Anticipating Critical Transitions with Nonlinearity, Periodicity and Heterogeneity

by

Shiyang Chen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in The University of Michigan
2018

Doctoral Committee:

Professor Bogdan Epureanu, Chair
Professor Carlos E. S. Cesnik
Professor Noel Perkins
Assistant Professor Bogdan Popa

Shiyang Chen

shychen@umich.edu

ORCID iD: 0000-0001-8974-3705

To my parents and my wife,

for their unconditional love and support.

ACKNOWLEDGEMENTS

Growing up in a small town in China, I have no idea what academic life is like. Till this day, I find it hard to explain my everyday life to my relatives, and my childhood friends. Four years ago, I stumbled into this new world out of curiosity, and it has been a truly incredible journey. I wouldn't be able to make it this far without the help of many people.

Professor Bogdan Epureanu took me in as a research assistant when I had little idea how research was done. He has been a great advisor ever since. I want to thank him for his guidance when I feel lost, and his patience when I wondered into directions that were not optimal. His rigorous attitude towards research will always be an inspiration for me. I would also like to thank my committee members, Professors Cesnik, Perkins and Popa for their valuable comments and suggestions.

I want to thank Dr. Caihao Weng and Dr. Xuning Feng for their guidance and friendship during the first year of my PhD life. I wouldn't be able to survive without those poker and video games after dinner, although their gaming skills are not remotely close to their ability to do research. I want to extend my gratitude also to my fellow advisees from the Applied Nonlinear Dynamics of Multi-Scale Systems Laboratory: Mainak Mitra, Jauching Lu, Xingyu Li, Yiqian Gan, Amin Ghadami, Ehsan Mirzakhalili, Andrea Lupini, Kejie Chen, Zhiwei Liu, and my friends at Michigan: Daning Huang, Xiwei Xia, Ruizw Hu, Yu Mao, Zeng Qiu, Yang Chen and many others.

Finally, I'd like to thank my family for all the love and support they have given

me. To my parents for encouraging and supporting me to see the bigger world and search my own meaning, to my grandparents for instilling kindness in me, and to my wife Jiaomin, for being my source of joy and pillar of strength.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

Many natural and engineering systems may switch abruptly from one stable state to another due to a small perturbation to the system's state or a small change in the underlining conditions. In ecosystems, for example, extinctions of species or desertification can occur rapidly. Therefore, critical transitions can be dangerous to a number of systems, and it could be very beneficial if monitoring or early warning methods were available while the system is still in the healthy regime. The approach of critical transitions in many natural and engineering systems is accompanied by a phenomenon called critical slowing down. Theoretical and experimental studies have suggested that responses to small perturbations become increasingly slow when these systems are near critical transitions. Statistics such as variance, autocorrelation calculated from time series data have been proposed as early warning signals to anticipate the system's approach to a transition point.

The problem of anticipating critical transitions becomes more complicated when other factors come into play. Factors such as nonlinearity, periodicity and heterogeneity can alter the behavior of the system, and thus affect the applicability of generic early warning signals. This thesis examines the effect of these factors on the critical transition of a system, and develops new data-driven approaches accordingly. To deal with and exploit the existence of nonlinearity in the system, recoveries from large instead of small perturbations are used to calculate the recovery rates of the system versus amplitudes. Under the circumstances of periodicity, recovery rates are calculated discretely via the Poincaré section. Using experimental and computational data, we show that a combination of using recoveries from large perturbations and

calculating recovery rates using the Poincaré section can be highly effective in terms of anticipating critical transitions for systems with parametric resonance. Moreover, this thesis develops new early warning signals for spatially extended systems based on the eigenvalues of the covariance matrix. We mathematically show that the dominance of the largest eigenvalue of the covariance matrix can be used as an early warning signal by establishing the relationship between the eigenvalues of the covariance matrix and the eigenvalues of the force matrix. This new set of early warning signals are especially useful when the system has strong spatial heterogeneity. Lastly, this thesis investigates the influence of the choice of hyper-parameters, such as moving window size, sample rate, detrending methods, on the robustness of several early warning signals. General rules regarding data preparation and hypothesis testing are proposed.

CHAPTER I

Introduction

1.1 Motivation

Many natural and engineering systems may switch abruptly from one stable state to another [5, 6, 7, 8] due to a small perturbation to the system's state or a small change in the underlining conditions. In ecosystems, for example, extinctions of species or desertifications can occur rapidly [9]. Other such examples include systemic market crashes [10], spontaneous medical failures [11, 12], ship capsizings [13] and so on. It is notably hard to predict critical transition because the state of the system may show little change prior to the transition due to small changes in conditions. Such critical transitions are in contrast to changes that are a result of large external shocks.

In theory, critical transitions are often studied in terms of bifurcations of dynamical systems [14, 15]. Bifurcation theory provides a useful way to consider the dynamics of the system near critical transitions. Research on critical transitions have shown the clear relationship between some rapid regime shifts from one branch to another and saddle-node bifurcation [16]. It is also suggested that some piecewise continuous transitions are associated with a transcritical bifurcation or a pitchfork bifurcation [17, 18]. Studying critical transitions from the point of view of bifurcation theory provides us with a key insight: a system will lose resilience close to the

bifurcation in the sense that it takes much longer for the system to recover from perturbations. This phenomenon is referred to as critical slowing down [19]. The critical slowing down phenomenon can be understood through linear stability analysis. The recovery rate of a system from small perturbations are characterized by the eigenvalues of the linearized system [20]. The system is stable when all the eigenvalues have negative real parts. This condition is violated when a bifurcation occurs. As a system approaches a bifurcation point, the real part of the dominant eigenvalue becomes increasingly close to 0, resulting in an increasingly slower recovery from perturbations. A number of indicators based on critical slowing down have been proposed to forecast critical transitions [5, 21, 7, 8, 22, 23]. Here we list a few of them that received most attention:

1. Recovery rate from perturbations becomes smaller.

2. Variance of the system's response to continuous stochastic excitations increases.

3. Autocorrelation of the system's response to continuous stochastic excitations increases.

These generic early warning signals have been proven useful in a broad range of empirical studies [16]. Nonetheless, major challenges remain in the robustness of early warning signals in face of real world complications. First, generic early warning signals typically rely on using a system's response to small random perturbations to infer its linear stability. However, most natural and engineering systems are nonlinear. Previous studies have shown that certain early warning signals only work with a limited type of nonlinearity, while others can be misleading when some parameters also have stochastic components [24]. Second, many systems are influenced by periodic fluctuations [1]. Some critical transitions can be triggered by a synchronization between a natural mode of the system and the periodic forcing or parameter variation [25, 8, 26, 27, 28, 20]. These types of critical transitions are especially hard

to forecast, because different time scales are involved in the dynamics, and critical slowing down is hard to detect. Third, real systems usually are spatially heterogeneous [29, 30]. Generic temporal early warning signals have been shown to have limitations for systems showing spatial patterns [31] due to oversimplification. Some spatial indicators - spatial variance, spatial skewness, spatial spectral density - are shown to be valid counterparts to temporal indicators [18, 22, 32]. These spatial indicators, however, only look at one snapshot at one time, thus limiting the information they can gather from the system. It is hard to find the (dynamical) pattern that is associated with the critical transition using only a single snapshot.

This thesis aims at developing new early warning signals and algorithms that can bridge these gaps. The first part of this thesis is dedicated to forecast both the bifurcation point and the bifurcation diagram of critical transitions caused by parametric resonance. Parametric resonance is a phenomenon that can occur when one or more parameters vary periodically in time at a frequency related to the resonance frequencies of the system. Such resonances are observed in a wide range of engineering and natural systems [59, 72, 73] where large amplitude parametric resonance can be triggered by a relatively small-amplitude oscillation in parameters. Two different methods were developed using recoveries in the pre-bifurcation regime. The applicability of these methods to both simulation and experimental data are investigated. In the second part of this thesis, a set of new early warning signals, namely the largest eigenvalue of the covariance matrix and the percentage it accounts for of the total variation, are proposed for high dimensional, heterogeneous systems.

1.2 Background

1.2.1 Generic early warning signals

A growing amount of literature has reported the observation that some systems undergo a critical transition from one stable state to another [9, 30, 33]. This type of transition is often the result of some gradually changing conditions, such as climate, ground water reduction, harvesting of certain species, and so on [34, 35]. An example of the bifurcation diagram of a system subject to critical transition can be found in Fig. 1.1. As the underlining conditions change, small variations in the mean state variables of the system are observed. However, the system can jump dramatically from one branch to another after passing some tipping points.

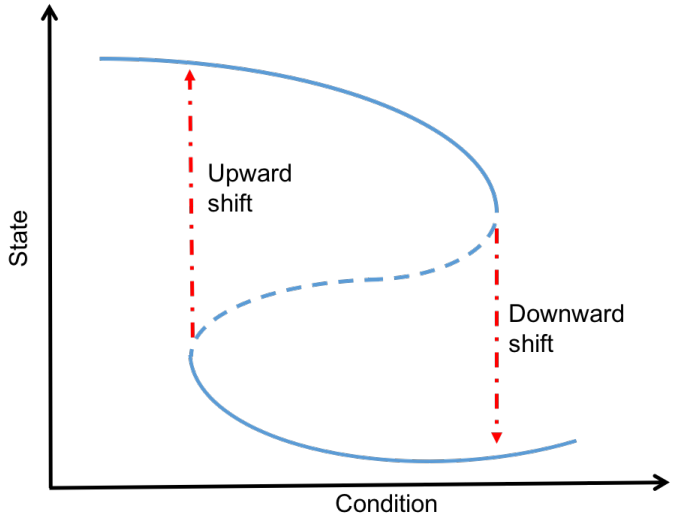

Figure 1.1: Critical transition can be caused by changes in the underling conditions.

Close to the threshold of such critical transitions, some systems lose resilience in the sense that it takes much longer for the system to recover from perturbations. Consequently, several indicators based on fluctuations around equilibrium values have been proposed as early warning signals for critical transitions [5, 21, 17, 36, 37, 38, 39, 8]. These methods are based on the presumption that as a system approaches

the threshold, one would expect to observe the increase of certain statistics such as variance, skewness and autocorrelation due to the critical slowing down phenomenon. It has been shown in models that this is because the real part of the dominant eigenvalue that characterizes the recovery rate around the equilibrium state becomes closer to zero as the system approaches the critical transition [19, 40, 41]. An example of variance and autocorrelation rising as early warning signals for a harvested population driven slowly across a bifurcation can be found in Fig. 1.2.

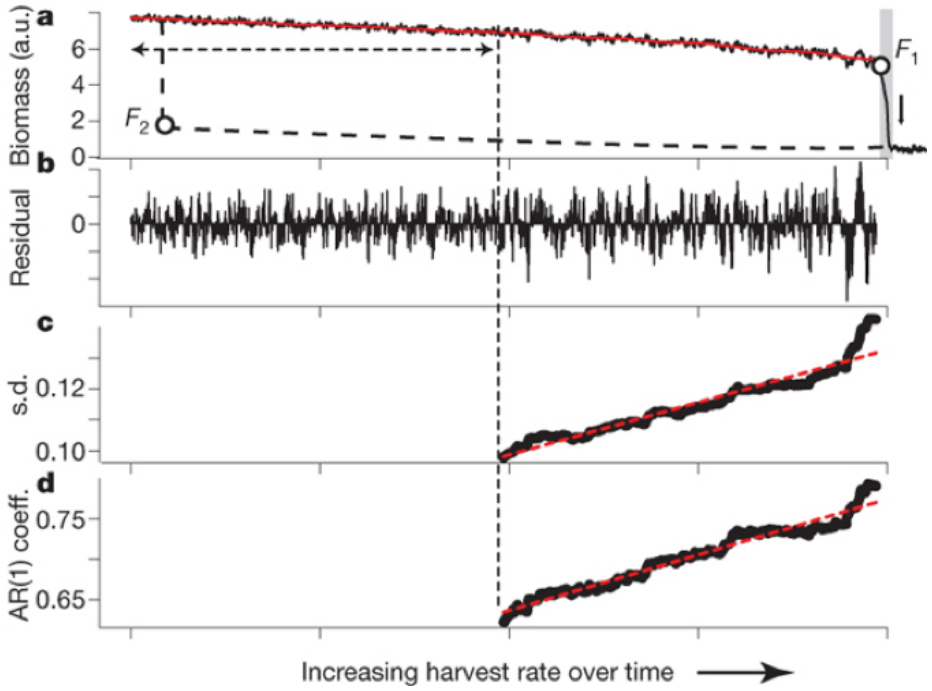

Figure 1.2: Early warning signals for a critical transition using data generated by a model of a harvested population77 driven slowly across a bifurcation. This figure is obtained from [1].

Apart from these temporal early warning signals, recent studies suggest that spatial patterns can also provide useful information [42, 22, 32, 23]. In particular, Dakos et al. [22] point out that an increase in spatial correlation can serve as an early warning signal for spatially extended systems with alternative steady states. This is because, as the system approaches the critical transition, the system becomes slow in recovering from perturbations, which might lead to stronger fluctuations of state vari-

ables around the equilibrium state under random environmental perturbations [43]. In such cases, the fluctuations of state variables around the spatial mean will also increase. Also, diffusive effects (exchanges between neighboring patches) become more dominant close to the critical transition, which means that spatial correlations, especially correlations between neighboring units will increase [31]. Therefore, spatial early warning signals, such as the spatial variance, spatial skewness and spatial correlation have been introduced to serve as indicators for critical transitions of spatially extended systems.

1.2.2 Bifurcation diagram forecasting methods

Generic early warning signals, however, can only be used as qualitative early warning signals because of their difficulty in quantitatively detecting the distance to critical transitions [44]. The model-less method available to anticipate the distance to a critical transition and the state of the system after such transition was proposed in [45] where tipping points in the dynamics of electro-mechanical systems were studied. In that method, recoveries of the system from large perturbations are used to calculate the rate of change of amplitude of the dynamics so as to predict not only the bifurcation point, but also the stable and unstable branches of the bifurcation diagram. This method was extended in [46, 36] to forecast bifurcations of multi-dimensional aeroelastic systems. The benefit of using recoveries from large perturbations is that nonlinearity of the system better manifest itself at large amplitudes. Therefore, by examining the recovery rates at large amplitudes, we can not only obtain information about the linear stability of the equilibrium, but also the type of critical transition (rapid or piecewise continuous) that the system might go through.

Procedures of the bifurcation diagram forecasting method is as follows:

1. A perturbation is applied to the system, and the recovery data from this perturbation is recorded as shown in Fig. 1.3 (1).

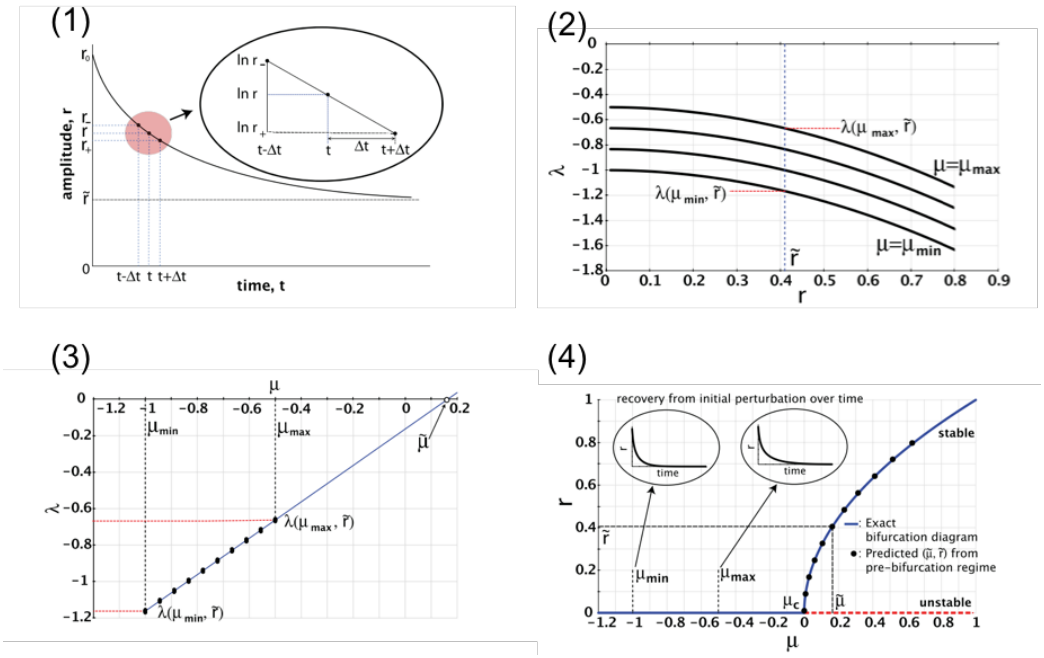

Figure 1.3: Procedures of forecasting bifurcation diagrams using recoveries from large perturbations.

2. Recovery rate is calculated from the recovery data and plotted against its corresponding amplitude as shown in Fig. 1.3 (2).

3. Repeat step 1 and 2 at several different parameter values prior to the bifurcation point.

4. For a particular amplitude value, calculate the recovery rate using data collected at the parameter values from step 3. Extrapolate the recovery rate versus parameter value curve until it crosses the x axis and record the parameter value at the crossing as shown in Fig. 1.3 (3).

5. Repeat step 4 at several amplitude values and construct the predicted bifurcation diagram as shown in Fig. 1.3 (4).

1.2.3 Inertial manifold

For high dimensional systems, it is sometimes unrealistic to monitor all the states to anticipate a potential critical transition. Moreover, some states might provide irrelevant or even wrong warning signals due to the dynamical structure of the system [47]. Therefore, the robustness and effectiveness of early warning algorithms for high dimensional systems rely on the ability to identify and monitor states that are strongly correlated with the critical transition. To that aim, we utilize inertial manifold theory [14].

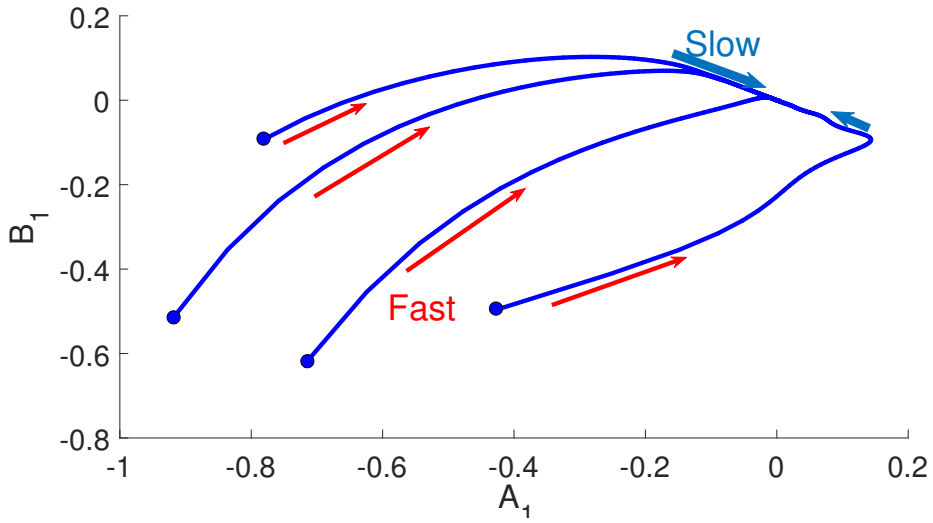

Figure 1.4: The slow manifold is low dimensional with dynamics slowest in time. When the system is perturbed, dynamics in the fast manifold will die out quickly. Recoveries from different perturbations converge to the low dimensional slow manifold.

When a bifurcation occurs, the equilibrium is non-hyperbolic, and the associated linearized system can be divided into a stable subspace and a center subspace [14]. For co-dimension one bifurcation, there exists a one or two dimensional invariant manifold tangent to the center subspace at the equilibrium. This manifold is defined as the center manifold. For a high dimensional system, there is a certain parameter-dependent one or two dimensional manifold in which the system exhibits a bifurcation, and this

manifold is referred to as the inertial manifold. The inertial manifold coincides with the center manifold when the system is at the bifurcation point. Dynamics in the stable manifold can be neglected because of their faster decay rate compared to the dynamics in the slow manifold. An example of the slow manifold is shown in Fig. 1.4. When a perturbation is applied to the system, the dynamics in the fast manifold quickly die out. All recoveries from different random perturbations converge to the low dimensional slow manifold. Therefore, we can identify the important state by examining the inertial manifold. The more aligned a state is with the inertial manifold, the more information it contains about the critical transition.

1.3 Dissertation Contributions and Outline

This thesis aims at developing new early warning signals and algorithms that can bridge the gaps between the current early warning signal theories and the challenges posed by nonlinearity, periodicity and spatial heterogeneity. In Chapter 2 and Chapter 3 we proposed two method to forecast the bifurcation diagram of systems with parametric resonance. In Chapter 4 we studied an epidemiology system that is subject to parametric resonance. We turn our focus to high dimensional, heterogeneous systems in Chapter 5, and proposed a set of new early warning signals, namely the largest eigenvalue of the covariance matrix and the percentage it accounts for of the total variation. Chapter 6 provides a summary of contributions, and suggestions for future work. In addition, Appendix A provides accompanying derivations for the work presented in Chapter 5. Appendix B examines the robustness of the trend statistic Kendalls τ used in most early warning signals.

The main contributions are summarized as follows:

1. We developed a new forecasting method that uses the transient recovery of parametrically excited systems from large perturbations to predict the bifur-

cation diagram. Recoveries are studied in a Poincaré section to address the challenge posed by periodic variation of the parameter. Details of this method are included in Chapter 2, where we also demonstrated the method using both simulation and experimental data. Chapter 2 is adapted from a published paper [8].

2. We developed a new forecasting method for a high dimensional parametrically excited systems that uses the recoveries of a single variable in the slow manifold to forecast the bifurcation diagram of that variable. Details of this method are included in Chapter 3, where we also demonstrated the method using both simulation and experimental data. Chapter 3 is adapted from a published paper [48].

3. We studied a classic SIR model with periodic forcing using the method of multiple scales. The SIR model is a compartmental model, where the population is divided into three compartments: S for the number susceptible, I for the number of infectious, and R for the number recovered (or immune). This study shows that large amplitude epidemics can take place if the system is moved into a resonance regime in the parameter plane by the change of any one of its parameters. Details of the analysis are reported in Chapter 4, which is aslo adapted from a published paper [25].

4. In Chapter 5, We proposed to use the largest eigenvalue of the covariance matrix as spatial early warning signals. By establishing the relationship between the eigenvalues of the covariance matrix and the eigenvalues of the force matrix, we mathematically show that the dominance of the largest eigenvalue of the covariance matrix can also be used as an early warning signal. We then demonstrated the proposed early warning signals using simulation data collected from three spatially correlated ecological models. This chapter is adapted from a

manuscript submitted to a scientific journal.

CHAPTER II

Forecasting bifurcations in parametrically excited systems

2.1 Introduction

A system is referred to as parametrically excited when one or some of its parameters vary with time. Parametric excitation can be observed in various engineered and physical systems such as elastic cables [49], rotating machines [50], electrostatic waves in a plasma with radiation [51], and others [52, 53, 54, 20]. Many natural systems are also subject to variations of parameters due to environmental fluctuations and human activities [55, 56, 57].

Many systems subject to parametric excitation exhibit critical transitions from one state to another as the amplitude and frequency of the fluctuating parameters or the values of other parameters change. Such critical transitions can either be caused by a change in the topological structure of the unforced system, or by a synchronization between a natural mode of the system and the parameter variation. While both cases are important, the second scenario has received particular interest because a small parametric excitation might produce a large response when the driving frequency is close to twice the natural frequency of the system [20]. This phenomenon is referred to as parametric resonance and has been widely observed in both engineered and natural

systems. For example, efforts have involved applications to encourage its onset, such as parametrically excited vibration energy harvester [58] and parametric resonance amplification in MEMS gyroscopes [59]. Other efforts have been made to limit the growth of parametric resonance, such as prevention of parametric rolling in ship engineering [13] and suppressing vortex induced vibration in offshore engineering [60].

Bifurcations can be dangerous because of the large amplitude response in the post-bifurcation regime. Thus, it could be very beneficial if monitoring or early warning were available while the system is still in the stable regime. Traditionally, Floquet theory is used to study the linear stability of time dependent systems [61]. In that theory, the eigenvalues of the monodromy matrix are calculated to determine the stability of the system. Carrying out such analysis requires accurate knowledge of the model of the system [62]. This may not be possible for a variety of complex systems. Thus, model-less methods are needed to predict the actual distance to the transition and the state of the system after the transition using only data obtained in the pre-bifurcation regime.

Recently, a number of model-less methods which use time series [16] have been proposed to anticipate critical transitions. These model-less methods are based on the observation that the dynamics of many systems near critical transitions have a common property regardless of the different details of each system. Such generic property is the phenomenon of critical slowing down, in which the system recovers from perturbations ever more slowly as the system approaches a critical point [5, 6, 63, 15, 64]. These methods rely on indicators such as an increased autocorrelation or variance [5, 65] obtained from measurements during the system's recovery from small perturbations. However, these methods can only be used as qualitative early warning signals because of their difficulty in quantitatively detecting the distance to critical transitions [44].

The model-less method available to anticipate the distance to a critical transition

and the state of the system after such transition is breached was proposed in [45] where tipping points in the dynamics of electro-mechanical systems were studied. In that method, recoveries of the system from large perturbations are used to calculate the rate of change of amplitude of the dynamics so as to predict not only the bifurcation point, but also the stable and unstable branches of the bifurcation diagram. This method was extended in [46, 36] to forecast bifurcations of multi-dimensional aeroelastic systems. Although methods based on the persistent response to persistent stochastic excitations [15, 64, 16], and methods utilizing transient recoveries from large perturbations [45, 46] are all using slowing down of the system response close to critical transitions, the latter method relies on recoveries from large perturbations instead of the persistent response to persistent stochastic excitations. Thus, the noise is considerably smaller than the system response, and noise filtering approaches can be use.

Despite the different utilities of the existing methods mentioned above, they were all developed for autonomous systems, without considering the effects of time variant inputs. To deal with time-varying parameters, researches have used Hilbert transform or Gabor transform to estimate both the instantaneous amplitude and frequency of the system response to large perturbations [66, 67]. Important parameters, such as damping and stiffness coefficients are then identified. Despite all their benefits, these identification methods rely on a relatively accurate model of the system, which limits their applicability. Thompson and Virgin [68] proposed to use transient frequencies to predict an incipient jump to resonance of a lightly damped nonlinear oscillator. However, that method is not able to predict the behavior of the system after the jump. In this work, we present another approach to predict the location of bifurcations related to parametric excitation only using recovery data in the pre-bifurcation regime. Moreover, the behavior of the system after the bifurcation can also be anticipated using the proposed approach, which provides an understanding of the post-bifurcation

dynamics.

Recoveries from large perturbations are exploited in the proposed method. These perturbations can be applied or can be caused by natural events. When a system recovers from large perturbations, the recovery rates at different amplitudes and parameter values can be used to obtain desired information about the bifurcation and about the post-bifurcation dynamics. Different from what has been done in [45, 69], the method proposed herein uses recoveries in a Poincaré section obtained by sampling the trajectory once for each forcing period. Also, different techniques are proposed for critical transitions caused by bifurcations of the unforced system, or by parametric resonance.

Numerical simulations and experimental results are provided to demonstrate the proposed method. In numerical simulations, a parametrically excited logistic equation and a parametrically excited Duffing oscillator are used to demonstrate the proposed forecasting method. These two types of systems show that the method can predict transitions induced by either bifurcation of the unforced system, or by parametric resonance. We further examine the robustness of the method to measurement and process noise by collecting recovery data from an electrical circuit system which exhibits parametric resonance as one of its parameters varies. We also discuss how the method can be applied to other time variant systems other than parametrically excited systems.

2.2 Theory

2.2.1 Method

The method proposed herein is an extension of work in [45]. In that work, co-dimension one Hopf bifurcations of autonomous systems were studied. The rate of change of amplitude was used as an indicator of critical slowing down. In this paper,

parametrically excited systems with critical transitions are studied. A new technique is introduced to deal with the explicit time input. The proposed method does not depend on the details of the physics, but has some requirements/limitations. One of the requirements is that the system must have a periodically varying parameter. Another requirement is that the bifurcation has to exhibit slowing down. Indeed, the ideas in the proposed method are relying on the fact that the dynamics in a phase portrait is topologically equivalent to that of a system in a normal form. The benefit is that the approach does not require a model of the system, i.e. it does not depend on the details of the underlying physics.

The proposed method is summarized in Fig. 2.1. The method relies on observations of system recoveries from perturbations before a bifurcation takes place. Distinct from previous work, the proposed method can predict the entire bifurcation diagram at all phases using data collected from the pre-bifurcation region, as shown in Fig. 2.1. Recoveries are collected at two or more parameter values before the bifurcation. The state-space recoveries at two such parameter values are shown in boxes A_1 and A_2. Note that it takes more cycles for the system in box A_2 to recover to the equilibrium than the system in box A_1, which indicates that the dynamics slow down as the system approaches the bifurcation. The fundamental idea of this method is that the region where the states have slowest recovery rates corresponds to the limit cycles of the nearest bifurcation in the parameter space.

Note that a limit cycle in the state space corresponds to a set of fixed points in a Poincaré section. Therefore, we relate the distance to a limit cycle in the parameter space to the local recovery rate in a Poincaré section. The states with slowest recovery rates in a Poincaré section are closest in the parameter space to a limit cycle.

In addition, a bifurcation of a parametrically excited system can be caused by two different mechanisms. In the first case, there is a bifurcation in the corresponding autonomous system where all the parameters are kept constant. In such cases, a

bifurcation exists even if the parametric excitation is removed. The second case is parametric resonance, where the bifurcation is caused a phase-lock between the periodic variation in state variables and the periodic variation in parameter values. Different methods are developed for these two different scenarios, and are discussed separately below.

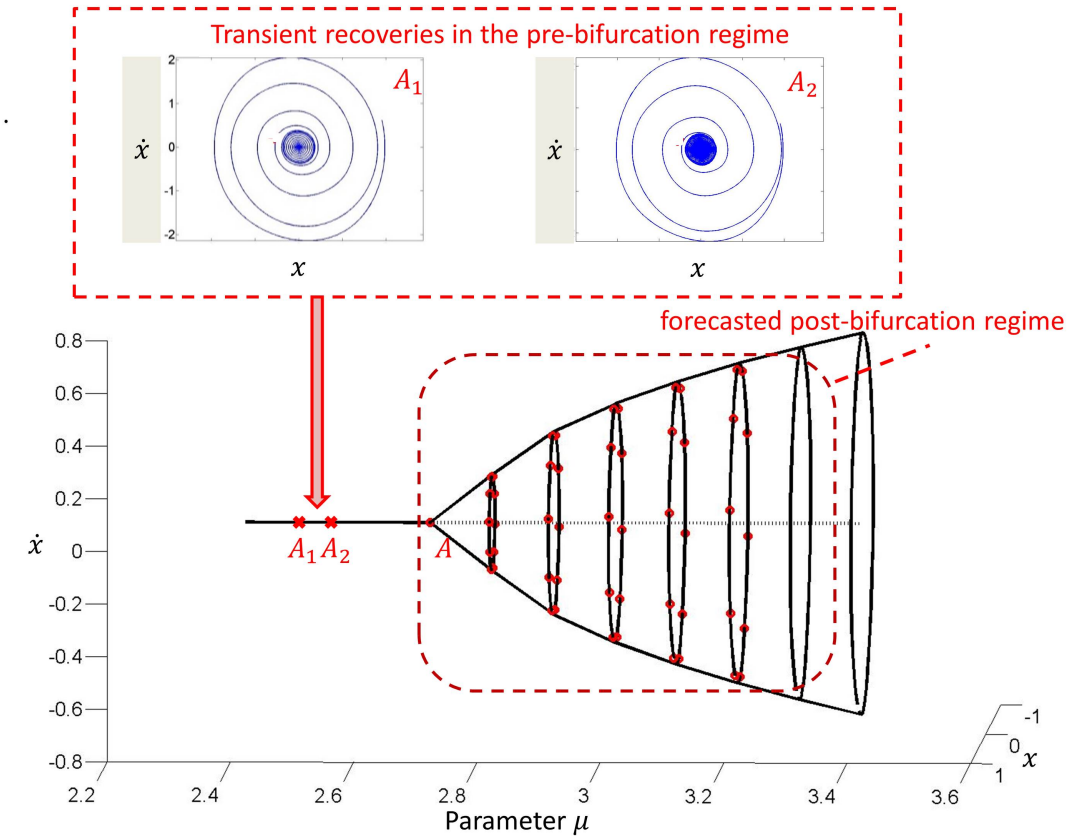

Figure 2.1: The bifurcation diagram of a parametrically excited system with both a stable branch (solid line) and an unstable branch (dash line) are shown. μ denotes the bifurcation parameter. x and $\dot{x}$ denote the states of the system. The method can predict both the bifurcation point A and the post-bifurcation regime (dots) using transient recoveries from perturbations collected in the pre-bifurcation regime. Examples of the state space transient recoveries are shown in boxes A_1 and A_2.

2.2.1.1 Bifurcations in the corresponding autonomous systems

Consider a parametrically excited system which experiences a critical transition as one of the parameters vary. When there is no resonance and the dynamics in the corresponding autonomous system exhibit a bifurcation at a critical parameter value μ_c, parametric excitation works merely as a driving force, creating a limit cycle which has the same frequency as the excitation. In this case, the dynamics can be characterized by a (physical, measurable) coordinate x, and a governing equation given by

$$\frac{dx}{dt} = xf(x, \mu, t) + g(x), \tag{2.1}$$

where μ is the bifurcation parameter, and where f and g are smooth near $x = 0$. In this new expression, parametric excitation is caused by $f(x, \mu, t)$ which is periodic in time. $g(x)$ is a function that is only related to the state variables. $f(x)$ and $g(x)$ together define the bifurcation of the system. Different from [45], in the current method, the recovery rate is studied in the Poincaré section obtained by sampling the trajectory once during each forcing period. Thus, Eq. (1) is first rewritten as

$$\frac{dx}{dt} = xf(x, \mu, \theta) + g(x), \tag{2.2}$$

where $\theta \in S^1$, and $S^1 = R/T$ is the time circle of length T [14]. Next, we define a cross section $\Sigma = \{(x, \theta)|\theta = \phi\}$, and the Poincaré map corresponding to that section as

$$x_{n+1} = p(x_n, \mu, \phi).$$

Here, we refer to ϕ as the forcing phase of the map. x_{n+1} represents the state that x_n is mapped to after one period T given a parameter value μ and a forcing phase ϕ.

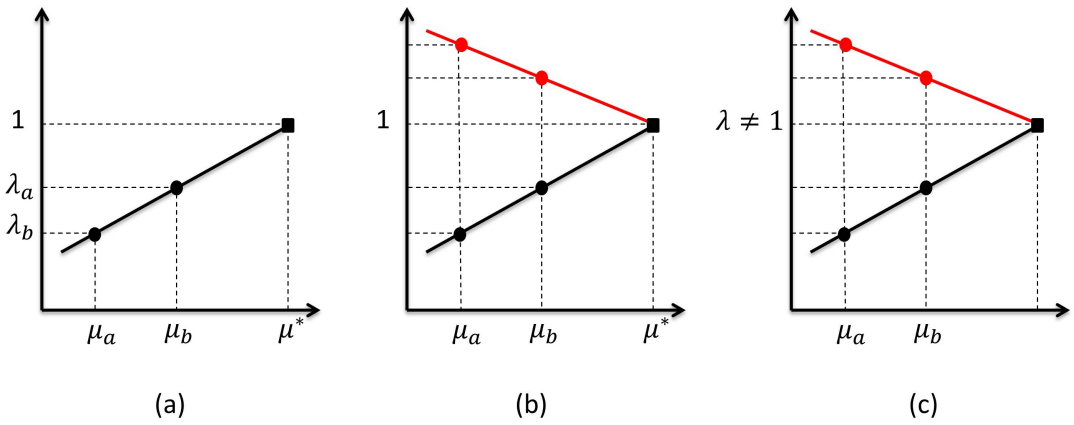

Figure 2.2: Examples of using the function λ versus μ at a fixed state x^* to find the corresponding parameter μ^* for the chosen state x^*. (a) Parametrically excited logistic equation. (b) Parametrically excited Duffing oscillator when the forcing phase is the corresponding forcing phase of the state. (c) Parametrically excited Duffing oscillator when the forcing phase is not the corresponding forcing phase of the state.

Thus, the recovery rate of state x in this map is defined as:

$$\lambda(x_n, \mu, \phi) = \frac{p(x_n, \mu, \phi)}{x_n}, \tag{2.3}$$

where $\lambda(\mu, x, \phi)$ is not known, but it is required to be smooth with respect to μ. Next, we denote the critical value of the bifurcation parameter as μ_c. The bifurcation occurs at $\mu = \mu_c$, which is point A in Fig. 2.1. Then, the Taylor series of λ with respect to μ near μ_c can be written as

$$\begin{aligned} \lambda(x_n, \mu, \phi) = \quad & \lambda(x_n, \mu_c, \phi) \\ & + (\mu - \mu_c) \left. \frac{d\lambda(x_n, \mu, \phi)}{d\mu} \right|_{\mu=\mu_c} + H.O.T. \end{aligned} \tag{2.4}$$

Note that the linearization in Eq. (2.4) is done in the parameter space rather than in the state space. Thus, no approximation is made in the state space, even when x_n is large. This helps to predict large amplitude limit cycles which exist after the bifurcation.

The higher order terms ($H.O.T.$) in Eq. (2.4) can be neglected when μ is close to μ_c. However, when the prediction is applied to recoveries with parameter values far from μ_c, or for a post-bifurcation dynamics far from μ_c, one can use the higher order terms in Eq. (2.4) to enhance the accuracy of the method. In the following, we focus on cases where μ is close to μ_c.

Consider a fixed point x^* and at a parameter value μ^* on the bifurcation diagram of the map $x_{n+1} = p(x_n, \mu, \phi)$. We obtain

$$\lambda(x^*, \mu^*, \phi) = \frac{p(x^*, \mu^*, \phi)}{x^*} = 1. \tag{2.5}$$

Consider two time series of measurements (denoted by x_n^a and x_n^b) of the system response to large perturbations for two different values of μ (denoted by μ_a and μ_b). An Example of such measurements is shown in Fig. 2.3. The recovery rate at these two different parameter values can be expressed using Eq. (2.3) as

$$\begin{aligned} \lambda(x_n^a, \mu_a, \phi) &= \frac{p(x_n^a, \mu_a, \phi)}{x_n^a}, \\ \lambda(x_n^b, \mu_b, \phi) &= \frac{p(x_n^b, \mu_b, \phi)}{x_n^b}. \end{aligned}$$

Denote by λ_a and λ_b the recovery rates of the system at state x^* with parameter values μ_a and μ_b, namely

$$\begin{aligned} \lambda_a &= \frac{p(x^*, \mu_a, \phi)}{x^*}, \\ \lambda_b &= \frac{p(x^*, \mu_b, \phi)}{x^*}. \end{aligned}$$

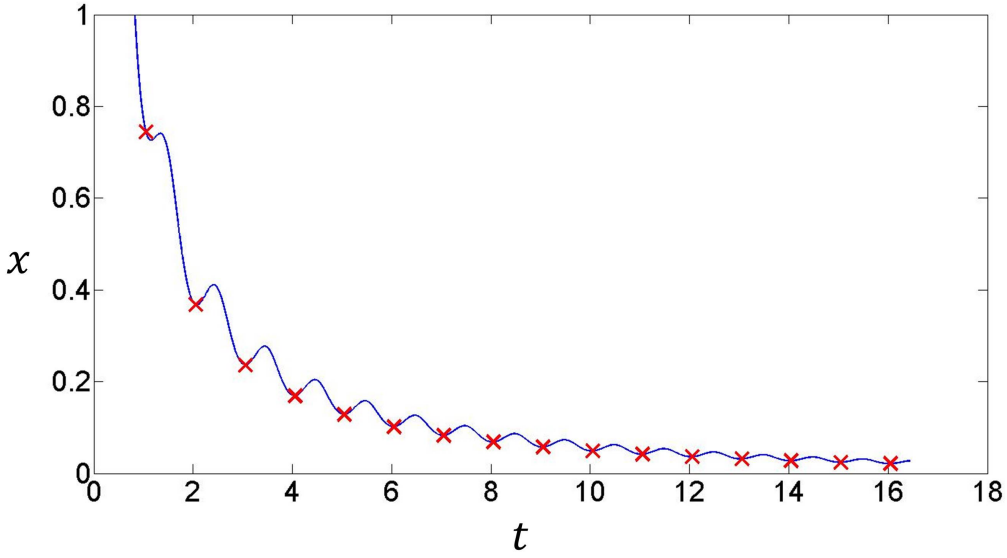

Figure 2.3: A recovery of the parametrically excited logistic equation from a perturbation. Symbols represent the samples with the same forcing phase $\phi = \pi$.

Because recovery data is collected in a Poincaré section, measurements may not be available exactly at state x^*. Thus, the values of λ_a and λ_b can be obtained through a local interpolation using available measurements which are at states close to x^* and at the forcing phase value of ϕ.

Using Eq. (4) it follows that at a fixed value of $x = x^*$ λ varies linearly with μ close to μ_c. A linear function of λ versus μ can be fitted using μ_a, μ_b and λ_a, λ_b at a fixed value of x^* as shown in Fig. 2.2 (a). Then, μ^* can be obtained by determining the parameter value at which Eq. (5) holds, namely $\lambda(x^*, \mu_c, \phi) = 1$.

If measurements at more than two parameter values are available, a linear or nonlinear function of recovery rate λ versus parameter value μ can be fitted. Then, μ^* can be obtained by determining the parameter value at which Eq. (5) holds, namely $\lambda(x^*, \mu^*, \phi) = 1$.

The method above shows how to find the parameter value μ^* which corresponds to a state x^*. Thus, the entire bifurcation diagram for a fixed forcing phase can be

constructed. The value of μ^* is affected by the value of the forcing phase ϕ for the same state x^*. Therefore, it is important that the system is studied in the Poincaré section corresponding to that value of the phase.

The method to predict the bifurcation diagram of a system with parametric excitation (when bifurcation is not caused by parametric resonance) can be summarized as the following steps:

1. Choose at least two parameter values that are close to the bifurcation and measure at least one system recovery from a large perturbation at each parameter value.

2. Apply noise filtering techniques to available transient recoveries.

3. Choose a forcing phase ϕ to predict the corresponding bifurcation diagram. For each value of the bifurcation parameter μ, compute λ for all the states x collected at the forcing phase ϕ. Approximate the dependence of λ on x by using a polynomial function (e.g., a linear function).

4. Choose a state x^* where λ can be calculated from measurements at all parameter values, and use the relationship between λ and μ to find the corresponding parameter value μ^* for which x^* is a fixed point (i.e. λ=1).

5. In a similar manner, choose different x^* values and repeat the procedure at steps 3 and 4 to predict the entire bifurcation diagram at the chosen forcing phase. Finally, choose another value of the forcing phase and repeat the procedure at steps 3 and 4 to predict the entire bifurcation diagram for all measured values of ϕ.

2.2.1.2 Parametric resonance

It is more complicated to predict the bifurcation diagram when the critical transition is caused by a parametric resonance. Parametric resonance typically occurs when

the parametric excitation is close to twice the natural frequency of one of the modes of the system. At parametric resonance, the amplitude variation is synchronized with the parameter variation, resulting in a sustained oscillation at a certain amplitude. Therefore, each state on a limit cycle corresponds to a fixed forcing phase value ϕ^*. Thus, to predict the parameter value μ^* corresponding to a state x^*, we also need to predict the corresponding forcing phase ϕ^*. Hence, predicting the bifurcation diagram of a system with parametric resonance is more complicated in the sense that both ϕ^* and μ^* have to be predicted. This contrasts the analysis in the previous section where x^* and ϕ are chosen and only μ^* has to be predicted.

Consider a system that experiences parametric resonance as one of its parameters varies. Parametric resonance typically occurs when the parametric excitation is close to twice the frequency of the dynamics in one of the manifolds (e.g., the frequency of one of the normal modes of the system). Consider that the dynamics of that mode are characterized by coordinates (x_1, x_2) and by the two dimensional Poincaré map

$$\begin{aligned} x_{1,n+1} &= p_1(x_{1,n}, x_{2,n}, \mu, \phi), \\ x_{2,n+1} &= p_2(x_{1,n}, x_{2,n}, \mu, \phi), \end{aligned} \tag{2.6}$$

where ϕ is a chosen forcing phase and $(x_{1,n+1}, x_{2,n+1})$ represents the state that $(x_{1,n}, x_{2,n})$ is mapped to after two periods T at a given parameter value μ and forcing phase ϕ. Two periods are used to construct the Poincaré map instead one because parametric resonance is a 2:1 resonance, which means that the frequency of the limit cycle is half the frequency of the excitation.

Recovery rates λ_1 and λ_2 are defined similar to Eq. (2.4),

$$\lambda_1(x_{1,n}, x_{2,n}, \mu, \phi) = \frac{p_1(x_{1,n}, x_{2,n}, \mu, \phi)}{x_1},$$
$$\lambda_2(x_{1,n}, x_{2,n}, \mu, \phi) = \frac{p_2(x_{1,n}, x_{2,n}, \mu, \phi)}{x_2}. \tag{2.7}$$

Denote by (x_1^*, x_2^*) a fixed point of the map at a parameter value μ^* and forcing phase ϕ^*. Therefore, $\lambda_1(x_1^*, x_2^*, \mu^*, phi^*) = \lambda_2(x_1^*, x_2^*, \mu^*, \phi^*) = 1$. To find the values of ϕ^* and μ^*, λ_1 and λ_2 are calculated for several forcing phase values from 0 to 2π. For each value of the forcing phase, λ_1 and λ_2 are calculated in the pre-bifurcation regime at (x_1^*, x_2^*) using at least two parameter values μ_a and μ_b. Polynomial functions of λ_1 and λ_2 versus μ are fitted and extrapolated into the post-bifurcation regime as shown in Fig. 2.2. This is a step similar to step 4 in the previous section.

In general, at a chosen phase value ϕ, the two recovery rates λ_1 and λ_2 become equal at some value μ, but they have values which are not equal to 1 at that value of μ as shown in Fig. 2.2 (c). This is because the chosen value of ϕ does not correspond to an actual fixed point. Thus, the process is repeated at other values of the phase. At one of the phase values, the intersection of λ_1 and λ_2 versus μ occurs at a common value of 1 as shown in Fig. 2.2 (b). Thus, two quantities are obtained, namely μ^* and ϕ^*, which correspond to two constraints $\lambda_1 = 1$ and $\lambda_2 = 1$.

If measurements at two or more parameter values are available, linear or nonlinear functions of λ_1 and λ_2 of μ can be fitted. Then, the values of μ when $\lambda_1 = \lambda_2$ can be determined using these polynomial functions at each phase value ϕ. The values of μ and ϕ corresponding to the fixed point are obtained when λ_1 and λ_2 are equal to 1.

This procedure identifies the parameter value μ^* and forcing phase ϕ^* which correspond to the fixed point (x_1^*, x_2^*). The process can be repeated at other values of the fixed point to obtain the entire bifurcation diagram, as shown in Fig. 2.10.

The method to predict the bifurcation diagram of a system with parametric resonance can be summarized as the following steps:

1. Choose at least two parameter values that are close to the bifurcation and measure at least one system recovery from a large perturbation at each parameter value.

2. Apply noise filtering techniques to available transient recoveries.

3. Choose a state (x_1^*,x_2^*). For each μ value, compute λ_1 and λ_2 using Eq. (2.7) for (x_1^*,x_2^*) with several chosen forcing phase ϕ from 0 to 2π. For some forcing phase values $\tilde{\phi}$, the combination of (x_1^*,x_2^*) and $\tilde{\phi}$ may not be encountered in the time series, which means that λ_1 and λ_2 can not be calculated directly from measurements. In this case, λ_1 and λ_2 are calculated for states close to (x_1^*, x_2^*) with a forcing phase $\tilde{\phi}$. Then, $\lambda_1(x_1^*, x_2^*, \mu, \tilde{\phi})$ and $\lambda_2(x_1^*, x_2^*, \mu, \tilde{\phi})$ can be interpolated from these results.

4. For all the chosen forcing phases from 0 to 2π, polynomial functions of λ_1 and λ_2 versus μ are fitted to find out the values of μ^* and ϕ^* where $\lambda_1(x_1^*, x_2^*, \mu^*, \phi^*) = \lambda_2(x_1^*, x_2^*, \mu^*, \phi^*) = 1$.

5. Choose different states (x_1^*,x_2^*) and repeat the procedure in steps 3 and 4 to predict the entire bifurcation diagram.

2.2.2 Discussion of the Method

For parametrically excited systems, recovery rate λ is calculated in a Poincaré section. The computation of λ does not require a model of the system. It only requires measurements of the state x of the system during its transient recovery from perturbations for at least two parameter values μ. λ values are thus calculated directly from measurements. These λ values are used to predict the critical parameter value μ^*. The extrapolation used in this process puts some constraints on the method. First is that the method only works for co-dimension one bifurcations. Second, dynamics of the system have to vary smoothly as parameter changes. Third, measurements

have to be collected at the parameter values close to the bifurcation for the dynamics to experience a slow down.

Due to such requirements of the system, the proposed approach bears some resemblance to the normal form approach from dynamical system theory. Normal forms provide key information about the qualitative nature of the dynamics, and specifically about the presence of slowing down. The specific expression of the normal form, however, can be distinct. The proposed method (including several past methods developed for autonomous systems) can be applied to a range of bifurcations, including transcritical, pitchfork and Hopf bifurcations. These different bifurcations have different normal forms. However, the method remains the same. Thus, we describe the method even more generically than a normal form, i.e. the system in Eq. (1). In addition, the proposed approach focuses on two types of critical transitions with parametric excitation. In the first type, there is a bifurcation in the corresponding autonomous system where all the parameters are kept constant. In such cases, a bifurcation exists even if the parametric excitation is removed. A normal form exists, although its expression can be one of several. In the second type, there is a parametric resonance, where the bifurcation is caused a phase-lock between the periodic variation in the state variables and the periodic variation in the parameter values. Of course, the causes of these two different types of bifurcations are different, but the proposed approach remains the same.

Since measurement noise is unavoidable, filtering techniques can be applied before calculating λ. However, there are some errors that cannot be filtered easily (e.g. by using a low pass filter). For some combinations of state values and forcing phase values, the λ value can not be calculated directly using the time series. Therefore, interpolation has to be applied, which introduced errors that cannot be easily filtered out. However, there are two methods to reduce this error. The first method is to gather more time series at each parameter value, which will help increase the accuracy

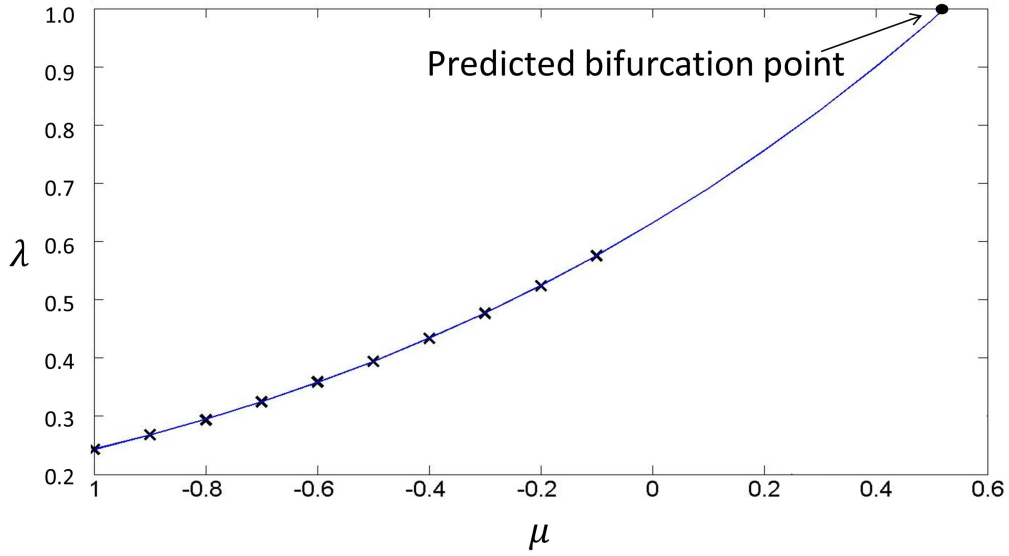

Figure 2.4: Nonlinear extrapolation is used to find the bifurcation parameter μ^* for a state x^*. Symbols $\times$ represent λ values calculated from the recovery. The solid line represents the fitted cubic function of λ vs. μ at state x^*.

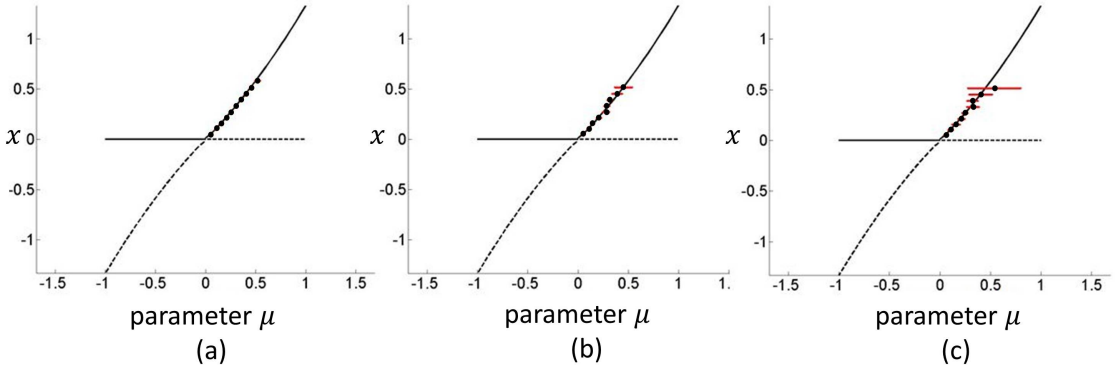

Figure 2.5: Prediction results for the bifurcation diagram of a parametrically excited logistic equation at a forcing phase of $\phi = \pi$. 50 recoveries were collected for 9 pre-bifurcation parameter values. The exact bifurcation diagram with both the stable (solid line) and unstable (dashed line) branches are shown. The predicted bifurcation diagram is shown with symbols together with standard deviation error bars for three cases: (a) no measurement noise, (b) 5% measurement noise, and (c) 10% measurement noise.

of the interpolation. Another method is to approximate the dependence of λ values on forcing phase values ϕ at the same state and parameter values using a periodic function as shown in Fig. 2.7. More details about this are provided in the next section.

2.3 Results

We demonstrate the proposed method using both computational data and experimental data to examine its application to data collected from a system's dynamics.

A parametrically excited logistic equation is used as an example of a critical transition induced by the bifurcation of the unforced system as a parameter varies. The logistic equation is a commonly used model in population dynamics where the growth rate is proportional to both population size and available resources [70]. Since population dynamics are constantly affected by seasonal forcing caused by environmental fluctuations and human activities, a logistic equation with parametric excitation is an important topic to study [71].

A parametrically excited Duffing Oscillator is used as an example system existing a critical transition caused by parametric resonance. The parametrically excited Duffing oscillator is realized physically through an analog electric circuit and experimental data are used for forecasting. This is a nonlinear dynamics system that has been studied using numerical simulation and perturbation methods [20].

2.3.1 Numerical Results

2.3.1.1 Parametrically excited logistic equation

The governing equation of a parametrically excited logistic equation can be written as

$$\dot{x} = \mu(1 + \epsilon sin(\omega t))x - x^2, \tag{2.8}$$

where $\epsilon = 2$, and $\Omega = 1$. μ is the bifurcation parameter.

An example of a recovery from a perturbation with $\mu = -0.1$ is shown in Fig. 2.3. Symbols represent states with forcing phase $\phi = \pi$. The measurements shown in Fig. 2.3 are used to predict the bifurcation diagram at forcing phase $\phi = \pi$.

λ values are estimated at parameter values in a range from $r_{min} = -1$ to $r_{max} = -0.1$ by using Eq. (5). These values are used to predict the corresponding parameter value μ^* of a chosen state x^* as shown in Fig. 2.4. Note that a cubic function of λ versus μ is used for extrapolation to find out the value μ^* of the parameter when the value of λ equals to 1. The entire bifurcation diagram with forcing phase $\phi = \pi$ is constructed by repeating this process for different states x^*.

Prediction results for the bifurcation diagram corresponding to forcing phase $\phi = \pi$ are shown in Fig. 2.5. Measurement noise of 0%, 5% and 10% was added to the observed data, and 50 noisy recoveries were collected at each value of μ. The prediction results anticipated the transcritical bifurcation diagram well in all three case. In general, only one recovery is needed at each parameter value, but because of the measurement noise and the need for local interpolation, additional recoveries can help to improve the accuracy of the forecasting results.

2.3.1.2 Parametrically excited Duffing oscillator

The governing equation of a parametrically excited Duffing oscillator can be written as

$$\ddot{x} + x + \epsilon(2c\dot{x} + \alpha x^3 + \mu x cos(\Omega t)) = 0, \tag{2.9}$$

where $\epsilon = 0.1$, $c = 0.5$, $\alpha = 1$, and $\Omega = 1.9$ are used to obtain numerical data. μ is the bifurcation parameter.

Examples of state space recoveries at three different μ values (of 1.8, 2.2 and 2.6)

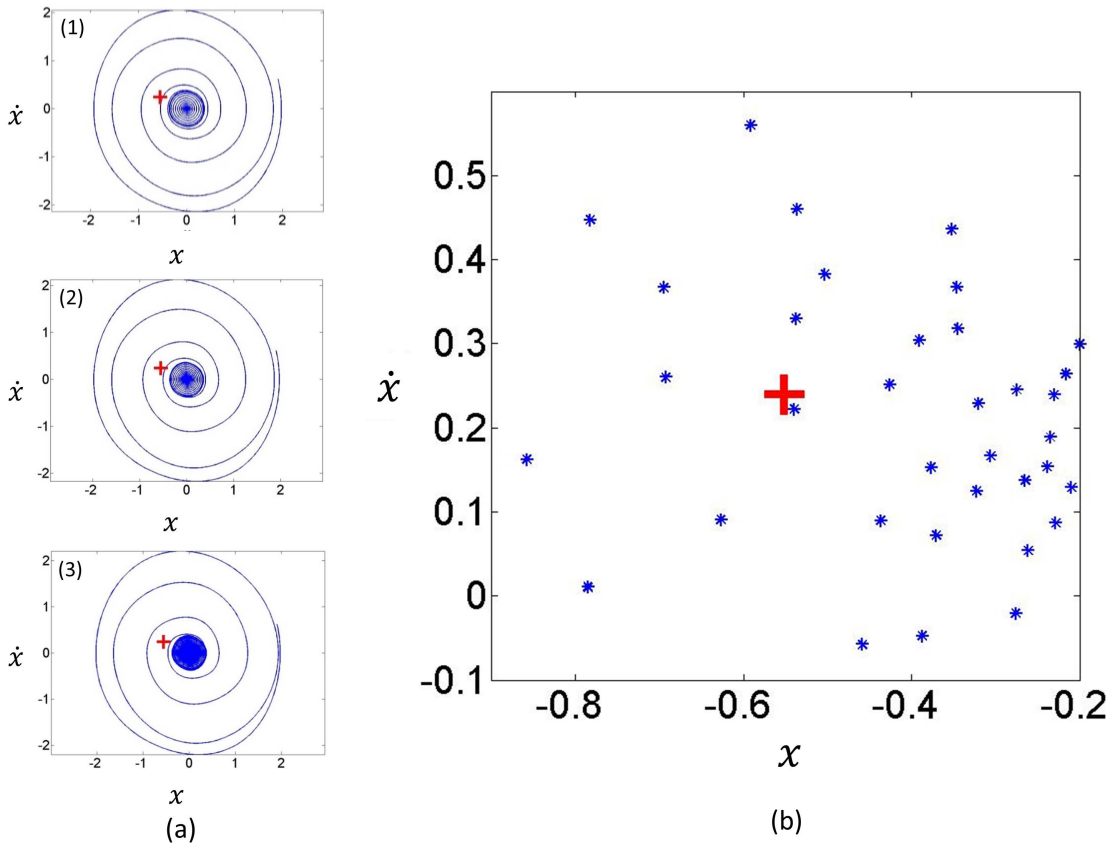

Figure 2.6: Recovery from perturbations of the parametrically excited Duffing oscillator. (a) state-space recoveries at three three different parameter values in the pre-bifurcation regime. Symbols show the chosen state (x_1^*, x_2^*). (b) λ_1 and λ_2 for the state (x_1^*, x_2^*) (+ symbol) and forcing phase ϕ are interpolated from λ_1 and λ_2 of states close by (dots) with the same forcing phase.

are shown in Fig. 2.6 (a).

Data with the same phase ϕ are extracted from recoveries, as shown in Fig. 2.6 (b). λ_1 and λ_2 of these data are calculated using Eq. (2.7). Thus, λ_1 and λ_2 of a selected state $(x, \dot{x}) = (x_1^*, x_2^*)$ can be locally interpolated from these measurements. To increase the accuracy of the local interpolation, and especially to deal with measurement noise, 50 time histories are collected at each bifurcation parameter value. Using multiple recoveries provides more data at the given phase value, which helps alleviate the effects of noise. Results of $\lambda_1(\phi)$ versus the forcing phase at the same bifurcation parameter value are shown in Fig. 2.7.

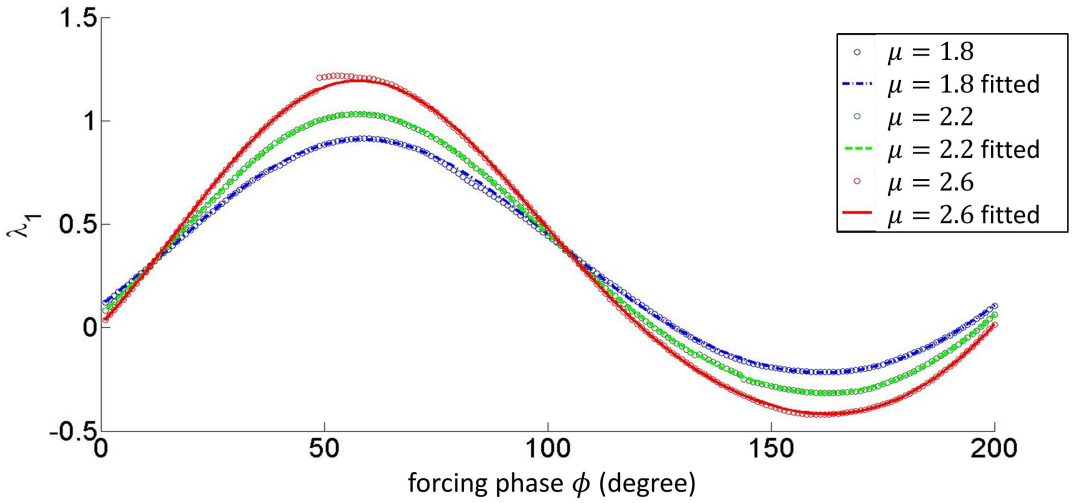

Figure 2.7: λ_1 versus the forcing phase. Solid and dashed lines represent the fitted functions of λ_1 versus forcing phase. Symbols represent the fitted data.

Whether a local interpolation is accurate depends on the amount of data with rich information close to the chosen state (x_1^*, x_2^*). Although 50 time histories are used for each bifurcation parameter, it is still possible that for some forcing phase, the local interpolation is not very accurate. Thus, to increase the accuracy of local interpolation for each chosen state (x_1^*, x_2^*), a periodic function of λ is fitted versus forcing phase ϕ, as shown in Fig. 2.7. Fitted values of λ_1 and λ_2 are used to forecast the corresponding parameter value μ^* of state (x_1^*, x_2^*) on the bifurcation diagram, and the forcing phase it corresponds to as shown in Fig. 2.8.

The approach forecasts the value of the critical parameter for the bifurcation, together with the states of the system after this transition (i.e., the entire bifurcation diagram). Results are shown in Figs. 2.9 and 2.10. Figure 2.9 shows prediction results for the bifurcation diagram corresponding to a single forcing phase with different levels of measurement noise. Although the prediction is better when the noise level is low, the result is still excellent even when there is a 10% measurement noise.Figure 2.10

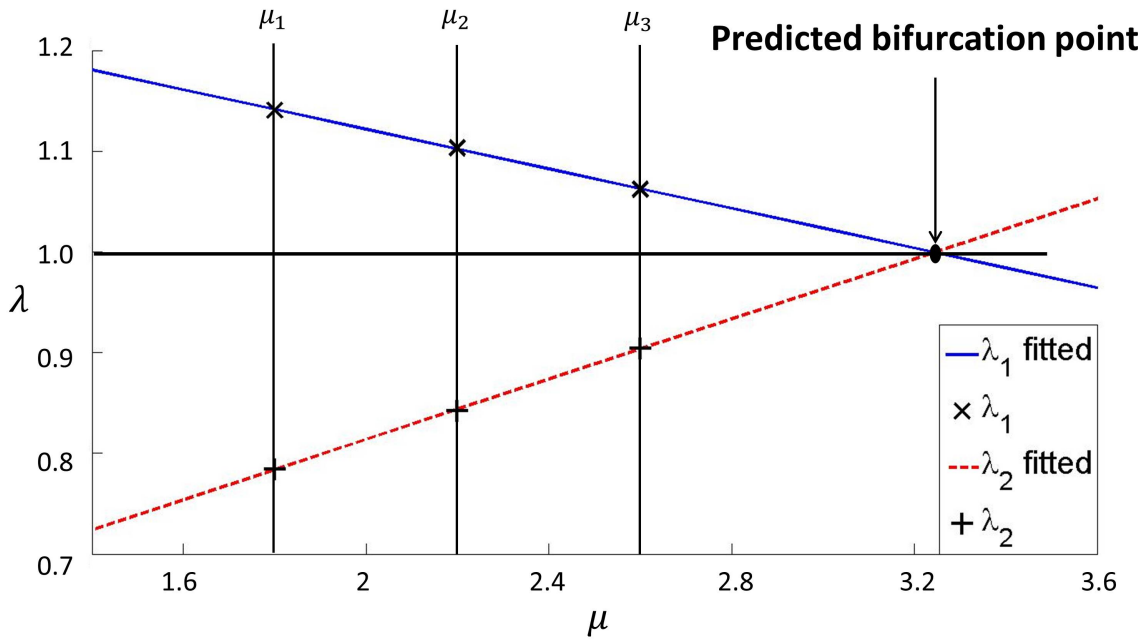

Figure 2.8: $\lambda_1(x_1^*, x_2^*, \mu, \phi)$ and $\lambda_2(x_1^*, x_2^*, \mu, \phi)$ equal to 1 at the same parameter value μ^* when the forcing phase is the corresponding forcing phase ϕ^* for the fixed point (x_1^*, x_2^*).

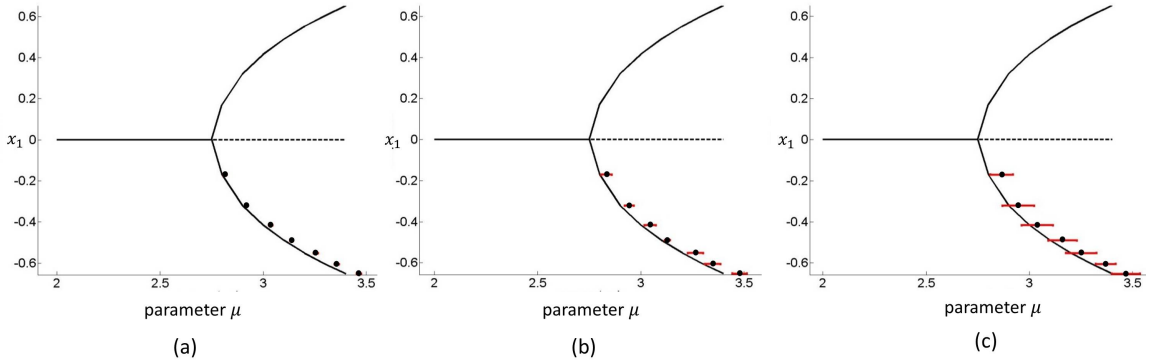

Figure 2.9: Prediction results for the bifurcation diagram corresponding to forcing phase $\phi = \pi$. 50 recoveries were collected at each of the three parameter values in the pre-bifurcation regime ($\mu = 1.8, 2.2$ and 2.6). The bifurcation diagram of both the stable (solid line) and unstable (dashed line) branches are shown. Prediction results are shown by symbols together with standard deviation error bars for three cases: (a) no measurement noise, (b) 5% measurement noise, and (c) 10% measurement noise.

shows prediction results for the entire bifurcation diagram with no measurement noise. The post-bifurcation dynamics are predicted well.

2.3.2 Experimental results

To demonstrate the applicability of the forecasting method to systems with both measurement and process noise, a parametrically excited Duffing oscillator is realized through a physical analog electric circuit shown schematically in Fig. 2.11 (a) and as constructed in Fig. 2.11 (b). The governing equation of the voltage of the circuit is that of Eq. (2.9). The value of the components in Fig. 2.11 are given in Tab. 1. The potentiometer allows control of parameters ϵ, α, c, while μ and Ω can be varied through a function generator responsible for the input signal. The operational amplifiers are model TL082 and the multipliers are model AD633.

Each experiment consists of three steps. The first step is to obtain recovery data from large perturbations in the pre-bifurcation region. The perturbation is caused by temporarily adding a strong white noise. After the perturbation, the dynamics

Table 2.1: Duffing circuit component values

Component	Value
C1, C2	10 nF
R8, R15, R16	1 KΩ pot
R5, R7, R13	10 KΩ
R2, R9	10 KΩ
R1, R3, R4, R6, R10, R12, R14	100 KΩ

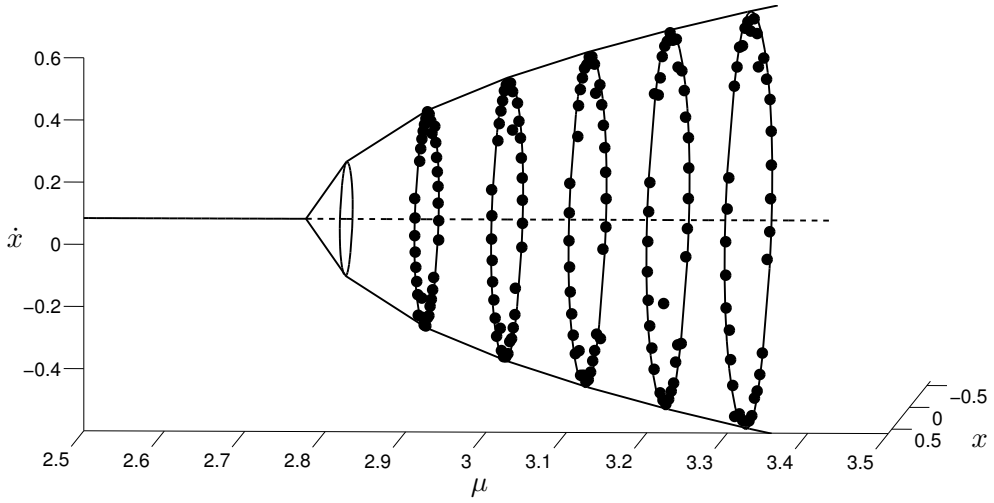

Figure 2.10: Prediction results for the bifurcation diagram. 50 recoveries were collected at each of three parameter values in the pre-bifurcation regime ($\mu = 1.8, 2.2$ and 2.6). Lines represent measured results with both the stable (solid) and the unstable (dashed) branches shown. Symbols represent prediction results at each parameter value. These results are obtained by first predicting the corresponding parameter value μ^* of state (x_1^*, x_2^*), and then interpolating the predicted results to the selected parameter values.

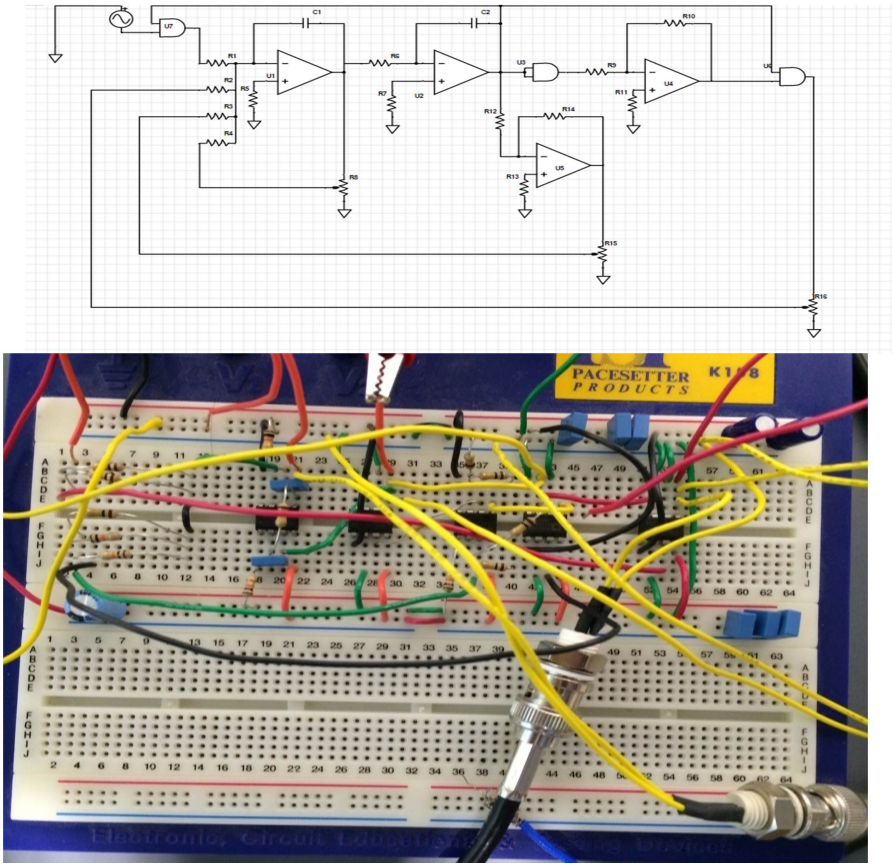

Figure 2.11: Schematic of an analog circuit which has the dynamics of the parametrically excited Duffing oscillator (top). The Duffing circuit as constructed in the laboratory (bottom).

during the recovery to equilibrium are recorded. Specifically, three states $(x, \dot{x}, \phi)$ are measured. x and $\dot{x}$ are represented by two voltages which can be measured directly. The forcing phase is measured directly by recording $\mu\cos(\Omega t)$. Different from numerical simulations, λ_1 and λ_2 cannot be obtain directly from these measurements due to the difference between the sampling frequency and the driving frequency Ω. Therefore, 201 forcing phases are selected, from 0 to 2π with an increment of $\frac{\pi}{100}$. State variables corresponding to these forcing phase values are interpolated from measurements collected close to them. Then, λ_1 and λ_2 can be calculated using data either interpolated or measured with all 201 forcing phase values.

The second step is to predict the bifurcation and the whole bifurcation diagram

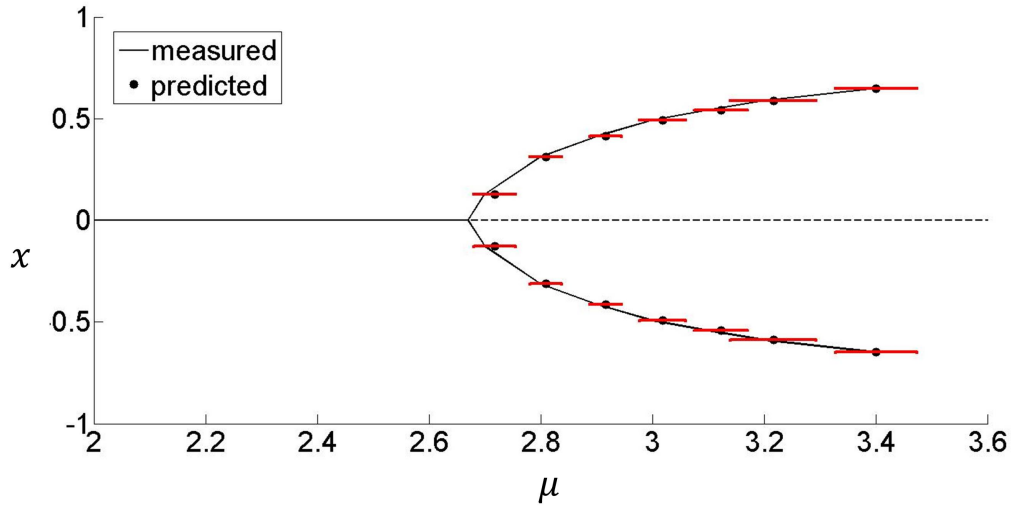

Figure 2.12: Prediction results for the bifurcation diagram of state x versus parameter μ for forcing phase $\phi = \frac{3}{2}\pi$. 50 recoveries were collected at each of two parameter values in the pre-bifurcation regime ($\mu = 2.4$ and 2.6). Solid lines represent the measured bifurcation diagram. Prediction results are shown by symbols together with standard deviation error bars.

using λ_1 and λ_2 calculated from the second step. Prediction results of the bifurcation diagram corresponding to one forcing phase are shown in Fig.2.12. Prediction results for the entire bifurcation diagram are shown in Fig. 2.13. To obtain this results, 50 recoveries were were collected at each of two parameter values. It is worth mentioning that multiple recoveries were collected to enhance the accuracy of the prediction results.

The third step is to obtain the actual bifurcation diagram by the classic method of parameter sweeping. This step is performed so that the predictions obtained using the proposed forecasting approach can be validated using the actual bifurcation diagram. The limit cycle oscillations are obtained and plotted versus the corresponding parameter value to obtain a bifurcation diagram.

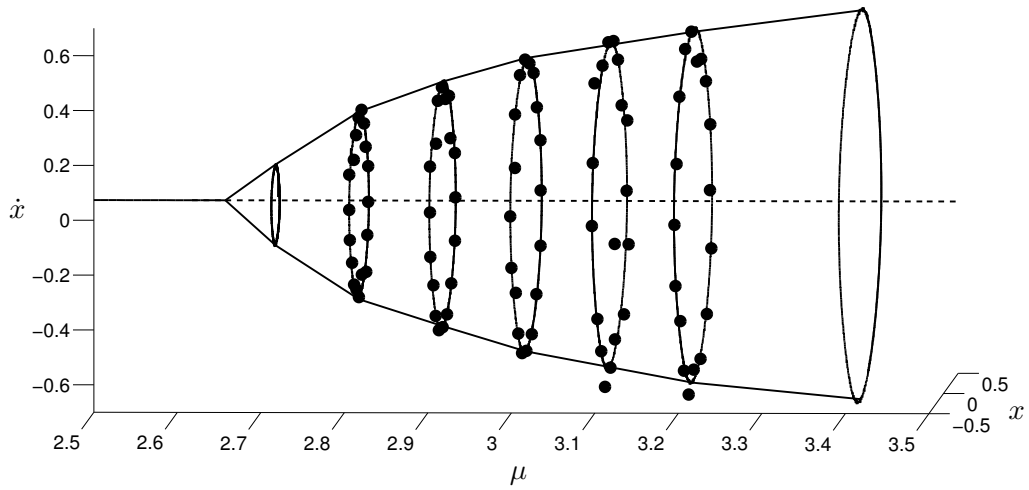

Figure 2.13: Prediction results for the entire bifurcation diagram. 50 recoveries were collected at each of two parameter values in the pre-bifurcation regime ($\mu = 2.4$ and 2.6). Lines represent measured bifurcation diagram with both the stable (solid) and the unstable (dashed) branches. Symbols represent prediction results at each parameter value. These results are obtained by first predicting the corresponding parameter values μ^* of certain states (x_1^*, x_2^*), and then interpolating the predicted results to the selected parameter values.

2.4 Discussion and Conclusions

We presented a new method that uses the transient recovery of parametrically excited systems from large perturbations to predict the bifurcation diagram. This forecasting method provides quantitative insights into parametrically excited systems despite the absence of actual models of these systems. As expected, the accuracy of forecasting increases when more data are collected closer to the critical transition. One of the important benefits of the method is that it is applicable in advance of the transition. The transition and the overall bifurcation diagram can be predicted without endangering the system by passing the transition point. We demonstrated experimentally and numerically that transitions caused by either bifurcation of the unforced system or parametric resonance can be predicted using this method.

This method method also works for subcritical bifurcations, although all the examples we presented in the paper have supercritical bifurcations. This is because we can still observe critical slowing down when subcritical bifurcation occurs. The difference is that for supercritical bifurcation, the dynamics is slowest at the zero equilibrium, while for subcritical bifurcation, the dynamics slow down the most at the amplitude of the saddle-node fixed point. Therefore, we need to use the relationship between the recovery rate and amplitudes, i.e. λ versus r, to study both the amplitude and the parameter of the jump point. Thus, two solutions that are created by one bifurcation can be predicted at the same time. The method, however, cannot predict multiple (more than two) solutions that are related to one bifurcation. In addition, our method cannot be applied to chaotic systems because we predict bifurcations by identifying steady state solutions through the critical slowing down phenomenon. Therefore, we use steady state solutions (fixed points or limit cycle oscillations) before and after the critical transition. A chaotic system does not meet these requirements.

The proposed method will work also for other periodically forced systems subject

to critical transition because of its model-less nature. Specifically, when a critical transition of a periodically forced system is caused by the bifurcation of the unforced system, a Poincaré section can be constructed by sampling recoveries once each excitation period. When the critical transition is caused by resonance, additional information about the system is required. The relationship between the frequency of the resonance dynamics and the excitation frequency is important for the construction of the Poincaré section.

Large perturbations are beneficial for this method because they provide information about the recovery rate of the system at large amplitudes, which enables the prediction of the bifurcation parameter corresponding to large-amplitude limit cycles or fixed points. Furthermore, operating with larger perturbations is beneficial to alleviate the effects of noise at lower amplitudes. Nevertheless, the proposed method can be used with small amplitudes also to predict the bifurcation point and the bifurcation diagram at small amplitudes when larger perturbations are not available.

Another important feature of the proposed method is its ability to predict the entire bifurcation diagram by studying the system dynamics in several different Poincaré sections. However, to gather the rich information provided by the method, more data will be needed in the pre-bifurcation regime, especially for large amplitudes. This is because interpolation is used to calculate λ_1 and λ_2 for the selected state with different values of the forcing phase. Recovery data is especially important to forecast the bifurcation diagram of subcritical bifurcations, because the stable region of the post bifurcation regime usually has large amplitudes. This also limits the method's applicability because for some systems, multiple perturbations at one parameter value may not be available, or they may be encountered only rarely. In addition, the existence of strong process noise might also limit the applicability of the method. Currently we are averaging over multiple recoveries to filter out the noise, but this method might not work for strong process noise. Currently the method can work out well when the

magnitude of the noise is one order of magnitude smaller than the damping in the system. However, when the noise is at the same order of magnitude as the damping, the prediction results are not very good. How to make this method more robust with process noise will be one of the things we will investigate in the future.

CHAPTER III

Forecasting bifurcations of multi-degree-of-freedom nonlinear systems with parametric resonance

3.1 Introduction

Parametric resonance is a phenomenon that can occur when one or more parameters vary periodically in time at a frequency related to the resonance frequencies of the system. Such resonances are observed in a wide range of engineering and natural systems [59, 72, 73] where large amplitude parametric resonance can be triggered by a relatively small-amplitude oscillation in parameters. There has been much interest in predicting the behavior of systems exposed to parametric resonance, because of the association between parametric resonance and large amplitude response, which can be dangerous in many engineering and natural systems [25, 74, 75]. Researchers have investigated this phenomenon in different applications to encourage or prevent its onset [13, 76].

Researchers have been using transient response to large perturbations for nonlinear system identification [66, 67]. This approach enables measurements at large amplitudes, where nonlinearities are more active. Periodic excitation is first applied so that the system oscillates at a certain amplitude. Excitation is then removed and free response data are collected. The transient data is then analyzed using Hilbert trans-

form or Gabor transform to estimate both the instantaneous amplitude and frequency of the system response, and thus identify both the bifurcation point and the bifurcation diagram of the system. Important parameters, such as damping and stiffness coefficient are then identified. Despite all the benefits, these identification methods rely on a relatively accurate model of the system, which limits their applicability.

To deal with the lack of an accurate model, model-less methods are proposed to identify backbone curves [77] or bifurcation diagrams [45, 7, 37] using transient recovery data. Bifurcation forecasting methods in particular, rely on the observation that the dynamics of many systems near critical transitions have a common property that the system recovers from perturbations ever more slowly as the system approaches critical points [16, 63, 18, 78]. Therefore, by monitoring the recovery rate of responses to large perturbations in the pre-bifurcation regime, these methods provide a prediction of the bifurcation point and the bifurcation diagram. The advantage of such methods is that they are model-less, which means that the bifurcation diagram can be extracted directly from the recovery data without the need for too many details of the system.

Bifurcation forecasting methods, however, are all focused on autonomous systems, without considering the effects of periodic excitation. For autonomous systems, the stability is determined by the eigenvalues of the linear part of the system at the equilibrium [63, 19]. For example, a critical transition of saddle-node type in an autonomous system is caused by the real part of an eigenvalue going from negative to positive. However, this is not necessary the same for systems with time varying input. The instability of a system with parametric resonance is not only associated with the linear part of the system, but also affected by nonlinearity [20, 79]. For multi-degree-of-freedom systems, the interaction between different linear modes of the system can also affect the onset of parametric resonance, and the bifurcation diagram [80]. A novel approach is proposed in this paper to deal with bifurcations

induced by near resonance parametric excitation. A procedure for the estimation of instantaneous amplitudes originated from the averaging theorem is proposed [14]. The approach transforms the non-autonomous parametrically excited system into an autonomous system, which enables the application of the method proposed in [45]. Transient recovery of the amplitudes of the oscillations are thus monitored and used to forecast not only the bifurcation point but also the whole bifurcation diagram. This is important for the understanding of how the amplitude of the oscillations changes with the parameter after the bifurcation. Quantitative knowledge of the branches of the bifurcation diagram can allow actions before the critical transition occurs, and can be helpful in designing systems that use/exhibit parametric resonance.

For multi-degree-of-freedom systems with parametric resonance, it is not realistic to monitor all of the degrees-of-freedom (DOFs). Therefore, the proposed method only collects data for one of the DOFs, and forecasts the bifurcation point and bifurcation diagram corresponding to that DOF. To that aim, we utilize center manifold theory [14]. When a codimension one bifurcation occurs, one or a pair of the eigenvalues of the equilibrium are on the imaginary axis. They are often referred to as the dominant eigenvalues. The dynamics associated with the dominant eigenvalues remain in an invariant set referred to as the center manifold. If the system satisfies the assumption that the manifold associated with the dominant eigenvalues varies slowly and smoothly with the parameter, when a system is randomly perturbed, then the dynamics in the stable manifold associated with other eigenvalues die out quicker than the dynamics in the slow manifold. Therefore, we can forecast the bifurcation diagram using the recoveries of a single variable in the slow manifold.

In this work, a double pendulum model is used to simulation data to demonstrate the method. We also have built a nonlinear electrical circuit that has a similar governing equation to that of the double pendulum. Two different types of bifurcation are investigated. The bifurcation parameter is the amplitude of parametric excitation

for the first type, and the detuning factor for the second type. Experimental data from the circuit is used to further examine the method.

3.2 Theory

This section presents the method to predict bifurcations of multi-DOF systems using time series collected in the pre-bifurcation regime. Consider a weakly nonlinear N-DOF system with parametric excitation and quadratic nonlinearity [79]. The governing equations are expressed as:

$$\begin{aligned}
\ddot{x}_1 + \omega_1^2 x_1 &= \epsilon[-\mu_1 \dot{x}_1 + \sum_{m,k}^{N} \alpha_{1mk} x_m x_k + \cos(\Omega t) \sum_{m=1}^{N} f_{1m} x_m] := f_1, \\
\ddot{x}_2 + \omega_1^2 x_2 &= \epsilon[-\mu_2 \dot{x}_2 + \sum_{m,k}^{N} \alpha_{2mk} x_m x_k + \cos(\Omega t) \sum_{m=1}^{N} f_{2m} x_m] := f_2, \\
&\cdots \\
\ddot{x}_N + \omega_N^2 x_N &= \epsilon[-\mu_N \dot{x}_N + \sum_{m,k}^{N} \alpha_{Nmk} x_m x_k + \cos(\Omega t) \sum_{m=1}^{N} f_{Nm} x_m] := f_N,
\end{aligned} \tag{3.1}$$

where ϵ is a small parameter; ω_n is the natural frequency of mode n; μ_n is the damping coefficient of mode n; α_{nmk} are the coefficients of the quadratic terms $x_m x_k$ of mode n; f_{nm} is the nth modal amplitude of parametric excitation and $n = 1, 2, ..., N$. Note that all the nonlinear and parametric excitation terms are small compared to the linear part of the system.

In this study, we focus on the case of a combination of a parametric resonance in one mode (chosen as mode 1, $\Omega \approx 2\omega_1$) and an internal resonance between two modes (chosen as models 1 and 2 $\Omega \approx \omega_2$). Hence, for $n \geq 3$, we assume that the nonlinear and parametrically excited terms in the governing equations do not produce resonance. Because this system is weakly nonlinear, we expect its response

to be approximately harmonic [14], expressed as

$$\begin{aligned}
x_1(t) &= A_1(t)\cos\left(\frac{\Omega}{2}t\right) + B_1(t)\sin\left(\frac{\Omega}{2}t\right) + O(\epsilon),\\
x_2(t) &= A_2(t)\cos(\Omega t) + B_2(t)\sin(\Omega t) + O(\epsilon),\\
x_3(t) &= A_3(t)\cos(\omega_3 t) + B_3(t)\sin(\omega_3 t) + O(\epsilon),\\
&\cdots\\
x_N(t) &= A_N(t)\cos(\omega_N t) + B_N(t)\sin(\omega_N t) + O(\epsilon),
\end{aligned} \tag{3.2}$$

for $n = 1, 2, ..., N$. For $n \geq 3$, there are no secular terms on the right hand side of the governing equations, which means there will be no resonance in those DOFs. Thus, the amplitudes $A_n(t)$ and $B_n(t)$ will tend to 0 as tends to infinity due to damping. Thus, after transients die out, dynamics in those DOFs will be small compared to dynamics in the first two DOFs. In this case, the stability of the steady state solutions of the first 2 DOFs will not be affected by modes higher than 2.

By applying the van der Pol transformation [14], we change coordinates for the first two DOFs to obtain

$$\begin{bmatrix} A_1 \\ B_1 \end{bmatrix} = \begin{bmatrix} \cos(\frac{\Omega t}{2}) & -\frac{2}{\Omega}\sin(\frac{\Omega t}{2}) \\ \sin(\frac{\Omega t}{2}) & \frac{2}{\Omega}\cos(\frac{\Omega t}{2}) \end{bmatrix} \begin{bmatrix} x_1 \\ \dot{x_1} \end{bmatrix}, \tag{3.3}$$

$$\begin{bmatrix} A_2 \\ B_2 \end{bmatrix} = \begin{bmatrix} \cos(\Omega t) & -\frac{1}{\Omega}\sin(\Omega t) \\ \sin(\Omega t) & \frac{1}{\Omega}\cos(\Omega t) \end{bmatrix} \begin{bmatrix} x_2 \\ \dot{x_2} \end{bmatrix}. \tag{3.4}$$

Using this coordinate transformation, the system (1) becomes

$$
\begin{aligned}
\dot{A}_1 &= -\frac{2}{\Omega}\left[\left(\frac{\Omega^2 - 4\omega_1^2}{4}\right)x_1 + \epsilon f_1\right]\sin\left(\frac{\Omega t}{2}\right), \\
\dot{B}_1 &= \frac{2}{\Omega}\left[\left(\frac{\Omega^2 - 4\omega_1^2}{4}\right)x_1 + \epsilon f_1\right]\cos\left(\frac{\Omega t}{2}\right), \\
\dot{A}_2 &= -\frac{1}{\Omega}\left[(\Omega^2 - \omega_2^2)x_2 + \epsilon f_2\right]\sin(\Omega t), \\
\dot{B}_2 &= \frac{1}{\Omega}\left[(\Omega^2 - \omega_2^2)x_2 + \epsilon f_2\right]\cos(\Omega t).
\end{aligned}
\tag{3.5}
$$

Because Ω is close to $2\omega_1$ and ω_2, let $\Omega^2 - 4\omega_1^2 = \epsilon 4\sigma_1$ and $\Omega^2 - \omega_2^2 = \epsilon\sigma_2$. Thus, according to the averaging theorem, the system (5) can be approximated using an autonomous system by averaging the dynamics over one period $T = \frac{2\pi}{\Omega}$ [14]. The associated autonomous averaged system is defined as

$$
\begin{aligned}
\dot{A}_1' &= \epsilon\frac{1}{T}\int_0^T -\frac{2}{\Omega}(\sigma_1 x_1 + f_1)\sin\left(\frac{\Omega t}{2}\right) := g_1(A_1', B_1', A_2', B_2', ...), \\
\dot{B}_1' &= \epsilon\frac{1}{T}\int_0^T \frac{2}{\Omega}(\sigma_1 x_1 + f_1)\cos\left(\frac{\Omega t}{2}\right) := g_2(A_1', B_1', A_2', B_2', ...), \\
\dot{A}_2' &= \epsilon\frac{1}{T}\int_0^T -\frac{1}{\Omega}(\sigma_2 x_1 + f_2)\sin(\Omega t) := g_3(A_1', B_1', A_2', B_2', ...), \\
\dot{B}_2' &= \epsilon\frac{1}{T}\int_0^T \frac{1}{\Omega}(\sigma_2 x_1 + f_2)\cos(\Omega t) := g_4(A_1', B_1', A_2', B_2', ...).
\end{aligned}
\tag{3.6}
$$

where A_n' and B_n' are the average of A_n and B_n over one period; g_i for $i = 1, 2, 3, 4, ...$ are autonomous functions of all the state variables A_n' and B_n' and all parameters. Therefore, if we can forecast the bifurcation of the system (6), we can find an approximation to the bifurcation of the system (1).

The bifurcation of the system (6) can be predicted using recovery data from perturbations in the pre-bifurcation regime. To obtain recovery data for the system,

perturbations are applied to system (1). For each period from time $T_{i-1} = \frac{2\pi(i-1)}{\Omega}$ to time $T_i = \frac{2\pi i}{\Omega}$ (where i is an integer), average amplitudes of the oscillatory recovery $A'_n(T_i)$ and $B'_n(T_i)$ can be estimated using least squares method to obtain,

$$\begin{bmatrix} A'_n(T_i) \\ B'_n(T_i) \end{bmatrix} = (X^T X)^{-1} X^T y, \tag{3.7}$$

where X and y are defined as

$$X = \begin{bmatrix} \cos(T_{i-1} + \delta t) & \sin(T_{i-1} + \delta t) \\ \cos(T_{i-1} + 2\delta t) & \sin(T_{i-1} + 2\delta t) \\ \cdots & \\ \cos(T_i) & \sin(T_i) \end{bmatrix},$$

$$y = \begin{bmatrix} x_n(T_{i-1} + \delta t) \\ x_n(T_{i-1} + 2\delta t) \\ \cdots \\ x_n(T_i) \end{bmatrix},$$

with $\frac{1}{\delta t}$ being the sampling frequency. The values of $A'_n(T_i)$ and $B'_n(T_i)$ are then used for forecasting.

Forecasting bifurcations of high dimensional autonomous systems using recovery data in the pre-bifurcation regime has been discussed in [37]. When a bifurcation occurs, the equilibrium is non-hyperbolic, and the associated linearized system can be divided into a stable subspace and a center subspace [14]. For co-dimension one bifurcation, there exists a one or two dimensional invariant manifold tangent to the center subspace at the equilibrium. This manifold is defined as the center manifold. For a high dimensional system, there is a certain parameter-dependent one or two

dimensional manifold in which the system exhibits a bifurcation, and this manifold is referred to as the slow manifold. The slow manifold coincides with the center manifold when the system is at the bifurcation point. Dynamics in the stable manifold can be neglected because of their faster decay rate compared to the dynamics in the slow manifold. An example of the slow manifold is shown in Fig. 3.1. When a perturbation is applied to the system, the dynamics in the fast manifold quickly die out. All recoveries from different random perturbations recover to the low dimensional slow manifold.

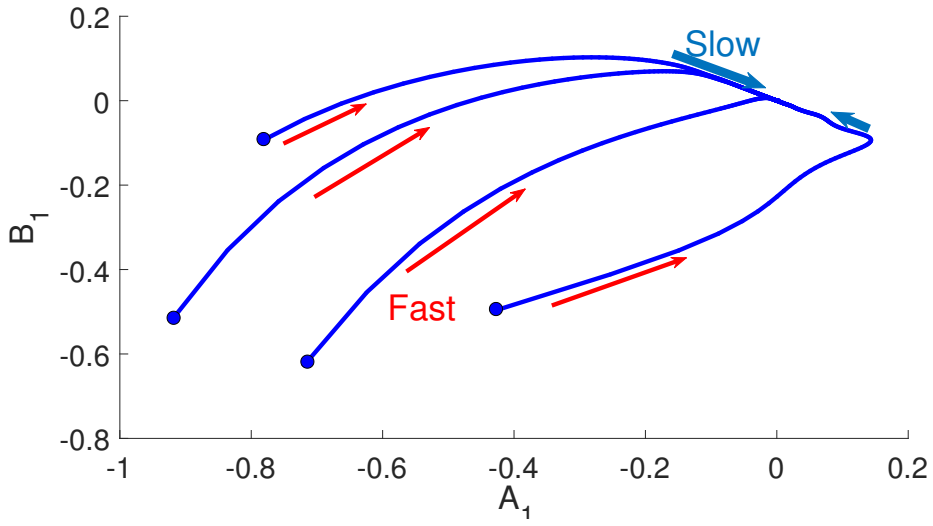

Figure 3.1: The slow manifold is low dimensional with dynamics slowest in time. When the system is perturbed, dynamics in the fast manifold will die out quickly. Recoveries from different perturbations converge to the low dimensional slow manifold. The diagram is generated by choosing a random initial condition for Eqns. 11, letting the system recover to equilibrium freely, collecting the recovery data of the first degree of freedom and applying the van der Pol transformation to obtain the recovery of amplitudes

$$A_1 \text{ and } B_1.$$

In [45], the method of forecasting bifurcations of one or two dimensional systems is developed. Therefore, to predict the bifurcations of high dimensional systems, we can first remove the transient recovery from perturbations that are not in the slow manifold, and then apply the method developed in [45] to forecast the bifurcation. An

important assumption for the proposed method is that the slow manifold associated with the dominant eigenvalue does not vary drastically with the parameter.

The forecasting method for a low dimensional system is briefly discussed next. Consider a one dimensional system

$$\dot{x} = f(x, \mu), \tag{3.8}$$

where x is the state variable, and μ is the bifurcation parameter.

The method is model-less, so the function $f(x, \mu)$ is not known. However, f is required to be smooth with respect to the bifurcation parameter μ. Next, define the recovery rate of the system at x as

$$\lambda(x, \mu) = \frac{\dot{x}}{x} = \frac{f(x, \mu)}{x}. \tag{3.9}$$

Since $f(x, \mu)$ varies smoothly with μ close to μ_c, λ also varies smoothly with μ close to μ_c. Therefore, a Taylor series of f with respect only to μ near the bifurcation point μ_c can be written as

$$\lambda(x, \mu) = \lambda(x, \mu_c) + (\mu - \mu_c)\frac{d\lambda(x, \mu_c)}{d\mu} + H.O.T, \tag{3.10}$$

where $H.O.T.$ represents high order terms in μ.

Note that $\lambda(0, \mu) = \lim_{x \to 0} \frac{\dot{x}}{x}$, is essentially the eigenvalue of the equilibrium at parameter value μ. Therefore, $\lambda(0, \mu_c) = 0$ at the bifurcation point. Meanwhile, for other points on the bifurcation diagram, denoted as (μ^*, x^*) as shown in Fig. 3.2 (c) , $\lambda(\mu^*, x^*)$ is also 0 because x^* is a fixed point at parameter value μ^*.

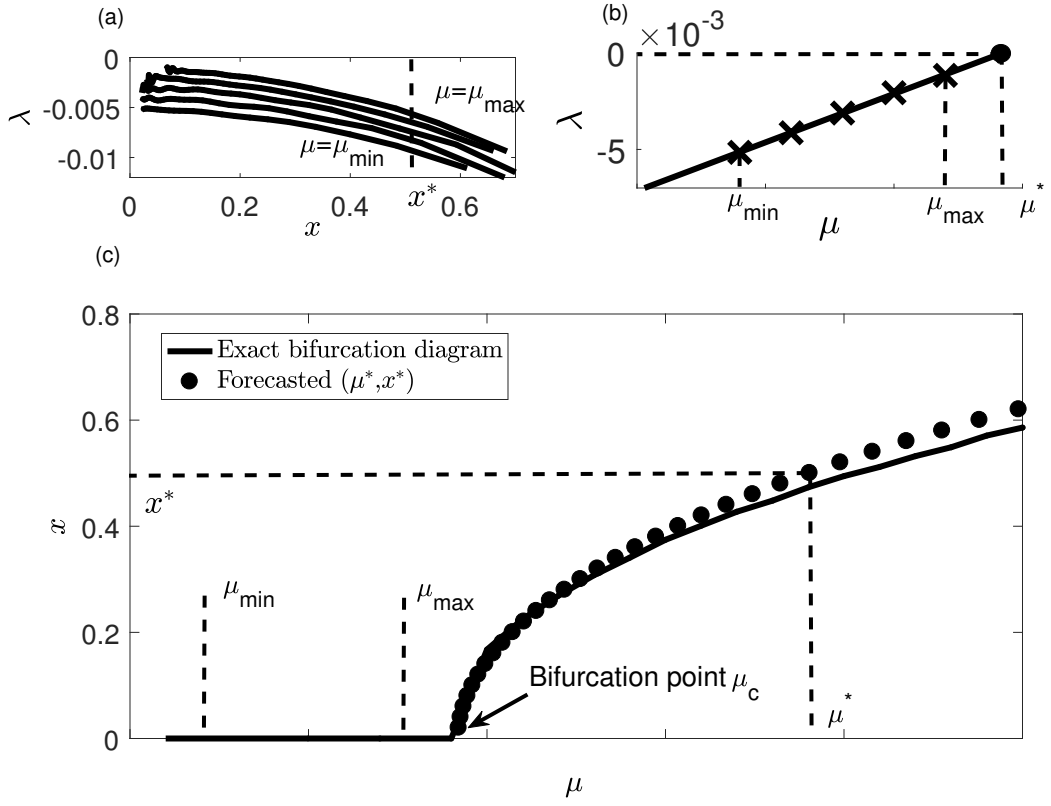

Figure 3.2: Recoveries from large perturbations can be used to forecast the bifurcation diagram. (a) Recovery rate $\lambda(x,\mu) = \frac{\dot{x}}{x}$ can be calculated for different amplitudes at different parameter values using the measured recovery data. (b) A linear or nonlinear function of λ versus μ can be fitted using λ for the same amplitude but different parameter values. Then, μ^* can be obtained by determining the parameter value at which $\lambda(x^*,\mu^*) = 0$. (c) The bifurcation diagram can be constructed by finding the corresponding parameter value μ^* for different values of x^*. The diagram is generated using Eqns (11) and the procedure outlined at the end of this section.

Thus, if we collect data from at least two parameters μ_1 and μ_2 in the pre-bifurcation regime, and calculate λ at x^* for μ_1 and μ_2, we can obtain

$$\begin{aligned}\lambda(x^*,\mu_1) &= \lambda(x^*,\mu_c) + (\mu_1-\mu_c)\frac{d\lambda(x,\mu_c)}{d\mu} + H.O.T.,\\ \lambda(x^*,\mu_2) &= \lambda(x^*,\mu_c) + (\mu_2-\mu_c)\frac{d\lambda(x,\mu_c)}{d\mu} + H.O.T.\end{aligned}$$

Neglecting $H.O.T.$, a linear function of λ versus μ can be fitted using μ_1, μ_2 and $\lambda(x^*, \mu_1)$, $\lambda(x^*, \mu_2)$. Then, μ^* can be obtained by determining the parameter value at which $\lambda(x^*, \mu^*) = 0$. Note that if data is collected at more than two parameter values, a linear or nonlinear function of recovery rate λ versus parameter μ can be fitted. Then, μ^* can be obtained by determining the parameter value at which $\lambda(x^*, \mu^*) = 0$ as shown in Fig. 3.2 (a) and Fig. 3.2 (b). The bifurcation diagram can be constructed by finding the corresponding parameter value μ^* for different state variable value x^* as shown in Fig.3.2 (c).

The whole method can be summarized as the following steps:

1. Choose at least two parameter values that are close to the bifurcation and at each parameter value measure the response of one of the state variables x_n as the system recovers from large perturbation in the pre-bifurcation regime.

2. Apply noise filtering techniques to the recovery of x_n. For each value of the bifurcation parameter μ, calculate the average of amplitude of oscillation $A'_n(t)$ and $B'_n(t)$ using the recovery data.

3. For each value of the bifurcation parameter μ, calculate λ for all the amplitudes $A'_n(t)$ and $B'_n(t)$. Approximate the dependence of λ versus $A'_n(t)$ and $B'_n(t)$ using a polynomial function.

4. Choose an amplitude value A'^*_n or B'^*_n where λ can be calculated from measurements at all parameter values, and use the relationship between λ and μ to find the corresponding parameter value μ^* for which A'^*_n or B'^*_n is a fixed point.

5. In a similar manner, choose different A'^*_n and B'^*_n values and repeat procedures at step 4 to predict the bifurcation diagram.

3.3 Results

We demonstrate the method using both computational data and experimental data to examine its applicability.

A double pendulum model [80, 2] is used as an example of a weakly nonlinear multi-degree-of-freedom system with parametric resonance. The resonance can be induced either by changing the amplitude of the parametric excitation or changing the detuning factor σ (i.e., the difference between the driving frequency and the resonance frequency). It is shown in the following sections how the forecasting method described in Section 2 can be used to predict the bifurcation point and the bifurcation diagram of the double pendulum system.

An electrical circuit is built also. The circuit mimics the behavior of the double pendulum model. Data in the form of voltages is collected from the circuit and the bifurcation diagram is forecasted.

3.3.1 Numerical Results

A sketch of the double pendulum system is shown in Fig. 3.3.

Let g be the gravitational acceleration, l_1 the moment of inertia of a hypothetical made of pendulum 1 and mass of pendulum 2 concentrated at P_2 about the pivot P_1, l_2 the moment of inertia of pendulum 2 about the pivot P_2, and m_1 and m_2 the mass of the whole system and the second pendulum. Then the governing equation of the double pendulum model is shown as follows [**?**],

$$\begin{aligned} &\ddot{\theta_1} + \omega_1\theta_1 + \epsilon[2\mu_1\dot{\theta_1} - \omega_1^2 F\theta_1\cos(\Omega t) + T_1(\dot{\theta_2}^2 + \theta_2\ddot{\theta_2} - \theta_1\ddot{\theta_2})], \\ &\ddot{\theta_2} + \omega_2^2\theta_2 + \epsilon[2\mu_2\dot{\theta_2} - \omega_2^2 F\theta_2\cos(\Omega t) - T_2(\dot{\theta_1}^2 - \theta_2\ddot{\theta_1} + \theta_1\ddot{\theta_1})]. \end{aligned} \tag{3.11}$$

where $\omega_1^2 = \frac{m_1 ag}{l_1}$, $\omega_2^2 = \frac{m_2 cg}{l_2}$, $T_1 = \frac{m_2 bc}{l_1}$, $T_2 = \frac{m_2 bc}{l_2}$, and ϵ is a small constant.

Here, a is the vertical distance of the center of mass of pendulum 1 below P_1, c is the vertical distance of the center of mass of pendulum 2 below P_2, and b is the horizontal separation of P_1 and P_2 when the system is at static equilibrium.

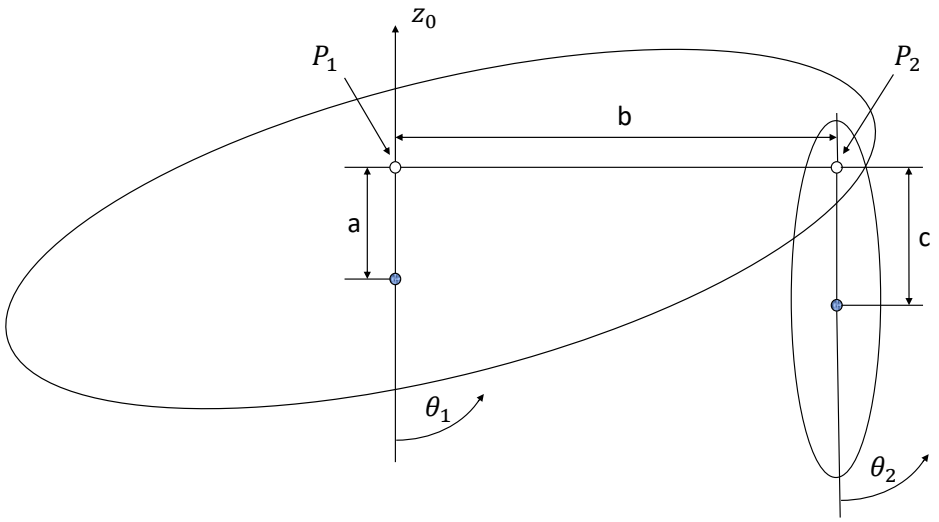

Figure 3.3: The double pendulum in [2]. Parametric excitation is applied through the vertical translation of the pivot P_1 of the dominant pendulum. P_1P_2, the line joining the two pivots, is horizontal at the equilibrium position.

For this system, different types of resonances can be obtained by choosing different relationships between the natural frequencies ω_1, ω_2 and the excitation frequency Ω. In this work, we choose $\omega_2 = 2\omega_1 + \epsilon\sigma_1$ and $\Omega = 2\omega_1 + \epsilon\sigma_2$ so that there exists both a parametric resonance, and a two-to-one internal resonance. In the simulation, parameter values are: $\mu_1 = 0.2$, $\mu_2 = 1.2$, $\sigma_1 = 0.2$, $\epsilon = 0.1$ and $T_1 = T_2 = 1$.

Two different types of bifurcations are studied in this section. For the first type of bifurcation, the amplitude of parametric excitation is chosen as the bifurcation parameter, in which case σ_2 is selected as 0.2. For the second type detuning factor σ_2 is the bifurcation parameter, and F is selected as 1.2.

3.3.1.1 Excitation amplitude as bifurcation parameter

The proposed method utilizes measurements of the response after perturbations in the pre-bifurcation regime to forecast the bifurcation diagram. The initial conditions of the transient response can be important for the forecasting results. When the perturbation occurs along or close to the slow manifold, the dynamics in the stable manifold die out quickly, in which case the system can reach the slow manifold at a large amplitude.

In practice, the system is at steady state when a perturbation occurs. To create a perturbation we change the bifurcation parameter value, which changes the steady state. The bifurcation parameter is then set back to its actual pre-bifurcation value, and the recovery is recorded. Note that the initial steady state response should be large enough to excite the nonlinearity which is essential for predicting the whole bifurcation diagram.

An example of a recovery is shown in Fig. 3.4. Bifurcation parameter F (pre-bifurcation regime) is 0.72. Then it is changes to 2.4 to create a perturbation. Finally, F is set back to 0.72. Data is recorded only when the parameter is 0.72 (pre-bifurcation). Using the recovery data of θ_1 and θ_2, we can calculate the average amplitude of variation using Eq. 3.8 as shown in Figs. 3.4(c) and (d). The recovery of θ_2 is just shown here for illustration. It is not used to predict the bifurcation diagram. The value of $\lambda = \frac{\dot{A}_1}{A_1}$ is then extracted from the recovery of A_1 and shown in Fig. 3.5(c). This process is repeated for 3 different parameter values, which are $F = 0.64$, $F = 0.72$ and $F = 0.8$. Values of λ corresponding to the same amplitude but different parameter values are used to forecast both the bifurcation point and the bifurcation diagram in the post-bifurcation regime. In Fig. 3.5(b), values of λ obtained for $A_1 = 0$ are used to forecast the bifurcation point.

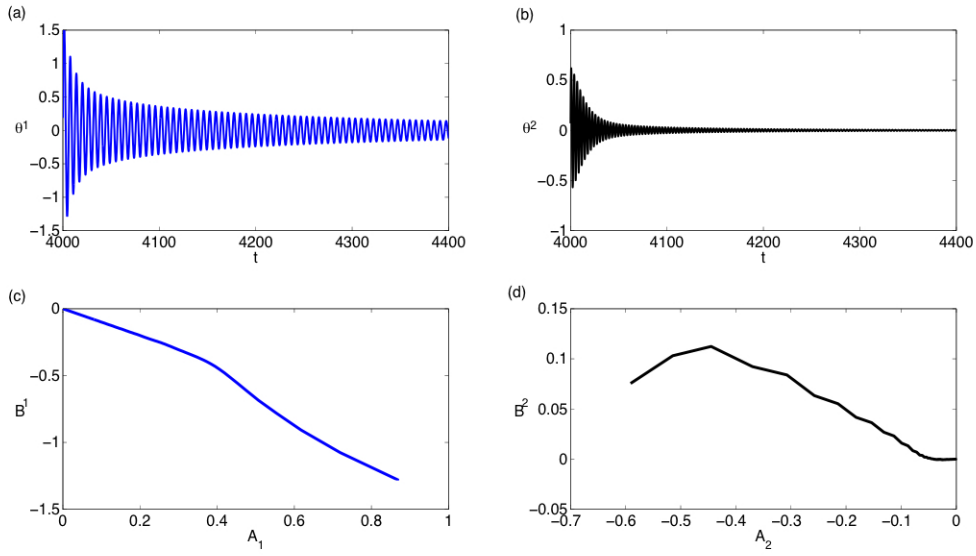

Figure 3.4: Recovery to the steady state response. Parameter value F is first set 0.72 (pre-bifurcation regime). Then it is changed to 2.4 to create a perturbation. Finally, F is set back to 0.72 to obtain the recovery data collected only at $F = 0.72$. (a) Recovery of θ_1 to the steady state response. (b) Recovery of θ_2 to the steady state response. (c) Phase portrait of recovery of A_1 and B_1. (d) Phase portrait of recovery of A_2 and B_2.

This process is then repeated by selecting a different amplitude value A_1 and forecasting its corresponding bifurcation parameter. The whole bifurcation diagram of A_1 versus excitation amplitude F is thus predicted and compared to the real bifurcation diagram as shown in Fig. 3.5(e). The bifurcation diagram of B_1 is also predicted in a similar fashion and shown in Fig. 3.5. The results reveal a strong agreement between the forecasted bifurcation diagram and the actual bifurcation diagram.

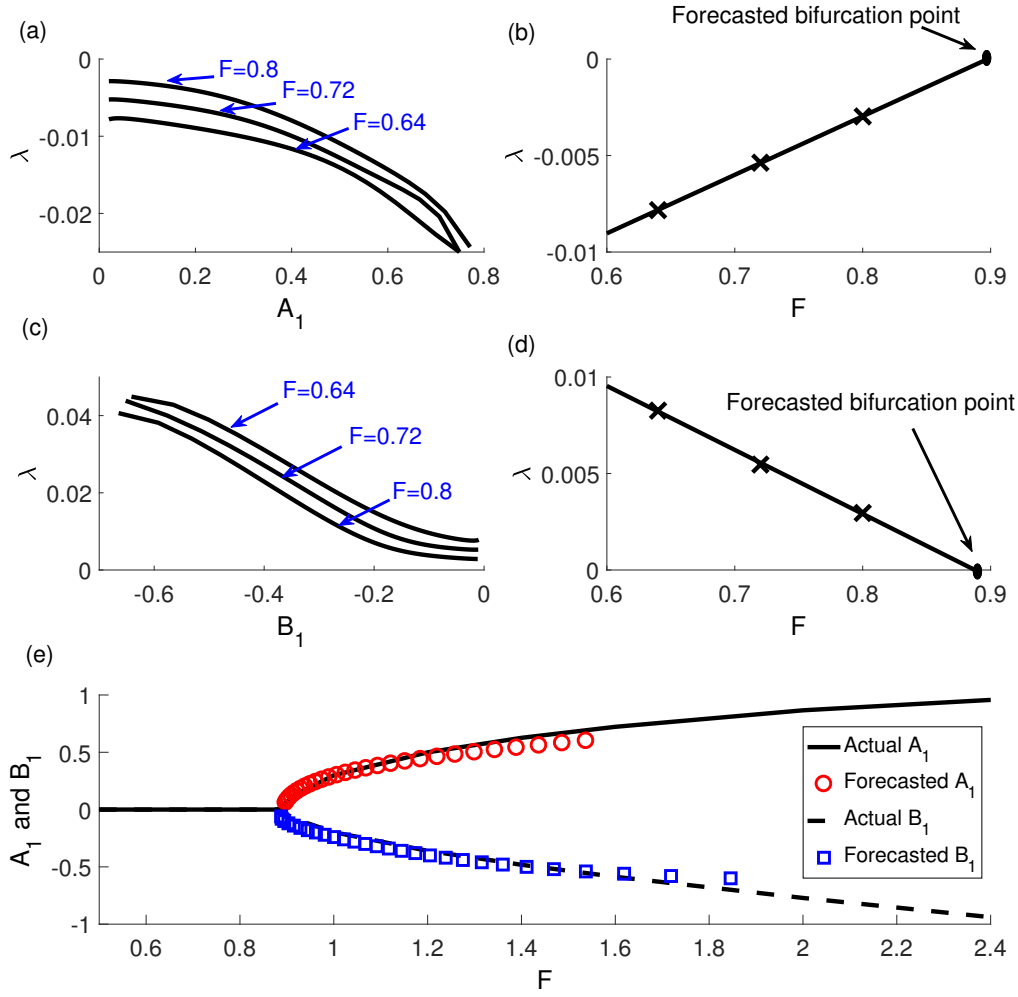

Figure 3.5: (a) λ versus A_1 at three different parameter values in the pre-bifurcation regime. (b) λ versus the parameter value F when $A_1 = 0$. The bifurcation point is predicted by fitting a function through the three points. (c) λ versus B_1 at three different parameter values in the pre-bifurcation regime. (d) λ versus the parameter value F when $B_1 = 0$. The bifurcation point is predicted by fitting a function through the three points. (e) Predicted bifurcation diagram of both A_1 and B_1 versus the parameter value F.

We also tested the method for systems with measurement noise. 5%, 10% and 20% relative noise were added to the observed data, and the same algorithm was used to forecast the bifurcation diagram. Forecasting results with noisy data are shown in Fig. 3.6. Only one noisy recovery was collected at each of the 3 parameter values F in the pre-bifurcation regime. Fig. 3.6(c) shows a relatively small variation of the

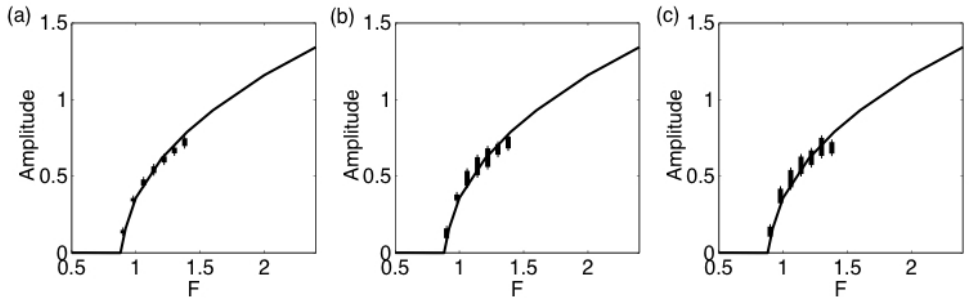

Figure 3.6: The exact bifurcation diagram is shown with a solid line in each plot. Transient data were collected in the pre-bifurcation regime. The predicted post-bifurcation regime is shown together with the standard deviation error bars. (a) 5%, (b) 10% and (c) 20% measurement noise were added to all recoveries.

forecasting results even when 20% relative measurement noise was added to the data.

3.3.1.2 Detuning factor as the bifurcation parameter

Parametric resonance can also be reached by changing the detuning factor. In this section, we use the same method to predict the bifurcation diagram but here we focus on the detuning factor as bifurcation parameter. This bifurcation diagram is essentially the amplitude frequency response curve for the system.

An example of a recovery from resonance response is shown in Fig. 3.7. Bifurcation parameter σ_2 is set to 0.5 (pre-bifurcation regime). Then σ_2 is changed to 0.3 to create a perturbation. Finally, σ_2 is set to 0.3 and the recovery data is collected. Using the recovery data of θ_1 and θ_2, we can calculate the average amplitude of variation using Eq. 3.8 as shown in Fig. 3.7(c) and Fig. 3.7(d).

The value of $\lambda = \frac{\dot{A}_1}{A_1}$ is then extracted from the recovery of A_1 and shown in Fig. 3.7(c). This process is repeated for 3 different parameter values, which are $\sigma_2 = 0.46$, $\sigma_2 = 0.48$ and $\sigma_2 = 0.5$. Values of λ corresponding to the same amplitude value but different parameter values are used to forecast both the bifurcation point and the bifurcation diagram in the post-bifurcation regime. In Fig. 3.7(b), values of λ for $A_1 = 0$ are used to forecast the bifurcation point.

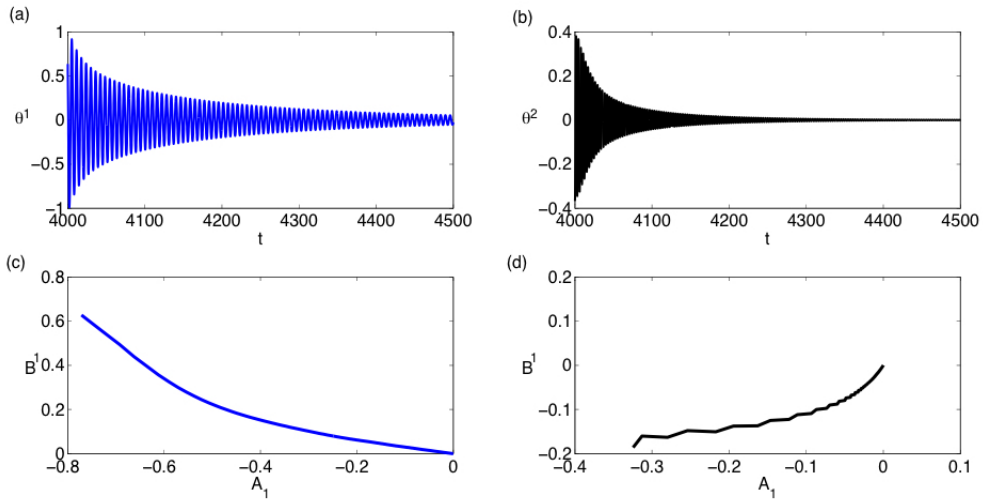

Figure 3.7: Recovery to the steady state response. Parameter value σ_2 is first set 0.5 (pre-bifurcation regime). Then it is changed to 0.3 to create a perturbation. Finally, σ_2 is set back to 0.5 to obtain the recovery data collected only at $\sigma_2 = 0.5$. (a) Recovery of θ_1 to the steady state response. (b) Recovery of θ_2 to the steady state response. (c) Phase portrait of recovery of A_1 and B_1. (d) Phase portrait of recovery of A_2 and B_2.

This process is then repeated by selecting a different amplitude value A_1 and forecasting its corresponding bifurcation parameter. The whole bifurcation diagram of A_1 versus detuning factor σ_2 is thus predicted and compared to the real bifurcation diagram as shown in Fig. 3.5(e). The bifurcation diagram of B_1 can be calculated in a similar fashion and shown in Fig. 3.5. The results reveal a strong agreement between the forecasted bifurcation diagram and the actual bifurcation diagram.

We also tested the method for systems with measurement noise. 5%, 10% and 20% relative noise were added to the observed data, and the same algorithm was used to forecast the bifurcation diagram. Forecasting results are shown in Fig. 3.9. Only one noisy recovery was collected at each of three parameter values σ_2 in the pre-bifurcation regime. The forecasted results have slightly larger amplitudes than the actual bifurcation diagram, but the bifurcation point and the shape of the bifurcation diagram is well captured in all cases.

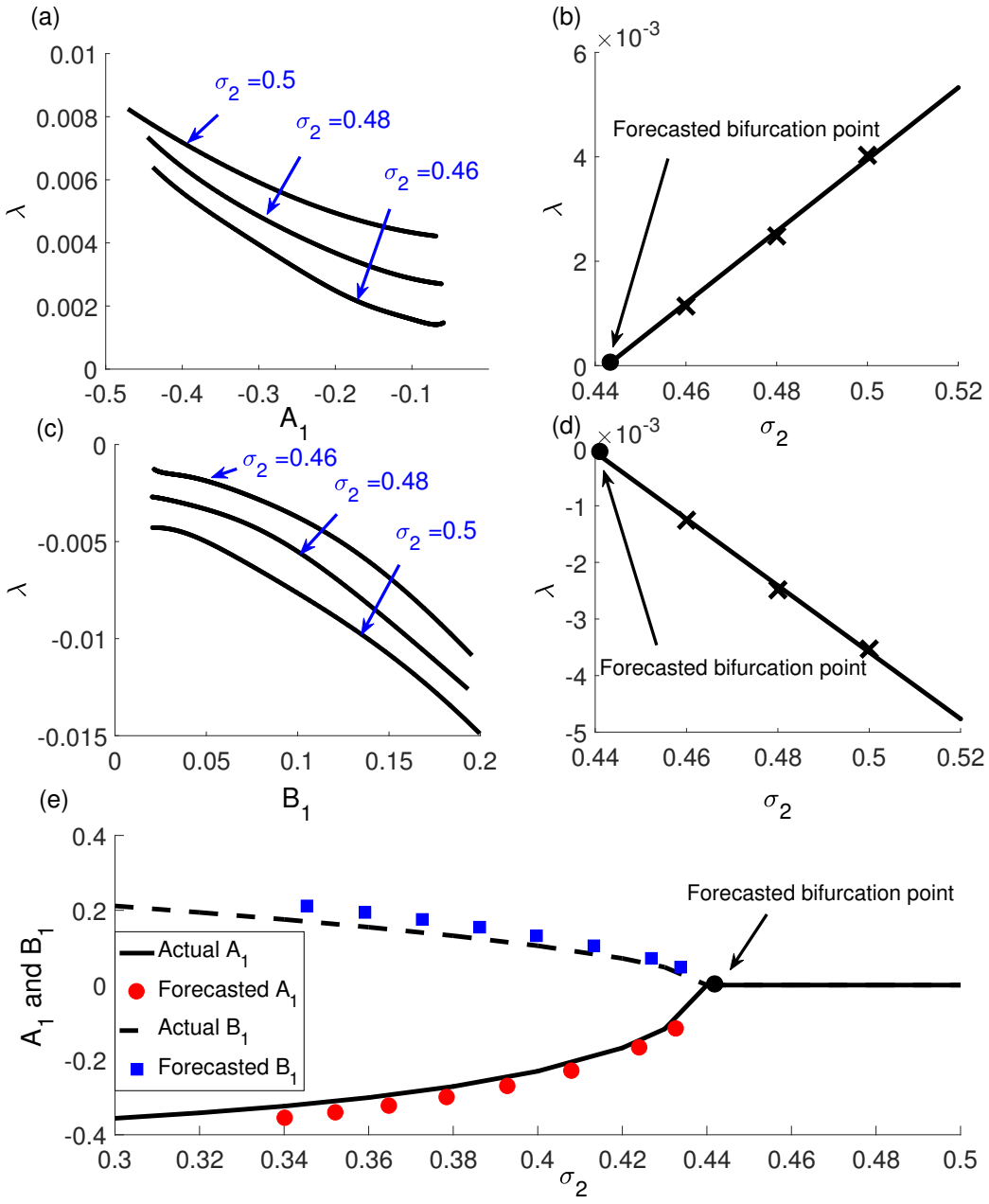

Figure 3.8: (a) λ versus A_1 at three different parameter values in the pre-bifurcation regime. (b) λ versus the parameter value σ_2 when $A_1 = 0$. The bifurcation point is predicted by fitting a function through the three points. (c) λ versus B_1 at three different parameter values in the pre-bifurcation regime. (d) λ versus the parameter value σ_2 when $B_1 = 0$. The bifurcation point is predicted by fitting a function through the three points. (e) Predicted bifurcation diagram of both A_1 and B_1 versus the parameter value σ_2.

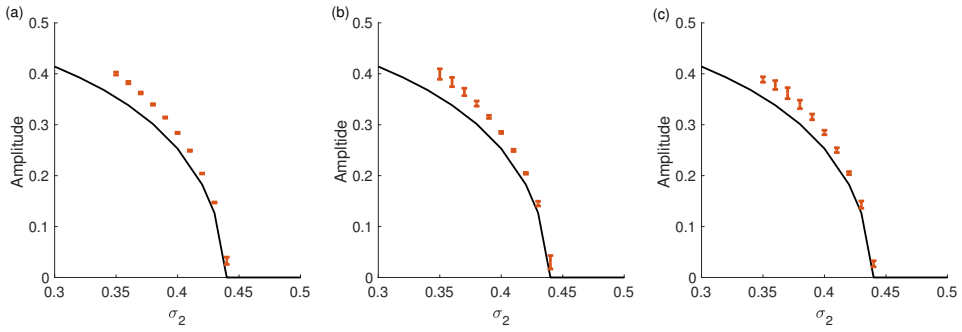

Figure 3.9: The exact bifurcation diagram is shown with a solid line in each plot. Transient data were collected in the pre-bifurcation regime. The predicted post-bifurcation regime is shown together with the standard deviation error bars. (a) 5%, (b) 10% and (c) 20% measurement noise were added to all recoveries.

3.3.2 Experimental results

To demonstrate the applicability of the forecasting method to systems with both measurement and process noise, an electrical circuit was built that mimics the behavior of the double pendulum model. The electrical circuit is shown schematically in Fig. 3.10, and as constructed in Fig. 3.11. The value of each component is provided in Fig. 3.11. Some potentiometers used in the circuit are not shown in Fig. 3.10 for clarity. They allow the control of several parameters such as the damping coefficient. The operational amplifiers are model TL082 and the multipliers are model AD633.

An example of recovery from a perturbation is shown in Fig. 3.12. The bifurcation parameter F is first set to 1.4 (pre-bifurcation). Then, F is changed to 2.0 to create a bifurcation. Finally, F is then set back to 1.4 in the pre-bifurcation regime and the recovery data is collected. In the experiment, only the recovery data of θ_1 is collected to show the method can predict the bifurcation diagram of one of the degrees of freedom without using data of other degrees of freedom. Using the recovery data of θ_1, we can calculate the average amplitude of variation using Eq. 3.8 as shown in Fig. 3.12(b).

The value of $\lambda = \frac{\dot{A}_1}{A_1}$ is then extracted from the recovery of A_1 and shown in Fig. 3.12(c). This process is repeated for 5 different parameter values, which are

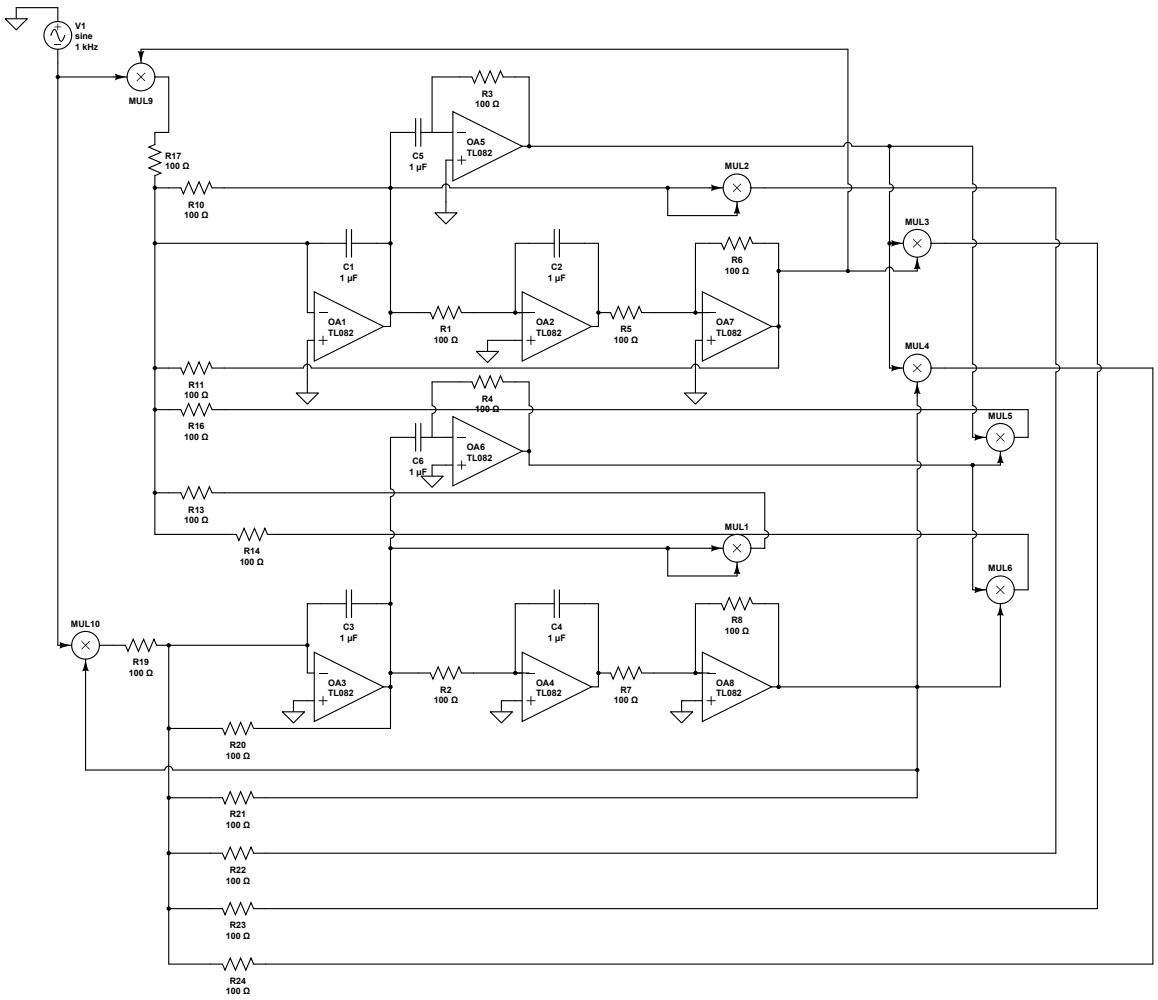

Figure 3.10: Schematic of an analog circuit realization of the double pendulum model.

$F = 1.34$, $F = 1.36$, $F = 1.38$, $F = 1.40$ and $F = 1.42$. Values of λ for the same amplitude value but different parameter values are used to forecast both the bifurcation point and the bifurcation diagram in the post-bifurcation regime. In Fig. 3.12(b), values of λ for $A_1 = 0$ are used to forecast the bifurcation point as shown.

This process is then repeated by changing the amplitude value A_1 and forecasting its corresponding bifurcation parameter. The whole bifurcation diagram of A_1 versus detuning factor F is thus predicted and compared to the real bifurcation diagram as shown in Fig. 3.13(e). The bifurcation diagram of B_1 can be calculated in a similar

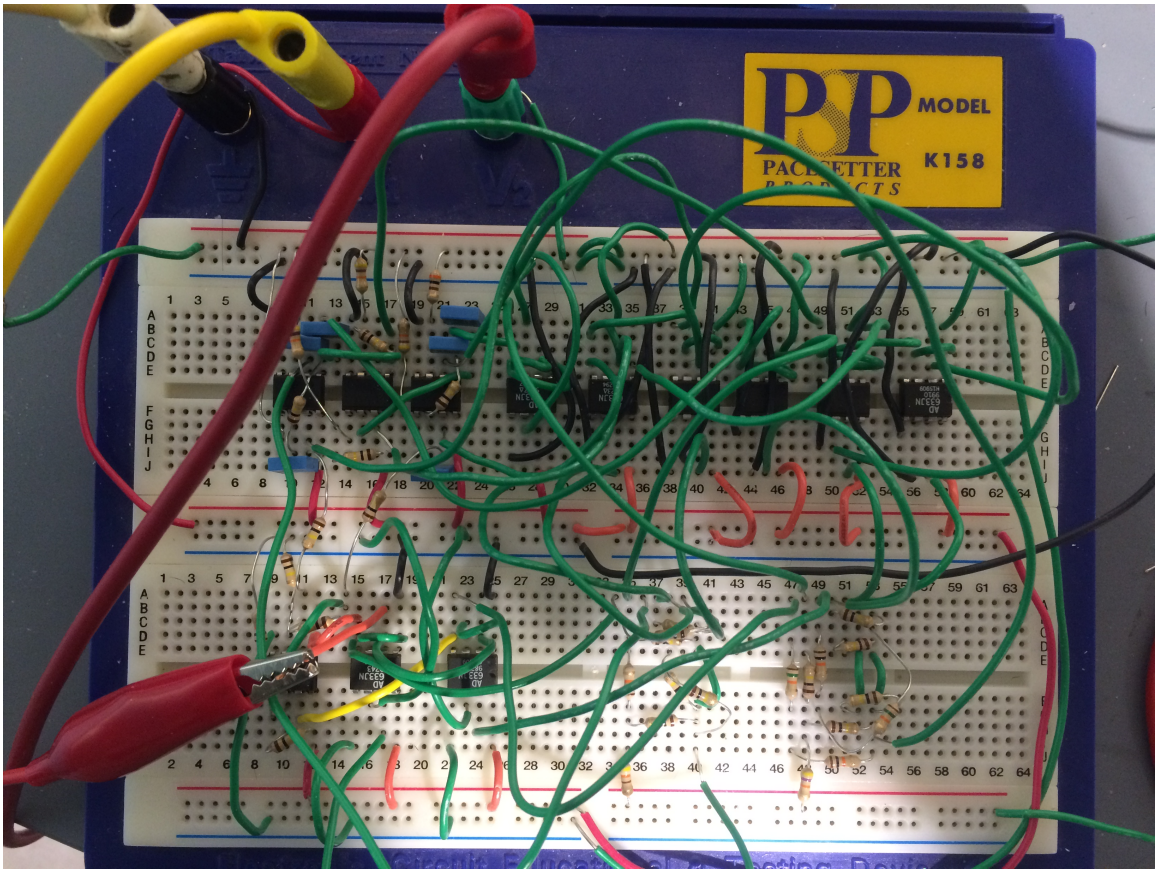

Figure 3.11: The analog circuit as constructed in the laboratory.

fashion and shown in Fig. 3.13.

3.4 Discussion

The proposed method relies on observations of transient recoveries from large perturbations in the pre-bifurcation regime to forecast the bifurcation point and the bifurcation diagram. After removing the recoveries perpendicular to the center manifold, the remaining part of the recoveries are along the slow manifold and can be used to forecast the bifurcation diagram. Therefore, the forecasted results, especially at large amplitudes, are more accurate if the initial condition is close to the slow manifold.

In the paper, a large response is used as an initial condition from which the system

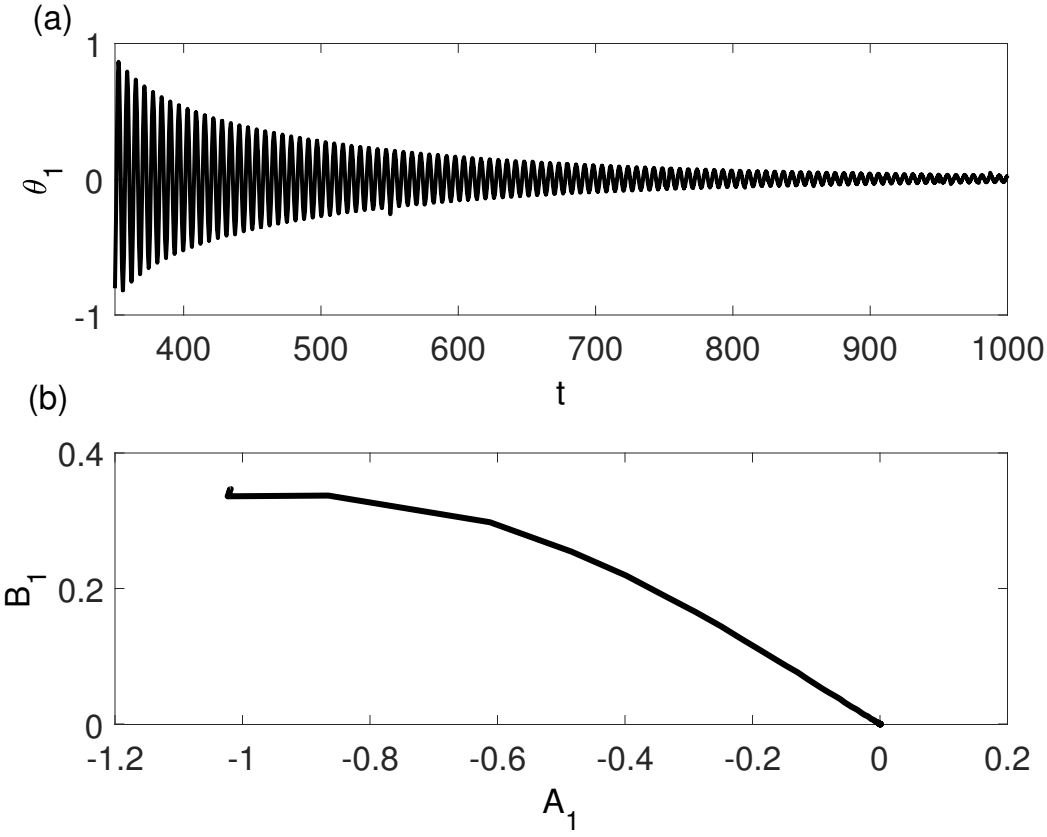

Figure 3.12: Recovery to the steady state response from a perturbation caused by parametric variation. Parameter value F is first set 1.4 (pre-bifurcation regime). Then it is changed to 2 to create a perturbation. Finally, F is set back to 1.4 to obtain the recovery data collected only at $F = 1.4$. (a) Recovery of θ_1 to the steady state response. (b) Phase portrait of recovery of A_1 and B_1.

will recovers to the dynamics close to the slow manifold. In fact, the large response can be exactly on the slow manifold if the perturbation is created by a parameter variation and the slow manifold does not change with the bifurcation parameter. A recovery from an initial condition is shown in Fig. 3.14(a) and Fig. 3.14(b). Several different phase portraits of recoveries of A_1 and B_1 from different initial conditions are shown in Fig. 3.14(c) and phase portraits of A_2 and B_2 in Fig. 3.14(d). It is clear that the system recovers to the one-dimensional manifold at quite small amplitudes, especially for A_2 and B_2. This makes it difficult to forecast the bifurcation diagram, especially at large amplitudes. Therefore, it is important that perturbed states created by

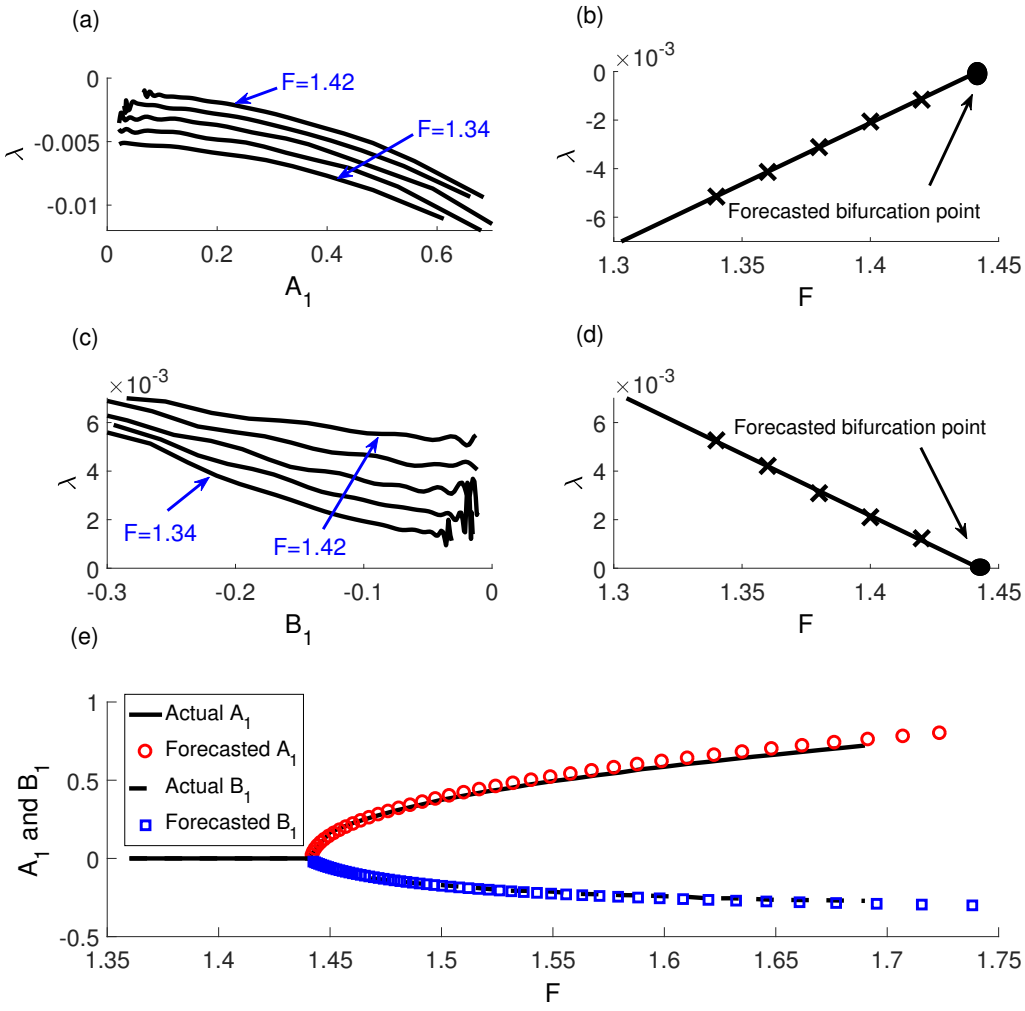

Figure 3.13: (a) λ versus A_1 at 5 different parameter values in the pre-bifurcation regime. (b) λ versus the parameter value F when $A_1 = 0$. The bifurcation point is forecasted by fitting a function through the 5 points. (c) λ versus B_1 at 5 different parameter values in the pre-bifurcation regime. (d) λ versus the parameter value F when $B_1 = 0$. The bifurcation point is forecasted by fitting a function through the 5 points. (e) Forecasted bifurcation diagram of both A_1 and B_1 versus the parameter value F.

parameter variation are used as initial conditions for optimal forecasting results.

When a resonance response is used as the initial condition, the system recovers to the slow manifold quickly. Therefore, the prediction is accurate except at amplitudes close to the initial condition, as shown in Figs. 3.5, 3.8 and 3.13.

There are certain key assumptions used in the proposed forecasting method. The first is that the bifurcation diagram varies smoothly with the bifurcation parameter. Additionally, if the slow manifold varies drastically with the bifurcation bifurcation parameter, the forecast does not work well either. It is also important to mention that the method currently only works for co-dimension one bifurcations.

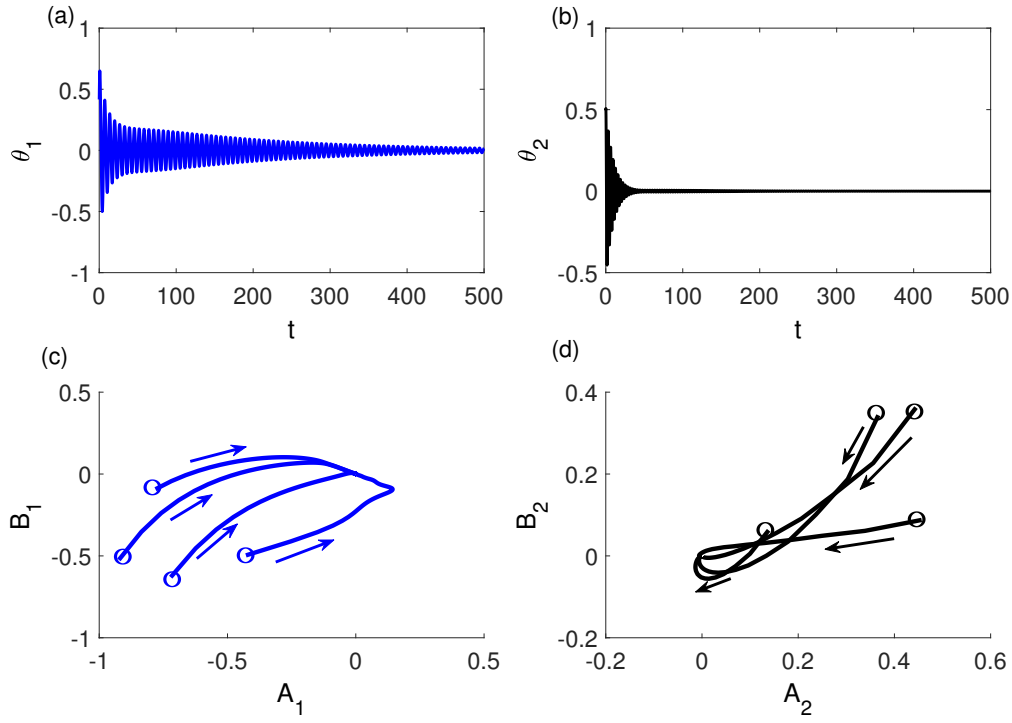

Figure 3.14: Recovery from different initial conditions. (a) Recovery of θ_1 from an initial condition. (b) Recovery of θ_2 from an initial condition. (c) Phase portrait of recovery of A_1 and B_1 from 4 different initial conditions. (d) Phase portrait of recovery of A_2 and B_2 from 4 different initial conditions.

One of the key advantages of the proposed forecasting method is that it is model-less. This feature contrasts all existing methods. For example, common approach used to predict the bifurcation diagram as to build a model of the system, identify its parameters and then simulate the model to compute the bifurcation diagram. That approach only works when an accurate model is available. However, that is not always the case. For example, we built a model using the values of the components of the

electrical circuit in Fig. 3.10. We then used that model to compute the bifurcation diagram. The bifurcation diagrams of the model-based approach and the experiment are not the same as shown in Fig. 3.15. The difference between these two diagrams is especially large at lower amplitudes. This difference is caused by the discrepancy between the theoretical and the real output of the multiplier used in the electrical circuit. The multiplier is more accurate at large amplitudes, and hence, the two bifurcation diagrams (model and actual) are in better agreement farther in the post-bifurcation regime, where amplitudes are larger. Note that this discrepancy between the experiment and the equation is not in the parameters used in the equation, but in the actual functional form of the behavior of the circuit at low amplitudes. Nonetheless, the forecasting approach proposed predict accurately the actual bifurcation diagram.

This difference in the bifurcation diagram essentially points out one advantage of the proposed forecasting method. For a multi-degree-of-freedom nonlinear system with parametric resonance, the bifurcation diagram, especially the bifurcation point, is strongly affected by the nonlinearity. When the model is not accurate enough, as in the case of the electrical circuit, the bifurcation point predicted by the model can be far away from the real bifurcation point. However, the proposed model-less forecasting method can accurately capture the nonlinearity at lower amplitudes, and predict the bifurcation point, as shown in Section 3.2.

3.5 Conclusions

The prediction of parametric resonance for multi-degree-of-freedom systems is an important topic. In this paper, a forecasting method is proposed to forecast both the bifurcation point and the bifurcation diagram in the post-bifurcation regime using recoveries in the pre-bifurcation regime.

Several assumptions are needed for the method to work. First, the system has to be a weakly nonlinear system for the averaging approximation to be valid. Second,

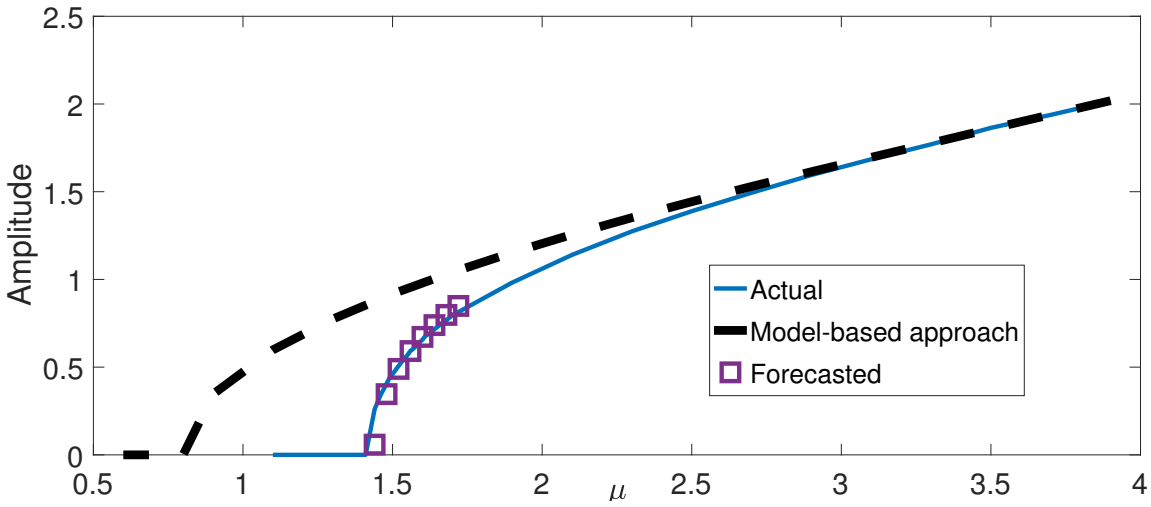

Figure 3.15: Bifurcation diagrams obtained from forecasting, experiments and model-based prediction.

data should be collected at parameter values close to the bifurcation point so that the extrapolation of λ will be accurate. Other assumptions include that the average system has a co-dimension one bifurcation, and the slow manifold varies smoothly with the parameter.

In the paper it is shown that if recovery can be obtained only for one of the DOFs, the bifurcation diagram of that degree of freedom can still be predicted. It is also discussed that initial perturbations play an important role in this method. The closer the condition after the perturbation is to the slow manifold, the better the forecasting results are.

Future work will include exploring the applicability method for systems with more complex nonlinear interaction and higher-DOFs.

CHAPTER IV

Regular biennial cycles in epidemics caused by parametric resonance

4.1 Introduction

Seasonal environmental changes can have strong effects on the dynamics of infectious diseases [81, 82, 4, 3]. Most seasonal variations are annual, and empirical evidence shows that such annual seasonality can cause oscillations ranging from annual cycles to multiyear cycles, and even chaotic dynamics [83, 3]. In the case of measles epidemics, both regular and chaotic cycles were observed during the past century in many cities throughout the world [84, 85, 3]. Regular biennial pattern was observed to be the most stable one, lasting for more than a decade in some cities, and causing major epidemics every other year. Such biennial cycles have aroused particular interest due to its large amplitude and persistence observed in disease case reports [86, 87, 88].

Large amplitude multiyear oscillations, which are also identified as subharmonic resonances, are formed from the instability caused by the interaction between the natural mode of the system and the external force [83]. The study of subharmonic resonances in childhood infectious diseases using SIR based models has resulted in a rich variety of interesting mathematical results that show good agreement with empir-

ical evidences [83, 4, 27, 89]. In particular, large biennial cycles in yearly forced SIR-based models usually appear when the external driving frequency is approximately twice the natural frequency of the system [83]. This phenomenon is commonly known as parametric resonance. Models from the deterministic SIR family all show qualitatively similar dynamical properties and parametric resonance might occur when parameters of the system fall into certain region [90, 91, 3]. One important question thus raised is: What determines the onset of biennial cycles?

The onset of parametric resonances in SIR based models have been studied in the past by varying certain parameters in numerical simulations. Dietz [91] observed a transition from an annual cycle to a biennial cycle by increasing the amplitude of seasonality. Aron and Schwarz [90] extended this result to an SEIR model. Earn et al. [3] varied the average contact rate that results in a transition from a biennial cycle to a small amplitude chaotic dynamics. Kuznetsov and Piccardi [89] studied the consequences of varying both amplitude of seasonality and transmission rate. Despite the rich results, most of these studies are carried out by direct numerical integration, which shows the dynamics but obscures the occurrence of parametric resonance caused by the interaction between model nonlinearity and external forcing. Black and McKane [92] carried out an analysis where the stochastic SIR model with periodic forcing is formulated as a master equation and studied using Van Kampen's expansion [93]. They show how parametric resonance, referred to as period-doubling bifurcation in the paper, occurs as R_0 varies. However, this method is essentially linear and the linear approximation breaks down near the bifurcation point and at large amplitudes. Therefore, the understanding of the role of the interaction between nonlinearity and periodic forcing in the onset of parametric resonance is incomplete.

Parametric resonances have been studied extensively for nonlinear mechanical and electrical systems [79, 20] using perturbation methods. In this paper, the method of multiple scales (MMS) is used to construct approximations to the solutions of SIR-

based models. MMS is carried out by formulating an independent variable (infectious fraction I for instance) as the sum of a fast-scale and several slow-scale variables, and treat them as if they are independent. Then, the system is studied at all scales, coupled one by one, so as to find an approximation solution. Using MMS, we can observe how resonance is caused by instability of the interaction between the natural mode of the system and the external force. Transition curves are also obtained to separate the parameter space into a resonance regime and a non-resonance regime. The onset of parametric resonance is caused by crossing the transition curve in the parameter space. We will show that cycles of large amplitudes can be caused by much lower excitations than commonly believed [91, 94, 57] when the driving frequency is close to twice the natural frequency of the system.

Sections 2.1 and 2.2 present the MMS approach to find approximate solutions for SIR-based models. Section 2.3 introduces the transition curves and periodic solutions obtained from analysis in different parameter regions. Section 3 presents the effects of varying different parameters on predicted periodic solutions. Section 4 broadens the discussion by introducing stochasticity to the problem and exploring the effect of noise on the onset of parametric resonance. It can be observed numerically that stochasticity can further lower the threshold of excitation amplitude that leads to larger amplitudes by pushing the system from one deterministic attractor onto another.

4.2 Modeling and analysis

In this section, we first describe the parametrically excited SIR model, and then show briefly how we use the MMS approach to obtain an approximate solution. More details of the approach can be found in the Supplementary Information.

4.2.1 SIR model

For simplicity we restrict our analysis to an SIR model, expressed as,

$$\begin{aligned} \frac{dS}{dt} &= \mu - \mu S - \beta IS, \\ \frac{dI}{dt} &= \beta SI - \gamma I - \mu I, \\ \frac{dR}{dt} &= -\mu R + \gamma I, \end{aligned} \tag{4.1}$$

where S, I and R denote susceptible, infectious and recovered fractions. Because $S + I + R = 1$, only two of the variables are independent, so the system can be reduced to the first two equations. Parameter β is the transmission rate, γ is the recovery rate, and μ is birth and death rates.

For any transmission rate β larger than $\gamma + \mu$, the SIR system in (1) has an endemic equilibrium at (S^*, I^*), where,

$$\begin{aligned} S^* &= \frac{1}{R_0}, \\ I^* &= \frac{\mu}{\gamma + \mu}(1 - \frac{1}{R_0}). \end{aligned}$$

Here $R_0 = \frac{\beta}{\gamma+\mu}$ is the basic reproductive ratio, commonly defined as the average number of secondary cases caused by an infectious individual in a totally susceptible population [57]. R_0 is one of the most important parameters in SIR-based models because the endemic equilibrium only exists when R_0 is larger than 1.

The Jacobian of the dynamics near the endemic fixed point is given by

$$J_{(S^*,I^*)} = \begin{bmatrix} -\mu R_0 & -\gamma - \mu \\ \mu(R_0 - 1) & 0 \end{bmatrix}. \tag{4.2}$$

Hence the eigenvalues of the system are

$$\lambda_{1,2} = -\mu R_0 \pm \sqrt{\mu^2 R_0^2 - 4\mu(\beta - \gamma - \mu)}. \tag{4.3}$$

Because $(\mu R_0)^2$ is much smaller than $4\mu(\beta - \gamma - \mu)$, the eigenvalues $\lambda_{1,2}$ consist of a real and an imaginary part. Because of the negative real part, the endemic equilibrium is stable. When the free system is perturbed away from the endemic equilibrium, it spirals inward, oscillating at the natural frequency of the system. The natural frequency is determined by the imaginary part of the eigenvalues,namely

$$\omega_0 = \sqrt{-\mu^2 R_0^2 + 4\mu(\beta - \gamma - \mu)}. \tag{4.4}$$

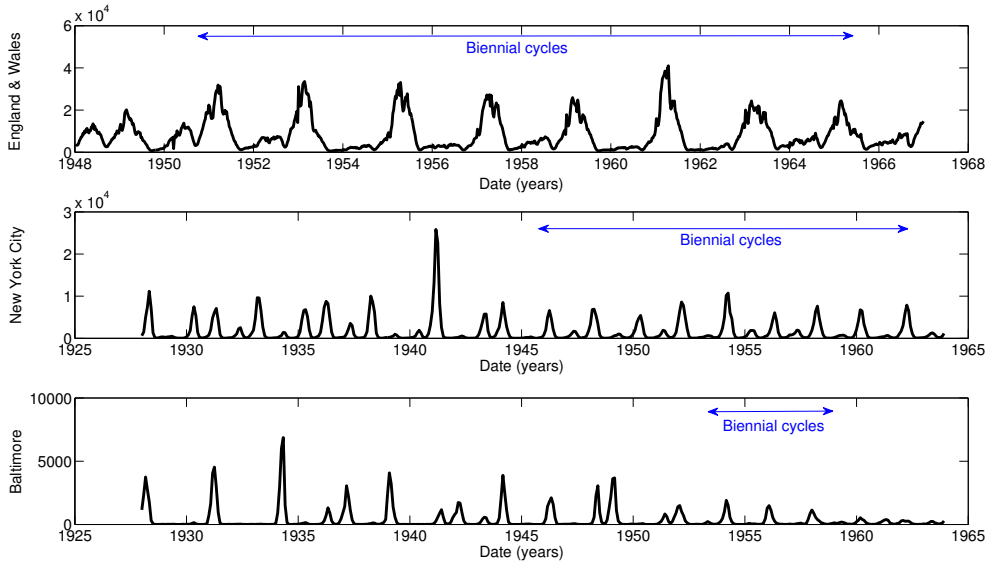

Figure 4.1: Time series of weekly case reports of measles in three regions: (a) England and Wales, (b) New York City, (c) Baltimore [3].

The free SIR model predicts damped oscillations when it is perturbed. However, observation of childhood disease case reports show regular cycles instead. Measles

case reports, in particular, show regular biennial cycles in different regions of the world as shown in Fig. 4.1. Such periodic dynamics can be caused by a transmission rate that varies annually, caused by environmental fluctuations and by human activity such as school terms. For simplicity, we assume that the transmission rate β varies sinusoidally, namely

$$\beta = \beta_0(1 + \epsilon cos(2\pi t)), \tag{4.5}$$

with β_0 and ϵ being constant, and time t measured in years. Although seasonality may have a period with a varying length, as in [95], we consider the case there the period is 1 year because in this analysis we focus on the effects of biological parameters and the effects of the amplitude of the seasonality (instead of the effects of the period of the seasonality).

Parametric resonance can occur when β varies at a frequency close to twice the natural frequency of the system. Because the frequency of the parameter variation is fixed (one oscillation per year), parametric resonance can occur when the natural period of the system is close to 2 years. In the case of measles, the value of β is shown in Table 4.1. The natural period is 2.1 years, which is close to 2 years. Therefore, large amplitude biennial cycles in measles can be caused by parametric resonance. The following sections show an MMS approach to analyze SIR model with periodic forcing to identify the criteria for the onset of parametric resonance.

4.2.2 Method of multiple scales

The method of multiple scales (MMS) is a technique to construct approximate solutions as corrections to the solution of the linearized system [87, 20]. This approach accounts for the nonlinearity and periodic forcing of the system. All the nonlinear

Disease (England Wales)	Measles
Transmission rate β (1/days)	17/13
Recovery rate γ (1/days)	1/13
Birth and Death rate μ (1/years)	1/70
Natural period (years)	2.1 years

Table 4.1: Epidemiological parameters of measles in England and Wales [4]. The natural period is calculated from the imaginary part of the eigenvalue at the fixed point.

parts of the system are assumed to be small compared to the linear part.

To apply MMS to our SIR model, several changes of variables are carried out. The first change of variable is applied so that the equilibrium of the transformed system is at the origin. Coordinates X_1 and X_2 are defined as $X_1 = S - S^*$ and $X_2 = I - I^*$. Thus, we obtain,

$$\begin{aligned}\frac{dX_1}{dt} =& \frac{-\mu\beta_0}{\gamma+\mu}X_1 - (\gamma+\mu)X_2 - \beta_0 X_1 X_2 - \epsilon\beta_0\cos(\Omega t)(X_1+S^*)(X_2+I^*),\\ \frac{dX_2}{dt} =& \frac{\mu(\beta_0-\gamma-\mu)}{\gamma+\mu}X_1 + \beta_0 X_1 X_2 + \epsilon\beta_0\cos(\Omega t)(X_1+S^*)(X_2+I^*),\end{aligned} \tag{4.6}$$

where Ω is the driving frequency of the system.

A second change of variables is applied to analyze the cases where the nonlinear term $\beta_0 X_1 X_2$ is small compared to the linear part of the system. Coordinates Y_1 and Y_2 are defined as $Y_1 = \epsilon X_1$ and $Y_2 = \epsilon X_2$. The governing equations become

$$\begin{aligned}\frac{dY_1}{dt} =& \frac{-\mu\beta_0}{\gamma+\mu}Y_1 - (\gamma+\mu)Y_2 - \beta_0\cos(\Omega t)S^*I^*\\ & -\epsilon\beta_0 Y_1 Y_2 - \epsilon\beta_0\cos(\Omega t)(\epsilon Y_1 Y_2 + I^*Y_1 + S^*Y_2),\\ \frac{dY_2}{dt} =& \frac{\mu(\beta_0-\gamma-\mu)}{\gamma+\mu}Y_1 + \beta_0\cos(\Omega t)S^*I^*\\ & +\epsilon\beta_0 Y_1 Y_2 + \epsilon\beta_0\cos(\Omega t)(\epsilon Y_1 Y_2 + I^*Y_1 + S^*Y_2),\end{aligned} \tag{4.7}$$

A third change of variables is applied so that the linear part of the system can be off-diagonalized. This step is necessary because MMS was originally developed in the

context of second order single-degree-of-freedom systems.

To that aim, we first define

$$\mathbf{A_x} = \begin{bmatrix} \frac{-\mu\beta_0}{2(\gamma+\mu)} & -(\mu+\gamma) \\ \frac{-\mu(\beta_0-\gamma-\mu)}{\gamma+\mu} & \frac{-\mu\beta_0}{2(\gamma+\mu)} \end{bmatrix},$$

so that $\begin{pmatrix} \dot{Y}_1 \\ \dot{Y}_2 \end{pmatrix} = \mathbf{A_x} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ is the linear part of Eq. (7) without damping and external forcing. Matrix $\mathbf{A_x}$ determines the natural frequency and mode shape of the system. The eigenvectors of $\mathbf{A_x}$ are $\mathbf{V_x} = \begin{bmatrix} -\frac{\mu\beta_0}{2(\beta+\mu)} & \omega \\ \frac{\mu(\beta-\gamma-\mu)}{\mu+\gamma} & 0 \end{bmatrix}$.

Next, we define

$$\mathbf{A_u} = \begin{bmatrix} 0 & -\omega_0^2 \\ 1 & 0 \end{bmatrix},$$

which has eigenvectors as $\mathbf{V_u} = \begin{bmatrix} 0 & \omega \\ 1 & 0 \end{bmatrix}$.

The coordinate transformation is defined as: $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \mathbf{M} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$, where $\mathbf{M} =$

$\mathbf{V_u}(\mathbf{V_x})^{-1}$. Equation (7) thus becomes:

$$\begin{aligned}\frac{du_1}{dt} =& -\omega_0^2 u_2 + (\mathbf{M}_{1,2} - \mathbf{M}_{1,1})\beta_0 \cos(\Omega t) S^* I^* \\ &+ \epsilon[\beta_0 Z_1 + (\mathbf{L}_{1,1}u_1 + \mathbf{L}_{1,2}u_2)\cos(\Omega t) + \mathbf{C}_{1,1}u_1 + \mathbf{C}_{1,2}u_2] \\ &+ \epsilon^2 \beta_0 \cos(\Omega t) Z_1, \end{aligned} \tag{4.8}$$

$$\begin{aligned}\frac{du_2}{dt} =& u_1 + (\mathbf{M}_{2,2} - \mathbf{M}_{2,1})\beta_0 \cos(\Omega t) S^* I^* \\ &+ \epsilon[\beta_0 Z_2 + (\mathbf{L}_{2,1}u_1 + \mathbf{L}_{2,2}u_2)\cos(\Omega t) + \mathbf{C}_{2,1}u_1 + \mathbf{C}_{2,2}u_2] \\ &+ \epsilon^2 \beta_0 \cos(\Omega t) Z_2, \end{aligned} \tag{4.9}$$

where ω_0 is the natural frequency of the system, $\mathbf{L} = \mathbf{M} \begin{bmatrix} -I^* & -S^* \\ I^* & S^* \end{bmatrix} \mathbf{M}^{-1}$, and $\mathbf{C} = \mathbf{M} \begin{bmatrix} -c & 0 \\ 0 & -c \end{bmatrix} \mathbf{M}^{-1}$. Here $\mathbf{L}$ is related to the linear parametric excitation terms, and $\mathbf{C}$ is related to damping terms (separated from the linear part of the system). Because this is a weakly damped system, a scalar c is defined as $c = \frac{\mu\beta_0}{2\epsilon(\gamma+\mu)}$ so that damping terms only appear in the slow time scales. Z_1 and Z_2 are nonlinear terms defined as

$$Z_1 = M_1 u_1^2 + M_2 u_1 u_2 + M_3 u_2^2,$$
$$Z_2 = N_1 u_1^2 + N_2 u_1 u_2 + N_3 u_2^2,$$

where parameters M_1, M_2, M_3, N_1, N_2 and N_3 are functions of entries of matrix $\mathbf{M}$. Details of the parameters can be found in the Supplementary Information.

The off-diagonalized system can be transformed into a second-order single-degree-of-freedom system by taking a time derivative of Eq. (8) and substituting Eq. (9)

into that to obtain

$$\begin{aligned}\frac{d^2u_1}{dt^2} =& -\omega^2 u_1 + (\mathbf{M}_{1,2} - \mathbf{M}_{1,1})\beta_0 \cos(\Omega t) S^* I^* \\ & - \omega_0^2(\mathbf{M}_{2,2}) - \mathbf{M}_{2,1})\beta_0 \cos(\Omega t) S^* I^* \\ & + \epsilon[\beta_0(2M_1 u_1 \dot{u}_1 + M_2 \dot{u}_1 u_2 + M_2 u_1 \dot{u}_2 + 2M_3 u_2 \dot{u}_2) \\ & + (\mathbf{L}_{1,1}\dot{u}_1 + \mathbf{L}_{1,2}\dot{u}_2)\cos(\Omega t) - \Omega(\mathbf{L}_{1,1}u_1 + \mathbf{L}_{1,2}u_2)\sin(\Omega t) \\ & + \mathbf{C}_{1,1}\dot{u}_1 + \mathbf{C}_{1,2}\dot{u}_2] + \omega_0^2\epsilon[\beta_0(N_1 u_1^2 + N_2 u_1 u_2 + M3u_2^2) \\ & + (\mathbf{L}_{2,1}u_1 + \mathbf{L}_{2,2}u_2)\cos(\Omega t) + \mathbf{C}_{1,1}u_1 + \mathbf{C}_{1,2}u_2] \\ & + \epsilon^2[\beta_0 \cos(\Omega t)(2M_1 u_1 \dot{u}_1 + M_2 \dot{u}_1 u_2 + M_2 u_1 \dot{u}_2 + 2M_3 u_2 \dot{u}_2) \\ & - \Omega\beta_0 \sin(\Omega t)(M_1 u_1^2 + M_2 u_1 u_2 + M_3 u_2^2)] \\ & - \omega_0^2\epsilon^2\beta_0(N_1 u_1^2 + N_2 u_1 u_2 + N_3 u_2^2)\cos(\Omega t).\end{aligned} \tag{4.10}$$

Because parametric resonance occurs when the driving frequency Ω is close to twice the natural frequency of the system, we introduce the detuning factor k to quantify how close Ω is to twice the natural frequency

$$\Omega = 2\omega_0(1 + k\epsilon).$$

The solution of Eq. (10) shall be represented at three different time scales, which are distinguished in their order of magnitude by the small parameter ϵ. The solution of the linear part of Eq. (10) is represented by the fast time scale $T_0 = \frac{\Omega}{2}t$ [20, 79, 96]. The nonlinear terms cause a deviation from the solution of the linearized system. This deviation is expressed as a variation in the amplitude and the phase of the system dynamics based on the slow time scales $T_1 = \epsilon t$ and $T_2 = \epsilon^2 t$.

The MMS is applied following several steps. First, the solution of Eq. (10) is

expanded into a power series of ϵ as

$$u_1 = u_{10}(T_0, T_1, T_2) + \epsilon u_{11}(T_0, T_1, T_2) + \epsilon^2 u_{12}(T_0, T_1, T_2) + H.O.T., \tag{4.11}$$

where $H.O.T.$ indicate higher order terms (of order ϵ^3 and higher).

Differential operators are introduced for the different time scales as

$$D_0 = \frac{\partial}{\partial T_0}, D_1 = \frac{\partial}{\partial T_1}, D_2 = \frac{\partial}{\partial T_2}.$$

Thus, we obtain three equations, one each power of ϵ as

order ϵ^0

$$\begin{aligned} \omega_0^2 D_0^2 u_{10} = &- \omega_2 u_{10} \\ &- 2\omega_0(\mathbf{M}_{1,2} - \mathbf{M}_{1,1})\beta_0 \sin(\Omega t) S^* I^* \\ &- \omega_0^2(\mathbf{M}_{2,2} - \mathbf{M}_{2,1})\beta_0 \cos(\Omega t) S^* I^*, \end{aligned} \tag{4.12}$$

order ϵ

$$\begin{aligned} \omega_0^2 D_0^2 u_{11} = &- \omega_0^2 u_{11} - 2\omega_0 D_0 D_1 u_{10} - 2\omega_0^2 k D_0^2 u_{10} \\ &+ H(u_{10}, D_0 u_{10}, u_{20}, D_0 u_{20}, \beta, \gamma, \mu, \cos(\Omega t), \sin(\Omega t)), \end{aligned} \tag{4.13}$$

order ϵ^2

$$\begin{aligned} \omega_0^2 D_0^2 u_{11} = &- \omega_0^2 u_{12} - k^2 \omega_0^2 D_0^2 u_{10} \\ &- 2\omega_0 D_0 D_1 u_{11} - 2\omega_0 k D_0 D_1 u_{10} \\ &- D_1^2 u_{10} - 2\omega_0 D_0 D_2 u_{10} \\ &+ G(u_{10}, D_0 u_{10}, u_{20}, D_0 u_{20}, u_{11}, D_0 u_{11}, u_{21}, D_0 u_{21}, \beta, \gamma, \mu, \cos(\Omega t), \sin(\Omega t)). \end{aligned} \tag{4.14}$$

Details of H and G can be found in the Supplementary Information.

The general solution of the order ϵ^0 equation (Eq. (12)) is of the form

$$\begin{aligned} u_{10} =& A(T_1, T_2)\cos T_0 + B(T_1, T_2)\sin T_0 \\ &+ P_1\cos(2T_0) + P_2\sin(2T_0), \\ u_{20} =& \frac{1}{\omega_0}[A(T_1, T_2)\sin T_0 - B(T_1, T_2)\cos T_0 \\ &+ 2P_1\sin(2T_0) - \frac{1}{2}P_2\cos(2T_0)], \end{aligned} \tag{4.15}$$

where P_1 and P_2 are constants, while A and B are functions of T_1 and T_2. The values of P_1 and P_2 are

$$\begin{aligned} P_1 &= \frac{1}{3}(\mathbf{M}_{2,2} - \mathbf{M}_{2,1})\beta_0 S^* I^*, \\ P_2 &= \frac{2}{3\omega_0}(\mathbf{M}_{1,2} - \mathbf{M}_{1,1})\beta_0 S^* I^*. \end{aligned}$$

Substituting Eq. (15) into Eq. (13) yields

$$\begin{aligned} \omega_0^2 D_0^2 u_{11} =& -\omega_0^2 u_{11} + C\cos(T_0) + D\sin(T_0) + Q_1\cos(2T_0) + Q_2\sin(2T_0) \\ &+ R_1\cos(3T_0) + R_2\sin(3T_0) + S_1\cos(4T_0) + S_2\sin(4T_0) + Z. \end{aligned} \tag{4.16}$$

All coefficients in Eq. (16) are functions of A, B, D_1A and D_1B, where $D_1A = \frac{\partial A}{\partial T_1}$ and $D_1B = \frac{\partial B}{\partial T_1}$. For the system to be stable, all the secular terms on the right side of Eq. (16) have to be 0. Therefore, $C = D = 0$.

With C and D being 0, we can solve Eq. (16) to obtain the solution of u_{11} as

$$\begin{aligned} u_{11} = &- \frac{1}{3\omega_0^2}(Q_1\cos(2T_0) + Q_2\sin(2T_0)) \\ &- \frac{1}{8\omega_0^2}(R_1\cos(3T_0) + R_2\sin(3T_0)) \\ &- \frac{1}{15\omega_0^2}(S_1\cos(4T_0) + S_2\sin(4T_0)) \\ &+ \frac{1}{\omega_0^2}Z. \end{aligned} \tag{4.17}$$

Next, u_{21} can be obtained using Eq. (8) and Eq. (9). Note that the $\cos(T_0)$ and $\sin(T_0)$ parts of u_{11} are omitted because their coefficients are not functions of A and B, which means they will not affect the growth of A and B.

Solutions of u_{10}, u_{20}, u_{11} and u_{21} are then substituted into Eq. (14), which yields

$$\omega_0^2 D_0^2 u_{12} = -\omega_0^2 u_{12} + E\cos(T_0) + F\sin(T_0) + N.S.T. \tag{4.18}$$

where $N.S.T.$ stands for terms that do not include secular terms. Again, eliminating secular terms, we obtain $E = F = 0$. Here, E and F are functions of A, B, D_1A, D_1B, D_2A and D_2B, where $D_2A = \frac{\partial A}{\partial T_2}$ and $D_2B = \frac{\partial B}{\partial T_2}$.

Enforcing that $C = D = E = F = 0$, we can obtain the relationship between D_1A, D_1B, D_2A, D_2B and A, B. Thus, the governing equations for the variation of A and B over time can be written as

$$\frac{dA(T_1,T_2)}{dt} = \frac{\partial A}{\partial T_1}\frac{dT_1}{dt} + \frac{\partial A}{\partial T_2}\frac{dT_2}{dt} = \epsilon D_1A + \epsilon^2 D_2A, \tag{4.19}$$

$$\frac{dB(T_1,T_2)}{dt} = \frac{\partial B}{\partial T_1}\frac{dT_1}{dt} + \frac{\partial B}{\partial T_2}\frac{dT_2}{dt} = \epsilon D_1B + \epsilon^2 D_2B. \tag{4.20}$$

Substituting Eq. (15) and Eq. (17) back into Eq. (11), we can obtain an approximation solution where A and B are governed by Eq. (19) and Eq. (20). Periodic steady state solutions can be obtained by solving $\frac{dA}{dt} = 0$ and $\frac{dB}{dt} = 0$, which gives a

relationship between the amplitudes A, B and, the detuning factor k and the seasonality ϵ.

Once the approximate solutions for u_1 and u_2 are obtained, S and I can be obtained by reversing the procedure of change of variables in Eqs. (6), (7) and (8).

4.2.3 Periodic solutions

The solution of the frequency response function (Eqs. (19) and (20)) yields either one, two or three coexisting solutions. The stability of these solutions can be determined using the eigenvalues of the Jacobian of Equations (19) and (20) at equilibrium. The trivial linear solution with oscillation frequency Ω is obtained when $A = B = 0$. When the trivial solution is unstable, A and B grow, and a non-trivial stable 2:1 subharmonic motion appears. In this case the system oscillated at frequency $\frac{\Omega}{2}$ that is roughly the natural frequency of the system ω_0. In childhood infectious disease models, the driving frequency Ω is 1/year. Therefore, the trivial solution corresponds to one year cycles, while the non-trivial solution corresponds to biennial cycles. The stable trivial solution and the non-trivial solution can also coexist, and in that case they are separated by another unstable solution.

Based on different characteristics of the solutions, the parameter space $\epsilon - \Omega/\omega_0$ can be divided into four regions as shown in Fig. 4.2.

When ϵ is below a certain threshold ϵ_0 and the system falls into region IV, there will be no resonance. In this case, the trivial solution is the only solution, and the system response oscillates at frequency Ω. Time series and phase plane plots from this region are shown in Fig. 5.9 (e) and (j).

When ϵ is above the threshold ϵ_0, whether the trivial solution ($A = B = 0$) is stable depends on both the amplitude of seasonality ϵ and the detuning factor k. The rest of the $\epsilon - \Omega/\omega_0$ plane can be further divided into three regions (regions I, II and III) as shown in Fig. 4.2.

When the detuning factor is positive and above a certain threshold, the system will fall into region III where only the trivial solution $A = B = 0$ exists. The system again oscillates at frequency Ω as shown in Fig. 5.9 (d) and (i). In this case, only small annual cycles can be observed.

For systems belonging to region II, two solutions exist. The trivial solution $A = B = 0$ is unstable. Thus, the system will always exhibit resonance, oscillating at frequency $\frac{\Omega}{2}$ with nonzero values for A and B, which corresponds to biennial cycles as shown in Fig. 5.9 (c) and (h).

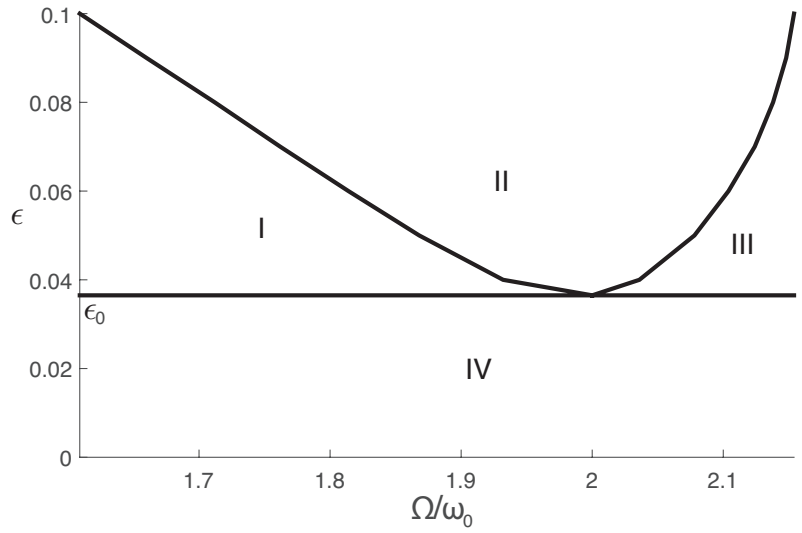

Figure 4.2: Transition curves of a periodically forced SIR model. Parameters of the SIR model are shown in Table 4.1. The parameter plane is separated into four regions based on different behaviors of the periodic solutions. In region I, three solutions coexist, including two stable solutions and one unstable solution. In region II, an unstable trivial solution and a stable non-trivial solution coexist. In region III and region IV, only a trivial solution exists.

As the detuning factor becomes smaller, system goes into region I where there are three coexisting solutions. The trivial solution is stable in this region, coexisting with one stable and one unstable resonance solutions. Time series and phase plane plots of the stable and unstable solutions are shown in Fig. 5.9 (a), (f) and (b), (g). In region I, annual cycles and biennial cycles are separated by an unstable solution. Transitions

from one solution to another can be triggered by perturbations in the state variables without changing the parameter. This suggests that transitions from annual cycles to biennial cycles might be caused by suddenly introducing a large enough amount of infectious individual to the region.

Biennial cycles can be observed in an annual forced SIR model in different parameter regions. Transitions from annual cycles to biennial cycles can thus be triggered in different ways. Using the transition curve obtained from analysis, we can explore the behavior of a periodically forced SIR model and the transitions between different behaviors in more detail.

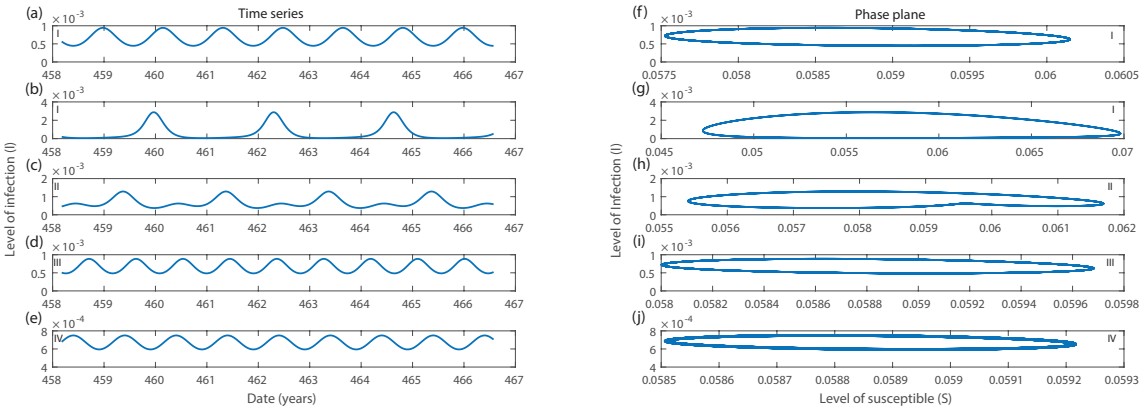

Figure 4.3: Time series (left column) and phase plane (right column) plots for an SIR model when parameters ϵ and Ω/ω_0 fall into different regions separated by transition curves. The corresponding region of parameters are shown on the plot. The same set of parameters is used in the first and second rows, showing two coexisting stable solutions. Parameter values of ϵ and Ω/ω_0 are as follows: first and second rows, $\epsilon = 0.05$, $\Omega/\omega_0 = 1.8$; third row, $\epsilon = 0.06$, $\Omega/\omega_0 = 2.1$; fourth row, $\epsilon = 0.02$, $\Omega/\omega_0 = 2.1$; final row, $\epsilon = 0.06$, $\Omega/\omega_0 = 2.3$.

4.3 Parametric analysis

The stability of trivial linear solution and the existence of non-trivial solution depends on the amplitude of seasonality ϵ and on the natural frequency of the system ω_0. Because natural frequency is affected by the transmission rate β, the recovery

rate γ and the death-birth rate μ, the transition from trivial linear solution (one year cycle) to non-trivial solution (biennial cycle) can be triggered by changes in any of these parameters.

In previous research, it was observed that transition from annual cycles to biennial cycles (referred to as period doubling bifurcation) could be triggered by increasing the amplitude of seasonality ϵ [90]. It was pointed out also that the transition from large amplitude biennial cycles to small amplitude oscillations can be caused by decreasing the transmission rate β, or increasing the birth rate [3]. In this section, we show that all transitions observed in previous studies are caused by crossing transition curves in the $\epsilon - \Omega/\omega_0$ plane due to changes in parameters. Effects of changing the amplitude of seasonality ϵ the transmission rate β and the birth rate are revisited both analytically and numerically. Results from both analyses are compared to show the accuracy of the approximate solution obtained from the analysis. The influence of other parameters, such as the recovery rate γ is also studied.

4.3.1 Amplitude of seasonality ϵ

Increasing the amplitude of seasonality ϵ can bring the system from region IV to region II when the driving frequency Ω is close to twice the natural frequency ω_0 as shown in Fig. 4.2. When system is in region IV, the trivial linear solution is stable. Thus, the system shows annual cycles with small amplitude. When ϵ is larger, the system transitions from annual cycles to biennial cycles. Fig. 4.4 shows the change in the maximum level of infection in this process. An abrupt increase in the amplitude of infectious population accompanies the period-doubling phenomenon. Results from both analytical and numerical analyses are compared to show the accuracy of the approximation solution obtained from analytical analysis.

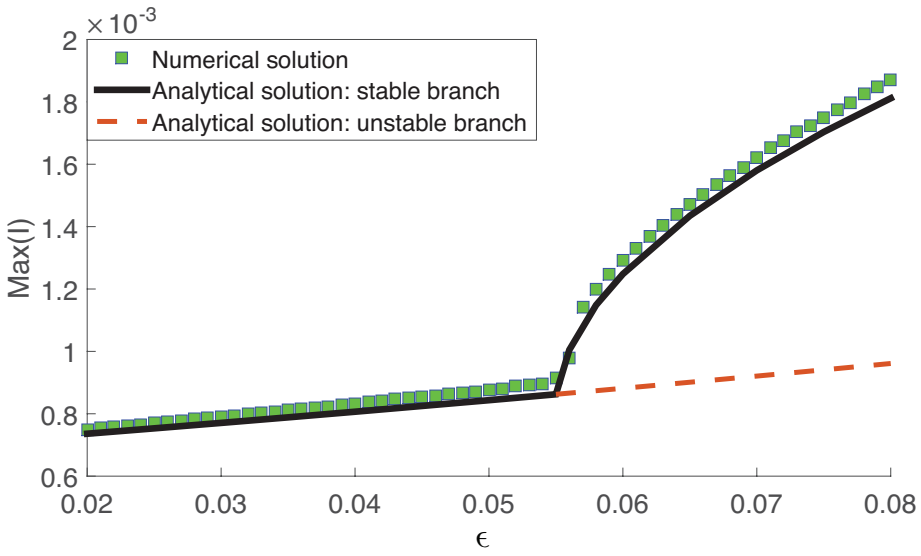

Figure 4.4: Maximum level of infection plotted against the amplitude of seasonal forcing ϵ as the system transitions from a non-resonance region (region IV) to a resonance region (region II) by increasing ϵ. The solid black line gives the stable branch of the analytical solution; the red dashed line gives the unstable branch of the analytical solution; the green squares show the results from numerical simulation. Parameters are same as in Table 4.1.

It is worth pointing out that the threshold ϵ_0 for the non-trivial solution to exist can be lower when the driving frequency Ω is closer to twice the natural frequency ω_0 of the system. The threshold is lowest when Ω is exactly $2\omega_0$. This threshold can be much lower than originally believed, and it shows that large biennial outbreaks are more likely to happen when the natural frequency is closer to 1/(2 years).

4.3.2 Transmission rate β

When ϵ is fixed, decreasing the transmission rate β causes a simultaneous shift of transition curve and of the natural frequency. Fig. 4.5 shows that decreasing β with fixed γ lowers and broadens the transition curve. Thus, it becomes easier to trigger parametric resonances. ω_0 decreases as β decreases because $\omega_0 = \sqrt{-\mu^2 R_0^2 + 4\mu(\beta - \gamma - \mu)}$ and $-\mu^2 R_0^2$ is much smaller than $\mu(\beta - \gamma - \mu)$. Because Ω is fixed as 1/ year, the decrease in β actually takes the system from left to right

in the parameter space in Fig. 5. Therefore, the transition may not take place as the transmission rate decreases, depending on the relative relationship between the position of the system in the parameter space and the transition curve.

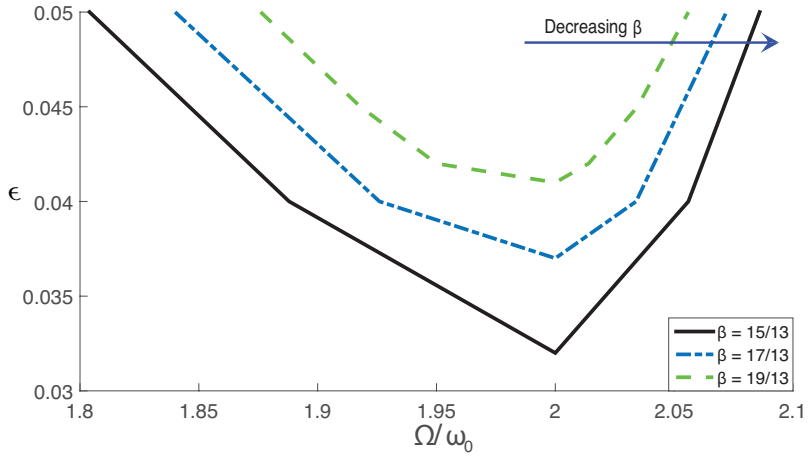

Figure 4.5: Change of the transition curve for parametric resonance as the transmission rate β is increased. The solid black line, blue and green dashed lines give the transition curve obtained using transmission rates of 15/13, 17/13 and 19/13, respectively. Parameters other than the transmission rate are the same as in Table 4.1.

Consider a scenario where the system starts in region I, where there are three coexisting solutions. As the transmission rate decreases, the system passes region II and finally reaches region III. In this process the solution transitions from large amplitude biennial cycles to annual cycles, as shown in Fig. 4.6. Analytical solutions are compared with numerical solutions (obtained using a shooting method [97, 98, 99]). Figure 4.6 shows that the analytical solution predicts the stable branch of the solution well. However, the unstable branch in region I has poor agreement with the numerical solution. In fact, region I can be further divided into two regions according to numerical simulation, because as β becomes larger than certain value, numerical simulation can only find one unique solution instead of three coexisting solutions. The discrepancy between analytical solutions and numerical simulations can be explained by the failure of the MMS approximation as the amplitude or the detuning factor

becomes large.

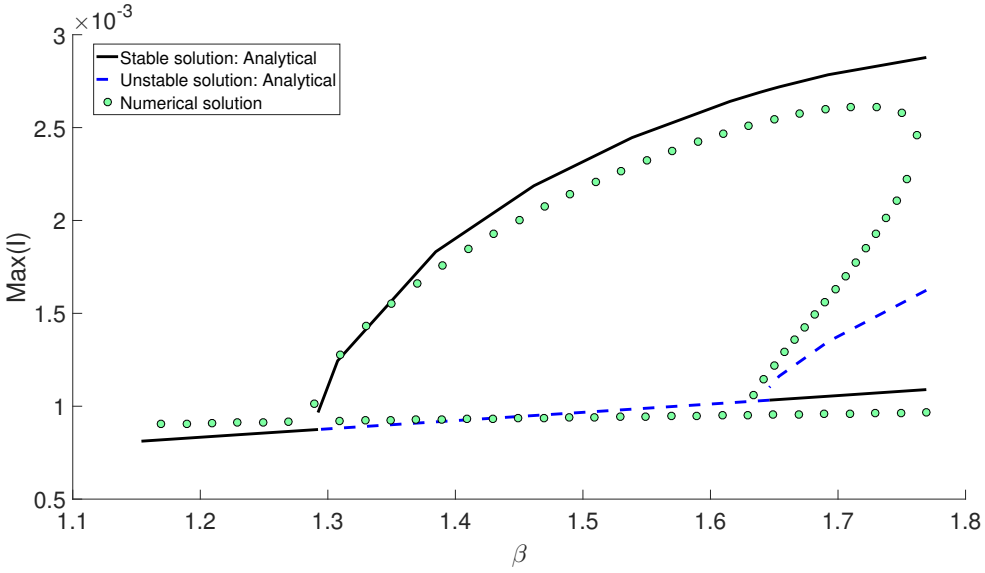

Figure 4.6: Maximum level of infection plotted against the transmission rate β as the system transitions from region I to region II and finally to region III by decreasing β. The black solid line shows the stable branch of the analytical solution; the dashed blue line shows the unstable branch of the analytic solution; the green circles show the results from the numerical simulations. Parameters other than the transmission rate are the same as in Table 4.1.

4.3.3 Birth rate

Changes in birth rate can also move the system from one section to another, causing changes on the periodicity of epidemics [3, 100, 101, 102]. It has been noted in [**?**] that for measles dynamics typically annual cycles take place when birth rate is high.

Changes in the birth rate are dynamically equivalent to changes in the transmission rate for SIR-based models as pointed out in [1]. Therefore, consider a scenario where the system starts in region II (similar to Section 3.2). The only stable solution in region II is the biennial solution. As the birth rate increases, the system shifts from region II and reaches region I. In this process the solution transitions from large amplitude biennial cycles to annual cycles, as shown in Fig. 4.7.

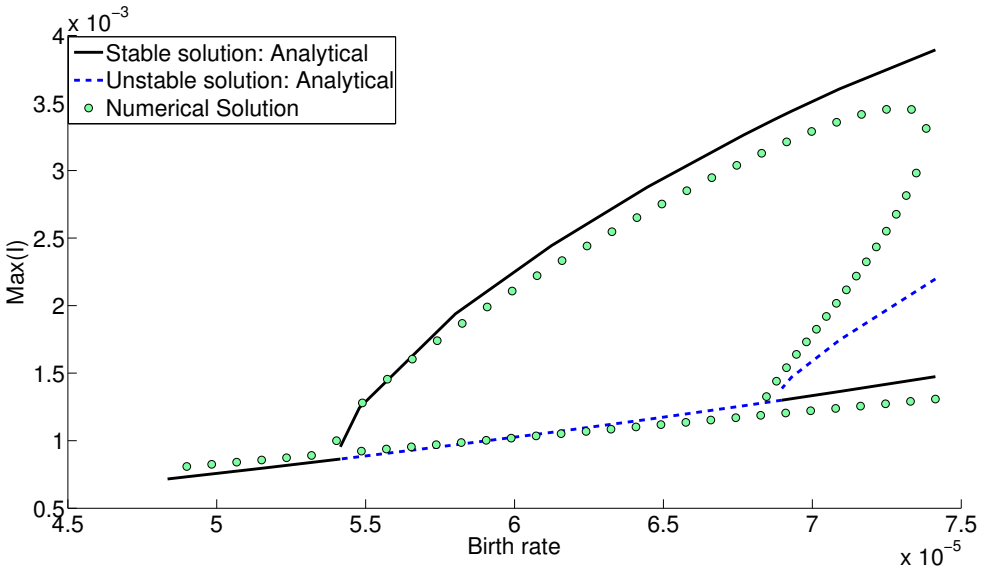

Figure 4.7: Maximum level of infection plotted against the birth rate as the system transitions from region I to region II and finally to region III by decreasing the birth rate. The black solid line shows the stable branch of the analytical solution; the dashed blue line shows the unstable branch of the analytic solution; the green circles show the results from the numerical simulations. Parameters other than the birth rate are the same as in Table 4.1.

Despite the limited accuracy of the analytical MMS method in predicting the unstable branch, this method still provides useful insight into biennial cycles. For the case of measles, the change from large biennial cycles to irregular annual cycles in the last century can be explained by the transition from one region to another induced by changes in the birth rate and vaccination rates.

4.3.4 Recovery rate γ

The recovery rate γ determines the average infectious period, which can be estimated from epidemiological data. The recovery rate typically does not change significantly over time, and it has a similar effect on the natural frequency of the system as the transmission rate. Thus, we do not focus on the effects of changing recovery rate over time on the transition from annual cycles to biennial cycles. However, different diseases have different transmission rate and recovery rate, which results in a different

thresholds ϵ_0 for the transition. By analyzing cases with the same transmission rate β yet different recovery rates, we understand what are the effects of recovery rate on the threshold.

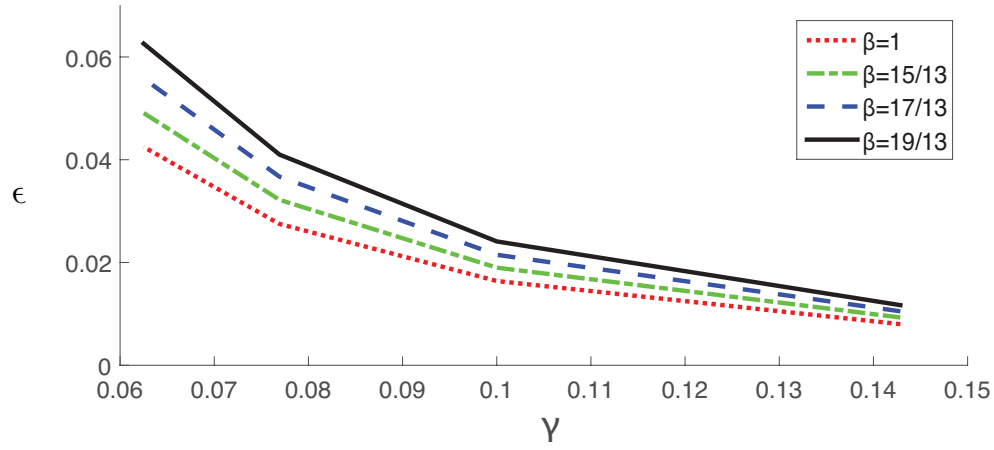

Figure 4.8: Threshold for amplitude of seasonal forcing ϵ plotted against recovery rate γ. The threshold ϵ_0 is defined as the maximum of region IV as shown in Fig. 4.2. All four lines from top to bottom are obtained using transmission rate values of 19/13, 17/13, 15/13, and 1, respectively.

Fig. 4.8 shows that with the same transmission rate, increasing recovery rate lowers the threshold for parametric resonance. When recovery rate γ becomes larger than 0.14, the threshold can be as low as 0.01, which is much lower than originally believed [57, 91, 94]. The situation is opposite with the transmission rate. For constant recovery rate and increasing the transmission rate, there will be an increase in the threshold. Therefore, it can be concluded that systems with larger recovery rates and small transmission rates have smaller threshold for parametric resonance even if they might have similar natural frequencies.

4.4 Effects of stochasticity

We have observed that small changes in the periodic forcing, either amplitude or frequency, can cause a qualitative change in the behavior of the system. Stochasticity

can have such disproportionate effects also. To understand the role of stochasticity in triggering parametric resonance, when periodic forcing and weak noise are present at the same time, the SIR model is transformed into a stochastic SIR model as in [103]

$$\begin{aligned} dS(t) =&(\mu - \beta S(t)I(t) - \mu S(t))dt - \sigma S(t)I(t)dW(t), \\ dI(t) =&(\beta S(t)I(t) - (\mu + \gamma)I(t))dt + \sigma S(t)I(t)dW(t), \\ dR(t) =&(\lambda I(t) - \mu R(t))dt, \end{aligned} \tag{4.21}$$

where σ is a positive constant, and W is a standard Wiener process over the time interval [0 T]. The model is solved numerically using the Euler-Maruyama method, and the results for different noise level are shown in Fig. 4.9. This model is chosen because it enables us to study how different levels of stochasticity can cause different types of behavior of the system. When $\sigma = 0$, no noise is added to the system, so the system remains at the trivial solution, oscillating at frequency Ω with a relatively low amplitude. However, as the level of noise increases, the system starts switching between annual cycles and biennial cycles. After noise reaches a certain level, such as $\sigma = 0.2$, the system stabilizes on biennial cycles with a much larger amplitude.

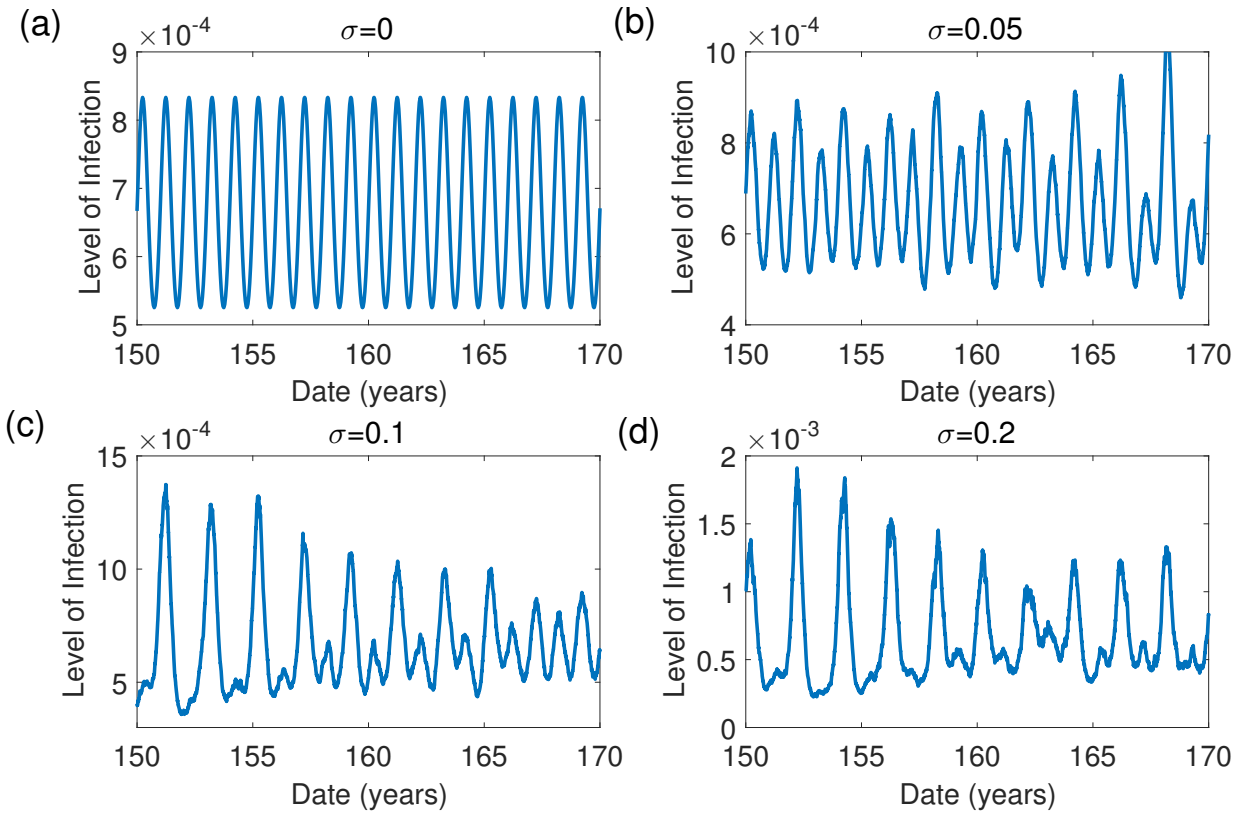

Figure 4.9: Stochastic simulations of an SIR model with periodic forcing and white noise. Levels of noise are 0, 0.05, 0.1 and 0.2 respectively for (a), (b), (c) and (d). All simulations start from the same initial condition. After a certain amount of simulation time, data are collected for 20 years to show different patterns under different levels of noise.

When the amplitude of the periodic forcing is close to the threshold, even weak noise can force the system out of the basin of attraction of the annual cycles. Therefore, the presence of stochasticity can significantly increase the amplitude of the system response by triggering the parametric resonance. In addition, the threshold of amplitude of periodic forcing can be further reduced when the system is accompanied by noise. This is important for understanding the effects of parametric resonance in epidemiology.

4.5 Conclusions

An analysis based on the method of multiple scales (MMS) is carried out to study a classic SIR model with periodic forcing. Analytical approximate solutions were obtained to show that parametric resonance can occur through the interaction between nonlinearity and seasonal forcing. The results reveal that large amplitude epidemics can take place if the system is moved into a resonance regime in the parameter plane by the change of any one of its parameters. This analysis can be extended to more complex epidemiological or biological models where more variables or other kinds of nonlinearity are present.

An important contribution of this analysis is that it shows that small order parametric excitations can synchronize with the system response, and trigger an order 1 parametric resonance. When parametric resonance is triggered, the excitation drives the growth of the biennial cycles until it is constrained by the nonlinearity.

This analysis also reveals that not only the amplitude of the parametric excitation, but also the relationship between natural frequency and excitation frequency matters. Therefore, all parameters can have an impact on triggering parametric resonance by changing the natural frequency of the system and changing the threshold of seasonality. Thus, this analysis unifies past research about the effects of different parameters on the biennial cycles of measles [3, 89, 90, 91]. It is also important to notice that large recovery rates and smaller transmission rates can lead to lower thresholds for parametric resonance.

Stochasticity also plays an important part in the dynamics. In particular, we showed numerically that stochasticity can trigger parametric resonance even when the system is in the non-resonance regime of the parameter space. This is because the stochasticity also contains components with the same frequency as the parametric excitation. When the intensity of stochasticity reaches a certain level, it can in fact push the system into the resonance regime, and induce the large amplitude biennial

cycles.

The dynamics of childhood infectious disease is affected by the interplay between nonlinearity, periodic forcing and stochasticity. Through this analysis, the interaction between nonlinearity and periodic forcing as the source of instability is can be understood more clearly. More importantly, this work introduced perturbation methods such as the MMS to epidemiological models with periodic excitation, and shows the considerable benefits of this approach. For example, it is much easier and more informative to study the effects of different parameters on the behavior of the system analytically rather than numerically. This type of methods can be applied to more sophisticated, higher dimensional epidemiological models, such as SIR model with multiple pathogens or pathogen strains [104], or spatial heterogeneity [105, 106]. In those cases, an analytical expression of the transition curve and the resonance peak as a function of the parameters can be quite useful [83].

CHAPTER V

Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems

5.1 Introduction

A growing amount of literature has reported the observation that some ecosystems undergo a critical transition from one stable state to another [9, 30, 33]. This type of transition is often the result of some gradually changing conditions, such as climate, ground water reduction, harvesting of certain species, and so on [34, 35]. Because of the gradual change in conditions, small variations in the mean state variables of the system are observed, which makes it hard to forecast critical transitions. Consequently, several indicators based on fluctuations around equilibrium values have been proposed as early warning signals for critical transitions.

Close to the threshold of such critical transitions, some systems lose resilience in the sense that it takes much longer for the system to recover from perturbations. This phenomenon is referred to as "critical slowing down" [63]. A number of indicators based on this phenomenon have been proposed to forecast critical transitions [5, 21, 17, 36, 37, 38, 39, 8]. These methods are based on the presumption that as a system approaches the threshold, one would expect to observe the increase of certain

statistics such as variance, skewness and autocorrelation due to the critical slowing down phenomenon. It has been shown in models that this is because the real part of the dominant eigenvalue that characterizes the recovery rate around the equilibrium state becomes closer to zero as the system approaches the critical transition [19, 40, 41].

Apart from these temporal early warning signals, recent studies suggest that spatial patterns can also provide useful information [42, 22, 32, 23]. In particular, Dakos et al. [22] point out that an increase in spatial correlation can serve as an early warning signal for systems with alternative steady states. This is because, as the system approaches the critical transition, the system becomes slow in recovering from perturbations, which might lead to stronger fluctuations of state variables around the equilibrium state under random environmental perturbations [43]. In such cases, the fluctuations of state variables around the spatial mean will also increase. Also, diffusive effects (exchanges between neighboring cells) become more dominant close to the critical transition, which means that spatial correlations, especially correlations between neighboring units will increase [31]. Therefore, spatial early warning signals, such as the spatial variance, spatial skewness and spatial correlation have been introduced to serve as indicators for critical transitions of spatially extended systems. These spatial early warning signals, however, have two major drawbacks. First, they are largely dependent on the deviation of state variables from their spatial mean. Thus, these indicators are affected by the change of the local equilibrium of each state variable as the system approaches a critical transition, especially for heterogeneous systems. Additional comments on this issue are included in the Discussion section. Second, these methods only look at one snapshot at one time, thus limiting the information they can gather from the system. It is hard, for instance, to find the (dynamical) pattern that is associated with the critical transition using only a single snapshot.

This paper proposes a set of indicators based on the fluctuations of several (local) state variables around their equilibrium values under stochastic excitations (random environmental perturbations). Such fluctuations can be characterized by the covariance matrix of these local variables. The proposed indicators are based on the observation that for systems with different characteristics at different locations (spatially heterogeneous systems), the largest eigenvalue of the covariance matrix generally increases much faster than other eigenvalues as the system approaches the critical transition. This is because the dynamics along the direction of the eigenvectors corresponding to the dominant eigenvalue become slower in that process. Therefore, the variance of the dynamics along that direction increases as is the case for temporal early warning signals used for low dimensional systems (not spatially distributed). Thus, we propose the largest eigenvalue of the covariance matrix and the percentage it accounts for of the total variation as spatial early warning signals. One important assumption has to be satisfied for the proposed early warning signals to work, that is the critical transition is the manifestation of a co-dimension one bifurcation. This assumptions means that the critical transition can be caused by the variation of a single parameter.

Eigenvalues of the covariance matrix has been proposed in the past as an early warning signal for high dimensional systems [47, 107]. It was established that the largest eigenvalue of the covariance matrix will increase close to the critical transition. In our paper, we take a step further to show that not only does the largest eigenvalue of the covariance matrix increase close to the critical transition, it also becomes dominant compared to other eigenvalues of the covariance matrix. The contributions of this paper are three fold: (1) it examines a linear high dimensional stochastic system modeled by a Fokker-Planck equation and shows the relationship between the eigenvalues of the covariance matrix and the eigenvalues of the system; (2) it shows how the eigenvalues of the covariance matrix and the percentage it accounts for of

the total variation can be estimated and used as early warning signals for spatially correlated ecological systems when an accurate model of the system is not available; and (3) it compares the proposed early warning signals with past spatial early warning signals and discusses advantages and drawbacks. It is important to note that early warning signals based on the critical slowing down phenomenon focus on critical transitions caused by the loss of linear stability. This also applies to the proposed early warning signals. Researchers have proposed other early warning signals (based on basin size and so on) to anticipate critical transitions not related to the loss of linear stability [108, 109].

5.2 Results

5.2.1 Main concept

Many high dimensional natural and physical systems [110, 111, 112] are constantly affected by random environmental perturbations, and can be modeled using first-order differential equations with noise terms [93]. Consider a nonlinear dynamical system with a vector $\mathbf{x}(t)$ of state variables described by first-order differential equations

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \delta(\mathbf{x}, t). \tag{5.1}$$

The force vector $\mathbf{f}(\mathbf{x}(t))$ models the deterministic evolution of the system and we assume $\mathbf{f}(\mathbf{x}(t))$ does not depend on time t explicitly for simplicity. It is expected to be continuous and differentiable in the vicinity of the equilibrium. The noise term, $\delta(\mathbf{x}, t)$ is modeled as Gaussian white noise with zero mean. The covariance matrix of the noise term is denoted by $\mathbf{D}(\mathbf{x})$.

Suppose the dynamics of this multi-dimensional system are characterized by small fluctuations around a steady state, and let $\mathbf{z}(t) = \mathbf{x}(t) - \boldsymbol{\mu}$ denote a vector of such small deviations from the steady state, where $\boldsymbol{\mu}$ denotes the steady state value of the

state variables in the absence of noise. Let $p(\mathbf{z}, t)$ be the probability density function that describes the likelihood of these deviations falling within a particular range of values. Because of the small deviation assumption, the force can be replaced by its linear approximation

$$\mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\boldsymbol{\mu}) + \mathbf{F}\mathbf{z} = \mathbf{F}\mathbf{z}, \tag{5.2}$$

where $\mathbf{f}(\boldsymbol{\mu}) = 0$.

Thus, the probability density function $p(\mathbf{x}, t)$ can be described using a linear Fokker-Planck equation [113] as

$$\frac{\partial p(\mathbf{z}, t)}{\partial t} = \sum_{i,j=1}^{N} -F_{ij} \frac{\partial (z_j p)}{\partial z_i} + \frac{1}{2} \sum_{i,j=1}^{N} D_{ij} \frac{\partial^2 p}{\partial z_i \partial z_j}, \tag{5.3}$$

where $\mathbf{F}$ is the force matrix that determines the expected trajectory of $\mathbf{z}$ toward zero, and $\mathbf{D}$ is the diffusion matrix that describes the covariance of a Gaussian white noise that acts on $\mathbf{z}$. The initial condition of this Fokker-Planck equation is set to be $p(\mathbf{z}, t) = \delta(\mathbf{0})$ and the boundary condition is $p(\pm\boldsymbol{\infty}, t) = 0$. The solution of this linear system is a multi-dimensional Gaussian density function [93]. Here, we consider the simple case where $\mathbf{D}$ is constant. We further consider that the fluctuations have reached a stable state around an equilibrium after the transients died out. In such cases, the covariance matrix $\boldsymbol{\Sigma}$ of the solution, which describes the correlation between different state variables, is solely dependent on the force matrix $\mathbf{F}$ and the diffusion matrix $\mathbf{D}$ [113]. Therefore, we can use the covariance matrix $\boldsymbol{\Sigma}$, which can be estimated directly from the measurements, to infer the characteristics of the force matrix $\mathbf{F}$ which is not necessarily available for a real dynamical system.

To calculate the covariance matrix $\boldsymbol{\Sigma}$, we can use the decomposition of Kwon and coau-

thors [114] which shows that $\mathbf{\Sigma}$ can be written as

$$\mathbf{\Sigma} = -\mathbf{F}^{-1}(\mathbf{D} + \mathbf{Q})/2, \tag{5.4}$$

where $\mathbf{Q}$ is an antisymmetric matrix with zeros on its diagonal which satisfies

$$\mathbf{FQ} + \mathbf{QF}^{\tau} = \mathbf{FD} - \mathbf{DF}^{\tau}, \tag{5.5}$$

where the superscript τ indicates the transpose.

Suppose $\mathbf{F}$ has distinct eigenvalues. Denote the eigenvalues of the covariance matrix $\mathbf{\Sigma}$ as $\sigma_1, \sigma_2, ..., \sigma_N$. In the Supplementary Material, we prove that under the assumption of a codimension one bifurcation, the largest eigenvalue of the covariance matrix σ_1 becomes much larger than the other eigenvalues if the real part of the dominant eigenvalue of the force matrix becomes closer to 0 compared to the rest of the eigenvalues. This is because the dynamics along the direction of the eigenvector corresponding to the dominant eigenvalue become slower as the dominant eigenvalue of the Jacobian matrix approaches zero. Because the other eigenvalues are not approaching zero at the same rate as the dominant eigenvalue, the variance of the dynamics along that direction increases at a much higher rate. Thus, the largest eigenvalue of the covariance matrix and the percentage it accounts for of the total variation can be used as early warning signals.

The Fokker-Planck equation has been used in a number of papers to study a system approaching the bifurcation [115, 15]. However, it is important to note that the solution to the linear Fokker-Planck equation (Eq. 5.3) is an approximation to the probability density function of the original system, which is the solution to a nonlinear Fokker-Planck equation that cannot be guaranteed to remain Gaussian as the system becomes arbitrarily close to the bifurcation point. In addition, at the bifurcation point, the probability density function of the original system will depend on time

(even at long times), and the variance will grow over time. Nevertheless, we observed that the covariance of the Gaussian distribution obtained from the linearization is a good approximation close to the bifurcation point, as shown in the Early warning signals section, and can be used as an indicator of the system approaching bifurcation.

For the case where the dominant eigenvalues of $\mathbf{F}$ are complex conjugate pairs, the relationship between the largest eigenvalues of the covariance matrix and the real part of the dominant eigenvalues are similar. The difference is that the subspace in which the variance of the dynamics increases at a much higher rate is now two-dimensional. For simplicity the example we give in the following sections only has real eigenvalues. Further comments about systems with complex conjugate pairs are also included in the Supplementary Material.

5.2.2 Spatial ecological model

We consider a general 2D spatial model under the assumption that the space is discrete and the dynamics take place in an $n \times n$ square lattice which consists of coupled cells [116, 29]. The dynamics at each location (i, j) are affected by a reaction-type process , a diffusion process and a random excitation modeled by a random walk process $dW_{i,j}$. The reaction-type process is described by the nonlinear deterministic function $g(X_{i,j}, r(i, j), c)$. Each cell is also connected to its neighbors through a diffusive process. The general form of this model is

$$dX_{i,j} = g(X_{i,j}, r(i, j), c) + R(X_{i+1,j} + X_{i-1,j} + X_{i,j+1} + X_{i,j-1} - 4X_{i,j}))dt + \sigma dW_{i,j} \quad (5.6)$$

where $X_{i,j}$ is the state variable at cell (i, j) in a 2D space, $r_{i,j}$ is the heterogeneous parameter that changes with location, c is the bifurcation parameter, R is the constant dispersion rate, and σ is the standard deviation of the white noise $dW_{i,j}$ at location (i, j). $i = 1, 2, ..., n$ and $j = 1, 2, ..., n$. We use a symmetric boundary condition so

that cells at the boundary only have exchange with its neighboring cells but not with the boundary.

In the neighborhood of the equilibrium, the reaction-type process can be approximated by its linear approximation. Thus the probability density function of the state variables in Eq. 5.6 can be approximated by a Fokker-Planck equation as

$$\frac{\partial p(\tilde{X}_i, t)}{\partial t} = \sum_{i,j=1}^{n^2} -J_{ij} \frac{\partial(\tilde{X}_j p)}{\partial \tilde{X}_i} + \frac{1}{2} \sum_{i,j=1}^{n^2} D_{ij} \frac{\partial^2 p}{\partial \tilde{X}_i \partial \tilde{X}_j}, \tag{5.7}$$

where $\tilde{\mathbf{X}}$ is all the state variables in Eq. 6 concatenated and detrended (i.e., $\tilde{X}_{n(j-1)+i} = X_{i,j} - \mu_{i,j}$), μ is the expected value of $\mathbf{X}$ under small random excitations, $\mathbf{J}$ is the Jacobian matrix of the force function evaluated at the equilibrium, and $\mathbf{D}$ is the matrix describing the covariance of the random excitation. In this study, $\mathbf{D}$ is set for simplicity to be an identity matrix, which means noise terms at different cells are independent of each other. In the supplementary material we've shown that the choice of $\mathbf{D}$ does not affect the results as long as the stochastic perturbations from the environment do not concentrate on one patch.

The equation for the elements of the Jacobian matrix $\mathbf{J}$ can be expressed as

$$J_{i,j} = \begin{cases} \frac{dg_i(\tilde{X}_i)}{d\tilde{X}_i}|_{\tilde{X}_i = \tilde{X}_{i,0}} - 4R, & \text{if } i = j \\ R, & \text{if } i - j = 1, i \bmod n \neq 1 \\ R, & \text{if } j - i = 1, j \bmod n \neq 1 \\ R, & \text{if } i - j = n \\ R, & \text{if } j - i = n \end{cases}$$

where g_i is the reaction function at cell $(i \mod n, \lfloor i/n \rfloor)$, and $\tilde{X}_{i,0}$ is the steady state value of $\tilde{X}_i$. Clearly $\mathbf{J}$ is a symmetric matrix containing real numbers. Thus, the

eigenvalues of **J** must be real.

To illustrate how the proposed early warning signals can be used to anticipate critical transitions of an ecological system, we adapt three well-studied models [22, 117, 118] to the form of Eq. 5.6 by adding the dispersion and noise terms. The first model describes the dynamics of biomass under harvesting. The second model describes the dynamics of nutrient of a eutrophic lake, and the third model describes the dynamics of macrophyte in a shallow lake. All three models have alternative stable states in their original form. For example, the amount of biomass of the harvesting model is stable at an equilibrium with high population at a low harvesting rate. As the harvesting rate increases, the system collapses to a low population equilibrium. Details about the three models we use can be found in Table. B.1. The parameters values are obtained from Dakos and coauthors [22]. Heterogeneous parameters ($r_{i,j}$) are introduced by randomly setting the value of $r_{i,j}$ within a certain range. $n^2 = 20 \times 20 = 400$ cells are used in this model. The values of $r_{i,j}$ we use in this study are included in the Supplementary Materials.

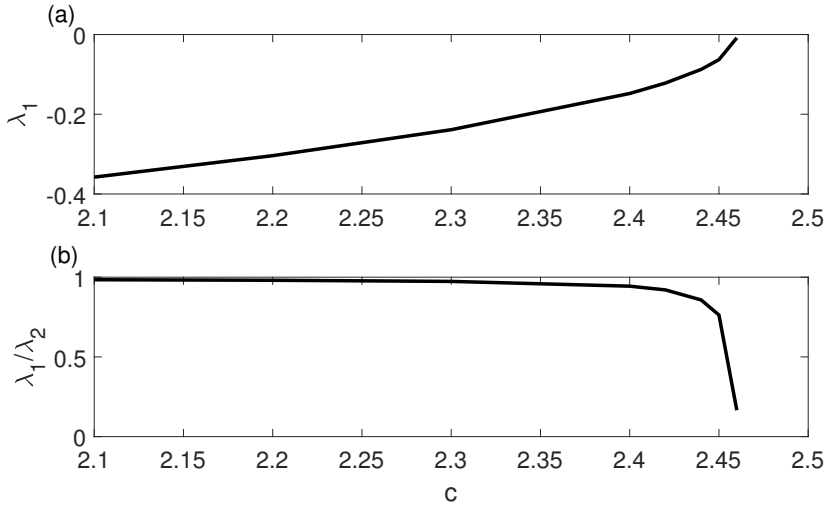

Figure 5.1: Variation of the eigenvalues of the spatial harvesting model; the largest eigenvalue (a) and the largest eigenvalue divided by the second largest eigenvalue (b) are shown versus the bifurcation parameter c.

The relationship between the dominant eigenvalue λ_1 of the Jacobian matrix **J**

Model and Parameter	Definition and value
Harvesting model	
$g(X_{i,j}, r(i,j), c) = r_{i,j}X_{i,j}(1 - \frac{X_{i,j}}{K}) - c\frac{X_{i,j}^2}{X_{i,j}^2+1}$	
$X_{i,j}$	Resource biomass at cell (i,j); state variable
c	Maximum harvesting rate; bifurcation parameter
$r_{i,j}$	Maximum growth rate at cell (i,j); heterogeneous parameter
K	Carrying capacity, 10
R	Dispersion rate, 0.2
σ	SD of white noise, 0.1
Eutrophication model	
$g(X_{i,j}, r(i,j), c) = a - r_{i,j}X_{i,j} + c\frac{X_{i,j}^8}{X_{i,j}^8+1}$	
$X_{i,j}$	Nutrient concentration at cell (i,j); state variable
c	Nutrient loading rate; bifurcation parameter
$r_{i,j}$	Nutrient loss rate at cell (i,j); heterogeneous parameter
a	Maximum recycling rate, 0.5
R	Dispersion rate, 0.2
σ	SD of white noise, 0.05
Vegetationturbidity model	
$g(X_{i,j}, r(i,j), c) = r_v X_{i,j}(1 - X_{i,j}\frac{r_{i,j}^4+E_{i,j}^4}{r_{i,j}^4})$	
$E_{i,j} = \frac{h_v c}{h_v+X_{i,j}}$	
$X_{i,j}$	Vegetation cover at cell (i,j); state variable
c	Background turbidity; bifurcation parameter
$r_{i,j}$	Half-saturation turbidity constant at cell (i,j); heterogeneous parameter
r_v	Maximum vegetation growth rate, 0.5
h_v	Half-saturation vegetation cover constant, 0.2
R	Dispersion rate, 0.2
σ	SD of white noise, 0.1

Table 5.1: Model details and parameter values used in the study

and the bifurcation parameter c of the harvesting model is shown in Fig. 5.1a as an example. As the system moves toward the bifurcation, the dominant eigenvalue becomes closer to zero. Meanwhile, the relationship between $\frac{\lambda_1}{\lambda_2}$ and the bifurcation parameter c is shown in Fig. 5.1b. When the system is far away from the bifurcation, λ_1 is close to λ_2. However, as the system moves toward the bifurcation, $\frac{\lambda_1}{\lambda_2}$ becomes closer to 0. Therefore, we expect to observe an increase of the largest eigenvalue of the covariance matrix as the system becomes closer to the bifurcation. The percentage the largest eigenvalue of the covariance matrix accounts for of the total variation is also expected to increase. The situation is similar for the other two models. To demonstrate how the proposed early warning signals can be used to forecast critical transitions, we apply the approach to the all three spatial models in the following section.

5.2.3 Early warning signals

We estimate the eigenvalues of the covariance matrix of the harvesting model at several different control parameter values as shown in Fig. 5.2. Each line shows all the eigenvalues of the covariance matrix in a descending order at a specific parameter value. The index is an integer which varies from 1 to 400, with index 1 meaning the largest eigenvalue, index 2 meaning the second largest eigenvalue, and so on.

As the system moves toward the bifurcation, the eigenvalue spectrum increases. Because the eigenvalues of the covariance matrix are closely related to the variance of principal components, this discovery is consistent with the observation that the variance increases as the system moves toward the bifurcation. In particular, the largest eigenvalue of the covariance increases at a much larger rate than the other eigenvalues. Consequently, the largest eigenvalue becomes much larger than other eigenvalues when the system becomes close to the bifurcation. Therefore, we propose two early warning signals:

- The first warning signal is the value of the largest eigenvalue of the covariance matrix σ_1,
- The second warning signal is the percentage the largest eigenvalue accounts for of the total variation $\frac{\sigma_1}{\sqrt{\sigma_1^2+...+\sigma_N^2}}$, where N is the number of states.

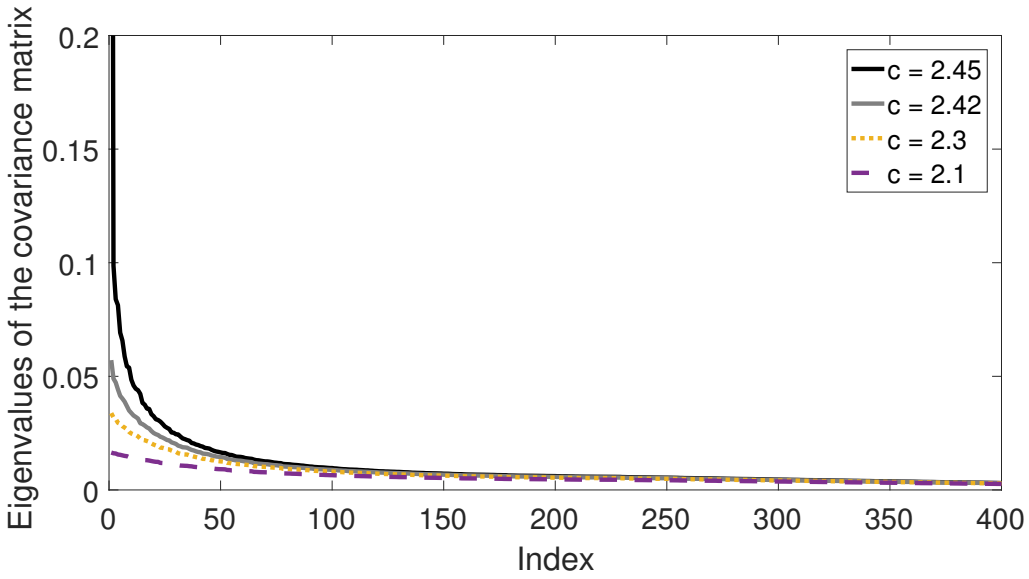

Figure 5.2: Change of the spectrum of the covariance matrix as the system moves toward the bifurcation at c = 2.47. The bifurcation point $c = 2.47$ is computed using the deterministic part of the harvesting model. Each line represents all the eigenvalues of the covariance matrix under a certain parameter value. The index is simply an integer which varies from 1 to 400, with index 1 meaning the largest eigenvalue, index 2 meaning the second largest eigenvalue, and so on.

The relationship between the largest eigenvalue and the bifurcation parameter for all three models is shown in Fig. 5.3b, e, h. The eigenvalues are first calculated analytically using the decomposition described in the Main Concept. We also estimated the eigenvalues of the covariance matrix using a moving window on data obtained from a system with a time-varying bifurcation parameter $c(t) = c_0 + \delta t$. In all simulations, time step δ is set to 1/16000. A total of 16000 snapshots are collected from the simulation. The moving window has a size of 1000 snapshots. As expected, the largest eigenvalue of the covariance matrix increases as the system moves toward the

bifurcation point. Moreover, the largest eigenvalue of the covariance matrix becomes dominant as the dominant eigenvalue approaches 0. Therefore, Fig. 5.3c, f, i suggest that the percentage that the largest eigenvalue accounts for of the total variation also increases as the system approaches the bifurcation. The percentage as an early warning signal is particularly useful because it has a clear upper limit of 100%. That is in contrast to the largest eigenvalue of the covariance matrix, which theoretically can increase without bound.

It is important to note that the analytical expression of the covariance matrix is generally not available for a real system because the forms of the force matrix and the diffusion matrix are seldom known. However, we can estimate the covariance matrix directly from the time series data without relying on any prior knowledge of the system. Results from Fig. 5.3 show that the estimated eigenvalues of the covariance matrix agree quite well with the analytical eigenvalues close to the bifurcation. For parameters extremely close to the bifurcation point, the estimated eigenvalues generally have higher values compared to the analytical eigenvalues. This is caused by the moving window used in the estimation. The estimated eigenvalue is essentially an average of eigenvalues in that moving window. As the system approaches the bifurcation, the largest eigenvalue of the covariance matrix increases quickly, and that causes the early takeoff of the estimation compared to the analytical results. More details about the estimation method are presented in the Methods section, where we show that eigenvalues estimated using the shrinkage method agree well with the analytical results when a fixed parameter is used for simulation.

5.2.4 Locations of tipping points

Another advantage of the proposed methods is that by examining the covariance matrix, we can obtain both the dominant eigenvalue, and its corresponding eigenvector (i.e., the first principal component). Examining the first principal component,

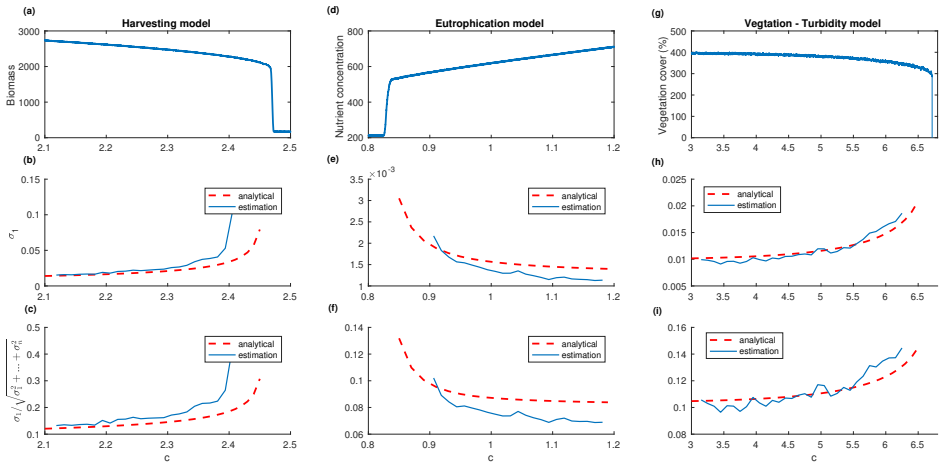

Figure 5.3: (a), (d), and (g) Sum of the state variables as the bifurcation parameter c changes over time. (b), (e), and (h) Largest eigenvalue σ_1 of the covariance matrix estimated using a moving window as the bifurcation parameter c changes over time. (c), (f), and (i) Largest eigenvalue of the covariance matrix over the Euclidean norm of a vector consisting of all the eigenvalues $\frac{\sigma_1}{\sqrt{\sigma_1^2+\ldots+\sigma_n^2}}$ estimated using a moving window as the bifurcation parameter c changes over time.

we can identify cells that have the largest variance and are the most vulnerable to the critical transition. This is because as the system becomes close to the critical transition, the dominant eigenvector of the covariance matrix becomes close to the dominant eigenvector of the force matrix. This can help identify the crucial areas of the system to monitor and possibly prevent the critical transition.

To examine the relationship between the first principle component and vulnerable regions, we calculate local eigenvalues corresponding to groups of cells. By local eigenvalues we mean the dominant eigenvalue of the force matrix of a single cell or a local group of cells without considering the diffusion effects of the neighboring cells.

We define a local group with size $i \times i$ at location (j, k) as cells:

$$\begin{bmatrix} (j,k) & (j,k+1) & \cdots & (j,k+i-1) \\ (j+1,k) & (j+1,k+1) & \cdots & (j+1,k+i-1) \\ \vdots & \vdots & \ddots & \vdots \\ (j+i-1,k) & (j+i-1,k+1) & \cdots & (j+i-1,k+i-1) \end{bmatrix}, \tag{5.8}$$

under the restriction that $j + i - 1,\ k + i - 1 \leq 20$.

In Fig. 5.4, we show the dominant local eigenvalues of all the 1 $\times$ 1, 2 $\times$ 2, 3 $\times$ 3 and 4 $\times$ 4 local groups of the harvest model. Because of the diffusion effects, the dominant eigenvalues tend to decrease as the number of cells in the local group increases. The diffusion effect, on the other hand, does not have a significant effect on cells that are too far away from the local group. Therefore, the eigenvalue map stabilizes as the size of the local group reaches a certain level.

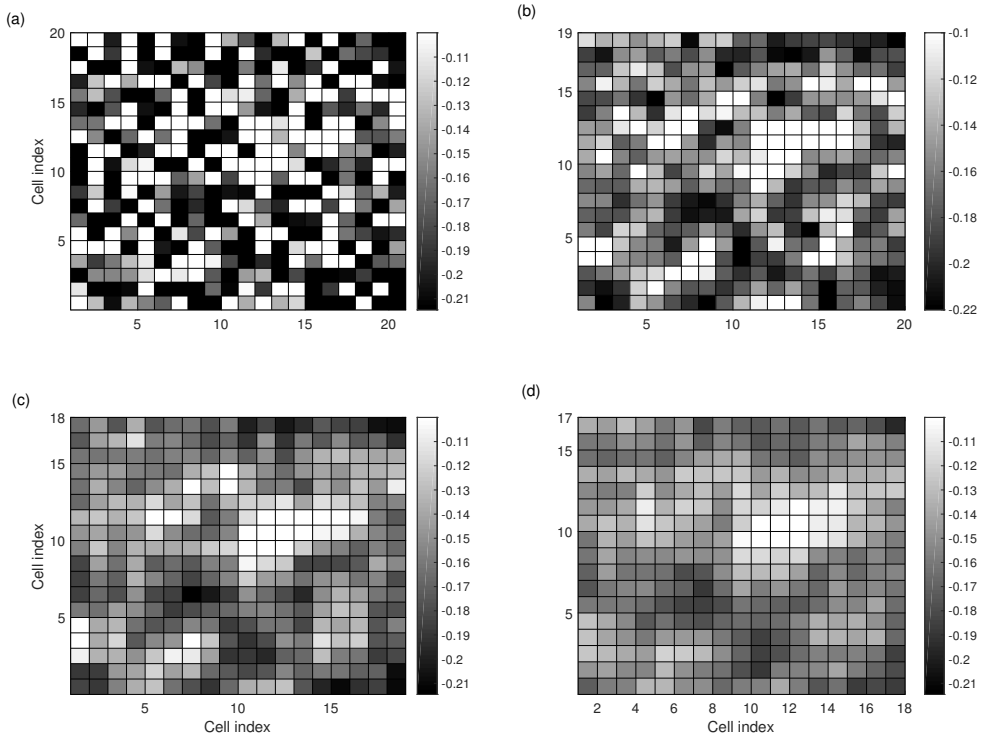

Figure 5.4: Local eigenvalues of the harvesting model. Only the dominant eigenvalue of the local area are plotted. For each local group, the dominant eigenvalue is plotted at the upper left cell of the group. (a) 1×1, (b) 2×2, (c) 3×3 and (d) 4×4 cells are used to construct the local groups.

Now we want to compare the dominant local eigenvalues of the harvesting model with its dominant principal component. Fig. 5.5 (a) exhibits the analytic first principal component of the covariance matrix, while Fig. 5.5 (b) shows the first principal component estimated from simulation data. It is obvious that they share a dominant peak in the upper right region. There are other smaller peaks in Fig. 5.5 (b) due to estimation errors. In Fig. 5.5 (c), we show the dominant local eigenvalues of the 4×4 groups. It is obvious that the area with the largest amplitudes in the first principal component coincides with the area with the largest local eigenvalues. Thus, regions with local eigenvalues closest to 0 are identified as regions most vulnerable to critical transition by the first principle component.

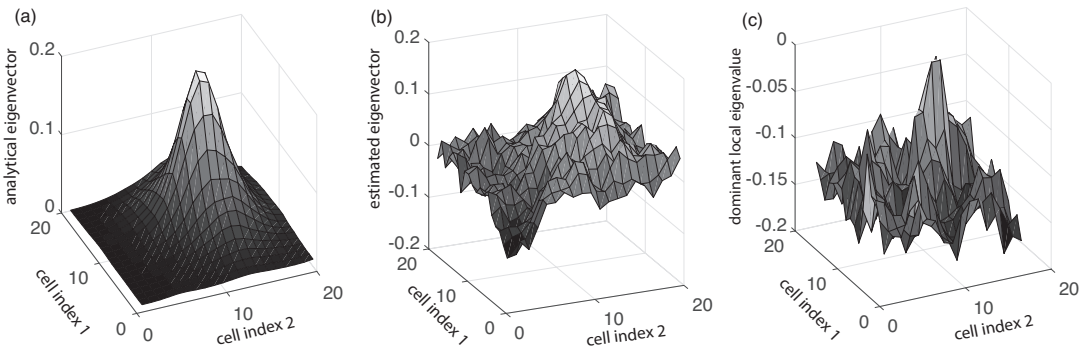

Figure 5.5: (a) The eigenvector corresponding to the dominant eigenvalue of the covariance matrix (analytical), (b) The eigenvector corresponding to the dominant eigenvalue of the covariance matrix (estimated using simulation data), (c) Dominant eigenvalues of the force matrix of the local cell groups.

5.3 Discussion

We proposed two spatial early warning signals. These signals are based on the eigenvalues of the covariance matrix, which indicate the variance of the principal components. Overall, our study suggests that an increase in the largest eigenvalue of the covariance matrix and the percentage it accounts for of the total variation can serve to create early warning signals for the critical transition of a system with spatial correlation.

Here, we compared the two proposed early warning signals with past early warning signals [42] including

1. spatial variance $\sigma^2 = \frac{1}{HL}\sum_{h=1}^{H}\sum_{l=1}^{L}(z_{h,l} - \bar{\mathbf{z}})^2$, where H and L are the number of cells in the vertical and horizontal direction, and $\bar{\mathbf{z}}$ is the mean value of $\mathbf{z}$ over the whole space.

2. spatial skewness $\gamma = \frac{1}{HL}\sum_{h=1}^{H}\sum_{l=1}^{L}\frac{(z_{h,l}-\bar{\mathbf{z}})^3}{\sigma^3}$, where σ is the square root of the spatial covariance.

3. spatial correlation function $C_2(r) = \frac{HL\sum_{i=1}^{H}\sum_{m=1}^{H}\sum_{j=1}^{L}\sum_{l=1}^{L} w_{i,j;m,l}(z_{i,j}-\bar{\mathbf{z}})(z_{m,l}-\bar{\mathbf{z}})}{W\sum_{l=1}^{L}\sum_{h=1}^{H}(z_{h,l}-\bar{\mathbf{z}})^2}$, where $w_{i,j;m,l}$ is 1 if spatial units $[i,j]$ and $[m,l]$ are separated by a distance r,

and is 0 otherwise, and W is the total number of ordered pairs of units separated by the distance r.

The two proposed early warning signals are different from past ones in the sense that it does not take the variation of expected values of the state variables into account. Past spatial early warning signals are all affected by the expected value of the state variables. To better explore this, we next take spatial variance as an example. The expected value of the spatial variance can be expressed as,

$$\begin{aligned}
\mathrm{E}(\hat{\sigma}^2_{\text{spatial}}) &= \mathrm{E}\left(\frac{1}{M}\sum_{i=1}^{M}(z_i-\bar{\mathbf{z}})^2\right) \\
&= \frac{M-1}{M^2}\mathrm{E}\left(\sum_{i=1}^{M} z_i^2\right) - \frac{1}{M^2}\sum_{i\neq j}\mathrm{E}\left(z_i z_j\right) \\
&= \frac{M-1}{M^2}\sum_{i=1}^{M}\left(\mu_i^2+\Sigma_{ii}\right) - \frac{1}{M^2}\sum_{i\neq j}\left(\mu_i\mu_j+\Sigma_{ij}\right),
\end{aligned}$$

where we used $\mathrm{E}(z_i z_j) = \Sigma_{ij} + \mu_i\mu_j$. Thus we have

$$\begin{aligned}
\mathrm{E}(\hat{\sigma}^2_{\text{spatial}}) &= \frac{1}{M^2}\left((M-1)\sum_{i=1}^{M}\mu_i^2 - \sum_{i\neq j}\mu_i\mu_j\right) \\
&\quad + \frac{1}{M^2}\left((M-1)\sum_{i=1}^{M}\Sigma_{ii} - \sum_{i\neq j}\Sigma_{ij}\right),
\end{aligned}$$

where $\hat{\sigma}^2_{\text{spatial}}$ is the empirical spatial variance, M is the number of cells in the grid, μ_i is the expected value of state i, and $\mathbf{\Sigma}$ is the covariance matrix. Therefore, the spatial variance is affected by the expected values of the state variables (first half of the right hand side of the equation), and the covariance matrix (second half of the equation). As the condition of the system changes gradually, the expected value of the state variables may also vary. The changes in the indicators caused by the expected values, however, do not necessarily provide any information about the stability of the system, especially for heterogeneous systems. Therefore, it is hard to predict in

general how these existing spatial early warning signals are affected by the eigenvalues of the system.

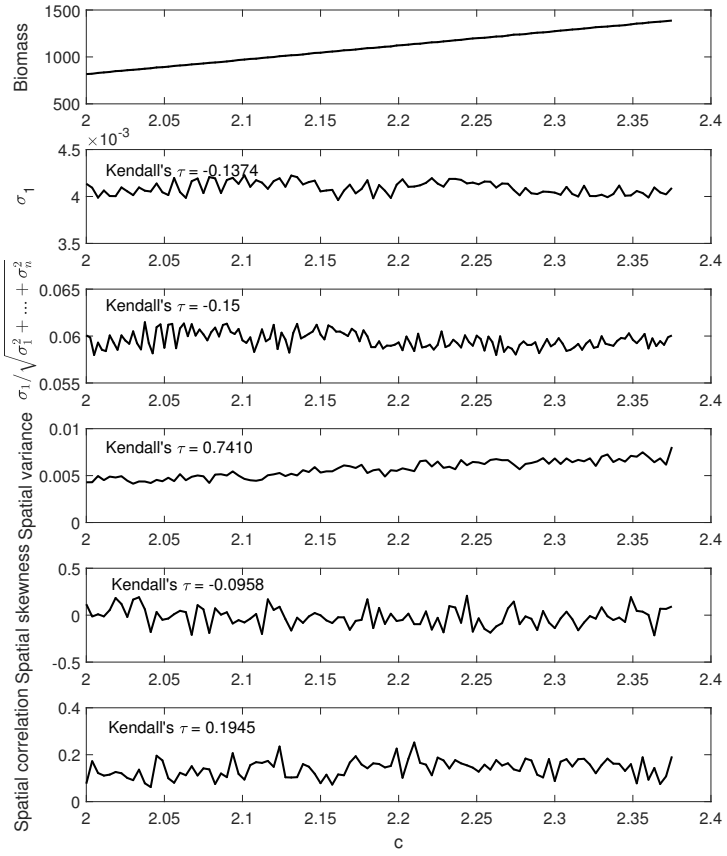

Figure 5.6: A comparison between the two proposed early warning signals, i.e. largest eigenvalue of the covariance matrix and the percentage it accounts for, with three past spatial early warning signals using simulation data obtained from system governed by 5.9.

Next, we use an example to illustrate how the changes in mean values can affect the performance of spatial variance as an early warning signal. Simulation data is collected from a simple spatial model

$$dX_{i,j} = (r_{i,j}c - X_{i,j} + R(X_{i+1,j} + X_{i-1,j} + X_{i,j+1} + X_{i,j-1} - 4X_{i,j}))dt + \sigma dW_{i,j}, \quad (5.9)$$

where **r** is the same as the **r** we used for the harvesting model, and c is the control parameter. Other parameters are the same as in Table B.1. In this simple model, there is no bifurcation as the control parameter c changes, therefore there is no critical slowing down.

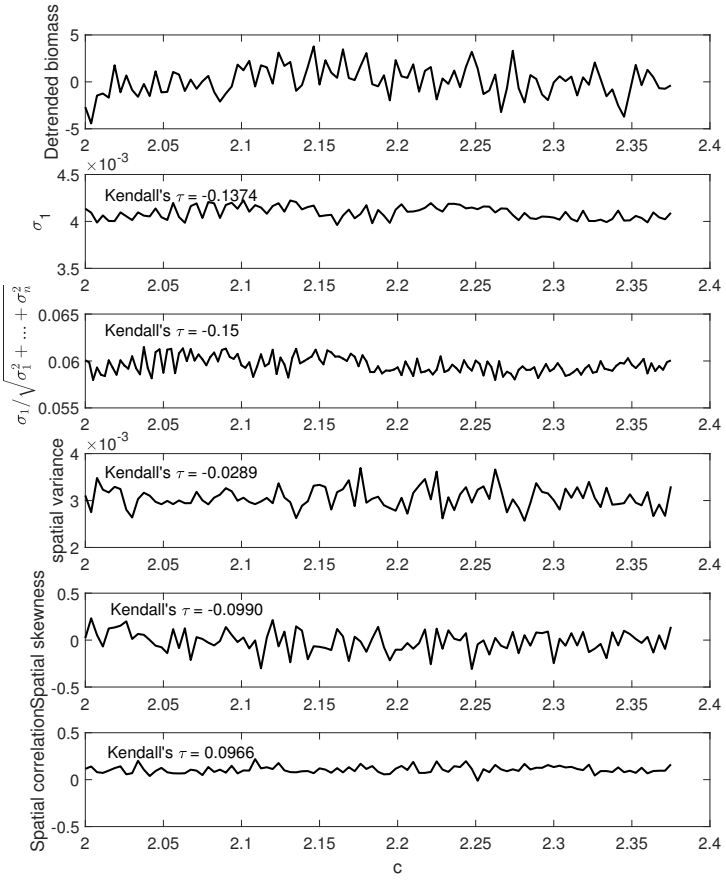

Figure 5.7: The comparison between the two proposed early warning signals, i.e. largest eigenvalue of the covariance matrix and the percentage it account for, with three past spatial early warning signals using detrended simulation data obtained from Model 5.9.

Spatial early warning signals are calculated using simulation data, and are shown in Fig 5.6. As the parameter c increases, values of the state variables also increase due to the change of their expected value. During this process, the largest eigenvalue of the covariance matrix remains stable due to a lack of critical slowing down. Spatial

variance, on the contrary, has a clear upward trend. To compare different spatial early warning signals, we quantified their trend using the nonparametric Kendall's τ rank correlation of the control parameter c and the spatial early warning signals. In this example, the spatial variance is strongly affected by the change of expected values with a Kendall's τ of over 0.74.

To better illustrate this point, all the spatial early warning signals are calculated again using detrended simulation data as shown in Fig. 5.7. As expected, the upward trend in spatial variance has disappeared after detrending the simulation data. Therefore, it is highly recommended to detrend the data before applying such spatial early warning signals.

Next, we apply the proposed and past early warning signals to the detrended simulation data from the harvesting model. The results are shown in Fig. 5.8.

As the system moves towards the bifurcation, both the spatial variance and the spatial correlation increase, while the spatial skewness does not have an obvious trend. The proposed early warning signals, i.e. the largest eigenvalue of the covariance matrix and the ratio $\sigma_1/\sqrt{\sigma_1^2 + ... + \sigma_n^2}$, have the largest Kendall's τ, which indicates that they have the most consistent upward trend compared to other spatial early warning signals. Moreover, the proposed early warning signals have sharp increases close to the bifurcation, while others increase linearly. Therefore, the sharp increase in the proposed early warning signals can be seen as an indication that the system is approaching the bifurcation.

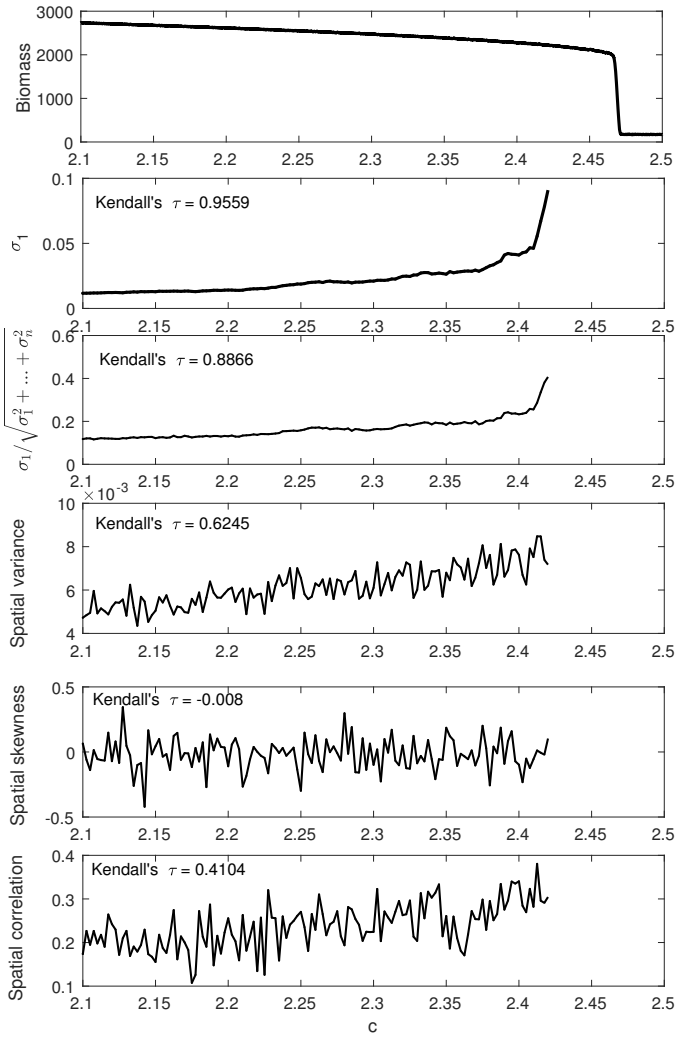

Figure 5.8: A comparison between the two proposed early warning signals, i.e. largest eigenvalue of the covariance matrix and the percentage it account for, with three past spatial early warning signals using detrended simulation data obtained from the harvesting model.

In conclusion, this study explored how eigenvalues and eigenvectors of a spatial covariance matrix related to critical transitions in spatially extended ecological systems.

Eigenvalues of the covariance matrix better capture critical slowing down due to their direct relationship with the eigenvalues of the force matrix that characterizes the dynamics, compared to past spatial early warning signals. We therefore propose

to use the largest eigenvalue of the covariance matrix as a spatial early warning signal. By establishing the relationship between the eigenvalues of the covariance matrix and the eigenvalues of the force matrix, we mathematically show that the dominance of the largest eigenvalue of the covariance matrix can also be used as an early warning signal. We compared these proposed signals with existing ones. The proposed early warning signals can potentially be applied to other high dimensional systems, such as multispecies systems [47], complex structures [119] and so on.

This approach may be used to identify also the vulnerable regions in a spatially correlated system. By studying the eigenvector corresponding to the dominant eigenvalue of the covariance matrix (the first principle component), we can obtain important clues regarding the region where the transition is most likely to occur. The correlation between the vulnerable regions and the first principle component will be stronger as the system becomes closer to the critical transition. This means that the percentage the largest eigenvalue of the covariance matrix account for of the total variation is not only an indicator of a potential critical transition, but also an indicator of a better opportunity for dimension reduction analyses.

5.4 Methods

5.4.1 Covariance matrix estimation

Here we show the process of covariance matrix estimation by choosing a parameter value close to the bifurcation, simulating the model to obtain a stationary response to the random excitation, and estimating the covariance matrix.

An example of simulation results is shown in Fig. 5.9a, where the sum of the biomass is plotted. Three different ways are used to estimate the covariance matrix, i.e.:

1. Unbiased empirical covariance matrix.

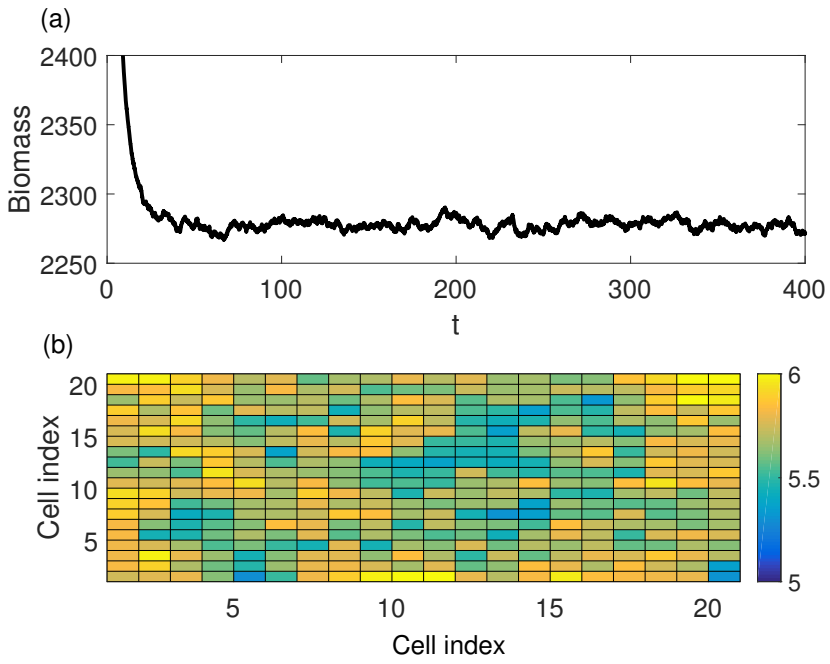

Figure 5.9: (a) An example of the time series of the total amount of biomass when the bifurcation parameter c is 2.4. (b) A snapshot of the state variable values at each cell.

2. Shrinkage approach.

3. Analytical method.

First, the unbiased empirical covariance matrix $\mathbf{S}$ is used to estimate the covariance matrix. Each entry of $\mathbf{S}$ is defined as

$$S_{jk} = \frac{1}{n-1} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k) \tag{5.10}$$

where x_{ij} is the i_{th} measurement collected at the j_{th} cell, and $\bar{x}_j$ is the average of measurements collected at the j_{th} cell over time, n is the number of snapshots used for the estimation.

However, the unbiased empirical covariance matrix is not a good estimate of the covariance matrix when the number of snapshots is small compared to the number of variables, as pointed out in [120, 121]. This is because the sample covariance matrix

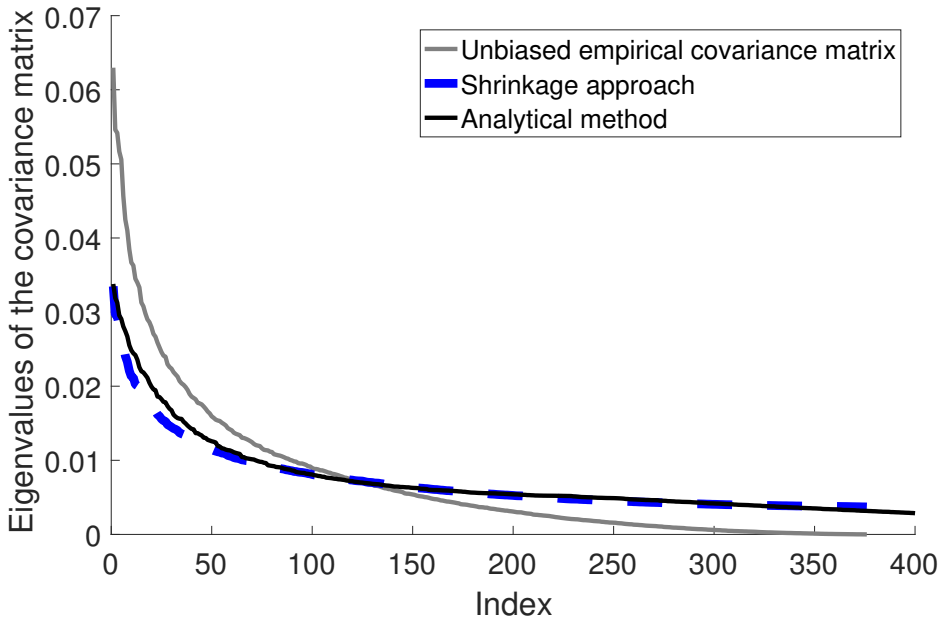

Figure 5.10: Estimation of eigenvalues of the covariance matrix using three different methods: analytical covariance matrix, sample covariance matrix, shrinkage estimation method. $c = 2.4$ is used to obtain the simulation data. 200 snapshot are used for the covariance matrix estimation.

$\mathbf{S}$ might not be positive definite anymore when only a small number of snapshots are available. In such cases, the sample covariance matrix tends to overestimate the value of its largest eigenvalues, while underestimate the rest. A shrinkage approach can be used to estimate the covariance matrix under such circumstances. The linear shrinkage approach suggests to combine a high-dimensional unconstrained estimate $\mathbf{S}$ and a low-dimensional constrained estimate $\mathbf{T}$ in a weighted average given by

$$\mathbf{S}^* = \eta\mathbf{S} + (1-\eta)\mathbf{T}, \tag{5.11}$$

where $\mathbf{S}^*$ is the regularized estimate, $1-\eta \in [0,1]$ is the shrinkage intensity. η is estimated by minimizing a squared error loss risk function which is a combination of mean square error and variance as shown in [120]. The low-dimensional constrained estimate $\mathbf{T}$ is chosen based on the presumed lower-dimensional structure in the data

set. In the case of the spatial harvesting model, the assumption is that the variance along different directions does not decrease dramatically after the first few principal components. Examples of values of $\mathbf{T}$ can be found in Schafer and coauthors [120].

The covariance matrix is obtained analytically also by calculating the Jacobian matrix $\mathbf{J}$ of the deterministic solution and using the decomposition method in Section 2. The analytical covariance matrix is used as a benchmark for the estimation results. Eigenvalues estimated by the three methods are calculated and plotted in Fig. 5.10. R package *corpcor* is used for the covariance matrix estimation. In particular, *cov.shrink* function is used for the shrinkage estimations. Default parameters are used. As expected, the shrinkage method out performs the sample covariance matrix in estimating the eigenvalues.

CHAPTER VI

Conclusions and Future Work

6.1 Conclusions

The main results of the dissertation can be summarized as follows:

- We presented a new method that uses the transient recovery of parametrically excited systems from large perturbations to predict the bifurcation diagram. This forecasting method provides quantitative insights into parametrically excited systems despite the absence of actual models of these systems. As expected, the accuracy of forecasting increases when more data are collected closer to the critical transition. One of the important benefits of the method is that it is applicable in advance of the transition. The transition and the overall bifurcation diagram can be predicted without endangering the system by passing the transition point. We demonstrated experimentally and numerically that transitions caused by either bifurcation of the unforced system or parametric resonance can be predicted using this method.

- We developed a new forecasting method for a high dimensional parametrically excited systems that uses the recoveries of a single variable in the slow manifold to forecast the bifurcation diagram of that variable. Several assumptions are needed for the method to work. First, the system has to be a weakly nonlinear

system for the averaging approximation to be valid. Second, data should be collected at parameter values close to the bifurcation point so that the extrapolation of recovery rate λ will be accurate. Other assumptions include that the average system has a codimension one bifurcation, and the slow manifold varies smoothly with the parameter. It is shown that if recovery can be obtained only for one of the DOFs, the bifurcation diagram of that degree of freedom can still be predicted, as long as this DOF is close to the inertial manifold of the system. It is also discussed that initial perturbations play an important role in this method. The closer the condition after the perturbation is to the slow manifold, the better the forecasting results are.

- An analysis based on the method of multiple scales (MMS) is carried out to study a classic SIR model with periodic forcing. Analytical approximate solutions were obtained to show that parametric resonance can occur through the interaction between nonlinearity and seasonal forcing. The results reveal that large amplitude epidemics can take place if the system is moved into a resonance regime in the parameter plane by the change of any one of its parameters. This analysis can be extended to more complex epidemiological or biological models where more variables or other kinds of nonlinearity are present. An important contribution of this analysis is that it shows that small order parametric excitations can synchronize with the system response, and trigger an order 1 parametric resonance. When parametric resonance is triggered, the excitation drives the growth of the biennial cycles until it is constrained by the nonlinearity.

- We showed how eigenvalues and eigenvectors of a spatial covariance matrix related to critical transitions in spatially extended ecological systems. Eigenvalues of the covariance matrix better capture critical slowing down due to their direct relationship with the eigenvalues of the force matrix that characterizes the dy-

namics, compared to past spatial early warning signals. We therefore propose to use the largest eigenvalue of the covariance matrix as a spatial early warning signal. By establishing the relationship between the eigenvalues of the covariance matrix and the eigenvalues of the force matrix, we mathematically show that the dominance of the largest eigenvalue of the covariance matrix can also be used as an early warning signal. We compared these proposed signals with existing ones. The proposed early warning signals can potentially be applied to other high dimensional systems, such as multispecies systems [47], complex structures [119] and so on.

- We also show that dominant eigenvector of the covariance matrix may be used to identify also the vulnerable regions in a spatially correlated system. By studying the eigenvector corresponding to the dominant eigenvalue of the covariance matrix (the first principle component), we can obtain important clues regarding the region where the transition is most likely to occur. The correlation between the vulnerable regions and the first principle component will be stronger as the system becomes closer to the critical transition. This means that the percentage the largest eigenvalue of the covariance matrix account for of the total variation is not only an indicator of a potential critical transition, but also an indicator of a better opportunity for dimension reduction analyses.

6.2 Future Work

The work presented in this dissertation can be extended in several directions.

1. Early warning signals for spatially extended system proposed in past research [22, 18, 32] and this dissertation focus on systems with homogeneous noise. However, there are a number of examples where random excitations from the environment are heterogeneous [122]. For example, wave induced disturbances are heteroge-

neous on mussel beds [123]. Such spatial heterogeneous disturbances can affect the reliability and robustness of current spatial early warning signals. One possible solution is to look at the local recoveries from perturbations instead of global spatial patterns [41]. This solution leads to another interesting problem, which is what region or state do we pick to monitor local resilience. It is pointed out in [65] that the state most aligned with the dominant eigenvector would be the best one to monitor. Therefore, identifying the best state to monitor under spatially heterogeneous noise will be another interesting topic to study in the future.

2. A lot of natural and engineering system have multiplicative noise where the noise level varies with the states [24]. Current research focuses on systems with additive noise [16, 15], or simple linear multiplicative noise [12, 124]. Some generic early warning signals, such as variance, can be affected by multiplicative noise. For example, when values of state variables decrease close to the bifurcation, variance of the response to multiplicative noise might decrease instead of increase. Therefore, it would be beneficial if the effect of multiplicative noise on generic early warning signals can be studied.

3. Current early warning signals are mostly developed in the context of autonomous systems. However, many critical transitions are accompanied by periodic fluctuation of the environment [125, 88, 28]. Such critical transitions can either be caused by a change in the topological structure of the unforced system, or by a synchronization between a natural mode of the system and the parameter variation. While both cases are important, the second scenario can be particularly dangerous because a small excitation might produce a large response when the driving frequency is close to the resonance frequency of the system. This type of critical transition, however, has not received enough attention from the

early warning signal community. Further research into reliable and robust early warning signals can be beneficial.

APPENDICES

APPENDIX A

Eigenvalues of the covariance matrix

We begin by stating our main assumptions and notation in the next section. Following that, our main analytic results are all stated without proof for the sake of conciseness. Proofs are provided in the final section, with the results restated for convenience.

A.1 Assumptions and notation

Consider a system of N state variables $z_1, \ldots, z_N$ that have a probability density $p(\mathbf{z}, t)$ that satisfies the linear Fokker-Planck equation

$$\frac{\partial p(\mathbf{z}, t)}{\partial t} = \sum_{i,j=1}^{N} -F_{ij} \frac{\partial (z_j p)}{\partial z_i} + \frac{1}{2} \sum_{i,j=1}^{N} D_{ij} \frac{\partial^2 p}{\partial z_i \partial z_j}$$

with a force matrix $\mathbf{F}$ and a diffusion matrix $\mathbf{D}$. Assume that all of the eigenvalues of $\mathbf{F}$ are distinct with negative real parts and are denoted $\lambda_1, \lambda_2, \ldots, \lambda_N$. These eigenvalues are indexed from smallest to largest in terms of the value of their real part (i.e., $|\mathrm{Re}(\lambda_1)| \leq \ldots \leq |\mathrm{Re}(\lambda_N)|$). The diffusion matrix $\mathbf{D}$ is assumed to have all positive eigenvalues. With these assumptions, the stationary distribution of $\mathbf{z}$ is Gaussian and we denote its covariance matrix as $\mathbf{\Sigma}$. Our results concern the relationship between the eigenvalues of $\mathbf{F}$ and $\mathbf{\Sigma}$ as the system undergoes a codimension-1 bifurcation.

In such a bifurcation, typically only one of $\mathbf{F}$'s eigenvalues—or the real part of one complex conjugate pair of $\mathbf{F}$'s eigenvalues—vanishes at the critical transition.

A.2 Eigenvalues of the covariance matrix

Lemma 1. Let the columns of a matrix $\mathbf{T}$ contain the eigenvectors of $\mathbf{F}$. Let $\tilde{\mathbf{\Sigma}}$ be the covariance of the state variables if the eigenvectors are used as their coordinate basis. That is, $\tilde{\mathbf{\Sigma}} = \mathbf{T}^{-1}\mathbf{\Sigma}\mathbf{T}^{-\tau}$. Then the elements of $\tilde{\mathbf{\Sigma}}$ satisfy

$$\tilde{\Sigma}_{ij} = -\frac{\tilde{D}_{ij}}{\lambda_i + \lambda_j}$$

where $\tilde{\mathbf{D}} = \mathbf{T}^{-1}\mathbf{D}\mathbf{T}^{-\tau}$.

Lemma 2. Suppose that the dominant eigenvalue of $\mathbf{F}$ is real and also suppose that $|\tilde{D}_{11}/(2\lambda_1)|\epsilon \geq |\max(\tilde{\mathbf{D}})/(\lambda_2)|$ with $0 < \epsilon \ll 1$. Then the dominant eigenvalue of $\mathbf{\Sigma}$ is equal to $|\tilde{D}_{11}/(2\lambda_1)| + O(\epsilon|\tilde{\Sigma}_{11}|)$ and all of the other eigenvalues are $O(\epsilon|\tilde{\Sigma}_{11}|)$.

Lemma 3. Suppose that the dominant eigenvalues of $\mathbf{F}$ form a complex conjugate pair. Further suppose that $|\tilde{D}_{12}/(2\mathrm{Re}(\lambda_1))|\epsilon \geq |\max(\tilde{\mathbf{D}})/\lambda_3|$, with $0 < \epsilon \ll 1$. Then the sum of the two largest eigenvalues of $\mathbf{\Sigma}$ is $O(|\tilde{D}_{12}/(2\mathrm{Re}(\lambda_1))| + |\tilde{D}_{11}/(2\lambda_1)|)$ and all of the other eigenvalues are $O(\epsilon|\tilde{\Sigma}_{12}|)$.

Theorem 4. Suppose that the magnitude of the real part of a dominant eigenvalue of $\mathbf{F}$ is small such that the assumptions of either Lemmas 2 or 3 are satisfied. Then as this real part approaches zero, if the assumptions of Lemma 2 are satisfied, then the largest eigenvalue of $\mathbf{\Sigma}$ becomes larger absolutely and larger relative to all of the other eigenvalues of $\mathbf{\Sigma}$. Alternatively, if the assumptions of Lemma 3 are satisfied, the sum of the largest two eigenvalues of $\mathbf{\Sigma}$ becomes larger absolutely and relative to all of the other eigenvalues of $\mathbf{\Sigma}$.

A.3 Proofs

Lemma 1. Let the columns of a matrix $\mathbf{T}$ contain the eigenvectors of $\mathbf{F}$. Let $\tilde{\mathbf{\Sigma}}$ be the covariance of the state variables if the eigenvectors are used as their coordinate basis. That is, $\tilde{\mathbf{\Sigma}} = \mathbf{T}^{-1}\mathbf{\Sigma}\mathbf{T}^{-\tau}$. Then the elements of $\tilde{\mathbf{\Sigma}}$ satisfy

$$\tilde{\Sigma}_{ij} = -\frac{\tilde{D}_{ij}}{\lambda_i + \lambda_j} \tag{A.1}$$

where $\tilde{\mathbf{D}} = \mathbf{T}^{-1}\mathbf{D}\mathbf{T}^{-\tau}$.

Proof. Kwon and coauthors[114] show that the covariance matrix $\mathbf{\Sigma}$ may be written as

$$\mathbf{\Sigma} = -\mathbf{F}^{-1}(\mathbf{D} + \mathbf{Q})/2, \tag{A.2}$$

where $\mathbf{Q}$ is an antisymmetric matrix with zeroes on its diagonal which satisfies

$$\mathbf{F}\mathbf{Q} + \mathbf{Q}\mathbf{F}^{\tau} = \mathbf{F}\mathbf{D} - \mathbf{D}\mathbf{F}^{\tau}. \tag{A.3}$$

Next, let $\tilde{\mathbf{Q}} = \mathbf{T}^{-1}\mathbf{Q}\mathbf{T}^{-\tau}$ and $\mathbf{\Lambda} = \mathbf{T}^{-1}\mathbf{F}\mathbf{T}$, the diagonalization of $\mathbf{F}$. Equation (A.3) repeated in terms of these matrices is

$$\mathbf{\Lambda}\tilde{\mathbf{Q}} + \tilde{\mathbf{Q}}\mathbf{\Lambda} = \mathbf{\Lambda}\tilde{\mathbf{D}} - \tilde{\mathbf{D}}\mathbf{\Lambda}. \tag{A.4}$$

Thus the elements of $\tilde{\mathbf{Q}}$ must satisfy

$$(\lambda_i + \lambda_j)\tilde{Q}_{ij} = (\lambda_i - \lambda_j)\tilde{D}_{ij}. \tag{A.5}$$

To use (A.5) to find elements of $\tilde{\Sigma}$, note that (A.2) holds in transformed coordinates as

$$\tilde{\mathbf{\Sigma}} = -\mathbf{\Lambda}^{-1}(\tilde{\mathbf{D}} + \tilde{\mathbf{Q}})/2. \tag{A.6}$$

Putting (A.5) and (A.6) together yields (A.1). □

Lemma 2. Suppose that the dominant eigenvalue of $\mathbf{F}$ is real and also suppose that $|\tilde{D}_{11}/(2\lambda_1)|\epsilon \geq |\max(\tilde{\mathbf{D}})/(\lambda_2)|$ with $0 < \epsilon \ll 1$. Then the dominant eigenvalue of $\boldsymbol{\Sigma}$ is equal to $|\tilde{D}_{11}/(2\lambda_1)| + O(\epsilon|\tilde{\Sigma}_{11}|)$ and all of the other eigenvalues are $O(\epsilon|\tilde{\Sigma}_{11}|)$.

Proof. Let $\mathbf{H} = \tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{\Sigma}}^*$, where the * indicates a conjugate transpose. $\mathbf{H}$ is a Hermitian matrix that can be factored as $\mathbf{H} = \mathbf{A}\ \mathrm{diag}(h_1^2, h_2^2, \ldots, h_N^2)\mathbf{A}^*$ with h_i^2 being distinct eigenvalues of $\mathbf{H}$ and $\mathbf{A}$ being a unitary matrix whose columns consist of the eigenvectors of $\mathbf{H}$ [126]. By hypothesis, all of the elements of $\tilde{\boldsymbol{\Sigma}}$ besides $\tilde{\Sigma}_{11}$ are at least a factor of $1/\epsilon$ smaller in magnitude than $\tilde{\boldsymbol{\Sigma}}_{11}$. We denote this difference in magnitude by noting that $\tilde{\Sigma}_{ij} = O(\epsilon|\tilde{\Sigma}_{11}|)$ for $(i,j) \neq (1,1)$. With this notation, we can write $\tilde{\boldsymbol{\Sigma}}$ as

$$\tilde{\boldsymbol{\Sigma}} = \begin{pmatrix} -\frac{\tilde{D}_{11}}{2\lambda_1} & O(\epsilon|\tilde{\Sigma}_{11}|) & \cdots & O(\epsilon|\tilde{\Sigma}_{11}|) \\ O(\epsilon|\tilde{\Sigma}_{11}|) & O(\epsilon|\tilde{\Sigma}_{11}|) & \cdots & O(\epsilon|\tilde{\Sigma}_{11}|) \\ \vdots & \vdots & \ddots & \vdots \\ O(\epsilon|\tilde{\Sigma}_{11}|) & O(\epsilon|\tilde{\Sigma}_{11}|) & \cdots & O(\epsilon|\tilde{\Sigma}_{11}|) \end{pmatrix} \tag{A.7}$$

and $\mathbf{H}$ as

$$\mathbf{H} = \begin{pmatrix} |\frac{\tilde{D}_{11}}{2\lambda_1}|^2 + O(\epsilon^2|\tilde{\Sigma}_{11}|^2) & O(\epsilon|\tilde{\Sigma}_{11}|^2) & \cdots & O(\epsilon|\tilde{\Sigma}_{11}|^2) \\ O(\epsilon|\tilde{\Sigma}_{11}|^2) & O(\epsilon^2|\tilde{\Sigma}_{11}|^2) & \cdots & O(\epsilon^2|\tilde{\Sigma}_{11}|^2) \\ \vdots & \vdots & \ddots & \vdots \\ O(\epsilon|\tilde{\Sigma}_{11}|^2) & O(\epsilon^2|\tilde{\Sigma}_{11}|^2) & \cdots & O(\epsilon^2|\tilde{\Sigma}_{11}|^2) \end{pmatrix}. \tag{A.8}$$

It is easy to verify that one of the eigenvalues of $\mathbf{H}$ is equal to $|\frac{\tilde{D}_{11}}{2\lambda_1}|^2 + O(\epsilon^2|\tilde{\Sigma}_{11}|^2)$ with an eigenvector of $(1 + O(\epsilon), O(\epsilon), \ldots, O(\epsilon))^\tau$. The other eigenvectors must be orthogonal to the dominant eigenvector given that $\mathbf{H}$ is a Hermitian matrix and $\mathbf{A}$ is a unitary matrix. Hence the first element of all of the other eigenvectors must be

$O(\epsilon)$. The trace of $\mathbf{H}$ is equal to $|\frac{\tilde{D}_{11}}{2\lambda_1}|^2 + O(\epsilon^2|\tilde{\Sigma}_{11}|^2)$. Also, all of the eigenvalues of $\mathbf{H}$ must be positive [126]. Hence all of the other eigenvalues of $\mathbf{H}$ are $O(\epsilon^2|\tilde{\Sigma}_{11}|^2)$. Thus we can write the unitary decomposition of $\mathbf{H}$ as

$$\begin{aligned}\mathbf{H} &= \mathbf{A}\ \mathrm{diag}(h_1^2, h_2^2, \ldots, h_N^2)\ \mathbf{A}^* \\ &= \begin{pmatrix} 1+O(\epsilon) & O(\epsilon) & \cdots & O(\epsilon) \\ O(\epsilon) & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ O(\epsilon) & A_{N2} & \cdots & A_{NN} \end{pmatrix} \mathrm{diag}\left(\left|\frac{\tilde{D}_{11}}{2\lambda_1}\right|^2 + O(\epsilon^2|\tilde{\Sigma}_{11}|^2), O(\epsilon^2|\tilde{\Sigma}_{11}|^2), \ldots, O(\epsilon^2|\tilde{\Sigma}_{11}|^2) \right) \times \\ &\quad \begin{pmatrix} 1+O(\epsilon) & O(\epsilon) & \cdots & O(\epsilon) \\ O(\epsilon) & \bar{A}_{22} & \cdots & \bar{A}_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ O(\epsilon) & \bar{A}_{2N} & \cdots & \bar{A}_{NN} \end{pmatrix}. \end{aligned} \tag{A.9}$$

The Takagi factorization [126] of $\tilde{\mathbf{\Sigma}}$ can be expressed as

$$\tilde{\mathbf{\Sigma}} = \mathbf{A}\mathbf{P}\ \mathrm{diag}(h_1, h_2, \ldots, h_N)\ \mathbf{P}^\tau \mathbf{A}^\tau \tag{A.10}$$

where $\mathbf{P}$ is diagonal matrix of phase factors that satisfies

$$\mathbf{P}^2 = \bar{\mathbf{A}}^\tau \tilde{\mathbf{\Sigma}} \bar{\mathbf{A}}\ \mathrm{diag}(1/h_1, 1/h_2, \ldots, 1/h_N). \tag{A.11}$$

Using (A.8), (A.9), and (A.11), we find that P_{11} will be $1+O(\epsilon)$ and the other diagonal

terms will be $O(1)$ as well. Thus the Takagi factorization of $\tilde{\mathbf{\Sigma}}$ is

$$\tilde{\mathbf{\Sigma}} = \begin{pmatrix} 1+O(\epsilon) & O(\epsilon) & \cdots & O(\epsilon) \\ O(\epsilon) & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ O(\epsilon) & A_{N2} & \cdots & A_{NN} \end{pmatrix} \operatorname{diag}\left(\left| \frac{\tilde{D}_{11}}{2\lambda_1} \right| + O(\epsilon|\tilde{\Sigma}_{11}|), O(\epsilon|\tilde{\Sigma}_{11}|), \ldots, O(\epsilon|\tilde{\Sigma}_{11}|) \right) \times \begin{pmatrix} 1+O(\epsilon) & O(\epsilon) & \cdots & O(\epsilon) \\ O(\epsilon) & A_{22} & \cdots & A_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ O(\epsilon) & A_{2N} & \cdots & A_{NN} \end{pmatrix}. \tag{A.12}$$

If we plug this factorization of $\tilde{\mathbf{\Sigma}}$ into $\mathbf{\Sigma} = \mathbf{T}\tilde{\mathbf{\Sigma}}\mathbf{T}^\tau$, we can write

$$\mathbf{\Sigma} = \left| \frac{\tilde{D}_{11}}{2\lambda_1} \right| \mathbf{T}_1 \mathbf{T}_1^\tau + \begin{pmatrix} O(\epsilon|\tilde{\Sigma}_{11}|) & \cdots & O(\epsilon|\tilde{\Sigma}_{11}|) \\ \vdots & \ddots & \vdots \\ O(\epsilon|\tilde{\Sigma}_{11}|) & \cdots & O(\epsilon|\tilde{\Sigma}_{11}|) \end{pmatrix}. \tag{A.13}$$

We scale our eigenvectors to have unit norms. Thus $\mathbf{T}_1^\tau \mathbf{T}_1 = 1$ and it is clear that $\mathbf{T}_1 + (O(\epsilon|\tilde{\Sigma}_{11}|), \ldots, O(\epsilon|\tilde{\Sigma}_{11}|))^\tau$ is an eigenvector of $\mathbf{\Sigma}$ with the corresponding eigenvalue of $|\tilde{D}_{11}/(2\lambda_1)| + O(\epsilon|\tilde{\Sigma}_{11}|)$. Because $\mathbf{\Sigma}$ is a covariance matrix all of its eigenvalues are non-negative and they sum to the trace of (A.13), which is $|\tilde{D}_{11}/(2\lambda_1)| + O(\epsilon|\tilde{\Sigma}_{11}|)$. Hence all of the other eigenvalues must be $O(\epsilon|\tilde{\Sigma}_{11}|)$. □

Lemma 3. Suppose that the dominant eigenvalues of $\mathbf{F}$ form a complex conjugate pair. Further suppose that $|\tilde{D}_{12}/(2\mathrm{Re}(\lambda_1))|\epsilon \geq |\max(\tilde{\mathbf{D}})/\lambda_3|$, with $0 < \epsilon \ll 1$. Then the sum of the two largest eigenvalues of $\mathbf{\Sigma}$ is $O(|\tilde{D}_{12}/(2\mathrm{Re}(\lambda_1))| + |\tilde{D}_{11}/(2\lambda_1)|)$ and all of the other eigenvalues are $O(\epsilon|\tilde{\Sigma}_{12}|)$.

Proof. Our proof uses a similar approach to the case where λ_1 is real. Using (A.1)

and letting $\mathbf{H} = \tilde{\mathbf{\Sigma}}\tilde{\mathbf{\Sigma}}^*$, we can write $\tilde{\mathbf{\Sigma}}$ as

$$\tilde{\mathbf{\Sigma}} = \begin{pmatrix} -\frac{\tilde{D}_{11}}{2\lambda_1} & -\frac{\tilde{D}_{12}}{2\mathrm{Re}(\lambda_1)} & O(\epsilon|\tilde{\Sigma}_{12}|) & \cdots & O(\epsilon|\tilde{\Sigma}_{12}|) \\ -\frac{\tilde{D}_{12}}{2\mathrm{Re}(\lambda_1)} & -\frac{\tilde{D}_{11}}{2\lambda_1} & O(\epsilon|\tilde{\Sigma}_{12}|) & \cdots & O(\epsilon|\tilde{\Sigma}_{12}|) \\ O(\epsilon|\tilde{\Sigma}_{12}|) & O(\epsilon|\tilde{\Sigma}_{12}|) & O(\epsilon|\tilde{\Sigma}_{12}|) & \cdots & O(\epsilon|\tilde{\Sigma}_{12}|) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ O(\epsilon|\tilde{\Sigma}_{12}|) & O(\epsilon|\tilde{\Sigma}_{12}|) & O(\epsilon|\tilde{\Sigma}_{12}|) & \cdots & O(\epsilon|\tilde{\Sigma}_{12}|) \end{pmatrix} \tag{A.14}$$

and $\mathbf{H}$ as

$$\begin{pmatrix} |\frac{\tilde{D}_{11}}{2\lambda_1}|^2 + |\frac{\tilde{D}_{12}}{2\mathrm{Re}(\lambda_1)}|^2 + O(\epsilon^2|\tilde{\Sigma}_{12}|^2) & \frac{\tilde{D}_{11}\tilde{D}_{12}}{4\lambda_1\mathrm{Re}(\lambda_1)} + \frac{\overline{\tilde{D}_{11}}}{\lambda_1}\frac{\tilde{D}_{12}}{4\mathrm{Re}(\lambda_1)} + O(\epsilon^2|\tilde{\Sigma}_{12}|^2) & O(\epsilon|\tilde{\Sigma}_{12}|^2) & \cdots & O(\epsilon|\tilde{\Sigma}_{12}|^2) \\ \frac{\tilde{D}_{11}\tilde{D}_{12}}{4\lambda_1\mathrm{Re}(\lambda_1)} + \frac{\overline{\tilde{D}_{11}}}{\lambda_1}\frac{\tilde{D}_{12}}{4\mathrm{Re}(\lambda_1)} + O(\epsilon^2|\tilde{\Sigma}_{12}|^2) & |\frac{\tilde{D}_{11}}{2\lambda_1}|^2 + |\frac{\tilde{D}_{12}}{2\mathrm{Re}(\lambda_1)}|^2 + O(\epsilon^2|\tilde{\Sigma}_{12}|^2) & O(\epsilon|\tilde{\Sigma}_{12}|^2) & \cdots & O(\epsilon|\tilde{\Sigma}_{12}|^2) \\ O(\epsilon|\tilde{\Sigma}_{12}|^2) & O(\epsilon|\tilde{\Sigma}_{12}|^2) & O(\epsilon^2|\tilde{\Sigma}_{12}|^2) & \cdots & O(\epsilon^2|\tilde{\Sigma}_{12}|^2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ O(\epsilon|\tilde{\Sigma}_{12}|^2) & O(\epsilon|\tilde{\Sigma}_{12}|^2) & O(\epsilon^2|\tilde{\Sigma}_{12}|^2) & \cdots & O(\epsilon^2|\tilde{\Sigma}_{12}|^2) \end{pmatrix} \tag{A.15}$$

It is straightforward to verify that two eigenvalues for $\mathbf{H}$ are

$$h_1^2 = \left(\frac{|\tilde{D}_{11}|}{2|\lambda_1|} + \frac{|\tilde{D}_{12}|}{2\ \mathrm{Re}(\lambda_1)}\right)^2 + O(\epsilon^2|\tilde{\Sigma}_{12}|^2), \tag{A.16}$$

$$h_2^2 = \left(\frac{|\tilde{D}_{11}|}{2|\lambda_1|} - \frac{|\tilde{D}_{12}|}{2\ \mathrm{Re}(\lambda_1)}\right)^2 + O(\epsilon^2|\tilde{\Sigma}_{12}|^2), \tag{A.17}$$

with the associated eigenvectors

$$\begin{pmatrix} \frac{\tilde{D}_{11}\tilde{D}_{12}}{2\lambda_1\ \mathrm{Re}(\lambda_1)n} + O(\epsilon) \\ \pm\frac{|\tilde{D}_{11}|\tilde{D}_{12}}{2|\lambda_1|\ \mathrm{Re}(\lambda_1)n} + O(\epsilon) \\ O(\epsilon) \\ \vdots \\ O(\epsilon) \end{pmatrix}, \tag{A.18}$$

where the normalizing constant n ensures that the eigenvectors have unit norms. Similar to what we have done before, we reason from the trace of $\mathbf{H}$ that all of the

other eigenvalues $h_i, i > 2$ must be $O(\epsilon^2|\tilde{\Sigma}_{12}|^2)$. The requirement that the eigenvectors of $\mathbf{H}$ are orthogonal tells us that the first two elements of the remaining eigenvectors are $O(\epsilon)$.

Now let $\mathbf{M} = \mathbf{TAP}\ \mathrm{diag}(\sqrt{h_1}, \sqrt{h_2}, \cdots, \sqrt{h_N})$ where $\mathbf{A}$ is the unitary matrix of eigenvectors of $\mathbf{H}$ and $\mathbf{P}$ is the diagonal matrix of phase factors from the Takagi factorization (A.10). Note that $\mathbf{\Sigma} = \mathbf{M}\mathbf{M}^\tau$. To obtain a simple equation for the eigenvalues of $\mathbf{\Sigma}$, we next decompose $\mathbf{M}$ into the product of an orthonormal matrix $\mathbf{V}$ and an upper triangular matrix $\mathbf{U}$. This can be done by applying the Gram-Shmidt process. The resulting first column of $\mathbf{V}$ is simply $\mathbf{M}_1/|\mathbf{M}_1|$ where $\mathbf{M}_1$ is the first column of $\mathbf{M}$. Let $\tilde{\mathbf{V}}_2 = \mathbf{M}_2 - \frac{\mathbf{M}_1^\tau \mathbf{M}_2}{\mathbf{M}_1^\tau \mathbf{M}_1}\mathbf{M}_1$. The second column of $\mathbf{V}$ is $\tilde{\mathbf{V}}_2/|\tilde{\mathbf{V}}_2|$. The upper left block of $\mathbf{U}$ is thus

$$\begin{pmatrix} |\mathbf{M}_1| & |\mathbf{M}_1|\frac{\mathbf{M}_2^\tau \mathbf{M}_1}{\mathbf{M}_1^\tau \mathbf{M}_1} \\ 0 & |\tilde{\mathbf{V}}_2| \end{pmatrix} \tag{A.19}$$

The rest of the upper triangle of $\mathbf{U}$ consists entirely of $O(\epsilon^{0.5}|\tilde{\Sigma}_{12}|^{0.5})$ terms. Now, since columns of $\mathbf{V}$ form an orthonormal basis, the eigenvalues of $\mathbf{\Sigma} = \mathbf{V}\mathbf{U}\mathbf{U}^\tau\mathbf{V}^\tau$ are identical to those of $\mathbf{U}\mathbf{U}^\tau$. Since all elements of $\mathbf{U}\mathbf{U}^\tau$ except the upper left 2×2 block are $O(\epsilon|\tilde{\Sigma}_{12}|)$, we can approximate to within $O(\epsilon|\tilde{\Sigma}_{12}|)$ two eigenvalues of $\mathbf{\Sigma}$ as those of the upper left 2×2 block. The trace of the upper left block of $\mathbf{U}\mathbf{U}^\tau$ turns out to be equal to

$$h_1+h_2+(h_1-h_2)\frac{\mathrm{Re}(\bar{\lambda}_1\tilde{D}_{11})}{|\tilde{D}_{11}||\lambda_1|}\ \mathrm{Re}(T_1\cdot\bar{T}_2)+(h_2-h_1)\frac{\mathrm{Im}(\bar{\lambda}_1\tilde{D}_{11})}{|\tilde{D}_{11}||\lambda_1|}\ \mathrm{Im}(T_1\cdot\bar{T}_2)+O(\epsilon|\tilde{\Sigma}_{12}|), \tag{A.20}$$

which is $O(|\tilde{D}_{12}/(2\mathrm{Re}(\lambda_1))|+|\tilde{D}_{11}/(2\lambda_1)|)$ due to the h_2 term (the positive square root of (A.17)). The diagonal of the full matrix $\mathbf{U}\mathbf{U}^\tau$ has only $O(\epsilon|\tilde{\Sigma}_{12}|)$ terms outside of the upper left 2×2 block. Since all of the eigenvalues of $\mathbf{\Sigma}$ must be non-negative and sum to the trace of $\mathbf{U}\mathbf{U}^\tau$, all of the other eigenvalues of $\mathbf{\Sigma}$ must be $O(\epsilon|\tilde{\Sigma}_{12}|)$. □

Theorem 4. Suppose that the magnitude of the real part of a dominant eigenvalue of $\mathbf{F}$ is small such that the assumptions of either Lemmas 2 or 3 are satisfied. Then as this real part approaches zero, if the assumptions of Lemma 2 are satisfied, then the largest eigenvalue of $\mathbf{\Sigma}$ becomes larger absolutely and larger relative to all of the other eigenvalues of $\mathbf{\Sigma}$. Alternatively, if the assumptions of Lemma 3 are satisfied, the sum of the largest two eigenvalues of $\mathbf{\Sigma}$ becomes larger absolutely and relative to all of the other eigenvalues of $\mathbf{\Sigma}$.

Theorem 4 follows straightforwardly from Lemmas 2 and 3.

APPENDIX B

Practical guide of using Kendall's tau as an indicator for critical transitions

B.1 Introduction

Critical transitions have been observed in numerous systems [1, 16, 5, 36, 8, 47]. The gradual change in some underlining conditions can bring the system from one stable state to an alternative state. The prediction of such critical transitions faces significant challenges because changes in the equilibrium state of the system are generally small prior to the transitions. To overcome such challenges, researchers have suggested the use of generic early warning signals [21, 15, 18, 63]. Such early warning indicators are based on the observation that the recovery from perturbations will become slower as the system approaches the critical transition. This phenomenon is referred to as "critical slowing down" [19]. It follows that certain statistics of the system such as variance, autocorrelation will change prior to such critical transitions.

Kendall's τ coefficient is often used to study the trend of statistics related to the critical slowing down phenomenon [21]. It is a measure of the correlation between the rank order of the observed values and their order in time [127]. A large Kendall's τ

typically means monotonic increase in the data. Because the sequence of the statistics are estimated from the time series, the value of Kendall's τ are affected by parameters such as window size, sample rate and etc. In this study, we are going to examine the effects of different parameters on the distribution of the trend statistic Kendall's τ.

A number of parametric and non-parametric methods have been proposed to understand the significance of the Kendall's τ value obtained from the time series [21, 44]. Parametric tests can be more powerful, but require more information about the system. On the other hand, non-parametric trend tests, such as Mann-Kendall test, require only that the data be independent and can tolerate outliers in the data [128]. In this study, we compared these two types of tests, and discussed their benefits and drawbacks.

B.2 Parametric study

To use Kendall's τ to detect critical slowing down, target statistics such as variance, autocorrelation are first calculated using a moving window. This sequence of statistics are then used to calculate Kendall's τ, which will be further used to make a decision about the system.

For a sequence of independent and randomly ordered data, i.e. there is no trend or serial correlation structure among the observations, the trend statistic Kendall's τ should tend to normal distribution for large n (number of observations in the sequence), with mean 0 and variance:

$$var(\tau) = \frac{2(2n+5)}{9n(n-1)}, \tag{B.1}$$

where n is the number of observations in the sequence.

Generic early warning signals, such as the Kendall's τ of autocorrelation, however, are calculated using a sequence of statistics that are obtained from the time series

Parameter	Value
r	1
K	10
σ	0.01

Table B.1: Parameter values used in the simulation

using a moving window. Positive correlation among the observations will increase the chance of obtaining a large trend statistic Kendall's τ, even in the absence of a trend [128]. Therefore, the choice of parameters, such as window size, sample rate, will affect the correlation in data, and thus affect the distribution of Kendall's τ.

Next we study how different parameters affect the distribution of Kendall's τ of systems either facing or not facing critical transition. Simulation data is obtained from the harvesting model:

$$dx = (rx(1 - \frac{x}{K}) - c\frac{x^2}{x^2 + 1})dt + \sigma dW, \tag{B.2}$$

where x is the amount of biomass, K is the carrying capacity, r is the maximum growth rate, c is the maximum grazing rate, and σ is the standard deviation of the white noise dW. Values of the parameters (except bifurcation parameter c) can be found in Table B.1

B.2.1 Window size

The choice of window size has a large influence on the distribution of Kendall's τ. This is because positive serial correlation exists when two consecutive moving windows have an overlap. This correlation is even stronger as the size of the moving window increases.

To show this, we collected 400 time series from the harvesting model Eq. B.2 with a fixed parameter $c = -1.1$, and calculated the Kendall's τ using a different window size for each time series. Fig. B.1 shows the relationship between the Kendall's τ

calculated using a smaller window size and a large window size. It is obvious that as the difference between two window sizes increases, the curve takes on an S shape. This means a large window size will inflate the value of the Kendall's τ calculated from the same time series data.

The inflation of Kendall's τ under large window size can also be observed using the distribution plot as shown in Fig. B.2. As window size increases, the distribution of Kendall's τ becomes flatter, and farther away from the normal distribution with variance calculated using Eq. B.1.

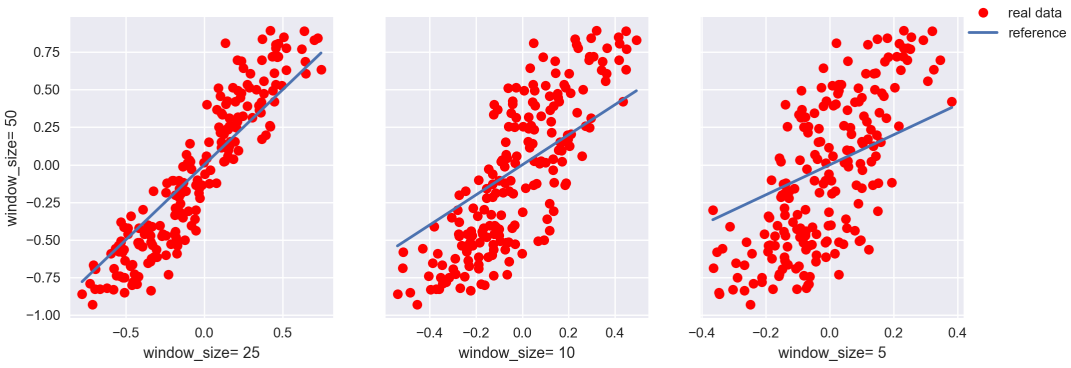

Figure B.1: The comparison between the Kendall's τ calculated using a smaller window size and a large window size

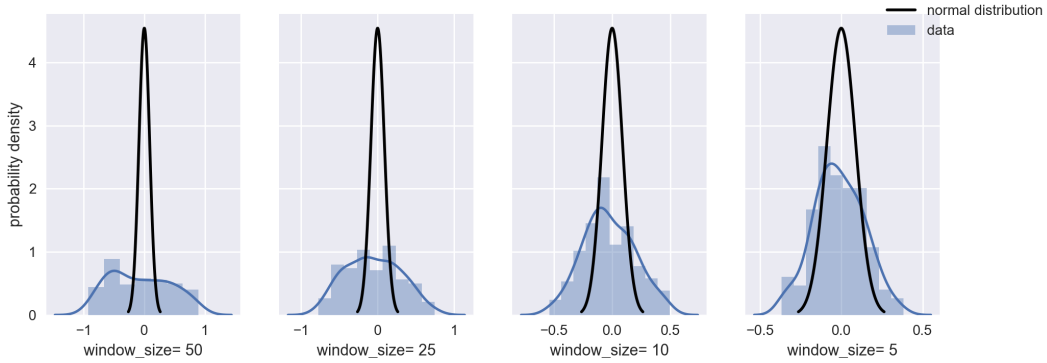

Figure B.2: Distribution of Kendall's τ calculated using different window sizes.

Therefore, a same Kendall's τ value will have a completely different meaning with a different window size. A 90 % percentile Kendall's τ value when window size is 5

% of the length of the time series is only 60 % percentile when the window size is 50 % of the length. Thus, merely looking at a Kendall's τ value is not enough to decide the probability of critical transition. A hypothesis test is necessary and we included discussions about that in Sec. 3.

B.2.2 Number of observations

The number of observations in the data is also important. This is because all the statistics will have a larger estimation error when only a limited amount of observations are available. Moreover, it is harder to detrend the time series data when only a limited amount of observations are available. Improper detrending will remove all the important low frequency information, leaving behind only high frequency random noise.

To understand how detrending can affect the result, one time series data is obtained from the harvesting model B.2. The time series data is then down sampled to obtained another time series data with a smaller number of observations. The spectrum of the time series is then calculated and show in Fig. B.3. Comparing the spectrum of time series data with or without detrending, it is obvious that detrending basically works as a low pass filter, removing low frequency components. Because of critical slowing down, we can see from Fig. B.3 that the spectrum closer to critical transition has larger low frequency components. This is known as the reddening effect. After detrending, as in Fig. B.3, the spectrum closer to critical transition still has larger low frequency components even if part of the low frequency components are removed. The reddening effect is still visible. The situation is quite different when only a small number of observations are available. When the low frequency components are removed by detrending, the reddening effect is not visible anymore. The detrending has effectively taken away the important parts of the spectrum that is crucial to the detection of critical slowing down.

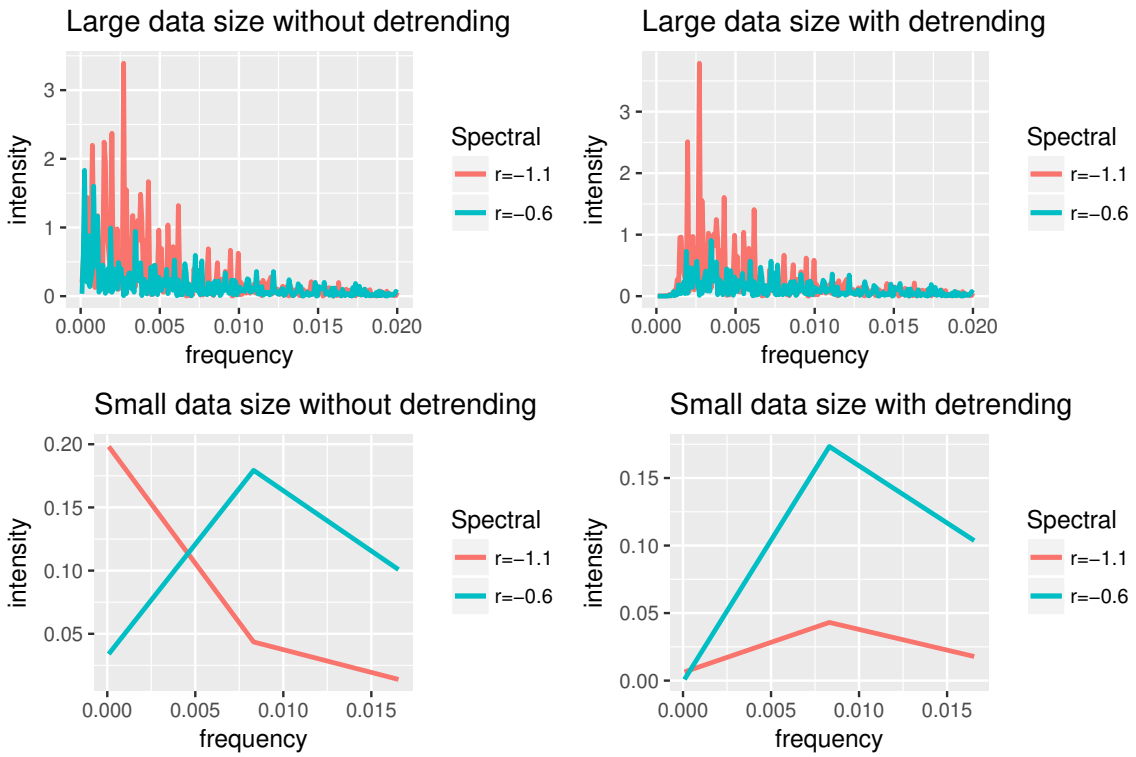

Figure B.3: Spectrum of the time series before and after detrending when either a large number of observations or a small number of observations are available.

This effect of the detrending can also be observed in the distribution of the Kendall's τ when different numbers of observations are available. We again collected 400 time series from the harvesting model, but this time with a parameter changing with time. Each time series has 20000 observations. These 400 time series are then down sampled to obtain time series with 20000, 200 and 100 observations. The distribution of the standard deviation over time and the distribution of Kendall's τ under different sample rate is shown in Fig B.4. When there are at least 2000 observations, the distribution of Kendall's τ is skewed towards left, which is accurate because the system is approaching the critical transition. However, when there are only 200 observations or fewer, the distribution becomes almost bell shape, which means the results are almost random. Therefore, it is important to have enough number of observations in the data when Kendall's τ is used as an early warning indicator, especially

when the equilibrium of the system is changing as the system approaches the critical transition and thus detrending is necessary.

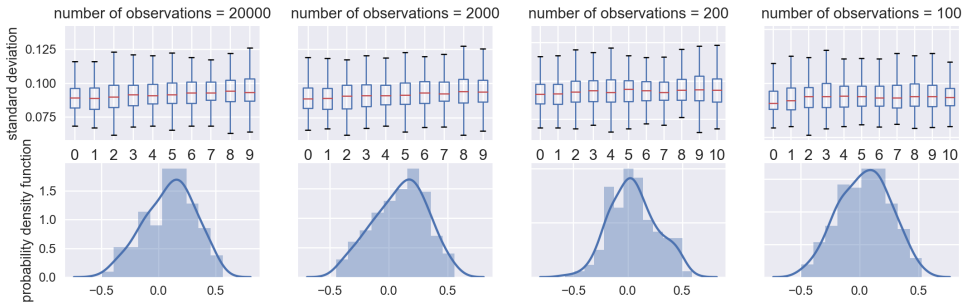

Figure B.4: The distribution of the standard deviation over time and the distribution of Kendall's τ when 20000, 2000, 200 and 100 observa

B.3 Parametric and non-parametric tests

Because the trend statistic Kendall's τ is affected by a lot of parameters, merely looking at it won't reveal much information about the system. A number of tests have been proposed to understand the significance of a certain value of the Kendall's τ. In this section, we introduced the non-parametric modified Mann-Kendall trend test and compared it to other parametric tests.

B.3.1 Non-parametric modified Mann-Kendall trend test

The null hypothesis for the traditional Mann-Kendall trend test is that there is no trend or serial correlation structure among the observations. This hypothesis is not entirely true for the series of statistic, such as standard deviation, autocorrelation and so on, that we obtained from the time series using a moving window.

[128] pointed out that we can use a modified Mann-Kendall trend test to study data with serial correlation structure. In the modified test, the null hypothesis is that there is no trend in the data, but there can be autocorrelation. If the null hypothesis

is true, the trend statistic Kendall's τ should follow a normal distribution with mean 0, and variance as

$$var(\tau) = \frac{2(2n+5)}{9n(n-1)}(1 + \frac{2}{n(n-1)(n-2)}\sum_{i=1}^{n-1}(n-i)(n-i-1)(n-i-2)\rho_S(i)) \quad \text{(B.3)}$$

where $\rho_S(i)$ is the autocorrelation of the ranks of the observation.

To better understand this, we compared the distribution of Kendall's τ to the normal distribution with variance calculated using Eq. B.1 and the modified distribution with variance calculated using Eq. B.3 in Fig. B.5. It is clear that the real distribution is much flatter than the normal distribution due to the positive correlation in data. The modified distribution, on the other hand, is much closer to the real one.

In practice, we typically only have one time series. Therefore, the real distribution is not available. In this case, we can use the time series to calculate the modified distribution of Kendall's τ, and use the distribution to calculate the percentile of the obtained Kendall's τ value.

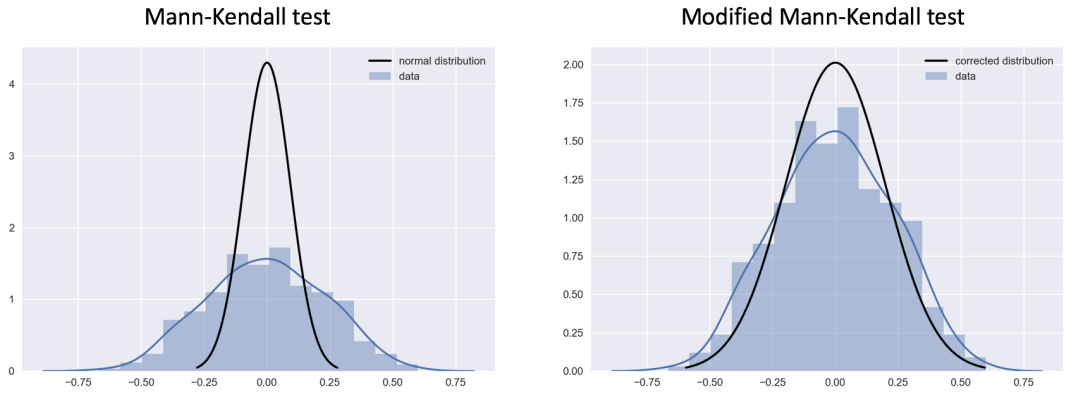

Figure B.5: (a) The comparison between the distribution of Kendall's τ in data and the Mann-Kendall test. (b) The comparison between the distribution of Kendall's τ in data and the modified Mann-Kendall test.

B.3.2 Comparison with parametric tests

Parametric tests [21, 44] have been proposed to study the significance of a Kendall's τ value. These methods use a general model to fit the data, and generate artificial data using the model to understand how significant the trend statistic value is.

[21] proposes to fit an ARMA model using the residual data after detrending. This test is designed to show that this data cannot come from a linear stationary process if a large Kendall's τ value is obtained from it. When this test gives a p-value as low as 0.1%, it doesn't mean that the probability of critical transition is as high as 99.9 %. It means that the probability that this time series data is generated using a linear stationary model is as low as 0.1 %.

[44] proposed to fit two nonlinear models that both have a normal form for the saddlenode bifurcation. The difference between these two models is that one of them has a fixed bifurcation parameter, while the other one has a changing parameter. The distributions of the test statistics are then generated from these two models and compared to determine if these two models are statistically different and why one of them better describes the data.

We compared the parametric test using ARMA model and the non-parametric Mann-Kendall test. 200 time series is first generated using the harvesting model (Eq. B.2). The distribution of Kendall's τ calculated from all the time series data is shown in Fig. B.6. Then we use one of the time series to fit an ARMA model, and generated another 200 time series using the fitted ARMA model. The distribution of Kendall's τ calculated from all the ARMA time series is also shown in Fig. B.6. Finally, we use the same time series to calculate the variance of the modified normal distribution using Eq. B.3. All three distributions are close to each other, as we can see from Fig B.6. Therefore, both the parametric ARMA test and the non-parametric modified Mann-Kendall can accurately approximate the distribution of Kendall's τ calculated from times series data of low-dimensional systems

with Gaussian noise. The benefit of the non-parametric Mann-Kendall test is that the distribution of Kendall's τ can be estimated directly from the time series, and thus no further simulation will be required. As a result it is much faster than the parametric ARMA test.

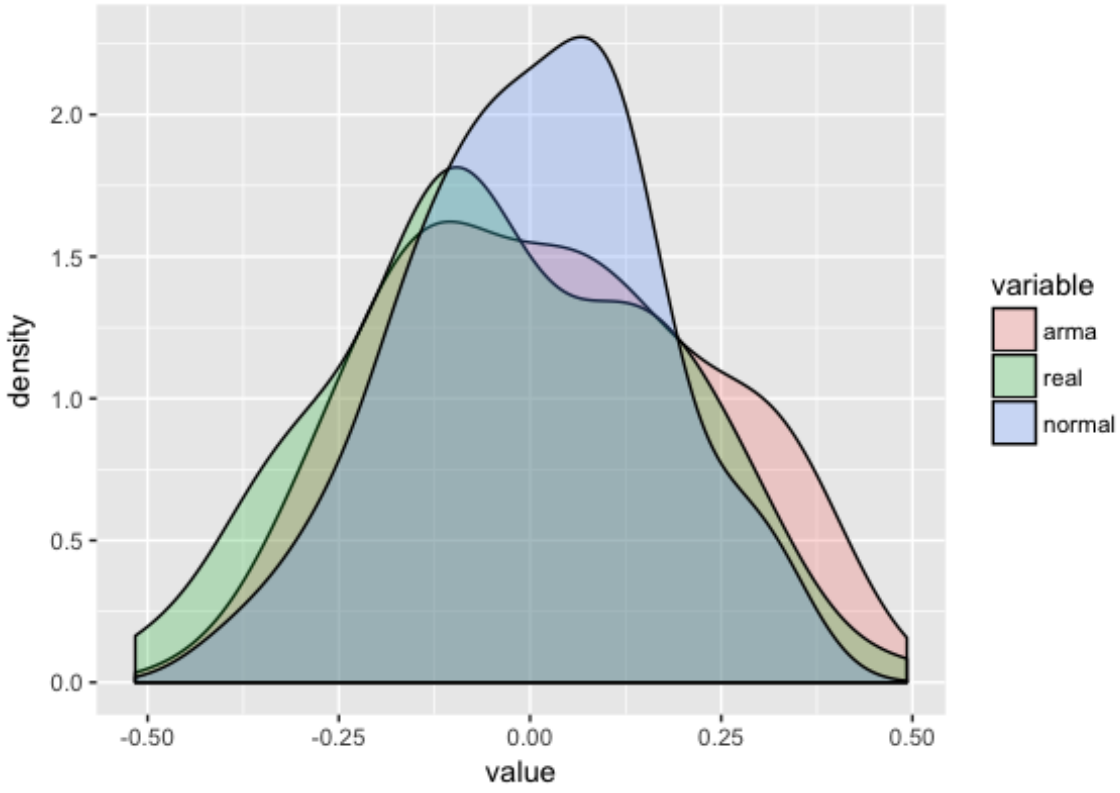

Figure B.6: The distribution of Kendall's τ calculated using the harvesting model, the fitted ARMA model, the modified Mann-Kendall test.

B.4 Conclusions

In this study we examined how different parameters can affect the value of Kendall's τ obtained from the same system. Specifically,

- A large window size can inflate the value of Kendall's τ compared to a small window size,

- Detrending can remove the redenning effect when only a small number of observations are available.

Therefore, it is highly advised that researchers keep the effect of these parameters in mind when they use the Kendall's τ of certain statistics as an early warning indicator for critical transition.

We also compared the non-parametric Mann-Kendall test and the parametric ARMA test. These two tests yield similar results for a low-dimensional system with Gaussian stochastic excitation. The benefit of the non-parametric Mann-Kendall test is that the distribution of Kendall's τ can be estimated directly from the time series, and thus no further simulation will be required. As a result it is much faster than the parametric ARMA test.

BIBLIOGRAPHY

[1] M. Scheffer, *Critical transitions in nature and society*. Princeton University Press, 2009.

[2] J. Miles, "Parametric excitation of an internally resonant double pendulum," *Zeitschrift für angewandte Mathematik und Physik ZAMP*, vol. 36, no. 3, pp. 337–345, 1985.

[3] D. J. Earn, P. Rohani, B. M. Bolker, and B. T. Grenfell, "A simple model for complex dynamical transitions in epidemics," *Science*, vol. 287, no. 5453, pp. 667–670, 2000.

[4] M. J. Keeling, P. Rohani, and B. T. Grenfell, "Seasonally forced disease dynamics explored as switching between attractors," *Physica D: Nonlinear Phenomena*, vol. 148, no. 3-4, pp. 317–335, 2001.

[5] J. M. Drake and B. D. Griffen, "Early warning signals of extinction in deteriorating environments," *Nature*, vol. 467, no. 7314, pp. 456–459, 2010.

[6] T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, "Tipping elements in the earth's climate system," *Proceedings of the National Academy of Sciences*, vol. 105, no. 6, p. 1786, 2008.

[7] A. Ghadami and B. I. Epureanu, "Bifurcation forecasting for large dimensional oscillatory systems: forecasting flutter using gust responses," *Journal of Computational and Nonlinear Dynamics*, vol. 11, no. 6, p. 061009, 2016.

[8] S. Chen and B. Epureanu, "Forecasting bifurcations in parametrically excited systems," *Nonlinear Dynamics*, vol. 91, no. 1, pp. 443–457, 2018.

[9] M. Scheffer, S. Carpenter, J. A. Foley, C. Folke, and B. Walker, "Catastrophic shifts in ecosystems," *Nature*, vol. 413, no. 6856, pp. 591–596, 2001.

[10] R. M. May, S. A. Levin, and G. Sugihara, "Complex systems: Ecology for bankers," *Nature*, vol. 451, no. 7181, p. 893, 2008.

[11] J. G. Venegas, T. Winkler, G. Musch, M. F. V. Melo, D. Layfield, N. Tgavalekos, A. J. Fischman, R. J. Callahan, G. Bellani, and R. S. Harris, "Self-organized patchiness in asthma as a prelude to catastrophic shifts," *Nature*, vol. 434, no. 7034, p. 777, 2005.

[12] I. A. van de Leemput, M. Wichers, A. O. Cramer, D. Borsboom, F. Tuerlinckx, P. Kuppens, E. H. van Nes, W. Viechtbauer, E. J. Giltay, S. H. Aggen, *et al.*, "Critical slowing down as early warning for the onset and termination of depression," *Proceedings of the National Academy of Sciences*, vol. 111, no. 1, pp. 87–92, 2014.

[13] G. Bulian, A. Francescutto, and C. Lugni, "On the nonlinear modeling of parametric rolling in regular and irregular waves," *International shipbuilding progress*, vol. 51, no. 2, 3, pp. 173–203, 2004.

[14] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, vol. 42. Springer Science & Business Media, 2013.

[15] C. Kuehn, "A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics," *Physica D: Nonlinear Phenomena*, vol. 240, no. 12, pp. 1020–1035, 2011.

[16] M. Scheffer, S. R. Carpenter, T. M. Lenton, J. Bascompte, W. Brock, V. Dakos, j. Van De Koppel, I. A. V. D. Leemput, S. A. Levin, E. H. V. Nes, *et al.*, "Anticipating critical transitions," *Science*, vol. 338, no. 6105, p. 344, 2012.

[17] S. M. ORegan and J. M. Drake, "Theory of early warning signals of disease emergence and leading indicators of elimination," *Theoretical Ecology*, vol. 6, no. 3, pp. 333–357, 2013.

[18] S. Kéfi, V. Dakos, M. Scheffer, E. H. Van Nes, and M. Rietkerk, "Early warning signals also precede non-catastrophic transitions," *Oikos*, vol. 122, no. 5, pp. 641–648, 2013.

[19] S. H. Strogatz, *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering*. Westview press, 2014.

[20] A. H. Nayfeh and D. T. Mook, *Nonlinear oscillations*. John Wiley & Sons, 2008.

[21] V. Dakos, S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, S. Kefi, V. Livina, D. A. Seekell, E. H. van Nes, *et al.*, "Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data," *PloS One*, vol. 7, no. 7, p. e41010, 2012.

[22] V. Dakos, E. H. van Nes, R. Donangelo, H. Fort, and M. Scheffer, "Spatial correlation as leading indicator of catastrophic shifts," *Theoretical Ecology*, vol. 3, no. 3, pp. 163–174, 2010.

[23] L. Dai, K. S. Korolev, and J. Gore, "Slower recovery in space before collapse of connected populations," *Nature*, vol. 496, no. 7445, p. 355, 2013.

[24] V. Dakos, E. H. Van Nes, P. D'Odorico, and M. Scheffer, "Robustness of variance and autocorrelation as indicators of critical slowing down," *Ecology*, vol. 93, no. 2, pp. 264–271, 2012.

[25] S. Chen and B. Epureanu, "Regular biennial cycles in epidemics caused by

parametric resonance," *Journal of theoretical biology*, vol. 415, pp. 137–144, 2017.

[26] N. C. Perkins, "Modal interactions in the non-linear response of elastic cables under parametric/external excitation," *International Journal of Non-Linear Mechanics*, vol. 27, no. 2, pp. 233–250, 1992.

[27] D. Alonso, A. J. McKane, and M. Pascual, "Stochastic amplification in epidemics," *Journal of the Royal Society Interface*, vol. 4, no. 14, pp. 575–582, 2007.

[28] P. B. Miller, E. B. ODea, P. Rohani, and J. M. Drake, "Forecasting infectious disease emergence subject to seasonal forcing," *Theoretical Biology and Medical Modelling*, vol. 14, no. 1, p. 17, 2017.

[29] E. H. van Nes and M. Scheffer, "Implications of spatial heterogeneity for catastrophic regime shifts in ecosystems," *Ecology*, vol. 86, no. 7, pp. 1797–1807, 2005.

[30] M. Scheffer and S. R. Carpenter, "Catastrophic regime shifts in ecosystems: linking theory to observation," *Trends in ecology & evolution*, vol. 18, no. 12, pp. 648–656, 2003.

[31] V. Dakos, S. Kéfi, M. Rietkerk, E. H. Van Nes, and M. Scheffer, "Slowing down in spatially patterned ecosystems at the brink of collapse," *The American Naturalist*, vol. 177, no. 6, pp. E153–E166, 2011.

[32] V. Guttal and C. Jayaprakash, "Spatial variance and spatial skewness: leading indicators of regime shifts in spatial ecological systems," *Theoretical Ecology*, vol. 2, no. 1, pp. 3–12, 2009.

[33] M. Rietkerk, S. C. Dekker, P. C. de Ruiter, and J. van de Koppel, "Self-organized patchiness and catastrophic shifts in ecosystems," *Science*, vol. 305, no. 5692, pp. 1926–1929, 2004.

[34] D. Tilman, J. Fargione, B. Wolff, C. D'Antonio, A. Dobson, R. Howarth, D. Schindler, W. H. Schlesinger, D. Simberloff, and D. Swackhamer, "Forecasting agriculturally driven global environmental change," *Science*, vol. 292, no. 5515, pp. 281–284, 2001.

[35] P. M. Vitousek, H. A. Mooney, J. Lubchenco, and J. M. Melillo, "Human domination of earth's ecosystems," *Science*, vol. 277, no. 5325, pp. 494–499, 1997.

[36] A. Ghadami, C. E. Cesnik, and B. I. Epureanu, "Model-less forecasting of hopf bifurcations in fluid-structural systems," *Journal of Fluids and Structures*, vol. 76, pp. 1–13, 2018.

[37] K. DSouza, B. I. Epureanu, and M. Pascual, "Forecasting bifurcations from large perturbation recoveries in feedback ecosystems," *PloS One*, vol. 10, no. 9, p. e0137779, 2015.

[38] L. Chen, R. Liu, Z.-P. Liu, M. Li, and K. Aihara, "Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers," *Scientific Reports*, vol. 2, 2012.

[39] T. Squartini, I. Van Lelyveld, and D. Garlaschelli, "Early-warning signals of topological collapse in interbank networks," *Scientific Reports*, vol. 3, 2013.

[40] C. Wissel, "A universal law of the characteristic return time near thresholds," *Oecologia*, vol. 65, no. 1, pp. 101–107, 1984.

[41] E. H. Van Nes and M. Scheffer, "Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift," *The American Naturalist*, vol. 169, no. 6, pp. 738–747, 2007.

[42] S. Kéfi, V. Guttal, W. A. Brock, S. R. Carpenter, A. M. Ellison, V. N. Livina, D. A. Seekell, M. Scheffer, E. H. van Nes, and V. Dakos, "Early warning signals of ecological transitions: Methods for spatial patterns," *PloS One*, vol. 9, no. 3, p. e92097, 2014.

[43] A. R. Ives, "Measuring resilience in stochastic systems," *Ecological Monographs*, vol. 65, no. 2, pp. 217–233, 1995.

[44] C. Boettiger and A. Hastings, "Quantifying limits to detection of early warning for critical transitions," *Journal of The Royal Society Interface*, vol. 9, no. 75, p. 2527, 2012.

[45] J. Lim and B. I. Epureanu, "Forecasting a class of bifurcations: Theory and experiment," *Physical Review E*, vol. 83, no. 1, p. 016203, 2011.

[46] A. Ghadami and B. I. Epureanu, "Forecasting the post-bifurcation dynamics of large-dimensional slow-oscillatory systems using critical slowing down and center space reduction," *Nonlinear Dynamics*, pp. 1–17.

[47] V. Dakos, "Identifying best-indicator species for abrupt transitions in multi-species communities," *Ecological Indicators*, 2017.

[48] S. Chen and B. Epureanu, "Forecasting bifurcations of multi-degree-of-freedom nonlinear systems with parametric resonance," *Nonlinear Dynamics*, pp. 1–16, 2017.

[49] N. C. Perkins, "Modal interactions in the non-linear response of elastic cables

under parametric/external excitation," *International Journal of Non-Linear Mechanics*, vol. 27, no. 2, pp. 233–250, 1992.

[50] G. Blankenship and A. Kahraman, "Steady state forced response of a mechanical oscillator with combined parametric excitation and clearance type nonlinearity," *Journal of Sound and Vibration*, vol. 185, no. 5, pp. 743–765, 1995.

[51] J. Bobin, M. Decroisette, B. Meyer, and Y. Vitel, "Harmonic generation and parametric excitation of waves in a laser-created plasma," *Physical Review Letters*, vol. 30, no. 13, p. 594, 1973.

[52] L. Panda and R. Kar, "Nonlinear dynamics of a pipe conveying pulsating fluid with parametric and internal resonances," *Nonlinear Dynamics*, vol. 49, no. 1, pp. 9–30, 2007.

[53] M. Belhaq and M. Houssni, "Quasi-periodic oscillations, chaos and suppression of chaos in a nonlinear oscillator driven by parametric and external excitations," *Nonlinear Dynamics*, vol. 18, no. 1, pp. 1–24, 1999.

[54] A. Raghothama and S. Narayanan, "Periodic response and chaos in nonlinear systems with parametric excitation and time delay," *Nonlinear dynamics*, vol. 27, no. 4, pp. 341–365, 2002.

[55] A. Taylor, J. A. Sherratt, and A. White, "Seasonal forcing and multi-year cycles in interacting populations: lessons from a predator–prey model," *Journal of Mathematical Biology*, vol. 67, no. 6-7, p. 1741, 2013.

[56] L. Bolzoni, A. P. Dobson, M. Gatto, and G. A. D. Leo, "Allometric scaling and seasonality in the epidemics of wildlife diseases," *The American Naturalist*, vol. 172, no. 6, p. 818, 2008.

[57] M. J. Keeling and P. Rohani, *Modeling infectious diseases in humans and animals*. Princeton University Press, 2008.

[58] S. Lei, W. Zhang, J. Lin, Q., D. Kennedy, and F. W. Williams, "Frequency domain response of a parametrically excited riser under random wave forces," *Journal of Sound and Vibration*, vol. 333, no. 2, p. 485, 2014.

[59] A. L. Oropeza-Ramos and L. K. Turner, "Parametric resonance amplification in a memgyroscope," in *Sensors, 2005 IEEE*, p. 4, IEEE, 2005.

[60] Y. Jia, J. Yan, K. Soga, and A. A. Seshia, "A parametrically excited vibration energy harvester," *Journal of Intelligent Material Systems and Structures*, p. 1045389X13491637, 2013.

[61] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential equations, dynamical systems, and an introduction to chaos*. Academic press, 2012.

[62] P. Sundararajan and S. T. Noahn, "Dynamics of forced nonlinear systems using shooting/arc-length continuation methodapplication to rotor systems," *Journal of Vibration and Acoustics*, vol. 119.1, p. 9, 1997.

[63] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara, "Early-warning signals for critical transitions," *Nature*, vol. 461, no. 7260, pp. 53–59, 2009.

[64] J. M. T. Thompson and J. Sieber, "Predicting climate tipping as a noisy bifurcation: a review," *International Journal of Bifurcation and Chaos*, vol. 21, no. 02, pp. 399–423, 2011.

[65] V. Dakos, M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held, "Slowing down as an early warning signal for abrupt climate change," *Proceedings of the National Academy of Sciences*, vol. 105, no. 38, p. 14308, 2008.

[66] D. Spina, C. Valente, and G. Tomlinson, "A new procedure for detecting nonlinearity from transient data using the gabor transform," *Nonlinear Dynamics*, vol. 11, no. 3, pp. 235–254, 1996.

[67] G. Kerschen, K. Worden, A. F. Vakakis, and J.-C. Golinval, "Past, present and future of nonlinear system identification in structural dynamics," *Mechanical systems and signal processing*, vol. 20, no. 3, pp. 505–592, 2006.

[68] J. M. T. Thompson and L. Virgin, "Predicting a jump to resonance using transient maps and beats," *International journal of non-linear mechanics*, vol. 21, no. 3, pp. 205–216, 1986.

[69] J. Lim and B. I. Epureanu, "Forecasting bifurcation morphing: application to cantilever-based sensing," *Nonlinear Dynamics*, vol. 67, no. 3, pp. 2291–2298, 2012.

[70] R. May, *Stability and complexity in model ecosystems*. Princeton University Press, 1973.

[71] T. Hallam and C. Clark, "Non-autonomous logistic equations as models of populations in a deteriorating environment," *Journal of Theoretical Biology*, vol. 93, no. 2, pp. 303–311, 1981.

[72] E. K. Akhmedov, "Parametric resonance of neutrino oscillations and passage of solar and atmospheric neutrinos through the earth," *Nuclear Physics B*, vol. 538, no. 1-2, pp. 25–51, 1999.

[73] V. Lednev, "Possible mechanism for the influence of weak magnetic fields on biological systems," *Bioelectromagnetics*, vol. 12, no. 2, pp. 71–75, 1991.

[74] Y. Shin, V. Belenky, J. Paulling, K. Weems, W. Lin, K. MCTAGGART, K. J. SPYROU, T. W. TREAKLE, M. LEVADOU, B. L. HUTCHISON, *et al.*, "Cri-

teria for parametric roll of large containerships in longitudinal seas. discussion," *Transactions-Society of Naval Architects and Marine Engineers*, vol. 112, pp. 14–47, 2004.

[75] F. Pellicano, A. Fregolent, A. Bertuzzi, and F. Vestroni, "Primary and parametric non-linear resonances of a power transmission belt: experimental and theoretical analysis," *Journal of Sound and Vibration*, vol. 244, no. 4, pp. 669–684, 2001.

[76] M. F. Daqaq, C. Stabler, Y. Qaroush, and T. Seuaciuc-Osorio, "Investigation of power harvesting via parametric excitations," *Journal of Intelligent Material Systems and Structures*, vol. 20, no. 5, pp. 545–557, 2009.

[77] J. M. Londoño, S. A. Neild, and J. E. Cooper, "Identification of backbone curves of nonlinear systems from resonance decay responses," *Journal of Sound and Vibration*, vol. 348, pp. 224–238, 2015.

[78] T. M. Lenton, "Early warning of climate tipping points," *Nature Climate Change*, vol. 1, no. 4, p. 201, 2011.

[79] A. Nayfeh, "The response of single degree of freedom systems with quadratic and cubic non-linearities to a subharmonic excitation," *Journal of Sound and Vibration*, vol. 89, no. 4, pp. 457–470, 1983.

[80] A. H. Nayfeh and B. Balachandran, "Modal interactions in dynamical and structural systems," *Applied Mechanics Reviews*, vol. 42, no. 11S, pp. S175–S201, 1989.

[81] S. Altizer, A. Dobson, P. Hosseini, P. Hudson, M. Pascual, and P. Rohani, "Seasonality and the dynamics of infectious diseases," *Ecology letters*, vol. 9, no. 4, pp. 467–484, 2006.

[82] R. M. Anderson, R. M. May, and B. Anderson, *Infectious diseases of humans: dynamics and control*, vol. 28. Wiley Online Library, 1992.

[83] J. Greenman, M. Kamo, and M. Boots, "External forcing of ecological and epidemiological systems: a resonance approach," *Physica D: Nonlinear Phenomena*, vol. 190, no. 1-2, pp. 136–151, 2004.

[84] P. E. Fine and J. A. Clarkson, "Measles in england and walesi: an analysis of factors underlying seasonal patterns," *International journal of epidemiology*, vol. 11, no. 1, pp. 5–14, 1982.

[85] L. Stone, R. Olinky, and A. Huppert, "Seasonal dynamics of recurrent epidemics," *Nature*, vol. 446, no. 7135, p. 533, 2007.

[86] B. Bolker and B. Grenfell, "Chaos and biological complexity in measles dynamics," *Proc. R. Soc. Lond. B*, vol. 251, no. 1330, pp. 75–81, 1993.

[87] D. Rand and H. Wilson, "Chaotic stochasticity: a ubiquitous source of unpredictability in epidemics," *Proc. R. Soc. Lond. B*, vol. 246, no. 1316, pp. 179–184, 1991.

[88] D. J. Earn, P. Rohani, and B. T. Grenfell, "Persistence, chaos and synchrony in ecology and epidemiology," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 265, no. 1390, pp. 7–10, 1998.

[89] Y. A. Kuznetsov and C. Piccardi, "Bifurcation analysis of periodic seir and sir epidemic models," *Journal of mathematical biology*, vol. 32, no. 2, pp. 109–121, 1994.

[90] J. L. Aron and I. B. Schwartz, "Seasonality and period-doubling bifurcations in an epidemic model," *Journal of theoretical biology*, vol. 110, no. 4, pp. 665–679, 1984.

[91] K. Dietz, "The incidence of infectious diseases under the influence of seasonal fluctuations," in *Mathematical models in medicine*, pp. 1–15, Springer, 1976.

[92] A. J. Black and A. J. McKane, "Stochastic amplification in an epidemic model with seasonal forcing," *Journal of Theoretical Biology*, vol. 267, no. 1, pp. 85–94, 2010.

[93] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, vol. 1. Elsevier, 1992.

[94] H. W. Hethcote and S. A. Levin, "Periodicity in epidemiological models," in *Applied mathematical ecology*, pp. 193–211, Springer, 1989.

[95] F. Halberg, G. Cornélissen, E. E. Bakken, R. B. Sothern, O. Schwartzkopff, and C. Hamburger, "Transyears: new endpoints for gerontology and geriatrics or confusing sources of variability?," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 59, no. 12, pp. 1344–1347, 2004.

[96] R. Rand, *Lecture notes on nonlinear vibrations (version 52)*. Available from: http://www.tam.cornell.edu/randdocs/nlvibe52.pdf, 2005.

[97] P. Sundararajan and S. Noah, "Dynamics of forced nonlinear systems using shooting/arc-length continuation methodapplication to rotor systems," *Journal of Vibration and Acoustics*, vol. 119, no. 1, pp. 9–20, 1997.

[98] P. Ribeiro, "Non-linear forced vibrations of thin/thick beams and plates by the finite element and shooting methods," *Computers & structures*, vol. 82, no. 17-19, pp. 1413–1423, 2004.

[99] G. Von Groll and D. J. Ewins, "The harmonic balance method with arc-length continuation in rotor/stator contact problems," *Journal of sound and vibration*, vol. 241, no. 2, pp. 223–233, 2001.

[100] D. A. Cummings, S. Iamsirithaworn, J. T. Lessler, A. McDermott, R. Prasanthong, A. Nisalak, R. G. Jarman, D. S. Burke, and R. V. Gibbons, "The impact of the demographic transition on dengue in thailand: insights from a statistical analysis and mathematical modeling," *PLoS medicine*, vol. 6, no. 9, p. e1000139, 2009.

[101] B. M. Althouse, J. Lessler, A. A. Sall, M. Diallo, K. A. Hanley, D. M. Watts, S. C. Weaver, and D. A. Cummings, "Synchrony of sylvatic dengue isolations: a multi-host, multi-vector sir model of dengue virus transmission in senegal," *PLoS neglected tropical diseases*, vol. 6, no. 11, p. e1928, 2012.

[102] H. Broutin, C. Viboud, B. Grenfell, M. Miller, and P. Rohani, "Impact of vaccination and birth rate on the epidemiology of pertussis: a comparative study in 64 countries," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 277, no. 1698, pp. 3239–3245, 2010.

[103] E. Tornatore, S. M. Buccellato, and P. Vetro, "Stability of a stochastic sir system," *Physica A: Statistical Mechanics and its Applications*, vol. 354, pp. 111–126, 2005.

[104] M. Kamo and A. Sasaki, "The effect of cross-immunity and seasonal forcing in a multi-strain epidemic model," *Physica D: Nonlinear Phenomena*, vol. 165, no. 3-4, pp. 228–241, 2002.

[105] V. A. Jansen, "The dynamics of two diffusively coupled predator–prey populations," *Theoretical Population Biology*, vol. 59, no. 2, pp. 119–131, 2001.

[106] B. Blasius, A. Huppert, and L. Stone, "Complex dynamics and phase synchronization in spatially extended ecological systems," *Nature*, vol. 399, no. 6734, p. 354, 1999.

[107] W. Brock and S. Carpenter, "Variance as a leading indicator of regime shift in ecosystem services," *Ecology and Society*, vol. 11, no. 2, 2006.

[108] P. J. Menck, J. Heitzig, N. Marwan, and J. Kurths, "How basin stability complements the linear-stability paradigm," *Nature physics*, vol. 9, no. 2, p. 89, 2013.

[109] B. C. Nolting and K. C. Abbott, "Balls, cups, and quasi-potentials: quantifying stability in stochastic systems," *Ecology*, vol. 97, no. 4, pp. 850–864, 2016.

[110] D. J. Wilkinson, "Stochastic modelling for quantitative description of heterogeneous biological systems," *Nature Reviews Genetics*, vol. 10, no. 2, p. 122, 2009.

[111] L. Gammaitoni, "Stochastic resonance and the dithering effect in threshold physical systems," *Physical Review E*, vol. 52, no. 5, p. 4691, 1995.

[112] M. G. Neubert and H. Caswell, "Alternatives to resilience for measuring the responses of ecological systems to perturbations," *Ecology*, vol. 78, no. 3, pp. 653–665, 1997.

[113] H. Risken, "Fokker-planck equation," in *The Fokker-Planck Equation*, pp. 63–95, Springer, 1996.

[114] C. Kwon, P. Ao, and D. J. Thouless, "Structure of stochastic dynamics near fixed points," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 37, pp. 13029–13033, 2005.

[115] E. Knobloch and K. Wiesenfeld, "Bifurcations in fluctuating systems: The center-manifold approach," *Journal of Statistical Physics*, vol. 33, no. 3, pp. 611–637, 1983.

[116] T. H. Keitt, M. A. Lewis, and R. D. Holt, "Allee effects, invasion pinning, and species borders," *The American Naturalist*, vol. 157, no. 2, pp. 203–216, 2001.

[117] S. R. Carpenter, D. Ludwig, and W. A. Brock, "Management of eutrophication for lakes subject to potentially irreversible change," *Ecological applications*, vol. 9, no. 3, pp. 751–771, 1999.

[118] M. Scheffer, "The story of some shallow lakes," in *Ecology of shallow lakes*, pp. 1–19, Springer, 2004.

[119] H. Sohn, J. A. Czarnecki, and C. R. Farrar, "Structural health monitoring using statistical process control," *Journal of Structural Engineering*, vol. 126, no. 11, pp. 1356–1363, 2000.

[120] J. Schäfer, K. Strimmer, *et al.*, "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics," *Statistical Applications in Genetics and Molecular Biology*, vol. 4, no. 1, p. 32, 2005.

[121] B. Efron, "Maximum likelihood and decision theory," *The Annals of Statistics*, pp. 340–356, 1982.

[122] A. Génin, S. Majumder, S. Sankaran, F. D. Schneider, A. Danet, M. Berdugo, V. Guttal, and S. Kéfi, "Spatially heterogeneous stressors can alter the performance of indicators of regime shifts," *Ecological Indicators*, 2018.

[123] F. Guichard, P. M. Halpin, G. W. Allison, J. Lubchenco, and B. A. Menge, "Mussel disturbance dynamics: signatures of oceanographic forcing from local interactions," *The American Naturalist*, vol. 161, no. 6, pp. 889–904, 2003.

[124] V. Dakos and J. Bascompte, "Critical slowing down as early warning for the onset of collapse in mutualistic communities," *Proceedings of the National Academy of Sciences*, vol. 111, no. 49, pp. 17546–17551, 2014.

[125] S. Bathiany, M. Scheffer, E. van Nes, M. Williamson, and T. Lenton, "Abrupt climate change in an oscillating world," *Scientific Reports*, vol. 8, no. 1, p. 5040, 2018.

[126] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 2012.

[127] M. G. Kendall, "Rank correlation methods.," 1948.

[128] K. H. Hamed and A. R. Rao, "A modified mann-kendall trend test for autocorrelated data," *Journal of Hydrology*, vol. 204, no. 1-4, pp. 182–196, 1998.