Alma Mater Studiorum
Università di Bologna

DOTTORATO DI RICERCA IN
SCIENZE STATISTICHE

CICLO XXX

Settore Concorsuale: 13/D1
Settore Scientifico Disciplinare: SECS-S/01

Statistical Inference for the Duffing Process

Presentata da: Michela Eugenia Pasetto

Coordinatore Dottorato
Prof.ssa Alessandra Luati

Supervisore
Prof.ssa Alessandra Luati
Co-Supervisore
Prof. Dirk Husmeier

Esame finale anno 2018

Alma Mater Studiorum
University of Bologna

PHD DEGREE IN
STATISTICAL SCIENCES

CYCLE XXX

Competition Field : 13/D1
Academic Discipline: SECS-S/01

Statistical Inference for the Duffing Process

Presented by: Michela Eugenia Pasetto

Ph.D. Director
Prof. Alessandra Luati

Supervisor
Prof. Alessandra Luati

Co-supervisor
Prof. Dirk Husmeier

Final exam year 2018

Abstract

The aim of the research concerns inference methods for non-linear dynamical systems. In particular, the focus is on a differential equation called Duffing oscillator. This equation is suitable to model non-linear phenomena like jumps, hysteresis, or subharmonics and it may lead to chaotic behaviour as control parameters vary. Such behaviour have been observed in many different real-world scenarios, as in economics or biology.

Inference in the Duffing process is performed with the unscented Kalman filter (UKF) by casting the system in state space form. In the context of ordinary differential equations, the uncertainty of the UKF estimates for chaotic systems is quantified by a simulation study. To overcome the limitations of the UKF when applied to the Duffing process, a new algorithm that matches Bayesian optimization (BO) and approximate Bayesian computation (ABC) within the UKF scheme is proposed. The novelty consists in (i) optimizing the sigma points location by means of maximization of the likelihood of observations with BO, and (ii) initialize the UKF with candidate parameters coming from the ABC scheme. The proposed algorithm can outperform the UKF in complex systems where the likelihood function is highly multi-modal.

Concerning stochastic differential equations, a massive simulation study is presented to evaluate the performance of the UKF for parameter estimation.

Finally, illustrations of the method with real data and further developments of the research are discussed.

Sommario

La presente ricerca ha l'obiettivo di sviluppare metodi d'inferenza per sistemi dinamici non lineari. In particolare, l'analisi è incentrata su una equazione differenziale chiamata l'oscillatore di Duffing. Tale equazione è utilizzata per modellare diversi fenomeni non lineari, quali salti, isteresi o subarmoniche, e, in generale, può mostrare comportamenti caotici al variare di parametri di controllo. Tali fenomeni sono diffusi in diversi scenari reali, sia in economia sia in biologia.

L'inferenza nel processo di Duffing è condotta tramite unscented Kalman filter (UKF) attraverso la riscrittura del sistema nella forma stato-spazio. Nel contesto di equazioni differenziali ordinarie, l'incertezza delle stime di UKF per sistemi caotici è quantificato tramite uno studio di simulazione. Per superare le limitazioni di UKF quando applicato al sistema di Duffing, viene proposto un nuovo algoritmo che unisce ottimizzazione bayesiana (BO) e approximate bayesian computation (ABC) all'interno dello schema UKF. Le novità del metodo consistono in: (i) ottimizzazione della posizione dei punti sigma tramite la massimizzazione della verosimiglianza delle osservazioni e (ii) inizializzazione di UKF con valori provenienti dallo schema ABC. L'algoritmo proposto può portare stime dei parametri migliori rispetto a UKF nel caso di sistemi complessi dove la funzione di verosimiglianza è altamente multimodale.

Per l'analisi di equazioni differenziali stocastiche, viene presentato un cospicuo studio di simulazione al fine di valutare i risultati del UKF per la stima dei parametri.

Infine, si illustra un'applicazione del metodo su dati reali e si discutono gli sviluppi futuri della ricerca.

Acknowledgements

This dissertation has materialized thanks to the contribution of many people to whom I am enormously indebted and I have the pleasure of expressing my gratitude.

I gratefully acknowledge the persistent support and encouragement from Professor Dirk Husmeier. He provided constant academic guidance and he had a strong influence on my scientific development. Dirk's help was not limited to statistical studies, but offered the best support in every aspect of my Ph.D. life.

Many thanks to Professor Alessandra Luati for her valuable comments on my research. She gave me many insightful feedbacks, provided financial funds and she is always enthusiastic in proposing new presentations of our work.

I thank the reviewers, Ernst Wit and Guido Sanguinetti, for helping me to improve the research with their accurate remarks.

I am very grateful to Umberto Noè. He has been offering his generous help since the beginning of my study in the UK. His patience was invaluable for me to go through some tough months at University of Glasgow. It is my hope to continue working together.

Sincere thanks to many friends with whom I share more than just an academic relationship: Elisa, Chiara, Francesco, Lorenzo, Ester, Francesca, Riccardo, and Maso. They were vital in making my stay in Bologna enjoyable.

I wish to express my deep gratitude to the founder of the typographical workshop "La tana della volpe" and his partner, not only for their editing suggestions, but mostly for their endless, brilliant and courteous support.

Finally, I would like to thank my parents for helping me throughout all my studies and for providing a home in which to complete my writing up.

Without the aid of these people, I am sure that I would have made many more mistakes. All remaining errors are mine alone.

Contents

1. Introduction

Differential equations are a powerful tool to describe dynamical processes: many systems studied in biology, social sciences, engineering and economics are described by ordinary or stochastic differential equations (ODEs or SDEs). Modelling ODEs and SDEs requires to reconstruct the hidden signal behind noisy observations and to estimate parameters that represent particular features of the system. Several algorithms have been proposed to face these necessities.

In this dissertation, the focus is on a non-linear differential equation called Duffing process. This equation bears the name of its detector and it has often been associated to the most known Van der Pol system. The Duffing equation describes many different non-linear phenomena, as jumps, hysteresis, and bifurcations. These behaviours are highly widespread in biological and economical processes and the aim of this research is to use the Duffing system to model data that shows non-linear dynamics.

Therefore, the goal of the project is the development of inference schemes for the Duffing equation. Analyses, both quantitative and qualitative (i.e. geometric), for the Duffing system has been extensively carried out in the physical and engineering literature, but so far the author is not aware of any statistical method to infer the signal and parameters for such oscillator.

The Duffing equation is described by a state space representation and a Kalman filter based method, known as the unscented Kalman filter (UKF), is utilized to conduct inference. The UKF is a non-linear extension of the Kalman filter, based on the unscented transform, which approximates Gaussian distributions on a deterministically chosen set of points, called in the literature as sigma points. In the evaluation of the UKF performance in the context of ODEs, two limits are faced.

The first limitation concerns the sigma points location in the likelihood space. Non-linear systems may show highly multi-modal likelihood: in this case the de-

terministic position of sigma points may result in a noticeably loss of accuracy in the estimates. Such limit is overcome inserting a Bayesian optimization (BO) algorithm inside the UKF steps that is able to understand the "best" sigma points location depending on the likelihood space the UKF is dealing with. A comparison between the performance of two acquisition functions (the expected improvement and the upper confidence bound) is also discussed.

The second limit is related to the choice of the starting values. The UKF convergence to the true parameters is highly affected by its initialization. Here, the class of approximate Bayesian computation (ABC) method is considered. The idea behind this approach is to find an approximate posterior distribution from which to pick candidate parameters to initialize the UKF.

The new algorithmic method that couples the ABC (performed with a sequential Monte Carlo sampling scheme) within the UKF with optimized sigma points according to the BO method is called Sequential ABC-UKF.

In the context of SDEs, the discussion focuses on numerical methods to approximate stochastic integrals. In this case, the UKF is modified to include the Euler-Maruyama approximation scheme to solve the SDE.

A large simulation study is developed, both for ODEs and SDEs.

Finally, two analyses for the U.S. gross domestic product (GDP) and the sunspot numbers are discussed. Many real-world time series may hide different generative parameters of the process depending on time intervals. To locate in the time domain the changes in these parameters, i.e. to identify the breakpoints in which the system switches among hidden regimes, a changepoint detection method based on wavelets is described. So far, the method is heuristic in the sense that it is a model-free scheme: the development of a model-driven changepoint detection strategy is one of the ongoing projects presented at the end of the thesis.

The dissertation is organized as follows. The Duffing process and a qualitative analysis on the non-linear phenomena that the system describes are presented in Chapter 2. The remainder of the thesis concerns inference methods. Chapter 3 is a review of the UKF, BO and ABC, and the proposed algorithm Sequential

ABC-UKF is discussed in the last Section. The simulation study for ODEs is shown in Chapter 4, while Chapter 5 is dedicated to simulation studies for SDEs. The changepoint detection method and the real data illustrations are in Chapters 6 and 7, respectively. Finally, conclusions and future research are discussed in Chapter 8.

2. The Duffing process

The aim of this Chapter is twofold. Initially, the main concepts for a quantitative and qualitative analysis of dynamical systems are described. Then, the behaviour of a non-linear second order differential equation called *Duffing equation* is discussed. I introduce the main mathematical tools developed in the literature to study the non-linearity of the Duffing system, but the reader is referred to more specialized books for a comprehensive discussion (e.g. Stoker, 1950 and Kovacic and Brennan, 2011). Here, the goal is to introduce non-linear phenomena.

Georg Duffing (1861 - 1944) was an engineer and his research comes from his personal practical experience of engineering systems. The most important Duffing's publication is "Forced oscillations with variable natural frequency and their technical significance" (Duffing, 1918), a book in which he described his studies on the pendulum and introduced a non-linear second order differential equation now bearing his name.

Nowadays, in the physical and engineering literature, the term *Duffing oscillator* indicates any equation with a cubic stiffness, regardless to the type of damping or excitation. In this work, however, following the original definition of Duffing's equation, free or forced harmonic vibration of an oscillator with linear viscous damping are studied.

The Duffing oscillator is suitable to describe several non-linear phenomena. Depending on the differential equation parameters, local bifurcations can occur in the system, arising from jump phenomena or sub-harmonics, and lead to chaotic responses (e.g. the period-doubling route to chaos). Bifurcations characterise a power spectra with the presence of frequencies beside the fundamental one. In the time series literature, the existence of different frequencies points out the presence of cycles nested one inside the other (Ramsey, 1990).

The Duffing equation sits alongside the most known Van der Pol equation (Kovacic and Brennan, 2011). Both the systems have been studied in the same historical period, and in the literature both equations are fundamental in non-linear dynamical

studies. In statistics, more attention has been devoted to the Van der Pol equation, with respect to the applications to time series (see Ramsey, 1990, Tong, 1990 and references therein) and parameter inference (Sitz et al., 2002). There is not a real motivation for which the Duffing system is less famous among statisticians. Some author claimed its importance (Tong and Lim, 1980), but, up to now, I am not aware of any comprehensive statistical discussion and inference scheme for non-linear phenomena arising from the Duffing process.

In this Chapter, Section 2.1 concerns the historical development of the theory of non-linear oscillations and the qualitative analysis for differential equations. Then, in Section 2.2, the non-linear phenomena that Duffing's equation models are described. The Sections 2.3 and 2.4 highlight the behaviour of the system depending on parameter values for, respectively, deterministic and stochastically excited vibrations. Finally, Section 2.5 gives an idea of the scientific influence of the Duffing's system.

2.1. *Non-linear dynamical systems*

The study of non-linear oscillations is vast, and it is behind the scope of this section to give a detailed description of the mathematical tools for the analysis of differential equations. In what follows, the focus is on the aspects related to Duffing's work, while for a review on the historical development of non-linear dynamical systems the reader is referred to Holmes (2005) and Shaw and Balachandran (2008). An outline on the methods to analyse dynamical systems, i.e. phase-space representation, is given in Section 2.1.2.

2.1.1. Historical perspective

The concept of non-linear vibrations have been known since Christiaan Huygens invented the pendulum clock (Huygens, 1673); however, Duffing was the one to tackle the problem of non-linear oscillators in a systematic way (see the Introduction in Kovacic and Brennan, 2011). The history started with a pendulum: in 1583, the 19-year-old Galileo timed the oscillation of a swinging chandelier, recording one of the early registration of oscillations in the history. Galileo noticed that

the *frequency*, that is the number of oscillations per unit of time, of the pendulum was independent of the *amplitude*, defined as the maximum numerical value of the vibration. Some years later, Huygens discovered that wide wings made the pendulum inaccurate, observing that the natural *period*, which is the time taken to go from one end to the other and return to the very beginning, was dependent upon the amplitude of motion: the pendulum is inherently non-linear. The relation between the amplitude and the period, more commonly called *amplitude-frequency dependence*, is a characteristic of non-linear oscillations. Huygens also found out that if the pendulum varied its length during the oscillation, then the frequency of oscillation became independent of the amplitude: today one may say that he linearised a non-linear system. Nevertheless, at that time, only rudimentary tools were available, until Leonhard Euler wrote down the *differential equation* of motion of an oscillator (Euler, 1750). Euler formally introduced the concept of non-dimensional *driving frequency* $\Omega = \omega/\omega_n$, where ω is the natural frequency and ω_n is the *input frequency*, i.e. the frequency of an external excitation. Euler noted that the response becomes infinite when $\Omega = 1$; hence, he was the first to explain the phenomenon of *resonance* in which an input force let oscillate another system with greater amplitude at specific frequencies. A century later, Hermann Von Helmholtz and Baron Rayleigh published their huge works on acoustics and vibrations (Helmholtz, 1885 and Rayleigh, 1896), closely linked with the technological (mainly mechanical and electrical) development of the nineteenth century.

The generic form of a differential equation of interest at the time was the following

$$m\ddot{x} + \phi(\dot{x}) + g(x) = F\cos\omega_n t, \tag{2.1}$$

where x is the position of the oscillation at time t, $\ddot{x} = d^2x/dt^2$ is the second derivative of x with respect to time t (or the acceleration), and $\dot{x} = dx/dt$ is the first derivative of x, that is the velocity. The coefficient m represents the mass of a body, and $m\ddot{x}$ is the *inertia force*, which is the resistance of an object to a change in its motion. The function $\phi(\cdot)$ depends on the velocity and it is called *damping force*. The latter quantifies the decay of the oscillation: the greater the damping, the faster the vibration reaches the state of rest. The term $g(x)$ is a *restoring force* or *spring force*, that is a force

that brings equation (2.1) toward equilibrium. The term $F\cos\omega_n t$ is the *external force* or *excitation*, and represents an input that perturbs the system. The amplitude F and the input frequency ω_n are fixed, i.e. they are not time-varying.

Let us imagine that a system is perturbed away from the equilibrium: the restoring force directs the oscillation back to its original position. For example, a non-swinging pendulum has all the forces in equilibrium at the bottom of the swing. If a force puts the pendulum in motion, the gravity brings the pendulum back down to the midpoint of the swing. In this case, the gravity can be seen as a restoring force. Another example may be described by the action of a spring. A spring exerts a force on an object proportional to the amount of deformation of the spring itself from its equilibrium length. When the spring is pulled to a greater extent, the restoring force let return the spring back to the initial equilibrium length.

Equation (2.1) may represent a *dynamical system*. This describes the time dependence of motion of a point in a space; in other words, a dynamical system models the displacement of a particle whose state varies over time.

In 1918 Duffing took part in the discussion on differential equations based on equation (2.1) by defining $\phi(\dot{x}) = c\dot{x}$ and $g(x) = \alpha x + \beta x^3$, and around the same time Balthasar Van der Pol developed the famous equation describing oscillations generated by a triode valve (Van der Pol, 1920). The meaning of the parameters α, β, c, is discussed in the remainder of this Chapter. The Duffing and the Van der Pol equations reveal many phenomena, such as jumps, frequency entrainment, limit cycles and amplitude-frequency dependence, that I will describe later.

In the development of the theory of non-linear oscillations, Henri Poincaré achieved crucial results (Poincaré, 1881). He introduced the *qualitative* (or geometric) analysis of non-linear systems and the revolutionary abstract idea of *limit cycle*. A limit cycle is a self-sustained oscillation and the first to identify this in practice was Aleksandr Aleksandrovich Andronov (Andronov and Khajkin, 1949). In 1929, Andronov pointed out that self-excited oscillations could be found in many different situations, from the vibrating strings of a violin to chemical reactions, from valves with electrodes to biological systems, etc. Andronov spent his entire life in the development of *phase space* analysis, introducing topological objects called *attractors* or *repellors*.

These latter, along with the main mathematical concepts of the studies of Poincaré and Andronov, will be explained in the following Section.

2.1.2. Phase-space representation

A phase space of a dynamical system is a space in which all possible states are represented. The evolution of the system through time is described by the path in the 2D space, with coordinates $(x, \dot{x})$. In such a space, it is possible to identify the points to which the system approaches as time t progresses independently of starting conditions, with t being a parameter. The points where $(x, \dot{x}) = (0, 0)$ are called *critical points*. These do not depend on the initial states but exclusively on the parameters of the system, that is they are intrinsic properties of a dynamical process. Critical points can be defined as *stables*, i.e. the points where the system converges, *unstables*, where the process diverges, or *saddles*, points in which the motion dynamics is attracted first and then it moves away. A set of stable critical points is often called an attractor, while a repellor in a set of unstable critical points.

A dynamical system may be analysed both in the time and phase-space domain in the following way. Let us imagine to deal with a periodic oscillation. The curves plotted in a phase-space correspond to the motion x: on the one hand, if x is periodic, the corresponding $x, \dot{x}$-curve is closed. On the other hand, if $x, \dot{x}$-curve is closed, the displacement and the velocity at time t are reached again after a certain time T, so that it is evident that the motion is periodic.

The critical points are classified depending on the eigenvalue of the system, as Section 2.3.1 will show. In the following Section, a qualitative analysis is discussed.

2.2. *Qualitative analysis on systems described by the Duffing oscillator*

Consider the displacement x at time t and a force g applied to a spring. In a system described by the Duffing oscillator, the source of the non-linearity between x and g is the *stiffness*. The latter is the rigidity of an object in resisting to deformations in response to an applied force. Therefore, the stiffness is a function of the applied force and the displacement resulting from the application of force. If the system is symmetric, that is the stiffness of a spring is the same in compression or in tension, the

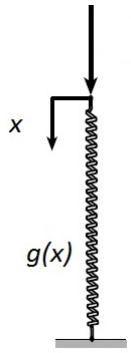

Figure 2.1: The motion x and the force applied to a spring given by equation (2.2). Here, the non-linear spring has a stiffness which is function of the displacement. Source: Figures 2.1 of Kovacic and Brennan (2011).

energy in one direction (e.g. in compression) is equivalent to the energy in another direction (in tension). In this case, the restoring force can be approximated by a series in x in which the exponents of x are odd integers (Kovacic and Brennan, 2011). If the series is truncated after the first two terms, the force-deflection relationship is

$$g(x) = \alpha x + \beta x^3, \quad \alpha > 0, \tag{2.2}$$

where α is the frequency at which the system oscillates and β is the coefficient of the stiffness. If $\beta > 0$, the spring is called *hard* because it becomes stiffer with increasing x. If the coefficient β associated to the cubic term is negative, the spring is *soft*, i.e. as more the displacement, as more the spring gets softer.

Equation (2.2) describes a non-linear restoring force, or non-linear change in the potential energy, that tends to bring the system toward equilibrium.

Differentiating equation (2.2) gives $dg(x)/dx = \alpha + 3\beta x^2$: it can be seen that the frequency α is independent of the x, while the term $3\beta x^2$ is function of displacement and the source of non-linearity. Figure 2.1 depicts the motion and the force applied to a spring.

Inserting equation (2.2) into (2.1) and setting $m = 1$, a differential equation with a non-linear restoring force g is

$$\ddot{x} + \phi(\dot{x}) + \left(\alpha x + \beta x^3\right) = F \cos \omega_n t. \tag{2.3}$$

When $\phi(\dot{x}) = c\dot{x}$, where c is a constant damping term, the above equation is called the Duffing equation:

$$\ddot{x} + c\dot{x} + \left(\alpha x + \beta x^3\right) = F\cos\omega_n t. \tag{2.4}$$

To proceed further with the analysis of systems described by the Duffing oscillator (2.4), the case of undamped (since $c\dot{x} = 0$) and free ($F\cos\omega_n t = 0$) oscillations is considered first. Hence, the equation of motion for undamped and free vibrations is

$$\ddot{x} = -\left(\alpha x + \beta x^3\right). \tag{2.5}$$

A closed form solution of equation (2.5) is available using elliptic integration (see e.g. Salas, 2014 and Marinca and Herişanu, 2011), but in the following sections I focus on two qualitative discussions. The first analysis mostly concerns the behaviour of the motion, while the second explains non-linear phenomena that can arise in the Duffing oscillator.

2.2.1. Periodic and non-periodic motion

In this Section, the type of oscillations coming from (2.5) are analysed.

Integrating equation (2.5), the motion can be rewritten as

$$\int \ddot{x}dx = -\int \left(\alpha x + \beta x^3\right) dx,$$

$$\frac{\dot{x}^2}{2} = -\frac{\alpha}{2}x^2 - \frac{\beta}{4}x^4 + h,$$

$$\dot{x}^2 + \alpha x^2 + \beta\frac{x^4}{2} = h = \text{constant}, \tag{2.6}$$

where the constant h represents twice the total energy of the system. In the neighbourhood of the origin, $x = 0$ and $\dot{x} = 0$ (for small value of displacement x), the term $\beta x^4/2$ can be neglected in comparison with αx^2, and (2.6) has the appearance of an ellipse. Hence, the constant h is small and positive and the closed curves representing the motion in the phase-space are elliptic near the origin. Setting the velocity to zero in equation (2.6), the maximum displacement $x_{\max} = A$ is merely the solution of the second-order polynomial in x:

$$A^2 = \frac{-\alpha + \sqrt{\alpha^2 + 2\beta h}}{\beta}, \tag{2.7}$$

in which the positive sign of the radical is taken for both $\beta > 0$ (hard spring) and $\beta < 0$ (soft spring) because the constant h and A^2 should be small and positive. Since the closed curves (2.6) are symmetric, the period T of motion can be obtained in the form

$$T = 4\int_0^A \frac{dx}{\dot{x}} = 4\int_0^A \frac{dx}{\sqrt{h - \left(\alpha x^2 + \beta x^4/2\right)}}. \tag{2.8}$$

Expression (2.8) is obtained after some mathematics starting from Newton's law of conservation of energy, see Stoker (1950) for a detailed derivation. The integral (2.8) can be re-written by changing variable of integration in the following way. From (2.6), if A^2 is the root of a polynomial $h - (\alpha z + \beta z^2/2) = 0$, then

$$h - \left(\alpha x + \beta x^2/2\right) = \frac{\beta}{2}\left(A^2 - x^2\right)\left(b^2 + x^2\right), \tag{2.9}$$

where

$$\frac{\beta}{2}\left(-b^2 + A^2\right) = -\alpha, \quad \text{or} \quad \beta b^2 = \beta A^2 + 2\alpha. \tag{2.10}$$

Let replace x in (2.8) with a new integration variable θ:

$$x = A\sin\theta. \tag{2.11}$$

Using (2.10) to eliminate the term b^2, the integral (2.8) takes the form

$$T = 4\sqrt{2}\int_0^{\pi/2} \frac{d\theta}{\sqrt{2\alpha + \beta A^2 + \beta A^2 \sin^2\theta}}. \tag{2.12}$$

Equation (2.12) highlights the *dependence between the period and the amplitude*. The non-linear relationship between T and A has a different behaviour depending on the sign of β.

In the case of hard spring, $\beta > 0$, the period of oscillation decreases (hence, the frequency increases) for increasing amplitude. The vibration in equation (2.6) is represented by a closed curve, and the motion is periodic, so that equation (2.12) is valid under all circumstances.

If $\beta < 0$ (soft spring), T grows as far as A increases: the frequency of the oscillation decreases with an increasing amplitude of motion. Equation (2.12) is meaningful

only if A given by (2.7) is small. Indeed, the curves in (2.6) are closed ellipses in a certain region of the phase-space only for small values of A. If the system has decreasing frequencies and increasing amplitudes over time, it may enter in a non-periodic regime represented by unstable focus or saddles in the phase-space.

Figure 2.2 sketches the relation between the amplitude and the circular frequency $\omega = 2\pi/T$ in the cases of linear, soft and hard springs. The frequency-amplitude plane is called *response diagram*. Setting $\beta = 0$, that is a linear spring force, in equation (2.12), the period becomes $T = 2\pi/\sqrt{\alpha}$, hence the frequency $\omega = \sqrt{\alpha}$ is the common tangent for all the curves. The amplitude is independent from the frequency for a linear spring, while it increases or decreases according to the growth or reduction of the frequency.

Figure 2.3 gives more insight of the behaviour of the oscillation depending on the type of spring in the phase-space. In the left panel of Figure 2.3, the hard spring case, the ellipses of equation (2.6) demonstrates the periodicity of the system. The arrows in the curves indicate the direction of motion, i.e. the direction of $(x, \dot{x})$ with increasing t. For $\beta < 0$, the features of the oscillation are more complicated. For ease of analysis, let set $\rho^2 = -\beta$. Hence, equation (2.6) can be written is the following form

$$\dot{x}^2 + \alpha x^2 - \rho^2 \frac{x^4}{2} = h. \tag{2.13}$$

For $h > 0$ and small x, the (2.13) represents ellipses around the origin (remember that the term x^4 is neglectable compared to x^2). In this case, when $x = 0$, the velocity takes a value $\dot{x}_0$. The quadratic polynomial in x in the right hand side of (2.13) has roots $x_{1,2}$

$$x_{1,2} = \frac{\alpha \pm \sqrt{\alpha^2 - 2\rho^2 h}}{\rho^2}. \tag{2.14}$$

Since the discriminant $\Delta = \alpha^2 - 2\rho^2 h$ is positive, the curves are ellipses that tend to stretch, i.e. the values in which the curves cross the x axis increase, until Δ approaches zero. The bold line in Figure 2.3 represents the case where the discriminant equals to zero. When $\alpha^2 = 2\rho^2 h$, or $h = \alpha^2/2\rho^2$, the curve crosses the axis with the following values: (i) if $x = 0$, the velocity is $\dot{x}_0 = \alpha/\rho\sqrt{2}$, (ii) if $\dot{x} = 0$, the position has roots $x_{1,2} = \pm\sqrt{\alpha}/\rho$. For $\dot{x}_0 > \alpha/\rho\sqrt{2}$ and $\Delta < 0$, open curves described by the quadratic

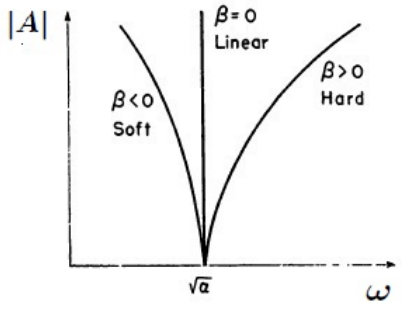

Figure 2.2: Amplitude-frequency dependence (here, the circular frequency $\omega = 2\pi/T$) for linear ($\beta = 0$), hard ($\beta > 0$) and soft ($\beta < 0$) spring forces. When $\beta = 0$ in equation (2.12), the frequency $\omega = \sqrt{\alpha}$ is the common tangent to all the curves. The amplitude is independent of the frequency in the linear case, while A increases with increasing or decreasing ω depending on the sign of the coefficient β. Source: Figure p. 22 of Stoker (1950).

polynomial in x arises. The points $(\pm\sqrt{\alpha}/\rho, 0)$ divide the space into three regions. The first area is represented by the periodic motion around the origin (the ellipses), the second is described by the open curves that do not cross the x axis but the $\dot{x}$ (when $\dot{x}_0 > \alpha/\rho\sqrt{2}$ and $\Delta < 0$), and the third is obtained for open curves that cross the x axis and not the velocity axis.

The qualitative discussion described so far for the system (2.5) emphasizes that *the amplitude-frequency dependence is related to non-periodic and unstable motion.* Phase-space analysis highlights that three critical points exist. The origin is a *centre*, in the sense that the system has a periodic motion, while the points $(\pm\sqrt{\alpha}/\rho, 0)$ are saddles.

2.2.2. Non-linear phenomena arising from amplitude-frequency dependence

In the preceding Section, the oscillation (2.5) has been discussed using phase-space analysis. Here, I show another qualitative study for the same motion suitable for the representation of many different non-linear phenomena.

In the evaluation of the free or forced oscillator (2.5), Duffing explained the following iterative method of approximations. The present discussion is heuristic, and the rigorous derivations can be found in Andronov and Khajkin (1949). Consider

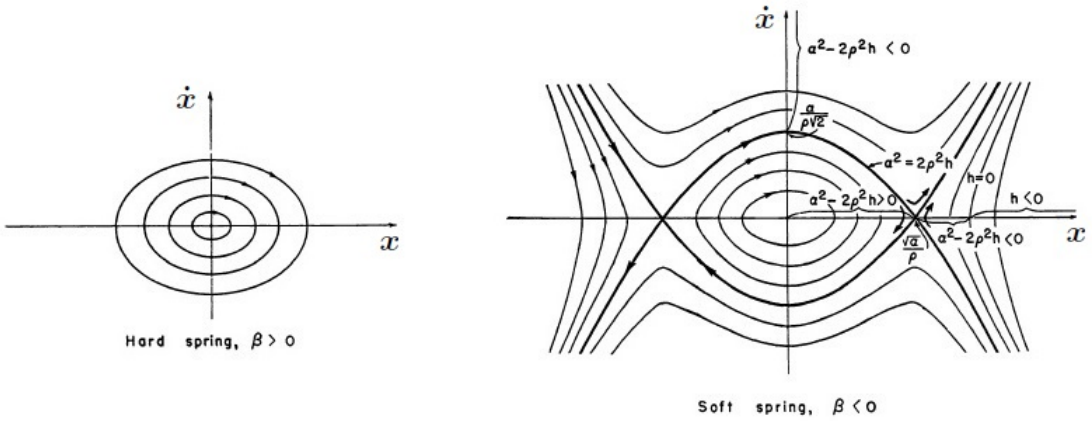

Figure 2.3: Phase-space for hard ($\beta > 0$) and soft ($\beta < 0$) spring forces. The axis are the position and the velocity. The motion is periodic when $\beta > 0$. For $\beta < 0$, the behaviour of the oscillation is more complex. In this case, the motion is periodic around the origin, but it becomes unstable as the frequency α varies. The points $(\pm\sqrt{\alpha}/\rho, 0)$ are saddles and divide the space into regions characterized by periodic or non-periodic oscillations. Source: Figure p. 23 of Stoker (1950).

the deterministically forced vibration

$$\ddot{x} = -\alpha x - \beta x^3 + F\cos\omega t, \tag{2.15}$$

where F and ω are real constants. If $\beta = 0$, the exact solution of (2.15) is $A\cos\omega t$. For small β, it is still possible to assume that $A\cos\omega t$ is a reasonable first approximation, and this is inserted in the right-hand side of (2.15) to obtain a second approximation x_1:

$$\ddot{x}_1 = -\left(\alpha A\cos\omega t + \beta A^3\cos^3\omega t\right) + F\cos\omega t. \tag{2.16}$$

Observing the identity

$$\cos^3\omega t = \frac{3}{4}\cos\omega t + \frac{1}{4}\cos 3\omega t, \tag{2.17}$$

equation (2.16) becomes

$$\ddot{x}_1 = -\left(\alpha A\cos\omega t + \beta A^3\left(\frac{3}{4}\cos\omega t + \frac{1}{4}\cos 3\omega t\right)\right) + F\cos\omega t$$

$$\ddot{x}_1 = -\left(\alpha A + \frac{3}{4}\beta A^3 - F\right)\cos\omega t - \frac{1}{4}\beta A^3\cos 3\omega t. \tag{2.18}$$

Integrating (2.18) twice (and setting the integration constants to zero, since only periodic solutions are of interest), the solution is

$$x_1 = \frac{1}{\omega^2}\left(\alpha A + \frac{3}{4}\beta A^3 - F\right)\cos\omega t + \frac{1}{36}\frac{\beta A^3}{\omega^2}\cos 3\omega t. \tag{2.19}$$

The iteration procedure based on reinserting successive approximation in the right-hand side of (2.15) requires that the constants α, β, ω and F be all sufficiently small in order to assure convergence. Figure 2.4 shows that for small β, the relationship between the amplitude and the frequency should lie in the vicinity of the linear oscillation. Thus, the successive iterations sketched above are good approximations to solve (2.15) only for small $|A|$. When $\omega \to \sqrt{\alpha}$, the amplitude $|A| \to \infty$. This phenomenon, that is the magnification of the displacement for certain values of the driving frequency, is called of *resonance*. To let $|A|$ remain small, the frequency ω should held fixed and $|A|$ becomes a function of it, but in this way it is quite impossible to obtain the essential features of the oscillation. Hence, Duffing took a bold step (Tong, 1990). The coefficient $A_1 = \alpha A + \frac{3}{4}\beta A^3 - F$ of $\cos\omega t$ in (2.19) is taken equal to A, on the basis that if $A\cos\omega t$ is a truly first approximation, A_1 should not differ much from A. Duffing's reasoning leads to

$$A = \frac{1}{\omega^2}\left(\alpha A + \frac{3}{4}\beta A^3 - F\right), \tag{2.20}$$

$$\omega^2 = \alpha + \frac{3}{4}\beta A^2 - \frac{F}{A}. \tag{2.21}$$

Equation (2.21) represents the *basic amplitude-frequency relation* and it is crucial for the remainder of the discussion. The dependence stated in (2.21) is the analogous of the period-amplitude relation of (2.12). The frequency ω is a function of A. Therefore, the amplitude is prescribed in advance, and the frequency has to be determined. To understand this procedure it is important to consider ω as depending upon A. Indeed, a significant point in (2.21) is the *multi-value aspect of the relation*: for certain frequencies there are three corresponding values of A. The iteration procedure represented in (2.20) may look like slightly unnatural since usually the frequency is prescribed in advance and does not depend on the amplitude. However, the apparently rather natural procedure of keep inserting approximate solutions in (2.15) could not yield

all the possible set of curves in the response diagram and the multi-valued relationship between A and ω. Duffing's prediction has been checked experimentally in a variety of cases, with a good agreement between the theory and experimental results (Tong, 1990).

In the case of free oscillation, the response relation is

$$\omega^2 = \alpha + \frac{3}{4}\beta A^2, \tag{2.22}$$

and the exact result $\omega^2 = \alpha$ corresponds to the linear oscillator.

The left panel in Figure 2.5 shows the response curves for hard, soft and linear springs. The non-linear stiffness parameter has the effects to stretch the response curve to the left or to the right with respect to the linear case, and so, for some ωs, the amplitude A can assume three values.

When a damping term $c\dot{x}$ is added to the left-hand side of (2.15), the response curves are similar to the (2.21) but closed in their top, since the role of the damping is to decrease the amplitude (right panel of Figure 2.5).

These response curves predict several phenomena, such as jumps and subharmonics, that are described later.

Notice that, from (2.19), the solution takes the generic form $x_1 = P\cos\omega t + Q\cos 3\omega t$. Thus, one may infer that a general solution of (2.15) involves all odd harmonics: in fact this is the case (Stoker, 1950). Going further with the iteration procedure, a second approximate solution has the form $x_2 = P\cos\omega t + Q\cos 3\omega t + R\cos 5\omega t$. Hence, the amplitude A in equation (2.21) can also be interpreted as the first Fourier coefficient. Notice that the occurrence of harmonics besides the fundamental frequency is a characteristic of non-linear dynamical systems.

Jump phenomena

The curves in Figure 2.5 lead to several conclusions. Let us consider an experiment in which the amplitude F is held constant, and the frequency slowly varies. The response of the amplitude is observed. Starting with the hard spring force, $\beta > 0$, the initial frequency is ω at point 1 in Figure 2.6. The frequency decreases from point 1 to point 2 until point 3 is reached. At the same time, the amplitude increases with

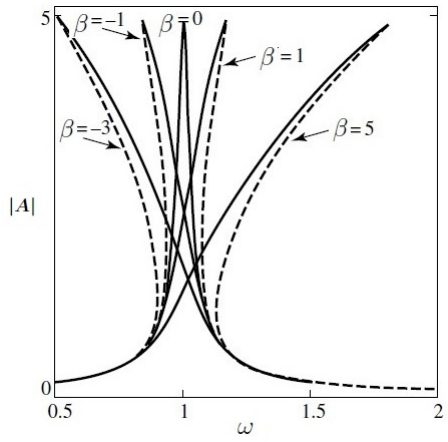

Figure 2.4: Response curves for varying strength of non-linear stiffness parameter β. The non-linear response diagrams between A and ω represented by equation (2.21) can arise from the linear response curve by bending the latter to the right (for $\beta > 0$) or to the left (for the soft spring $\beta < 0$). The more the coefficient β is far from zero, the more the system becomes non-linear. Source: Figure 5.3 of Kovacic and Brennan (2011).

decreasing frequencies. A further decrease of ω from point 3 would let the amplitude to abruptly *jump* from point 3 to point 4. Then, the value of A slowly decreases from point 4 to point 5. It is now clear that the term *jump phenomenon* indicates a dramatic change of the steady-state behaviour due to the transition from one stable solution to another stable solution as a control parameter varies. The same experiment can be performed in the opposite direction, that is starting from point 5 and increasing the frequency. The amplitude follows the increasing trajectory of points 5-4-6 and suddenly shrinks to point 2 and slowly decreases afterwards. The jump phenomenon in the soft spring case can be described in a similar way, but the jumps take place in the reverse direction.

The jump phenomena in the direction of points 1-2-3-4 for $\beta > 0$ (or the one of points 5-4-6-2 with $\beta < 0$) could be explained even in the absence of damping, but the latter is essential to explain the jump from point 6 to 2 or from 3 to 4 in the case, respectively, of hard and soft spring forces.

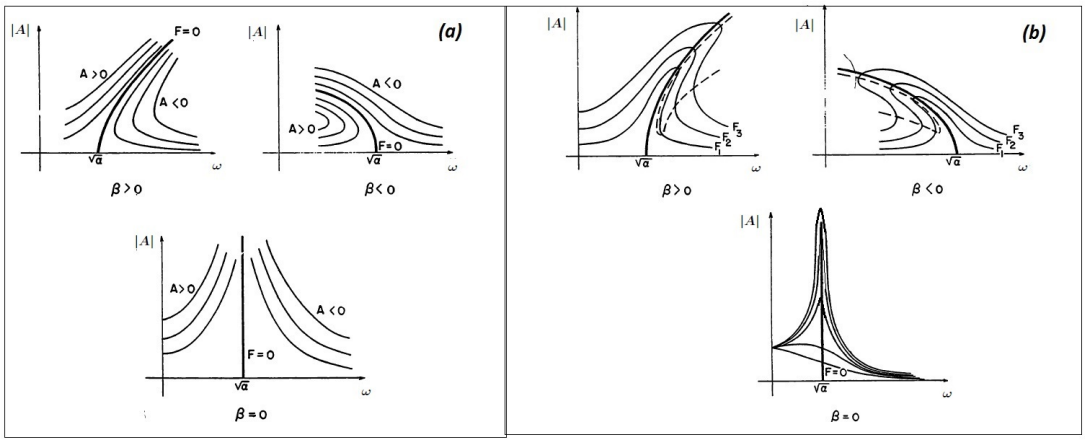

Figure 2.5: Response diagrams for linear ($\beta = 0$), hard ($\beta > 0$) and soft ($\beta < 0$) spring forces of equation (2.21). Notice that when ω approaches $\sqrt{\alpha}$, the response becomes infinite ($|A| \to \infty$), i.e. the phenomenon of resonance occurs. (a) Response curves without damping. (b) Response curves when damping is present: the latter closes the top of the curves. The common tangent for all the curves is the frequency $\omega = \sqrt{\alpha}$ (see Figure 2.2) and the amplitude of the input F. The different curves represent different values of the external amplitude F_1, F_2, F_3. Source: Figure pp. 88 and 92 of Stoker (1950).

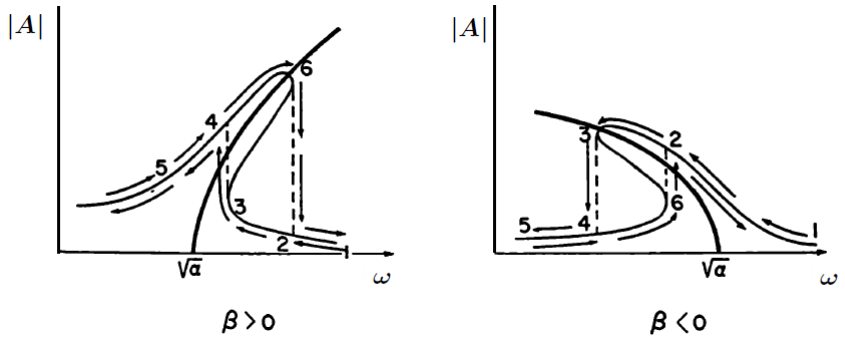

Figure 2.6: Jump phenomena for hard and soft spring forces. Following the trajectory 1-2-3-4 (for $\beta > 0$) or the direction of points 5-4-6-2 (when $\beta < 0$), the amplitude abruptly changes its value for, respectively, decreasing or increasing ω. Notice that the jumps in the two directions happen at different frequencies. Source: Figure p. 95 of Stoker (1950).

HYSTERESIS

So far, the frequency ω has been considered as a control parameter, being fixed the input amplitude F. An alternative way to analyse the dynamics of the system

is looking at the amplitude A as a function of F, while ω is constant. The response curve of A for varying F is shown in Figure 2.7. As F increases up to value F_2, the amplitude A has a jump and then gradually grows. On the contrary, when F decreases until point F_1 is reached, the value of A dramatically shrinks. There are three coexisting solutions for $F_1 < F < F_2$ and only one solution outside this region. This phenomenon is called *hysteresis*.

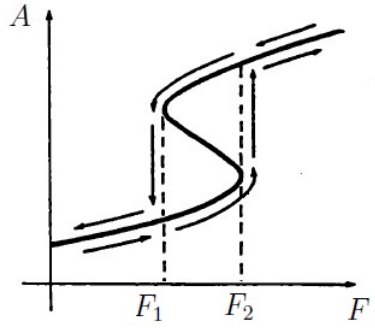

Figure 2.7: Response curve showing the phenomenon of hysteresis. In this experiment the frequency ω is held constant and the amplitude A jumps upwards when the input amplitude F increases up to point F_2. The A jumps down when F decreases until the value F_1. In the interval $F_1 < F < F_2$ three coexisting solutions exist. Source: Figure p. 95 of Stoker (1950).

Subharmonic response

Up to now, only harmonic solutions of (2.15) have been considered. However, oscillations whose frequency is a fraction of the fundamental frequency are characteristics of non-linear systems. The term *subharmonic response* is usually applied to define these phenomena (Tong, 1990). The approximate solution (2.19) shows how the (2.15) can be expressed in terms of Fourier series, that are now used to understand the qualitative behaviour of subharmonics.

In this Section, I will not present in all generality the problem and the solutions for many different subharmonic responses. Rather, I treat the special case with the subharmonic oscillation of order 1/3, and the analysis of higher harmonics can be found in Kovacic and Brennan (2011) or Stoker (1950).

For ease of discussion, the damping is neglected. Defining $\theta = \omega t$, the oscillation (2.15) can be re-written as

$$\omega^2 \ddot{x} + \alpha x + \beta x^3 = F \cos\theta. \tag{2.23}$$

A subharmonic of frequency 1/3 of the new variable θ can be developed in a Fourier series of the form

$$x = \sum_{n=1}^{\infty} a_n \cos\frac{n\theta}{3} + b_n \sin\frac{n\theta}{3}. \tag{2.24}$$

On the assumption of small β, and remembering that the sine terms and even multiples of $\theta/3$ in the cosines become zero, the general form of the solution in terms of Fourier series is

$$x = A_{1/3} \cos\frac{\theta}{3} + A_1 \cos\theta + A_{5/3} \cos\frac{5\theta}{3} + \dots. \tag{2.25}$$

Equation (2.25) can be substituted into (2.23), using some algebra manipulations and the following identities

$$\begin{aligned}
\cos^3\frac{\theta}{3} &= \frac{3}{4}\cos\frac{\theta}{3} + \frac{1}{4}\cos\theta,\\
\cos^2\frac{\theta}{3}\cos\theta &= \frac{1}{4}\cos\frac{\theta}{3} + \frac{1}{2}\cos\theta + \dots,\\
\cos\frac{\theta}{3}\cos^2\theta &= \frac{1}{2}\cos\frac{\theta}{3} + \dots,\\
\cos^3\theta &= \frac{3}{4}\cos\theta + \dots,\\
\cos^2\frac{\theta}{3}\sin\theta &= \frac{1}{4}\sin\frac{\theta}{3} + \frac{1}{2}\sin\theta,\\
\cos\frac{\theta}{3}\sin^2\theta &= \frac{1}{2}\cos\frac{\theta}{3} + \dots.
\end{aligned}$$

Hence, the following relations are obtained

$$\left(\alpha - \frac{\omega^2}{9}\right) A_{1/3} + \frac{3}{4}\beta\left(A_{1/3}^3 + A_{1/3}^2 A_1 + 2A_{1/3}A_1^2\right) = 0, \tag{2.26}$$

$$\left(\alpha - \omega^2\right) A_1 + \frac{1}{4}\beta\left(A_{1/3}^3 + 6A_{1/3}^2 A_1 + 3A_1^3\right) = F. \tag{2.27}$$

The equations (2.26)-(2.27) are the analogous of equation (2.21) that was crucial for the harmonic case, but here two harmonics are considered. As in the above iteration

procedure, β is set to zero in (2.26) and (2.27). The first Fourier coefficient A_1 is zero unless $\alpha - \omega^2/9 = 0$; hence, A_1 is non-zero only if $\omega = 3\sqrt{\alpha}$. For an arbitrary value of A_1 (when $\omega = 3\sqrt{\alpha}$), the second Fourier coefficient in (2.27) turns out to be $A_1 = -F/8\alpha$. Thus, the amplitude associated to the subharmonic 1/3 has to be prescribed in advance and held fixed, while the second amplitude follows as function of $A_{1/3}$. Since $A_{1/3} \neq 0$, equations (2.26)- (2.27) lead to

$$\omega^2 = 9\alpha + \frac{27}{4}\beta\left(A_{1/3}^2 + A_{1/3}A_1 + 2A_1^2\right),$$
$$-8\alpha A_1 = F + \left(\omega^2 - 9\alpha\right) - \frac{1}{4}\beta\left(A_{1/3}^3 + 6A_{1/3}^2A_1 + 3A_1^3\right),$$

and the elimination of ω^2 is the last equation yields

$$\omega^2 = 9\alpha + \frac{27}{4}\beta\left(A_{1/3}^2 + A_{1/3}A_1 + 2A_1^2\right), \tag{2.28}$$
$$-8\alpha A_1 = F - \frac{1}{4}\beta\left(A_{1/3}^3 - 21A_{1/3}^2A_1 - 27A_{1/3}A_1^2 - 51A_1^3\right). \tag{2.29}$$

With $A_{1/3}$ prescribed, $\beta = 0$, $\omega = 3\sqrt{\alpha}$ and $A_1 = -F/8\alpha = A$, the following relations represent the response curves in the subharmonic case:

$$\omega^2 = 9\alpha + \frac{27}{4}\beta\left(A_{1/3}^2 + A_{1/3}A + 2A^2\right), \tag{2.30}$$
$$A_1 = A + \frac{1}{32}\frac{\beta}{\alpha}\left(A_{1/3}^3 - 21A_{1/3}^2A - 27A_{1/3}A^2 - 51A^3\right). \tag{2.31}$$

The second relation (2.31) determines the value of the second Fourier coefficient for the subharmonic vibration as function of the first Fourier coefficient. When $A_{1/3} = 0$, the dependence (2.26) is satisfied and (2.27) is reduced to the relation for the harmonic case (2.21). This means that the subharmonic oscillation is the result of *bifurcation* from the harmonic solution. Bifurcations represent the changes in the dynamical behaviour of an oscillation, leading to topological changes in the phase-space. These changes involve either the birth and the destruction of attractors and the change in their size or shape. The bifurcation occurs when

$$A_1 = A - \frac{51}{32}\beta A^3, \tag{2.32}$$

with $A = -F/8\alpha$.

Figure 2.8 depicts the response curves when bifurcations arise. Equation (2.30) is an ellipse or hyperbola in the plane ω-$A_{1/3}$. The ω-A_1 plane is similar to the corresponding diagram in the harmonic case: if no subharmonics occur, relation (2.31) is the same of (2.21), but in this case point B in Figure 2.8 represents the chance that a bifurcation happen. Since the response curve in the ω-$A_{1/3}$ space is an hyperbola, the minimum (maximum) of ω when $\beta > 0$ ($\beta < 0$) is reached when $A_{1/3} = -A/2$, and the value is

$$\omega^2 = 9\left(\alpha + \frac{21}{16}\beta A^2\right). \tag{2.33}$$

Hence, the bifurcation exists only when the frequency is less or greater than its minimum or maximum:

$$\omega < 3\sqrt{\alpha + \frac{21}{16}\beta A^2}, \quad \beta < 0, \tag{2.34}$$

$$\omega > 3\sqrt{\alpha + \frac{21}{16}\beta A^2}, \quad \beta > 0. \tag{2.35}$$

Bifurcations and jump phenomena are highly related. The analysis developed so far has highlighted that more than one solution satisfying system (2.15) can exist, i.e. coexisting solutions may occur. The fixed points R and T in Figure 2.8 represents two stable branch of solutions, while the point S correspond to a saddle point. The latter divides the phase-space into two regions (Figure 2.9), each of them being a basin of attraction of the stable points R and T.

In this Section, only the presence of one subharmonics has been discussed, but other subharmonics or ultraharmonics can occur, leading the Duffing equation to become more and more complex. The qualitative change of bifurcation associated to jump phenomena leads the system to chaotic behaviour.

2.3. *Free Duffing oscillator*

In this Section, the effect of varying the parameter values of the Duffing oscillator is investigated. The *deterministic* Duffing model is

$$\ddot{x} + c\dot{x} + \alpha x + \beta x^3 = 0, \tag{2.36}$$

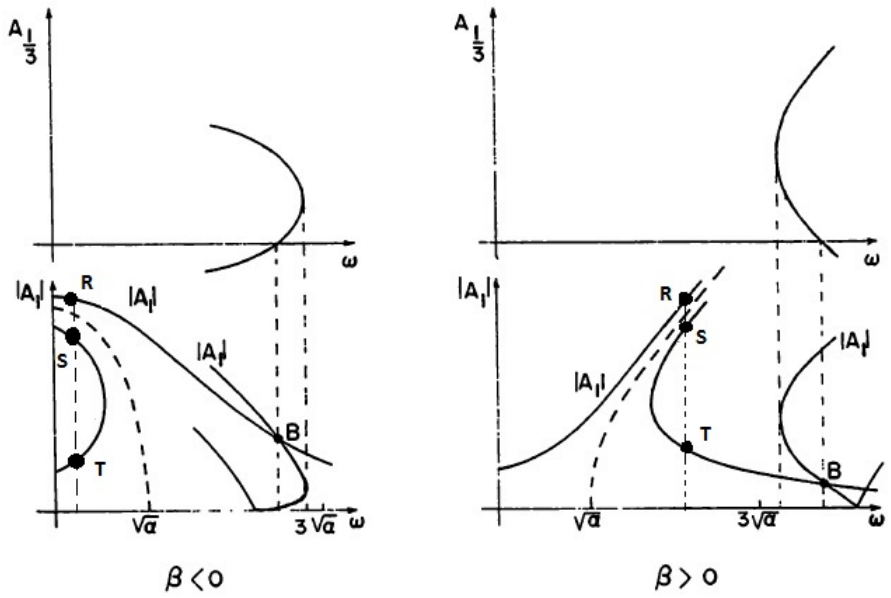

Figure 2.8: Response diagrams for subharmonic oscillations for hard ($\beta > 0$) and soft ($\beta < 0$) spring forces. The point B is the bifurcation point: here the subharmonic and harmonic oscillations are identical. The dotted lines around B highlight the bifurcation area: in this region, the amplitude A_1 can follow two different curves. The point R and T correspond to stable branch solutions, while the point S is a saddle. Both points R and T are coexisting solutions. Source: Figure p. 107 of Stoker (1950).

where, as previously stated, α is the natural frequency of the vibration, β is the mode of the restoring force (hard or soft spring), and c is the damping term. The parameters have the following effects. When α becomes negative, the system may diverge. In the case of $c < 0$, self-excited oscillations arise. The parameter β associated to the non-linear cubic stiffness affects the existence of *non-trivial*, i.e. non-zero, critical points and their stability.

2.3.1. Critical points

The critical points of equation (2.36) are evaluated as follows.

Equation (2.36) can be re-written in terms of state variables $x_1 = x$, $x_2 = \dot{x}$ as

$$\dot{x}_1 = x_2 \tag{2.37}$$

$$\dot{x}_2 = -\left(cx_2 + \alpha x_1 + \beta x_1^3\right). \tag{2.38}$$

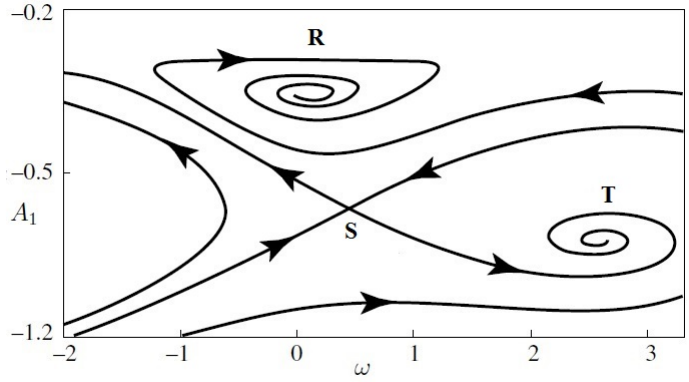

Figure 2.9: Phase-space for the frequency ω and the amplitude A_1. The arrows depict the direction of motion. The points R, T and S correspond to the ones of Figure 2.8. The saddle S separates the phase-space in domains of attraction. Source: Figure 5.9 of Kovacic and Brennan (2011).

In matrix form, equations (2.37)-(2.38) are

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{G}(\boldsymbol{x}), \tag{2.39}$$

where

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \boldsymbol{G}(\boldsymbol{x}) = \begin{bmatrix} g_1(x_1, x_2) \\ g_2(x_1, x_2) \end{bmatrix}, \tag{2.40}$$

and

$$g_1(x_1, x_2) = x_2, \quad g_2(x_1, x_2) = -\alpha x_1 - c x_2 - \beta x_1^3. \tag{2.41}$$

Fixing $\dot{x}_1 = \dot{x}_2 = 0$, gives the *equilibrium* equations

$$0 = x_{2st} \tag{2.42}$$

$$0 = -\left(c x_{2st} + \alpha x_{1st} + \beta x_{1st}^3\right), \tag{2.43}$$

where x_{1st} and x_{2st} denotes the critical points. At $x_{2st} = 0$,

$$\alpha x_{1st} + \beta x_{1st}^3 = x_{1st}\left(\alpha + \beta x_{1st}^2\right) = 0. \tag{2.44}$$

Hence, the determinant of (2.44) depends on the sign of the product $\alpha\beta$. If $\alpha\beta > 0$, only the trivial fixed point $(x_1,\ x_2) = (0,\ 0)$ exists. In the case $\alpha\beta < 0$ there are two

nontrivial fixed points in addition to the trivial one, that are $(x_1,\ x_2) = (-\sqrt{-\alpha/\beta},\ 0)$, and $(x_1,\ x_2) = (\sqrt{-\alpha/\beta},\ 0)$.

The stability of these points is evaluated by computing the eigenvalues of the system (2.40). The Jacobian $\boldsymbol{J}$ of $\boldsymbol{G}$ is

$$\boldsymbol{J} = \begin{bmatrix} 0 & 1 \\ -(\alpha + 3\beta x_{1st}^2) & -c \end{bmatrix}, \tag{2.45}$$

and the respective characteristic equation is

$$\lambda^2 + c\lambda + \alpha + 3\beta x_{1st}^2 = 0. \tag{2.46}$$

The roots of equation (2.46) are the eigenvalues λ_1 and λ_2 that determine the stability of the critical points.

In what follows, I refer to positive (negative) linear stiffness if $\alpha > 0$ ($\alpha < 0$), positive or negative non-linear stiffness when $\beta > 0$ or $\beta < 0$.

2.3.2. Existence of the trivial fixed point only

If $\alpha\beta > 0$, the non-trivial critical points do not exist and only the stability of the solution $(x_{1st}, x_{2st}) = (0, 0)$ has to be evaluated. In this case, the characteristic equation becomes

$$\lambda^2 + c\lambda + \alpha = 0. \tag{2.47}$$

If both α and β are positive, the damping c defines the kind of stability, while α affects the magnitude of the eigenvalues and β does not affect neither the stability nor the eigenvalues. Instead, if α and β are both negative, the trivial critical point is a saddle, since the eigenvalues are one positive and one negative independently on the value of c. Hence, the changes in c are crucial in the analysis of stability of the trivial fixed point only if $\alpha > 0$ and $\beta > 0$.

Unstable focus

When $\sqrt{\alpha} \leq c < 0$ the eigenvalues are complex conjugate with a positive real part, and diverging oscillations are produced. The trivial fixed point is an unstable focus, as shown in Figure 2.10. In this case, once c has been fixed, α represents the speed at which the system diverges: as more the frequecy increases, as the more the curves in the phase-space move away rapidly (Figure 2.11).

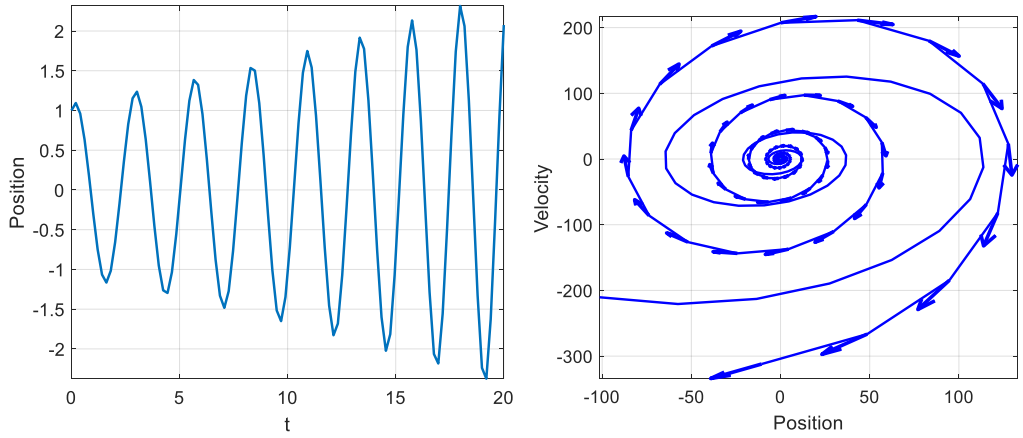

Figure 2.10: Oscillation and phace space representation for an unstable focus (complex conjugate eigenvalue with positive real part).

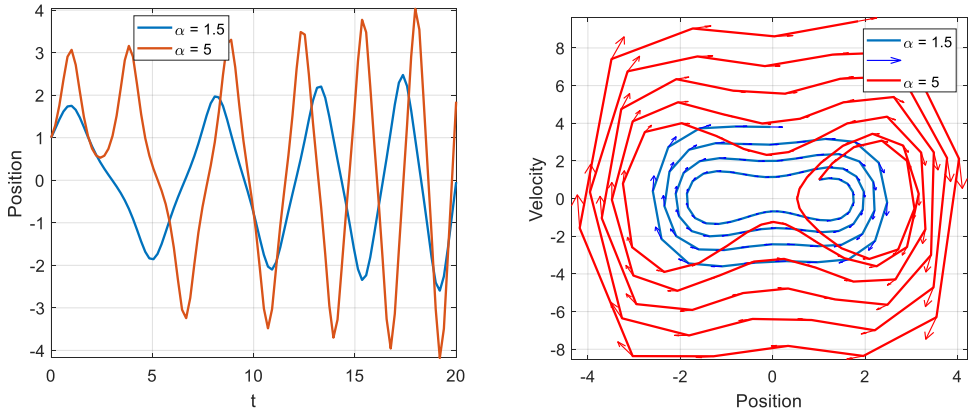

Figure 2.11: Unstable focus for different values of α, fixed β and c..

Stable focus

Stable focus arises when $0 < c \leq \sqrt{\alpha}$: the eigenvalues are complex conjugate with negative real part. The oscillation is damped; it approaches zero more slowly with increasing α. (Figures 2.12 and 2.13).

Centre

The origin is a centre if $c = 0$. The eigenvalues are pure imaginary complex conjugate, and the oscillations are self-sustained, as plotted in Figure 2.14.

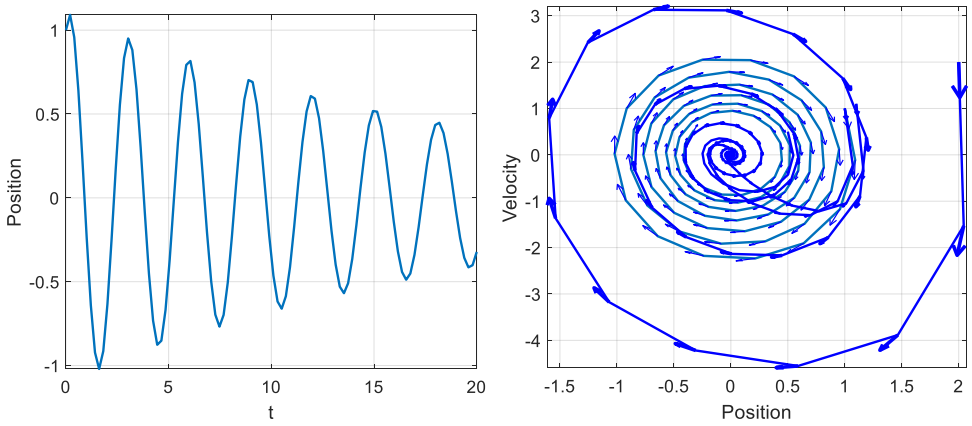

Figure 2.12: Oscillation and phace space representation in the case of stable focus (complex conjugate eigenvalue with negative real part).

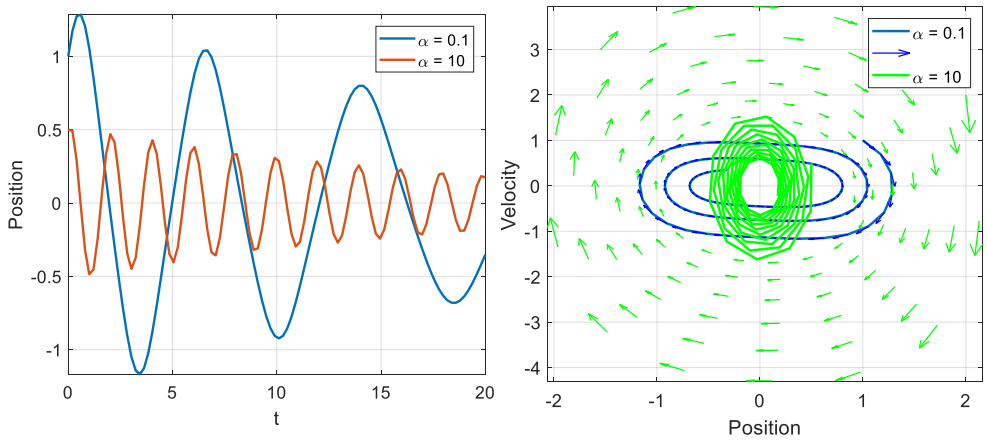

Figure 2.13: Stable focus in time and phase-space domain for different values of α, fixed β and c.

2.3.3. Case when the non-trivial fixed points exist

When $\alpha\beta < 0$, the non-trivial solutions $(x_1,\ x_2) = (-\sqrt{-\alpha/\beta},\ 0)$, and $(x_1,\ x_2) = (\sqrt{-\alpha/\beta},\ 0)$ exist as well as the trivial fixed point. In this case, $\alpha > 0$ and $\beta < 0$ or viceversa.

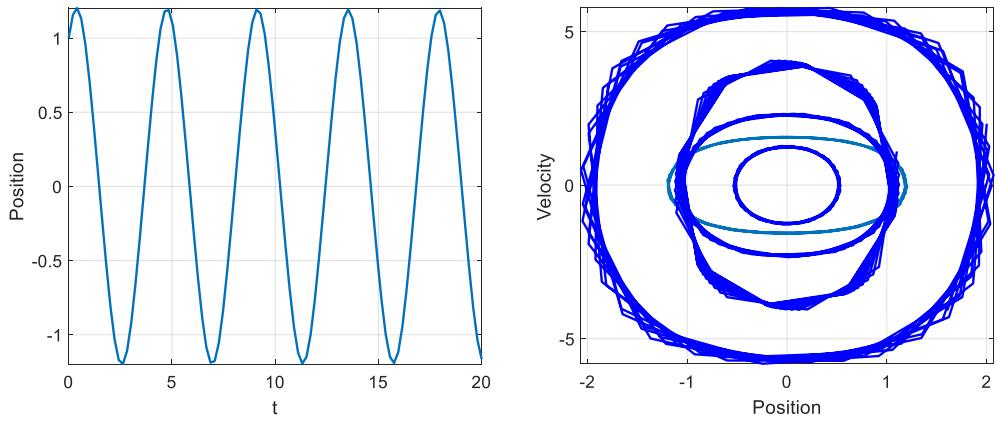

Figure 2.14: Oscillation and phase space representation for a centre (pure imaginary complex conjugate eigenvalue).

Positive α and negative β

The non-trivial fixed points are saddles, and the kind of stability of $(x_1\ x_2) = (0, 0)$ depends on c, as in Section 2.3.2. The parameter β influences the number of zero-crossings of the displacement. A small value of β let the oscillation to move away, so that x diverges from zero (Figures 2.15 - 2.16).

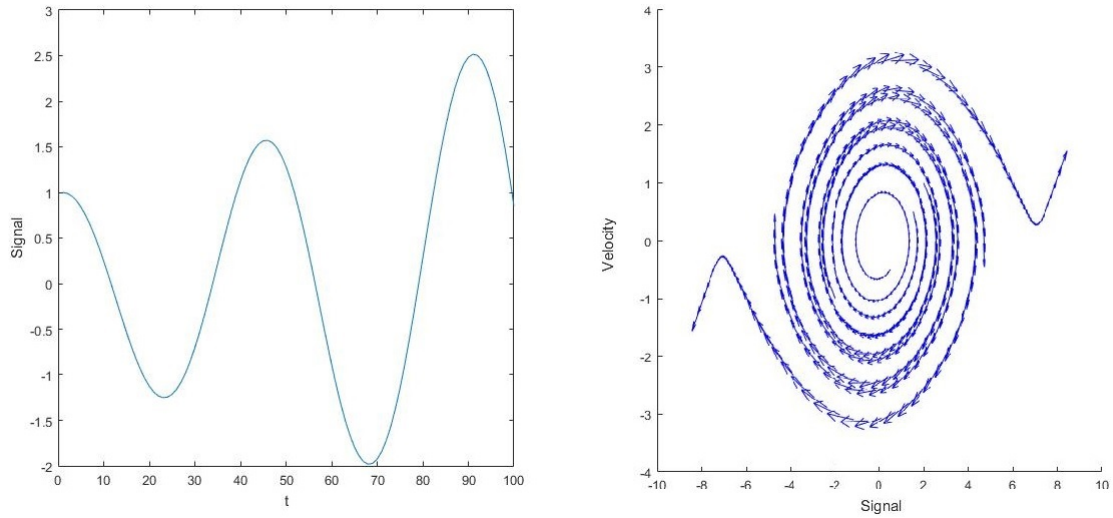

Figure 2.15: Oscillation and phace space representation for a saddle.

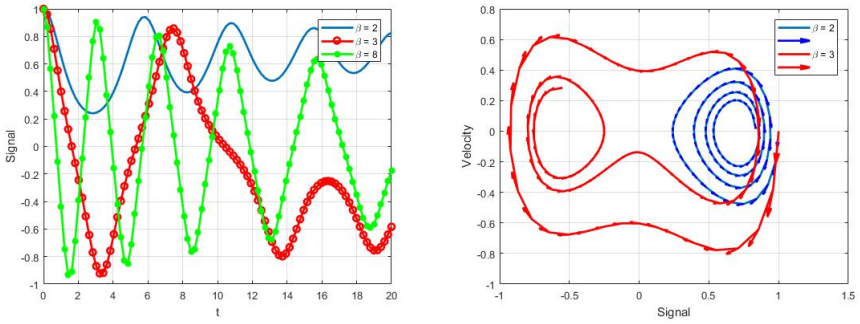

Figure 2.16: Time domain and phase-space portrait for different values of β, and c and α fixed.

Negative α and positive β

This case is the opposite of the previous one: here, the trivial fixed point is a saddle, while the stability of the non-trivial points depends on c.

2.4. *Forced oscillation*

The *stochastic* Duffing process is defined as

$$\ddot{x} + c\dot{x} + \alpha x + \beta x^3 = \sigma_\varepsilon \varepsilon_t \tag{2.48}$$

where ε_t is a stationary, zero-mean Gaussian white noise with variance σ_ε^2.

In this Section, the stability of the stochastic differential equation (SDE) (2.48) is analysed.

Equation (2.48) can be expressed in form of state variable, as previously done in (2.37)-(2.38): $x_1 = x$ and $x_2 = \dot{x}$,

$$\dot{x_1} = x_2, \tag{2.49}$$

$$\dot{x_2} = -\left(cx_2 + \alpha x_1 + \beta x_1^3\right) + \sigma_\varepsilon \varepsilon_t. \tag{2.50}$$

Taking the increments of the state variables, the (2.48) can be expressed as a stochastic differential equation in the first-order Itô form

$$d\dot{x}_1 = x_2 dt \tag{2.51}$$

$$d\dot{x}_2 = -\left(cx_2 + \alpha x_1 + \beta x_1^3\right) dt + \sigma_\varepsilon dW_t, \tag{2.52}$$

where dW_t represents the increment of a Brownian motion.

The stability of the critical points of system (2.51)-(2.52) are estimated in several ways in the literature. One of the possible strategies is to distinguish between the critical points coming from the underlying behaviour of the SDE and the ones arising from stochastic perturbations. Bifurcations in noisy systems may occur due to topological (i.e. dynamical) changes in the phase-space trajectories (as the bifurcations described in Section 2.1.2) or as a result of phenomenological changes associated with the probabilistic structure of the long term behaviour of the state variables (Kumar et al., 2016). The first type of bifurcation can be analysed computing the Lyapunov exponents (LE) (see Wolf et al., 1985 and Thomsen, 2013), while "stochastic" bifurcations can be evaluated through the Fokker-Planck equation (Kumar and Narayanan, 2010).

The LE Λ gives the rate of divergence or convergence, respectively for $\Lambda > 0$ and $\Lambda < 0$, of trajectories in phase space. In general, for a system with N first-order differential equations, exactly N LEs can be evaluated, with $\Lambda_1 \geq \Lambda_2 \geq \ldots \geq \Lambda_N$. Thus, for system (2.51)-(2.52), two LEs are computed. Qualitative changes in the nature of the Lyapunov exponents as the parameters of (2.51)-(2.52) vary are indicative of bifurcations. Assuming that $\boldsymbol{x}_0$ is a stable solution for (2.51)-(2.52), a small perturbation $\boldsymbol{u}$ to the solution $\boldsymbol{x}_0$ is governed by the linearised equation

$$\dot{\boldsymbol{u}} = \boldsymbol{J}(t)\boldsymbol{u}, \tag{2.53}$$

where the Jacobian $\boldsymbol{J}$ is evaluated at the solution $\boldsymbol{x}_0$. The Lyapunov exponents are defined as

$$\Lambda_i = \lim_{t\to\infty} \mathrm{E}\left[\frac{1}{t}\log\frac{\|\boldsymbol{u}(t)\|}{\|\boldsymbol{u}(0)\|}\right], \tag{2.54}$$

where the set $\{u(t), t > 0\}$ are the solution trajectories of linear differential equations when (2.51)-(2.52) is linearised around a solution $\boldsymbol{x}_t$, and $\|\cdot\|$ denotes the Euclidean norm. The largest Lyapunov exponent (LLE) indicates the stability of the dynamical system: a change in the sign of the LLE reflects a bifurcation. The presence of noise implies that the state space trajectories inherit the time-varying fluctuations of the system and hence the Lyapunov exponents can only be interpreted in terms of the long-term temporal mean (Kumar et al., 2016). The evaluation of the LEs is com-

putationally expensive, since it requires to solve for $\boldsymbol{u}_t$ coupled with (2.51)-(2.52). This have required the development of several numerical algorithms, described in Ramasubramanian and Sriram (2000) or Wolf et al. (1985).

Bifurcations due to changes in the probabilistic structure of the stationary joint probability density function (PDF) of the state variables may occur in a stochastically excited dynamical system. The joint PDF of the state variables represents a measure of the time spent by a solution in an area of the phase-space and gives an indication of the spatial extent of a stochastic attractor. The joint PDF does not explicitly take into account the system dynamics and its time evolution is governed by the Fokker-Planck (FP) equation:

$$\frac{\partial p}{\partial t} = -x_2 \frac{\partial p}{\partial x_1} - \frac{\partial}{\partial x_2}\left(\alpha x_1 + c x_2 - \beta x_1^3\right) + \frac{\sigma^2}{2}\frac{\partial^2 p}{\partial x_2^2}. \tag{2.55}$$

The computation of the FP equation is not straightforward due to discontinuity in the joint PDF and these difficulties may be overcome using numerical methods as described in Schenk-Hoppé (1996).

Figure 2.17 shows how, even for small variations of the level of variance, the system undergoes through instability.

2.5. *Disciplines influenced by the Duffing equation*

Duffing's original work is most devoted to the analysis of the pendulum. Some decades later, many engineering systems have been described by the Duffing's equation, such as beam and magnet systems, isolators and electrical circuits. For a discussion on these systems, the reader may refers to Kovacic and Brennan (2011). The attention of the scientific community on Duffing's equation started with James Stoker, who wrote in 1950 a seminal book on non-linear vibrations (Stoker, 1950). Following Stoker's book, the Duffing equation sat alongside the Van der Pol equation, one of the most known equations in non-linear vibrations. From 1970s, several papers related to the Duffing equation have been published, since in those years digital computers started to be used to solve differential equations. Figure 2.18 shows a survey carried out via SCOPUS to record the number of papers with the word

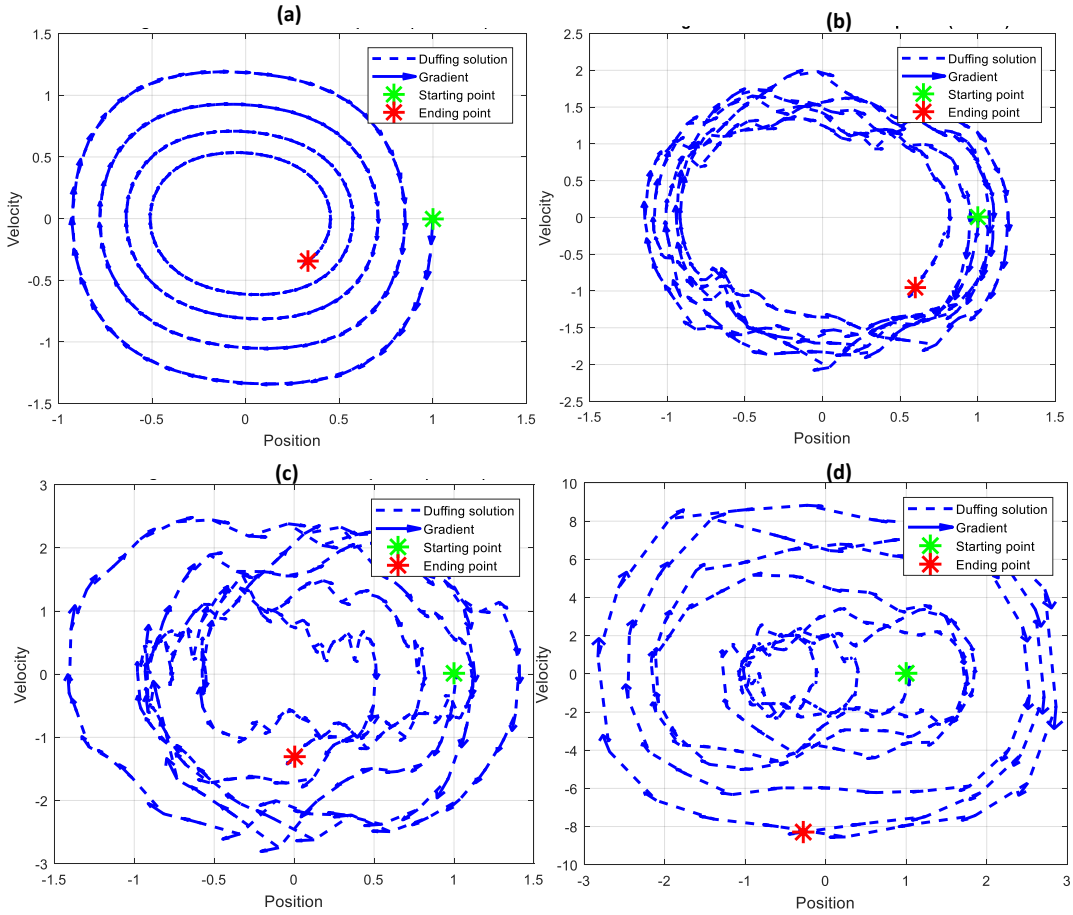

Figure 2.17: Variation of the dynamics of the stochastic Duffing process for different sizes of standard deviation σ_ε. (a) $\sigma_\varepsilon = 0.001$. (b) $\sigma_\varepsilon = 0.5$. (c) $\sigma_\varepsilon = 1.0$. (d) $\sigma_\varepsilon = 2.0$.

"Duffing" in the title, abstract or keywords, demonstrating the increasing interest by various communities. In 1976 Holmes and Rand (1976) published a paper on bifurcations of the Duffing's equation and its application to the catastrophe theory, while in 1980s Yoshisuke Ueda published his research on chaos (Ueda, 1979 and Ueda, 1985). Ueda stated that the study of the cubic Duffing's oscillator has inspired the discovery of chaotic behaviours.

In statistics and time series analysis, non-linear behaviours modelled by differential equations may be reproduced in terms of *difference* equations. However, mathematical explanations of complex phenomena in discrete time are not always easy

to construct (Tong, 1990). The main efforts in the reproduction of jump phenomena, amplitude-frequency dependence or subharmonics in the time series literature can be found in Tong and Lim (1980), Haggan and Ozaki (1981) and Ozaki (1981). These authors mainly studied non-linear oscillations through time series models such as Threshold Autoregressive (TAR). Figure 2.19 shows a simulation study of the amplitude-frequency relation coming from autoregressive models.

Concerning real data analysis, several applications in biology or economics of the non-linear phenomena described in Section 2.2 can be found. Ramsey (1990) applied the Duffing system to the reconstruction of the time series of the U.S. money supply, while Ball (2009) and Blanchard and Summers (1986) modelled unemployment rates through hysteresis. In biology, Tong and Lim (1980) studied Canadian lynx (the annual record of the number of the Canadian lynx trapped in the North-West Canada for the period 1821–1934) and sunspots data with models arising from the exploration of the Duffing oscillator. The above mentioned time series are characterized by asymmetric cycles with sharp and large peaks following relatively small trough. These kind of features of time series need to be deepened and the analysis of the Duffing process can shed some light on still unexplored behaviours.

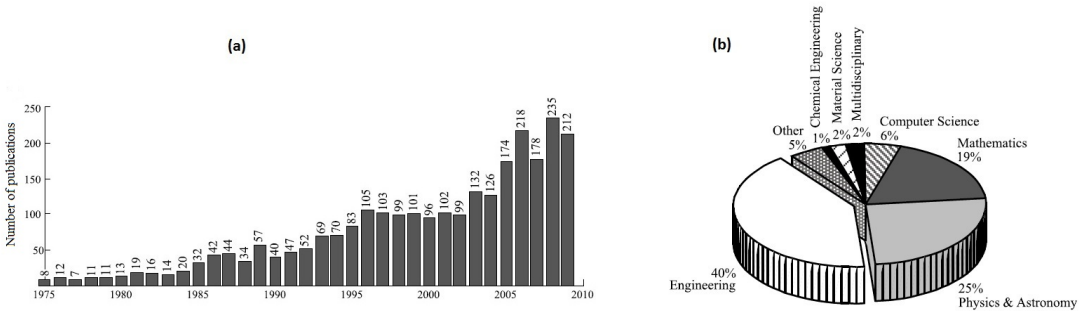

Figure 2.18: Publications with the word "Duffing" in the title, abstract or keyword. (a) Number of publications per years 1975-2009. (b) Percentage of publications per disciplines in the period 1950-2009. Survey carried out by Elsevier Scopus ^{TM} updated 30 March 2010. Source: Figures 1.4 and 1.5 of Kovacic and Brennan (2011).

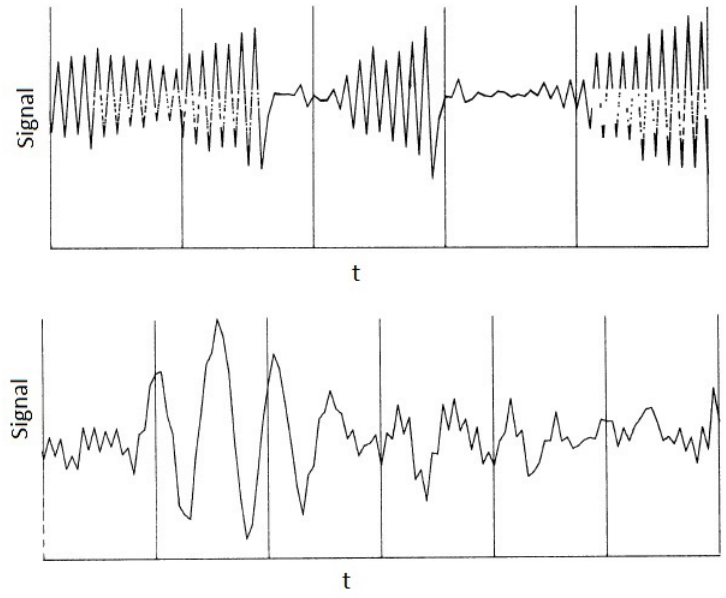

Figure 2.19: Simulation study of an amplitude-frequency relation in the time series literature. The top panel shows that a high amplitude is associated to high frequency, while, in the bottom panel, sharp peaks occur at small frequencies. Source: Figure 2.26 of Tong (1990).

3. Inference for ordinary differential equations

This Chapter focuses on inference for ordinary differential equations (ODEs); in particular, two approaches for signal and parameters estimation are discussed. The first concerns Kalman-filter based methods, while the second is an approximate Bayesian computation scheme.

Signal extraction

A general approach to extract the signal from noisy time series that hide unobserved components relies in a state space description.

A state space representation is a model that assumes that the observations (the data) depend on a hidden component (the state). The dynamics of the process evolves over time accordingly to a set of state space equations. A well established method for inference in linear Gaussian state space models is the Kalman filter (Kalman, 1960). The Kalman filter (KF) is an algorithm to perform exact (non-approximated) Bayesian filtering (inference) for linear systems with an additive white noise component. The KF is based on an iterative procedure of prediction and update and it requires the estimation of a Gaussian filter density; the latter represents the probability of the state given the previous observations. As Gaussian densities are fully described by the first two moments, the estimation of the filtering density needs the computation of the mean and the variance. As shown by Kalman and Bucy (1961), the time evolution of the mean (often called *estimation*) and variance (*estimation error*) of a Gaussian filter density can be expressed analytically. Hence, since everything is Gaussian, the prediction-correction structure and the likelihood function of the KF have a closed form solution. These properties let the KF yield unbiased and consistent estimates, but for the signal reconstruction in linear models only. For the estimation of parameters, even for linear systems, embodying the parameter dynamics in the state space description leads to non-linear state space equations. In the case of non-linear models, Dempster et al. (1979) and Cox (1964) warn about the direct use of the KF.

For inference in non-linear systems, several extensions of the KF have been proposed, as the extended KF (EKF, see e.g. Maybeck, 1979 and Ljung, 1979) and the unscented KF (UKF, for which the reader is referred to Julier and Uhlmann, 1997 and Julier and Uhlmann, 2004). Both EKF and UKF approximate the filter density necessary in the KF steps by Gaussian distributions. The two methods differ in the strategy of the computation of the approximations. On the one hand, EKF linearises the equations using a first or second order Taylor expansion and then applies the standard KF procedure. Thus, the EKF truncates the state space functions up to the second order moments and it requires the computation of derivatives. Such characteristics may lead the EKF to perform unconsistent and computationally expensive estimates. On the other hand, the UKF is based on the unscented transform (UT). The UT is a non-linear transformation of a probability density function: the transform takes a set of points from the density function and on these points evaluates a Gaussian approximation. The latter is the basis for the KF recursion steps (i.e. the time evolution of the filtering density). Some of the potentiality of the UKF are explored in Sitz et al. (2002).

For estimation of chaotic dynamical systems, the UKF approach is superior than EKF (Julier and Uhlmann, 2004). Indeed, the UKF, without involving derivatives, is suitable to manage non differentiability (as in the determination of excitation responses in engineering problems), and implicit forms of non-linearities (in example, in the analysis of artificial neural networks). Nevertheless, the UKF has two main limits. First, if the initial values of the algorithm are chosen far from the true unknown parameters, the convergence is not guaranteed. Second, the unscented transform is subject to the location of a deterministically chosen set of points (called sigma points). The position of sigma points, due to the complexity of the system, may be sub-optimal and mislead the Gaussian approximations. Therefore, the starting points of the iterative prediction-correction steps and the sigma points placement have a crucial impact on the overall inference results.

The KF and UKF algebraic expressions, as well as the discussion on initialization and sigma points position, are reviewed next.

Parameter estimation

To infer parameters of complex systems, one has to account for the possibility of mathematical intractability of real-word dynamics. In particular, as long as realistic models are developed (e.g. agent-based models or unobserved genealogy in population genetics), the likelihood functions are analytically unavailable or computationally costly. The class of approximate Bayesian computation (ABC), also called *likelihood-free* technique, has been developed to avoid the likelihood evaluation. The ABC method is a simulation-based procedure and in the last decade many efforts have been devolved to the development of this algorithm: the reviews of Beaumont (2010), Hartig et al. (2011), and Marin et al. (2012) give an insight of the increasing importance of this technique.

In a generic way, the ABC method produces random samples by means of a simulation and quantifies the distance between simulated and original data. The procedure can be briefly described as follows. A set of *candidate parameters* are picked from a prior distribution and a sample set is simulated according to a model. Then, the distance among simulated and original data is quantified depending on a distance function. In the case of "vicinity" between sample sets, the candidate parameters constitute a posterior distribution of parameter values.

Hence, the ABC approach crucially depends on the choice of a metric and on statistics that faithfully reproduce the key features of a given sample. Unfortunately, the quantification of a distance among datasets and the definition of the main characteristics of data are not a trivial task. Below, a more detailed description of ABC-based algorithms along with a review of distance functions are presented.

Inferring both signal and parameters

After considering the advantages and limitations of filtering methods and likelihood free algorithms, my strategy is to estimate signal and parameters of a state space model in the UKF framework with the novelty of using the ABC sampler and the optimization of sigma points location within UKF. The aim is to provide an algorithm that keeps safe the merits of the UKF and ABC and overcomes their disadvantages. This method is called Sequential ABC-UKF.

This Chapter is organized as follows. The mathematical description of state space models in presented in Section 3.1. Reviews of the Kalman filtering and UKF are in Sections 3.2 - 3.5, while the optimization problem is explained in Section 3.6. In Section 3.7 the ABC method is described. The novel algorithm that matches ABC within UKF is shown in Section 3.8.

3.1. *State space representation*

A state space model is described as

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{F}\left(\boldsymbol{x}(t), \boldsymbol{\lambda}, \boldsymbol{\varepsilon}(t)\right), \qquad \boldsymbol{y}(t) = \boldsymbol{H}\left(\boldsymbol{x}(t)\right) + \boldsymbol{\eta}(t), \tag{3.1}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^d$ is the latent state that evolves over time accordingly to the transition function $\boldsymbol{F}$, while $\boldsymbol{y}(t) \in \mathbb{R}^D$ is the observation at time t and $\boldsymbol{H}$ is the measurement function that maps the state $\boldsymbol{x}(t)$ to the data $\boldsymbol{y}(t)$. The parameter vector is $\boldsymbol{\lambda}$, embodied into the transition function $\boldsymbol{F}$. The system and observation noise, respectively, $\boldsymbol{\varepsilon}(t) \sim N(0, \Sigma_\varepsilon)$ and $\boldsymbol{\eta}(t) \sim N(0, \Sigma_\eta)$, are independently and identically Gaussian distributed over time. Both $\boldsymbol{\varepsilon}(t)$ and $\boldsymbol{\eta}(t)$ are mutually independent and independent from $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ for $t = 1, \ldots, T$.

In a general setting, the function $\boldsymbol{F}$ of model (3.1) is non-linear and $\boldsymbol{\lambda}$ can be a time-varying vector, but, for ease of notation, the time dependence is omitted. The stochastic term $\boldsymbol{\varepsilon}(t)$ represents the rapid fluctuations of the hidden dynamics so that the state $\boldsymbol{x}(t)$ becomes a random variable. The measurement noise, instead, incorporates the distortions occurred during the observation process.

Since in real-world applications measurements have a finite time interval, the time-continuous model (3.1) has to be discretized. The state space equations are transformed into the difference equations

$$\boldsymbol{x}_t = \boldsymbol{f}\left(\boldsymbol{x}_{t-\Delta t}, \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_t\right), \qquad \boldsymbol{y}_t = \boldsymbol{h}\left(\boldsymbol{x}_t\right) + \boldsymbol{\eta}_t, \tag{3.2}$$

where $\Delta t > 0$ is the sampling time step and $\boldsymbol{f}$ and $\boldsymbol{h}$ are, respectively, the time-discrete transition and observation functions. The $\boldsymbol{x}_{t-\Delta t}$ and $\boldsymbol{x}_t$ are the discretized states, while $\boldsymbol{y}_t$ are the measurements. The discrete process and observation noise are represented by $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$, respectively.

Equation (3.2) describes a *discrete state space model*. As in the continuous time, $\boldsymbol{x}_t$ and $\boldsymbol{y}_t$ are random variables: they create a set of time-discrete stochastic processes $\boldsymbol{x} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_t\}$ and $\boldsymbol{y} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_t\}$, for $t_1 = 1, t_2 = 2, \ldots, t_N = t$. For notational convenience, the time step is setted to $\Delta t = 1$, unless otherwise specified.

The statistical properties of processes (3.1) and (3.2) are fully described by the joint probability density $p_{XY}(\boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Y} = \boldsymbol{y})$.

In this Chapter, as a deterministic autonomous system is considered, the process noise is $\boldsymbol{\varepsilon}_t = 0$. The transition function $\boldsymbol{f}$ is defined as

$$\boldsymbol{f}\left(\boldsymbol{x}_{t-\Delta t}, \boldsymbol{\lambda}\right) = \boldsymbol{x}_{t-\Delta t} + \int_{t-\Delta t}^{t} \left(\boldsymbol{F}(\boldsymbol{x}(T), \boldsymbol{\lambda})\, dT. \right. \tag{3.3}$$

Equation (3.3) is not analytically tractable and it has to be evaluated through numerical approximation methods.

The stochastic version of models (3.1) and (3.2) are described in Chapter 5.

3.2. *The Kalman filter*

This Section describes the Kalman filter iterative scheme and the notation of Sitz et al. (2002) is followed.

In the Kalman filter framework, the extraction of the signal from noisy time series consists in the estimation of a time-varying *filter density* $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$. Consider the linear and Gaussian state space model:

$$\boldsymbol{x}_t = \boldsymbol{f}\boldsymbol{x}_{t-1} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{y}_t = \boldsymbol{h}\boldsymbol{x}_t + \boldsymbol{\eta}_t. \tag{3.4}$$

The Kalman filter performs inference of model (3.4) by estimating the evolution over time of the mean and the covariance of the states and observations in an unbiased and consistent way (for the mathematical proof of consistency consider Kalman and Bucy, 1961). The filtering method consists of a prediction and a correction step. The *prediction* step predicts the observed space $p(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1})$ using the *predictive distribution* $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1})$:

$$p\left(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1}\right) = \int p\left(\boldsymbol{y}_t|\boldsymbol{x}_t\right) p\left(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1}\right) d\boldsymbol{x}_t. \tag{3.5}$$

In other words, at each time step t, the mean of the filter density $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$ is extrapolated thanks to the information of prior observations. In order to obtain a prediction for both the state and observation at the next time point, the KF computes the conditional expectations:

$$\hat{\boldsymbol{x}}\,(t|t-1) = \boldsymbol{m}\,(t|t-1) = \mathrm{E}\,[\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1}] = \mathrm{E}\,[\boldsymbol{f}\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1}]\,, \tag{3.6}$$

$$\hat{\boldsymbol{y}}\,(t|t-1) = \boldsymbol{m}_y\,(t|t-1) = \mathrm{E}\,[\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1}] = \mathrm{E}\,[\boldsymbol{h}\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1}]\,. \tag{3.7}$$

These expectations, due to the linearity of the system, have a closed form solution. The solution for the state is

$$\boldsymbol{m}\,(t|t-1) = \mathrm{E}[\boldsymbol{f}\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t}] = \boldsymbol{f}\mathrm{E}[\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t}] = \boldsymbol{f}\boldsymbol{m}(t-1|t-1), \tag{3.8}$$

where $\hat{\boldsymbol{x}}(t-1|t-1) = \boldsymbol{m}(t-1|t-1)$ is the mean of the filtering density at time step $t-1$. The covariances are:

$$\boldsymbol{P}(t|t-1) = \mathrm{E}[(\boldsymbol{x}_t - \boldsymbol{m}(t|t-1))(\boldsymbol{x}_t - \boldsymbol{m}(t|t-1))'\,|\boldsymbol{y}_{1:t}], \tag{3.9}$$

$$\boldsymbol{P}_y(t|t-1) = \mathrm{E}[(\boldsymbol{y}_t - \boldsymbol{m}_y(t|t-1))(\boldsymbol{y}_t - \boldsymbol{m}_y(t|t-1))'\,|\boldsymbol{y}_{1:t}], \tag{3.10}$$

$$\boldsymbol{P}_{xy}(t|t-1) = \mathrm{E}[(\boldsymbol{x}_t - \boldsymbol{m}(t|t-1))(\boldsymbol{y}_t - \boldsymbol{m}_y(t|t-1))'\,|\boldsymbol{y}_{1:t}], \tag{3.11}$$

where the symbol $'$ denotes transposition. The *correction* step updates the prediction and the estimation of the errors when a new observation arrives into the system. Essentially, the KF updates the estimation of the filtering density $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$ applying the Bayes rule:

$$p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) \propto p(\boldsymbol{y}_t|\boldsymbol{x}_t)p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1}). \tag{3.12}$$

Since model (3.4) is a Gaussian linear system, the posterior distribution $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$ is a Gaussian with the following posterior mean and precision (the proof can be found in Theorem 4.4.1 of Murphy, 2012):

$$\boldsymbol{m}(t|t) = \boldsymbol{P}(t|t)\boldsymbol{h}\Sigma_\eta^{-1}\boldsymbol{y}_t + \boldsymbol{P}(t|t)\boldsymbol{P}^{-1}(t|t-1)\boldsymbol{m}(t|t-1), \tag{3.13}$$

$$\boldsymbol{P}(t|t)^{-1} = \boldsymbol{P}^{-1}(t|t-1) + \boldsymbol{h}'\Sigma_\eta^{-1}\boldsymbol{h}. \tag{3.14}$$

The posterior precision, using the matrix inversion lemma (Corollary 4.3.1 of Murphy, 2012), can be rewritten as

$$\boldsymbol{P}(t|t) = \boldsymbol{P}(t|t-1) - \boldsymbol{K}_t\boldsymbol{P}_y(t|t-1)\boldsymbol{K}_t', \tag{3.15}$$

where $\boldsymbol{K}_t$ is the Kalman gain matrix, defined as

$$\boldsymbol{K}_t = \boldsymbol{P}_{xy}(t|t-1)\boldsymbol{P}_y^{-1}(t|t-1). \tag{3.16}$$

The posterior mean, after some algebra manipulations described in Murphy (2012), is given by

$$\boldsymbol{m}(t|t) = \boldsymbol{m}(t|t-1) + \boldsymbol{K}_t(\boldsymbol{y}_t - \boldsymbol{m}_y(t|t-1)). \tag{3.17}$$

Looking at the updating equations (3.15) - (3.17), notice that the mean is the old mean plus a correction factor, that is the Kalman gain times a measurement residual. Hence, the Kalman gain weights the innovation $\boldsymbol{y}_t - \boldsymbol{m}_y(t|t-1)$. Let us consider the ratio $\boldsymbol{P}(t|t-1)\boldsymbol{P}_y^{-1}(t|t-1)$ that stresses the relation at time t between the prior knowledge of the state (up to time $t-1$), $\boldsymbol{x}_t - \boldsymbol{m}(t|t-1)$, and the observation residual. The ratio decreases as far as the denominator is high or the numerator is small. Since the difference between the state (the observations) and the mean $\boldsymbol{m}(t|t-1)$ (or $\boldsymbol{m}_y(t|t-1)$) represents the amount of noise in the prediction-correction steps, a high denominator indicates noisy measurements, while a small numerator points out little noise in the state prediction, thus a "strong" prior knowledge. In this case, a small correction to the innovation is sufficient because the noise variance $\boldsymbol{P}_{xy}(t|t-1)$ is low. On the contrary, a big $\boldsymbol{K}_t$ represents a bigger correction to the innovation, necessary in the case of weak prior knowledge of the state or high measurement precision.

3.2.1. Practical implementation of the KF

There are two dominant computational costs in the Kalman filter: the matrix inversion to compute the Kalman gain matrix, which takes $O(|\boldsymbol{y}_t|^3)$ time, and the matrices multiplication to compute $\boldsymbol{P}(t|t-1)$, which takes $O(|\boldsymbol{x}_t|^2)$ time. When the latter cost dominates (e.g. robotic mapping), sometimes sparse approximations are used (see Thrun et al., 2006). Instead, in the cases where $|\boldsymbol{y}_t| \gg |\boldsymbol{x}_t|$, the matrix $\boldsymbol{K}_t$ can be precomputed, since (surprisingly!) it does not depend on the actual observations $\boldsymbol{y}_{1:t}$ (an unusual property that is specific to linear Gaussian systems).

In practice, for reasons of numerical stability, more sophisticated implementations of the Kalman filter should be used. Further details can be found in various books, such as Simon (2006) and Murphy (2012).

3.2.2. Main disadvantage

The main limit of the KF concerns the choice of the starting values of the algorithm. Indeed, if the method is initialized far from the true values, the Gaussian distributions in the KF recursion steps may be affected by a high variance and a mean not centered around the signal at time t. In such a case, the KF can result in a very poor prediction. The initialization-dependence of the KF affects every algorithm based on the prediction-correction procedure: this limit and the proposed solution of this dissertation are discussed in Section 3.5.4.

3.3. *Parameter inference in the context of Kalman filtering*

Section 3.2 describes the KF steps to estimate the hidden signal $\boldsymbol{x}_t$. In analogy with the KF estimation strategy, performing parameter inference means that the vector $\boldsymbol{\lambda}$ has to considered a hidden state, with its own trajectory over t to be inferred. If the parameters represent a hidden component as well as the signal, they have to be included into the state function $\boldsymbol{f}$ of equation (3.2). The embodying of ODE parameters into the state vector breaks the linear dynamics assumed in equation (3.4) among $\boldsymbol{x}_t$, $\boldsymbol{x}_{t-1}$ and $\boldsymbol{y}_t$. Indeed, the inclusion of parameters with different behaviours among the state and space equations violates the linearity assumption. Figure 3.1 gives a simple representation of a state space model and the KF steps: behind each measurement, a hidden component is considered. The linear dynamics among unobserved variables over time, $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$, and the relationship between $\boldsymbol{x}_t$ and the observations, $p(\boldsymbol{y}_t|\boldsymbol{x}_t)$, is represented by arrows. The filtering density $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$ is used to make a prediction on the next point in the space. Once a new measurement is available, the prediction is corrected in the update step.

To estimate the signal and parameters from model (3.2) in the case of non-linear functions, approximate inference is necessary; approximate methods are reviewed next, with emphasis on the UKF.

3.4. *Kalman filter-based methods*

The approximate inference algorithms discussed below approximate the posterior distribution of the KF by a Gaussian. In general, if $y = f(x)$, where x has a

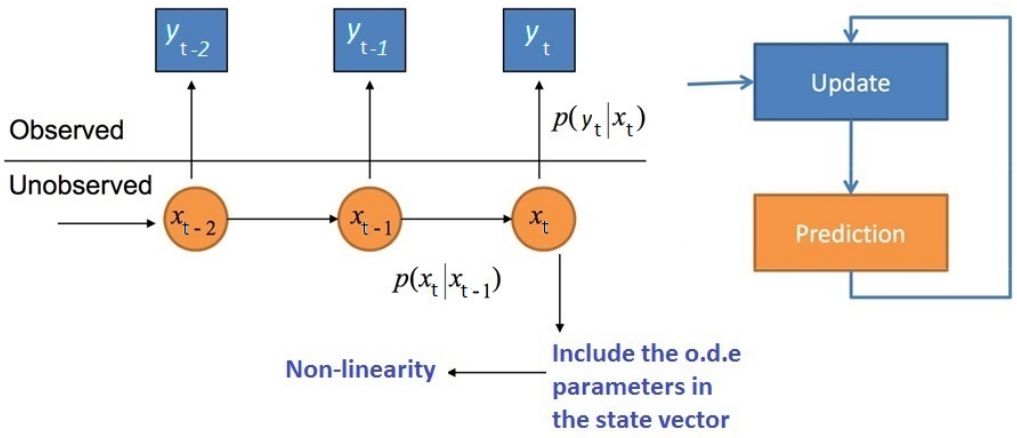

Figure 3.1: Illustration of a state space model within the KF scheme. Behind observations $\boldsymbol{y}_t$ there is a hidden component $\boldsymbol{x}_t$. As time goes on, a new measurement enters into the system and corrects the predictions on the future state $\boldsymbol{x}_{t+1}$. The inclusion into the state $\boldsymbol{x}_t$ of the ODE parameters corrupts the linearity of the model. Source: re-editing of a Wikipedia's figure on Bayesian filtering.

Gaussian distribution and f is a non-linear function, there are two main ways to approximate $p(y)$ by a Gaussian. The first is to use a first-order approximation of f. The second is to use the exact f, but to project $f(x)$ onto the space of Gaussians (Murphy, 2012). The EKF is based on the first method of approximation, while the UKF relies on the second one.

The EKF can be applied to non-linear Gaussian dynamical systems with differentiable state space functions. The idea is to linearise $\boldsymbol{f}$ and $\boldsymbol{h}$ about the previous state estimate using a first order Taylor series expansion, and then apply the standard Kalman filter equations. The intuition behind the EKF approach is shown in Figure 3.2, which shows what happens when one passes a Gaussian distribution $p(x)$ (on the bottom right panel), through a non-linear function $y = f(x)$ (top right). The resulting distribution is shown in the shaded gray area in the top left corner. The best Gaussian approximation to this, computed from $\mathrm{E}[f(x)]$ and $\mathrm{Var}[f(x)]$, is shown by the solid black line. The EKF approximates this Gaussian as follows: it linearises the $\boldsymbol{f}$ function at the current mode, μ, and then passes the Gaussian distribution $p(x)$ through this linearised function. In this example, the result is quite a

good approximation to the first and second moments of $p(y)$. However, there are two cases when the EKF works poorly: (i) when the prior covariance is large and (ii) when the function is highly non-linear near the current mean. In the first case the prior distribution is broad, so the EKF sends a lot of probability mass through different parts of the function far from the mean, that is the point where the function has been linearised. In the second case, much information is lost cutting the $\boldsymbol{f}$ around the mean; in this case, the very assumption of the EKF (i.e. the reliability of the first-order Taylor series approximation) gets worse as one moves far from the mean.

In both of these settings, the unscented Kalman filter is a better version of the EKF (Julier and Uhlmann, 1997). The key intuition behind the UKF is that it is easier to approximate a Gaussian than to approximate a function (Murphy, 2012). In other words, there is not a linear approximation to $\boldsymbol{f}$, but a deterministically chosen set of points, known as *sigma points*, pass through the function, and then a Gaussian is fitted to the resulting transformed points. This scheme is called the unscented transform, and it is sketched in Figure 3.3. The mathematical details are given below.

Basically, the UKF has two main merits. First, using the *whole density* of $\boldsymbol{x}_t$, the UKF method does not truncate the functions of model (3.2) but the filter density to its higher order moments. Second, due to the construction of a pre-defined set of points, the UKF can fit the whole state density on a sample small in size that is not computationally expensive. Consider that both the UKF and EKF perform $O(d^3)$ operations per time step. Nevertheless the UKF is accurate to at least second order, whereas the EKF is only a first order approximation.

3.5. *Unscented Kalman Filter*

In this Section, the unscented transform is explained first, and then the whole UKF iterations are discussed.

3.5.1. The unscented transform

As stated above, the UT transforms the probability density function through the set of sigma points. The sigma points are not drawn at random but they are

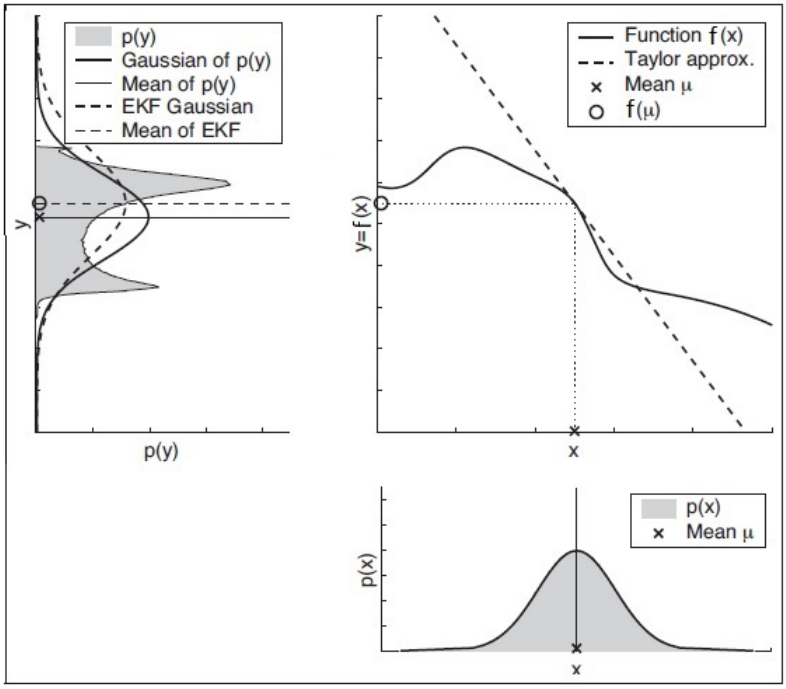

Figure 3.2: Illustration of a non-linear transformations of a Gaussian. The prior $p(x)$ is shown on the bottom right. The function $y = f(x)$ is on the top right. A non-linear transform $f(x)$ induces a complex distribution, while a linear function creates a Gaussian distribution. The transformed distribution $p(y)$ is shown in the top left. The dotted line is the EKF approximation, the solid line is the best Gaussian approximation to $p(y)$. Source: Figure 3.4 of Thrun et al. (2006).

deterministically chosen so that they exhibit certain specific properties. In particular, the sigma points are chosen so that their mean and covariance is the same of the state, $\boldsymbol{m}(t|t-1)$ and $\boldsymbol{P}(t|t-1)$. In this way, high-order information about the distribution can be captured with a fixed, small number of points (Julier and Uhlmann, 2004). A symmetric set of $2d+1$ points, $\boldsymbol{\chi}(t-1|t-1)$, where d is the dimension of the system, that satisfies the above conditions and lies in the d-th covariance contour is derived in Julier and Uhlmann (2004).

In the Kalman filter-based technique, the sigma points location is parametrised

by three scalar values $\boldsymbol{\theta} = (\alpha_{\text{ukf}}, \beta_{\text{ukf}}, k_{\text{ukf}})$ and are given by

$$\chi_0\left(t-1|t-1\right) = \boldsymbol{m}\left(t-1|t-1\right), \tag{3.18}$$

$$\chi_i\left(t-1|t-1\right) = \left\{\boldsymbol{m}\left(t-1|t-1\right) + \sqrt{(d+\lambda_{\text{ukf}})\,\boldsymbol{P}_{:i}}\right\}_{i=1}^{d}, \tag{3.19}$$

$$\chi_{i+d}\left(t-1|t-1\right) = \left\{\boldsymbol{m}\left(t-1|t-1\right) - \sqrt{(d+\lambda_{\text{ukf}})\,\boldsymbol{P}_{:i}}\right\}_{i=1}^{d}, \tag{3.20}$$

for $i = 1, \ldots, d$, $\lambda_{\text{ukf}} = \alpha_{\text{ukf}}^2(d + k_{\text{ukf}}) - d$ is a scaling parameter, $\boldsymbol{P}_{:i}$ represents the i-th column of matrix $\boldsymbol{P}$.

UKF passes the sigma points through $\boldsymbol{f}$ and $\boldsymbol{h}$ to obtain two new sample sets,

$$\chi_i\left(t|t-1\right) = \boldsymbol{f}\left(\chi_i(t-1|t-1)\right), \tag{3.21}$$

$$\Upsilon_i\left(t|t-1\right) = \boldsymbol{h}\left(\chi_i(t|t-1)\right). \tag{3.22}$$

The means and covariances for the measurements, derived in Murphy (2012), are

$$\boldsymbol{m}_y(t|t-1) = \sum_{i=0}^{2d} w_m^{(i)} \Upsilon_i(t|t-1), \tag{3.23}$$

$$\boldsymbol{P}_y(t|t-1) = \sum_{i=0}^{2d} w_c^{(i)} \left(\Upsilon_i(t|t-1) - \boldsymbol{m}_y(t|t-1)\right)\left(\Upsilon_i(t|t-1) - \boldsymbol{m}_y(t|t-1)\right)', \tag{3.24}$$

where the weights w are given by

$$w_m^{(0)} = \frac{\lambda_{\text{ukf}}}{d+\lambda_{\text{ukf}}}, \tag{3.25}$$

$$w_c^{(0)} = \frac{\lambda_{\text{ukf}}}{d+\lambda_{\text{ukf}}} + (1 - \alpha_{\text{ukf}}^2 + \beta_{\text{ukf}}), \tag{3.26}$$

$$w_m^{(i)} = w_c^{(i)} = \frac{1}{2(d+\lambda_{\text{ukf}})}, \quad \text{for } i = 1, \ldots, 2d. \tag{3.27}$$

3.5.2. The UKF method

The UKF uses the unscented transform twice, once to approximate passing the system through model $\boldsymbol{f}$, and once to approximate through the measurement model $\boldsymbol{h}$.

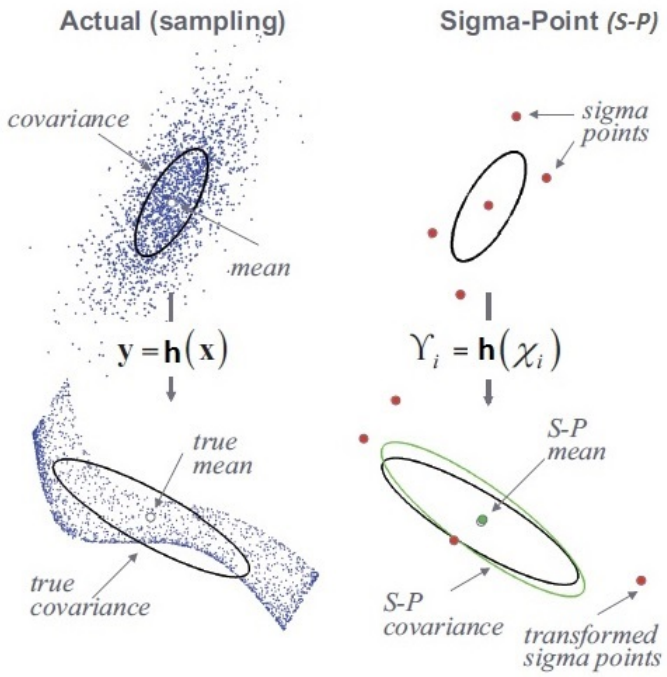

Figure 3.3: An example of the unscented transform in two dimensions. Instead of considering all the dataset size, the unscented transform computes the mean and the covariance only on the small sample of sigma points. Source: Figure 18.10 of Murphy (2012).

For the state estimation, the mean and the covariance are

$$\boldsymbol{m}(t|t-1) = \sum_{i=0}^{2d} w_m^{(i)} \chi_i(t|t-1), \tag{3.28}$$

$$\boldsymbol{P}(t|t-1) = \sum_{i=0}^{2d} w_c^{(i)} \left(\chi_i(t|t-1) - \boldsymbol{m}(t|t-1)\right)\left(\chi_i(t|t-1) - \boldsymbol{m}(t|t-1)\right)'. \tag{3.29}$$

Finally, the covariance between state and measurement is straightforward

$$\boldsymbol{P}_{xy}(t|t-1) = \sum_{i=0}^{2d} w_c^{(i)} \left(\chi_i(t|t-1) - \boldsymbol{m}(t|t-1)\right)\left(\Upsilon_i(t|t-1) - \boldsymbol{m}_y(t|t-1)\right)'. \tag{3.30}$$

The set of equations (3.18) - (3.30) describes the *unscented transform* for the UKF. Thus, the predictive distribution $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1}) \approx N(\boldsymbol{x}_t|\boldsymbol{m}(t|t-1), \boldsymbol{P}(t|t-1))$ is approximated by the old belief state $N(\boldsymbol{x}_{t-1}|\boldsymbol{m}(t-1|t-1), \boldsymbol{P}(t-1|t-1))$, while the likelihood

is $p(\boldsymbol{y}_t|\boldsymbol{x}_t) \approx N(\boldsymbol{y}(t|t-1)|\boldsymbol{m}_y(t|t-1), \boldsymbol{P}(t|t-1))$. The iterations of the UKF are based on the recursion of the KF (see equations (3.15) - (3.17)). The marginal likelihood of observations is defined as

$$p(\boldsymbol{y}_{1:T}|\boldsymbol{\theta}) = \prod_{t=1}^{T} p(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1},\ \boldsymbol{\theta}) = \prod_{t=1}^{T} N(\boldsymbol{y}_t|\boldsymbol{m}(t|t-1),\ \boldsymbol{P}(t|t-1)), \tag{3.31}$$

where the terms in the product are computed recursively as

$$p(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1},\ \boldsymbol{\theta}) \approx N(\boldsymbol{y}_t|\boldsymbol{m}_y(t|t-1), \boldsymbol{P}_y(t|t-1), \boldsymbol{\theta}) \times N(\boldsymbol{x}_t|\boldsymbol{m}(t|t-1), \boldsymbol{P}(t|t-1), \boldsymbol{\theta}). \tag{3.32}$$

3.5.3. Parameter estimation

The method described in Sections 3.2 and 3.5.1 mainly concerns the signal extraction of a dynamical process. To handle with parameter estimation, as suggested in Sitz et al. (2002), the parameter vector is treated through the evolution equation

$$\boldsymbol{\lambda}_t = \boldsymbol{\lambda}_{t-1}. \tag{3.33}$$

The parameter $\boldsymbol{\lambda}$ is constant in the state dynamics but it is updated at each measurement step since the value at time t converges to the true values. As for the parameter vector, to estimate the process and observation noise, the latter two are treated as state variables. A joint state vector $\boldsymbol{j}_t$ with the following time dynamics is defined

$$\boldsymbol{j}_t = \begin{pmatrix} \boldsymbol{x}_t \\ \boldsymbol{\lambda}_t \\ \boldsymbol{\varepsilon}_t \\ \boldsymbol{\eta}_t \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}(\boldsymbol{x}_{t-1}, \boldsymbol{\lambda}_{t-1}) + \boldsymbol{\varepsilon}_{t-1} \\ \boldsymbol{\lambda}_{t-1} \\ \boldsymbol{\varepsilon}_{t-1} \\ \boldsymbol{\eta}_{t-1} \end{pmatrix} = \boldsymbol{f}^j(\boldsymbol{j}_{t-1}), \tag{3.34}$$

and measurement function

$$\boldsymbol{y}_t = \boldsymbol{h}^j(\boldsymbol{j}_t) = \boldsymbol{h}(\boldsymbol{x}_t) + \boldsymbol{h}^\eta(\boldsymbol{\eta}_t). \tag{3.35}$$

If not stated otherwise, in the following additive and uncorrelated observation noise is considered, such that $\boldsymbol{h}^\eta(\boldsymbol{\eta}_t)$ is the identity function.

3.5.4. Disadvantages of the UKF

As previously described, the unscented transform does not require the analytic evaluation of derivatives (it is called a *derivative free filter*), making it simpler to implement and more widely applicable than the EKF, at the price that it highly relies on the sigma points position. Moreover, as discussed in Section 3.2.2, the UKF is affected by the choice of the starting values of the algorithm. Both these UKF limits are detailed below.

Sigma points location

The values of the sigma points parameter $\boldsymbol{\theta} = (\alpha_{\text{ukf}}, \beta_{\text{ukf}}, k_{\text{ukf}})$ are heuristically set by the algorithm. The default recommended values for the parameters are $\alpha_{\text{ukf}} = 1$, $\beta_{\text{ukf}} = 0$, $k_{\text{ukf}} = 3 - d$. Depending on the system, the pre-defined values of $\boldsymbol{\theta}$ lead the UKF to poor predictions. In the most extreme case, the UKF may estimate a prediction density function with null variance. Such a case can be called *sigma points collapse* and it is shown in Figure 3.4. The prior distribution is plotted in the bottom panel; the function $\boldsymbol{f}$ (blue line) passes through the sigma points taken by the prior distribution. In the case of "bad" assignment of sigma points, represented by the red crosses, the approximation can not reproduce the true oscillation of the state function and the resulting posterior distribution is a peaked Gaussian with a small variance. If the sigma points are the green circles (this case may be called "good" assignment), the function $\boldsymbol{f}$ is better approximated and the corresponding posterior is shown in the left panel. Anytime collapse happens during the UKF training phase, the marginal likelihood of observations $p(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1})$ becomes lower.

The strategy proposed here is to develop a *learning algorithm* to find $\boldsymbol{\theta}$ so that collapse becomes unlikely. This means that during the UKF iterations, the new learned parameters have to maximise the marginal likelihood of measurements and avoid its decrease. From a certain point a view, this approach implies that the UKF is considered as a model, not merely as an approximated method. Hence, if the UKF is a model, the researcher has to identify the underlying assumptions of the model (i.e. the position of points in the space) and learn the parameter $\boldsymbol{\theta}$ from training data (Turner and Rasmussen, 2012).

As Section 3.6 will describe, the optimization of the location of sigma points may let inference difficult due to the multi-modality of the likelihood.

Initialization of the UKF

As all the KF-based methods, the overall inference performance of the UKF and its convergence may depend on the initialization of the algorithm. The research of Sitz et al. (2002) has proved the wide applicability and the inference performances of the UKF on several dynamical systems, but Giurghita and Husmeier (2016) have shown that the results highly depend on the initialization of the method. Indeed, the starting values of an algorithm reflect the researcher's belief on the state of the system. With real-world data, it may often happen to set an initialization far from the true unknown parameters governing the hidden state. At the same time, even if a comprehensive pre-inference analysis can be carried out to guess the dynamics of the process, there is the need of a method which does not rely on "tuning" parameters or hand-crafting of the algorithm before inference.

For such a reason, my strategy is to initialize the UKF with the posterior distribution approximation of the ABC algorithm. The ABC approach is presented in Section 3.7; in order to obtain credible intervals of parameters, the ABC will work as a prelude to the UKF estimation of the state space model.

3.6. *Sigma points optimization*

As stated in the previous Section, the sigma points collapse implies the decrease of the likelihood of measurements. Hence, for a good sigma point assignment the likelihood $p(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1})$ has to be maximised into the UKF steps.

Mathematically, the problem concerns with the research of a global optimizer, or minimizer, of an unknown objective function $\mathscr{L}$:

$$\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}\in\Theta} \mathscr{L}(\boldsymbol{\theta}), \tag{3.36}$$

where $\Theta \in \mathbb{R}^d$ and $\boldsymbol{\theta}$ is any arbitrary query point in the time domain. The *black-box* function $\mathscr{L}$ is assumed to have no closed form but may be evaluated at point $\boldsymbol{\theta}$: this evaluation produces noise-corrupted outputs $\boldsymbol{y}$ such that $\mathrm{E}[\boldsymbol{y}|\boldsymbol{\theta}] = \mathscr{L}(\boldsymbol{\theta})$. In other

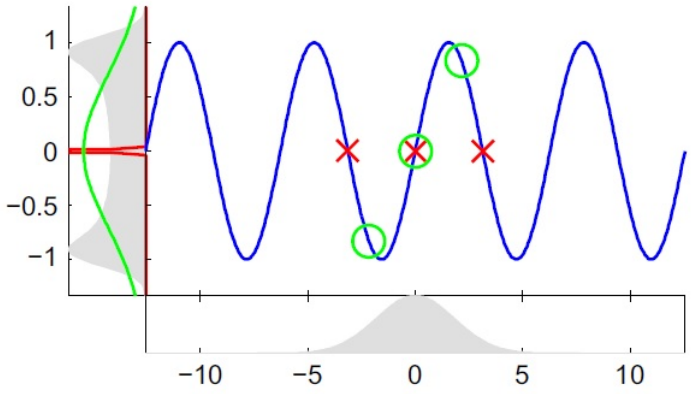

Figure 3.4: An example of good and bad assignment of sigma points. The bottom panel shows the true input distribution. The center panel shows the state function $\boldsymbol{f}$ (in blue) and the sigma points for $\alpha_{\text{ukf}} = 1$ (red crosses) and for $\alpha_{\text{ukf}} = 0.68$ (green circles. This value comes from a simulation study). The parameters β_{ukf} and k_{ukf} are fixed to default. Using the set of sigma points given by the red crosses one gets a degenerate solution, while for the different set a near optimal approximation is reached. Source: Figure 1 of Turner and Rasmussen (2012)

.

words, the function $\mathscr{L}$ can not be observed directly but only its noisy point-wise observations $\boldsymbol{y}$.

To solve problem (3.36), the following sequential search method is considered: at the j-th iteration, the algorithm selects a location point in the parameter space $\boldsymbol{\theta}_j$ at which to query $\mathscr{L}$ and observe $\boldsymbol{y}_j$. After J queries, the final recommendation $\boldsymbol{\theta}^*$ represents the best estimate of the optimizer.

In the case of sigma points location in the UKF framework, the objective (*loss*) function is the negative log marginal likelihood of measurements

$$\mathscr{L}(\boldsymbol{\theta}) = -\log p(\boldsymbol{y}_{1:T}|\boldsymbol{\theta}) = -\sum_{t=1}^{T} \log p(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1},\ \boldsymbol{\theta}). \tag{3.37}$$

To find the parameter $\boldsymbol{\theta}$ that minimizes equation (3.37), the problems of unavailability of a closed-form solution and non-convexity and multi-modality of the $\mathscr{L}$ function is faced.

I focus on the implementation and comparison of three optimizing methods: the loss function $\mathscr{L}(\boldsymbol{\theta})$ is minimized by using a discrete search method and Bayesian optimisation (BO).

3.6.1. Bayesian optimization

Bayesian optimization is a sequential model-based approach to solve problem (3.36). BO is able to take advantage of the full information provided by the history of the optimization to make the search of $\boldsymbol{\theta}^*$ efficient (Shahriari et al., 2016). Basically, the BO method can be summarized by two main components: a probabilistic surrogate model and a loss function. The first consists of a prior distribution that captures the beliefs about the behaviour of the unknown objective $\mathscr{L}$ and an observation model representing the data generating process; the second, the loss function, describes how optimal a sequence of queries $\boldsymbol{\theta}_j$ are. A simple description of the algorithm may be as follows. A prior belief over the possible objective functions is prescribed and then the expected loss function is minimized to select an optimal sequence of queries. After observing the output of each query of $\mathscr{L}$, the prior is updated to produce a more informative posterior distribution over the space of objective functions. In other words, the BO model is sequentially refined as data are observed via Bayesian posterior updating. Indeed, the Bayesian posterior represents the updated beliefs, given the observed data, on the likely objective function to optimize. One problem with this minimum expected risk framework is that up to the full evaluation budget, the true sequential risk is computationally intractable.

This intractability has led to the introduction of heuristics such as *acquisition functions*, so that the problem of directly optimizing the computationally expensive loss function is shifted to the maximization of the computationally cheaper acquisition functions (Shahriari et al., 2016 give a comprehensive review of acquisition functions such as Thompson sampling, probability of improvement, expected improvement, upper confidence bound). Usually, these acquisition functions trade off *exploration* and *exploitation*: their optima are located where the uncertainty in the surrogate model is large (exploration) and/or when the model prediction is high (exploitation). BO algorithms selects the next query point by maximizing such acquisition functions, which are analytically easier to evaluate or at least approximate than the

original objective function. In other words, the acquisition function $\mathscr{A}_j : \Theta \to \mathbb{R}$ evaluates the utility of candidate points for the next evaluation of $\mathscr{L}$ and leverage the uncertainty in the posterior to guide exploration. The point $\boldsymbol{\theta}_{j+1}$ is selected by maximizing $\mathscr{A}_j$, where j highlights the implicit dependence on the currently available data, where "data" refers to previous locations (*training dataset*) where $\mathscr{L}$ has been evaluated, and the corresponding noisy outputs. Figure 3.5 and Algorithm 1 sketch the BO procedure. For an extensive review of Bayesian optimization, the reader may see Shahriari et al. (2016) and Jones et al. (1998).

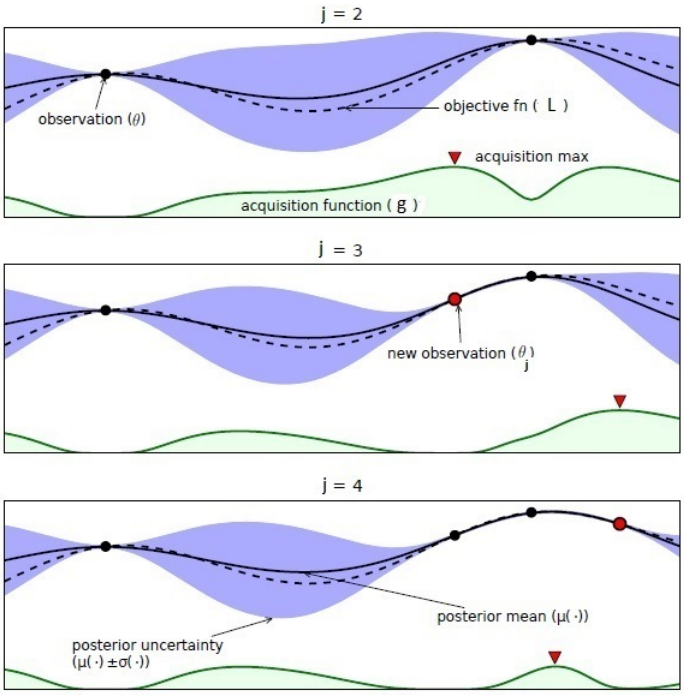

Figure 3.5: Illustration of the BO method over three iterations. Although the objective function $\mathscr{L}$ is plotted (dashed line), in real applications it is unknown. The solid black line is the probabilistic model (the estimation of the loss) and the blue area represents the confidence interval. The acquisition functions are the lower shaded plots. The more the objective takes high values (exploitation) and the prediction uncertainty increases (exploration), the more the acquisition function is high. Source: Fig.1 of Shahriari et al. (2016).

Method of optimization

Let us define the *simulator* as the state function of model (3.2). Standard approaches consist in emulating the simulator through a Gaussian process (GP) prior. Here, the strategy is different: the objective function $\mathscr{L}$ itself is emulated. Following the method of Rasmussen and Williams (2006), the emulator g of the loss is given by a GP with constant mean function and Matérn kernel function k. Notice that the covariance structure of the GP dictates the structure of the response function to fit; in example, if one expects that the response function to model is periodic, the researcher may find suitable the use of a periodic kernel. Here a stationary (shift invariant) kernel of the Matérn class is chosen. The latter kind of kernel is parametrized by a *smoothness parameter* $\nu > 0$, which indicates that the samples from the GP are differentiable $\lfloor \nu - 1 \rfloor$ times. The most commonly used Matérn kernels (Shahriari et al., 2016) have $\nu = \{1/2, 3/2, 5/2, \text{Sq-Exp}\}$, where Sq-Exp is the Squared Exponential kernel with smoothness $\nu \to \infty$. The use of the Squared Exponential kernel would lead to infinitely-differentiable functions, which is an unrealistic assumption in many scenarios. As in Snoek et al. (2012), I fix $\nu = 5/2$, which leads to twice differentiable sample paths.

Consider the following hierarchical Bayesian non-parametric regression model:

$$\mathscr{L} | \boldsymbol{g}, \sigma^2 \sim N(\boldsymbol{g}, \sigma^2 I) \tag{3.38}$$

$$g(\boldsymbol{\theta}) | m, k \sim GP(m(\boldsymbol{\theta}), k(\boldsymbol{\theta}, \boldsymbol{\theta}')), \tag{3.39}$$

where $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta$, $\mathscr{L} = [\mathscr{L}(\boldsymbol{\theta}_1), \dots, \mathscr{L}(\boldsymbol{\theta}_J)]'$, $\boldsymbol{g} = [g(\boldsymbol{\theta}_1), \dots, g(\boldsymbol{\theta}_J)]'$ and I assume that $m(\boldsymbol{\theta}) = c, \ \forall \boldsymbol{\theta} \in \Theta$ and the scalar c has to be estimated from data (this assumption is equivalent to the standard literature approach of placing a zero-mean Gaussian prior to zero-centred data). The GP parameters are estimated by maximum log marginal likelihood.

At the initial point of the optimization scheme the observation variance σ^2 is fixed. The model's prediction and uncertainty in the objective function at the point

$\boldsymbol{\theta}_j$ are represented by, respectively, the mean and variance posterior:

$$\mu_j(\boldsymbol{\theta}) = m(\boldsymbol{\theta}) + k(\boldsymbol{\theta})'(k(\boldsymbol{\theta}, \boldsymbol{\theta}') + \sigma^2 \boldsymbol{I})^{-1}(\mathscr{L} - m) \tag{3.40}$$

$$\sigma_j^2(\boldsymbol{\theta}) = k(\boldsymbol{\theta}, \boldsymbol{\theta}) - k(\boldsymbol{\theta})'(k(\boldsymbol{\theta}, \boldsymbol{\theta}') + \sigma^2 \boldsymbol{I})^{-1} k(\boldsymbol{\theta}), \tag{3.41}$$

where $k(\boldsymbol{\theta})$ is a vector of covariance terms between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_{1:j}$.

The starting point of the optimization scheme should think out an initial representation of the objective function, obtained by conditioning the GP on a set of design points in the parameter (input) space. To minimize the evaluation-costly $\mathscr{L}$, the efficient global optimization (EGO) algorithm proposed by Jones et al. (1998) is utilized, which iteratively selects the point with the highest expected improvement over the *incumbent* minimum (best feasible solution known up to the j-th iteration). The EGO has been extended to GPs by Huang et al. (2006), while its convergence is established in the article of Vazquez and Bect (2007). Here, following Jones et al. (1998), a space filling Latin Hypercube design, with $10 \times d$ initial input points is used.

In the framework of sigma points placement, the Bayesian optimisation algorithm iteratively maintains a statistical emulator of the objective function $\mathscr{L}$ and chooses the next "best" point $\boldsymbol{\theta} = (\alpha_{\text{ukf}}, \beta_{\text{ukf}}, k_{\text{ukf}})$ by maximising an auxiliary acquisition function derived from the current emulator. Given the GP at the current iteration, $\hat{\mathscr{L}} \sim \text{GP}(m, s)$, I compare the performance of the expected improvement (EI) and the upper confidence bound (UCB) acquisition functions. Both acquisition functions balance *exploitation*, where the GP mean $m(\boldsymbol{\theta})$ predicts a low function value, and *exploration* where the GP predicts high uncertainty $s^2(\boldsymbol{\theta})$. The EI has been first introduced by Mockus et al. (1978), while the convergence of the UCB is proved in Srinivas et al. (2010). Even if the EI or UCB functions may be highly multi-modal, they can be efficiently optimized using multiple restarts of standard state-of-the-art global optimization solvers (e.g. Perttunen et al., 1993), since the costs for the computation of the acquisition functions are negligible to those required for the evaluation of the loss. Once the minimum of the $\mathscr{A}$ is reached at point $\bar{\boldsymbol{\theta}}$, the objective function is computed at the next best candidate $\bar{\boldsymbol{\theta}}$ obtaining the output $\bar{\mathscr{L}} = \mathscr{L}(\bar{\boldsymbol{\theta}})$. Then, the new point $(\bar{\boldsymbol{\theta}}, \bar{\mathscr{L}})$ is added to the training dataset $\mathscr{D}$ and a

new iteration starts by re-fitting the GP. The process is iterative until the maximum budget number of function evaluations has been exceeded. The point $(\boldsymbol{\theta}_j, \mathscr{L}_j)$ having minimum observed objective $\mathscr{L}_j = \mathscr{L}(\boldsymbol{\theta}_j)$ is the output of the EGO minimizer of $\mathscr{L}$.

Result: Best optimizing point $\boldsymbol{\theta} = \arg\min_{\boldsymbol{\theta}\in\Theta} \mathscr{L}(\boldsymbol{\theta})$

Define the acquisition function $\mathscr{A}$;
for $j = 1, \ldots, J$ **do**
  select new $\boldsymbol{\theta}_{j+1}$ by optimizing $\boldsymbol{\theta}_{j+1} = \arg\max_{\boldsymbol{\theta}}(\mathscr{A})$;
  query the loss function to compute $\boldsymbol{y}_{j+1}$;
  augment the training dataset $\mathscr{D}_{j+1} = \{\mathscr{D}_j, (\boldsymbol{\theta}_{j+1}, \boldsymbol{y}_{j+1})\}$;
  update the statistical model;
end

Algorithm 1: The Bayesian optimization procedure

Expected Improvement The *improvement* at the point $\boldsymbol{\theta}$ is defined as

$$I(\boldsymbol{\theta}) = \begin{cases} \boldsymbol{\theta}; & \boldsymbol{\theta} < \bar{\boldsymbol{\theta}} \\ \bar{\boldsymbol{\theta}}; & \boldsymbol{\theta} \geq \bar{\boldsymbol{\theta}} \end{cases} = \max(\mathscr{L}_{\min} - g(\boldsymbol{\theta}), 0), \tag{3.42}$$

where $\bar{\boldsymbol{\theta}}$ is the incumbent solution at iteration j, $g(\boldsymbol{\theta}) \sim N(m(\boldsymbol{\theta}), s^2(\boldsymbol{\theta}))$ is the marginal GP at point $\boldsymbol{\theta}$, $\mathscr{L}_{\min} = \mathscr{L}(\boldsymbol{\theta}_{\min})$ is the best function value known so far. In other words, $g(\boldsymbol{\theta})$ is the random variable that models the uncertainty about the function's value at $\boldsymbol{\theta}$ and hence the improvement is a random variable itself. To obtain the expected improvement, the integral to compute is

$$\mathrm{E}\left[I(\boldsymbol{\theta})\right] \equiv \mathrm{E}\left[\max(\mathscr{L}_{\min} - g(\boldsymbol{\theta}), 0)\right] = \int I(\boldsymbol{\theta}) g(\boldsymbol{\theta}|\boldsymbol{\theta}_{j+1}, I_0) d\boldsymbol{\theta}, \tag{3.43}$$

in which the GP is conditional to the next evaluation location $\boldsymbol{\theta}_{j+1}$ and to the all available information I_0 (the data seen so far). Equation (3.43) means that the expected

loss is the expected lowest value of the function after observing the next evaluation. Following Jones et al. (1998) and applying some integration by parts, the right-hand side of (3.43) is a Gaussian integral and can be expressed in closed form

$$\mathrm{EI}(\boldsymbol{\theta}) = (\mathscr{L}_{\min} - m(\boldsymbol{\theta}))\Phi\left(\frac{\mathscr{L}_{\min} - m(\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right) + s(\boldsymbol{\theta})\phi\left(\frac{\mathscr{L}_{\min} - m(\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right), \tag{3.44}$$

where Φ and ϕ denote the cumulative distribution function (CDF) and probability density function (PDF) of a $\mathrm{N}(0, 1)$ random variable evaluated at $\boldsymbol{\theta}$. The weights of the EI acquisition (3.44) are the probabilities of a successful objective function evaluation. This allows us to account for failure in the evaluation of $\mathscr{L}$ due to matrix singularities, and still optimise it when standard optimization algorithms would fail to. The balance between exploration and exploitation of (3.44) is expressed by the contribution $\mathscr{L}_{\min} - m(\boldsymbol{\theta})$ (that is higher when the prediction is smaller than the incumbent minimum) and by $s(\boldsymbol{\theta})$, which increases the acquisition function value as far as the GP uncertainty is high at $\boldsymbol{\theta}$.

Upper Confidence Bound The UCB optimizes the maximum of a confidence interval on the loss function value:

$$\mathrm{UCB}(\boldsymbol{\theta}) = \mathrm{E}[\mathscr{L}(\boldsymbol{\theta})] + r\sqrt{\mathrm{Var}[\mathscr{L}(\boldsymbol{\theta})]}, \tag{3.45}$$

where r is a constant which controls the exploration exploitation trade-off. The guiding principle behind the acquisition function (3.45) is to be optimistic in the face of uncertainty (Shahriari et al., 2016). Using the upper confidence for every query $\boldsymbol{\theta}$ corresponds to effectively using a fixed probability best case scenario according to the model. The optimizer based on (3.45) is implemented following the approach of Turner and Rasmussen (2012). A simple comparison between the EI and UCB acquisition functions is shown in Figure 3.6.

3.6.2. Discrete search

The Bayesian optimization method with a GP interpolation is compared with a a discrete grid search scheme with a cubic spline interpolation. I create an interval of

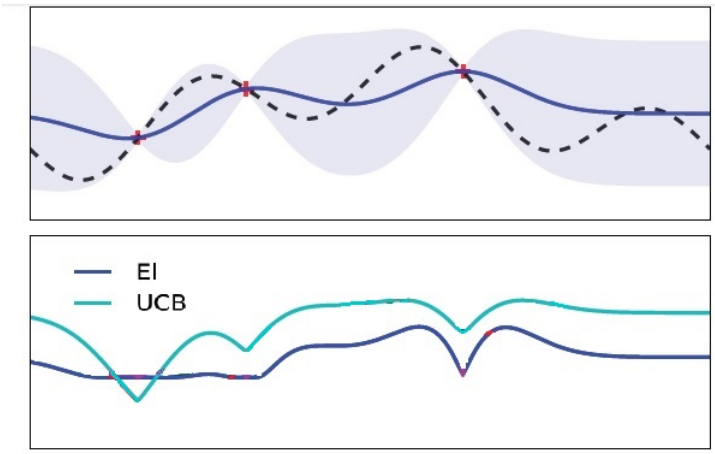

Figure 3.6: Illustration of the surrogate regression model (top panel) and the EI and UCB acquisition functions (bottom panel). The true loss is the dashed line, while the solid line represents the probabilistic regression model with the shaded region of confidence intervals. The observations are the red crosses. The UCB acquisition function is more optimistic than the EI. Source: Fig.5 of Shahriari et al. (2016).

$\boldsymbol{\theta}$ in which the surface of the loss $\mathscr{L}$ is monitored. The function $\mathscr{L}(\boldsymbol{\theta})$ is interpolated with a cubic convolution taking as many knots as the size of the grid,

$$s_{3,j}(\boldsymbol{\theta}) = \sum_{i=0}^{3} a_{ij}(\boldsymbol{\theta} - \boldsymbol{\theta}_j)^i, \tag{3.46}$$

where the coefficient a_{ij} are determined following the standard continuity conditions listed in Bowman and Azzalini (1997). Finally, the maximum of the interpolated function and the corresponding parameters are taken.

The grid search has the advantage that it is easier to implement than BO and it is considered as a matter of comparison with the GP interpolation. Indeed, Bayesian optimization may be more accurate in the evaluation of the behaviour of $\mathscr{L}(\boldsymbol{\theta})$ but has higher computational costs.

3.7. *Likelihood free inference*

As more realistic and complex dynamical models are created, the likelihood surfaces of such models can be computationally intractable or too costly to evaluate.

More recently, much literature has been devoted to the development of a simulation-based technique known as ABC to infer posterior distributions in complicated scenarios (see the reviews of Beaumont, 2010, Sisson and Fan, 2011, Marin et al., 2012, Hartig et al., 2011).

Essentially, the idea behind the ABC algorithm is that the evaluation of the likelihood is replaced by a comparison between observed and simulated data. Figure 3.7 shows a representation of the ABC intuition. An unknown process generates observed data, summarized by a statistics S. At the same time, a stochastic simulation gives rise to a simulated sample. The summary statistics S is calculated on simulated data, and compared with the one of the observed dataset. If $S_{obs} = S_{sim}$, the stochastic simulation is able to reproduce real data. If $S_{obs} \neq S_{sim}$, another simulation has to be evaluated.

If data are discrete, it is possible to sample from the posterior density of the parameters without an explicit likelihood function evaluation and without any approximation. If data are continuous, the probability of exact matching between the simulated data and the real sample is zero; hence, the method relies on approximate parameter distributions. The posterior distribution from which the parameters are drawn is the probability distribution that could have originated the observations.

In this Section, the theory underlying the ABC, its extensions and some difficulties are discussed.

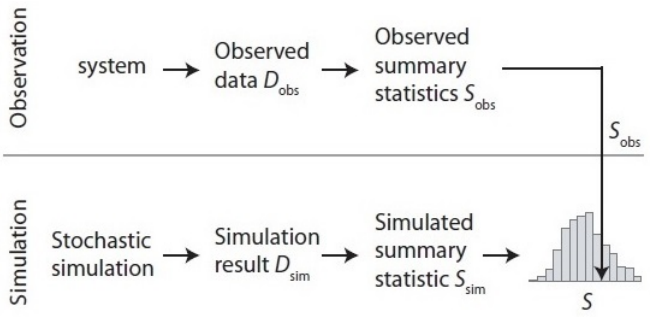

Figure 3.7: Graphical representation of the idea behind the ABC method: observed and simulated data are compared through summary statistics. Source: Figure 4 of Hartig et al. (2011).

3.7.1. The ABC method

Consider a vector of parameters $\boldsymbol{\lambda}$, an observation set $\boldsymbol{y}_0$ and a model M. In general, given the prior distribution $\pi(\boldsymbol{\lambda})$, the posterior distribution $\pi(\boldsymbol{\lambda}|\boldsymbol{y}) \propto l(\boldsymbol{y}|\boldsymbol{\lambda})\pi(\boldsymbol{\lambda})$ is evaluated approximating the likelihood $l(\boldsymbol{y}|\boldsymbol{\lambda})$ of $\boldsymbol{\lambda}$ for data $\boldsymbol{y}$. A generic form of the ABC rejection sampler is sketched in Algorithm 2. A tolerance is defined by ζ and N parameter values, called *particles* are drawn from the prior $\pi(\boldsymbol{\lambda})$. For each i-th particle, a sample $\boldsymbol{y}^*$ is simulated following model M. The simulated dataset is compared with observed data $\boldsymbol{y}_0$ using a distance function D: if $D(\boldsymbol{y}^*, \boldsymbol{y}_0) \leq \zeta$, it means that the candidate parameter $\boldsymbol{\lambda}^{(i)}$ can reproduce experimental data accordingly to the level of agreement defined by the threshold ζ, and particle $\boldsymbol{\lambda}^{(i)}$ is accepted.

Result: Posterior distribution $\pi(\boldsymbol{\lambda}|D(\boldsymbol{y}^*, \boldsymbol{y}_0) \leq \zeta)$

Define the threshold ζ;
for $i = 1, \ldots, N$ **do**
  while $D(\boldsymbol{y}^*, \boldsymbol{y}_0) \geq \zeta$ **do**
    sample $\boldsymbol{\lambda}^{(i)}$ from the prior distribution $\pi(\boldsymbol{\lambda})$;
    simulate a dataset $\boldsymbol{y}^*$ accordingly to model $M(\boldsymbol{y}|\boldsymbol{\lambda}^{(i)})$;
  end
end

Algorithm 2: The ABC rejection sampler

The ABC outcome is a sample of parameters. For univariate or bivariate parameters, it is possible to visualize the distribution drawing an histogram (Hartig et al., 2011). If the threshold is sufficiently small, the distribution

$$\pi(\boldsymbol{\lambda}|D(\boldsymbol{y}^*, \boldsymbol{y}_0) \leq \zeta) \tag{3.47}$$

is a good approximation for the posterior $\pi(\boldsymbol{\lambda}|\boldsymbol{y}_0)$. The threshold has to be fixed to a small value in order to achieve a good approximation of the posterior, meaning that a lot of computational time is required. Moreover, the acceptance rate is very low

if the prior distribution is very different from the ABC posterior approximation of equation (3.47). To skip these problems, two main methods exist in the literature: (i) the ABC based on Markov chain Monte Carlo (ABC-MCMC) and (ii) the ABC in a Sequential Monte Carlo sampling scheme (ABC-SMC). Figure 3.8 highlights the difference between the ABC, ABC-MCMC and ABC-SMC. Toni et al., 2009 and Beaumont, 2010 give a detailed review of ABC-MCMC; here the focus is on the ABC-SMC.

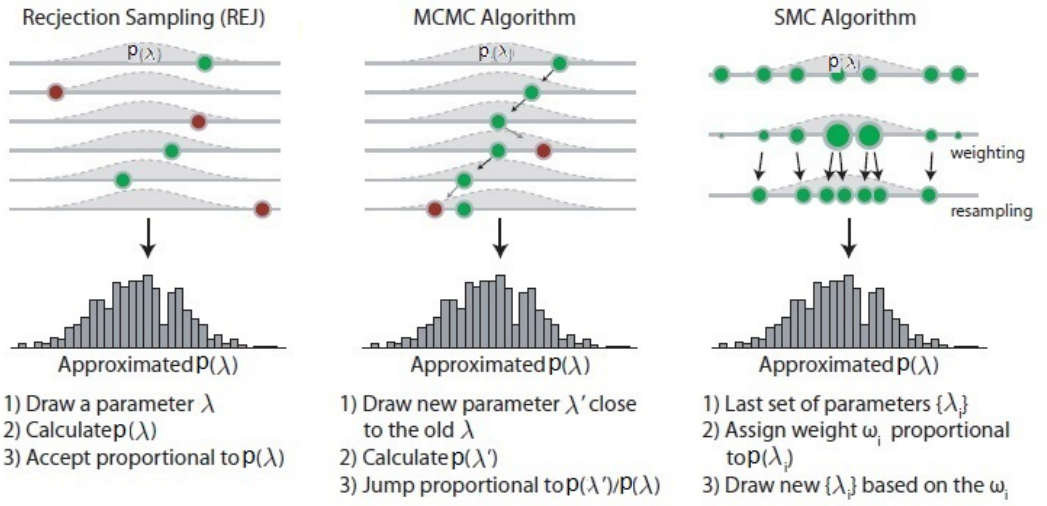

Figure 3.8: Differences between ABC rejection sampler, ABC-MCMC and ABC-SMC. The circles represent parameter combination within the algorithm. Red circles are rejected parameters, where the rejection depends on the version of the ABC. Green circles are accepted parameters. In the rejection sampling (left panel), a particle is accepted on the basis of the ABC posterior approximation (equation (3.47)). The MCMC sampler (central panel) draws a new parameter value depending on the ratio between posterior approximations. SMC (right panel) begins with a set of candidate parameters and accepts the particles according to their weights. Source: Figure 7 of Hartig et al. (2011).

The MCMC is a standard approach in Bayesian methods that constructs a Markov chain of parameter values. At each time step, the parameter combination is chosen by a random movement conditional on the distribution of parameters at the previous time. A parameter value is accepted in the chain conditional on the ratio of the

posterior distribution approximations. The ABC-MCMC converges to the target posterior distribution, with a time usually much shorter than the simple ABC rejection sampler. Despite this, the correlated nature of samples of the MCMC scheme, together with the low acceptance rate of ABC, may result in very long chains with the chance that the chain gets stuck in regions with a low probability for a long time.

The ABC-SMC, at least in part, avoids the limits of the ABC and ABC-MCMC. Indeed, the ABC-SMC has two main features. First, it computes a weighted resampling from the set of points already drawn; second, the tolerances decrease over time. According to the weighted resampling, each point has a given weight (*importance weight*) that takes into account that the points are not picked from the prior. Since the tolerance is smaller, the weighted set of points improves the approximation to the posterior. In the next Section, the ABC-SMC is described in detail.

3.7.2. The ABC-SMC scheme

In the ABC-SMC procedure the posterior distribution is approximated sequentially for $t = 1, \dots, T$ by constructing intermediate distributions that converge to the posterior distribution (Toni et al., 2009). To perform the ABC-SMC, a tolerance schedule $\zeta_1 > \zeta_2 > \dots > \zeta_T$ is defined first. For the first parameter distribution, $t = 1$, the ABC rejection sampler brings a first parameter population $\boldsymbol{\lambda}_1$. For the other intermediate distributions, $t > 1$, N particles are sampled from the previous population $\boldsymbol{\lambda}_{t-1}$ according to a set of weights W_t. Each particle is perturbed according to a kernel K_t, and a dataset $\boldsymbol{y}^*$ is simulated through a model $M(\boldsymbol{y}|\boldsymbol{\lambda}^*)$. The ABC-SMC proceeds as described in Algorithm 3.

The SMC sampling scheme approximates the belief distribution using a weighted set of particles. In other words, the algorithm samples from a *proposal distribution* assigning *importance weights* to the particles: this method is called *importance sampling* (Murphy, 2012). The importance weights are given by

$$W_t^{(*i)} = \pi(\boldsymbol{\lambda}_t^{(i)}) / K_t(\boldsymbol{\lambda}_{t-1}^{(i)} | \boldsymbol{\lambda}_t^{(i)}; \tau_t^2), \tag{3.48}$$

where $\pi(\boldsymbol{\lambda}_t^{(i)})$ represents the approximated belief state and $K_t(\boldsymbol{\lambda}_{t-1}^{(i)} | \boldsymbol{\lambda}_t^{(i)}; \tau_t^2)$ is the proposal distribution. The SMC can fail after few steps if most of particles have negligible

Result: Posterior distribution $\pi(\boldsymbol{\lambda}|D(\boldsymbol{y}^*, \boldsymbol{y}_0) \leq \zeta_T)$

Define $\zeta_1, \zeta_2, \ldots, \zeta_T$;
for $t = 1$ **do**
  for $i = 1, \ldots, N$ **do**
    while $D(\boldsymbol{y}^*, \boldsymbol{y}_0) \geq \zeta_1$ **do**
      sample $\boldsymbol{\lambda}^{(i)}$ from the prior distribution $\pi(\boldsymbol{\lambda})$;
      simulate a dataset $\boldsymbol{y}^*$ accordingly to model $M(\boldsymbol{y}|\boldsymbol{\lambda}^{(i)})$;
    end
  end
  Define τ_2^2 as twice the empirical variance of the population $\boldsymbol{\lambda}_1$;
  Set $W_1 = 1/N$;
end
for $t = 2, \ldots, T$ **do**
  for $i = 1, \ldots, N$ **do**
    while $D(\boldsymbol{y}^*, \boldsymbol{y}_0) \geq \zeta_t$ **do**
      sample $\boldsymbol{\lambda}^{(i,**)}$ from the previous population $\boldsymbol{\lambda}_{t-1}$ with weights W_{t-1};
      perturb the particle $\boldsymbol{\lambda}^{(i,*)} \sim K_t(\boldsymbol{\lambda}|\boldsymbol{\lambda}^{**}; \tau_t^2)$;
      simulate a dataset $\boldsymbol{y}^*$ accordingly to model $M(\boldsymbol{y}|\boldsymbol{\lambda}^{(i,*)})$;
      Set $W_t^{(i)} = \pi(\boldsymbol{\lambda}_t^{(i)}) / \sum_{j=1}^{N} W_{t-1}^{(j)} K_t(\boldsymbol{\lambda}_{t-1}^{(j)}|\boldsymbol{\lambda}_t^{(j)}; \tau_t^2)$;
    end
  end
  Normalize the weights;
  Take τ_{t+1}^2 as twice the empirical variance of the population $\boldsymbol{\lambda}_t$.
end

Algorithm 3: The ABC-SMC algorithm

weights. This event is called the *degeneracy problem* and occurs in high-dimensional space (Murphy, 2012). The degree of degeneracy is quantified by the effective sam-

ple size:

$$N_{\text{eff}} = \frac{N}{1 + \text{Var}(W_t^{(*i)})}, \tag{3.49}$$

where $W_t^{(*i)}$ for particle i is defined as in equation (3.48). Since the effective sample size is not quantified, it is approximated by

$$\hat{N}_{\text{eff}} = \frac{N}{1 + \sum_{i=1}^{N} (W_t^{(i)})^2}. \tag{3.50}$$

The idea is that if the variance of weights is large, the algorithm is wasting resources in updating particles with negligible weight which do not add informations to the posterior estimates. The problem is solved adding a resampling step. The resampling step consists in monitoring the effective sample size of particles, according to equation (3.50), eliminating particles with a low weight and replicating the surviving particles.

3.7.3. Choice of summary statistics

As presented in Sections 3.7 and 3.7.2, the ABC and its sophisticated versions, both MCMC and SMC, are based on the choice of a metric D. The definition of a suitable distance function between datasets is not trivial. Hence, the distance can be based on summary statistics $S(\boldsymbol{y}_0)$ and $S(\boldsymbol{y}^*)$, such that $D(\boldsymbol{y}^*, \boldsymbol{y}_0) = D^*(S(\boldsymbol{y}^*), S(\boldsymbol{y}_0))$, where D^* is a distance function defined on the summary statistics space.

The key challenge in ABC approaches relies both on the definition of a distance function that quantifies the difference between the simulated and observed datasets and on the choice of a summary statistic that best captures the key features of data (Jones et al., 2015 show an interesting comparison between summary statistics and distance functions).

Through summary statistics the dimensionality of the data is reduced. At the same time, to avoid the risk of loss of information, the sufficiency of the statistic should be tested with respect to the inferential task. One should consider that the choice of summary statistics affects the inference scheme, and may bias the outcome because of the relation among summary statistics and data (Beaumont, 2010). As more and more summary statistics are used, they should be jointly sufficient for the

likelihood. With many summary statistics, and the varying degrees of correlation to each other and the parameters, the influence of any one on the outcome decreases. Unfortunately, the increasing number of summary statistics rises the dimensionality of the system as well. For complex model, a researcher may decide to use summary statistics that are only close to sufficiency. To a certain extent, the finding of summary statistics close to sufficiency as possible depends on the experience and intuition of the researcher. It may happen to work with a summary statistics without a necessarily strong theory relating the statistics to parameters.

The choice of summary statistics is very crucial and definitely affects the ABC posterior approximation, so that in this dissertation I do not completely rely on the ABC method to infer the ODE parameters.

Result: Signal extraction and $\boldsymbol{\lambda}$ estimate

---

Define the starting value $\boldsymbol{\lambda}$;
Define a suitable choice of summary statistics $S(\boldsymbol{y}_0)$;
Step 1: run ABC-SMC scheme;
- *Result*: Posterior distribution $\pi(\boldsymbol{\lambda}|D(\boldsymbol{y}^*, \boldsymbol{y}_0) \leq \zeta_T)$;
Step 2: run UKF where the starting value of $\boldsymbol{\lambda}$ is the median of the posterior distribution of ABC-SMC;
- *Result*: Learn the sigma points location from the process;
Step 3: optimize sigma points placement;
- *Result*: Estimate of $\boldsymbol{\theta}$;
Step 4: run UKF with optimized sigma points;

Algorithm 4: The Sequential ABC-UKF algorithm

3.8. *Sequential ABC - UKF*

Because of the ABC limitations discussed in Section 3.7.3 I do not perform the ABC-SMC on $\boldsymbol{\lambda}$, but I sample from the approximation (3.47) the starting values for

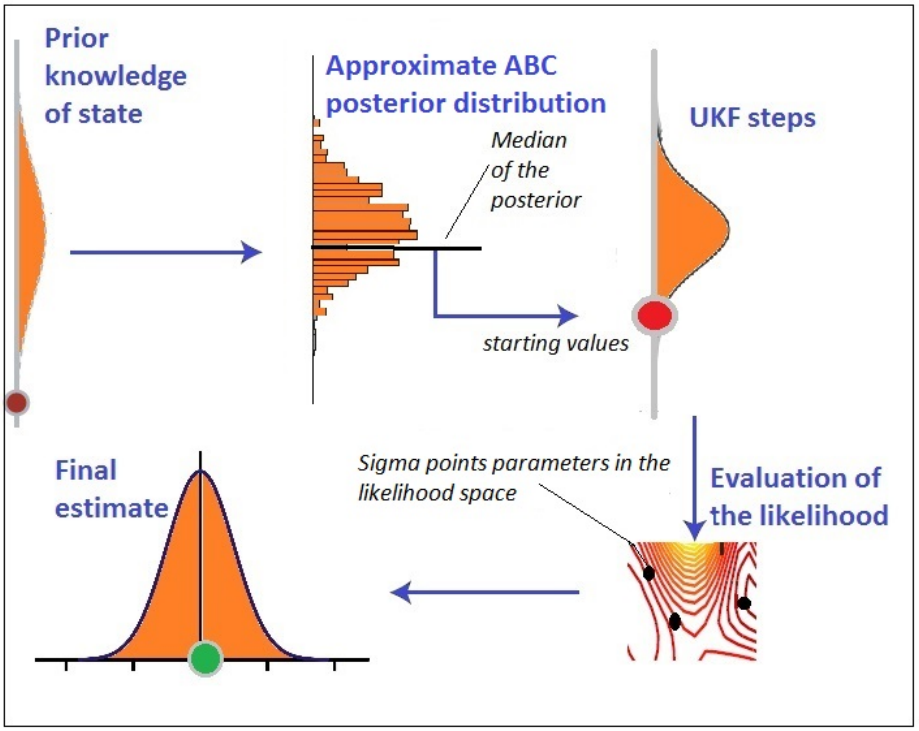

Figure 3.9: An illustration of the Sequential ABC-UKF. Through the ABC-SMC, the method obtains starting values to initialize the UKF from the distribution of equation (3.47), and after the evaluation of the likelihood and the optimization of sigma points location, the UKF estimates converge. The red circles are the initial values, while the green circle is the true parameter vector.

UKF. Briefly recall that the UKF is a powerful method to infer parameters of a non-linear system, but its convergence depends on the initialization and on the sigma points placement.

The proposed strategy is to perform Bayesian filtering through the UKF where the starting values come from the ABC-SMC and to learn the sigma point parameter vector $\boldsymbol{\theta}$ in the likelihood space. To boost UKF performance with respect to the initialization, the UKF method is merged with the ABC-SMC scheme, and this algorithm is called Sequential ABC-UKF. In such a way, the ABC works as a prelude to UKF training and the outcome is a domain in which the true parameter $\boldsymbol{\lambda}$ lies.

The Sequential ABC-UKF works as follows. The ABC-SMC is initialized with

a Uniform prior with a large domain in which I guess the true parameters lie. The ABC-SMC estimates a posterior distribution considered as a credible domain of parameters. From this distribution, the median value to initialize the UKF is taken. The UKF runs twice. First, the UKF evaluates the marginal likelihood of observations $p(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1})$. As the unscented transform calculates the first two moments of the measurements distribution with weights depending on the sigma points parameter $\boldsymbol{\theta}$ (see Section 3.5), the likelihood function depends on the sigma set location. The sigma points position is learnt as described in Section 3.6. Finally, the UKF is performed with starting values coming from ABC-SMC and with optimized sigma set. The proposed inference scheme solves the UKF limits and at the same time preserves the properties of the filtering method. Algorithm 4 and Figure 3.9 illustrate the Sequential ABC-UKF in a generic way and suggests how different algorithms can work together to gain a better knowledge of complex systems.

4. Simulation study

Through a simulation study, the performance of the Sequential ABC-UKF discussed in Chapter 3 to learn the signal and the parameters of the non-linear Duffing oscillator is evaluated. The study is divided into three steps. First, the dependence of the UKF on the signal to noise ratio (SNR), sample size and offsets is quantified (Section 4.2). Second, Section 4.3 concerns the evaluation of the impact of sigma points optimization in the UKF scheme. Third, the improvement of the Sequential ABC-UKF in the convergence to the true values is discussed in Section 4.4.

4.1. *State space Duffing system*

The state space representation of the Duffing ODE is

$$dx_{1t}/dt = x_{2t}, \qquad dx_{2t}/dt = -\left(cx_{2t} + \alpha x_{1t} + \beta x_{1t}^3\right), \tag{4.1}$$

where x_{1t} and x_{2t} are the position and the velocity, respectively, of the oscillation at time t and the parameters of interest are $\boldsymbol{\lambda} = (\alpha, \beta, c)'$. Data are simulated with the `ode23` MATLAB function, with stepsize of integration $\delta t = 0.01$ and starting values for the numerical integration $[1, 0]$. To fix the stepsize $\delta t = 0.01$ several simulations have been performed, in order to assure that a lower stepsize would not affect the accuracy of the ODE approximation. The true parameters are $\boldsymbol{\lambda} = (\alpha, \beta, c)' = (1, 2, 0.1)'$. The function $\boldsymbol{f}$ of model (3.2) is given by the numerical solution of equation (4.1) and $\boldsymbol{h}$ is the identity function. Measurements y_t are obtained by sampling n data points from the first component, x_{1t}, and adding observational noise $\eta_t \sim N(0, \sigma_\eta^2)$ with known variance. The time interval is $t = 1, \ldots, 20$, and the initial covariance of the system is set to $2\mathbb{I}_5$, where $\mathbb{I}_5$ is the 5×5 identity matrix. As previously discussed in Section 3.5.3, a joint state vector merges the true unknown signal with the parameter vector as $\boldsymbol{j}_t = [\boldsymbol{x}_t,\ \boldsymbol{\lambda}_t]' = [(f(\boldsymbol{x}_{t-1}, \boldsymbol{\lambda}_{t-1}) + \boldsymbol{\varepsilon}_t),\ \boldsymbol{\lambda}_{t-1}]'$, and $\boldsymbol{y}_t = h(\boldsymbol{j}_t) + \boldsymbol{\eta}_t$. Recall that, in the deterministic case, $\boldsymbol{\varepsilon}_t = 0$. Simulations are coded in MATLAB, and the UKF is implemented in the `EKF/UKF` toolbox of Hartikainen et al. (2011). To avoid numerical instability of the Cholesky decomposition

of the matrices in UKF training, the default `inv` function of the `EKF/UKF` toolbox is substituted with the backslash operator. A jitter is added when necessary.

The measures of convergence are the following: the Euclidean norm for each parameter before and after inference, the standardized Euclidean norm (SEN) in the parameter space before and after the UKF training, the root mean square (RMS) error, the parameter residual sum of squares (RSS), the average relative bias (ARB), the Signal RSS and Solution RSS:

$$\text{SEN before} = \sqrt{\left(\frac{\hat{\alpha}^{(\text{before})} - \alpha}{\alpha}\right)^2 + \left(\frac{\hat{\beta}^{(\text{before})} - \beta}{\beta}\right)^2 + \left(\frac{\hat{c}^{(\text{before})} - c}{c}\right)^2}, \tag{4.2}$$

$$\text{SEN after} = \sqrt{\left(\frac{\hat{\alpha}^{(\text{after})} - \alpha}{\alpha}\right)^2 + \left(\frac{\hat{\beta}^{(\text{after})} - \beta}{\beta}\right)^2 + \left(\frac{\hat{c}^{(\text{after})} - c}{c}\right)^2}; \tag{4.3}$$

$$\text{Parameter RSS} = \left(\frac{\hat{\alpha}_i - \alpha}{\alpha}\right)^2 + \left(\frac{\hat{\beta}_i - \beta}{\beta}\right)^2 + \left(\frac{\hat{c}_i - c}{c}\right)^2 \text{ for } i = 1, \ldots, n; \tag{4.4}$$

$$\text{Signal RSS} = \sum_{i=1}^{n} \left(\hat{\boldsymbol{y}}_i^{(\text{ukf})} - \boldsymbol{y}_i\right)^2; \quad \text{Solution RSS} = \sum_{i=1}^{n} \left(\hat{\boldsymbol{y}}_i^{(\text{ode})} - \boldsymbol{y}_i\right)^2; \tag{4.5}$$

$$\text{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\hat{\boldsymbol{y}}_i^{(\text{ukf})} - \boldsymbol{y}_i\right)^2}; \tag{4.6}$$

$$\text{ARB} = \frac{\hat{\boldsymbol{y}}_i^{(\text{ukf})} - \boldsymbol{y}_i}{\boldsymbol{y}_i}; \quad \text{ARB} = \frac{\hat{\boldsymbol{\lambda}}_i^{(\text{ukf})} - \boldsymbol{\lambda}_i}{\boldsymbol{\lambda}_i}, \tag{4.7}$$

where $\hat{\boldsymbol{y}}_i^{(\text{ukf})}$ and $\hat{\boldsymbol{\lambda}}_i = (\hat{\alpha}_i, \hat{\beta}_i, \hat{c}_i)$ are, respectively, the signal and the parameter estimates of the UKF for $i = 1, \ldots, n$. The $\hat{\boldsymbol{y}}_i^{(\text{ode})}$ is the numerical solution of the ODE with estimated parameters $\hat{\boldsymbol{\lambda}}$ after the UKF training. Both the Signal RSS and the Solution RSS in equation (4.5) are considered as a differential equation may be reconstructed through a filter (that is the Kalman filtering idea) or from a set of parameters (the ones estimated by the UKF). In my study, I investigate the precision of filter reconstruction and the numerical approximation of the vibration given the estimated parameters. This strategy enables to quantify the sensitivity of the system to a slight change in the parameter settings.

4.2. *Dependence of UKF on noise, sample size and starting values*

To investigate the behaviour of the Duffing process and the UKF performance, several scenarios have been simulated, varying the level of noise, quantified by the SNR, SNR $\in \{30, 10, 1\}$, and the sample size, $n \in \{1000, 100, 50\}$ (Pasetto et al., 2017a).

The evaluation of the impact of initialization is carried out by considering different offsets as starting values for the parameters. The offsets are sampled randomly from a Gaussian distribution in which the mean is defined by a percentage deviation from the true parameter values and the variance is 10% of the mean. The final offset used to initialize the UKF is the average of the offsets sampled from the Gaussian. The results are averaged over 50 independent datasets.

Figures 4.1 - 4.5 show that the UKF successfully learns the parameters from the noisy data, and that at the end of the filtering phase the true parameters always lie within the predicted standard error around the estimate. This suggests that Bayesian filtering offers a successful paradigm for inference in chaotic dynamical systems. The prediction uncertainty depends on the sample size n, and the level of noise, quantified by the SNR. As one would expect, the uncertainty increases with decreasing n and decreasing SNR, i.e. as information in the data is lost, and this study allows a quantification of this trend. The increase in uncertainty particularly affects the parameter β, which is associated with the nonlinear term and the source of the chaotic behaviour.

Table 4.1 shows the effect of the initialization, measured in terms of the Euclidean distance in parameter space. In the case of a small or medium percentage deviation from the true parameters, respectively, offset = 100% and offset = 250%, the distance between estimates and true values is consistently reduced in the filtering process, and the posterior distance (after filtering) is always smaller than the prior distance (before filtering). However, the posterior distance increases with the prior distance, suggesting that a good initialisation will improve the inference results. In the case of "bad" initialization, that is offsets greater than 250%, the final estimates of parameters are very poor, as depicted by the increasing RMS errors in Figure 4.6.

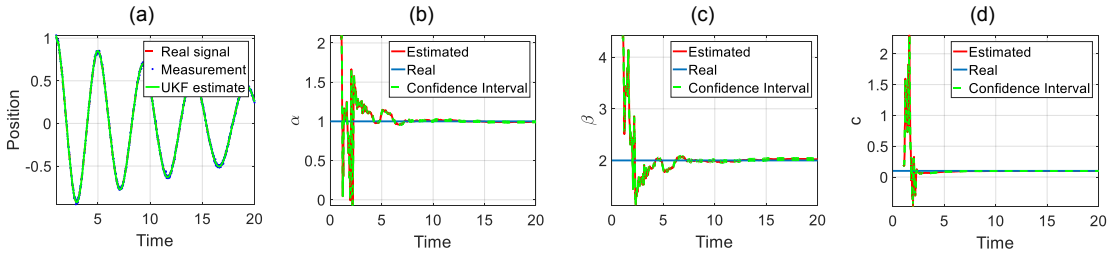

Figure 4.1: UKF estimates for the deterministic Duffing system with SNR = 31 and n = 1000. (a) Signal estimate. (b) Estimate of parameter α. (c) Estimate of parameter β. (d) Estimate of parameter c.

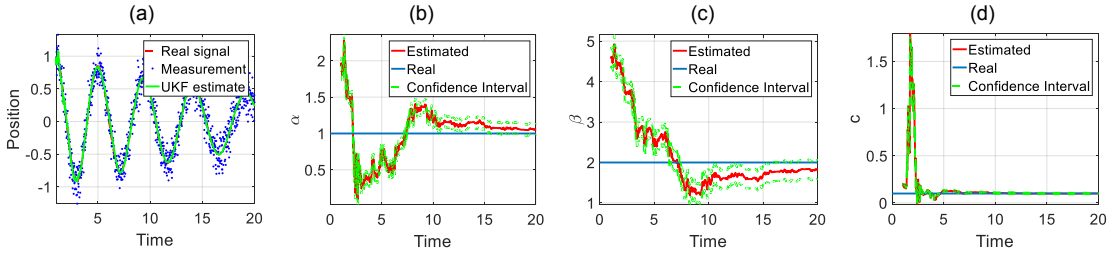

Figure 4.2: UKF estimates for the deterministic Duffing system with SNR = 10 and n = 1000. (a) Signal estimate. (b) Estimate of parameter α. (c) Estimate of parameter β. (d) Estimate of parameter c.

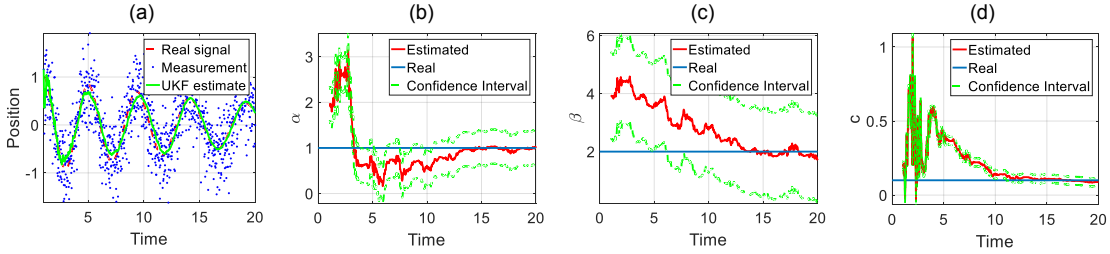

Figure 4.3: UKF estimates for the deterministic Duffing system with SNR = 1 and n = 1000. (a) Signal estimate. (b) Estimate of parameter α. (c) Estimate of parameter β. (d) Estimate of parameter c.

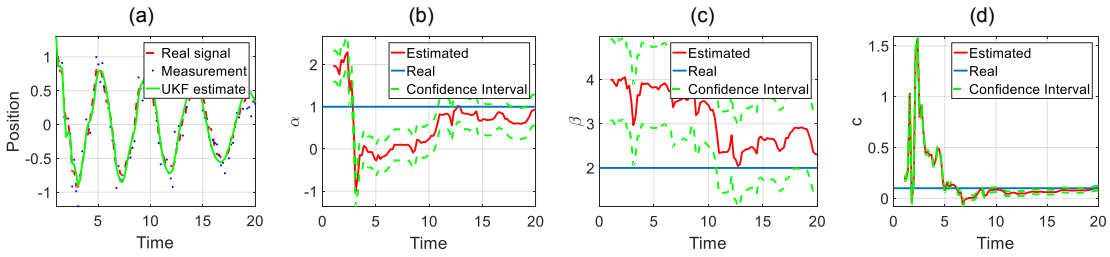

Figure 4.4: UKF estimates for the deterministic Duffing system with SNR = 10 and n = 100. (a) Signal estimate. (b) Estimate of parameter α. (c) Estimate of parameter β. (d) Estimate of parameter c.

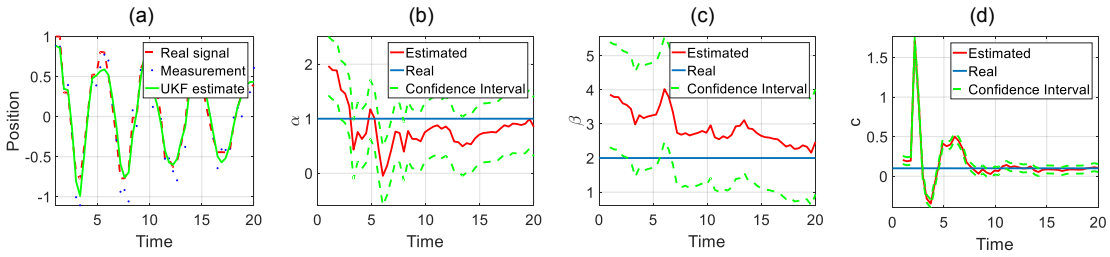

Figure 4.5: UKF estimates for the deterministic Duffing system with SNR = 10 and n = 50. (a) Signal estimate. (b) Estimate of parameter α. (c) Estimate of parameter β. (d) Estimate of parameter c.

4.3. *The UKF and sigma points optimization*

To test the optimization procedures presented in Chapter 3, Section 3.6, the starting values are picked from a 400% offset: the latter may be called a UKF "bad" initialization. The likelihood of measurement of simulated data is shown in Figure 4.7. In every case, even with a small percentage deviation from the true values (i.e. bottom panel of Figure 4.7), the likelihood surface is unstable and multi-modal. The search of the absolute maximum of such a function is a tricky task, and Figure 4.7 demonstrates that gradient-based optimizer are not a wise choice.

In this study, the investigation concerns the necessity of a learning algorithm to choose the "best" sigma set $\boldsymbol{\theta}$. Moreover, what follows compares three optimizing methods with two acquisition functions in the BO scheme and the discrete grid search with a cubic spline interpolation (Table 4.2).

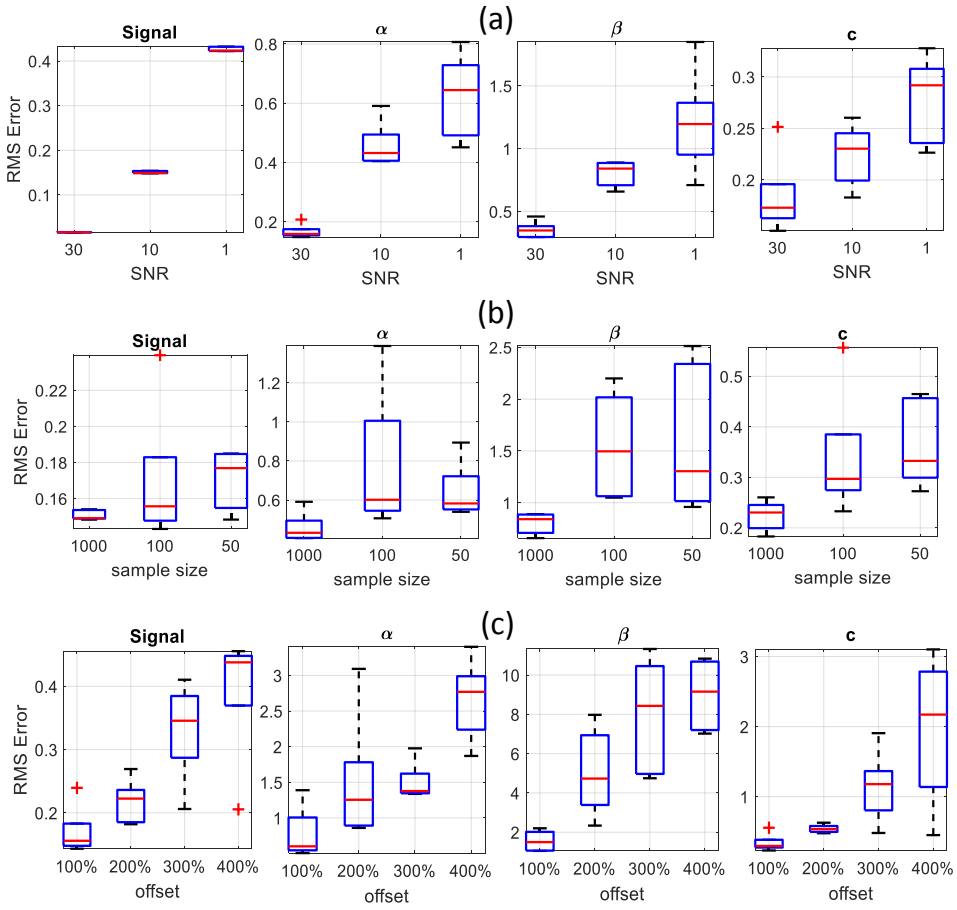

Figure 4.6: Deterministic Duffing system: RMS error of UKF estimates for different sizes of SNR, n and offsets. (a) fixed $n = 1,000$, offset = 100%, and varying SNR; (b) fixed SNR = 10, offset = 100% and varying n; (c) fixed SNR = 10 and $n = 1,000$, varying offsets. Results are averaged over 50 independent datasets.

With respect to BO, the acquisition functions are optimized using the LH design, with the number of initial input points set to $10 \times d$. The GP is defined with constant mean function and Matérn $\nu = 5/2$ covariance function, which leads to twice differentiable sample paths (Pasetto et al., 2017b). For the discrete grid, the loss function $\mathscr{L}(\boldsymbol{\theta})$ is interpolated with a cubic convolution taking as many points as the size of the grid, and get the maximum of the interpolated function.

The results in Table 4.2 suggest that the optimization of the likelihood function

	α		β		c	
	Before	After	Before	After	Before	After
100%	1.00	0.36	2.04	1.20	0.10	0.02
150%	1.52	0.12	3.02	0.50	0.15	0.01
200%	2.03	0.23	3.90	0.94	0.21	0.02
250%	2.48	0.69	5.04	9.36	0.25	0.04
300%	2.72	1.01	6.00	9.62	0.32	0.99
400%	3.95	1.65	7.89	10.99	0.40	1.56

Table 4.1: Deterministic Duffing system: impact of the initialization for different offsets (as percentage deviation from the true parameter values) in term of Euclidean norm before and after inference. Results are averaged above 50 independent datasets.

improves the poor outcome of the default sigma points location. All the methods reach the aim, i.e. the final parameter estimates are closer to the true values than the default UKF sigma points assignment in the case of "bad" initialization. The EI acquisition function achieves the better performance compared to the other methods, since the UCB is a too optimistic acquisition function. Figures 4.8 and 4.9 show the computational costs of the EI and UCB.

4.4. *Sequential ABC-UKF estimates*

Following the method described in Section 3.8 of Chapter 3, the ABC-SMC inputs are defined as follows. The other steps of the Sequential ABC-UKF, i.e. the UKF initialization, the data generating process and the optimization methods settings, are the same of Sections 4.1 - 4.3. The number of particles is fixed to $N = 1,000$. The prior distribution for the parameters is a Uniform and the extremes of the interval are the 400% deviation from the true values. To avoid instability, the parameter $\beta \in \mathbb{R}^+$, and, hence, the left extreme value of the Uniform for β is truncated at zero. The perturbation kernel is a Gaussian distribution centered at $\boldsymbol{\lambda}^{**}$ and with variance τ_t^2. The distance function is the Euclidean norm, while the thresholds

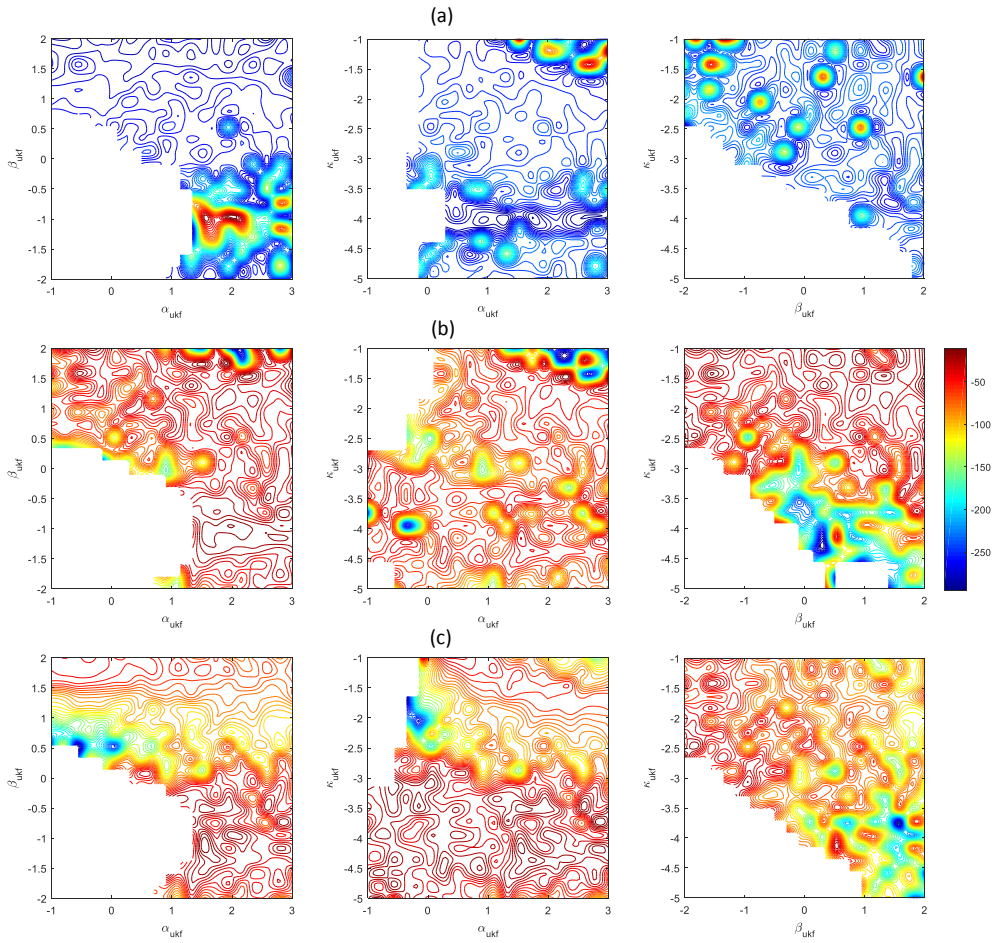

Figure 4.7: Deterministic Duffing system: log-likelihood of measurements for different offset. (a) High offset (400% of deviation from the true values). (b) Medium offset (250% of deviation from the true values). (c) Low offset (100% of deviation from the true values). The white spaces are due to numerical instability when inverting the Kalman gain matrix.

are generated as linearly spaced vector in the summary statistics space.

Let us consider three summary statistics: (i) the Ratio of the Peaks (RP), which is the ratio between the amplitude of the first peak and the amplitude of the last peak of the oscillation, (ii) the dominant frequency (DF) of the fast Fourier transform (FFT), and (iii) the number of zero crossings (ZC) of the signal. The choice of the statistics relies on the characteristics of the three Duffing parameters. The DF detects the predominant frequency of oscillation, which is captured by α, while the RP is more suitable to identify the damping term. The ZC statistics is evaluated to find

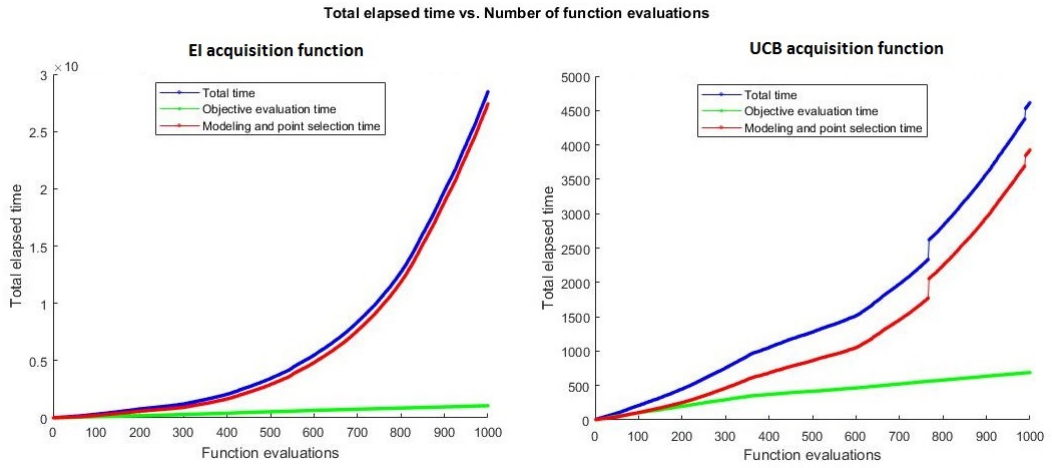

Figure 4.8: Computational costs of EI and UCB acquisition functions.

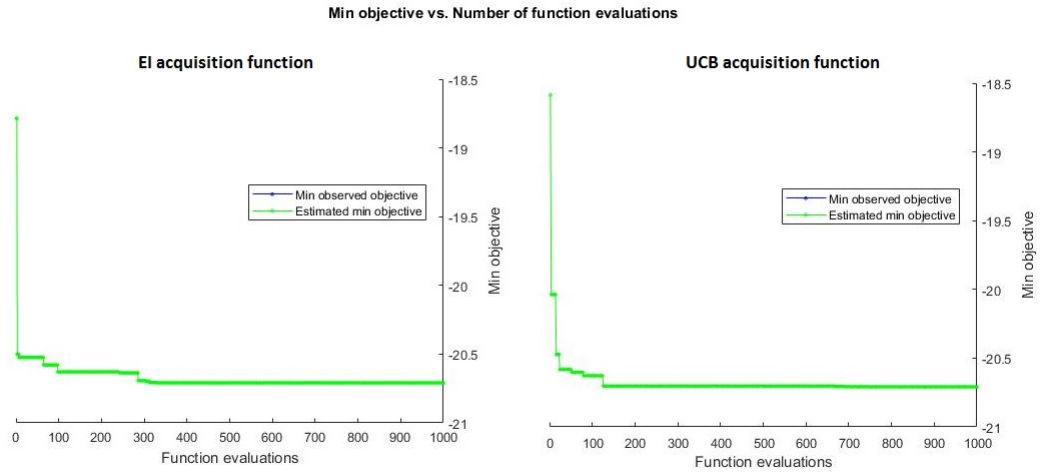

Figure 4.9: Objective function evaluation and number of acquisition function computations. Comparison between EI and UCB.

β, which influences the stability of the trivial fixed points in the phace-space and, so, the convergence to zero of the oscillation in the time domain, as pointed out in Chapter 2.

The univariate and bivariate posterior distributions of the ABC-SMC are shown in Figures 4.10 and 4.11. The ABC-SMC posterior distributions for α and c are

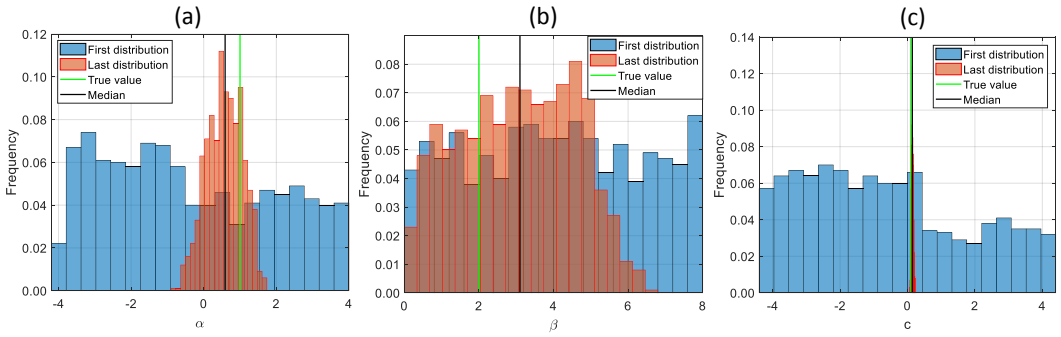

Figure 4.10: Deterministic Duffing system: univariate first and posterior distributions of parameters in the ABC-SMC scheme initialized with 400% offset. (a) Estimate of parameter α. (b) Estimate of parameter β. (c) Estimate of parameter c.

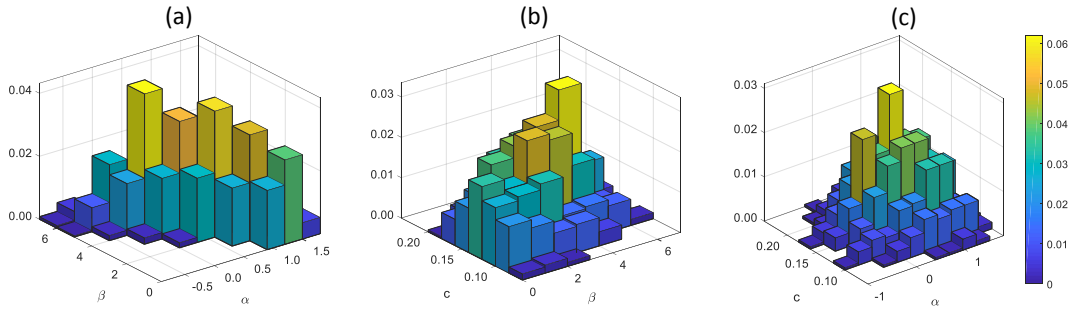

Figure 4.11: Deterministic Duffing system: bivariate posterior distributions of parameters after ABC-SMC scheme with 400 % offset. (a) Distribution of (α, β). (b) Distribution of (c, β). (c) Distribution of (c, α).

peaked near their true values. The most difficult parameter to infer is β, the term associated to the chaotic behaviour. In order to obtain 1,000 particles per population in the SMC intermediate distributions, the overall generating process consists of about 500,000 particles and the final acceptance rate is 0.04. The number of effective particles at the end of the intermediate distributions is the 95% of the original number of particles. The median values of the posterior distributions of the ABC-SMC are the starting values for the UKF.

Table 4.3 summarizes the inference achievements with the Sequential ABC-UKF in the parameter space. The new method converges to the true values in the sense that the true parameters lie in the estimated confidence interval (Figure 4.12). Com-

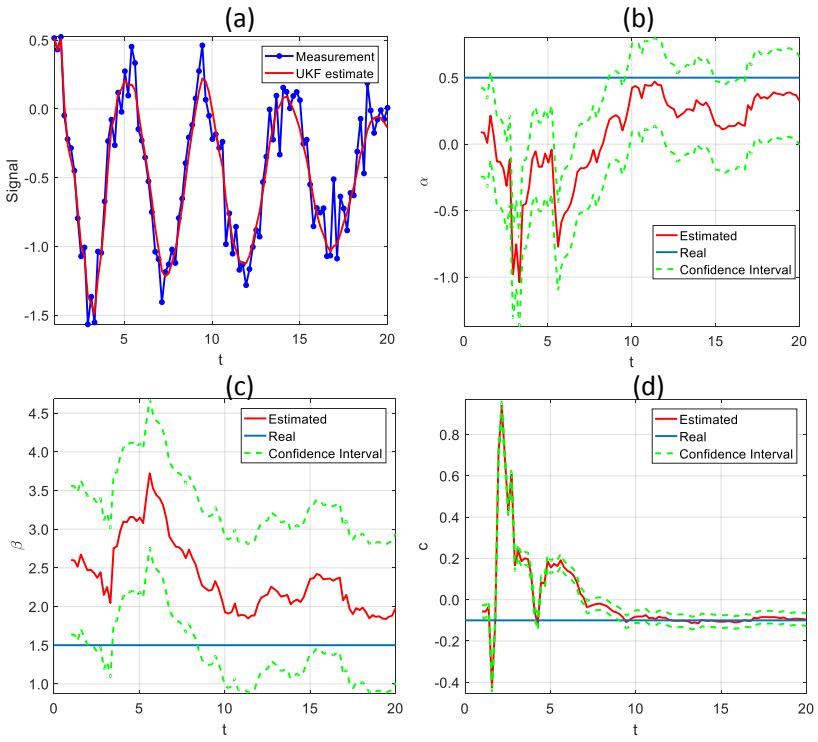

Figure 4.12: Deterministic Duffing system: Sequential ABC-UKF estimates in the time domain. (a) Signal estimate. (b) Estimate of parameter α. (c) Estimate of parameter β. (d) Estimate of parameter c.

pared with the default UKF, the Sequential ABC-UKF shows a massive improvement: the first let increase the distance between the estimates and the true values of 239%, while the latter cuts the same distance of 6.65. The estimates of α and c at the end of the filtering phase always lie within the predicted standard error around the estimate. Instead, the value of β is more affected by uncertainty. Figure 4.13 shows the RSS in functional and parameter space. The Signal RSS is low but not equal to zero; indeed, from the top-left panel of Figure 4.12 it is noticeable that the signal reconstruction is smooth and it reproduces the observation path, except for some peaks in the measurements. The Solution RSS is higher than the Signal RSS, meaning that inserting the final parameter estimates of the filtering phase in the ordinary differential equation comes out in more jagged path.

As far as the comparison between optimizing methods is concerned, Table 4.3

	Norm in parameter space	
Optimization method	Before inference	After inference
Default UKF	6.90	16.50
UKF with EI	6.90	3.60
UKF with UCB	6.90	4.43
UKF with Grid	6.90	4.09

Table 4.2: Deterministic Duffing system: standardized Euclidean norm in the parameter space before and after UKF filtering with 400% offset. Comparison between default sigma points location and optimizing methods.

	Norm in parameter space	
Method	Before inference	After inference
Default UKF	6.90	16.50
Sequential ABC-UKF with EI	6.90	0.25
Sequential ABC-UKF with UCB	6.90	0.27
Sequential ABC-UKF with Grid	6.90	0.26

Table 4.3: Deterministic Duffing system: standardized Euclidean norm in the parameter space before and after UKF filtering. Comparison between the default UKF and Sequential ABC-UKF method with different optimization schemes. Initialization is 400% offset.

points out that there is little difference among them. Tables 4.2 and 4.3 highlight that the discrepancies between optimization algorithm disappear as the UKF reaches a good initialization. The EI acquisition function comes out to have the smallest Euclidean distance in the parameter space but its predominance on the UCB and discrete grid shrinks if compared with the results in Table 4.2. The evaluation of the UKF filtering performance coupled with optimized sigma points location and with respect to a thoughtful research of starting values brings out the conclusion that the initialization is more determinant than sigma points placement to convergence to the true parameters.

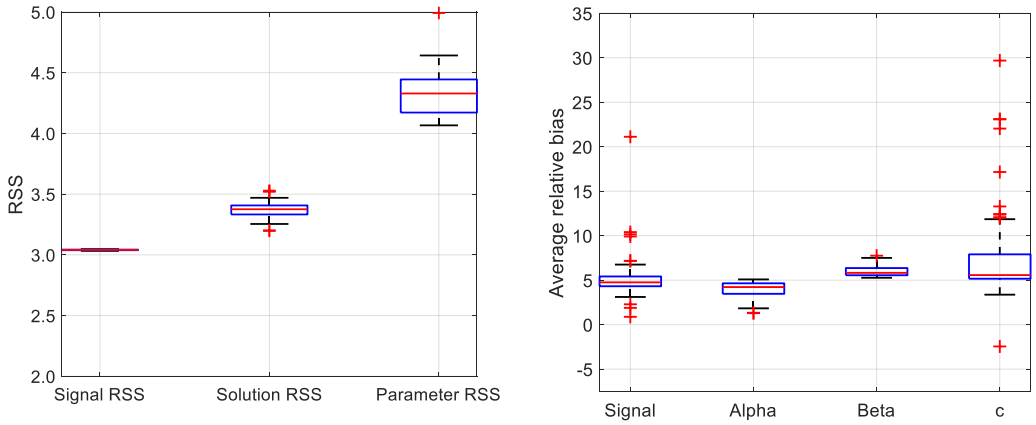

Figure 4.13: Deterministic Duffing system: RSS and ARB of Sequential ABC-UKF estimates in functional and parameter space.

4.5. *Discussion*

In the previous Chapter, two approaches to the signal and parameter estimation for ODEs have been presented, that are the filtering methods and the likelihood free schemes. After the resume of the main concepts and statistical theories behind these techniques, the Achille's heel of each method was discussed and to overcome such limits I proposed a new algorithm in the context of the Duffing oscillator, called Sequential ABC-UKF. I studied three optimization algorithms in a highly multimodal likelihood space and investigated their results in terms of inference within the UKF framework.

The contribution to the methodology to deal with the estimation of chaotic ODEs concerns a UKF-based method with the novelty of the ABC-SMC procedure nested in it. The proposed algorithm is able to infer a credible domain in which the true parameters of a system lie and then executes filtering methods to pursue convergence.

The presented simulation study evaluates the UKF sensibility in terms of accuracy of inference depending on (i) the level of noise and sample size, (ii) the sigma point position in the likelihood space, (iii) the starting values. Three methods of optimization are compared in terms of improvement in the UKF results.

However, the optimizing procedure is not as crucial as the initialization is. Indeed, I tested the better performance the Sequential ABC-UKF has with respect to the default UKF scheme, proving that the more the starting values go away from the true values, the more UKF results are poor. A slight change in the starting values affects the UKF more than the sigma points location and the method here proposed encompasses the UKF limits.

5. Inference for stochastic differential equations

Stochastic differential equations (SDEs) represent dynamical systems with a random component. Stochastic perturbations lead to different oscillatory behaviours; thus, a unique solution for an SDE does not exist. In the remainder of this Chapter, the term "signal" is used but it refers to one possible instantiation of the random number seed.

Inference for stochastic processes is performed with the UKF. The latter is slightly modified to include into the filtering phase the approximate solution of the SDE. For the evaluation of the transition function of the state model, intractable stochastic integrals may occur when constructing high order integration schemes, like the Runge-Kutta method of fourth order. In this case, only low order integration method, as the Euler-Maruyama scheme, can be applied. The Euler-Maruyama numerical solution is inserted into the UKF steps to solve the SDE for obtaining the predictive density. The mathematics behind this method is fully explained in what follows.

The presence of a random disturbance into the process affects the sigma set. To include the effect of noise, the UKF implemented in the same way of the ODE case should draw sigma points for the state and observations, and then redraw another set when adding the process noise. A repeated evaluation of sigma points can be risky, since, as already shown in Chapter 3, the default deterministic choice of sigma points location may result in less accurate estimates. As will be discussed later on, to avoid multiple computations of sigma points, the covariance matrix of the state is augmented to account for the variance of the process noise. The "augmented" UKF is used for all the following simulation studies.

The simulations presented below focus on three aspects. First, an evaluation of the Euler-UKF method for inference in the stochastic Duffing system is presented. The algorithm is also compared on an independent dataset and on a new method developed in the particle filtering literature. Second, the variance of the process noise is estimated: this is a crucial step to move to real data applications where usually the

level of noise is unknown.

The Chapter is organized as follows. Section 5.1 describes the UKF for SDEs. Section 5.2 carries out a simulation study to compare the stochastic simulation setting with the deterministic one of Chapter 4. A description of the augmented UKF and a comparison with the non-augmented version is presented in Section 5.3, while the estimation of the noise variance is discussed in Section 5.4. Comparisons between two simulation settings and with an independent dataset are shown in Section 5.5 and Section 5.6, respectively. Section 5.7 presents a preliminary study to initialize the UKF for SDEs. Finally, results are discussed in Section 5.8.

5.1. *Stochastic state space models*

5.1.1. Stochastic state models

Consider a stochastic differential equation of the form

$$d\boldsymbol{x}(t) = \boldsymbol{F}(\boldsymbol{x}(t), \boldsymbol{\lambda})\, dt + \sigma_\varepsilon dW(t), \tag{5.1}$$

where $W(t)$ is a Wiener process, $\boldsymbol{x}(t) \in \mathbb{R}^d$ is the hidden state at time t, $\boldsymbol{\lambda}$ is a constant parameter vector, $\boldsymbol{F}$ is a transition function and σ_ε is the linear diffusion term. The initial conditions are $\boldsymbol{x}(0) = \boldsymbol{x}_0$ in the time interval $0 \leq t \leq T$. If $\sigma_\varepsilon = 0$ and $\boldsymbol{x}_0$ is constant, equation (5.1) reduces to an ODE with $d\boldsymbol{x}(t) = \boldsymbol{F}(\boldsymbol{x}(t), \boldsymbol{\lambda})dt$.

Occasionally, in some applications, non-mathematical papers may divide all the terms of equation (5.1) by dt, obtaining the following

$$d\boldsymbol{x}(t)/dt = \boldsymbol{F}(\boldsymbol{x}(t), \boldsymbol{\lambda}) + \sigma_\varepsilon \boldsymbol{\varepsilon}(t), \tag{5.2}$$

where $\boldsymbol{\varepsilon}(t) = dW(t)/dt$ is the process noise, $\boldsymbol{\varepsilon} \sim N(0, \sigma_\varepsilon^2)$. Note that, since Brownian motion is nowhere differentiable with probability 1, the ratio $dW(t)/dt = \boldsymbol{\varepsilon}(t)$ is not mathematically meaningful. However, in the practical implementations of the UKF described later, the term $\boldsymbol{\varepsilon}(t)$ is considered because it can be computed through the Itô calculus.

Equation (5.1) can be written in integral forms like

$$\boldsymbol{x}(t) = \boldsymbol{x}_0 + \int_0^t \boldsymbol{F}(\boldsymbol{x}(s), \boldsymbol{\lambda})ds + \int_0^t \sigma_\varepsilon dW(s). \tag{5.3}$$

The second integral in the right-hand side of (5.3) is a *stochastic integral*, i.e. it is taken with respect to the Brownian motion. This can be evaluated with the Itô calculus, which is sketched in the next Section.

5.1.2. Stochastic integrals

Consider an interval $[0, T]$ to be partitioned into L equal subintervals of width $\delta t > 0$, such that $0 < \delta t < 2\delta t < \ldots < L\delta t = T$. In what follows, the term δt is called the *stepsize of integration*.

Given a function h, the integral $\int_0^T h(t)dt$ can be approximated by the Riemann sum

$$\sum_{j=0}^{L-1} h(t_j)(t_{j+1} - t_j), \tag{5.4}$$

where the discrete time points $t_j = j\delta t$ are the steps of integration. The Riemann integral may be defined as the limit of $\delta t \to 0$ of sum (5.4).

By analogy with equation (5.4), the stochastic integral

$$\int_0^T h(t)dW(t) \tag{5.5}$$

is approximated by

$$\sum_{j=0}^{L-1} h(t_j)\left(W(t_{j+1}) - W(t_j)\right), \tag{5.6}$$

where the function h is integrated with respect to the Brownian motion. The sum $\sum_{j=0}^{L-1}(W(t_{j+1}) - W(t_j))$ in (5.6) is known as the *Itô integral*. When $h(t) \equiv W(t)$, in the

limiting case $\delta t \to 0$, this sum is

$$\begin{aligned}\sum_{j=0}^{L-1} W(t_j)\left(W(t_{j+1}) - W(t_j)\right) &= \sum_{j=0}^{L-1}\left(W(t_j)W(t_{j+1}) - W(t_j)^2\right)\\ &= \sum_{j=0}^{L-1}\left(W(t_j)W(t_{j+1}) - W(t_j)^2 + W(t_{j+1})^2 - W(t_{j+1})^2\right)\\ &= \frac{1}{2}\sum_{j=0}^{L-1}\left(W(t_{j+1})^2 - W(t_j)^2 - \left(W(t_{j+1}) - W(t_j)\right)^2\right)\\ &= \frac{1}{2}\left(W(T)^2 - W(0)^2 - \sum_{j=0}^{L-1}\left(W(t_{j+1}) - W(t_j)\right)^2\right). \qquad (5.7)\end{aligned}$$

The term $\sum_{j=0}^{L-1}\left(W(t_{j+1}) - W(t_j)\right)^2$ can be shown to have expected value T and variance $O(\delta t)$ (Higham, 2001): for small δt, this random variable is close to the constant T. In particular, when $\delta t \to 0$, the distance between the computed sum (5.6) and the limiting case (5.7), i.e. the Itô error, is

$$\int_0^T W(t)dW(t) = \frac{1}{2}W(T)^2 - \frac{1}{2}T. \qquad (5.8)$$

The Itô calculus is crucial in the analysis of SDEs and in mathematical modelling. Indeed, the SDE (5.3) can be evaluated only with the tools of stochastic calculus. The underlying theory on Itô integrals and Brownian motion is deepened in Karatzas and Shreve (1991).

5.1.3. Discrete stochastic state space models

The measurement model is

$$\boldsymbol{y}(t) = \boldsymbol{H}\left(\boldsymbol{x}(t)\right) + \boldsymbol{\eta}(t), \qquad (5.9)$$

where $\boldsymbol{y}(t) \in \mathbb{R}^D$ are the observations, $\boldsymbol{H}$ is the measurement function and $\boldsymbol{\eta}(t) \sim N(0, \sigma_\eta^2)$ represents the observation noise. The noise assumptions are the same as listed in Chapter 3, Section 3.1.

As previously discussed, since in real-world scenarios data are time-discrete, the system (5.1) has to be discretized and then integrated over the sampling time interval $\Delta t > 0$. Remember from Chapter 3 that the difference equations for the state

equation (5.1) and the space model (5.9) are

$$\boldsymbol{x}_t = \boldsymbol{f}\left(\boldsymbol{x}_{t-\Delta t}, \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_t\right), \qquad \boldsymbol{y}_t = \boldsymbol{h}\left(\boldsymbol{x}_t\right) + \boldsymbol{\eta}_t, \tag{5.10}$$

where $\boldsymbol{f}$ and $\boldsymbol{h}$ are, respectively, the discrete transition and observation functions, $\boldsymbol{x}_{t-\Delta t}$ and $\boldsymbol{x}_t$ are the discretized states variables. The measurements at sampling time t are $\boldsymbol{y}_t$, while the time-discrete process noise and the observation disturbance are $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$.

Integrating the transition function of equation (5.10) gives

$$\boldsymbol{f}\left(\boldsymbol{x}_{t-\Delta t}, \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_t\right) = \boldsymbol{x}_{t-\Delta t} + \int_{t-\Delta t}^{t} \left(\boldsymbol{F}\left(\boldsymbol{x}(T), \boldsymbol{\lambda}\right) + \boldsymbol{\varepsilon}(T)\right) dT. \tag{5.11}$$

Intractable stochastic integrals may occur in the right-hand side of (5.11) when using high order integration schemes like the Runge-Kutta method of fourth order. Hence, to numerically integrate equation (5.11), only low order methods can be applied, such as the Euler-Maruyama scheme (Sitz et al., 2002), which is briefly described in the following.

5.1.4. The Euler-Maruyama method

The Euler-Maruyama method (the stochastic extension of the Euler method) is a scheme for the approximate numerical solution of SDEs.

Consider the following general form of the SDE:

$$d\boldsymbol{x}(t) = \boldsymbol{f}(\boldsymbol{x}(t))dt + \boldsymbol{g}(\boldsymbol{x}(t))dW(t), \tag{5.12}$$

where the $\boldsymbol{f}$ and $\boldsymbol{g}$ are the drift and diffusion terms, respectively. The SDE (5.12) has to be evaluated in an interval $[0, T]$, with initial conditions $\boldsymbol{x}(0) = \boldsymbol{x}_0$.

Once the interval $[0, T]$ has been divided into subintervals like in Section 5.1.2, the Euler-Maruyama approximation recursively defines

$$\boldsymbol{x}_j = \boldsymbol{x}_{j-1} + \boldsymbol{f}(\boldsymbol{x}_{j-1})\delta t + \boldsymbol{g}(\boldsymbol{x}_{j-1})\left(W(j\delta t) - W((j-1)\delta t)\right), \quad j = 1, \ldots, L, \tag{5.13}$$

where $\boldsymbol{x}_j$ denotes the numerical approximation to $\boldsymbol{x}(j)$. To understand where equation (5.13) comes from, consider equation (5.3). In terms of the discrete small step-size, δt, the integral (5.3) can be

$$\boldsymbol{x}_{j\delta t} = \boldsymbol{x}_{\delta t(j-1)} + \int_{\delta t(j-1)}^{j\delta t} \boldsymbol{f}(\boldsymbol{x}_s, \boldsymbol{\lambda})ds + \int_{\delta t(j-1)}^{j\delta t} \sigma_\varepsilon dW(s). \tag{5.14}$$

Each of the three terms on the right-hand side of equation (5.13) approximates the corresponding term of the right-hand side of (5.14).

5.1.5. The Euler-Maruyama scheme within the UKF framework

In order to obtain a reliable numerical approximation, the use of the Euler-Maruyama scheme requires an integration step δt considerably smaller than the sampling time interval Δt. Therefore, several integration steps in the interval $[t - \Delta t, t]$ are necessary to predict the statistics of the state function at time t. The type of equations like (5.11) are called Langevin equations (Sitz et al., 2002), and the Euler-Maruyama method in this case reads

$$\boldsymbol{x}_{t-\Delta t+\delta t} = \boldsymbol{x}_{t-\Delta t} + \delta t \boldsymbol{F}\left(\boldsymbol{x}_{t-\Delta t}, \boldsymbol{\lambda}\right) + \sqrt{\delta t}\boldsymbol{\varepsilon}_{t-\Delta t}. \tag{5.15}$$

Let us highlight that the random variable $\boldsymbol{x}_{t-\Delta t}$ is completely characterized by its probability density. Looking at the evolution of this density, it is possible to predict the state $\boldsymbol{x}_{t-\Delta t}$. The reconstruction of the time evolution of the density function can be done using the respective discrete Fokker-Planck equation or through simulations utilizing a discretization of the density by a finite and representative set of sample $\{\boldsymbol{x}_{i,t-\Delta t}\}$. The discrete density at time $t - \Delta t + \delta t$ is obtained by propagating the following i-th sample

$$\boldsymbol{x}_{i,t-\Delta t+\delta t} = \boldsymbol{x}_{i,t-\Delta t} + \delta t \boldsymbol{F}\left(\boldsymbol{x}_{i,t-\Delta t}, \boldsymbol{\lambda}\right) + \sqrt{\delta t}\boldsymbol{\varepsilon}_{i,t-\Delta t}, \tag{5.16}$$

where $\boldsymbol{\varepsilon}_{i,t-\Delta t}$ is the i-th sample of the stationary Gaussian white noise.

The sigma points for the state and process noise are, respectively,

$$\{\chi_i\left(t - \Delta t | t - \Delta t\right)\}_{i=1}^{2d+1}, \tag{5.17}$$

$$\{\mathsf{E}_i\left(t - \Delta t | t - \Delta t\right)\}_{i=1}^{2d+1}. \tag{5.18}$$

The set of points $\{\boldsymbol{x}_{i,t-\Delta t}\}$, $\{\boldsymbol{\varepsilon}_{i,t-\Delta t}\}$ and the sigma points in equations (5.17)-(5.18) are chosen as described in Section 3.5.1 of Chapter 3. Thus, even when using the sets $\{\boldsymbol{x}_{i,t-\Delta t}\}$ and $\{\boldsymbol{\varepsilon}_{i,t-\Delta t}\}$, only the information provided by the mean and covariance of the full state density is considered during each integration step of the Euler-Maruyama method. The numerical integration step over δt of equation (5.16) is

performed by propagating the sigma points through the Euler-Maruyama scheme. The latter is applied to the following SDE expressed in terms of sigma points

$$\chi_i\left(t-\Delta t+\delta t|t-\Delta t\right)=\boldsymbol{f}_{\text{Euler }(\delta t)}\left(\chi_i\left(t-\Delta t|t-\Delta t\right),\boldsymbol{\lambda},\mathsf{E}_i\left(t-\Delta t|t-\Delta t\right)\right), \quad (5.19)$$

where $\boldsymbol{f}_{\text{Euler }(\delta t)}$ is the state function of equation (5.16) but with respect to the sigma points.

From the set of predicted sigma points $\{\chi_i\left(t-\Delta t+\delta t|t-\Delta t\right)\}_{i=1}^{2d+1}$, the mean and the covariance at time $t-\Delta t+\delta t$ are computed accordingly to equations (3.28)-(3.29) of Section 3.5.2 of Chapter 3. Then, these mean and covariance are used to construct a new set of sigma points $\{\chi_i\left(t-\Delta t+\delta t|t-\Delta t+\delta t\right)\}_{i=1}^{2d+1}$ which is used for the next Euler-Maruyama step performed at time $t-\Delta t+\delta t$.

A stationary noise dynamics is assumed:

$$\{\mathsf{E}_i\left(t-\Delta t+\delta t|t-\Delta t\right)\}_{i=1}^{2d+1}=\{\mathsf{E}_i\left(t-\Delta t|t-\Delta t\right)\}_{i=1}^{2d+1},\ i=1,\ldots,d. \quad (5.20)$$

Equation (5.20) indicates that the predictions for the sigma points that represent the density of the process noise is constant over δt.

5.1.6. Joint state space representation

The dynamics of the parameter vector $\boldsymbol{\lambda}$ and the stationary measurement noise $\boldsymbol{\eta}$ are considered within the Euler-Maruyama method in the UKF framework to construct a joint state space representation like

$$\begin{pmatrix}\chi_i\left(t-\Delta t+\delta t|t-\Delta t\right)\\ \boldsymbol{\lambda}_i\left(t-\Delta t+\delta t|t-\Delta t\right)\\ \mathsf{E}_i\left(t-\Delta t+\delta t|t-\Delta t\right)\\ \boldsymbol{\eta}_i\left(t-\Delta t+\delta t|t-\Delta t\right)\end{pmatrix}=$$

$$=\begin{pmatrix}\boldsymbol{f}_{\text{Euler }(\delta t)}\left(\chi_i\left(t-\Delta t|t-\Delta t\right),\boldsymbol{\lambda}_i\left(t-\Delta t|t-\Delta t\right),\mathsf{E}_i\left(t-\Delta t|t-\Delta t\right)\right)\\ \boldsymbol{\lambda}_i\left(t-\Delta t|t-\Delta t\right)\\ \mathsf{E}_i\left(t-\Delta t|t-\Delta t\right)\\ \boldsymbol{\eta}_i\left(t-\Delta t|t-\Delta t\right)\end{pmatrix}. \quad (5.21)$$

The unscented transform is repeated over all subsequent integration steps in each interval $[t-\Delta t,t]$, for $t=1,\ldots,T$. At time t, a predicted mean and covariance for

the joint state (5.21) is updated by observation $\boldsymbol{y}_t$, accordingly to equations (3.15) and (3.17) of Section 3.2 in Chapter 3.

Compared with the UKF for ODEs, this version of the UKF matching the Euler-Maruyama scheme is computationally more expensive, since in each sampling time, several integration steps have to be computed.

This method is used for simultaneous signal and parameter estimation for the stochastic Duffing system.

5.1.7. Stochastic Duffing system

The stochastic Duffing system is

$$\ddot{x} + c\dot{x} + \alpha x + \beta x^3 = \sigma_\varepsilon \varepsilon_t, \tag{5.22}$$

where the parameters of interest are α, the frequency of oscillation, β, the mode of the restoring force, c, the damping term and σ_ε, the process noise variance. Equation (5.22) can be written in form of state space equations:

$$dx_{1t}/dt = x_{2t}, \quad dx_{2t}/dt = -\left(cx_{2t} + \alpha x_{1t} + \beta x_{1t}^3\right) + \sigma_\varepsilon \varepsilon_t, \tag{5.23}$$

where the second component dx_{2t}/dt is driven stochastically by an uncorrelated noise $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$.

With respect to the Euler-Maruyama scheme within the UKF method described above, the (5.23) reads

$$x_{1,t-\Delta t+\delta t} = x_{1,t-\Delta t} + \delta t x_{2,t-\Delta t}, \tag{5.24}$$

$$x_{2,t-\Delta t+\delta t} = x_{2,t-\Delta t} - \delta t \left(cx_{2,t-\Delta t} + \alpha x_{1,t-\Delta t} + \beta x_{1,t-\Delta t}^3\right) + \sqrt{\delta t}(\sigma_\varepsilon \varepsilon_t). \tag{5.25}$$

In the remainder of this Chapter, several simulation studies are presented for system (5.23).

5.2. *Simulation study of SDE*

In what follows, the UKF method described in Section 5.1 is used to estimate the signal and parameters of the SDE (5.23). Notice that, due to the stochasticity of the

system, the term "signal" means one instantiations. Indeed, while for an ODE only one solution exists (either in closed form or approximation), the SDEs, since they incorporate a stochastic element, have different oscillations for different random seeds. Therefore, the simulated data are obtained through the Euler-Maruyama approximation method with stepsize of integration $\delta t = 0.001$ and the default MATLAB random number generator. The value of δt has been chosen after several trials. On the one hand, as long as the stepsize grows, the Euler-Maruyama method gets worse. This has been demonstrated taking as reference an ODE and setting to zero the noise term in the SDE: if the ODE solver coincides with the Euler-Maruyama method when $\sigma_\varepsilon = 0$, than the Euler-Maruyama approximation is reliable. Hence, the bigger the stepsize, the stronger the deviation with respect to the ODE. On the other hand, small values of δt drive the computational costs in the UKF iterative scheme. Setting $\delta t = 0.001$ is a compromise between the accuracy of the numerical approximation and the expensive algorithmic steps. This choice also finds support in the paper of Sitz et al. (2002), who use the same value for the analysis of the stochastic Van der Pol system.

Measurements y_t, for $t = 1, \ldots, T = 20$ and sample size $n = 100$, are generated by corrupting the first component of the state, x_{1t}, with independent and identically distributed Gaussian noise of SNR $= 10$. The true parameters are $\alpha = 1$, $\beta = 2$, $c = 0.1$. In this simulation study, the process noise variance is known and generated as a linearly spaced vector between 0.001 and 2. The value of $\sigma_\varepsilon = 0.001$ is considered a reference to quantify how the process noise affects the UKF estimates with respect to a quasi-deterministic setting.

For ease of comparison, the starting values are defined as in Chapter 4: the offsets are a percentage deviation from the true values, and they are 100% (low offset), 250% (medium offset), 400% (high offset). The code is implemented in MATLAB with the `EKF/UKF` toolbox of Hartikainen et al. (2011) for the computation of the UKF, but the Euler-Maruyama method inside the UKF steps and the corrections for numerical stability of matrices are hand-coded.

The UKF performance is evaluated with the average relative bias (ARB) and the

root mean squared (RMS) errors, defined like in Chapter 4:

$$\mathrm{ARB}^{(\mathrm{ukf})} = \frac{\hat{\boldsymbol{y}}_i^{(\mathrm{ukf})} - \boldsymbol{y}_i}{\boldsymbol{y}_i}, \tag{5.26}$$

$$\mathrm{ARB} = \frac{\hat{\boldsymbol{\lambda}}_i - \boldsymbol{\lambda}_i}{\boldsymbol{\lambda}_i}, \tag{5.27}$$

$$\mathrm{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_i\right)^2}, \tag{5.28}$$

$$\mathrm{RMS}^{(\text{Signal-UKF})} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\boldsymbol{y}}_i^{(\mathrm{ukf})} - \boldsymbol{y}_i\right)^2}, \tag{5.29}$$

$$\mathrm{RMS}^{(\text{Signal-SDE})} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\boldsymbol{y}}_i^{(\mathrm{sde})} - \boldsymbol{y}_i\right)^2}, \tag{5.30}$$

where "Signal UKF" indicates the oscillation estimate of the UKF, $\hat{\boldsymbol{y}}_i^{(\mathrm{ukf})}$, and "Signal SDE" ($\hat{\boldsymbol{y}}_i^{(\mathrm{sde})}$) refers to the signal reconstruction obtained when the final UKF parameter estimates, $\hat{\boldsymbol{\lambda}}_i$, are inserted back into the SDE. Since each SDE solution varies stochastically, the approximation is computed 10 times, i.e. the SDE is numerically solved 10 times from different random seeds with the parameter estimates, and then the $\hat{\boldsymbol{y}}_i^{(\mathrm{sde})}$ is the average of 10 approximations. The UKF estimates are averaged over 50 independent datasets.

The ARB is depicted in Figures 5.1-5.5, while Figures 5.6-5.10 show the RMS of the simulations. As expected, the larger the system noise, the poorer the performance. In particular, the tails of the boxplots point out that as far as the process noise variance increases, the independent evaluations of the UKF on different datasets are more spread. Notice that in the deterministic case the most difficult parameter to infer was the non-linear stiffness β. Here, instead, the damping term c is affected by a higher RMS and seems more crucial to estimate than the other parameters. This can be due to the fact that the process noise mostly affects the decay of the oscillation. In Chapter 2 it has been heuristically shown that a slight change in the variance value affects the stability of the trivial fixed point: the parameter governing this kind of stability is precisely the damping term.

The Signal-SDE has a poorer reconstruction (a bigger RMS) than the Signal-UKF. The latter gets worse as the offsets increases, while the Signal-SDE is constantly farer from the reconstruction of the generated path independently on the level of offsets.

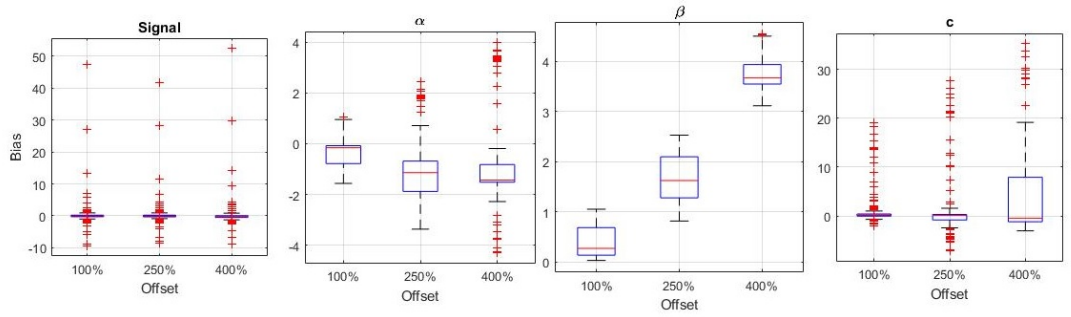

Figure 5.1: Average relative bias of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$, $\sigma_\varepsilon^2 = 0.001$.

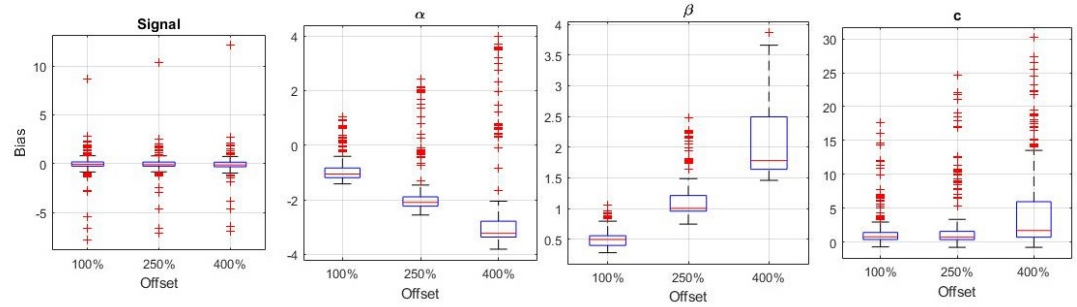

Figure 5.2: Average relative bias of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$, $\sigma_\varepsilon^2 = 0.5$.

5.3. *Augmented vs. Non-Augmented UKF*

The simulation study presented in Section 5.2 has been carried out with the *augmented* UKF. Essentially, the idea behind the "augmentation" of the UKF consists in considering the covariance of the state along with the variance of the process noise in a unique covariance matrix. In the context of ODEs, the state noise is $\boldsymbol{\varepsilon}_t = 0$ and

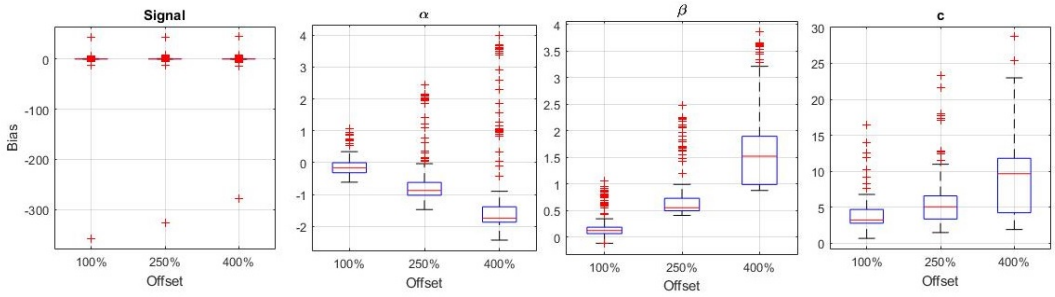

Figure 5.3: Average relative bias of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$, $\sigma_\varepsilon^2 = 1.0$.

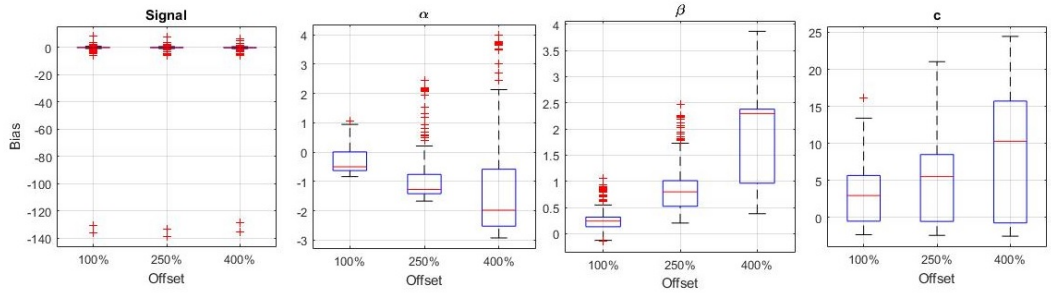

Figure 5.4: Average relative bias of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$, $\sigma_\varepsilon^2 = 1.5$.

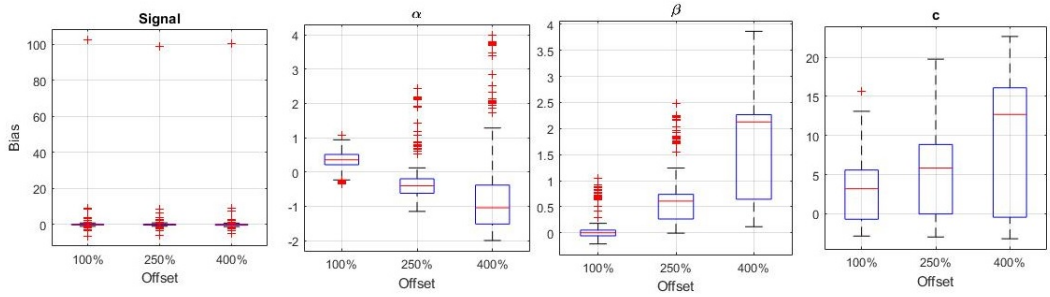

Figure 5.5: Average relative bias of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$, $\sigma_\varepsilon^2 = 2.0$.

there is no need to include its variance in the filtering steps, but when the process disturbance exists, it has to be considered into the inference scheme.

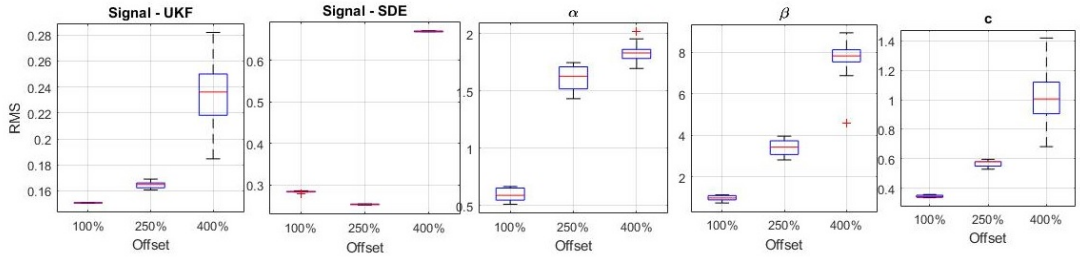

Figure 5.6: RMS of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$, $\sigma_\varepsilon^2 = 0.001$.

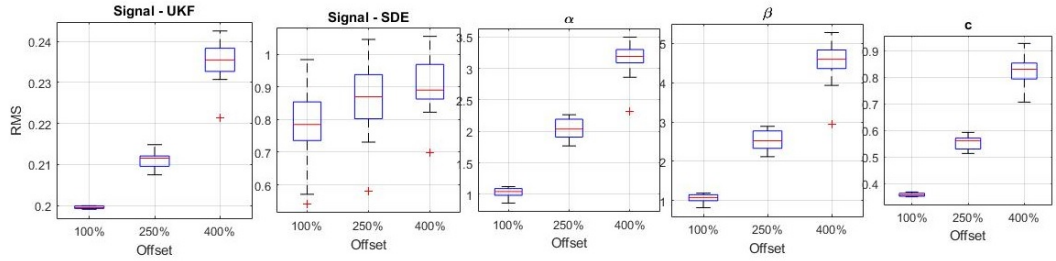

Figure 5.7: RMS of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$, $\sigma_\varepsilon^2 = 0.5$.

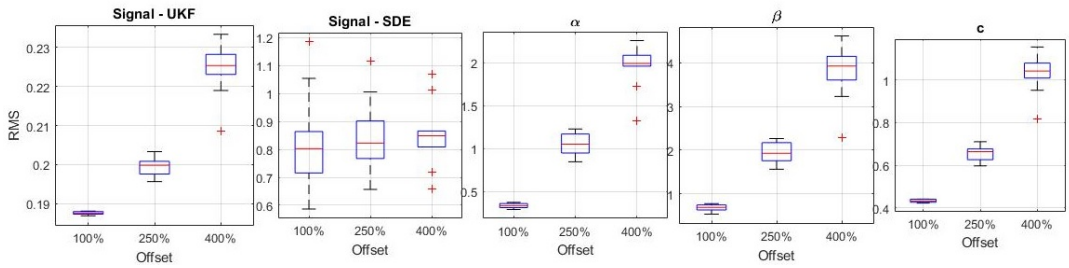

Figure 5.8: RMS of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$, $\sigma_\varepsilon^2 = 1.0$.

The augmented UKF draws sigma set only once within a filtering recursion, while the non-augmented UKF, i.e. the UKF computed exactly as in the ODE case, has to redraw a new set of sigma points during the prediction phase to incorporate the effect of additive process noise (Wu et al., 2005). As stated in Section 3.5.4 of Chapter 3, the sigma points location in the likelihood space is so crucial that drawing

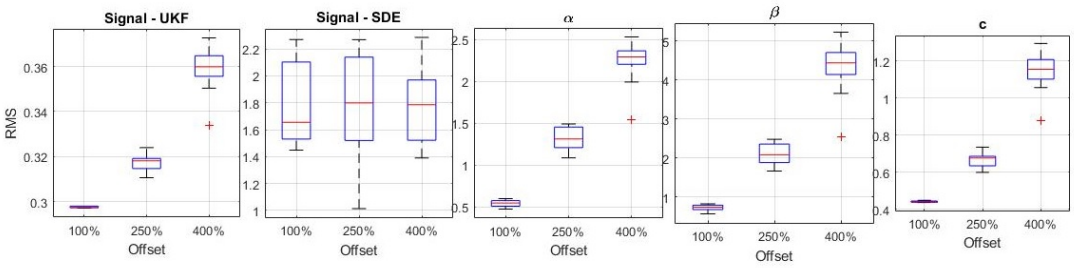

Figure 5.9: RMS of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$, $\sigma_\varepsilon^2 = 1.5$.

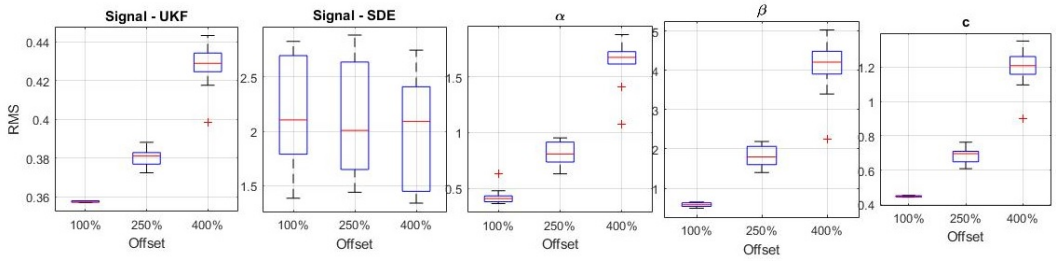

Figure 5.10: RMS of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$, $\sigma_\varepsilon^2 = 2.0$.

twice the sigma set may lead to noticeable losses in accuracy.

In the mathematical description of the non-augmented and augmented UKF, a general setting of equation (5.10) is considered, that is the process and observation noise can be multivariate:

$$\boldsymbol{x}_t = \boldsymbol{f}\left(\boldsymbol{x}_{t-\Delta t}, \boldsymbol{\lambda}, \boldsymbol{\varepsilon}_t\right), \qquad \boldsymbol{y}_t = \boldsymbol{h}\left(\boldsymbol{x}_t\right) + \boldsymbol{\eta}_t, \tag{5.31}$$

where $\boldsymbol{\varepsilon}_t \sim N(0, \Sigma_\varepsilon)$ and $\boldsymbol{\eta}_t \sim N(0, \Sigma_\eta)$. The pseudo-codes for the non-augmented and augmented unscented transforms follow in the next Sections.

5.3.1. Non-augmented unscented transform

Step 1. The vector $\boldsymbol{x}_t$ is approximated by $2d + 1$ sigma points computed as discussed in Section 3.5.1 of Chapter 3. For ease of description, the sigma set is reported again:

$$\chi_0\left(t-1|t-1\right) = \boldsymbol{m}\left(t-1|t-1\right), \tag{5.32}$$

$$\chi_i\left(t-1|t-1\right) = \left\{\boldsymbol{m}\left(t-1|t-1\right) + \sqrt{\left(d+\lambda_{\mathsf{ukf}}\right)\boldsymbol{P}_{:i}}\right\}_{i=1}^{d}, \tag{5.33}$$

$$\chi_{i+d}\left(t-1|t-1\right) = \left\{\boldsymbol{m}\left(t-1|t-1\right) - \sqrt{\left(d+\lambda_{\mathsf{ukf}}\right)\boldsymbol{P}_{:i}}\right\}_{i=1}^{d}, \tag{5.34}$$

for $i = 1, \ldots, d$, the scaling parameter is $\lambda_{\mathsf{ukf}} = \alpha_{\mathsf{ukf}}^2(d + k_{\mathsf{ukf}}) - d$, and $\boldsymbol{P}_{:i}$ represents the i-th column of matrix $\boldsymbol{P}$.

The weights associated to the sigma points are

$$w_m^{(0)} = \frac{\lambda_{\mathsf{ukf}}}{d+\lambda_{\mathsf{ukf}}}, \tag{5.35}$$

$$w_c^{(0)} = \frac{\lambda_{\mathsf{ukf}}}{d+\lambda_{\mathsf{ukf}}} + (1 - \alpha_{\mathsf{ukf}}^2 + \beta_{\mathsf{ukf}}), \tag{5.36}$$

$$w_m^{(i)} = w_c^{(i)} = \frac{1}{2(d+\lambda_{\mathsf{ukf}})}, \quad \text{for } i = 1, \ldots, 2d. \tag{5.37}$$

Step 2. Each sigma points is instantiated and yields the sets for the state and observations:

$$\chi_i\left(t|t-1\right) = \boldsymbol{f}\left(\chi_i(t-1|t-1)\right), \tag{5.38}$$

$$\Upsilon_i\left(t|t-1\right) = \boldsymbol{h}\left(\chi_i(t|t-1)\right). \tag{5.39}$$

Step 3. The means and covariances for $\boldsymbol{x}_t$ and $\boldsymbol{y}_t$ are:

$$\boldsymbol{m}(t|t-1) = \sum_{i=0}^{2d} w_m^{(i)}\chi_i(t|t-1), \tag{5.40}$$

$$\boldsymbol{P}(t|t-1) = \sum_{i=0}^{2d} w_c^{(i)}\left(\chi_i(t|t-1) - \boldsymbol{m}(t|t-1)\right)\left(\chi_i(t|t-1) - \boldsymbol{m}(t|t-1)\right)' + \Sigma_\varepsilon, \tag{5.41}$$

$$\boldsymbol{m}_y(t|t-1) = \sum_{i=0}^{2d} w_m^{(i)}\Upsilon_i(t|t-1), \tag{5.42}$$

$$\boldsymbol{P}_y(t|t-1) = \sum_{i=0}^{2d} w_c^{(i)}\left(\Upsilon_i(t|t-1) - \boldsymbol{m}_y(t|t-1)\right)\left(\Upsilon_i(t|t-1) - \boldsymbol{m}_y(t|t-1)\right)' + \Sigma_\eta. \tag{5.43}$$

5.3.2. Augmented unscented transform

Augmenting the UKF means that (i) a new state covariance matrix is generated, $\boldsymbol{P}_a$ which includes both $\boldsymbol{P}$, and Σ_ε, and (ii) the sigma points are computed in the space of the new covariance $\boldsymbol{P}$.

The measurement model of equation (5.31) can be reformulated as

$$\boldsymbol{y}_{a,t} = \boldsymbol{h}_a\left(\boldsymbol{x}_{a,t}\right), \tag{5.44}$$

where the augmented vector $\boldsymbol{x}_{a,t}$ is

$$\boldsymbol{x}_{a,t} = \left[\boldsymbol{x}_t \ \boldsymbol{\eta}_t\right]', \tag{5.45}$$

and the non-linear function $\boldsymbol{h}_a$ is defined as

$$\boldsymbol{h}_a\left(\boldsymbol{x}_{a,t}\right) = \boldsymbol{h}_a\left(\left[\boldsymbol{x}_t \ \boldsymbol{\eta}_t\right]'\right) = \boldsymbol{h}\left(\boldsymbol{x}_t\right) + \boldsymbol{\eta}_t. \tag{5.46}$$

Step 1. Augment the mean vector as

$$\boldsymbol{m}_a(t|t-1) = \left[\boldsymbol{m}(t|t-1)\ 0\right]'. \tag{5.47}$$

The augmented covariance is

$$\boldsymbol{P}_a(t|t-1) = \begin{bmatrix} \boldsymbol{P}(t|t-1) & 0 \\ 0 & \Sigma_\varepsilon \end{bmatrix} = \tag{5.48}$$

$$= \begin{cases} \begin{bmatrix} \boldsymbol{P}(t|t-1) \\ 0 \end{bmatrix}; & \text{for } i = 1, \ldots, d, \\ \begin{bmatrix} 0 \\ \Sigma_\varepsilon \end{bmatrix}; & i = d+1, \ldots, 2d, \end{cases} \tag{5.49}$$

Step 2. Since the sigma points are evaluated taking the i-th column of the covariance state matrix, as in equations (5.32)-(5.34), when the latter is augmented, it follows that the set of sigma points has to be augmented as well. Hence, the augmented set of sigma points has the form

$$\chi_{a,0}\left(t-1|t-1\right) = \boldsymbol{m}_a\left(t-1|t-1\right), \tag{5.50}$$

$$\chi_{a,i}\left(t-1|t-1\right) = \left\{\boldsymbol{m}_a\left(t-1|t-1\right) + \sqrt{\left(d+\lambda_{\mathsf{ukf}}\right)\boldsymbol{P}_{a,:i}}\right\}_{i=1}^{d}, \tag{5.51}$$

$$\chi_{a,i+d}\left(t-1|t-1\right) = \left\{\boldsymbol{m}_a\left(t-1|t-1\right) - \sqrt{\left(d+\lambda_{\mathsf{ukf}}\right)\boldsymbol{P}_{a,:i}}\right\}_{i=1}^{d}. \tag{5.52}$$

Substituting equations (5.49) into equations (5.50)-(5.52) yields

$$\chi_{a,0}\left(t-1|t-1\right)=\begin{bmatrix}\boldsymbol{m}\left(t-1|t-1\right)\\0\end{bmatrix}, \tag{5.53}$$

$$\chi_{a,i}\left(t-1|t-1\right)=\begin{bmatrix}\left\{\boldsymbol{m}\left(t-1|t-1\right)+\sqrt{\left(d+\lambda_{\mathsf{ukf}}\right)\boldsymbol{P}_{:i}}\right\}_{i=1}^{d}\\0\end{bmatrix}, \tag{5.54}$$

$$\chi_{a,i+d}\left(t-1|t-1\right)=\begin{bmatrix}\left\{\boldsymbol{m}\left(t-1|t-1\right)-\sqrt{\left(d+\lambda_{\mathsf{ukf}}\right)\boldsymbol{P}_{:i}}\right\}_{i=1}^{d}\\0\end{bmatrix}, \tag{5.55}$$

$$\chi_{a,i+2d}\left(t-1|t-1\right)=\begin{bmatrix}\boldsymbol{m}\left(t-1|t-1\right)\\\left\{\sqrt{\left(d+\lambda_{\mathsf{ukf}}\right)\Sigma_{\varepsilon}}\right\}_{i=1}^{d}\end{bmatrix}, \tag{5.56}$$

$$\chi_{a,i+2d}\left(t-1|t-1\right)=\begin{bmatrix}\boldsymbol{m}\left(t-1|t-1\right)\\-\left\{\sqrt{\left(d+\lambda_{\mathsf{ukf}}\right)\Sigma_{\varepsilon}}\right\}_{i=1}^{d}\end{bmatrix}, \tag{5.57}$$

Step 3. Instantiate the points,

$$\chi_{a,i}\left(t|t-1\right)=\boldsymbol{f}\left(\chi_{a,i}(t-1|t-1)\right), \tag{5.58}$$

$$\Upsilon_{a,i}\left(t|t-1\right)=\boldsymbol{h}_{a}\left(\chi_{a,i}(t|t-1)\right). \tag{5.59}$$

Step 4. Compute the measurement mean

$$\boldsymbol{m}_{a,y}(t|t-1)=\sum_{i=0}^{2d}w_{a,m}^{(i)}\Upsilon_{a,i}(t|t-1), \tag{5.60}$$

and the covariance:

$$\boldsymbol{P}_{a,y}(t|t-1)=\sum_{i=0}^{2d}w_{a,c}^{(i)}\left(\Upsilon_{a,i}(t|t-1)-\boldsymbol{m}_{a,y}(t|t-1)\right)\left(\Upsilon_{a,i}(t|t-1)-\boldsymbol{m}_{a,y}(t|t-1)\right)', \tag{5.61}$$

where the weights $w_{a,m}^{(i)}$ and $w_{a,c}^{(i)}$ are the respective counterparts of weights (5.35)-(5.37).

5.3.3. Comparisons

Using the same simulation setting of Section 5.2, Figure 5.11 demonstrates the poorer performance of the non-augmented UKF compared with the augmented results of Section 5.2. With a low offset and an intermediate variance noise, $\sigma_\varepsilon^2 = 1$, the parameter estimates in the non-augmented UKF considerably diverge from the true values.

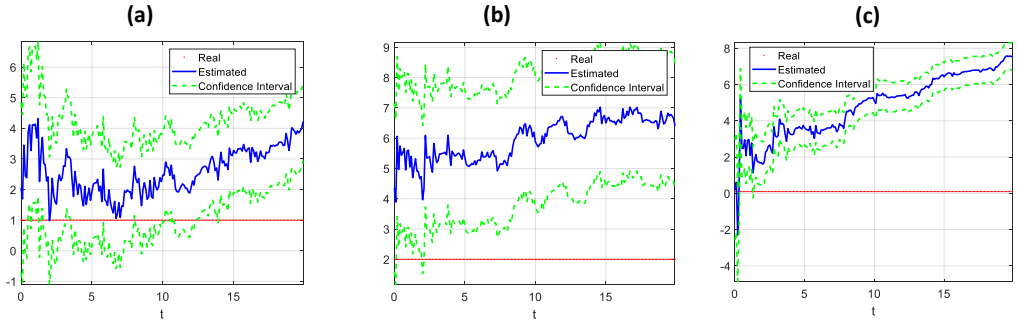

Figure 5.11: Parameter estimates of the non-augmented UKF, $\sigma_\varepsilon^2 = 1$ and 100% offset. (a) Estimates of α. (b) Estimates of β. (c) Estimates of c.

5.4. *Estimation of the state noise variance*

In this Section, with the same data generating process of Section 5.2, the process noise variance σ_ε^2 is estimated along with the SDE parameters. Therefore, the parameter vector is $\boldsymbol{\lambda} = (\alpha, \beta, c, \sigma_\varepsilon^2)'$ and the dimension of the system grows to $d = 6$. The observational noise is considered fixed $\eta \sim N(0, \sigma_\eta^2)$ to maintain a SNR = 10.

Inferring the variance noise is crucial. In real world applications, the latter is one of the generative parameter of a dynamical process and is unknown. Hence, investigating if the UKF is able to infer σ_ε^2 is important to move to the further step of the analysis of a real case scenario.

Figure 5.12 and 5.13 show the ARB and the RMS for the variance noise estimates. Since the ARB is the ratio between the difference among the estimates and the true value divided by σ_ε^2, the left panel of Figure 5.12 is affected by numerical instability. As far as the level of noise increases, the impact of the offsets in more definitive. In the case of low offset, the bias is the same for every level of noise, while higher

offsets have the effect of increasing the uncertainty, in the sense that the estimates on independent datasets do not result to a unique estimate. The RMS, instead, reflects the same behaviour for every level of noise. The bigger the offsets, the more spread the results around the median estimate.

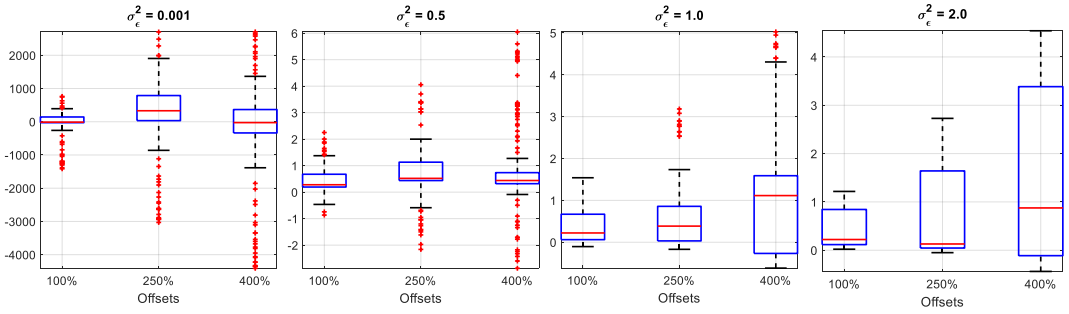

Figure 5.12: Average relative bias of the UKF estimation for several levels of the state noise variance and different offsets, with SNR = 10, n = 100.

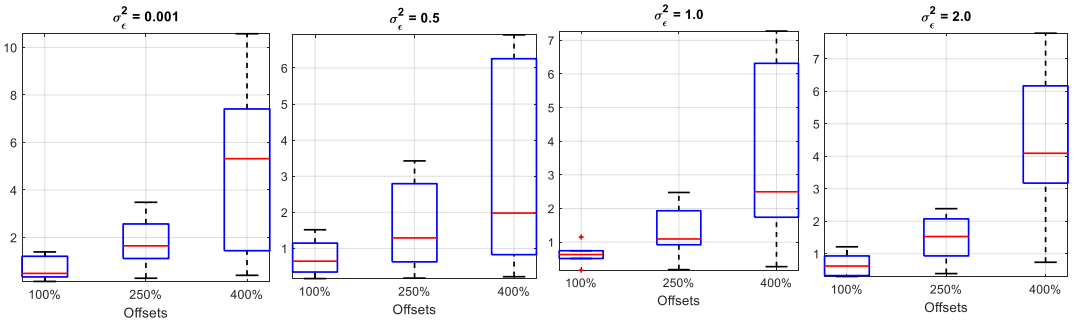

Figure 5.13: RMS of UKF estimates for several levels of the state noise variance and different offsets, with SNR = 10, n = 100.

The convergence of the estimates in the time domain is represented in Figures 5.14-5.16. Even with increasing level of variance, the UKF converges to the true values, but the uncertainty quantified by the confidence intervals increases. The UKF is able to infer the variance of the process noise, and this result will be used in the following real data applications.

The signal extraction is plotted in Figure 5.17. The Signal-UKF is more accurate than the Signal-SDE. This has a poor performance as a consequence of the chaoticity. Indeed, chaotic systems are highly sensitive to initial conditions and, even for a small

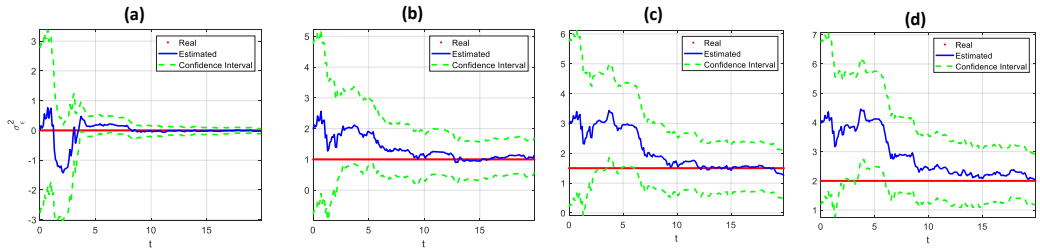

Figure 5.14: UKF estimates for the state noise variance for low offset. (a) $\sigma_\varepsilon^2 = 0.001$. (b) $\sigma_\varepsilon^2 = 1.0$. (c) $\sigma_\varepsilon^2 = 1.5$. (d) $\sigma_\varepsilon^2 = 2.0$.

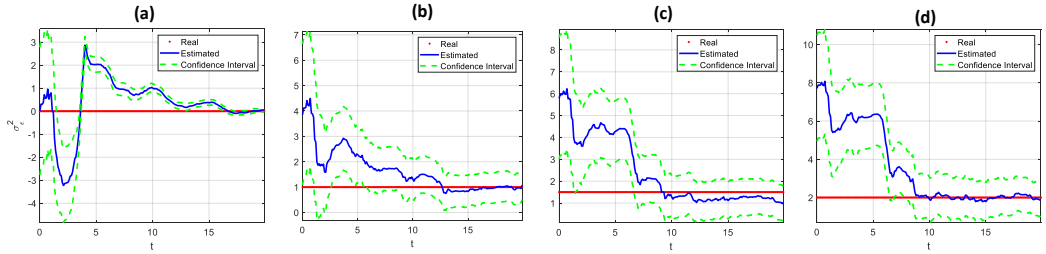

Figure 5.15: UKF estimates for the state noise variance for medium offset. (a) $\sigma_\varepsilon^2 = 0.001$. (b) $\sigma_\varepsilon^2 = 1.0$. (c) $\sigma_\varepsilon^2 = 1.5$. (d) $\sigma_\varepsilon^2 = 2.0$.

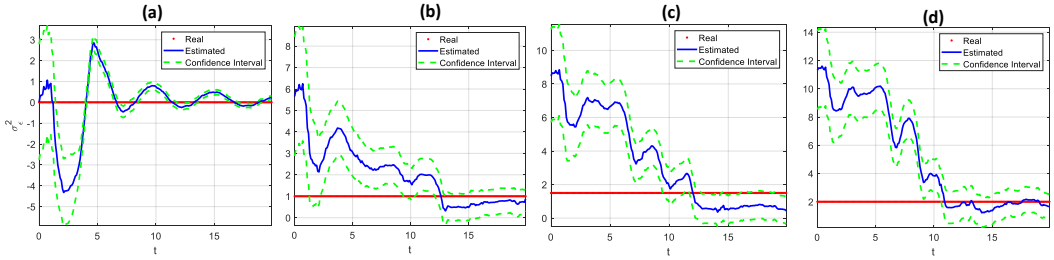

Figure 5.16: UKF estimates for the state noise variance for high offset. (a) $\sigma_\varepsilon^2 = 0.001$. (b) $\sigma_\varepsilon^2 = 1.0$. (c) $\sigma_\varepsilon^2 = 1.5$. (d) $\sigma_\varepsilon^2 = 2.0$.

mismatch of parameter values, the SDE undergoes different behaviours. If the final estimates $\hat{\boldsymbol{\lambda}}$ do not exactly match the true values, the motion reconstruction may differ from the measurements. In such a case, the filter rebuilding of the oscillatory data is more precise since when a new observation enters into the algorithm, it is able to reproduce, step-by-step, the whole underlying process.

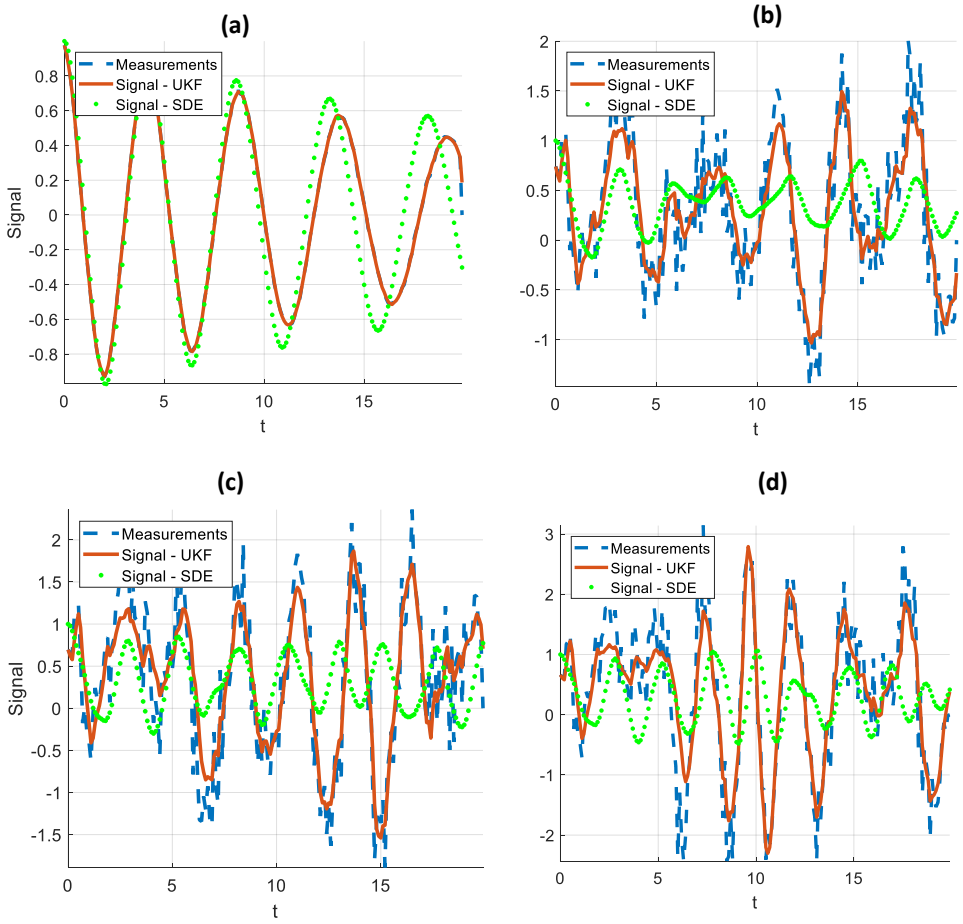

Figure 5.17: Comparison between the UKF filter and the SDE reconstruction. (a) $\sigma_\varepsilon^2 = 0.001$. (b) $\sigma_\varepsilon^2 = 1.0$. (c) $\sigma_\varepsilon^2 = 1.5$. (d) $\sigma_\varepsilon^2 = 2.0$.

5.5. *Comparison between known and unknown state noise variance*

The results of Section 5.2 and 5.4 are compared in Figures 5.18 - 5.22. In the case of 100% offset, the RMS errors between the two settings (known and unknown state noise variance) coincide. In the case of a "bad" initialization, the RMS for α, β and c is higher for the unknown variance framework than for the known process noise. Moreover, the distance between the two settings in terms of RMS increases with growing variance value. In particular, the divergence is higher in the param-

eter space than in the functional space. As already discussed (in Figure 5.17), the extraction of the hidden signal is fairly accurate during the filtering steps even with unknown process variance. Unknowing the variance affects more the UKF performance in terms of parameter estimates than in terms of oscillation reconstruction.

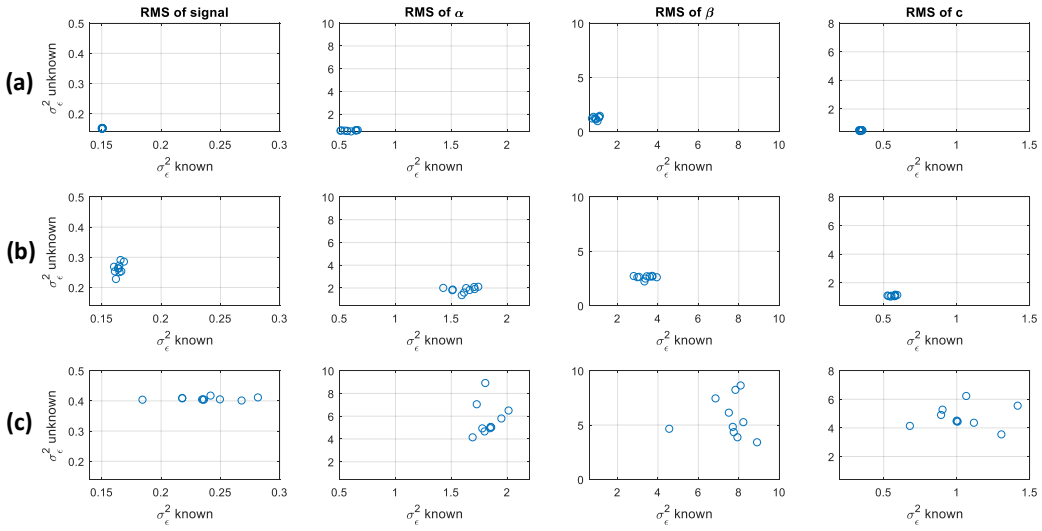

Figure 5.18: RMS of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$ and $\sigma^2_\varepsilon = 0.001$. (a) Low offset (100% of deviation from the true values). (b) Medium offset (250% of deviation from the true values). (c) High offset (400% of deviation from the true values).

5.6. *Comparison on an independent dataset*

This Section questions if the choice of the UKF as a method to infer the parameters for SDEs is adequate through a comparison on an independent dataset and a particle filter algorithm.

Consider a model for pharmacokinetics dynamics, that could be used to study the Theophylline drug pharmacokinetics. This model has a long history in the literature (e.g. Pinheiro and Bates, 1995 and Donnet and Samson, 2008), mostly in longitudinal data with mixed-effects models, but here the mixed-effects are not considered.

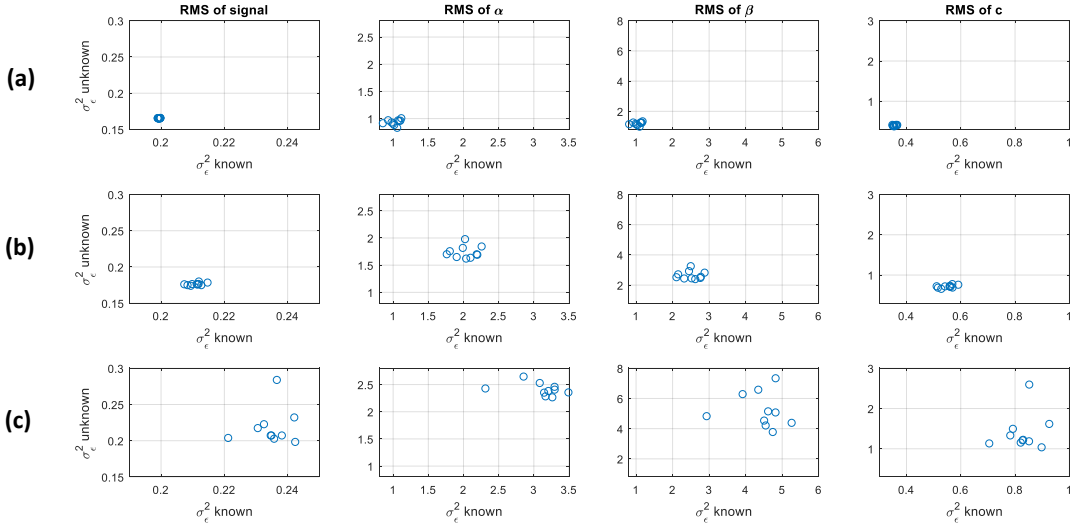

Figure 5.19: RMS of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$ and $\sigma_\varepsilon^2 = 0.5$. (a) Low offset (100% of deviation from the true values). (b) Medium offset (250% of deviation from the true values). (c) High offset (400% of deviation from the true values).

Instead, a recent article of Picchini and Samson (2017) is the matter of comparison and the same data generating process is used.

The SDE describing the dynamics of x_t, that is the level of drug concentration in blood at time t (hours), is

$$dx_t = \left(\frac{Dose \cdot K_a \cdot K_e}{Cl} e^{K_a t} - K_e x_t\right) dt + \sigma_\varepsilon \sqrt{x_t} dW_t, \tag{5.62}$$

for $t \geq t_0$, where *Dose* is the known drug oral dose received by a subject, K_e is the elimination rate constant, K_a is the absorption rate constant, Cl the clearance of the drug and σ_ε the intensity of the intrinsic stochastic noise.

For model (5.62), data are simulated at $n = 100$ equispaced sampling times where the time interval is $\Delta t = 1$: $\{t_1, t_2, \ldots, t_{100}\} = \{1, 2, \ldots, 100\}$. The drug oral dose is chosen to be 4 mg. After the administration of the drug, at $t_0 = 0$, the concentration first reaches $x_{t_0} = x_0 = 8$.

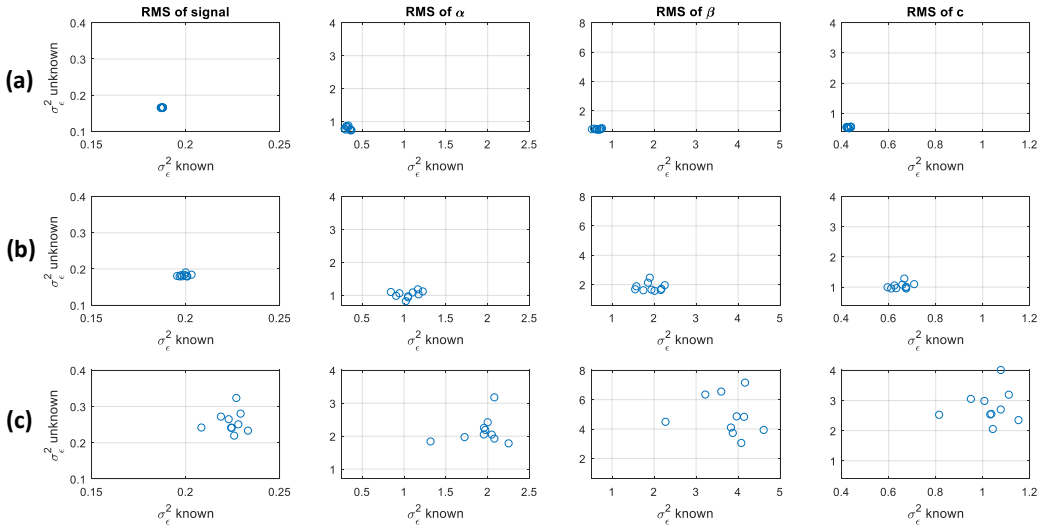

Figure 5.20: RMS of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$ and $\sigma_\varepsilon^2 = 1.0$. (a) Low offset (100% of deviation from the true values). (b) Medium offset (250% of deviation from the true values). (c) High offset (400% of deviation from the true values).

The measurement model is linear:

$$y_j = x_j + \eta_j, \tag{5.63}$$

where the $\eta_j \sim N(0, \sigma_\eta^2)$ are independent and identically distributed over the sample $j = 1, \ldots, n$. The parameter K_a is assumed known, and the parameters of interest are $\boldsymbol{\lambda} = \left(K_e, Cl, \sigma_\varepsilon^2, \sigma_\eta^2\right)$.

Equation (5.62) has no closed-form solution, and data are simulated using the Euler-Maruyama method with stepsize $\delta t = 0.05$ in the time interval $[t_0, 100]$. The Euler-Maruyama scheme is

$$x_{t+\delta t} = x_t + \left(\frac{Dose \cdot K_a \cdot K_e}{Cl} e^{K_a t} - K_e x_t\right) \delta t + \left(\sigma_\varepsilon \sqrt{\delta t \cdot x_t}\right) Z_{t+\delta t}, \tag{5.64}$$

where the $\{Z_t\}$ are identically distributed and independent Gaussian $N(0, 1)$. The generated values from equation (5.64) are linearly interpolated at sampling time $\{t_1, t_2, \ldots, t_{100}\}$ to give $\boldsymbol{x}_{1:n}$. Finally, accordingly to model (5.63), a residual error η is added. Since the measurement errors are independent, the observations

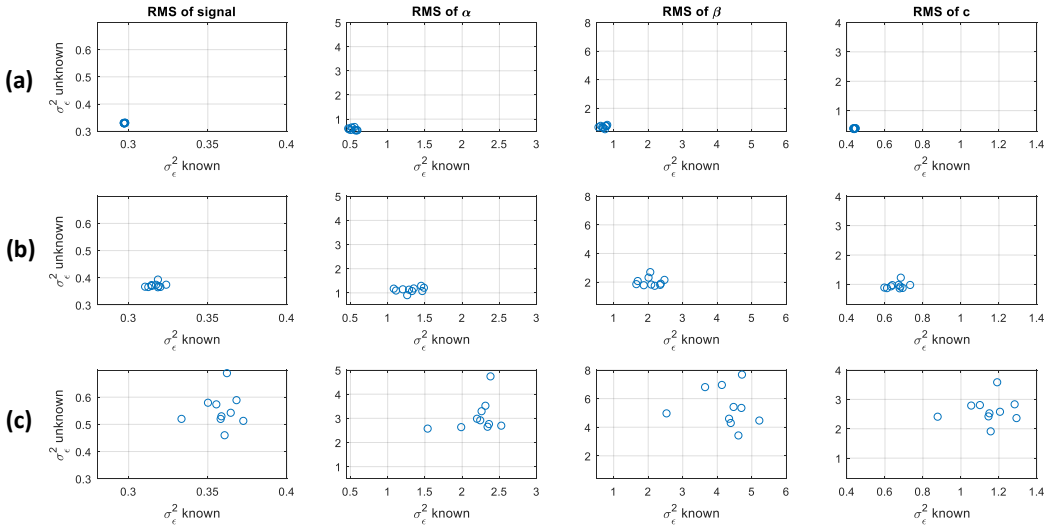

Figure 5.21: RMS of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$ and $\sigma_\varepsilon^2 = 1.5$. (a) Low offset (100% of deviation from the true values). (b) Medium offset (250% of deviation from the true values). (c) High offset (400% of deviation from the true values).

y_j are conditionally independent given the latent process x_t. The experiment is carried out with 50 independent datasets of length $n = 100$ and the true parameters $(K_e, K_a, Cl, \sigma\varepsilon, \sigma_\eta) = (0.05, 1.492, 0.04, 0.1, 0.1)$.

Picchini and Samson (2017) estimate model (5.62) with a new algorithm based on a stochastic approximation expectation-maximization (SAEM) coupled with the ABC scheme (SAEM-ABC). The authors look for a maximum likelihood estimation of parameters with the expectation-maximization (EM) method to compute the conditional expectation for the pair state-observation $[x, y]$. Since the process x_t is stochastic, the EM is used in a stochastic approximation extension. The problem consists in generating, conditionally on the current value of parameters $\boldsymbol{\lambda}$ during the maximization step, an appropriate "proposal" for the state x_t. The chosen path for x_t to feed SAEM with comes from the ABC-SMC inserted in the filtering phase. Since this method shares some connections with signal processing and the UKF, it is used to investigate the UKF performance with respect to another dataset. For the

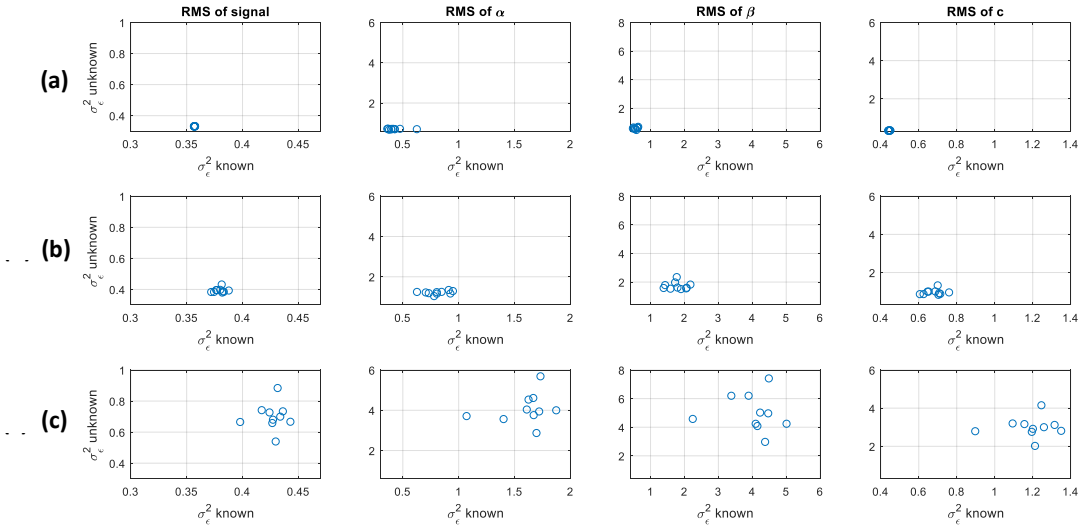

Figure 5.22: RMS of UKF estimates for the stochastic Duffing system with SNR = 10, $n = 100$ and $\sigma^2_\varepsilon = 2.0$. (a) Low offset (100% of deviation from the true values). (b) Medium offset (250% of deviation from the true values). (c) High offset (400% of deviation from the true values).

whole description of the SAEM-ABC, the reader is referred to Picchini and Samson (2017).

Comparisons for the parameter estimates are in Table 5.1. The SAEM-ABC re-

Parameters	True values	UKF	SAEM-ABC
K_e	0.05	0.058 [0.019, 0.097]	0.049
Cl	0.04	0.005 [0, 2.783]	0.045
$\sigma\varepsilon$	0.10	0.112 [0, 2.087]	0.710
σ_η	0.10	0.115 [0, 2.090]	0.030

Table 5.1: Comparison between the UKF and the SAEM-ABC algorithms for the pharmacokinetics model (5.62). In square brackets the confidence intervals truncated at zero, since the parameters can not reach negative values. The SAEM-ABC estimates come from Picchini and Samson (2017) and they do not give confidence intervals.

sults are closer to the true value for Cl, but the parameter K_e and the process and measurements variances are estimated with precision with the UKF. Even if the confidence intervals of parameters highlight uncertainty around the estimates, the UKF coupled with the Euler-Maruyama method for SDE challenges another filtering methods.

5.7. *Initialization of the UKF for SDEs*

In Chapter 3 the UKF limitation concerning the initialization has been extensively discussed: the method proposed in Section 3.8 tries to overcome the problem of the choice of the starting values for ODEs using the ABC-SMC scheme. The simulation study of Chapter 4, Section 4.4 shows that the ABC-UKF outperforms the UKF in the case of an initialization far from the true values.

Naturally, the ABC-SMC could be applied for SDEs as well. The only difference from the ODE case may concern the dispersion of the approximate posterior distributions of parameters. In other words, for SDEs, since a stochastic term perturbs the oscillation, one can expect that the posterior approximations are not as peaked as for the ODE system, that is these distributions could be flatter. Comparing the ABC-SMC results for ODEs and SDEs should be of interest, but it is not yet developed.

The comparison carried out in this Section does not focus on the evaluation of the ABC-SMC for SDEs and ODEs, but on the method to initialize the UKF, i.e. the aim consists in finding an alternative strategy to choose the starting values of the algorithm.

The method here proposed is based on successive reinsertions of the UKF parameter estimates in an iterative way. When the UKF is initialized with a high offset (in the preceding Sections, the deviation is fixed as 400%), the final parameter estimates do not converge to the true values (from Chapter 4, the reader knows that these estimates are far from the true parameters) but they can be reinserted in a successive UKF computation. The idea is that the final estimates become the new starting points for another UKF evaluation: the UKF can be computed many times, and each calculation starts with the final estimates of the preceding UKF. This strategy may be called “Iterative UKF”.

If $\hat{\boldsymbol{\lambda}}_j^{\mathrm{f}}$ denotes the final estimate of the parameter vector $\boldsymbol{\lambda}$ at the j-th UKF computation, for $j = 1, \ldots, J$, the initial values of the next $j + 1$ UKF are

$$\boldsymbol{\lambda}_j^{\mathrm{s}} = \hat{\boldsymbol{\lambda}}_j^{\mathrm{f}}, \tag{5.65}$$

where $\boldsymbol{\lambda}_j^{\mathrm{s}}$ is the starting vector.

Simulations have been carried out for $J = 50$. At the first iteration, $j = 1$, the location of sigma points is optimized with the BO scheme. For the successive iterations, $j = 2, \ldots, 50$, the optimization is not carried out because, as already shown in Chapter 4, it highly improves the UKF performance but does not assure the convergence to the true values. Furthermore, the BO drives consistently the computation costs if iterated for every $j = 2, \ldots, J$.

Preliminary results of the iterative UKF are shown in Figure 5.23.

The improvement is obtained up to the 13th iteration, but after the Iterative UKF suffers of numerical instability and reaches a steady solution which does not match the true values. The source of the numerical problems happen when inverting the Kalman gain matrix. Recall from Chapter 4 that I already modified the `EKF/UKF` toolbox of Hartikainen et al. (2011) used for the UKF computation with the addition of a jitter to invert the matrices. The numerical instability of the UKF is more pronounced for the SDE system than for the ODE (due to the random perturbations), and, after a certain point, the jitter the MATLAB functions need to add is too high and the method crashes.

This alternative strategy to initialize the UKF for SDEs has to be extended but it may offer new research topics to deal with the choice of the starting values.

5.8. *Discussion on SDEs*

The UKF performance in the context of SDEs has been analysed. To manage the numerical approximation of the transition function, the Euler-Maruyama method has been inserted into the UKF steps. The Euler-Maruyama schemes within the UKF has been shown to give accurate estimates both in the signal reconstruction and parameter estimates. Several levels of process noise variance has been evaluated

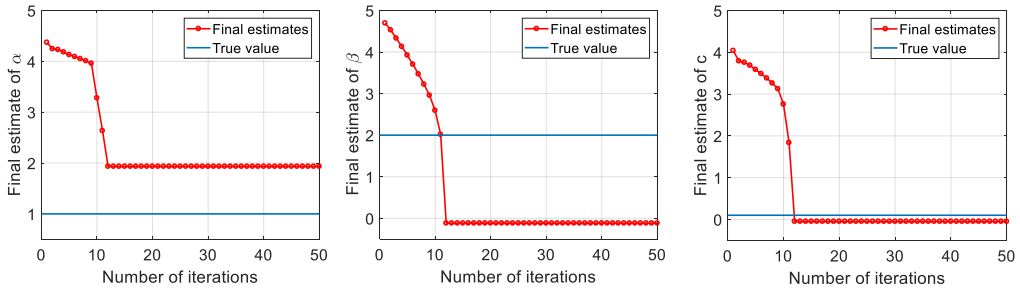

Figure 5.23: Preliminary simulations of the Iterative UKF. The blue lines are the true values and the red dotted lines represent the final estimates of each UKF computation.

and the UKF results to be a powerful method for inference for the stochastic Duffing system. Moreover, the proposed inference scheme is able to estimate the process noise variance; this is a crucial step to move further to real data analysis. The comparison between the settings of known and unknown state noise variance shows that the RMS and the bias are higher and grows faster for the case of unknown variance.

The mathematical difference and comparisons between augmented and non-augmented UKF have been described, concluding that the augmented version is more accurate than the non-augmented UKF. To avoid the risk related to a double drawing of sigma points, the augmented version should be preferred.

Finally, in the comparison on an independent dataset, the Euler-Maruyama method coupled with the UKF demonstrates that it may challenge other inference schemes.

6. Changepoint detection method

In the preceding Chapters, inference methods for the Duffing system have been discussed through simulation studies. Before moving to real data applications of the inference schemes presented so far, it should be worth to consider that the oscillatory behaviour of a time series may vary over time, e.g. the frequency or the amplitude can differ from one period of time to another. The change in the nature of oscillation can be attributed to an alteration of the generative parameters of the process. This shift can mean that two or more hidden regimes may exist behind the observed data, and each regime is characterized by different parameter values. To investigate when the hidden process switches to one regime to another, there is the need of a method that locates the breakpoints in the generative parameter values in the time domain.

The changepoint detection strategy developed is *heuristic* in the sense that it is model-free and based on wavelets analysis. Wavelets are a tool to decompose time series and represent a signal through a finite-length wave. The approach discussed below consists in finding the breakpoints in the parameters of the process when there is a change in the variance of the wavelets.

This Chapter starts with an introduction to wavelet transforms and the main difference between wavelets and the Fourier analysis (in Section 6.1), and Section 6.2 mathematically defines the wavelets transform. Characteristics and difficulties of wavelets are discussed in Section 6.3 while some classes of wavelets are mentioned in Section 6.4. The changepoint detection strategy used in real applications is presented in Section 6.5.

6.1. *Wavelets and Fourier series*

Wavelets share some similarities with the Fourier series, but there are also important points of difference. The idea behind both wavelets and the Fourier transform consists by projecting a signal into a basis space. They differ in the bases used.

The Fourier representation projects trigonometric functions into the signal, assuming stationarity over the whole time series. The Fourier analysis is defined in

the $L^2(0, 2\pi)$ space and the decomposition of the oscillation is achieved in the frequency domain. All the frequencies of data are obtained by projecting the signal into a sequence of series like

$$\left\{e^{-\iota s\omega_0}\right\}, \tag{6.1}$$

where ι is the imaginary number, ω_0 is the fundamental frequency and s is the scaling. There are many excellent expositions in the literature of the Fourier analysis, e.g. Percival and Walden (2006), to which the interested reader is recommended.

Wavelets are functions defined over Besov spaces, that is a complete quasinormed space, and constitute bases for functions defined in such spaces. Each basis function $h(\cdot)$ can be expressed as

$$h(t) = \frac{1}{\sqrt{s}} h\left(\frac{t-k}{s}\right), \quad t = 1, \ldots, T, \tag{6.2}$$

where $k \in \mathbb{Z}$ is the time domain index and $s \in \mathbb{N}$ is the scale at which $h(\cdot)$ is evaluated. Equation (6.2) means that a sequence of functions is constructed in the time domain: the functions $h(\cdot)$ are doubly indexed by (i) location in time k and (ii) the scale. In other words, each function is centred at k with a dilation of s. Therefore, the main difference between the bases (6.1) and (6.2) concerns the domain: the former, i.e. the Fourier transform, is in terms of frequency, while the latter, a wavelet basis, is scaled in the time domain.

6.2. *Wavelets definition*

Wavelets can be defined as a sequence of filters, or as splines satisfying certain properties: in what follows, both definitions are provided. However, only discrete time wavelets are considered.

6.2.1. Definition based on splines

In terms of splines, *father* and *mother* wavelets should first be defined. Father wavelets are used to represent the very long scale smooth component of the signal and integrate to one. They generate the scaling coefficients and act as a low pass filter. Instead, mother wavelets represent deviations from the smooth components

and integrate to zero. The latter generate differencing coefficients and act as a high pass filter.

Given a function $\Phi(\cdot)$, and restricting the scale parameter s in equation (6.2) to the dyadic scale 2^j, for $j = 1, \ldots, J \in \mathbb{N}$, the corresponding father wavelet $\Phi_{J,k}$ is

$$\Phi_{J,k} = 2^{-\frac{J}{2}} \Phi \left(\frac{t - 2^J k}{2^J} \right), \tag{6.3}$$

$$\int \Phi(t) dt = 1. \tag{6.4}$$

The respective mother wavelet, $\Psi_{j,k}$, is

$$\Psi_{j,k} = 2^{-\frac{j}{2}} \Psi \left(\frac{t - 2^j k}{2^j} \right), \tag{6.5}$$

$$\int \Psi(t) dt = 0. \tag{6.6}$$

Given the basis functions of equations (6.3) and (6.5), a sequence of coefficients that represent the projections of the signal into the basis can be defined. The coefficients for the father wavelet at 2^J, i.e. the maximal scale, are called "smooth coefficients" and are

$$s_{J,k} = \int f(t) \Phi_{J,k} dt. \tag{6.7}$$

The detail coefficients obtained from the mother wavelet are evaluated at all scales $j = 1, \ldots, J$ and determined as

$$d_{j,k} = \int f(t) \Psi_{j,k} dt. \tag{6.8}$$

The function $f(\cdot)$ in equations (6.7)-(6.8) is a wavelet basis if it is an *orthonormal* basis for $L^2(\mathbb{Z})$ of the form

$$f(t) = \sum_{k=0}^{K} s_{J,k} \Phi_{J,k}(t) + \sum_{k=0}^{K} d_{J,k} \Psi_{J,k}(t) + \ldots + \sum_{k=0}^{K} d_{j,k} \Psi_{j,k}(t) + \ldots + \sum_{k=0}^{K} d_{1,k} \Psi_{1,k}(t), \tag{6.9}$$

where K is an even integer representing the number of *vanishing moments*. The latter indicates the degree of the polynomial generated by the scaling function. A K-th vanishing moment of a wavelet points out that a polynomial up to degree $K - 1$ is passed through the mother wavelets. When the wavelet has K vanishing moments,

the wavelet transform can be interpreted as a multiscale differential operator of order K. This yields a relation between the differentiability of f and its wavelet transform decay at fine scales (Ramsey, 2002). Indeed, this is equivalent to saying that the first K derivatives of the Fourier transform of the wavelet filter all are zero when evaluated at 0. The projection from the mother wavelet integrates to zero and the polynomial component is captured by the father wavelet, i.e. the scaling function alone can be used to represent functions. In other words, if the signal contains a polynomial component, the appropriate number of vanishing moments decomposes the time series giving insight of the characteristics of data.

In an alternative way, equation (6.9) can be re-written as

$$f(t) = S_J + D_J + D_{J-1} + \ldots + D_j + \ldots + D_1, \tag{6.10}$$

where

$$S_J = \sum_{k=0}^{K} s_{J,k}\Phi_{J,k}(t), \tag{6.11}$$

$$D_j = \sum_{k=0}^{K} d_{j,k}\Psi_{j,k}(t),\ j = 1, \ldots, J. \tag{6.12}$$

To easily visualize the above description, let us imagine a sequence of topographical maps: S_J provides a smooth outline and higher levels of detail are given by each D_j. The multiresolution analysis (MRA) of the signal is highlighted by the complete derivation of function $f(\cdot)$ in equation (6.10). Naturally, one can obtain less detailed representations of the signal examining $S_j = S_J + D_J + \ldots + D_{j+1}$ or only $S_j = S_J + D_{j+1}$.

The maximum number of coefficients D_j is the nearest integer less than $\log_2(n)$, where n is the sample size.

6.2.2. Definition based on filters

Wavelets can also be defined in terms of low and high pass filters. In this case, the functions $\Phi(t)$ and $\Psi(t)$ are defined, respectively, as

$$\Phi(t) = \sqrt{2} \sum_{k=0}^{K} l(k)\Phi(2t - k), \tag{6.13}$$

$$\Psi(t) = \sqrt{2} \sum_{k=0}^{K} L(k)\Phi(2t - k), \tag{6.14}$$

where $l(k)$ is a linear lowpass filter and $L(k)$ is a linear highpass filter. Analogously, the low and high pass filters can be derived from the father and mother wavelet like:

$$l(k) = \frac{1}{\sqrt{2}} \int \Phi(t)\Phi(2t - k)dt, \tag{6.15}$$

$$L(k) = \frac{1}{\sqrt{2}} \int \Psi(t)\Phi(2t - k)dt = (-1)^k l(k). \tag{6.16}$$

Following the approach of wavelets definition in terms of filters, it can be seen that the low pass filter averages, the high pass filter differences (Ramsey, 2002). In the signal processing research area, the filters l and L constitute *filter banks*, i.e. an array of band-pass filters that separates the signal into multiple components, where each one carries out a single frequency sub-band of the original vibration. The relationship between wavelets and filter banks is developed by many authors, as Strang and Nguyen (1996) or Percival and Walden (2006) and references therein. Several classes of wavelets are created by specifying particular properties for the filter banks.

6.3. *Characteristics and difficulties of wavelets*

Wavelets may be characterized by the symmetry or the smoothness of the basis functions. The choice of the kind of wavelet to use depends on the weights that one places on various criteria. Hence, a researcher may choose the class of wavelets most suitable to represent some properties of the function.

Symmetry of the function $h(\cdot)$ is one criterion and it is seldom satisfied (Ramsey, 2002). Symmetry of wavelets is useful to represent signals that exhibit local symmetries: the Haar wavelet, discussed below, is an example of this.

Another important property is the *smoothness* of a wavelet basis. The degree of smoothness is represented by the number of continuous derivatives of the basis function.

However, the introduction of filter banks reveals the difficulty of dealing with boundary conditions (Percival and Walden, 2006). To solve this problem several approaches have been proposed in the literature. One solution consists by taking advantage of any periodicity of the data and use polynomials with some regularities that can capture the period at the end of the observed series. Another solution involves the extension of data by means of a reflection and the assumption of periodic boundaries.

6.4. *Family wavelets*

The wavelets transform described in Section 6.2 is known as discrete wavelet transform (DWT). Some classes of DWTs are presented in what follows.

6.4.1. Haar wavelets

The Haar wavelet is the simplest possible wavelet, it is smooth and often used in the representation of Poisson processes. The technical disadvantage of the Haar wavelet is that it is not continuous, and therefore not differentiable. The mother function of the Haar wavelet is described by

$$\Psi(t) = \begin{cases} 1; & 0 \leq t < \frac{1}{2} \\ -1; & \frac{1}{2} \leq t < 1 \\ 0; & \text{otherwise.} \end{cases} \tag{6.17}$$

The Haar wavelet has several properties:

- Any continuous real function with compact support can be approximated uniformly by linear combinations of father wavelets $\Phi(t), \Phi(2t), \Phi(4t), \ldots, \Phi(2^n t)$.
- Any continuous real function on $[0, 1]$ can be approximated uniformly on $[0, 1]$ by linear combinations of the constant function $1, \Psi(t), \Psi(2t), \Psi(4t), \ldots, \Psi(2^n t)$.

- Orthogonality of the form

$$\frac{1}{\left(\sqrt{2}\right)^{(n+n_1)}} \int \Psi\left(2^n t - k\right) \Psi\left(2^{n_1} t - k_1\right) dt = \delta_{n,n_1} \delta_{k,k_1}, \tag{6.18}$$

  where $\delta_{i,j}$ represents the Kronecker delta.

Since the Haar wavelets has length of filters equals to two, the equations (6.15) and (6.16) are

$$l(k) = \left\{\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\}, \tag{6.19}$$

$$L(k) = \left\{\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right\}. \tag{6.20}$$

6.4.2. Daubechies wavelets

The Daubechies (DB) wavelets generalize the Haar transform and constitute a family of orthogonal wavelets characterized by a maximal number of vanishing moments for some given support. In this class of wavelets, the father wavelet generates an orthogonal MRA.

The DB wavelets is constructed in order to satisfy some special properties, as

- Orthogonality.
- Each wavelet is compactly supported.
- The regularity of wavelets increases linearly with the support width.

However, since the DB wavelets are not defined in terms of scaling function, a closed form expression for these wavelets does not exist.

These wavelets are neither symmetric nor anti-symmetric around any axis. Indeed, satisfying symmetry conditions cannot go together with all other properties of the Daubechies wavelets. To overcome this limit, the Symlet wavelets have been proposed as a modified version of the DB with increased symmetry.

A part from DWTs, there are other generalizations that one could examine. The maximum overlap discrete wavelet transform (MODWT) gains a resolution of the

signal by losing the property of orthogonality. Furthermore, The MODWT is translation invariant and the transform can be applied to datasets whose length is not divisible by 2^J (Ramsey, 2002). In the computation of wavelets of the next Chapter, the MODWT is used.

6.5. *Changepoint detection strategy*

The strategy to detect the changepoints in the generative parameters of a process is the following. The transforms belonging to the class of wavelets presented in Section 6.4 are computed and a MRA is performed. The choice of the wavelets to use is made on the basis of the maximum distance between the MRA and the data. The Shannon entropy of the MRA is also evaluated, and the quantification of how critical is this value will be discussed in the next Chapter depending on the dataset. Once the type of wavelet transform is decided, the number of coefficients of the mother wavelets is fixed by calculating the contribution of each coefficient on the total variance of data. Finally, the changepoints in the variance of each wavelet coefficient gives an insight on where the breakpoints in the generative parameter of the process happen.

Below, the pseudo-code of this heuristic strategy is presented.

Step 1. Goal: choose the kind of wavelet and the number of vanishing moments.

1. Compute several MODWTs;
2. Carry out a MRA for each wavelet;
3. Compare wavelets by evaluating the maximum distance between the MRA reconstruction and the original data, keeping, at the same time, the smallest value of the Shannon entropy.

Step 2. Goal: define the level of detail of the wavelet, i.e. the number of coefficients of the mother wavelet.

1. Calculate the contribution of each level D_j into the total variance of the data;

2. Eliminate the levels with the smallest contribution.

Step 3. Goal: find the changepoints in the variance of the wavelet transform.

1. Compute the variance changepoint in each wavelet level.

The computation is performed using the MATLAB Wavelet Toolbox, based on the fast wavelet transform (FWT).

7. Real data analysis

This Chapter presents results of the estimation methods discussed in the preceding Chapters for real world datasets. First, wavelets analysis to detect changing points of the parameters is performed, according with the heuristic approach of Chapter 6. Then, the time series of the U.S. business cycle and the sunspot numbers are modelled by the Duffing equation and the inference is carried out with the UKF for stochastic systems, as discussed in Chapter 5.

7.1. *The U.S. business cycle*

This application focuses on the quarterly growth rate of real gross domestic product (GDP) for the U.S., available for the period 1957:2 – 2017:1 (the notation 1957:2 indicates the second quarter of the year 1957). The series and its autocovariance function are plotted in Figure 7.1.

The left panel of Figure 7.1 shows two main oscillatory behaviours. The first one starts from 1957 to the early 1980s, where the GDP growth rates are characterized by sharp volatility. The dynamicity of the business cycle is described by high frequencies associated to sharp amplitudes. Instead, the second behaviour in the oscillation spans from the second half of the first decade of the second millennium to nowadays. In this period, small frequencies match a reduction of the amplitudes.

Such features may lead to think about two regimes, i.e. hidden states, behind the data: one described by high parameter values, and the other associated to lower values.

The period from the second half of the 1980s to the late 2000s can be considered as a transition period between the time of high volatility to the years of low growth of the business cycle. The transition phase highlights a gradual decline in the volatility of the business cycle fluctuations, and in the literature it is sometimes known as the "great moderation" (Mojon, 2007).

Observing the business cycle quarterly percentage change, a researcher could imagine that at least three changepoints (or regimes) in the parameter values exist.

The first can be economically explained by the "economic boom", the second may correspond to the great moderation, and, finally the contemporary period could be interpreted according to a new economic paradigm for a system with low volatility.

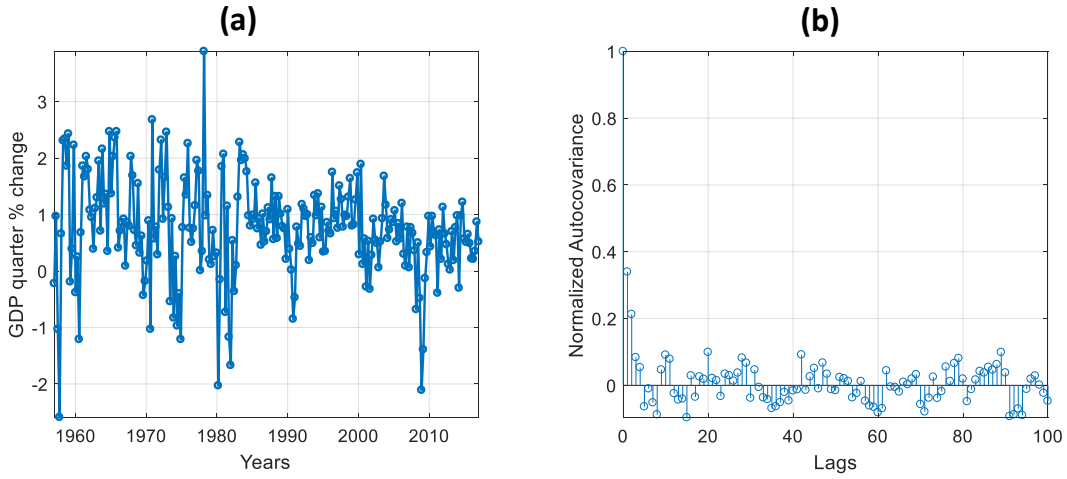

Figure 7.1: (a) U.S. GDP quarter over quarter percentage change from 1957:2 to 2017:1. Data source: Thompson Reuters Datastream. (b) Autocovariance function of the U.S. GDP time series.

Before moving to wavelet transforms, a descriptive analysis of the changes in the data is performed through the computation of the RMS and standard deviation (SD) statistics. These statistics, using Gaussian log-likelihood, are search criteria to minimize the log weighted dispersion of data according to the maximum number of changepoints allowed. Figure 7.2 depicts the dispersion of the observations versus the number of breakpoints. In other words, if the number of changing points is allowed to be no more than five, the dispersion of data is the smallest by fixing breaks in the years indicated in the boxes of Figure 7.2. When four or five changing points are admitted, the RMS statistic detects the years 1979, 1980, 1984 and 2009, while the SD indicates 1983, 1984, 2008 and 2009.

Even if the dispersion reaches its minimum, the first three years shown by the RMS, and the first and last couple of years of the SD are too close together to consider all of them as breakpoints. Rather, these points may point out that a change, that is

a shift from one hidden regime to another, happens in terms of variability around the half of the 1980s and in 2008–2009. The difference in the volatility in the years 1979–1984 or 2008–2009 can be ascribed to an adjustment period in the economic system following the switching of regime.

The 1984:2 is the year detected in each number of breakpoints allowed, both by the RMS and the SD. In Figure 7.3 the GDP quarterly growth rate is plotted with changing points corresponding to 1984:2 and 2009:1.

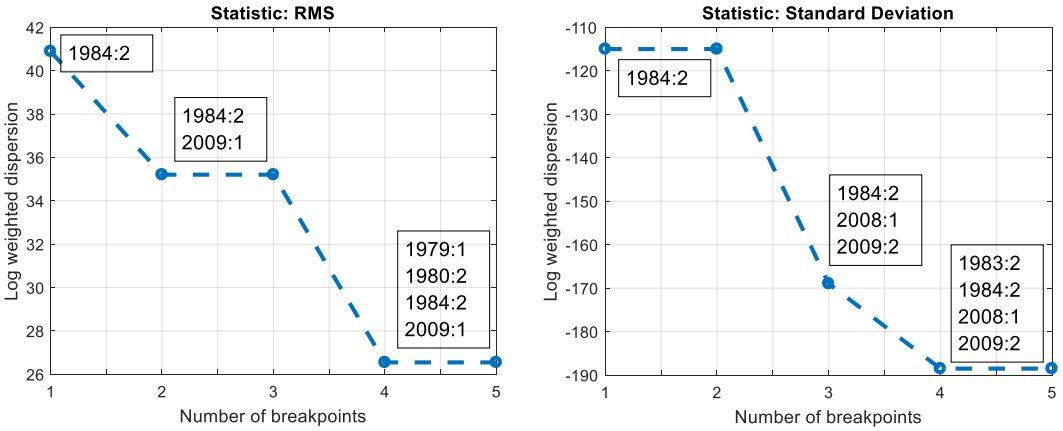

Figure 7.2: RMS and SD statistics to detect changepoints. The years in the boxes indicate the changepoint detected.

7.1.1. Changepoints of wavelets variance

The wavelet transforms discussed in Chapter 6 (the Haar, the DB and the Symlets) are compared on the U.S. GDP data to choose the number of vanishing moments. Figure 7.4 gives a graphical representation of the comparison.

In terms of maximum distance between the MRA and the original observations, the Symlets and the DB transforms share the same behaviour until the 4th vanishing moment. The Haar, instead, is a straight line: this transform is smooth and thus the MRA based on the Haar is not reliable. Let us recall from Chapter 6 that the Haar wavelets correspond to the DB with one vanishing moment, while the Symlets is a version of the DB with increasing symmetry. Hence, when a MRA is carried out, at

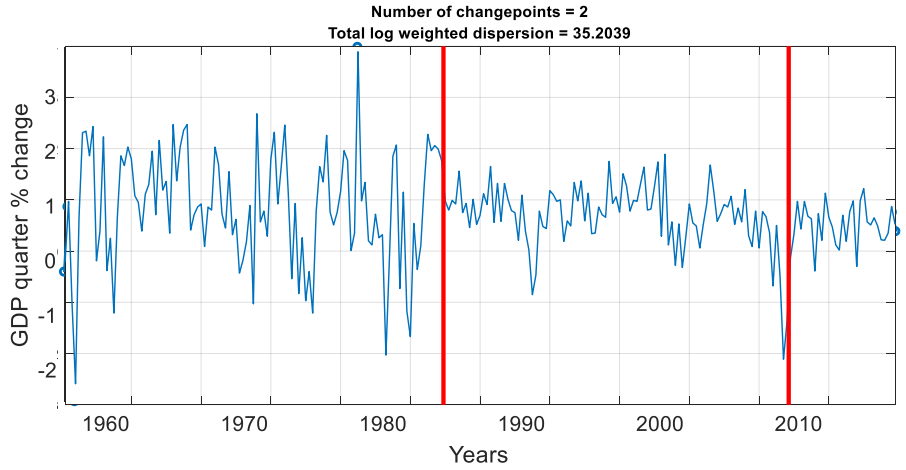

Figure 7.3: Changepoints of the parameter values in the data detected by the RMS and SD statistics of Figure 7.2. The blue line is the U.S. GDP quarterly growth rate, the red lines depict the changepoints corresponding to 1984:2 and 2009:1.

the first vanishing moments, the Haar, the DB and the Symlets are equivalent (that is the starting point in panel (b) of Figure 7.4 is the same).

With respect to the entropy curves for increasing number of vanishing moments (left panel of Figure 7.4), the Haar wavelets are characterized by greater entropy with increasing vanishing moments. On the contrary, the entropy growth is less pronounced for the DB and the Symlets transforms.

Considering that (i) the Haar wavelet shows a strong growing in the entropy level and (ii) the Symlets is based on the DB transform with more symmetry, but the latter is not a requirement for the business cycle data (the observations are not symmetric), the DB wavelet with two vanishing moments may be the most appropriate choice. This choice also finds support in a similar analysis discussed in Gallegati and Semmler (2014).

As previously stated in Chapter 6, the maximum number of wavelet coefficients depends on the sample size and is the nearest integer less than $\log_2(n)$, where n is the sample size. In this case, with the U.S. business cycle dataset, the maximum number of wavelet coefficients, the "levels" D_j in equation (6.12), is 7. The contribution of

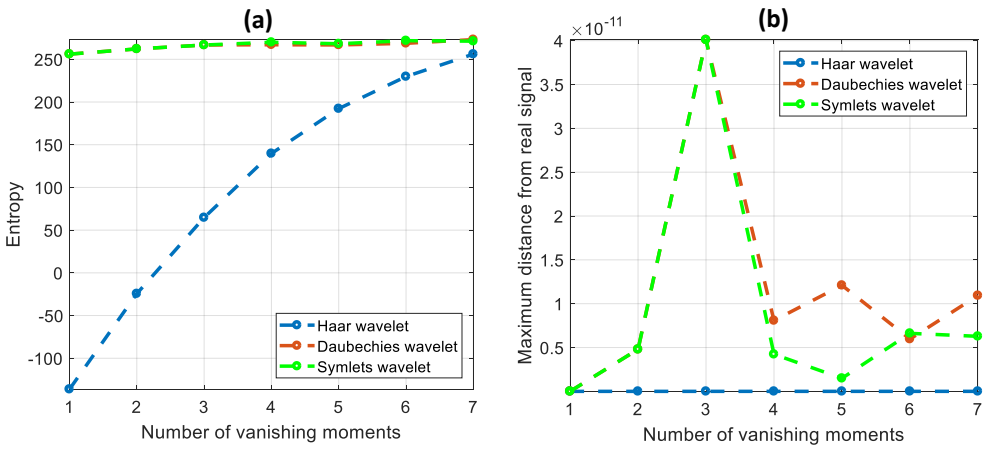

Figure 7.4: Comparison between wavelets to choose the number of vanishing moments. (a) Comparison in terms of entropy. (b) Comparison in terms of maximum distance from the MRA and the observations.

the 7 coefficients in the total variance is shown in Figure 7.5. Since these data are quarterly, the first scale captures variations between the 1st and the 3rd quarters, the second scale between the 3rd and the 6th quarters, the third scale between the 8th and the 16th quarters and so on. The bar plot highlights that cycles between the 1st and the 12th quarters account for the largest variability in the GDP data.

However, since the sixth and the seventh wavelet coefficients give a small contribution to the overall decomposition of the series, they are eliminated. The breakpoints in the variance of the wavelets are evaluated on the DB with two existing moments (DB2) and 5 coefficients.

The details, or the levels, of the DB2 wavelets are plotted in Figures 7.6 - 7.7. The black straight lines indicate the variance changepoints that happen in the year written in the boxes. The first three levels of decomposition of the DB wavelets, which are the ones that most account for the all variability of data, put in evidence that breakpoints are located between 1983-1986 and, from these years, a reduction in the variance of the oscillation appears. The last two levels, that contribute less to the total explanation of the variance, have changing points in the 1990 (Level 4) and 1982 and 2002 (Level 5).

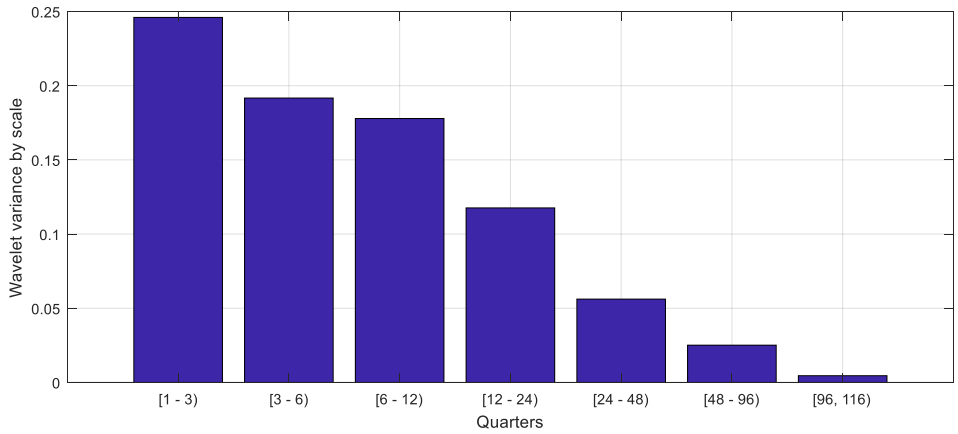

Figure 7.5: Contribution to the total variability of the data of the DB wavelet with two vanishing moments and 7 coefficients (the maximum number allowed) of the mother wavelets.

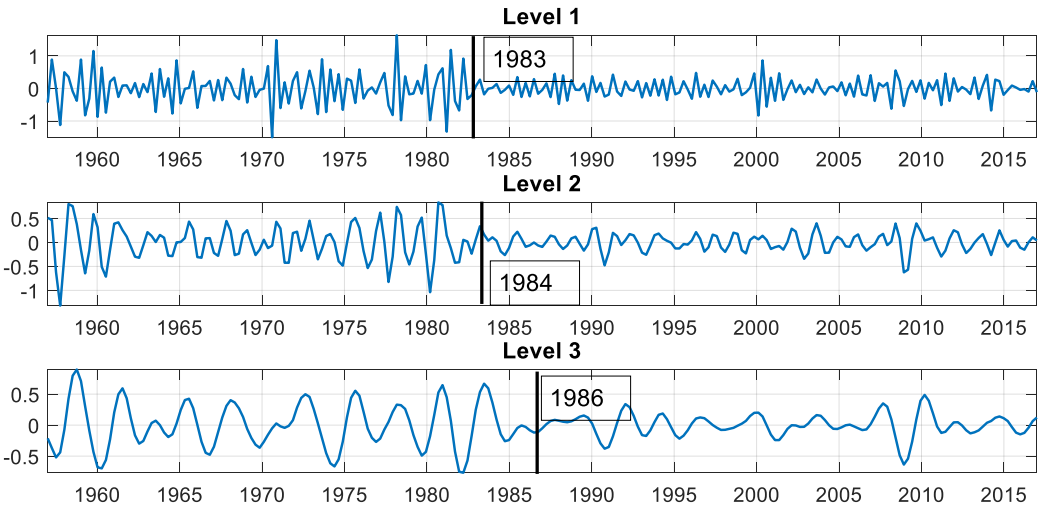

Figure 7.6: First three level of details of the DB2 with five coefficients. The black straight lines indicate the changepoints in the variance of the wavelets that happen in the years written in the boxes. Notice that these are the coefficients that account for the most variability of the data.

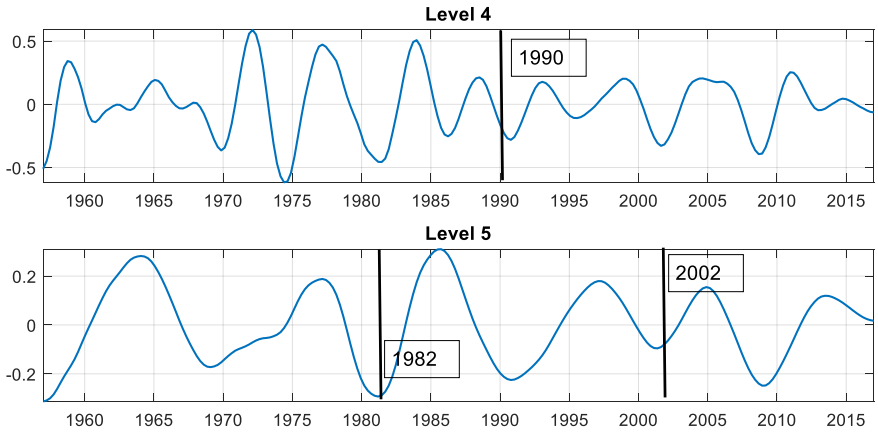

Figure 7.7: Last two level of details of the DB2 with five coefficients. The black straight lines indicate the changepoints in the variance of the wavelets that happen in the years written in the boxes. Notice that these are the coefficients that do not show a strong contribution in the decomposition of the variability of the data.

7.1.2. Estimates for the business cycle

Data are modelled according to equation (5.23) of Chapter 5. The estimation with the UKF is carried out for three periods of time, that is, following the changepoint detection strategy in Section 7.1.1, three different regimes are considered. The first spans in the period 1957-1984, the second considers 1985-2008, and the third is 2009-2017. The choice of the first breakpoint, the 1984, is made on the basis of the wavelet transforms heuristic method. Furthermore, in the economic literature, some authors identify the 1984 as the starting year of the great moderation (Stock and Watson, 2002, Kim and Nelson, 1999, Blanchard and Simon, 2001).

The parameter estimates are shown in Table 7.1, while Figure 7.8 depicts the signal reconstruction of the UKF.

The Duffing process is able to reproduce the volatility of the time series and the parameter estimates give an insight on the nature of the business cycle fluctuations. The chaoticity of the system, represented by β, abruptly shrinks over the regimes, meaning that the underlying complexity of the system is completely changing characteristic. The parameter α describes the frequency of oscillation and highlights the

decreasing in the volatility from the economic boom to the great moderation period. However, in the third regime, the one from 2009 to nowadays, the UKF estimates a higher value of α with respect of the preceding periods. This growth of α can be interpreted along with the value of c. Chapter 2 has shown that the kind of stability of the trivial fixed point depends on the relation between α and c. In particular, when c is bigger than α, the focus is unstable, while for $c \leq \alpha$ a stable solution arises. This interpretation is supported by the evidence that from 2010 the U.S. business cycle seems to decay toward a system with small amplitudes since it oscillates among values in the range $[-0.5, 1]$.

Years	α	β	c
1957-1984	29.45	87.25	38.46
1985-2008	19.94	21.38	22.68
2009-2017	40.53	3.89	10.45

Table 7.1: Parameter estimates for the U.S. GDP modelled by the Duffing process and estimated with the UKF method.

7.2. *Sunspots data*

Sunspots are areas of cooler zones protected by magnetic fields on the surface of the Sun. Near the sunspot, hotter areas of the Sun react with the magnetic field outside the sunspot and create a solar flare which project x-rays or energy particles toward the Earth's atmosphere in the form of a geomagnetic storm. Since it is still debated if and how the sunspots affects the climate on the Earth, scientists and solar observers are collecting an enormous amount of data on solar cycles to predict the number of sunspots per cycle.

The analysis presented in this Section aims to model the well known time series of the yearly sunspot numbers for the period 1700 - 1988 whose source is Tong (1990). The sunspot cycle has been analysed in many different textbook on time series analysis which can provide testbeds for the present study. However, a con-

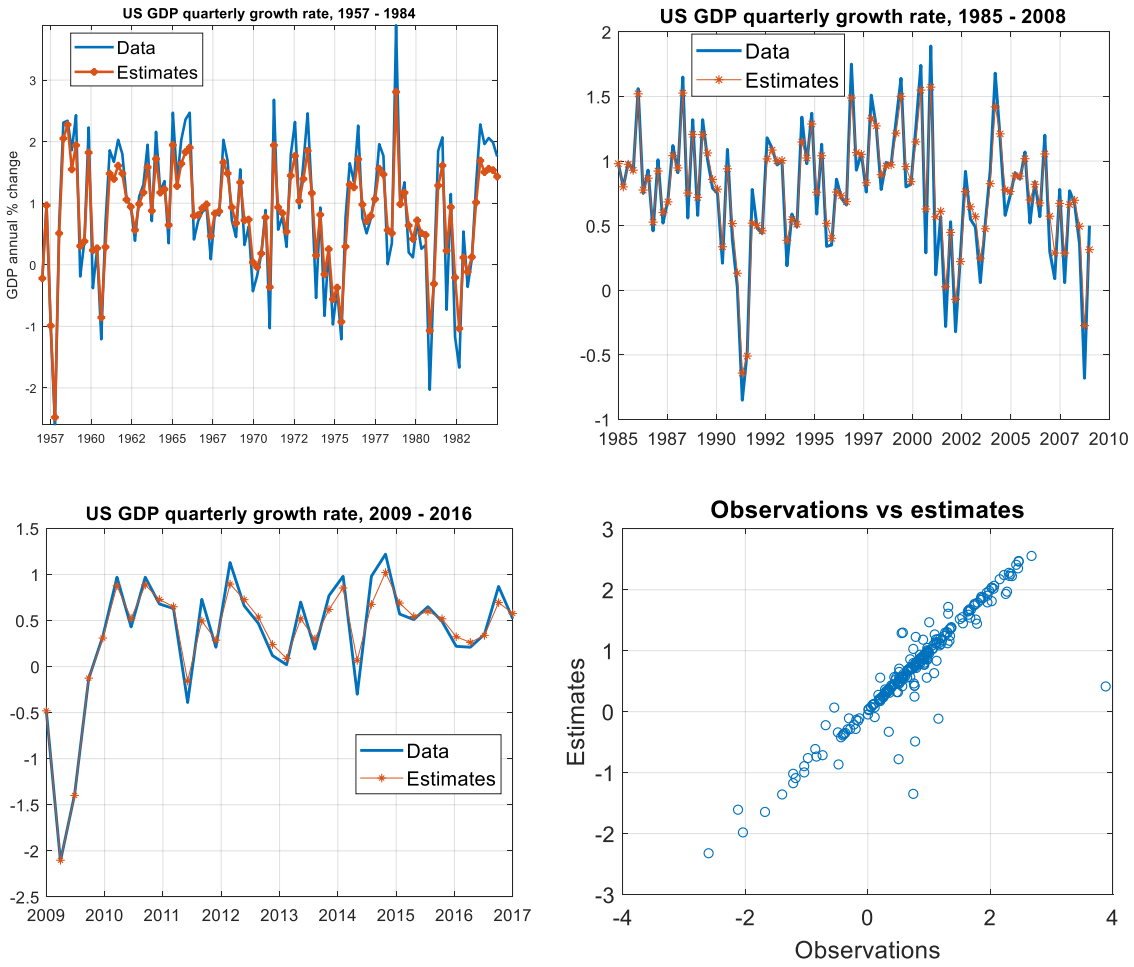

Figure 7.8: Signal reconstruction of the UKF for the U.S. business cycle modelled by the Duffing process. The RSS for the 1st, 2nd and 3rd regime is, respectively, 5.41, 4.17 and 1.12.

tinuous updating of sunspot records is offered by the Sunspot Index and Long-term Solar Observations (SILSO) Centre of the Royal Observatory of Brussels, Belgium, and the next development of the current research will consider more recent data.

The solar cycle has roughly a 11-years period with asymmetric cycles characterized by rising period shorter than the descending period. The sunspot numbers are shown in Figure 7.9: the time series is evidently non-stationary and the main

goal for scientists is the prediction of future sunspot numbers. The SILSO centre has implemented many techniques to predict the nature of the solar cycle and, among the other, also a Kalman-filter based method is included; for a review of the different methods used to deal with sunspot numbers, the reader is referred to Tong (1990) and the SILSO web page.

This preliminary analysis does not predict the future sunspot number but aims to discuss the modelling skills of the Duffing process.

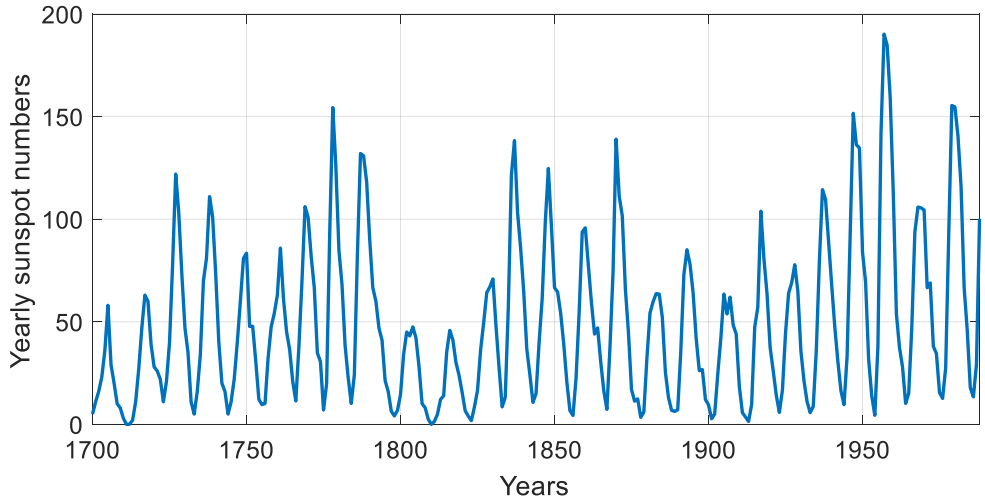

Figure 7.9: Annual mean of sunspot numbers from 1700 to 1988. Data source: Tong (1990).

7.2.1. Wavelets and solar cycles

Rebuilding the heuristic approach of Chapter 6 for sunspot numbers, the DB and Symlets wavelet transforms are compared in Figure 7.10. The DB7 has the lowest entropy value, but it shows a bigger distance from the real signal in comparison both to the same transform with less number of vanishing moments and the Symlets wavelets. On the contrary, since the Sym4 has the smallest distance between the MRA reconstruction and the observations, this is the wavelet chosen to detect the changepoints in the nature of solar cycles.

The maximum number of coefficients of the mother wavelets allowed for the sunspots data is 8: the contribution of these coefficients in the total variance of data

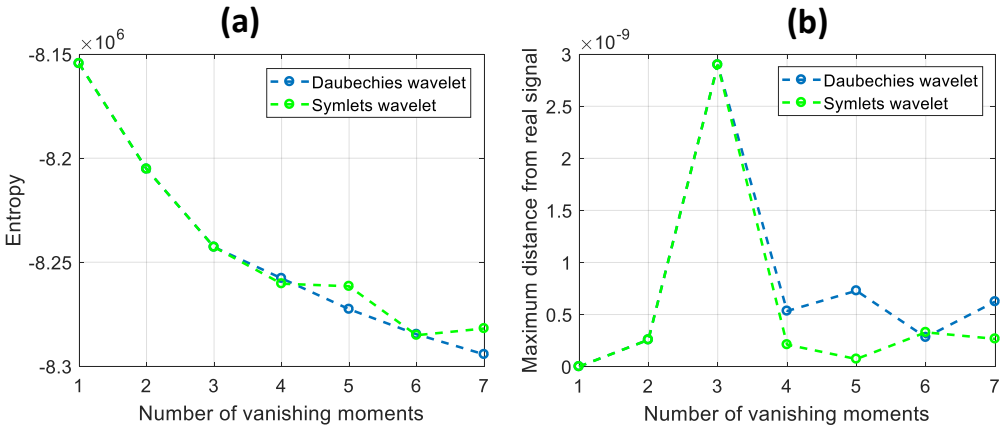

Figure 7.10: Sunspot data: comparison between wavelets to choose the number of vanishing moments. (a) Comparison in terms of entropy. (b) Comparison in terms of maximum distance from the MRA and the observations.

is shown in Figure 7.11. The most explanatory coefficient is the third, while the others account less in the total variability of the time series and lead the conclusion that three coefficients could be sufficient in the computation of the Symlets wavelet transform.

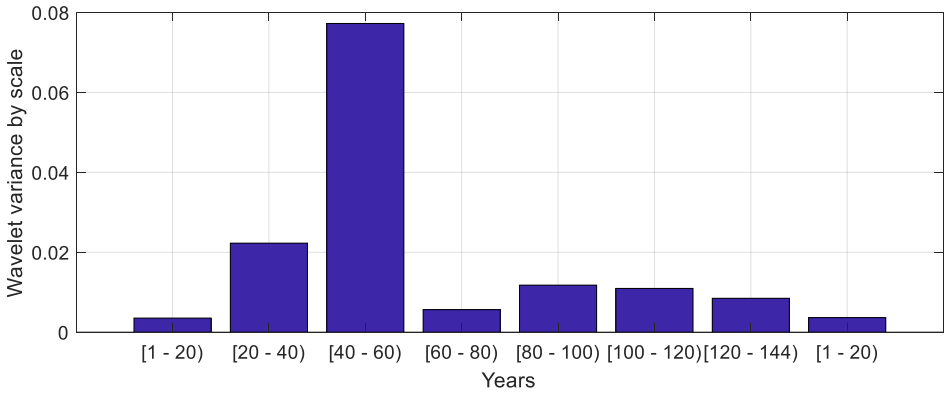

Figure 7.11: Contribution to the total variance of sunspot numbers of the data of the Symlets wavelet with four vanishing moments and 8 coefficients (the maximum number allowed) of the mother wavelets.

At the end of the heuristic procedure, the changepoints of the variance of wavelets indicate that the breakpoints, i.e. the changes in the nature of the solar cycle, occur in the years 1776-1777, 1791-1800, 1837-1839 and 1962 (Figure 7.12). These years are located in the sunspot numbers time series in Figure 7.13. The period 1800-1837 has the smallest amplitudes of the oscillation while the very antecedent and immediately following period show some of the highest values of the whole series.

The years depicted by the red lines in Figure 7.13 are considered the changepoints in the inference presented below.

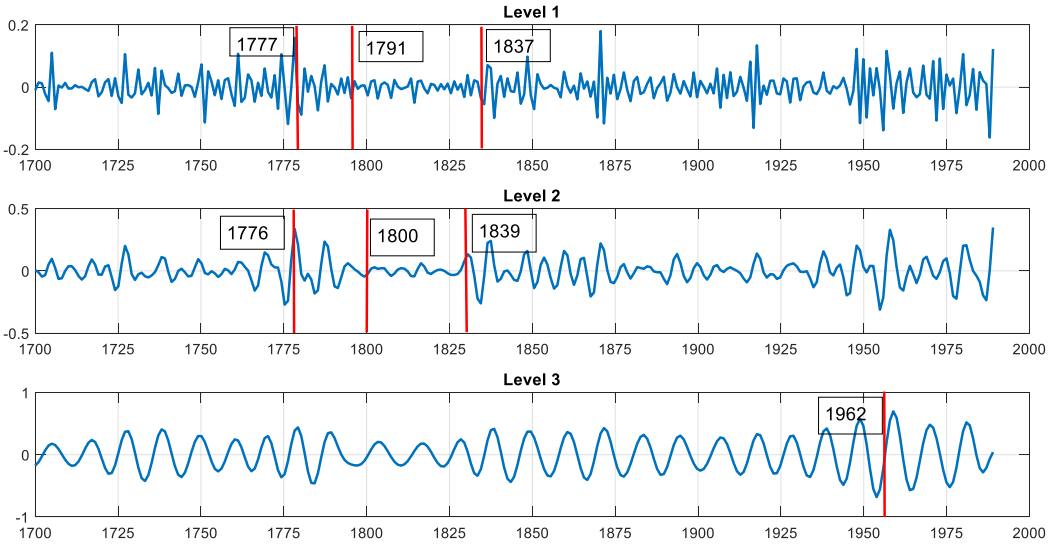

Figure 7.12: The three level of details of the Sym4 with 3 coefficients. The red straight lines indicate the changepoints in the variance of the wavelets that happen in the years written in the boxes.

7.2.2. Estimates for sunspot numbers

The UKF signal and parameter estimates of sunspots data modelled by the Duffing system are shown in Figure 7.14 and Table 7.2, respectively. The RSS of the filter reconstruction: (i) for the first regime (1700-1771) is 0.98, (ii) for the years 1778-1791 is 1.71, (iii) in the period 1792-1837 is 4.14, (iv) in 1838-1962 is 4.37 and finally (v) in 1963-1988 is 0.17.

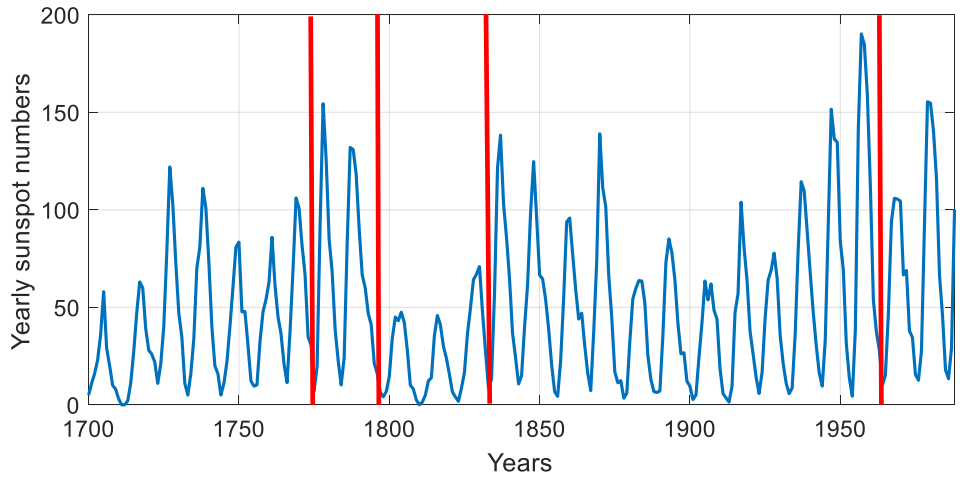

Figure 7.13: Changepoints of the parameter values for the sunspot data. The blue line is the mean of sunspot numbers, the red lines depict the changepoints corresponding to the years 1777, 1791, 1837, 1962.

Years	α	β	c
1700–1777	39.50	−0.01	0.20
1776–1791	34.70	1.06	−0.08
1792–1837	40.61	0.58	0.25
1838–1962	42.10	−0.50	−0.18
1963–1988	40.75	0.37	−0.28

Table 7.2: Parameter estimates for the sunspot data modelled by the Duffing process and estimated with the UKF method.

At first glance, one could immediately notice that β and c alternate between negative and positive values. The two regimes with the shorter time span, 1778–1791 and 1963–1988, share the feature of a negative damping term, while a negative β is estimated for the regimes with the longest time interval. In the phase space analysis, these parameter values indicate the existence of saddles which can arise both in the

trivial and non-trivial fixed points and may lead to bifurcations. The next steps in the analysis of sunspot numbers will focus on estimates for phase portraits, i.e. the phase space depicted by the estimation of the position and the velocity of the Duffing model (the variables x_{1t} and x_{2t} of equation (5.23) of Chapter 5). Such study of attractors and repellors in the sunspot numbers can give an alternative point of view and may gain more insight into the solar cycles.

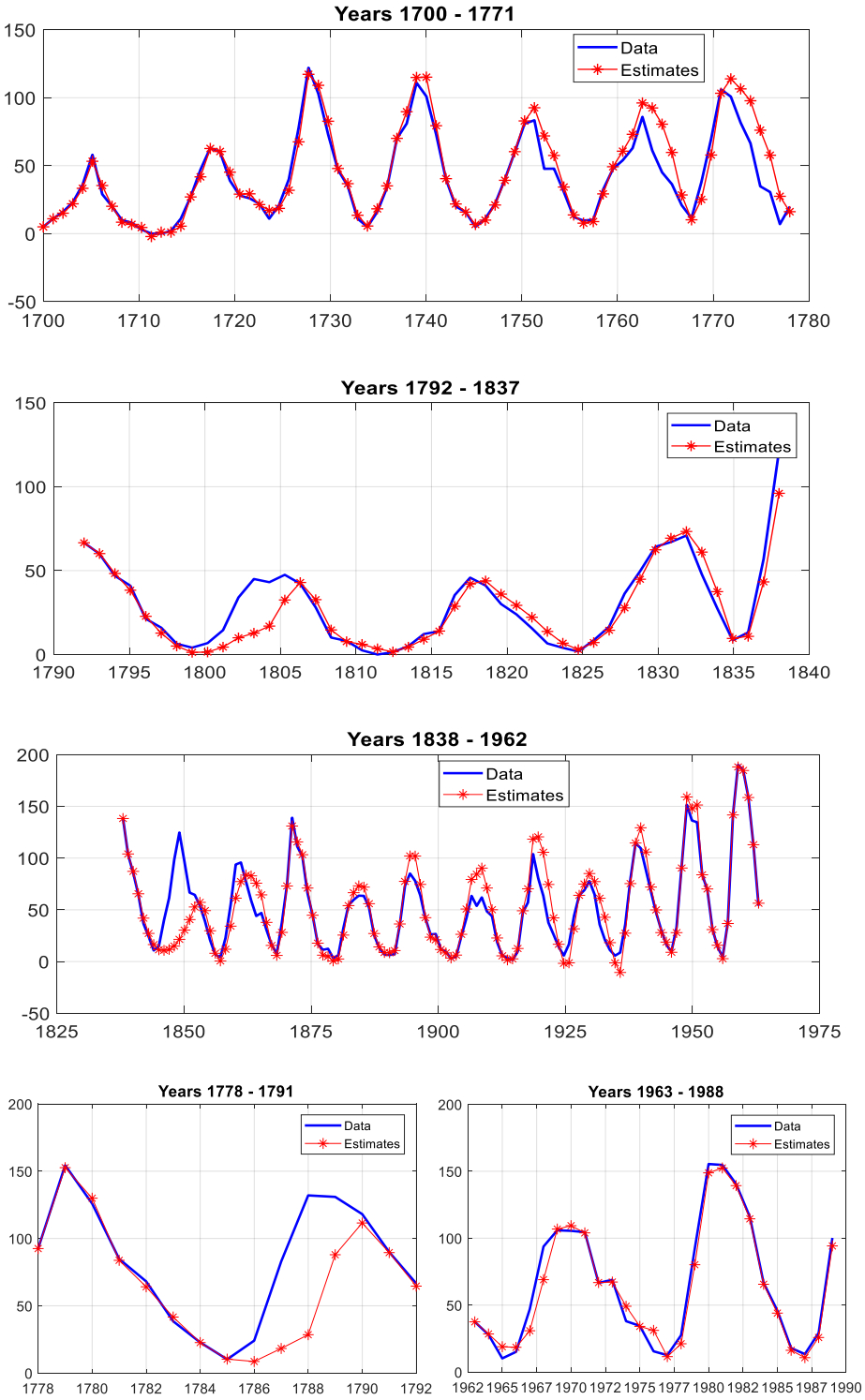

Figure 7.14: Signal reconstruction of the UKF for the sunspot numbers modelled by the Duffing process.

8. Conclusions and outlook on future research

This Chapter summarizes the dissertation, discusses the main findings and contributions, points out limitations of the current work, and finally outlines directions for future research.

8.1. *Concluding remarks*

8.1.1. General framework

This dissertation deals with inference methods for a differential equation called the Duffing system, both in the deterministic and stochastic case. The thesis focuses on the Duffing equation, a process that describes many non-linear and chaotic phenomena but real data modelling with such system is still quite unexplored.

The UKF is the main algorithm used to conduct parameter inference and signal reconstruction for the Duffing equation, but also other schemes have been investigated. The UKF is a non-linear version of the Kalman filter and it is based on the unscented transform developed by Julier and Uhlmann (2004) that fits Gaussian approximations on a set of points, known as sigma points.

After a comprehensive review of state space models, Kalman filter and Kalman filter non-linear extensions, the discussion concentrates on the UKF limits, with respect to the location of sigma points and the choice of the starting values of the algorithm. In particular, the more the UKF is initialized far from the true parameters, the more the UKF estimates lose accuracy, i.e. the method do not converge to the true values.

8.1.2. Contribution for ODE inference

The first part of the study concerns inference techniques for the deterministic Duffing equation. Once the uncertainty of the UKF estimates has been quantified and the UKF proved its inference power in the ODE context, its limitations are overcome in the following way.

To find a sigma points placement in the marginal likelihood space that let the UKF to result in accurate estimates, the likelihood function has to be maximized. Due to the chaoticity of the Duffing system, the likelihood is highly multi-modal and the search for its maximum is made with Bayesian optimization. BO is a sequential model based approach that allows to face the problems of non-convexity and multi-modality of objective functions. The strategy to optimize the sigma set location consists by developing the BO algorithm so that it can identify the underlying assumptions of the model (i.e. the position of points in the space) and learn the sigma points placement from training data. The BO method returns a sigma set used to compute the unscented transform into the UKF steps; this approach is validated through a simulation study. In the context of the Duffing equation, the performance of the EI and the UCB acquisition functions is also compared. The results in terms of convergence to the true values of the UKF with optimized sigma points location are quantified by the Euclidean norm in functional and parameter space. Optimizing the sigma points in the filtering phase consistently reduces the Euclidean distance between the final estimates and the true values and the EI is the acquisition function that gives the best convergence.

However, even if the BO strategy consents to get more precise estimates than the deterministic sigma points assignment, the improvement gained with the BO decreases with a "worse" UKF initialization. To overcome the limit related to the choice of the starting values, the class of the ABC methods is used as a prelude for the UKF inference. The ABC approach, also called likelihood free, has been chosen thanks to its characteristic of avoiding the computation of the likelihood function by means of comparison between observed and simulated data. The summary statistics for the Duffing equation utilized into the ABC, with a SMC sampling scheme, are defined on the grounds of the phase space analysis. The idea behind the use of the ABC coupled with the UKF consists by sampling from the approximate posterior distributions of the ABC-SMC scheme the starting values for the UKF. The proposed algorithm is called Sequential ABC-UKF and it outperforms the default UKF giving a massive improvement in the signal reconstruction and parameter estimates. The parameter most affected by uncertainty, quantified by the flatness of the ABC

approximate posterior distribution, is β, the coefficient associated to the cubic term and the source of chaos in the system.

8.1.3. Contribution for SDE inference

The second part of the research focuses on the stochastic Duffing system.

To allow the UKF to carry out parameter inference for stochastic processes, the prediction steps of the algorithm should include a numerical integration method as the Euler-Maruyama scheme to find an approximate solution of the transition function. A simulation study evaluates the UKF estimates when the Euler-Maruyama integration method is inserted within the filtering phase for several sizes of the variance of the process noise.

The UKF associated with the Euler-Maruyama is able to infer the parameters. Contrarily to the deterministic case, in the SDE context, the most difficult parameter to infer is the damping term c. This is due to a characteristic of the Duffing system discussed in the geometric analysis at the beginning of this dissertation. In the phase space representation, it has been shown that the damping values drives the stability of the fixed points and that an attractor may become a repellor if a small random perturbation affects the system.

A successive simulation study compares the UKF results in the case of known and unknown state noise variance in terms of RMS error. As one would expect, when the variance of the process noise belongs to the parameter vector the UKF has to estimate, the RMS error is higher and grows faster than the case of knowledge of the variance.

To treat the UKF limits with respect to the sigma points assignment and the initialization in the context of SDEs, a different approach from the one used for ODEs is taken in order to investigate other methods.

The sigma set is computed for the augmented and non-augmented versions of the UKF and the two methods are compared. The addition of the process noise in the state space model, indeed, requires a double computation of the sigma points, one for the evaluation of the transition and measurement functions, and one for the process noise: this is what the non-augmented UKF does. The augmented UKF, instead, adds the noise to the covariance matrix (in other words, it augments the

size of the matrices) to avoid a repeated calculation of sigma points. A simulation study that will be extended in the near future shows the poor performance of the non-augmented UKF even for a small size of process noise variance.

The search of an alternative strategy to the ABC-SMC prelude to deal with the initialization of the UKF is presented in a preliminary study. The idea behind this other approach, called Iterative UKF, consists in successive UKF evaluations by inserting the final estimates of one UKF as starting points of the next UKF. Until now, this method suffers of numerical instability of matrices but an improvement is achieved for the first iterations.

Finally, the UKF performance in compared on a independent dataset simulated by a model for pharmacokinetics dynamics. The UKF estimates are compared with a new algorithmic methodology (SAEM-ABC) developed by Picchini and Samson (2017). This comparison proves that the UKF coupled with the Euler-Maruyama challenges alternative approaches for inference on SDEs.

8.1.4. Real data illustrations

The final part of the thesis concentrates on real data analysis.

Many time series may show different behaviours depending on the period of time the measurements are made; e.g., an oscillation in a certain time interval can be described by some features that are not present in another interval. In example, an economic time series can reduce its volatility after a period of time. Different characteristics of an oscillation may be ascribed by a change in the generative parameter values of the underlying process. In such a case, one could imagine that hidden states, or regimes, exist behind the observed data and each regime is defined by its own parameter values.

Modelling a time series that switches to one regime to another requires to locate the breakpoints of parameters in time by using a changepoint detection method.

The changepoint strategy developed so far is heuristic in the sense that is model free but based on wavelet transforms. In particular, the change in the variance of the wavelets is used as an indicator of a regime shift.

The heuristic approach to identify the breakpoints is used in real data analysis. Once the changing points are found, the Duffing process is used for modelling two

well-known time series, the U.S. GDP and the sunspot numbers data, that have been extensively analysed in the literature and that could provide a benchmark for the proposed model. The Duffing equation is able to reconstruct the oscillation and states its competitiveness as a new alternative approach for modelling non-linear dynamical systems.

8.2. *Future work*

While this study has demonstrated the potentiality of the Duffing process and the UKF as inference method for ODEs and SDEs, many opportunities for extending the scope of this dissertation remain. This Section presents some of these directions.

8.2.1. Simulation extensions

More simulations will be carried out to further discuss some topics which have not been fully developed in the thesis. In particular, the following points are the near future steps.

Geometric analysis of UKF estimates. The UKF convergence of estimates will be evaluated in terms of stability of the fixed points by analysing if the estimated position and velocity give back the same phase portrait of the simulated data.

Comprehensive comparison between augmented and non-augmented UKF. In the context of SDEs, the comparison between augmented and non-augmented UKF needs to be carried out in a more extensive manner by evaluating the results for many different values of the process noise variance.

Alternative strategy of initialization. An alternative strategy to initialize the UKF can be developed with the use of gradient matching techniques, which are approximate methods that have recently gained much attention in the literature (see e.g. Dondelinger et al., 2013 and Niu et al., 2016). The gradient matching scheme aims to minimize the discrepancy between the slope of a data interpolant and the derivatives predicted from the differential equations by using a surrogate cost function with Gaussian processes. Such a study will not only give an alternative method

to find the starting values of the UKF but will also provide a comparison with the ABC approach.

Critical issues on the choice of wavelets. For a more comprehensive validation of the heuristic changepoint approach based on wavelets, more simulations are necessary in order to quantify how the type of wavelets affects (i) the changepoints in the variance of wavelets and (ii) the detection of breakpoints location in time of the generative parameters of the process.

8.2.2. Model based changepoint detection method

A model-based approach to detect the changepoints in real time series will substitute the heuristic approach to gain more robust estimates.

Adams and MacKay (2007) have proposed a Bayesian online changepoint detection strategy and more recently the study of Mavrogonatou and Vyshemirsky (2016) extends this method by substituting the conjugate prior requirement with a sequential importance sampling scheme.

The next step will aim to use a changepoint model based on the Duffing equation. In other words, the Duffing SDE will be used to model data and the technique proposed by Mavrogonatou and Vyshemirsky (2016) will be the method to infer the chagepoints. In this way, when the Duffing system models real datasets, the researcher will be able to infer the SDE parameters along with the detection of changepoints.

With a model based inquiry of breakpoints in the parameter values, the real data analysis will be extended and will include comparisons with the already existing literature.

Bibliography

Adams, R. P., and MacKay, D. J. (2007). Bayesian online changepoint detection. *arXiv preprint arXiv:0710.3742*.

Andronov, A. A., and Khajkin, S. E. (1949). *Theory of oscillations*. Princeton University Press.

Ball, L. M. (2009). Hysteresis in unemployment: old and new evidence. *National Bureau of Economic Research*, no. w14818.

Beaumont, M. A. (2010). Approximate Bayesian Computation in Evolution and Ecology. *Annu. Rev. Ecol. Evol. Syst.*, 41, 379-406.

Blanchard, O. and J. Simon (2001). The long and large decline in U.S. output volatility. *Brookings Papers on Economic Activity*, 135-164.

Blanchard, O. J., and Summers, L. H. (1986). Hysteresis and the European unemployment problem. *NBER macroeconomics annual*, 1, 15-78.

Bowman, A. W. and Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis* . Oxford Science Publications.

Cox, H. (1964). On the estimation of state variables and parameters for noisy dynamic systems. *IEEE Transactions on Automatic Control*, 9(**1**), 5-12.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, Series B, 1-38.

Dondelinger, F., Husmeier, D., Rogers, S., and Filippone, M. (2013). ODE parameter inference using adaptive gradient matching with Gaussian processes. In: *Proceedings of the 16th International Conference on Artificial Intelligence and Statistics*, Scottsdale, AZ, USA, 216-228.

Donnet, S. and Samson, A. (2008). Parametric inference for mixed models defined by stochastic differential equations. *ESAIM: Probability and Statistics*, 12, 196–218.

Duffing, G. (1918). *Forced Oscillations with Variable Natural Frequency and their Technical Significance*. Braunschweig, Freidrich Vieweg and Son.

Euler, L. (1750). De novo genere oscillationum. *Comment. acad. sc. Petrop*, 11(**1739**), 128–149.

Gallegati, M. and Semmler, W. (Eds.) (2014). *Wavelet Applications in Economics and Finance*. Springer.

Giurghita, D., and Husmeier, D., (2016). Inference in Nonlinear Systems with Unscented Kalman Filters. In: *22nd International Conference on Computational Statistics*, Oviedo, Spain, 23-26 Aug 2016, 383–393.

Haggan, V., and Ozaki, T. (1981). Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika*, 68(**1**), 189–196.

Hartig, F., Calabrese, J, M., Reineking, B., Wiegand, T., Huth, A. (2011). Statistical Inference for stochastic simulation models – theory and application. *Ecology Letters*, 14, 816–827.

Hartikainen, J., Solin, A., Särkkä, S. (2011). EKF/UKF Toolbox for MATLAB. `http://becs.aalto.fi/en/research/bayes/ekfukf/`.

Helmholtz, H. L. F. (1885). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Translated by Ellis A. J. (1895). Longmans Green and Co.

Higham, D. J. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM review*, 43(**3**), 525–546.

Holmes, P. (2005). Ninety plus thirty years of nonlinear dynamics: Less is more and more is different. *International Journal of Bifurcation and Chaos*, 15(**09**), 2703–2716.

Holmes, P. J., and Rand, D. A. (1976). The bifurcations of Duffing's equation: An application of catastrophe theory. *Journal of Sound and Vibration*, 44(**2**), 237-253.

Huang, D., Allen, T. T., Notz, W. I., and Zeng, N. (2006). Global optimization of stochastic black-box systems via sequential kriging meta-models. *Journal of global optimization*, 34(**3**), 441-466.

Huygens, C. (1673). *Horologium Oscillatorium sive de motu pendulorum*. Muguet.

Jones, D. R., Schonlau, M., and Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(**4**), 455-492.

Jones, P. J., Sim, A., Taylor, H. B., Bugeon, L., Dallman, M. J., Pereira, B., and Liepe, J. (2015). Inference of random walk models to describe leukocyte migration. *Physical Biology*, 12(**6**), 066001.

Julier, S. J., and Uhlmann, J, K. (2004). Unscented Filtering and Nonlinear Estimation. *Proceedings of the IEEE*, 92(**3**), 401-422.

Julier, S. J., and Uhlmann, J. K. (1997). A new extension of the Kalman filter to nonlinear systems. In: *International Symposium Aerospace/Defense Sensing, Simulations and Controls*, 3(**26**), 182-193.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(**1**), 35-45.

Kalman, R. E. and Bucy, R. S. (1961). New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering*, 83(**1**), 95-108.

Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*. Springer-Verlag, Berlin.

Kim, C.-J., and C. R. Nelson (1999). Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *The Review of Economics and Statistics*, 81, 608-616.

Kovacic, I., and Brennan, M. J. (Eds.) (2011). *The Duffing equation: nonlinear oscillators and their behaviour*. John Wiley & Sons.

Kumar, P., Narayanan, S., and Gupta, S. (2016). Stochastic bifurcations in a vibro-impact Duffing–Van der Pol oscillator. *Nonlinear Dynamics*, 85(**1**), 439-452.

Kumar, P., and Narayanan, S. (2010). Modified path integral solution of Fokker–Planck equation: response and bifurcation of nonlinear systems. *Journal of Computational and Nonlinear Dynamics*, 5(**1**), 011004.

Ljung, L. (1979). Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24(**1**), 36-50.

Marin, J. M., Pudlo, P., Robert, C. P., and Ryder, R. J. (2012). Approximate Bayesian computational methods. *Statistics and Computing*, 1-14.

Marinca, V., and Herişanu, N. (2011). Explicit and exact solutions to cubic Duffing and double-well Duffing equations. *Mathematical and Computer Modelling*, 53(**5**), 604-609.

Mavrogonatou, L., and Vyshemirsky, V. (2016). Sequential Importance Sampling for Online Bayesian Changepoint Detection. In: *22nd International Conference on Computational Statistics*, Oviedo, Spain, 23-26 Aug 2016, 73-84.

Maybeck, P. S. (1979). *Stochastic models, estimation, and control*. Academic Press.

Mojon, B. (2007). Monetary policy, output composition and the Great Moderation. *Federal Reserve Bank of Chicago*. WP 2007-07.

Mockus, J., Tiesis, V., and Zilinskas, A. (1978). Toward global optimization. Volume 2, Chapter in *Bayesian methods for seeking the extremum*, 117-128.

Murphy K., P. (2012). *Machine Learning: A probabilistic Perspective*. MIT Press.

Niu, M., Rogers, S., Filippone, M., and Husmeier, D. (2016). Fast Inference in Non-linear Dynamical Systems using Gradient Matching. In: *Proceedings of the 33rd International Conference on Machine Learning*, New York, NY, USA.

Ozaki, T. (1981). Non-linear threshold autoregressive models for non-linear random vibrations. *Journal of Applied Probability*, 18(**2**), 443-451.

Pasetto M. E., Husmeier D., Noè U., and Luati A. (2017a). Statistical Inference in the Duffing System with the Unscented Kalman Filter. In: *32nd International Workshop on Statistical Modelling*, Groningen, The Netherlands, 3-7 July 2017, 119-122.

Pasetto M. E., Noè U., Luati A., and Husmeier D. (2017b). Inference with Unscented Kalman Filter and Optimization of Sigma Points. In: *Conference of the Italian Statistical Society*, Florence, Italy, 28-30 June 2017, 767-772.

Percival, D. B., and Walden, A. T. (2006). *Wavelet methods for time series analysis* (Vol. 4). Cambridge university press.

Perttunen, C. D., Jones, D. R., and Stuckman, B. E. (1993). Lipschitzian optimization without the Lipschitz constant. *Journal of Optimization Theory and Applications*, 79(**1**), 157-181.

Picchini, U., and Samson, A. (2017). Coupling stochastic EM and Approximate Bayesian Computation for parameter inference in state-space models. *Computational Statistics*, `https://doi.org/10.1007/s00180-017-0770-y`.

Pinheiro J., and Bates, D. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*, 4(**1**), 12-35.

Poincaré, H. (1881). Mémoire sur les courbes définies par une équation différentielle (I). *Journal de Mathématiques Pures et Appliquées*, 7, 375-422.

Rayleigh, J. W. S. B. (1896). *The theory of sound*. Macmillan.

Ramasubramanian, K., and Sriram, M. S. (2000). A comparative study of computation of Lyapunov spectra with different algorithms. *Physica D: Nonlinear Phenomena*, 139(**1**), 72-86.

Ramsey, J. B. (2002). Wavelets in economics and finance: Past and future. *Studies in Nonlinear Dynamics & Econometrics*, 6(**3**).

Ramsey, J. B. (1990). Economic and financial data as nonlinear processes. In: *The Stock Market: Bubbles, Volatility, and Chaos*, Springer Netherlands, 81-139.

Rasmussen, C. E., and Williams, C. K. (2006). *Gaussian processes for machine learning*. MIT press.

Salas, A. H. (2014). Exact solution to Duffing equation and the pendulum equation. *Applied Mathematical Sciences*, 8(**176**), 8781-8789.

Schenk-Hoppé, K. R. (1996). Bifurcation scenarios of the noisy Duffing-van der Pol oscillator. *Nonlinear Dynamics*, 11(**3**), 255-274.

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and de Freitas, N. (2016). Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(**1**), 148-175.

Shaw, S. W., and Balachandran, B. (2008). A review of nonlinear dynamics of mechanical systems in year 2008. *Journal of System Design and Dynamics*, 2(**3**), 611-640.

Simon, D. (2006). *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons.

Sisson, S. A., and Fan, Y. (2011). *Handbook of Markov Chain Monte Carlo*, chapter Likelihood-free MCMC. CRC Press.

Sitz, A., Schwarz, U., Kurths, J., and Voss, H. U. (2002). Estimation of parameters and unobserved components for nonlinear systems from noisy time series. *Physical Review E*, 66(**1**), 016210.

Snoek, J., Larochelle, H., and Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In: *Advances in Neural Information Processing Systems*, 2951-2959.

Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. (2009). Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*.

Stock, J. H., and Watson, M. W. (2002). Has the business cycle changed and why?. *NBER macroeconomics annual*, 17, 159-218.

Stoker, J. J. (1950). *Nonlinear vibrations in mechanical and electrical systems*. Interscience Publishers.

Strang, G., and Nguyen, T. (1996). *Wavelets and filter banks*. SIAM.

Thomsen, J. J. (2013). *Vibrations and stability: advanced theory, analysis, and tools*. Springer Science & Business Media.

Thrun, S., W. Burgard, and D. Fox (2006). *Probabilistic Robotics.* MIT Press.

Tong, H. (1990). *Non-linear time series. A Dynamical System Approach*. Oxford Science Publications.

Tong, H., and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society*. Series B, 245-292.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., Stumpf, M. P. H. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface* 6(**31**), 187-202.

Turner, R. and Rasmussen, C. E. (2012). Model based learning of sigma points in unscented Kalman filtering. *Neurocomputing*, 80, 47-53.

Ueda, Y. (1985). Random phenomena resulting from non-linearity in the system described by Duffing's equation. *International Journal of Non-Linear Mechanics*, 20, 481-491.

Ueda, Y. (1979). Randomly transitional phenomena in the system governed by Duffing's equation. *Journal of Statistical Physics*, 20(**2**), 181-196.

Van der Pol, B. (1920). A theory of the amplitude of free and forced triode vibrations. *Radio Review*, 1, 701-710.

Vazquez, E., and Bect, J. (2007). On the convergence of the expected improvement algorithm. *No. arXiv: 0712.3744*.

Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. (1985). Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(**3**), 285-317.

Wu, Y., Hu, D., Wu, M., and Hu, X. (2005). Unscented Kalman filtering for additive noise case: augmented vs. non-augmented. In: *American Control Conference*, Portland, USA, 08-10 June 2005, 4051-4055.