D. (1960). The “visual cliff.” *Scientific American*, 202(4), 64–71.
- Gilmore, R. (1993). *Catastrophe theory for scientists and engineers*. Courier Corporation.
- Glauber, R. J. (1963). Time-dependent statistics of the Ising model. *Journal of Mathematical Physics*, 4(2), 294–307. <https://doi.org/10.1063/1.1703954>
- Goldstein, J. (1999). Emergence as a construct: History and issues. *Emergence*, 1(1), 49–72. https://doi.org/10.1207/s15327000em0101_4
- Granic, I. (2005). Timing is everything: Developmental psychopathology from a dynamic systems perspective. *Developmental Review*, 25(3), 386–407. <https://doi.org/10.1016/j.dr.2005.10.005>
- Granic, I., & Hollenstein, T. (2003). Dynamic systems methods for models of developmental psychopathology. *Development and Psychopathology*, 15(3), 641–669. <https://doi.org/10.1017/S0954579403000324>
- Granic, I., Lewis, M. D., & Lichwarck-Aschoff, A. (2018). Complex pathways to psychopathology. *Psychological Inquiry*, 29(3), 145–150. <https://doi.org/10.1080/1047840X.2018.1513682>
- Gruppen, D. (2023). *Modelling the psychopathology of panic disorder—Creating a consistent, simple model for panic disorder and exploring its system dynamics to enhance our understanding of the psychopathology and treatment of panic disorder* [Master’s Research Project 2, Biology, University of Groningen]. <https://fse.studenttheses.ub.rug.nl/31838/>
- Guastello, S. J. (1982). Moderator regression and the cusp catastrophe: Application of two-stage personnel selection, training, therapy, and policy evaluation. *Behavioral Science*, 27(3), 259–272. <https://doi.org/10.1002/bs.3830270305>

- Guidoum, A. C., & Boukhetala, K. (2020). *Sim.DiffProc: Simulation of Diffusion Processes* (Version 4.6) [Computer software]. <https://CRAN.R-project.org/package=Sim.DiffProc>
- Guttal, V., Raghavendra, S., Goel, N., & Hoarau, Q. (2016). Lack of critical slowing down suggests that financial meltdowns are not critical transitions, yet rising variability could signal systemic risk. *PLOS ONE*, 11(1), e0144198.
<https://doi.org/10.1371/journal.pone.0144198>
- Haken, H. (2011). *Synergetics: Introduction and advanced topics*. Springer.
- Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, 51(5), 347–356.
<https://doi.org/10.1007/BF00336922>
- Hamaker, E. L. (2023). The curious case of the cross-sectional correlation. *Multivariate Behavioral Research*, 1–12. <https://doi.org/10.1080/00273171.2022.2155930>
- Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research*, 53(6), 820–841. <https://doi.org/10.1080/00273171.2018.1446819>
- Hamaker, E. L., Ceulemans, E., Grasman, R. P. P. P., & Tuerlinckx, F. (2015). Modeling affect dynamics: State of the art and future challenges. *Emotion Review*, 7(4), 316–322.
<https://doi.org/10.1177/1754073915590619>
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384. <https://doi.org/10.2307/1912559>
- Hamilton, J. D. (1994). *Time series analysis* (Princeton). Princeton University Press.
- Hao, N., Feng, Y., & Zhang, H. H. (2018). Model selection for high-dimensional quadratic regression via regularization. *Journal of the American Statistical Association*, 113(522), 615–625. <https://doi.org/10.1080/01621459.2016.1264956>
- Hao, N., & Zhang, H. H. (2014). Interaction screening for ultrahigh-dimensional data. *Journal of the American Statistical Association*, 109(507), 1285–1301.
<https://doi.org/10.1080/01621459.2014.881741>
- Hao, N., & Zhang, H. H. (2017). A note on high-dimensional linear regression with interactions. *The American Statistician*, 71(4), 291–297. <https://doi.org/10.1080/00031305.2016.1264311>

-
- Hardy, G. E., Cahill, J., Stiles, W. B., Ispan, C., Macaskill, N., & Barkham, M. (2005). Sudden gains in cognitive therapy for depression: A replication and extension. *Journal of Consulting and Clinical Psychology*, 73(1), 59–67.
<https://doi.org/10.1037/0022-006X.73.1.59>
- Haslbeck, J., Borsboom, D., & Waldorp, L. J. (2021). Moderated network models. *Multivariate Behavioral Research*, 56(2), 256–287. <https://doi.org/10.1080/00273171.2019.1677207>
- Haslbeck, J., Bringmann, L. F., & Waldorp, L. J. (2021). A tutorial on estimating time-varying vector autoregressive models. *Multivariate Behavioral Research*, 56(1), 120–149.
<https://doi.org/10.1080/00273171.2020.1743630>
- Haslbeck, J. M. B., Epskamp, S., Marsman, M., & Waldorp, L. J. (2021). Interpreting the Ising model: The input matters. *Multivariate Behavioral Research*, 56(2), 303–313.
<https://doi.org/10.1080/00273171.2020.1730150>
- Haslbeck, J. M. B., & Ryan, O. (2022). Recovering within-person dynamics from psychological time series. *Multivariate Behavioral Research*, 57(5), 735–766.
<https://doi.org/10.1080/00273171.2021.1896353>
- Haslbeck, J. M. B., Ryan, O., & Dablander, F. (2021). The sum of all fears: Comparing networks based on symptom sum-scores. *Psychological Methods*.
<https://doi.org/10.1037/met0000418>
- Haslbeck, J. M. B., Ryan, O., Robinaugh, D., Waldorp, L., & Borsboom, D. (2022). Modeling psychopathology: From data models to formal theories. *Psychological Methods*, 27(6), 930–957. <https://doi.org/10.1037/met0000303>
- Haslbeck, J., Ryan, O., & Dablander, F. (2023). Multimodality and skewness in emotion time series. *Emotion*, 23(8), 2117–2141. <https://doi.org/10.1037/emo0001218>
- Haslbeck, J., & Waldorp, L. J. (2020). mgm: Estimating time-varying mixed graphical models in high-dimensional data. *Journal of Statistical Software*, 93(8), 1–46.
<https://doi.org/10.18637/jss.v093.i08>
- Hasselmann, F. (2022). Early warning signals in phase space: Geometric resilience loss indicators from multiplex cumulative recurrence networks. *Frontiers in Physiology*, 13, 1–18.
<https://doi.org/10.3389/fphys.2022.859127>
- Hasselmann, F. (2023a). Going round in squares: Theory-based measurement requires a theory of measurement. *Theory & Psychology*, 33(1), 145–152.
<https://doi.org/10.1177/09593543221131511>

- Hasselman, F. (2023b). Understanding the complexity of individual developmental pathways: A primer on metaphors, models, and methods to study resilience in development. *Development and Psychopathology*, 1–13. <https://doi.org/10.1017/S0954579423001281>
- Hasselman, F., & Bosman, A. M. T. (2020). Studying complex adaptive systems with internal states: A recurrence network approach to the analysis of multivariate time-series data representing self-reports of human experience. *Frontiers in Applied Mathematics and Statistics*, 6. <https://doi.org/10.3389/fams.2020.00009>
- Hayes, A. M., & Andrews, L. A. (2020). A complex systems approach to the study of change in psychotherapy. *BMC Medicine*, 18, 197. <https://doi.org/10.1186/s12916-020-01662-2>
- Hayes, A. M., Laurenceau, J.-P., Feldman, G., Strauss, J. L., & Cardaciotto, L. (2007). Change is not always linear: The study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review*, 27(6), 715–723. <https://doi.org/10.1016/j.cpr.2007.01.008>
- Hayes, A. M., & Yasinski, C. (2015). Pattern destabilization and emotional processing in cognitive therapy for personality disorders. *Frontiers in Psychology*, 6, 107.
- Hayes, A. M., Yasinski, C., Ben Barnes, J., & Bockting, C. L. H. (2015). Network destabilization and transition in depression: New methods for studying the dynamics of therapeutic change. *Clinical Psychology Review*, 41, 27–39. <https://doi.org/10.1016/j.cpr.2015.06.007>
- Heino, M. T. J., Proverbio, D., Resnicow, K., Marchand, G., & Hankonen, N. (2022). *Attractor landscapes: A unifying conceptual model for understanding behaviour change across scales of observation*. 17(4), 655–672. <https://doi.org/10.1080/17437199.2022.2146598>
- Hekler, E. B., Klasnja, P., Chevance, G., Golaszewski, N. M., Lewis, D., & Sim, I. (2019). Why we need a small data paradigm. *BMC Medicine*, 17(1), 133. <https://doi.org/10.1186/s12916-019-1366-x>
- Helmich, M. A. (2024). Duration-adjusted Reliable Change Index (DaRCI): Defining clinically relevant symptom changes of varying durations. *Assessment*, 31(7), 1493–1507. <https://doi.org/10.1177/10731911231221808>
- Helmich, M. A., Olthof, M., Oldehinkel, A. J., Wichers, M., Bringmann, L. F., & Smit, A. C. (2021). Early warning signals and critical transitions in psychopathology: Challenges and recommendations. *Current Opinion in Psychology*, 41, 51–58. <https://doi.org/10.1016/j.copsyc.2021.02.008>

-
- Helmich, M. A., Schreuder, M. J., Bringmann, L. F., Riese, H., Snippe, E., & Smit, A. C. (2024). Slow down and be critical before using early warning signals in psychopathology. *Nature Reviews Psychology*, 3, 767–780. <https://doi.org/10.1038/s44159-024-00369-y>
- Helmich, M. A., Smit, A., Bringmann, L., Schreuder, M., Oldehinkel, A., Wichers, M., & Snippe, E. (2022). Detecting impending symptom transitions using early warning signals in individuals receiving treatment for depression. *Clinical Psychological Science*, 11(6), 994–1010. <https://doi.org/10.1177/21677026221137006>
- Helmich, M. A., Wichers, M., Olthof, M., Strunk, G., Aas, B., Aichhorn, W., Schiepek, G., & Snippe, E. (2020). Sudden gains in day-to-day change: Revealing nonlinear patterns of individual improvement in depression. *Journal of Consulting and Clinical Psychology*, 88(2), 119–127. <https://doi.org/10.1037/ccp0000469>
- Heymann, M., & Vanden-Eijnden, E. (2008). Pathways of maximum likelihood for rare events in nonequilibrium systems: Application to nucleation in the presence of shear. *Physical Review Letters*, 100(14), 140601. <https://doi.org/10.1103/PhysRevLett.100.140601>
- Hitchcock, P., Fried, E. I., & Frank, M. (2022). Computational psychiatry needs time and context. *Annual Review of Psychology*, 73, 243–270.
<https://doi.org/10.1146/annurev-psych-021621-124910>
- Hoekstra, R. H. A., De Ron, J., Epskamp, S., Robinaugh, D., & Borsboom, D. (2024). *Mapping the dynamics of idiographic network models to the network theory of psychopathology using stability landscapes*. PsyArXiv. <https://doi.org/10.31234/osf.io/9sguw>
- Holtzheimer, P. E., & Mayberg, H. S. (2011). Stuck in a rut: Rethinking depression and its treatment. *Trends in Neurosciences*, 34(1), 1–9. <https://doi.org/10.1016/j.tins.2010.10.004>
- Hosenfeld, B., Bos, E. H., Wardenaar, K. J., Conradi, H. J., van der Maas, H. L. J., Visser, I., & de Jonge, P. (2015). Major depressive disorder as a nonlinear dynamic system: Bimodality in the frequency distribution of depressive symptoms over time. *BMC Psychiatry*, 15(1).
<https://doi.org/10.1186/s12888-015-0596-5>
- Houben, M., Vansteelandt, K., Claes, L., Sienaert, P., Berens, A., Sleuwaegen, E., & Kuppens, P. (2016). Emotional switching in borderline personality disorder: A daily life study. *Personality Disorders: Theory, Research, and Treatment*, 7(1), 50–60.
<https://doi.org/10.1037/per0000126>
- Hulsmans, D. H. G., J. W. Oude Maatman, F., Otten, R., A. P. Poelen, E., & Lichtwarck-Aschoff, A. (2024). Idiographic personality networks: Stability, variability and when they become

- problematic. *Journal of Research in Personality*, 104468.
<https://doi.org/10.1016/j.jrp.2024.104468>
- Hulsmans, D. H. G., Otten, R., Poelen, E. A. P., van Vonderen, A., Daalmans, S., Hasselman, F., Olthof, M., & Lichtwarck-Aschoff, A. (2024). A complex systems perspective on chronic aggression and self-injury: Case study of a woman with mild intellectual disability and borderline personality disorder. *BMC Psychiatry*, 24(1), 378.
<https://doi.org/10.1186/s12888-024-05836-7>
- Iacus, S. M. (2022). *sde: Simulation and inference for stochastic differential equations* [Manual].
<https://CRAN.R-project.org/package=sde>
- Ising, E. (1925). Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31(1), 253–258.
<https://doi.org/10.1007/BF02980577>
- Iwanski, J. S., & Bradley, E. (1998). Recurrence plots of experimental data: To embed or not to embed? *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 8(4), 861–871.
<https://doi.org/10.1063/1.166372>
- Jones, B. (2023). Breaking boundaries with generative AI: Good practice of unleashing the power of ChatGPT for inter- and transdisciplinary breakthroughs in the age of complexity. *Journal of the European Honors Council*, 6. <https://doi.org/10.31378/jehc.177>
- Jordan, D. G., Winer, E. S., & Salem, T. (2020). The current status of temporal network analysis for clinical science: Considerations as the paradigm shifts? *Journal of Clinical Psychology*, 76(9), 1591–1612. <https://doi.org/10.1002/jclp.22957>
- Kalis, A., & Borsboom, D. (2020). Folk psychology as a causal language. *Theory & Psychology*, 30(5), 723–728. <https://doi.org/10.1177/0959354320933940>
- Kalokerinos, E. K., Russo-Batterham, D., Koval, P., Moeck, E. K., Grewal, K. K., Greenaway, K. H., Shrestha, K. M., Garrett, P., Michalewicz, A., Garber, J., & Kuppens, P. (in preparation). *The EMOTE Database: An open, searchable database of experience sampling data mapping everyday life*.
- Kane, M. J., Emerson, J., & Weston, S. (2013). Scalable strategies for computing with massive data. *Journal of Statistical Software*, 55(14), 1–19. <https://doi.org/10.18637/jss.v055.i14>
- Kazdin, A. E., & Wilcoxon, L. A. (1976). Systematic desensitization and nonspecific treatment effects: A methodological evaluation. *Psychological Bulletin*, 83(5), 729–758.
<https://doi.org/10.1037/0033-2909.83.5.729>

-
- Keener, J., & Sneyd, J. (2009). Biochemical reactions. In J. Keener & J. Sneyd (Eds.), *Mathematical physiology: I: Cellular physiology* (pp. 1–47). Springer.
https://doi.org/10.1007/978-0-387-75847-3_1
- Kelso, J. A. S., Scholz, J. P., & Schöner, G. (1986). Nonequilibrium phase transitions in coordinated biological motion: Critical fluctuations. *Physics Letters A*, 118(6), 279–284.
[https://doi.org/10.1016/0375-9601\(86\)90359-2](https://doi.org/10.1016/0375-9601(86)90359-2)
- Kendler, K. S., Aggen, S. H., Werner, M., & Fried, E. I. (2022). A topography of 21 phobic fears: Network analysis in an epidemiological sample of adult twins. *Psychological Medicine*, 52(13), 2588–2595. <https://doi.org/10.1017/S0033291720004493>
- Kendler, Kenneth S., & Prescott, Carol A. (1999). A population-based twin study of lifetime major depression in men and women. *Archives of General Psychiatry*, 56(1), 39–44.
<https://doi.org/10.1001/archpsyc.56.1.39>
- Kloek, T., & van Dijk, H. K. (1978). Bayesian estimates of equation system parameters: An application of integration by Monte Carlo. *Econometrica*, 46(1), 1–19.
<https://doi.org/10.2307/1913641>
- Kogo, N., Galli, A., & Wagemans, J. (2011). Switching dynamics of border ownership: A stochastic model for bi-stable perception. *Vision Research*, 51(18), 2085–2098.
<https://doi.org/10.1016/j.visres.2011.08.010>
- Korotkikh, V. (2001). Kolmogorov complexity. In C. A. Floudas & P. M. Pardalos (Eds.), *Encyclopedia of optimization* (pp. 1191–1196). Springer US.
https://doi.org/10.1007/0-306-48332-7_241
- Kossakowski, J., Groot, P., Haslbeck, J., Borsboom, D., & Wichers, M. (2017). Data from ‘Critical slowing down as a personalized early warning signal for depression’ *Journal of Open Psychology Data*, 5, Article 1. <https://doi.org/10.5334/jopd.29>
- Koval, P., Brose, A., Pe, M. L., Houben, M., Erbas, Y., Champagne, D., & Kuppens, P. (2015). Emotional inertia and external events: The roles of exposure, reactivity, and recovery. *Emotion*, 15(5), 625–636. <https://doi.org/10.1037/emo0000059>
- Koval, P., & Kuppens, P. (2024). Changing feelings: Individual differences in emotional inertia. In A. C. Samson, D. Sander, & U. Kramer, *Change in emotion and mental health* (pp. 3–21). Academic Press. <https://doi.org/10.1016/B978-0-323-95604-8.00007-1>

- Kramer, J., & Ross, J. (1985). Stabilization of unstable states, relaxation, and critical slowing down in a bistable system. *The Journal of Chemical Physics*, 83(12), 6234–6241.
<https://doi.org/10.1063/1.449571>
- Kroc, E., & Olvera Astivia, O. L. (2023). The case for the curve: Parametric regression with second- and third-order polynomial functions of predictors should be routine. *Psychological Methods*. <https://doi.org/10.1037/met0000629>
- Kuehn, C. (2011). A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics. *Physica D: Nonlinear Phenomena*, 240(12), 1020–1035.
<https://doi.org/10.1016/j.physd.2011.02.012>
- Kuehn, C., Lux, K., & Neamțu, A. (2022). Warning signs for non-Markovian bifurcations: Colour blindness and scaling laws. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 478, 20210740. <https://doi.org/10.1098/rspa.2021.0740>
- Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological Science*, 21(7), 984–991.
<https://doi.org/10.1177/0956797610372634>
- Kuppens, P., Oravecz, Z., & Tuerlinckx, F. (2010). Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology*, 99(6), 1042–1060. <https://doi.org/10.1037/a0020962>
- Kuznetsov, Y. A. (2023). *Elements of applied bifurcation theory* (Vol. 112). Springer International Publishing. <https://doi.org/10.1007/978-3-031-22007-4>
- Lakens, D. (2025). Concerns about replicability across two crises in social psychology. *International Review of Social Psychology*, 38(1), 5. <https://doi.org/10.5334/irsp.1036>
- Lamothe, K. A., Somers, K. M., & Jackson, D. A. (2019). Linking the ball-and-cup analogy and ordination trajectories to describe ecosystem stability, resistance, and resilience. *Ecosphere*, 10(3), e02629. <https://doi.org/10.1002/ecs2.2629>
- Lane, S. T., & Gates, K. M. (2017). Automated selection of robust individual-level structural equation models for time series data. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(5), 768–782. <https://doi.org/10.1080/10705511.2017.1309978>
- Lee, C., & Hu, X. (2022). Sex differences in depressive symptom networks among community-dwelling older adults. *Nursing Research*, 71(5), 370–379.
<https://doi.org/10.1097/nnr.0000000000000601>

-
- Lenton, T. M. (2011). Early warning of climate tipping points. *Nature Climate Change*, 1(4), 201–209. <https://doi.org/10.1038/nclimate1143>
- Levins, R. (1966). The strategy of model building in population biology. *American Scientist*, 54(4), 421–431.
- Lewars, E. (Ed.). (2003). The concept of the potential energy surface. In *Computational chemistry: Introduction to the theory and applications of molecular and quantum mechanics* (pp. 9–41). Springer US. https://doi.org/10.1007/0-306-48391-2_2
- Lewis, P. A. W., & Stevens, J. G. (1991). Nonlinear modeling of time series using multivariate adaptive regression splines (MARS). *Journal of the American Statistical Association*, 86(416), 864–877. <https://doi.org/10.2307/2290499>
- Li, C., & Wang, J. (2013). Quantifying cell fate decisions for differentiation and reprogramming of a human stem cell network: Landscape and biological paths. *PLOS Computational Biology*, 9(8), e1003165. <https://doi.org/10.1371/journal.pcbi.1003165>
- Li, C., & Ye, L. (2019). Landscape and flux govern cellular mode-hopping between oscillations. *The Journal of Chemical Physics*, 151(17), 175101. <https://doi.org/10.1063/1.5125046>
- Li, R., Root, T. L., & Shiffman, S. (2006). A local linear estimation procedure for functional multilevel modeling. In T. A. Walls & J. L. Schafer (Eds.), *Models for intensive longitudinal data* (pp. 63–83). Oxford University Press.
- Lichtwarck-Aschoff, A., & van Geert, P. (2004). A dynamic systems perspective on social cognition, problematic behaviour, and intervention in adolescence. *European Journal of Developmental Psychology*, 1(4), 399–411. <https://doi.org/10.1080/17405620444000157>
- Liu, Q., & Wang, L. (2021). T-Test and ANOVA for data with ceiling and/or floor effects. *Behavior Research Methods*, 53(1), 264–277. <https://doi.org/10.3758/s13428-020-01407-2>
- Livina, V. N., Kwasniok, F., & Lenton, T. M. (2010). Potential analysis reveals changing number of climate states during the last 60 kyr. *Climate of the Past*, 6, 77–82.
- Lloyd, S. (2001). Measures of complexity: A nonexhaustive list. *IEEE Control Systems Magazine*, 21(4), 7–8.
- Loossens, T., Mestdagh, M., Dejonckheere, E., Kuppens, P., Tuerlinckx, F., & Verdonck, S. (2020). The Affective Ising Model: A computational account of human affect dynamics. *PLOS Computational Biology*, 16(5), e1007860. <https://doi.org/10.1371/journal.pcbi.1007860>

- López-Ratón, M., Rodríguez-Álvarez, M. X., Suárez, C. C., & Sampedro, F. G. (2014). OptimalCutpoints: An R package for selecting optimal cutpoints in diagnostic tests. *Journal of Statistical Software*, 61(8), 1–36. <https://doi.org/10.18637/jss.v061.i08>
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley.
- Lunansky, G., Bonanno, G. A., van Borkulo, C. D., Blanken, T. F., Cramer, A. O. J., & Borsboom, D. (2024). Bouncing back from life's perturbations: Formalizing psychological resilience from a complex systems perspective. *Psychological Review*. <https://doi.org/10.1037/rev0000497>
- Luo, S., & Chen, Z. (2021). Sequential interaction group selection by the principle of correlation search for high-dimensional interaction models. *Statistica Sinica*, 31(1), 197–221.
- Lutz, W., Ehrlich, T., Rubel, J., Hallwachs, N., Röttger, M.-A., Jorasz, C., Mocanu, S., Vocks, S., Schulte, D., & Tschitsaz-Stucki, A. (2013). The ups and downs of psychotherapy: Sudden gains and sudden losses identified with session reports. *Psychotherapy Research*, 23(1), 14–24. <https://doi.org/10.1080/10503307.2012.693837>
- Ma, J., Zhao, J., Tian, J., Bai, X., & Tu, Z. (2013). Regularized vector field learning with sparse approximation for mismatch removal. *Pattern Recognition*, 46(12), 3519–3532. <https://doi.org/10.1016/j.patcog.2013.05.017>
- MacKinnon, D. F., & Pies, R. (2006). Affective instability as rapid cycling: Theoretical and clinical implications for borderline personality and bipolar spectrum disorders. *Bipolar Disorders*, 8, 1–14. <https://doi.org/10.1111/j.1399-5618.2006.00283.x>
- Mallinckrodt, B., Abraham, W. T., Wei, M., & Russell, D. W. (2006). Advances in testing the statistical significance of mediation effects. *Journal of Counseling Psychology*, 53(3), 372–378. <https://doi.org/10.1037/0022-0167.53.3.372>
- Mansueto, A. C., Wiers, R. W., van Weert, J. C. M., Schouten, B. C., & Epskamp, S. (2023). Investigating the feasibility of idiographic network models. *Psychological Methods*, 28(5), 1052–1068. <https://doi.org/10.1037/met0000466>
- Margalit, D., & Rabinoff, J. (2019). Eigenvalues and eigenvectors. In *Interactive linear algebra*. <https://textbooks.math.gatech.edu/ila/chap-eigenvalues.html>
- Marwan, N., Carmen Romano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5), 237–329. <https://doi.org/10.1016/j.physrep.2006.11.001>

-
- Mattera, A., Pagani, M., & Baldassarre, G. (2020). A computational model integrating multiple phenomena on cued fear conditioning, extinction, and reinstatement. *Frontiers in Systems Neuroscience*, 14. <https://doi.org/10.3389/fnsys.2020.569108>
- Matteson, D. S., & James, N. A. (2014). A nonparametric approach for multiple change point analysis of multivariate data. *Journal of the American Statistical Association*, 109(505), 334–345. <https://doi.org/10.1080/01621459.2013.849605>
- McLeod, A. I. (2011). *Kendall: Kendall rank correlation and Mann-Kendall trend test* [Manual]. <https://CRAN.R-project.org/package=Kendall>
- Meehl, P. E. (1990a). Appraising and amending theories: The strategy of lakatosian defense and two principles that warrant it. *Psychological Inquiry*, 1(2), 108–141.
- Meehl, P. E. (1990b). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports*, 66(1), 195–244. <https://doi.org/10.2466/pr0.1990.66.1.195>
- Meisel, C., & Kuehn, C. (2012). Scaling effects and spatio-temporal multilevel dynamics in epileptic seizures. *PLoS ONE*, 7(2), e30371. <https://doi.org/10.1371/journal.pone.0030371>
- Mendoza, N. B., Frondozo, C. E., Dizon, J. I. W. T., & Buenconsejo, J. U. (2022). The factor structure and measurement invariance of the PHQ-4 and the prevalence of depression and anxiety in a Southeast Asian context amid the COVID-19 pandemic. *Current Psychology*. <https://doi.org/10.1007/s12144-022-02833-5>
- Meucci, A. (2009). *Review of statistical arbitrage, cointegration, and multivariate Ornstein-Uhlenbeck* (1404905). SSRN. <https://doi.org/10.2139/ssrn.1404905>
- Microsoft, & Weston, S. (2020). *foreach: Provides foreach looping construct* [Manual]. <https://CRAN.R-project.org/package=foreach>
- Milad, M. R., & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: Ten years of progress. *Annual Review of Psychology*, 63(1), 129–151. <https://doi.org/10.1146/annurev.psych.121208.131631>
- Miller, W. R. (2004). The phenomenon of quantum change. *Journal of Clinical Psychology*, 60(5), 453–460. <https://doi.org/10.1002/jclp.20000>
- Milnor, J. W. (2006). Attractor. *Scholarpedia*, 1(11), 1815. <https://doi.org/10.4249/scholarpedia.1815>

- Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives*, 2(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1
- Moore, C. M., Stieha, C. R., Nolting, B. C., & Cameron, M. K. (2016). *QPot: An R package for stochastic differential equation quasi-potential analysis*. 8, 20.
- Moore, J. W., Stanitski, C. L., & Jurs, P. C. (2006). Thermodynamic and kinetic stability. In *Chemistry: The molecular science* (3rd ed., pp. 905–908). Thomson Brooks/Cole.
- Morr, A., & Boers, N. (2024). Detection of approaching critical transitions in natural systems driven by red noise. *Physical Review X*, 14(2), 021037.
<https://doi.org/10.1103/PhysRevX.14.021037>
- Morrow, S. L. (2005). Quality and trustworthiness in qualitative research in counseling psychology. *Journal of Counseling Psychology*, 52(2), 250–260.
<https://doi.org/10.1037/0022-0167.52.2.250>
- Moulder, R. G., Martynova, E., & Boker, S. M. (2022). Extracting nonlinear dynamics from psychological and behavioral time series through HAVOK analysis. *Multivariate Behavioral Research*, 0(0), 1–25. <https://doi.org/10.1080/00273171.2021.1994848>
- Müller, K. (2021a). *Jointprof* [Manual]. <https://github.com/r-prof/jointprof/>
- Müller, K. (2021b). *profile: Read, manipulate, and write profiler data* [Manual].
- Murray, J. D. (1989). Biological waves: Single species models. In J. D. Murray (Ed.), *Mathematical biology* (pp. 274–310). Springer.
https://doi.org/10.1007/978-3-662-08539-4_11
- Myin-Germeys, I., & Kuppens, P. (Eds.). (2021). *The open handbook of experience sampling methodology: A step-by-step guide to designing, conducting, and analyzing ESM studies*. Center for Research on Experience Sampling and Ambulatory Methods Leuven.
- Nelder, J. A. (1977). A reformulation of linear models. *Journal of the Royal Statistical Society. Series A (General)*, 140(1), 48–77. <https://doi.org/10.2307/2344517>
- Ninot, G., Fortes, M., & Delignieres, D. (2001). A psychometric tool for the assessment of the dynamics of the physical self. *European Review of Applied Psychology / Revue Européenne de Psychologie Appliquée*, 51, 205–216.
- Noble, D. (2011). A theory of biological relativity: No privileged level of causation. *Interface Focus*, 2(1), 55–64. <https://doi.org/10.1098/rsfs.2011.0067>

-
- Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, 26(5), 1596–1618.
<https://doi.org/10.3758/s13423-019-01645-2>
- Okino, M. S., & Mavrovouniotis, M. L. (1998). Simplification of mathematical models of chemical reaction systems. *Chemical Reviews*, 98(2), 391–408.
<https://doi.org/10.1021/cr950223l>
- Oliva, T. A., & Capdevielle, C. M. (1980). Sussman and zahler: Throwing the baby out with the bath water. *Behavioral Science*, 25(3), 229–230. <https://doi.org/10.1002/bs.3830250308>
- Olthof, M., Bunge, A., Maciejewski, D., Hasselman, F., & Lichtwarck-Aschoff, A. (2024). Transitions and resilience in ecological momentary assessment: A multiple single-case study. *Journal for Person-Oriented Research*, 10(2), 89–99.
<https://doi.org/10.17505/jpor.2024.27102>
- Olthof, M., Hasselman, F., Aas, B., Lamoth, D., Scholz, S., Daniels-Wredenhagen, N., Weinans, E., Strunk, G., Schiepek, G., Bosman, A., & Lichtwarck-Aschoff, A. (2023). The best of both worlds? General principles of psychopathology in personalized assessment. *Journal of Psychopathology and Clinical Science*, 132(7), 808–819.
<https://doi.org/10.1037/abn0000858>
- Olthof, M., Hasselman, F., & Lichtwarck-Aschoff, A. (2020). Complexity in psychological self-ratings: Implications for research and practice. *BMC Medicine*, 18(1), 317.
<https://doi.org/10.1186/s12916-020-01727-2>
- Olthof, M., Hasselman, F., Oude Maatman, F., Bosman, A. M. T., & Lichtwarck-Aschoff, A. (2023). Complexity theory of psychopathology. *Journal of Psychopathology and Clinical Science*, 132(3), 314–323. <https://doi.org/10.1037/abn0000740>
- Olthof, M., Hasselman, F., Strunk, G., Aas, B., Schiepek, G., & Lichtwarck-Aschoff, A. (2020). Destabilization in self-ratings of the psychotherapeutic process is associated with better treatment outcome in patients with mood disorders. *Psychotherapy Research*, 30(4), 520–531. <https://doi.org/10.1080/10503307.2019.1633484>
- Olthof, M., Hasselman, F., Strunk, G., van Rooij, M., Aas, B., Helmich, M. A., Schiepek, G., & Lichtwarck-Aschoff, A. (2020). Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clinical Psychological Science*, 8(1), 25–35. <https://doi.org/10.1177/2167702619865969>

- Olthof, M., Mestdagh, M., Hasselman, F., Maciejewski, D., Bunge, A., Dejonckheere, E., Schat, E., Ceulemans, E., & Lichtwarck-Aschoff, A. (2025). *Micro interventions based on real-time monitoring of affective instability: A proof-of-concept study*. PsyArXiv. https://doi.org/10.31234/osf.io/e7djf_v1
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. <https://doi.org/10.1126/science.aac4716>
- Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2009). A hierarchical Ornstein–Uhlenbeck model for continuous repeated measurement data. *Psychometrika*, 74(3), 395–418. <https://doi.org/10.1007/s11336-008-9106-8>
- Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2011). A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods*, 16(4), 468–490. <https://doi.org/10.1037/a0024375>
- Oude Maatman, F. (2020). Reformulating the network theory of mental disorders: Folk psychology as a factor, not a fact. *Theory & Psychology*. <https://doi.org/10.1177/0959354320921464>
- Oude Maatman, F. (2021). *Psychology's theory crisis, and why formal modelling cannot solve it*. PsyArXiv. <https://doi.org/10.31234/osf.io/puqvs>
- Oude Maatman, F., & Eronen, M. (2025). *Unraveling networks: The conceptual incoherence of the network approach*. PsyArXiv. <https://doi.org/10.31234/osf.io/h4jv2>
- Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Basic Books.
- Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, 6(1), 7–11.
- Powers, M. B., Halpern, J. M., Ferenschak, M. P., Gillihan, S. J., & Foa, E. B. (2010). A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clinical Psychology Review*, 30(6), 635–641. <https://doi.org/10.1016/j.cpr.2010.04.007>
- Proverbio, D., Kemp, F., Magni, S., & Gonçalves, J. (2022). Performance of early warning signals for disease re-emergence: A case study on COVID-19 data. *PLOS Computational Biology*, 18(3), e1009958. <https://doi.org/10.1371/journal.pcbi.1009958>
- Proverbio, D., Skupin, A., & Gonçalves, J. (2023). Systematic analysis and optimization of early warning signals for critical transitions using distribution data. *iScience*, 26(7), 107156. <https://doi.org/10.1016/j.isci.2023.107156>

-
- Puth, M.-T., Neuhäuser, M., & Ruxton, G. D. (2015). On the variety of methods for calculating confidence intervals by bootstrapping. *Journal of Animal Ecology*, 84(4), 892–897.
<https://doi.org/10.1111/1365-2656.12382>
- Qiu, X., Zhang, Y., Martin-Rufino, J. D., Weng, C., Hosseinzadeh, S., Yang, D., Pogson, A. N., Hein, M. Y., Min, K. H. (Joseph), Wang, L., Grody, E. I., Shurtleff, M. J., Yuan, R., Xu, S., Ma, Y., Replogle, J. M., Lander, E. S., Darmanis, S., Bahar, I., ... Weissman, J. S. (2022). Mapping transcriptomic vector fields of single cells. *Cell*, 185(4), 690-711.e45.
<https://doi.org/10.1016/j.cell.2021.12.045>
- R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software]. <https://www.R-project.org/>
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922.
<https://doi.org/10.1162/neco.2008.12-06-420>
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281.
<https://doi.org/10.1016/j.tics.2016.01.007>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.
- Richardson, M., Paxton, A., & Kuznetsov, N. (2017). Nonlinear methods for understanding complex dynamical phenomena in psychological science. *APA Psychological Science Agenda*. <http://www.apa.org/science/about/psa/2017/02/dynamical-phenomena.aspx>
- Rigdon, E. E., Schumacker, R. E., & Wothke, W. (1998). A comparative review of interaction and nonlinear modeling. In *Interaction and nonlinear effects in structural equation modeling*. Routledge.
- Ritchie, P. D. L., Alkhayouon, H., Cox, P. M., & Wieczorek, S. (2023). Rate-induced tipping in natural and human systems. *Earth System Dynamics*, 14(3), 669–683.
<https://doi.org/10.5194/esd-14-669-2023>
- Ritchie, P. D. L., & Sieber, J. (2016). Early-warning indicators for rate-induced tipping. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(9), 093116.
<https://doi.org/10.1063/1.4963012>

- Ritchie, P. D. L., & Sieber, J. (2017). Probability of noise- and rate-induced tipping. *Physical Review E*, 95(5), 052209. <https://doi.org/10.1103/PhysRevE.95.052209>
- Robinaugh, D. J., Haslbeck, J. M. B., Waldorp, L. J., Kossakowski, J. J., Fried, E. I., Millner, A. J., McNally, R. J., Ryan, O., de Ron, J., van der Maas, H. L. J., van Nes, E. H., Scheffer, M., Kendler, K. S., & Borsboom, D. (2024). Advancing the network theory of mental disorders: A computational model of panic disorder. *Psychological Review*, 131(6), 1482–1508. <https://doi.org/10.1037/rev0000515>
- Robinaugh, D. J., Haslbeck, J., Ryan, O., Fried, E. I., & Waldorp, L. J. (2021). Invisible hands and fine calipers: A call to use formal theory as a toolkit for theory construction. *Perspectives on Psychological Science*, 16(4), 725–743. <https://doi.org/10.1177/1745691620974697>
- Robinaugh, D. J., Hoekstra, R. H. A., Toner, E. R., & Borsboom, D. (2020). The network approach to psychopathology: A review of the literature 2008–2018 and an agenda for future research. *Psychological Medicine*, 50(3), 353–366. <https://doi.org/10.1017/S0033291719003404>
- Rodríguez-Sánchez, P., Nes, E. H. van, & Scheffer, M. (2020). Climbing Escher's stairs: A way to approximate stability landscapes in multidimensional systems. *PLOS Computational Biology*, 16(4), e1007788. <https://doi.org/10.1371/journal.pcbi.1007788>
- Rottenberg, J. (2005). Mood and emotion in major depression. *Current Directions in Psychological Science*, 14(3), 167–170. <https://doi.org/10.1111/j.0963-7214.2005.00354.x>
- Rowland, Z., & Wenzel, M. (2020). Mindfulness and affect-network density: Does mindfulness facilitate disengagement from affective experiences in daily life? *Mindfulness*, 11(5), 1253–1266. <https://doi.org/10.1007/s12671-020-01335-4>
- Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review*, 5(1), 2–14. https://doi.org/10.1207/S15327957PSPR0501_1
- Rubino, G., & Tuffin, B. (2009). *Rare event simulation using monte carlo methods*. John Wiley & Sons, Incorporated. <https://doi.org/10.1002/9780470745403>
- Ryan, O., & Hamaker, E. L. (2022). Time to intervene: A continuous-time approach to network analysis and centrality. *Psychometrika*, 87(1), 214–252. <https://doi.org/10.1007/s11336-021-09767-0>
- Ryan, O., Haslbeck, J. M. B., & Robinaugh, J. D. (2025). Improving treatments for mental disorders using computational models. *Behaviour Research and Therapy*, 189, 104706. <https://doi.org/10.1016/j.brat.2025.104706>

Ryan, O., Kuiper, R., & Hamaker, E. (2018). A continuous-time approach to intensive longitudinal data: What, why, and how? In K. L. Montfort, J. H. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 29–57). Springer.
https://doi.org/10.1007/978-3-319-77219-6_2

Sakurambo. (2005). *An illusion of an endless staircase using false perspective* [Digital painting]. Wikimedia Commons.
https://commons.wikimedia.org/wiki/File:Impossible_staircase.svg

Sayama, H. (2015a). Continuous-time models II: Analysis. In *Introduction to the modeling and analysis of complex systems* (pp. 111–129).

Sayama, H. (2015b). *Introduction to the modeling and analysis of complex systems*. Open SUNY Textbooks.

Schat, E., Tuerlinckx, F., Smit, A. C., De Ketelaere, B., & Ceulemans, E. (2023). Detecting mean changes in experience sampling data in real time: A comparison of univariate and multivariate statistical process control methods. *Psychological Methods*, 28(6), 1335–1357.
<https://doi.org/10.1037/met0000447>

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M., & Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, 461(7260), Article 7260. <https://doi.org/10.1038/nature08227>

Scheffer, M., Bolhuis, J. E., Borsboom, D., Buchman, T. G., Gijzel, S. M. W., Goulson, D., Kammenga, J. E., Kemp, B., van de Leemput, I. A., Levin, S., Martin, C. M., Melis, R. J. F., van Nes, E. H., Romero, L. M., & Olde Rikkert, M. G. M. (2018). Quantifying resilience of humans and other animals. *Proceedings of the National Academy of Sciences*, 115(47), 11883–11890. <https://doi.org/10.1073/pnas.1810630115>

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., van de Koppel, J., van de Leemput, I. A., Levin, S. A., van Nes, E. H., Pascual, M., & Vandermeer, J. (2012). Anticipating critical transitions. *Science*, 338(6105), 344–348.
<https://doi.org/10.1126/science.1225244>

Schelldorfer, J., Meier, L., & Bühlmann, P. (2014). GLMMlasso: An algorithm for high-dimensional generalized linear mixed models using ℓ_1 -penalization. *Journal of Computational and Graphical Statistics*, 23(2), 460–477.

Schiepek, G., & Strunk, G. (2010). The identification of critical fluctuations and phase transitions in short term and coarse-grained time series—A method for the real-time

- monitoring of human change processes. *Biological Cybernetics*, 102(3), 197–207.
<https://doi.org/10.1007/s00422-009-0362-1>
- Schiepek, G., Tominschek, I., & Heinzel, S. (2014). Self-organization in psychotherapy: Testing the synergetic model of change processes. *Frontiers in Psychology*, 5.
<https://doi.org/10.3389/fpsyg.2014.01089>
- Schiepek, G., & Tschacher, W. (1992). Application of synergetics to clinical psychology. In W. Tschacher, G. Schiepek, & E. J. Brunner (Eds.), *Self-organization and clinical psychology: Empirical approaches to synergetics in psychology* (pp. 3–31). Springer.
https://doi.org/10.1007/978-3-642-77534-5_1
- Schiepek, G., Viol, K., Aichhorn, W., Hütt, M.-T., Sungler, K., Pincus, D., & Schöller, H. J. (2017). Psychotherapy is chaotic—(Not only) in a computational world. *Frontiers in Psychology*, 8.
<https://doi.org/10.3389/fpsyg.2017.00379>
- Schöller, H., Viol, K., Aichhorn, W., Hütt, M.-T., & Schiepek, G. (2018). Personality development in psychotherapy: A synergetic model of state-trait dynamics. *Cognitive Neurodynamics*, 12(5), 441–459. <https://doi.org/10.1007/s11571-018-9488-y>
- Schreuder, M. J., Hartman, C. A., George, S. V., Menne-Lothmann, C., Decoster, J., van Winkel, R., Delespaul, P., De Hert, M., Derom, C., Thiery, E., Rutten, B. P. F., Jacobs, N., van Os, J., Wigman, J. T. W., & Wichers, M. (2020). Early warning signals in psychopathology: What do they tell? *BMC Medicine*, 18(1), 269. <https://doi.org/10.1186/s12916-020-01742-3>
- Schreuder, M. J., Wigman, J. T. W., Groen, R. N., Weinans, E., Wichers, M., & Hartman, C. A. (2022). Anticipating the direction of symptom progression using critical slowing down: A proof-of-concept study. *BMC Psychiatry*, 22, 49.
<https://doi.org/10.1186/s12888-022-03686-9>
- Schumacher, L., Lunansky, G., Cui, J., Marsman, M., Borsboom, D., & Kriston, L. (in preparation). *How does the stability of mental states change as a result of treatment? Treatment evaluation with the help of Stability Landscapes*.
- Schumacher, S., Miller, R., Fehm, L., Kirschbaum, C., Fydrich, T., & Ströhle, A. (2015). Therapists' and patients' stress responses during graduated versus flooding in vivo exposure in the treatment of specific phobia: A preliminary observational study. *Psychiatry Research*, 230(2), 668–675. <https://doi.org/10.1016/j.psychres.2015.10.020>
- Scott, A. (Ed.). (2005). *Encyclopedia of nonlinear science*. Routledge.
- Serès, P. (Ed.). (2020). *Computational psychiatry: A primer*. The MIT Press.

-
- Shalom, J. G., & Aderka, I. M. (2020). A meta-analysis of sudden gains in psychotherapy: Outcome and moderators. *Clinical Psychology Review*, 76, 101827.
<https://doi.org/10.1016/j.cpr.2020.101827>
- Shao, J. (1997). An asymptotic theory for linear model selection. *Statistica Sinica*, 7(2), 221–242.
- Shi, J., Li, T., & Chen, L. (2016). Towards a critical transition theory under different temporal scales and noise strengths. *Physical Review E*, 93(3), 032137.
<https://doi.org/10.1103/PhysRevE.93.032137>
- Sievert, C. (2020). *Interactive web-based data visualization with R, plotly, and shiny*. Chapman and Hall/CRC. <https://plotly-r.com>
- Simmering, V. R., Schutte, A. R., & Spencer, J. P. (2008). Generalizing the dynamic field theory of spatial cognition across real and developmental time scales. *Brain Research*, 1202, 68–86. <https://doi.org/10.1016/j.brainres.2007.06.081>
- Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, 106(6), 467–482.
- Smit, A. C., & Snippe, E. (2023). Real-time monitoring of increases in restlessness to assess idiographic risk of recurrence of depressive symptoms. *Psychological Medicine*, 53(11), 5060–5069. <https://doi.org/10.1017/S0033291722002069>
- Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The brief resilience scale: Assessing the ability to bounce back. *International Journal of Behavioral Medicine*, 15(3), 194–200. <https://doi.org/10.1080/10705500802222972>
- Smith, B. W., Epstein, E. M., Ortiz, J. A., Christopher, P. J., & Tooley, E. M. (2013). The foundations of resilience: What are the critical resources for bouncing back from stress? In S. Prince-Embury & D. H. Saklofske (Eds.), *Resilience in children, adolescents, and adults: Translating research into practice* (pp. 167–187). Springer. https://doi.org/10.1007/978-1-4614-4939-3_13
- Spencer, J. P., Smith, L. B., & Thelen, E. (2001). Tests of a dynamic systems account of the A-not-B error: The influence of prior experience on the spatial memory abilities of two-year-olds. *Child Development*, 72(5), 1327–1346. <https://doi.org/10.1111/1467-8624.00351>
- Stephen, D. G., & Van Orden, G. (2012). Searching for general principles in cognitive performance: reply to commentators. *Topics in Cognitive Science*, 4(1), 94–102.
<https://doi.org/10.1111/j.1756-8765.2011.01171.x>

- Stiles, W. B., Leach, C., Barkham, M., Lucock, M., Iveson, S., Shapiro, D. A., Iveson, M., & Hardy, G. E. (2003). Early sudden gains in psychotherapy under routine clinic conditions: Practice-based evidence. *Journal of Consulting and Clinical Psychology*, 71(1), 14–21.
<https://doi.org/10.1037/0022-006X.71.1.14>
- Stone, A. A., Schneider, S., & Smyth, J. M. (2023). Evaluation of pressing issues in ecological momentary assessment. *Annual Review of Clinical Psychology*, 19, 107–131.
<https://doi.org/10.1146/annurev-clinpsy-080921-083128>
- Stone, M. (1977). An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1), 44–47.
- Stumpf, P. S., Arai, F., & MacArthur, B. D. (2021). Modeling stem cell fates using non-markov processes. *Cell Stem Cell*, 28(2), 187–190. <https://doi.org/10.1016/j.stem.2021.01.009>
- Sun, R., & Zhou, X. (2023). Differences in posttraumatic stress disorder networks between young adults and adolescents during the COVID-19 pandemic. *Psychological Trauma: Theory, Research, Practice, and Policy*, 15(Suppl 1), S29–S36.
<https://doi.org/10.1037/tra0001252>
- Sussmann, H. J., & Zahler, R. S. (1978). Catastrophe theory as applied to the social and biological sciences: A critique. *Synthese*, 37(2), 117–216.
- Takens, F. (1981). Detecting strange attractors in turbulence. In D. A. Rand & L.-S. Young (Eds.), *Dynamical systems and turbulence* (Vol. 898, pp. 366–381). Springer-Verlag.
<https://doi.org/10.1007/BFb0091924>
- Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, 67(6), 894–904.
<https://doi.org/10.1037/0022-006X.67.6.894>
- Thelen, E., & Smith, L. B. (1998). Dynamic systems theories. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development, Volume 1, 5th ed* (pp. 563–634). John Wiley & Sons Inc.
- Thom, R. (1975). *Structural stability and morphogenesis: An outline of a general theory of models*. W. A. Benjamin.
- Thompson, J. M. T., & Sieber, J. (2011a). Climate tipping as a noisy bifurcation: A predictive technique. *IMA Journal of Applied Mathematics*, 76(1), 27–46.
<https://doi.org/10.1093/imamat/hxq060>

-
- Thompson, J. M. T., & Sieber, J. (2011b). Predicting climate tipping as a noisy bifurcation: A review. *International Journal of Bifurcation and Chaos*, 21(02), 399–423.
<https://doi.org/10.1142/S0218127411028519>
- Thulin, M. (2023). *boot.pval: Bootstrap p-Values* [Manual].
<https://CRAN.R-project.org/package=boot.pval>
- Tong, H., & Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 42(3), 245–292.
- Tredicce, J. R., Lippi, G. L., Mandel, P., Charasse, B., Chevalier, A., & Picqué, B. (2004). Critical slowing down at a bifurcation. *American Journal of Physics*, 72(6), 799–809.
<https://doi.org/10.1119/1.1688783>
- Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment*, 21, 457–462.
<https://doi.org/10.1037/a0017653>
- Tschacher, W., & Haken, H. (2020). Causation and chance: Detection of deterministic and stochastic ingredients in psychotherapy processes. *Psychotherapy Research*, 30(8), 1075–1087. <https://doi.org/10.1080/10503307.2019.1685139>
- Vachon, H., Viechtbauer, W., Rintala, A., & Myin-Germeys, I. (2019). Compliance and retention with the experience sampling method over the continuum of severe mental disorders: Meta-analysis and recommendations. *Journal of Medical Internet Research*, 21(12), e14475.
<https://doi.org/10.2196/14475>
- van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific Reports*, 4(1), Article 1. <https://doi.org/10.1038/srep05918>
- van Borkulo, C. D., & Epskamp, S. (2023). *IsingFit: Fitting ising models using the ELasso method* [Manual]. <https://CRAN.R-project.org/package=IsingFit>
- van Borkulo, C. D., van Bork, R., Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2022). Comparing network structures on three aspects: A permutation test. *Psychological Methods*, 28(6), 1273–1285.
<https://doi.org/10.1037/met0000476>
- van de Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., van Nes, E. H., Viechtbauer, W., Giltay, E. J., Aggen, S. H., Derom, C., Jacobs, N., Kendler, K.

- S., van der Maas, H. L. J., Neale, M. C., Peeters, F., Thiery, E., Zachar, P., & Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*, 111(1), 87–92.
<https://doi.org/10.1073/pnas.1312114110>
- Van den Broeck, C., Parrondo, J. M. R., & Toral, R. (1994). Noise-induced nonequilibrium phase transition. *Physical Review Letters*, 73(25), 3395–3398.
<https://doi.org/10.1103/PhysRevLett.73.3395>
- van der Maas, H. L. J. (2024). *Complex-systems research in psychology*. The SFI Press.
- van der Maas, H. L. J., Kolstein, R., & van der Pligt, J. (2003). Sudden transitions in attitudes. *Sociological Methods & Research*, 32(2), 125–152. <https://doi.org/10.1177/0049124103253773>
- van Dongen, N., Finnemann, A., de Ron, J., Tiokhin, L., Wang, S., Algermissen, J., Altmann, E. C., Bahník, Š., Chuang, L.-C., Dumbravă, A., Fünderich, J. H., Geiger, S. J., Gerasimova, D., Golan, A., Herbers, J., Jekel, M., Kunnari, A., Lin, Y.-S., Moreau, D., ... Borsboom, D. (2025). Practicing theory building in a many modelers hackathon: A proof of concept. *Metapsychology*, 9. <https://doi.org/10.15626/MP.2023.3688>
- van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, 98(1), 3–53. <https://doi.org/10.1037/0033-295X.98.1.3>
- van Minnen, A., Hendriks, L., & Olff, M. (2010). When do trauma experts choose exposure therapy for PTSD patients? A controlled study of therapist and patient factors. *Behaviour Research and Therapy*, 48(4), 312–320. <https://doi.org/10.1016/j.brat.2009.12.003>
- van Geert, P. (2019). Dynamic systems, process and development. *Human Development*, 63(3–4), 153–179. <https://doi.org/10.1159/000503825>
- Vatiwutipong, P., & Phewchean, N. (2019). Alternative way to derive the distribution of the multivariate Ornstein–Uhlenbeck process. *Advances in Difference Equations*, 2019(1), 276. <https://doi.org/10.1186/s13662-019-2214-1>
- Viol, K., Schöller, H., Kaiser, A., Fartacek, C., Aichhorn, W., & Schiepek, G. (2022). Detecting pattern transitions in psychological time series – A validation study on the Pattern Transition Detection Algorithm (PTDA). *PLOS ONE*, 17(3), e0265335. <https://doi.org/10.1371/journal.pone.0265335>
- Visser, I. (2011). Seven things to remember about hidden Markov models: A tutorial on Markovian models for time series. *Journal of Mathematical Psychology*, 55(6), 403–415. <https://doi.org/10.1016/j.jmp.2011.08.002>

-
- Waddington, C. H. (1966). *Principles of development and differentiation*. Macmillan.
- Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, 11(1), 192–196. <https://doi.org/10.3758/BF03206482>
- Wagenmakers, E.-J., van der Maas, H. L. J., & Farrell, S. (2012). Abstract concepts require concrete models: Why cognitive scientists have not yet embraced nonlinearly coupled, dynamical, self-organized critical, synergistic, scale-free, exquisitely context-sensitive, interaction-dominant, multifractal, interdependent brain-body-niche systems. *Topics in Cognitive Science*, 4(1), 87–93. <https://doi.org/10.1111/j.1756-8765.2011.01164.x>
- Wallot, S., & Kelty-Stephen, D. G. (2018). Interaction-dominant causation in mind and brain, and its implication for questions of generalization and replication. *Minds and Machines*, 28(2), 353–374. <https://doi.org/10.1007/s11023-017-9455-0>
- Wang, C., Jiang, B., & Zhu, L. (2021). Penalized interaction estimation for ultrahigh dimensional quadratic regression. *Statistica Sinica*, 31(3), 1549–1570.
- Wang, H., Li, G., & Jiang, G. (2007). Robust regression shrinkage and consistent variable selection through the LAD-lasso. *Journal of Business & Economic Statistics*, 25(3), 347–355. <https://doi.org/10.1198/073500106000000251>
- Wang, J., Xu, L., & Wang, E. (2008). Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. *Proceedings of the National Academy of Sciences*, 105(34), 12271–12276. <https://doi.org/10.1073/pnas.0800579105>
- Wang, J., Zhang, K., Xu, L., & Wang, E. (2011). Quantifying the Waddington landscape and biological paths for development and differentiation. *Proceedings of the National Academy of Sciences*, 108(20), 8257–8262. <https://doi.org/10.1073/pnas.1017017108>
- Wang, L., Zhang, Z., McArdle, J. J., & Salthouse, T. A. (2008). Investigating ceiling effects in longitudinal data analysis. *Multivariate Behavioral Research*, 43(3), 476–496. <https://doi.org/10.1080/00273170802285941>
- Wang, S., Genugten, R. van, Yacoby, Y., Pan, W., Bentley, K., Bird, S., Buonopane, R., Christie, A., Daniel, M., DeMarco, D., Haim, A., Follet, L., Fortgang, R., Kelly, F., Kleiman, E., Millner, A., Obi-Obasi, O., Onnela, J. P., Ramlal, N., ... Nock, M. (2024). Idiographic prediction of suicidal thoughts: Building personalized machine learning models with real-time monitoring data. *Nature Mental Health*, 2, 1382–1391. <https://doi.org/10.1038/s44220-024-00335-w>

- Wang, S., Robinaugh, D., Millner, A., Fortgang, R., & Nock, M. K. (2023). *Mathematical and computational modeling of suicide as a complex dynamical system*. PsyArXiv.
<https://doi.org/10.31234/osf.io/b29cs>
- Watson, D., Clark, L., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology*, 97(3), 346–353.
<https://doi.org/10.1037/0021-843X.97.3.346>
- Weinans, E., Lever, J. J., Bathiany, S., Quax, R., Bascompte, J., van Nes, E. H., Scheffer, M., & van de Leemput, I. A. (2019). Finding the direction of lowest resilience in multivariate complex systems. *Journal of The Royal Society Interface*, 16(159), 20190629.
<https://doi.org/10.1098/rsif.2019.0629>
- Weinans, E., Quax, R., van Nes, E. H., & Leemput, I. A. van de. (2021). Evaluating the performance of multivariate indicators of resilience loss. *Scientific Reports*, 11, 9148.
<https://doi.org/10.1038/s41598-021-87839-y>
- Wichers, M., Groot, P. C., & Psychosystems, ESM Group, EWS Group. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*, 85(2), 114–116. <https://doi.org/10.1159/000441458>
- Wichers, M., Schreuder, M. J., Goekoop, R., & Groen, R. N. (2019). Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*, 49(3), 380–387.
<https://doi.org/10.1017/S0033291718002064>
- Wichers, M., Smit, A. C., & Snippe, E. (2020). Early warning signals based on momentary affect dynamics can expose nearby transitions in depression: A confirmatory single-subject time-series study. *Journal for Person-Oriented Research*, 6(1), Article 1.
<https://doi.org/10.17505/jpor.2020.22042>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.
<https://ggplot2.tidyverse.org>
- Wickham, H. (2019). Improving performance. In *Advanced R* (2nd edition, pp. 531–546). Chapman and Hall/CRC.
- Wickham, H., & Seidel, D. (2020). *scales: Scale functions for visualization*. <https://CRAN.R-project.org/package=scales>
- Wolfram Research, Inc. (2022). *Mathematica, Version 13.2*.
<https://www.wolfram.com/mathematica>

-
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R, Second Edition* (2nd ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315370279>
- Wright, A. G. C., Gates, K. M., Arizmendi, C., Lane, S. T., Woods, W. C., & Edershile, E. A. (2019). Focusing personality assessment on the person: Modeling general, shared, and person specific processes in personality and psychopathology. *Psychological Assessment, 31*(4), 502–515. <https://doi.org/10.1037/pas0000617>
- Wright, A. G. C., & Woods, W. C. (2020). Personalized models of psychopathology. *Annual Review of Clinical Psychology, 16*(1), 49–74.
<https://doi.org/10.1146/annurev-clinpsy-102419-125032>
- Wright, D. B., London, K., & Field, A. P. (2011). Using bootstrap estimation and the plug-in principle for clinical psychology data. *Journal of Experimental Psychopathology, 2*(2), 252–270. <https://doi.org/10.5127/jep.013611>
- Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress on Genetics, 1*, Article 8.
- Xiang, S., Romero, D. M., & Teplitskiy, M. (2025). Evaluating interdisciplinary research: Disparate outcomes for topic and knowledge base. *Proceedings of the National Academy of Sciences, 122*(16), e2409752122. <https://doi.org/10.1073/pnas.2409752122>
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18*(5), 459–482.
<https://doi.org/10.1002/cne.920180503>
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*(1), 32–35.
[https://doi.org/10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3)
- Zahler, R. S., & Sussmann, H. J. (1977). Claims and accomplishments of applied catastrophe theory. *Nature, 269*(5631), 759–763. <https://doi.org/10.1038/269759a0>
- Zeeman, E. C. (1976). Catastrophe Theory. *Scientific American, 234*(4), 65–83.
<https://doi.org/10.1038/scientificamerican0476-65>
- Zhang, X., Chong, K. H., & Zheng, J. (2020). A Monte Carlo method for in silico modeling and visualization of Waddington's epigenetic landscape with intermediate details. *Biosystems, 198*, Article 104275. <https://doi.org/10.1016/j.biosystems.2020.104275>

- Zhou, J. X., Aliyu, M. D. S., Aurell, E., & Huang, S. (2012). Quasi-potential landscape in complex multi-stable systems. *Journal of The Royal Society Interface*, 9(77), 3539–3553.
<https://doi.org/10.1098/rsif.2012.0434>
- Zhou, P., & Li, T. (2016). Construction of the landscape for multi-stable systems: Potential landscape, quasi-potential, A-type integral and beyond. *The Journal of Chemical Physics*, 144(9), 094109. <https://doi.org/10.1063/1.4943096>
- Zilcha-Mano, S., Errázuriz, P., Yaffe-Herbst, L., German, R. E., & DeRubeis, R. J. (2019). Are there any robust predictors of “sudden gainers,” and how is sustained improvement in treatment outcome achieved following a gain? *Journal of Consulting and Clinical Psychology*, 87(6), 491–500. <https://doi.org/10.1037/ccp0000401>
- Zou, Y., Donner, R. V., Marwan, N., Donges, J. F., & Kurths, J. (2019). Complex network approaches to nonlinear time series analysis. *Physics Reports*, 787, 1–97. <https://doi.org/10.1016/j.physrep.2018.10.005>

Dutch Summary (Nederlandse Samenvatting)

Psychologische processen resulteren uit een groot aantal biologische, psychologische en sociale factoren die op complexe wijze in de tijd op elkaar inwerken. Om de menselijke geest en psychische stoornissen beter te kunnen bestuderen, is het van essentieel belang om deze complexiteit te erkennen en te begrijpen. In de afgelopen jaren zijn steeds meer onderzoekers begonnen met het toepassen van het perspectief van complexe dynamische systemen. Hoewel veel concepten uit de complexiteitstheorie hun weg hebben gevonden naar de gedragswetenschappen, is verdere technische ontwikkeling nodig om conceptuele ideeën beter te kunnen verbinden met kwantitatieve toepassingen. In dit proefschrift ontwikkel ik formele analysetools om complexiteit in de gedragswetenschappen beter te begrijpen. Mijn werk richt zich specifiek op drie belangrijke onderwerpen: het kwantificeren van de stabiliteit van een psychologisch systeem met behulp van potentieellandschappen, het bestuderen van abrupte veranderingen en hun voorspellers in psychologische systemen, en het beschrijven van de niet-lineaire dynamiek van psychologische systemen.

Deel 1. Methoden voor stabiliteitslandschappen

Complexe interacties in een dynamisch systeem kunnen leiden tot verschillende fasen met uiteenlopende kenmerken (bijvoorbeeld een depressieve fase en een gezonde fase). Vaak blijft het systeem in de buurt van één zo'n fase, maar onder bepaalde voorwaarden kan het overgaan naar een andere. In theoretische artikelen wordt dit vaak verbeeld met de bal-in-landschap-metaphoer (zie Figuur 1 in Hoofdstuk 2), waarbij de bal de toestand van het systeem represeneert en het landschap de stabiliteit van die toestanden. Kwantitatieve methoden om stabiliteitslandschappen van psychologische systemen daadwerkelijk te berekenen ontbraken echter nog.

In **Hoofdstuk 2** beschrijf ik een nieuwe methode om stabiliteitslandschappen van formele modellen te verkrijgen. Ik laat me daarbij inspireren door biologische modellen en maak gebruik van de stationaire verdeling van modelsystemen. Als voorbeeld pas ik de methode toe op een formeel model van paniekstoornis. Dit model heeft twee fasen: een gezonde fase en een paniekfase. Hoe stabieler de paniekfase, hoe ernstiger de paniekstoornis. Ik laat zien hoe onze methode gebruikt kan worden om de stabiliteit van verschillende psychologische fasen te illustreren, de invloed van modelparameters te analyseren en modellen te verbeteren.

In **Hoofdstuk 3** geef ik verdere technische details van de methode uit Hoofdstuk 2 en introduceer ik het R-pakket *simlandr*, waarin deze methode geïmplementeerd is. De analyse bestaat uit drie stappen: (1) het model simuleren tot het is geconvergeerd, (2) de stationaire

verdeling schatten met een kernel-benadering en deze omzetten naar een potentieelfunctie, en (3) de *minimum energy path* (een ‘weg van de minste weerstand’ in het landschap) berekenen om de barrièrehoge tussen fasen te verkrijgen. Deze hoogte geeft aan hoe moeilijk het is voor het systeem om overgangen te maken. Ik illustreer de methoden van het pakket met een biochemisch en een psychologisch modelsysteem.

In **Hoofdstuk 4** beschrijf ik een methode om stabiliteitslandschappen te berekenen op basis van empirische data, zoals data verkregen via herhaalde zelfrapportages van deelnemers. De methode werkt als volgt. Eerst schatten we een vectorveld dat de veranderingstendens van het systeem op verschillende punten weergeeft. Hiervoor gebruiken we een niet-lineaire schattingsmethode om complexe dynamische patronen goed te kunnen vastleggen. Vervolgens gebruiken we de methode uit Hoofdstukken 2 en 3 om een potentieellandschap op te bouwen op basis van het vectorveld. Ik illustreer deze aanpak met gesimuleerde en empirische datasets.

In **Hoofdstuk 5** beschrijf ik een methode om stabiliteitslandschappen af te leiden uit Ising-netwerken, een type netwerkmodel dat gebaseerd is op binaire symptoomdata. Het aantal actieve symptomen (de knopen in het netwerk) wordt in dit netwerkmodel gezien als een maat voor de ernst van de stoornis. Ik maak gebruik van een intrinsieke energiemaat van Ising-netwerken, de Hamiltoniaan, en vat deze samen om een stabiliteitslandschap op te stellen in relatie tot het aantal actieve knopen. Daarnaast introduceer ik stabiliteitsmaten om de stabiliteit van zowel de gezonde als de pathologische fase te kwantificeren. Ik laat zien hoe netwerkconnectiviteit en knoopdremels (*thresholds*) de stabiliteit beïnvloeden, en hoe groepen hiermee vergeleken kunnen worden.

Deel 2. Vroege waarschuwingssignalen en transities

In klinische trajecten worden vaak abrupte veranderingen waargenomen, zowel verbeteringen als verslechtering. Eerdere onderzoekers hebben dergelijke abrupte veranderingen geïnterpreteerd als kritieke transities in niet-lineaire dynamische systemen. Vanuit eerder onderzoek is gehypothetiseerd dat de stijging van een aantal statistische indicatoren, zoals variantie en autocorrelatie, kan dienen als een vroeg waarschuwingssignaal voor een transitie. De empirische resultaten zijn echter wisselend. In dit deel verdiep ik mij in de theorie van kritieke transities en onderzoek ik hoe deze theorie kan bijdragen aan het begrijpen van abrupte veranderingen in de klinische psychologie.

In **Hoofdstuk 6** bespreek ik de wiskundige theorie achter vroege waarschuwingssignalen en identificeer ik drie belangrijke aannames voor het detecteren ervan: (1) het systeem moet

vertrekken vanuit een puntattractor, (2) de waarschuwingssignalen moeten verschijnen vóór de transitie, en (3) waarschuwingssignalen en transities moeten zich voordoen in dezelfde variabelen. Op basis van deze aannames evaluateer ik gangbare onderzoeksmethoden en identificeer ik een aantal inconsistenties met de theorie, wat kan verklaren waarom empirische resultaten tegenstrijdig zijn. Ik geef ook theorie gestuurde aanbevelingen voor toekomstig onderzoek.

In **Hoofdstuk 7** neem ik een breder perspectief om de verschillende typen transities in dynamische systemen te onderzoeken en bespreek ik hoe deze getypeerd kunnen worden in een klinische context. Hoewel er veel factoren zijn die tot een transitie kunnen leiden, zijn er op abstract niveau een beperkt aantal mechanismen met unieke kenmerken te onderscheiden. Ik introduceer vier soorten veranderingen uit de dynamische systeemtheorie: bifurcatie-geïnduceerde kantelingen (B-tipping), ruis-geïnduceerde kantelingen (N-tipping), snelheid-geïnduceerde kantelingen (R-tipping), en ruis-geïnduceerde diffusie (N-diffusion). Aan de hand van twee klinische scenario's laat ik zien waarom het belangrijk is om het niveau van het systeem in de onderzoeksvervraag te specificeren om het type verandering te kunnen bepalen. Ik sluit af met praktische suggesties voor empirisch onderzoek.

Deel 3. Niet-lineaire dynamiek

Psychologische variabelen beïnvloeden elkaar vaak op niet-lineaire wijze, wat aanleiding geeft tot vele interessante fenomenen die in eerdere delen van dit proefschrift besproken zijn. Het modelleren en begrijpen van deze niet-lineaire interacties is echter moeilijker dan in het lineaire geval, en vereist geavanceerdere methoden en technische zorgvuldigheid. In dit deel behandel ik dit probleem vanuit twee invalshoeken: niet-lineaire schatting op basis van empirische data en analysemethoden voor computationele modellen.

In **Hoofdstuk 8** introduceer ik een nieuw niet-lineair model voor individuele tijdreeksdata: het kwadratische vectorautoregressieve model. Dit model breidt het lineaire model uit door ook kwadratische termen toe te voegen aan de regressie. Een specifieke regularisatiemethode wordt toegepast om overfitting te voorkomen, en een linearisatiemethode wordt ingezet om de interpretatie van resultaten te vergemakkelijken. Aan de hand van simulaties en empirische datasets laat ik zien dat het model, ondanks een matige algemene prestatie, in staat is belangrijke niet-lineaire verbanden op te sporen. Het model is daarom geschikt voor exploratief onderzoek.

In **Hoofdstuk 9** richt ik mij op formele dynamische modellen en bespreek ik twee grafische analysemethoden: fasevlakanalyse en bifurcatieanalyse. Met fasevlakanalyse kunnen stabiele

toestanden en veranderingsdynamiek in kaart worden gebracht; met bifurcatieanalyse kunnen we systematisch de invloed van modelparameters onderzoeken. Aan de hand van een paniekstoornismodel en een model van suïcidale gedachten laat ik zien hoe deze methoden gebruikt kunnen worden om stabiliteit en invloedrijke factoren te analyseren in psychologische systemen.

Reflectie en toekomstige richtingen

In **Hoofdstuk 10** sluit ik af met een algemene reflectie op complexiteitsonderzoek in de gedragswetenschappen en een blik op toekomstige ontwikkelingen. Ondanks het feit dat complexe systemen vaak gedeelde kenmerken hebben, kunnen ze onderling sterk verschillen. Deze heterogeniteit vereist voorzichtigheid bij het kiezen van geschikte analysemethoden en bij het overnemen van concepten uit andere disciplines. Werken met het complexe-systeem-perspectief in de psychologie is uitdagend, mede door de vaagheid van kernconcepten en meetmethoden. Tegelijkertijd biedt dit perspectief unieke kansen, zoals het combineren van verschillende methoden. Toekomstig werk zou zich kunnen richten op het systematiseren van beschikbare methoden, het ontwerpen van studies gebaseerd op complexiteitstheorie, en op meer interdisciplinaire samenwerking.

List of Publications

Peer-Reviewed Articles

- Cui, J., Olthof, M., Lichtwarck-Aschoff, A., Li, T., & Hasselman, F. (in press). simlandr: Simulation-based landscape construction for dynamical systems. *The R Journal*. Preprint available at PsyArXiv: <https://doi.org/10.31234/osf.io/pzva3>
- Cui, J., Olthof, M., Hasselman, F., & Lichtwarck-Aschoff, A. (2025). Examining the research methods of early warning signals in clinical psychology through a theoretical lens. *BMC Psychiatry*, 25, 261. <https://doi.org/10.1186/s12888-025-06688-5>
- Cui, J., Hasselman, F., Olthof, M., & Lichtwarck-Aschoff, A. (2025). Understanding types of transitions in clinical change: An introduction from the complex dynamic systems perspective. *Journal of Psychopathology and Clinical Science*, 134(4), 469–482.
<https://doi.org/10.1037/abn0000991>
- Cui, J., Hasselman, F., & Lichtwarck-Aschoff, A. (2023). Unlocking nonlinear dynamics and multistability from intensive longitudinal data: A novel method. *Psychological Methods*.
<https://doi.org/10.1037/met0000623>
- Cui, J., Lichtwarck-Aschoff, A., Olthof, M., Li, T., & Hasselman, F. (2023). From metaphor to computation: Constructing the potential landscape for multivariate psychological formal models. *Multivariate Behavioral Research*, 58(4), 743–761.
<https://doi.org/10.1080/00273171.2022.2119927>

Preprints under Review/Revision

- Cui, J., Wagenmakers, D., van Doorn, G.S., Hasselman, F., & Lichtwarck-Aschoff, A. (2025). *Analyzing formal dynamic models in psychology: A tutorial using graphical tools*. PsyArXiv.
https://doi.org/10.31234/osf.io/fhstw_v1
- Hasselman, F., Olthof, M., Cui, J., & Lichtwarck-Aschoff, A. (2024). *A Recursive partitioning approach for detecting nonstationarity of levels and trends in time series data*. PsyArXiv.
<https://doi.org/10.31234/osf.io/asv5f>
- Cui, J., Lichtwarck-Aschoff, A., & Hasselman, F. (2024). *Examining the feasibility of nonlinear vector autoregressions for psychological intensive longitudinal data*. PsyArXiv.
<https://doi.org/10.31234/osf.io/5s8jc>
- Cui, J.*, Lunansky, G.*., Lichtwarck-Aschoff, A., Mendoza, N., & Hasselman, F. (2023). *Quantifying the stability landscapes of psychological networks*. PsyArXiv.
<https://doi.org/10.31234/osf.io/nd8zc>

Non-Peer-Reviewed Articles

Cui, J., Lichtwarck-Aschoff, A., & Hasselman, F. (2023). *Comments on “Climbing Escher’s stairs: A way to approximate stability landscapes in multidimensional systems”*. arXiv.
<https://doi.org/10.48550/arXiv.2312.09690>

Open-Source R Packages

Cui, J. (2025). *quadVAR: Quadratic Vector Autoregression* (Version 0.1.2) [R package]. Comprehensive R Archive Network (CRAN).

<https://doi.org/10.32614/CRAN.package.quadVAR>

Cui, J. (2024). *NVAR: Nonlinear Vector Autoregression Models* (Version 0.1.0) [R package]. Comprehensive R Archive Network (CRAN).

<https://doi.org/10.32614/CRAN.package.NVAR>

Cui, J. (2023). *Isinglandr: Landscape Construction and Simulation for Ising Networks* (Version 0.1.1) [R package]. Comprehensive R Archive Network (CRAN).

<https://doi.org/10.32614/CRAN.package.Isinglandr>

Cui, J. (2023). *fitlandr: Fit Vector Fields and Potential Landscapes from Intensive Longitudinal Data* (Version 0.1.0) [R package]. Comprehensive R Archive Network (CRAN).

<https://doi.org/10.32614/CRAN.package.fitlandr>

Cui, J. (2022). *SparseVFC: Sparse Vector Field Consensus for Vector Field Learning* (Version 0.1.2) [R package]. Comprehensive R Archive Network (CRAN).

<https://doi.org/10.32614/CRAN.package.SparseVFC>

Cui, J. (2021). *simlandr: Simulation-Based Landscape Construction for Dynamical Systems* (Version 0.4.0) [R package]. Comprehensive R Archive Network (CRAN).

<https://doi.org/10.32614/CRAN.package.simlandr>

* These authors contributed equally.

About the Author

Jingmeng Cui was born on February 12th, 1997, in Shijiazhuang, Hebei, China. In 2019, he received his bachelor's degrees in Material Chemistry and Psychology from Peking University, Beijing, China. He moved to the Netherlands afterward and completed a Research Master's degree in Behavioural Science (cum laude) in 2021 at Radboud University, Nijmegen.

Since 2021, he has conducted his PhD research at the Nieuwenhuis Institute of Educational Research, University of Groningen, supported by the university's PhD scholarship program. During his PhD project, Jingmeng was an active member of the inter-university research group Complexity in Behavioural Science, and served as the group's coordinator for over 3 years. He has also participated in various research collaborations, such as the Dynamical Network and Time Series Models (DynaNeT) meetings and the Measurement Is The New Black (MITNB) consortium. He has presented his work at various international conferences, such as the Conference on Complex Systems, the International Convention of Psychological Science, and the Society for Ambulatory Assessment Conference. Jingmeng has also contributed to a range of courses, including bachelor's and master's-level classes, as a lecturer, guest lecturer, and teaching assistant for computer practicals.

In December 2025, Jingmeng will join the Research Group of Quantitative Psychology and Individual Differences at KU Leuven, Belgium, as a postdoctoral researcher.



Acknowledgments

First of all, this project couldn't have been completed without my supervisors, Anna and Fred. Since my master's major project, you have not only been my supervisors but also my role models for what it means to be professional researchers and great supervisors, and the past years have definitely proved how much chemistry I have with both of you!

Anna, I still remember the first meeting we had about my Master's thesis, which was on the day right before the COVID lockdown, and I'm so glad we had that one in person before everything had to go remote. When I moved with you to Groningen, you showed Rineke and me around the maze of faculty buildings, invited us to your home where we saw your boy coming out of the swimming pool, made beautiful dishes during the writing week that I can't even name, and organized a huge party after your oratie. You never felt like a serious professor with a poker face to me; instead, you're such a nice person that I always felt comfortable chatting with you about anything. But as a (cool) professor, you also know research so well. You can always think along and provide helpful suggestions, tell me clearly what can be improved, and help me keep the whole thing structured while still giving me a lot of room to explore. You have always been so supportive through all the ups and downs of my PhD. And your passion for pursuing the research direction you believe is meaningful has always inspired me. Thank you for your guidance, support, and everything!

Fred, I remember seeing you for the first time at a seminar shortly after I arrived in the Netherlands, and you were one of the first people I recognized from an online photo. Your course introduced me, like many others, to the world of complex systems. We didn't see each other that often during my PhD years, as half the vast landscape of the Netherlands lies between Groningen and Nijmegen. Still, you appeared on my screen each week in a Hawaiian shirt, either from your metal music studio or a Cyberpunk virtual world. I've always enjoyed our in-person meetings, from onsite group meetings to conferences, from the party at Merlijn's home to the Pho restaurant in Leiden. You always have great stories to tell, whether they're about complexity, behavioral science, or something completely unexpected. In research, you are both an idea generator and a guardian of principles. Whatever topic I brought up, you always had something worth looking into, and your insistence on not overlooking important aspects has always been inspiring to me. Thank you so much!

Next, I would like to express my gratitude to the assessment committee of my thesis, **Prof. Ralf Cox**, **Prof. Denny Borsboom**, and **Prof. Eva Ceulemans**. It is a great honor for me to have

you on the committee, especially as you have long been academic role models of mine. Thank you for taking the time to assess my thesis and provide your evaluations.

Then, I would like to thank all my colleagues. I would start with my colleagues in the Complexity in Behavioral Science (CiBS) group. **Merlijn**, my co-supervisor for my Master's thesis, my third major co-author during the PhD, and my academic brother (I would literally call you this if we were speaking Chinese), you've supported me from the very beginning, starting with my PhD application proposal, all the way to the end, with the Dutch summary of this thesis. Your own PhD work has always been a lighthouse for me, shaping my idea of what a great dissertation should be. Beyond that, your philosophical perspective on research and your human movement pattern at parties have both inspired me in different ways. **Rineke**, my officemate all this time and my academic sister, you've been the colleague I've spent the most time with at work over these past few years. From our countless morning chats, lunch talks, and tea breaks, we've shared both the joys and the struggles of this PhD journey. You brought not only ideas from clinical practice and qualitative research into the office, but also warmth, encouragement, microwave walks, and plenty of tasty gluten-free cookies, all of which I've enjoyed so much. **Daan**, we didn't meet offline that often, and mostly we see each other at the CiBS virtual meeting room, but it was a pleasure every time we had a chance to chat. Your work, especially the case studies and idiographic research grounded in clinical practice, has given me many insights and intuition about the real-world side of things. Also, special thanks for being one of my paronyms! It's really great to have someone as reliable as you by my side on this important day. **Andrea**, thank you for the contagious energy you always bring to me and the whole CiBS group! You came up with so many great ideas in the past years, from the handmade landscape to the group logo. Working with you has always been both inspiring and a lot of fun. **Edmund**, I've always been impressed by your enthusiasm for research on emotion, and I've really enjoyed our conversations, not only because of the interesting stories you share, but also because it's wonderful to have someone who understands even the subtleties of Chinese expressions (despite the small gap between Mandarin and Cantonese)! **Freek**, the talkative philosopher in the group. You have always been a unique thought-provoker to me ever since your course in my research master's. Thank you for all the food for thought you provided, and I also enjoyed our chess time and your pasta-making crash course a lot. **Dominique**, I knew you from the SMR meetings in Nijmegen before we left in different (geographical) directions. You have such a recognizable personality, warm and full of enthusiasm (just like the color of your hair), that each time it's great to see you and work together. Your work in Track Your Mood and response processes is very enlightening to me. And to **Anna B, Benjamin, Matti, Maarten**,

Nessa, Leonie, Oliver, Rafael, Yingjie, and many other colleagues from CiBS, thank you for all the ideas, help, and support. I've really enjoyed being part of the CiBS group with you!

Gaby L, Prof. Tiejun Li, Dieta, Sander, Norman, Lea, thank you for collaborating with me on various research projects during my PhD, and for bringing your expertise in methodology, data collection, and research ideas, as well as your creativity and open minds from across different disciplines. I've learned a great deal through working with you, and I've really enjoyed not only our research collaborations but also each meeting and interaction we've had along the way. **Dr. Willem Frankenhuys and Dr. Eiko Fried**, thank you for giving me the opportunity to present some of my preliminary research ideas in your lab groups and for sharing your valuable suggestions. I would also like to thank all the **researchers dedicated to open science** who generously shared their code and data. This thesis would not have been possible without their contributions.

And of course, I would like to thank all my lovely colleagues in Groningen. Although many of you work in very different disciplines from mine, your variety of perspectives has broadened my thinking so much, and you've always made me feel at home here. **Arjen**, it's been great having you as an office neighbor who has always been kind, helpful, and considerate. Thank you for all your help, from the small PhD practicals to your valuable suggestions on my manuscripts. **Anne-Marie**, I really enjoyed our (at least) once-a-year talks. Thank you for being such a kind listener and for your care for me and so many other PhD students. **Rodrigo**, thank you for your warm welcome when I first arrived in Groningen. The plant you gave us (I only recently learned it's called Chinese Money Plant) accompanied me for four years in the office. **Ymke, Lysbert, Lise, Lotte, Maartje, Kjell, Gaby MJ, Sarai, Wendy, Katrien, and many other colleagues at Ortho and/or PedOn**, thank you for organizing and coming together with our *gezellige* group meetings, teaching days, lunches, walks, and all those fun activities. **Laura Bringmann, Yong**, thank you for your valuable insights from a statistical perspective, which helped me a lot. **Ruud, Niklas, and Jur**, thank you for all the sporty science you've shared. **Marijtje, Wendy, Arjen, Yong, Niek, Muirne, Letty, Klaas, and Laura Bringmann**, it was great to teach with you for all these courses. Thank you for all your guidance and collaboration. **Miranda, Cilia, Susanne, Dianne, Ingrid, Margreet**, I feel very fortunate to have been supported by such a kind, helpful, and cheerful group of secretaries and coordinators in our research unit throughout my PhD. Thank you all for your patience and assistance. I would also like to thank the **PhD support staff and trainers** at the university, those who organized workshops, well-being sessions, and health checks, and those who maintained the workspace and computational resources. Your work is deeply appreciated.

My friends in our Groningen “only 4 people here are worrying about their theses” group, **Zhe** (“Bro. Dong”), **Jinghao** (“Haohao”), and **Mingmin** (“Xiaoming”), we had so much fun together in the past years, cooking together, traveling around, crocheting, and playing Sichuan Mahjong. We will no longer stay in the same town, yet no matter where you go, I wish you all the best with graduation and everything afterwards. Also, special thank you to Zhe and Mingmin, my paronymph and photographer for the defense, it’s great to have you around in this “no longer worrying about my thesis” time. **Zijun**, thank you for inviting me to your home board game days! Our murder mystery games, as well as the experience sampling discussion after that, have been a lot of fun. **Shuai** (“Bro. Shuai”), thank you for always leading the way in the Netherlands, from research career to fun activities. I’ve always admired not only your passion for research but also the lively and interesting spirit you bring to everything. **Kun**, my old friend in Nijmegen, we didn’t see each other so much in the past years, but each time we met, it was always a joy to catch up and chat. **Zhiyuan** (“jls”), **Yiwen** (“Yvan”), I’m really glad to have had you as my “local guides” in Noord- and Zuid-Holland. Thank you for taking me on all those fun outings and inviting me to stay over and enjoy your amazing dishes. I’ve really enjoyed the time with you! And **Luojian** (“sls”), **Shuang** (“tyls”), **Yingjian** (“lls”), **Linyun** (“Rls”), it was really great to have all the chats and explore so many different places in the Netherlands with you. To my teammates from **Donitas volleyball teams H9 (2023) and H8 (2024)**, the **SPR badminton group**, and coach **Rik**, it was a real joy to play sports with you! Thank you for keeping me active and also for all the Dutch culture crash courses that came along.

To my friends who are further away. We don’t have a lot of time to meet in person, but it is a lot of fun each time I meet you or travel with you. **Mingdi** (“dd”), **Qiliang** (“Black Kitten”), **Yueyu** (“Fish Queen”), **Yanxing** (“Arctic Fox”), although we’ve gone our separate ways from Beijing, we’ve shared so many adventures over the past years, from Iceland to Spain, America, and beyond, as well as all the fun chats and price guesses we have even when we are apart. I’ve truly enjoyed our moments together! **Junren** (“Little Dino”), **Jiaming** (“kk”), **Zichun** (“Changhe”), **Zhe** (“zls”), **Shuchang** (“Haji”), **Jiahao**, it was great to talk with you about all the things, from everyday moments to life’s bigger questions. Every time I got the chance to go back to China, it was wonderful to have a reunion, go out for meals, sing karaoke, and play board games. Even though we’re in different places now, I really hope we’ll have the chance to catch up and have a good chat whenever we find ourselves in the same place again. Also, many other online friends from “Boy@BDWM”, “Boy@BDWM Pass the Pot”, “Zone G, College of Chemistry”, “Neo-Physics Institute, School of Physics”, and “Black Kitten’s Fan Club” (yes, we have many creative nicknames and group names), as well as all my other friends who’ve

shared moments and chats along the way, thank you for all your funny quotes, interesting stories, and for accompanying and cheering me on throughout my journey.

I would like to thank the **FanC club** for organizing all the fun murder mystery and board game sessions, and **Feiwen** (sosad.fun), a Chinese literature website, for providing a friendly and inspiring writing platform full of creative ideas. Special thanks also to the singer **Zhou Shen**. I've been listening to your songs for more than 10 years, and it was fascinating to attend your concert in Nanjing in 2024. Thank you for your music that accompanied me through so many happy and sad times during my journey, for always reminding me to take care of myself, and for showing through your career path that sticking to your interests and expertise with diligence will eventually pay off.

Mom, Dad, as well as my extended family, especially my **uncle, aunt, and Xuanxuan**, thank you for your love, encouragement, and support throughout these years. I feel so lucky to have a family that always trusts my choices and abilities, supports me wherever I want to go, and reminds me that no matter the difficulty, there are always dear people I can rely on and a safe home base I can return to after any adventure. Your presence, whether close by or far away, has always been a great source of strength. Thank you.

My *liefste Haoming*, we've been together for six years! During my journey in a foreign land, you've always been my steady light through all seasons, a tree hole for sharing little joys and falls, and the attractor of my heart. Even though we've spent much of our time apart, every time we meet, in Beijing, Shijiazhuang, Hengyang, Groningen, and so many other places, it's full of happiness. With every moment we share and every challenge we face, I feel more certain about us. Thank you for always being there for me. With you by my side, there's nothing I fear.