POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Aerospaziale

Time-periodic sensitivity analysis methods for chaotic systems

Relatori:
Prof. Davide Lasagna
Prof. Gaetano Iuso

Candidato:
Andrea Bongiovanni

2017/2018

Sommario

In questo progetto viene presentato un nuovo approccio per eseguire analisi di sensitività su sistemi dinamici caotici. L'analisi di sensitività è uno strumento ben noto ed efficiente utilizzato per comprendere come le variabili di progetto che caratterizzano un dato sistema, ad esempio forma e geometria di un corpo aerodinamico, influiscano su un determinato osservabile del sistema, quale ad esempio la resistenza aerodinamica del corpo stesso. L'analisi di sensitività è in grado di fornire il gradiente dell'osservabile del sistema rispetto a queste variabili e dunque risulta uno strumento estremamente efficace per risolvere problemi di ottimizzazione. Oggigiorno in applicazioni aeronautiche, ad esempio, l'analisi di sensitività può essere accoppiata a simulazioni Reynolds-Averaged-Navier-Stokes (RANS) per la progettazione di corpi aerodinamici del gruppo ala-fusoliera o per applicazioni nel campo delle turbomacchine. La necessità di sviluppare nuovi metodi per eseguire analisi di sensitività risiede nel fatto che, applicata a sistemi dinamici con comportamento caotico come la turbolenza, i classici algoritmi falliscono, fornendo gradienti che non rispecchiano la fisica del problema in esame. Tale fenomeno si verifica poiché un'infinitesima perturbazione nelle variabili di progetto di un sistema caotico, ad esempio la viscosità del fluido, si riflette in un'evoluzione dello spazio degli stati che diverge esponenzialmente dalla traiettoria di riferimento caratterizzante il sistema non perturbato. Dato che le simulazioni su flussi turbolenti assumeranno un ruolo sempre più importante nei cicli di progettazione dell'avvenire e dato che gli attuali algoritmi di analisi di sensitività risultano completamente inefficaci quando accoppiati a tali simulazioni, risulta comprensibile la necessità di sviluppare dei metodi per colmare questa mancanza.
Il metodo per l'analisi di sensitività su sistemi caotici che viene presentato in questa tesi prende il nome di *Periodic Shadowing* (Lasagna, *arXiv:1806.02077*, 2018). In particolare lo scopo di questo progetto è quella di validare l'efficacia del metodo applicandolo a due sistemi caotici a bassa dimensionalità di complessità crescente, derivanti da problemi di convezione naturale in celle fluide delimitate da due piastre poste a temperature differenti. In particolare verrà valutata la sensitività della media temporale di una quan-

tità correlata al trasferimento di calore nelle celle fluide rispetto ad un parametro fisico proporzionale alla differenza di temperatura tra le piastre.
Dai risultati è stato possibile osservare che mentre la sensitività media converga per traiettorie di lungo periodo, il confronto con l'approssimazione del gradiente ottenuto mediante differenze finite presenta un bias coerente e riproducibile. Si congettura che questo errore sia una caratteristica dei sistemi a bassa dimensionalità e che tuttavia diverrebbe trascurabile qualora venissero considerati sistemi caotici con molteplici gradi di libertà, come i flussi turbolenti ad alto numero di Reynolds. Lo sviluppo di una piena e solida comprensione di questo aspetto è attualmente oggetto di ricerca.

Abstract

This project, partially carried out at the Politecnico di Torino and at the University of Southampton, describes a new approach for sensitivity analysis of chaotic dynamical systems. Sensitivity analysis is a well-known and efficient tool used to understand how design and control variables characterizing a given system, e.g. the shape of an aerodynamic body, affect a specific output of interest, e.g. the drag of the body. Sensitivity analysis provides the gradient of the output with respect to these parameters and, for this reason, is very useful to solve optimization problems. Nowadays, for instance, in aeronautical applications, sensitivity analysis is coupled to steady Reynolds-Averaged-Navier-Stokes (RANS) for aerodynamic shape design of wing-fuselage bodies or for turbomachinery applications. The need to develop new methods for sensitivity analysis lies in the fact that, applied to chaotic systems like turbulence, classical algorithms fail, providing non-physical gradients. This occurs because in a chaotic system a small perturbation of the design variables, e.g. the fluid viscosity, is reflected in an evolution of the trajectory that diverges exponentially from the reference trajectory characterizing the non-perturbed system. Since scale-resolving turbulent flow simulations will take an increasingly larger role in the engineering design cycle and since the current generation of sensitivity methods is completely ineffective when coupled to unsteady turbulent flow simulation, the development of new methods to make up for this lack is necessary.

In this thesis a new sensitivity analysis method suitable for chaotic systems, called Periodic Shadowing (Lasagna, *arXiv:1806.02077*, 2018), is presented. The aim of this project is to validate the effectiveness of this method by applying it to two low-dimensional chaotic systems of increasing complexity, arising from natural convection problems in fluid cells bounded by two plates and heated from below. We evaluate the sensitivity of the time average of a quantity related to the heat transfer in the fluid cells with respect to a physical parameter proportional to the temperature difference between the plates.

We observed that while the time average sensitivity converges for long time trajectories, the comparison with the finite-difference approximation of the gradient shows a consis-

tent bias error. It is conjectured that this error is a feature of these low-dimensional systems, but it would be negligible for chaotic systems with many degrees of freedom, like turbulent flows at high Reynolds number. Developing a fundamental understanding of this aspect is currently subject of research.

Acknowledgements

I would like to thank my supervisor, Dr. Davide Lasagna for the patient guidance, encouragement and advice he has provided throughout my time as his student. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. The door of his office has always been open and he was always been available to discuss new results or ideas. I feel really proud of helping his research with my thesis project.
I would also like to thank my co-supervisor Prof. Gaetano Iuso for having him as professor and for his constant presence in the project.
I must express my gratitude to my mother, my father and my sister for their continued support and encouragement. They experienced all of the ups and downs of my research and academic path, finding the best word on every occasion. I would like to thank my girlfriend for supporting and keeping me motivated a throughout this work. Sweet thanks also for my brother-in-law and my nephews.
All my friends I met during my academic journey are valuable people who have always wished well and I have enjoyed the time that I spent with them.

A. B.

Contents

List of Figures

Chapter 1

Introduction

"Optimization" term is a leitmotif widely used in the modern approach industry. Taking a look at the Cambridge Dictionary, optimization represents the process of making something work as effectively as possible, which is another way to say better solution and, in the engineering environment, more saving. The typical optimization problem in the aeronautic field is to find the optimal configuration, shape or structure which minimizes/maximizes a certain cost functional and which satisfies some given constraints. To achieve these results, sensitivity analysis can be used coupled to other numerical tools in order to exploit gradient information and rapidly obtain the optimal solution at a cost smaller then what is required to explore large parameter spaces.
In practice, a wide range of optimization problems requires an estimate of the sensitivity of model-simulated quantities to changes in initial conditions, boundary conditions, or external forcing. Sensitivity analysis achieves this by evaluating the gradient of the time averaged observables of interest with respect to design and control variables, e.g. how the shape of an aerodynamic body influences the drag of the body itself. With the significant advances in computer technology over the last thirty years, industries which design and manufacture high-performance products are increasingly interested in exploiting the advantages of computer-aided design, numerical analysis and optimal design methods. Classical examples that can be found in literature refer to turbulent combustion or shape design optimization for wing-fuselage bodies and turbomachinery applications [1, 2]. In figure 1.1, we present some examples of mentionable and current fields of application of sensitivity analysis, starting from geometry optimization of the free-form surfaces in a water pump, panel (a), to aerodynamic shape optimization, panels (b) and (c). In particular, considering panel (c) of figure 1.1, the example refers to an aerodynamic shape design process where sensitivity analysis is carried out coupled to steady Reynolds-Averaged-Navier-Stokes (RANS) equations. The drag coefficient is

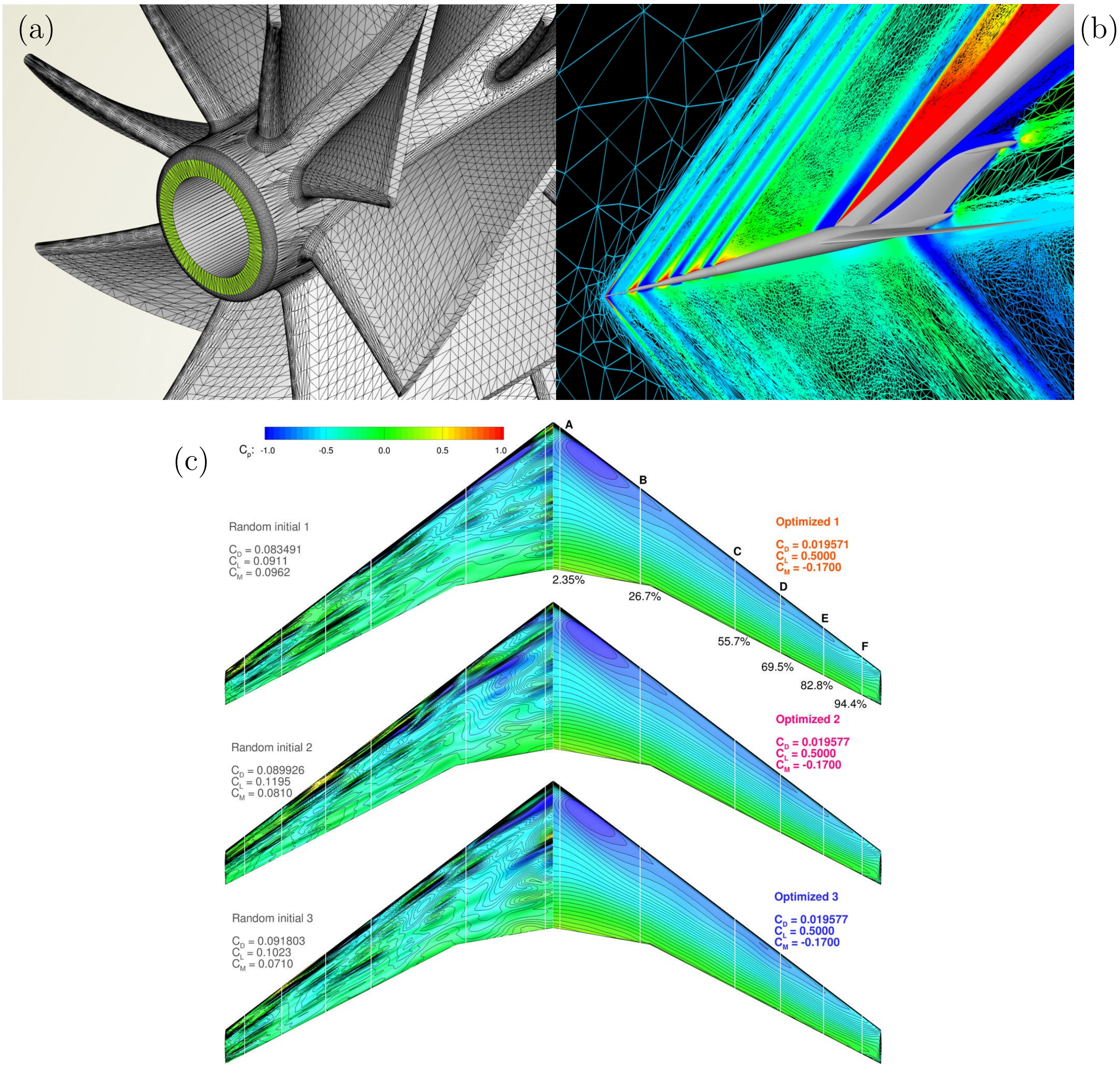

Figure 1.1: Shape optimization of water pumps with shrouded impellers using commercial software (a), RANS simulation of supersonic business jet configuration with blade sting mount using mesh adaptation based on the adjoint solution for the off-body pressure signature [*Eric Nielsen, The FUN3D Develompent Team, NASA*] (b) and wing geometry optimisation [3] (c).

minimized subject to lift, pitching moment and geometric constraints. As stated, all the optimization processes can be performed through a computational path before to any time-consuming, expensive and labour physical experiments. For that reason, computational methods for sensitivity analysis are an efficient tool in modern computational science and engineering.

1.1 Tangent and Adjoint Sensitivity Analysis

The way to approach the sensitivity analysis depends on the type of problem under examination. In fact, to select the optimal sensitivity analysis method to be applied, it is necessary to consider the number of design variables with respect to which we need to evaluate the sensitivity of the observable of interest. To provide a very broad view of the topic, we can subdivide the sensitivity analysis methods in tangent or adjoint approaches. The tangent method allows calculating the sensitivity of the observable of interest with respect to a variation in one (or few) design parameters. If the sensitivity with respect to several parameters is required, the tangent algorithm must be repeated for each perturbation applied. To make it clearer, the typical optimization problem has a single objective function (possibly combining multiple objectives through suitable weights) that has to be minimized or maximized with respect to a large number, or even a continuum, of input variables. In such a context, the tangent method becomes computationally expensive and, therefore, the adjoint methods are preferable. A clarifying conceptual scheme is presented in the figure 1.2.

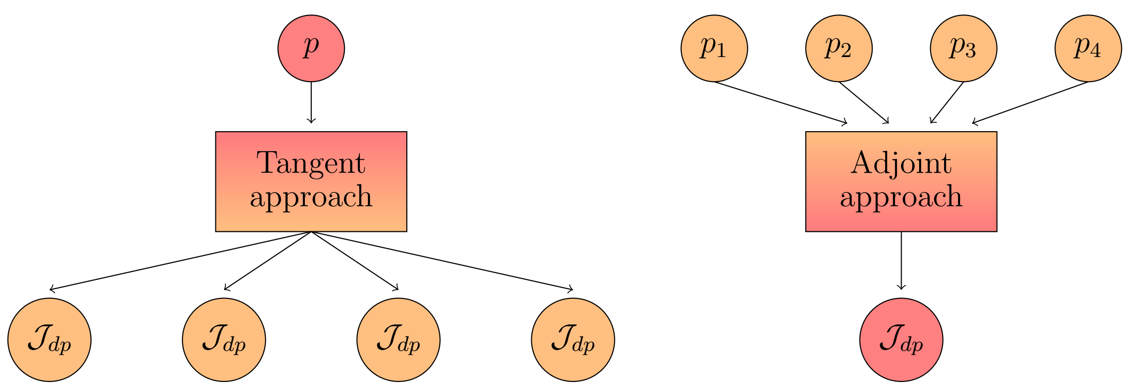

Figure 1.2: Tangent (left) and adjoint (right) sensitivity analysis methods conceptual schemes. The p represents an input parameter whilst the $\mathcal{J}_{dp}$ represents the gradient of the observable of interest with respect to a perturbation in parameter p.

In figure 1.2 we have indicated with p the generic design parameter whilst with $\mathcal{J}_{dp}$ the gradient of the observable of interest with respect to the parameter (or parameters if an adjoint approach is considered). The strength of adjoint methods lies in the fact that instead of computing the gradient explicitly, the adjoint methods allow computing gradients with the same cost regardless of the number of parameters perturbations.

1.2 Limitations of Current Approaches

The aforementioned examples are just some of currently sensitivity analysis applications coupled to steady simulations. However, trends in computer resources suggest that scale-resolving, unsteady turbulent flow simulations will take in the next decades an increasingly larger role in the engineering design cycle. This will dramatically improve prediction ability of unsteady multi-scale flow phenomena, potentially leading to more cost-effective greener design.
The need to develop a new method for sensitivity analysis lies in the fact that, applied to an unsteady simulation, classical sensitivity algorithms fail. When the equation is marched in time, in order to obtain the convergence of the long-time statistics of the observable of interest, the sensitivity does not converge but rather grows exponentially in time [4], resulting in non-physical gradients. Unfortunately, this characteristic is in conflict with the modern trend in computing resources and in industrial demands. In fact, obtaining an optimized solution of the problem relatively quickly is a crucial factor for the industries of the future. However, these insights on possible developments in the industry would be in vain if it is not possible to couple unsteady turbulent flow simulations with an effective sensitivity analysis method to reduce the computational cost of the optimization process.
More precisely, the current generation of tangent and adjoint sensitivity methods are completely ineffective when coupled to unsteady turbulent flow simulation due to the chaotic nature of the turbulence. Chaotic systems, by definition, present high sensitivity to initial conditions and evaluating effect of the parameter perturbations is anything but a trivial process. For instance, this phenomenon is clearly visible in figure 1.3a, where a small finite perturbation (10^{-4}) in the initial condition of the chaotic system arising from the Lorenz equations [5], leads to a completely different state-space evolution of the attractor after a small time relative to the time scale of the system. In panel (b) of figure 1.3 we observe that the time evolution of both third components do not show the difference in initial conditions up to a threshold value (this is not absolute value but specific for parameters and system considered). In panel (c) it is depicted how the distance between the reference and perturbed trajectories evolves over time. The distance, starting from the perturbation introduced in the problem, evolves exponentially and saturates when the attractor is completely developed. We will discuss this particular in depth in next chapters. Summing up, this breakdown is a serious limitation because many aerospace applications involve physical phenomena that exhibit chaotic dynamics. For this reason, it is necessary to develop new approaches to take a step forward in sensitivity analysis field and in unsteady turbulent flow simulations

comprehension.

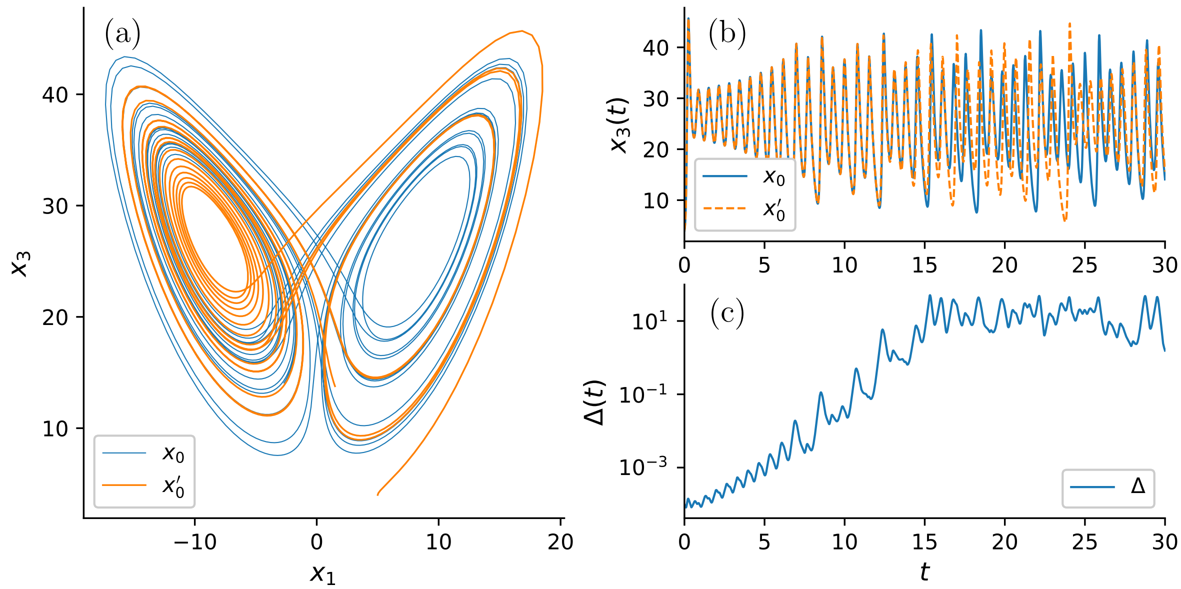

Figure 1.3: Lorenz attractor for a perturbation $\delta x_3 = 10^{-4}$ (a), time-evolution of third components (b) and time evolution of the distance between the trajectories (c). The simulation was performed using 4^{th} order Runge-Kutta scheme with initial condition $\mathbf{x_0} = (1, 1, 1)$, $\mathbf{x_0} = (1, 1, 1 + 10^{-4})$ and $(\sigma, \beta, \rho) = (10, 8/3, 28)$ over $T = 40$ time units.

Thus, given that most of the major engineering observables of interest, e.g. the drag of aerodynamic bodies, in chaotic turbulent flows are long-time averaged quantities and given the problem of not being able to carry out long-term integration of sensitivity analyses on them, the frontiers of fluid mechanics have been pushed more and more to find accurate and computationally cheap methods to obtain gradient informations.

1.3 Recent Advances

Over the years, different methods have been proposed, both from the standpoint of the formulation and the computational cost. However, as discussed by *Lea et al* in Ref. [4], all the proposed adjoint methods are revealed (some less than others) wrong. Recently, to overcome this problem and evaluate the sensitivity of an observable of interest with respect to parameter perturbations, has been proposed to use the *Shadowing Lemma* [6]. The starting point of the development of this kind of algorithms has been given by *Bowen* in Ref. [6] more then 40 years ago, but only in the last years this property was used in sensitivity analysis applications. This lemma guarantees that every pseudo-orbit stays uniformly close to a true trajectory (with slightly altered initial position). In other words, a pseudo-trajectory is "shadowed" by a true one. This Lemma is better known in the computational sciences community for its use in justifying finite-precision calculations of trajectories affected by round-off error. It is important to underline that the lemma assumes hyperbolic dynamic. In the context of sensitivity analysis of dynamical systems, the Shadowing Lemma can be used to show the existence of a trajectory of the perturbed system that starts at a different initial condition and remains uniformly close in time to the trajectory of the unperturbed system, as presented in figure 1.4. In particular, we depict a trajectory and its shadow of a generic state-space evolution of the Lorenz attractor (see equations 6.6) and of the time-evolution of them third components, in panel (a) and (b), respectively.

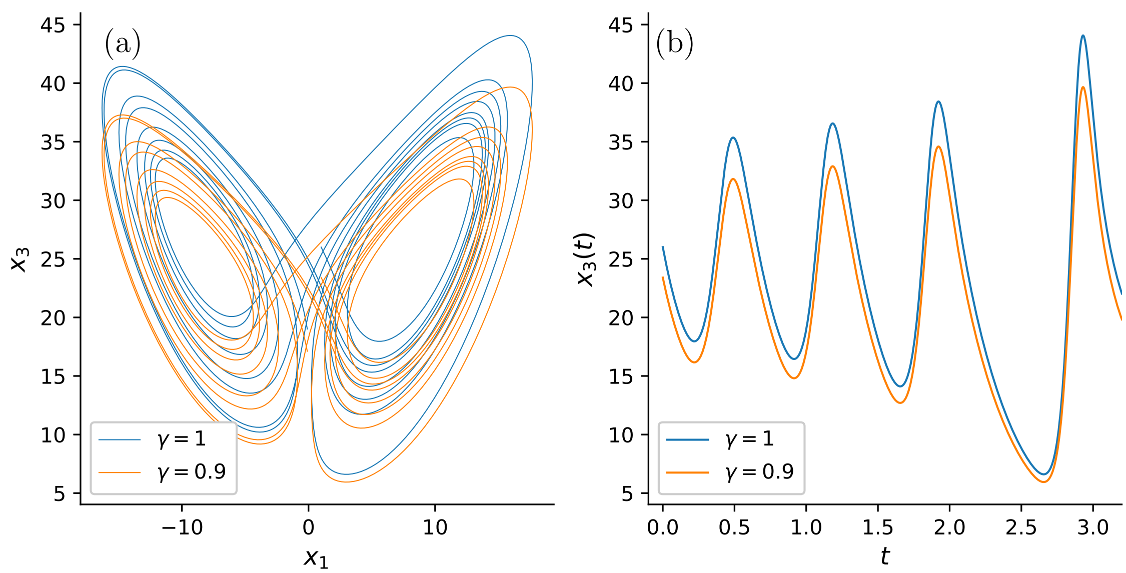

Figure 1.4: Shadowing for chaotic trajectories of the Lorenz equations for the perturbation parameter γ, defined in section 6, equation (6.6). The shadow trajectory ($\mathbf{x}_0'$), solution of the perturbed equations at $\gamma = 0.9$, stays uniformly close in time to the reference trajectory $\mathbf{x}_0$, solution of the unperturbed equations at $\gamma = 1$.

Since the two trajectories remain uniformly close to each other, the linearisation concerning the time evolution of the distance between the trajectories holds throughout and accurate gradients can be obtained. This is important because conventional sensitivity methods based on a linearised problem fail to obtain a gradient that reflect the actual behaviour of the problem. The main cause of this error are the growing modes

that are not considered during linearisation process.
Although the *Shadowing Lemma* guarantees the existence of the shadowing direction, it does not suggest practical algorithms to determine it. If for periodic trajectories the shadowing direction is also periodic in time, and thus it is possible to derive periodic boundary conditions in time for the sensitivity equations, for chaotic trajectories, however, it currently appears unlikely that an efficient strategy exists. In fact, so far, it is impossible to provide exact initial/boundary conditions and ones need to rely on approximations.
Among the first work which exploited the shadowing-direction in the sensitivity analysis field, there is Ref. [7] presented by *Wang and Hu* and named Least Square Shadowing (LSS). The peculiarity of this method is that it determines the solution of the sensitivity analysis problem with the least square average norm over the time span T. An another remarkable step in recent developments is the one presented by *Lasagna* in Ref. [8], where a well-behaved sensitivity technique is formulated on unstable time periodic trajectories. All these contributions were important for understanding the problem and allowed us to develop the new method here proposed.

The project presents a novel shadowing-based algorithm, which an alternative approximation to find the shadowing direction. The key idea of the present method is to enforce periodic boundary condition in time to the sensitivity equations, leading to the name Periodic Shadowing.

1.4 Organisation of the Thesis

The present work is subdivided as follows. Chapter 2 reviews the past work on approaches used to perform sensitivity analysis on chaotic systems. In particular, it focuses on the merits and limitations of the various methods developed. Chapter 3 describes the aim and the objectives that will be accomplished. In chapter 4 the tangent and the adjoint Periodic Shadowing methods will be presented from the conceptual standpoint and they will be analytically derived. Chapter 5 describes the numerical algorithm used to solve the boundary value problems formulated in chapter 4. Chapter 6 focuses on describing two low-dimensional chaotic systems arising from natural convection problems which will be used as application models. The results of the sensitivity analysis are presented in chapter 7 whilst in chapter 8 we outline conclusions and list few outstanding issues for future developments.

1.5 Programming Language

All results and graphs presented in this project were obtained using the Julia programming language. Julia is a high-level general-purpose dynamic programming language that was originally designed to address the needs of high-performance numerical analysis. Julia provides ease and expressiveness for high-level numerical computing, in the same way as languages such as R, MATLAB, and Python. Since the 2012 launch, the Julia community has grown, with over 2 million downloads as of August 2018.

Figure 1.5: Official Julia logo.

Chapter 2

Literature Review

The purpose of this chapter is to present the best contributions provided to the development of modern sensitivity analysis methods. The digression will follow a chronological order, initially giving an overview about the state of the art of sentivity analysis perfomered on chaotic systems and then focusing on recent developments obtained by exploiting the Shadowing Lemma [6].

2.1 Historiography of the methods developed

Since the great advancement of new computing technologies and the rapid development of new technique, numerical simulation become an indispensable tool for aerodynamic optimizations. However, it is still very time-consuming to obtain large amounts of data when high-fidelity methods are used for flow solutions. Nowadays, to decrease this computational cost, numerical simulations such as RANS-based methods are coupled to sensitivity analysis to exploit the gradient information. As stated previously, the limitations of this pairing lie in the fact that currently sensitivity analysis methods are no applicable to unsteady turbulent flow simulations due to the chaotic behaviour of turbulence. Before going into modern approaches it is necessary to present what were the most important contributions provided on this topic.

2.1.1 Direct Method and Chaotic Systems

The Direct Method (DM) was the first approach that managed to obtain practical applicability for the sensitivity analysis of chaotic systems. This approach has been demonstrated in details by *Lea et al.* in Ref. [4] for the chaotic system arising from Lorenz equations [5] (see chapter 6, equations (6.1)). In particular, the sensitivity of an observable of interest with respect to a physical parameter related to the temperature has been demonstrated convergent to the true (or so-called macroscopic) sensitivity of the problem if the integration time is very large relative to the time scale of the system. In figure 2.1 the averaged sensitivity as a function of the parameter is presented for different integration time.

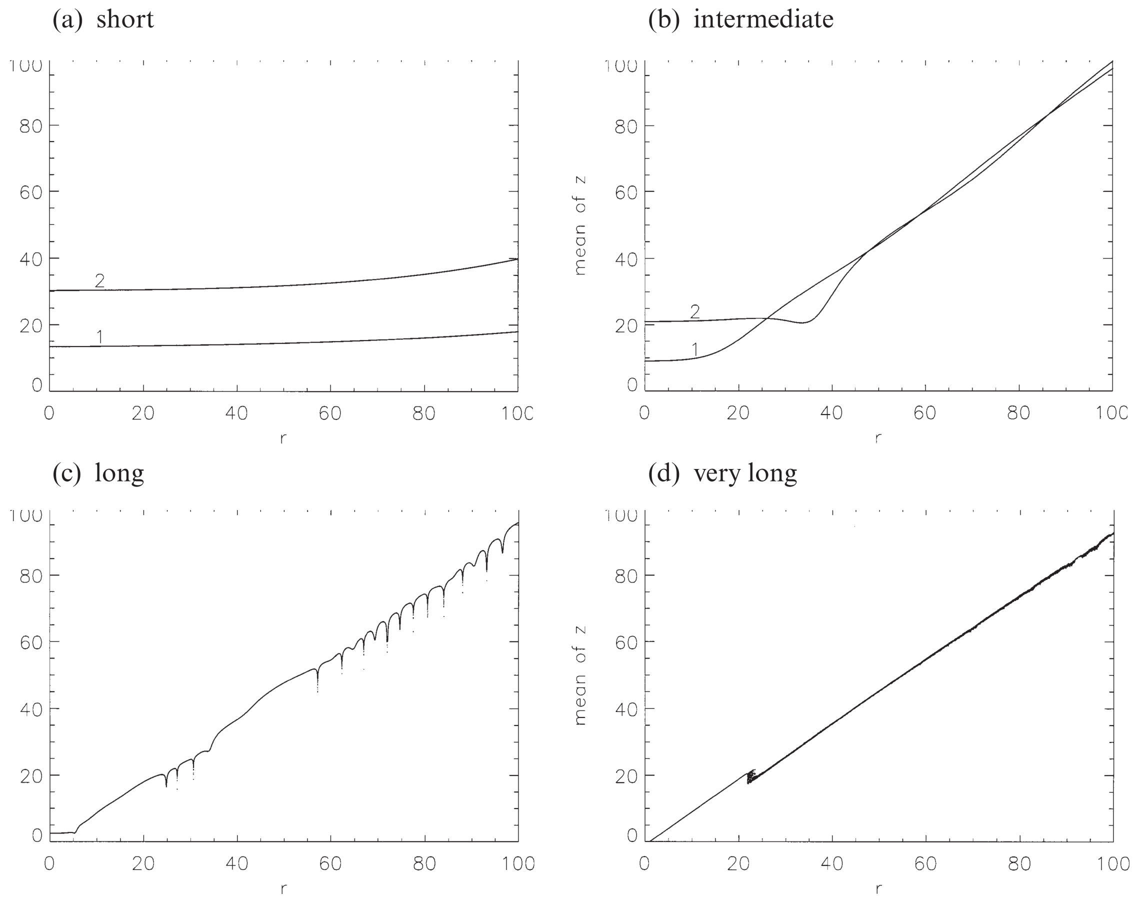

Figure 2.1: Averaged sensitivity of an observable of interest related to the heat transfer with respect to physical parameter r evaluated for increasing time span. Descriptions placed at the top left of each plot refer to the integration time considered: $T = [0.1, 0.44, 2.26, 131.36]$ for short, intermediate, long and very long time span, respectively. Values of z are evaluated at intervals of $r = 0.005$. [4].

The time spans considered are been chosen with a very specific purpose. In particular, the short integration only traverses part of an orbit around the Lorenz attractor, the intermediate integration completes a full orbit, the long integration completes several such orbits and the very long integrations travel around the attractor $\mathcal{O}(100)$ times [4]. Considering panel (d) of figure 2.1, we observe the nearly linear dependence of averaged z on r with a slope of around 0.96, which is the macroscopic sensitivity.

Limitations of Direct Approaches

The direct method, also called forward or tangent, has shown good results in the application to a low-dimensional chaotic system such as the ones arising from Lorenz equations. However, from a practical point of view, evaluate the sensitivity with respect to only one design variable it is a countercurrent approach if we consider the contemporary industrial progress. In fact, if the sensitivity of many parameters is required, the direct method becomes prohibitively from the computational cost since the sensitivity must be evaluated separately for each design variable. It is foreseeable that this problem, coupled to the ever increasing engineering requirements, became the bottleneck for a multi-parametric sensitivity analysis: for this reason, it is preferable to go through the adjoint methods.

2.1.2 Adjoint Sensitivity Methods

Instead of computing the gradient explicitly, the adjoint methods allow computing gradients with the same cost regardless of the number of parameters perturbations. Actually, this approach has not been invented recently and specifically for sensitivity analysis: in fact, the first one who used this term was *Lagrange* in 1763 in his "*Mélanges de Philosophie et de Mathématique*", who used "équation adjointe" to describe a method to lower the order of a general linear ordinary differential equation and applied it to problems as diverse as fluid motion, vibrating strings and the orbit of planets [9]. Certainly, over the years the applications of this algorithm have been multiple, starting from a general approach to perturbation theory in neutronics [10] passing to electromagnetism [11] and geophysical system [4], but all of them had the common goal of performing sensitivity analysis. The adjoint formulation is useful when is seeking to obtain the sensitivity of few outputs of interest of a given system for a wide range of design and control variables. In fact, the typical optimization problem has a single objective function (possibly combining multiple objectives through suitable weights) that has to be minimized or maximized with respect to a large number, or even a continuum, of input variables [9]. This approach has also changed the point of view of sensitivity analysis since it allows to examine an ensemble of perturbations.

Outcomes and Limitations of Classical Adjoint Methods on Chaotic Systems

Important findings in the application of the adjoint method to chaotic systems are presented in the aforementioned Ref. [4]. In that work, the authors compare the sensitivity results evaluated through the direct method and adjoint formulation. The chaotic system arising from Lorenz equations and the sensitivity evaluated for an observable of interest with respect to the temperature are once more chosen. The adjoint method, applied for four increasing integration intervals, provides results that are highly unrepresentative of how a perturbation in design variable really affects the sensitivity. Results are presented in figure 2.2.

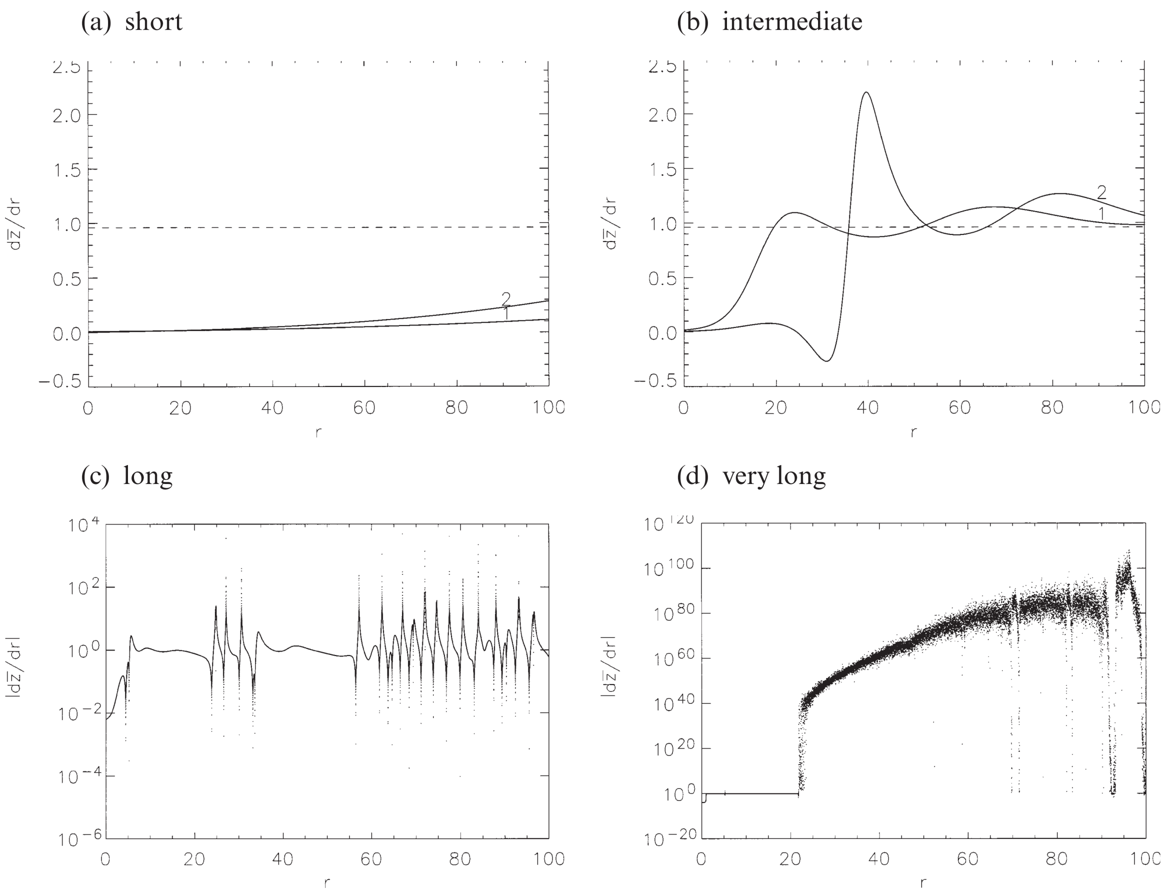

Figure 2.2: Sensitivity of an observable of interest related to the heat transfer with respect to the physical parameter r evaluated for increasing time span. Descriptions placed at the top left of each plot refer to the integration time considered: $T = [0.1, 0.44, 2.26, 131.36]$ for short, intermediate, long and very long time span, respectively. The dashed line, in (a) and (b), shows an estimate of the macroscopic climate sensitivity, equals to 0.96 [4].

The short integrations in panel (a), (b) produce a stable (i.e. only weakly dependent on initial condition) but incorrect estimate of the sensitivity, exactly as in the direct method. The long integration time in panel (c) shows gradients that peak near 10^4, completely wrong if compared to ≈ 0.96 value of the true sensitivity evaluated with the direct method. The results appear strongly dependent on initial conditions and this

behaviour is amplified considering ever larger integration times as in panel (d).
Several remedies have been proposed to apply an adjoint approach on chaotic systems, such as the ensemble average methods presented in Refs. [12, 13, 14, 15]. In particular, one of the first techniques developed to reduce the errors is to average results from an ensemble of intermediate length integration as proposed in Ref. [14]. However, the length of the integration segment must be chosen carefully, and in practical simulation even several segments are tested for seek stable results. Subsequently, *Lea et al.* in Ref. [15] suggest that averaging an ensemble of adjoint gradients (over N samples) might provide a better estimate of the sensitivity. Supported by the evidence, the authors proposed that adjoint sensitivity consists of two components: an underlying "parametric sensitivity", and a random state space dependent sensitivity. By ensemble averaging it is possible to remove the state-space dependent part and reveal the sensitivity [15]. In fact, it is proven in Ref. [12] that using $N = 299$ samples produced a sensitivity accurate to about 10%, whereas a single adjoint calculation for a time-average of duration 299 time units produced an estimate too large by 100 orders of magnitude (as presented in figure 2.2d). However, to obtain 1% accuracy $N \approx 6.4 \times 10^{14}$ samples are required. In figure 2.3, taken from the discussion section of Ref. [15], the convergence between the median of $N = 10000$ samples and the macroscopic sensitivity for the Lorenz system is presented.

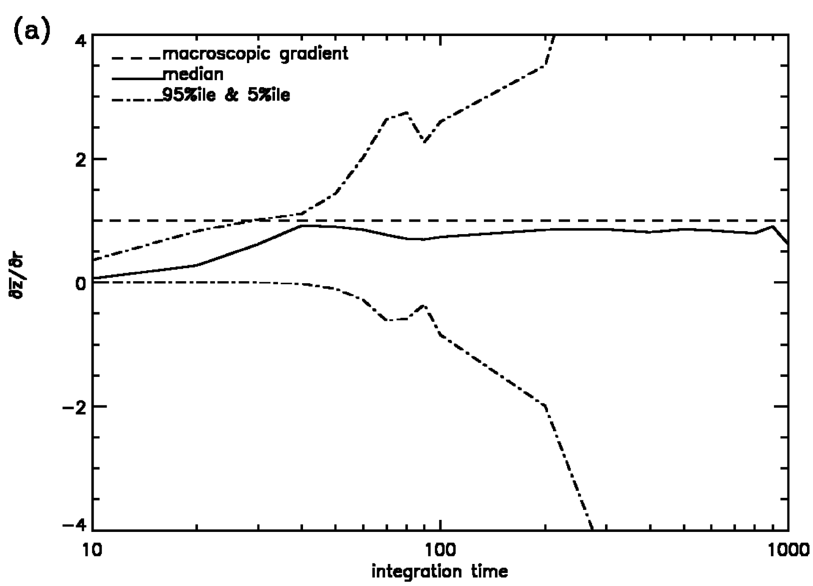

Figure 2.3: An ensemble-adjoint plot for the Lorenz (1963) system. This plot shows that the median (solid line) of the ensemble-adjoint converges to the macroscopic climate sensitivity (dashed line). The dash-dot lines represent the 95 and 5 percentile, respectively. [15].

In a subsequent work [12], *Eyink et al.* proposed a further ensemble-adjoint approach which has a convergence rate essentially the same of the one aforementioned. However, the importance of this work lies in the fact that they focused the attention on the statistical distribution of the sensitivity. In fact, the authors showed that both methods are very slowly converging in the number of samples and therefore a very large ensemble for even modest accuracy improvement in the sensitivity is required. This occurs because the probability distributions of the adjoint gradient presents a Lévy-type power-law tail, as it can be observed in figure 2.4, and thus the central limit theorem breaks down. In this way, the error bars in the ensemble-average gradient are not governed by the standard central limit theorem and are much more slowly decreasing than would be expected.

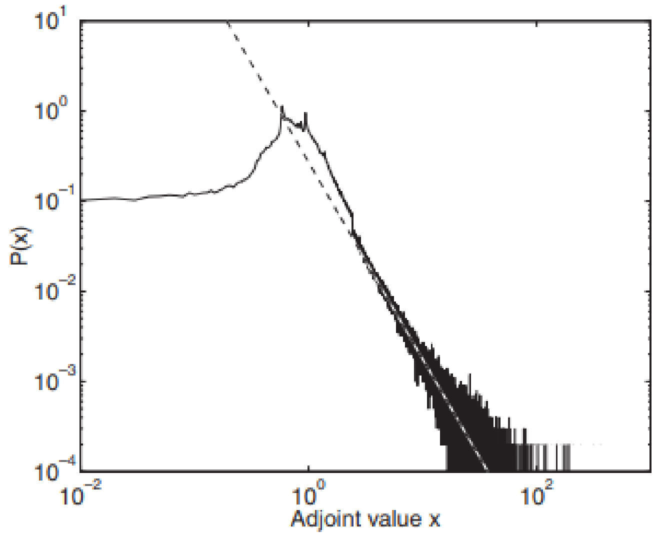

Figure 2.4: Histogram of the adjoint gradient values for $N = 10^6$ samples. The dashed line is the best power-law fit to the tail in loglog plot $\approx 0.2718x^{-(1+1.185)}$ [12].

The physical interpretation can be traced back to the fact that certain rare dynamical trajectories of the chaotic system under consideration, which pass very close to the unstable fixed point at the origin in phase space, present an enormous adjoint gradient. Another approaches to solve the problem related to the use of an adjoint formulation on chaotic systems are based on the analysis of the invariant probability density function and its adjoint, which are presented, for instance, in Refs. [16, 17]. However, both method suffers from the curse of dimensionality, in fact, the computational costs grow explosively with the increase of the attractor dimensions.

2.2 Shadowing-based Methods

In recent years, major advances and more refined solutions have been obtained, i.e. in Refs. [18, 19], by exploiting the so-called Shadowing Lemma, a theoretical result proved by *Bowen* in Ref. [6], that exclusively applies to systems with hyperbolic dynamics. The subject of Shadowing Lemma concerns the existence of true orbits near pseudo-orbits, where the latter are defined as orbits that arise when a stochastic perturbation is applied to the system or when a round-off error is present in the numerical computation [20]. The property asserts that the original and the pseudo-orbit (it shadow), initially spaced due to small parameters perturbation, remains uniformly close for the whole integration time if the system under consideration is uniformly hyperbolic.

2.2.1 Shadowing Lemma and Sensitivity Analysis

Transporting this idea into the context of the sensitivity analysis for chaotic systems, the Shadowing Lemma can be used to show the existence of a trajectory of the perturbed system that starts at a different initial condition and remains uniformly close in time to the trajectory of the unperturbed system [18, 21]. This concept is depicted in figure 2.5.

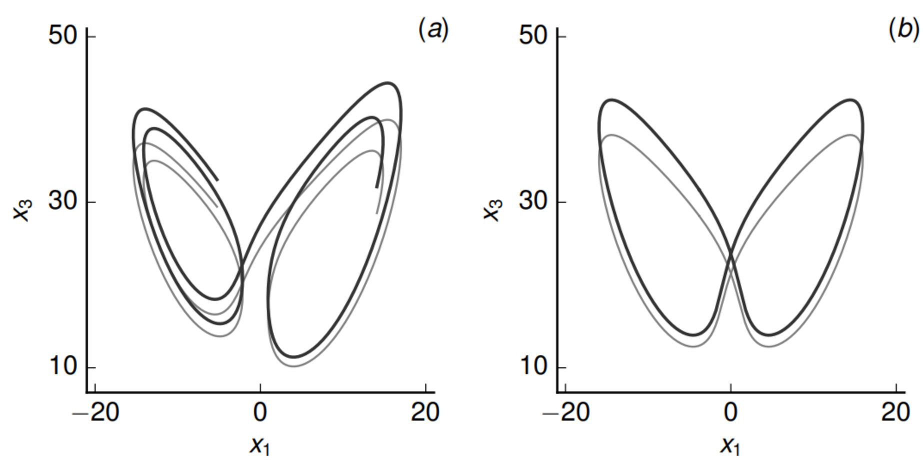

Figure 2.5: Shadowing for chaotic (a) and periodic (b) trajectories of the Lorenz equations. The shadow trajectory stays uniformly close in time to the reference trajectory [21].

Since the two trajectories remain uniformly close to each other, the linearisation concerning the time evolution of the distance between the trajectories holds throughout and accurate gradients can be obtained. The special case depicted in figure 2.5b represents the shadowing direction applied to a period trajectory. The use of this theory allowed to observe that if the pseudo-orbit is periodic, then the true orbit which shadows it is unique and periodic with the same period [22]. This intuition was at the base of Ref. [8] where the adjoint problem is formulated on time periodic rather than on open trajectories. Coupling periodic boundary conditions to the sensitivity equation, the solution remains bounded over time span considered and does not exhibit the typical unbounded exponential growth observed (remember figure 2.2) on open unstable trajectories. Applying such constraints allow passing from an ill-conditioned initial values problem to a well-conditioned boundary value problem.

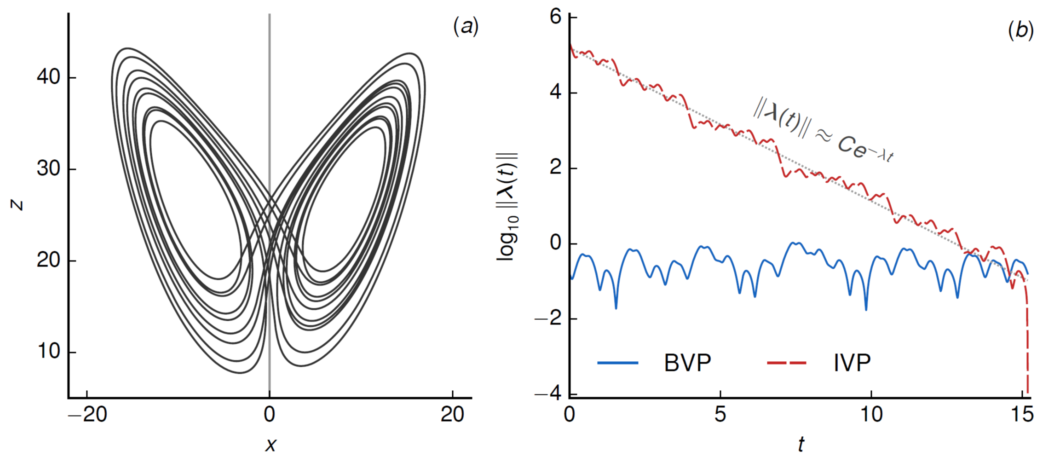

Figure 2.6: Panel (a): The $n = 20$ Unstable Periodic Orbit (UPO) ($T \approx 15.1827$) of the Lorenz attractor used as reference trajectory to solve the adjoint boundary value problem (BVP) and initial value problem (IVP) presented in Ref. [8]. Panel (b) represents the norm of the adjoint variables from solution of these two problems.

In figure 2.6a the long UPO ($T \approx 15.1827$) is depicted whilst in panel (b) the time histories of the norm of the adjoint variables from the solution of BVPs and IVPs are reported. In this way the exact sensitivity of period averaged statistics with respect to problem parameters can be evaluated. In fact, applying this method on the chaotic system arising from the Lorenz equations, the sensitivity with respect to the physical parameter related to the temperature results ≈ 1.01847, representing physically meaningful value. Applying the analogous concept of determining the exact initial (or boundary) conditions, however, is unlikely if chaotic systems are considered, and thus, it has been necessary to develop methods based on approximations. *Wang* in Ref. [18] proposed to divide the perturbed coordinate system into stable, neutral and unstable components, each one corresponding to one of Lyapunov's exponents, and solve the sensitivity equations forward/backward, respectively. This approach, computationally expensive, is then superseded by the same author through a new method [7] called Least Square Shadowing (LSS).

2.2.2 Least Square Shadowing

This method approximates the unknown shadowing direction by determining the solution of the sensitivity equations with the least square average norm over the time span T. The minimisation ensures that exponentially growing modes that would highly contribute to the solution norm are effectively controlled, so that the optimal solution remains bounded, providing useful gradients [21]. In figure 2.7 the resulting sensitivity of the Lorenz system computed with Least Square Shadowing is depicted.

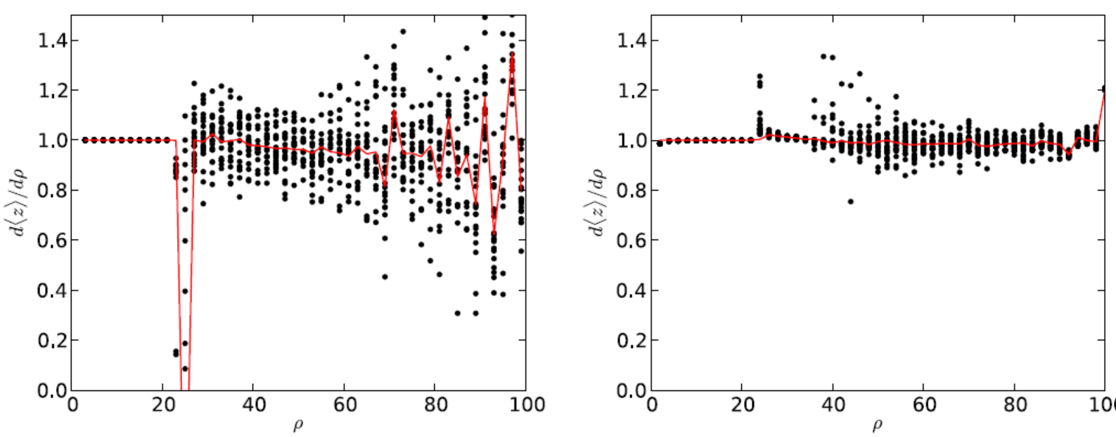

Figure 2.7: Finite difference ($\Delta\rho = 2$) approximation of the gradient using (left) and the sensitivity evaluated with Least Square Shadowing (right) [7].

Since the results obtained with LSS and presented in panel (b) are very similar to ones obtained by finite difference approximation of the gradient, panel (a), we can assert the good response of the method on this low-dimensional chaotic system. Recently, an improvement of the method using multiple-shooting strategies and suitable for high-dimensional systems has been presented in Ref. [19]. As we can understand, however, this method modifies the nature of the problem, switching to an optimization problem [21].

2.3 Periodic Shadowing

By combining all the contributions presented up to now, it has been possible to develop a new, potentially simpler, method for performing sensitivity analysis on chaotic systems. The key idea of the presented method is to approximate the shadowing direction enforcing periodic boundary condition in time to the sensitivity equation. For this reason, the here presented method is called Periodic Shadowing. Rather than formulating an optimisation problem, as in the Least Square Shadow method, the Periodic Shadowing solves a boundary value problem in time from which it is possible to obtain bounded (periodic) solutions almost always, resulting in physically meaningful sensitivity. The boundary value problem requires appropriate numerical methods for the solution, such as the multiple-shooting approach implemented for this project.

2.4 Open Questions

The main purpose of developing a new shadowing-based algorithm is to provide to the scientific landscape an alternative method to perform sensitivity analysis on chaotic system. Furthermore, by observing the results and the numerical features of the Least Square Shadowing, some questions remain nowadays still open. First of all, the Least Square Shadowing formulation implies the multiplication of the left-hand side matrix of the corresponded linear system for its transpose. During the resolution process, i.e. when the matrix inversion is required, this feature, even if the condition number remains bounded with respect to classical adjoint methods, leads to compute a matrix whose condition number is the square of the previous one. The issues related to it are well-known. Another aspect that we intend to improve with this project is to increase the knowledge on the statistical quantities of the computed sensitivities. In fact, if in figure 2.7b the sensitivity is computed for only 10 different initial conditions, we will concentrate particularly on the probability distribution of more than 10^6 evaluated sensitivity. In this way, we can also understand how the number of samples affects the convergence rate of the results.

Chapter 3

Research Objectives

3.1 Research Aims and Objectives

The primary aim of this project is to verify and understand the Periodic Shadowing approach for sensitivity analysis of chaotic systems. The method will be applied to two low-dimensional chaotic system arising from natural convection problems in fluid cells bounded by two plates and heated from below. In particular, we select the chaotic system arising from Lorenz equation [5] and from the nine-states problem presented by Reiterer et al. [23]. For both systems, we select an observable of interest related to the heat transfer in the fluid cells whilst the parameter perturbation will be applied separately on two different physical design variables. In order to obtain the convergence of the long-time statistics of the observable of interest, the equations are marched until the integration time is highly longer relative to the time scale of the problem. The increasing complexity of the systems allows us to obtain sensitivity results on problems that have the same roots, but different dynamics.
The achievement is accomplished by fulfilling the following research objectives:

- consider the Lorenz system, we evaluate the sensitivity of the time averaged observable of interest with respect to a parameter describing the state evolution under a coordinate transformation. The statistical results are compared to the gradient obtained with by finite-difference approximation;

- consider the Lorenz system we evaluate the sensitivity with respect to the parameter related to the temperature difference between the plates of the fluid cells. In this way we computed statistical quantities that can be compared with the Least Square Shadowing sensitivity results and finite-difference approximation of the gradient;

- analyse the nine-states model representing the three-dimensional cells in dissipative Rayleigh-Bénard convection and investigate the response of the statistical quantities computed with respect to a coordinate transformation;
- replicate the previous analysis evaluating the sensitivity of the same observable of interest with respect to the physical parameter related to the temperature and compare the results with the finite-difference approximation of the gradient.

It is important to underline that all the aforementioned analyses were subordinated to the comprehension of the probability distribution of sensitivity results. In particular we focused on the decay law of the right tails of the full distribution, in order to verify which statistical quantities is the most representative for each systems.

Chapter 4

Sensitivity Analysis of Chaotic Systems

Conventional sensitivity analysis methods can lead to incorrect solutions if they are applied to chaotic systems like those presented in chapter 6. This occurs because the property of being dynamically unstable leads to an initial value problem inherently ill-conditioned. This manifests itself in a great sensitivity to parameters perturbation, especially when the time interval is expanded, i.e. when is required the sensitivity of long-term statistics. The computed sensitivity grows exponentially over time and leads to a non-physical gradient that does not coincide in the limit with the required sensitivity of the time average. The chapter will be divided as follows. In the first section, the conventional methods will be presented whilst in the remaining ones the tangent and adjoint formulations of the Periodic Shadowing will be described.

4.1 Conventional Methods

The choice on which approach, tangent or adjoint, to be used for the analysis of sensitivity depends on how the problem arises and what our goals are. As stated previously, the tangent (or direct) method finds its practical response in problems where the number of observables of interest is higher than design and control variables. If the situation were reversed, the direct method would become extremely computationally expensive and therefore the adjoint method is preferable.

4.1.1 Tangent Method

Consider the following dynamical system as the interpretation of a time dependent simulation

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{p}) \tag{4.1}$$

where $\mathbf{x}(t) \in \mathbb{R}^N$, $\mathbf{f}(\mathbf{x}(t), \mathbf{p}) : \mathbb{R}^N \times \mathbb{R}$ is a vector function of $\mathbf{x}(t)$ and $\mathbf{p}$ is the set of design and control variables (for simplicity in the following we will consider single scalar parameter p). To take a practical example, for a computational fluid dynamics (CFD) simulation, the $\mathbf{x}(t)$ represents the vector containing the conserved quantities and $\mathbf{p}$ could include, for instance, geometry or flow properties [19]. The trajectory originated at $\mathbf{x}_0$ is denoted as $\mathbf{x}(t; \mathbf{x_0}, p)$. We assume that the vector function on the right-hand side is sufficiently smooth with respect to its arguments, so that existence and uniqueness of solutions are formally guaranteed. The observables of interest, represented by the scalar-valued function $J(\mathbf{x}(t), p)$, is a function of the initial condition, parameters value and model state $\mathbf{x}(t)$, which itself evolves in time in a way dependent on the parameter p. Generally, since such simulations are performed with time-varying forces, e.g. in turbulent flows applications, the goals of the optimisation processes are focused on the finite-time averaged observables of interest

$$\mathcal{J}^T(\mathbf{x_0}, p) = \frac{1}{T}\int_0^T J(\mathbf{x}(t; \mathbf{x_0}, p), p)dt \tag{4.2}$$

The dependence on the initial condition can be avoided assuming ergodicity, so

$$\mathcal{J}^\infty(p) = \lim_{T\to\infty} \mathcal{J}^T(\mathbf{x_0}, p) \tag{4.3}$$

The sensitivity analysis aims at the knowledge of how observables of interest modify due to a parameter perturbation $\delta p = p' - p$. For small perturbations around some reference p, this information, encoded by the gradient $\mathcal{J}^\infty_{dp}(p)$, is formally defined as

$$\begin{aligned}\mathcal{J}^\infty_{dp}(p) &= \lim_{\delta p\to 0} \frac{1}{\delta p}\left[\mathcal{J}^\infty(p') - \mathcal{J}^\infty(p)\right] \\ &= \lim_{\delta p\to 0} \frac{1}{\delta p}\left[\lim_{T\to\infty} \mathcal{J}^T(\mathbf{x_0}, p') - \lim_{T\to\infty} \mathcal{J}^T(\mathbf{x_0}, p)\right]\end{aligned} \tag{4.4}$$

Since the superscript was used to indicate the perturbed parameter, the related dynamic system is

$$\dot{\mathbf{x}}'(t) = \mathbf{f}(\mathbf{x}'(t), p') \tag{4.5}$$

Terming the difference between the two trajectories as

$$\delta\mathbf{x}(t) = \mathbf{x}'(t) - \mathbf{x}(t) \tag{4.6}$$

and considering a small perturbation δp, we can write the following linearisation

$$\delta \mathbf{x}(t) = \mathbf{y}(t)\delta p + \mathcal{O}(\delta p^2) \tag{4.7}$$

where $\mathbf{y}(t)$ is considered the sensitivity of the system. Linearising the observable of interest around the reference trajectory we obtain

$$\begin{aligned} J(\mathbf{x}'(t), p') &= J(\mathbf{x}(t), p) + J_{\partial \mathbf{x}}(\mathbf{x}(t), p)(t)\delta \mathbf{x}(t) + J_{\partial p}(\mathbf{x}(t), p)\delta p + \mathcal{O}(\delta p^2) \\ &\overset{(4.7)}{=} J(\mathbf{x}(t), p) + J_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t)\delta p + J_{\partial p}(\mathbf{x}(t), p)\delta p + \mathcal{O}(\delta p^2) \end{aligned} \tag{4.8}$$

Exploiting (4.2), it is possible to rewrite (4.4) as

$$\begin{aligned} \mathcal{J}^{\infty}_{dp}(p) &= \lim_{\delta p \to 0} \lim_{T \to \infty} \frac{1}{\delta p} \left[\frac{1}{T} \int_0^T J(\mathbf{x}'(t; \mathbf{x_0}, p'), p') dt - \frac{1}{T} \int_0^T J(\mathbf{x}(t; \mathbf{x_0}, p), p) dt \right] \\ &\overset{(4.8)}{=} \lim_{\delta p \to 0} \lim_{T \to \infty} \frac{1}{\delta p} \frac{1}{T} \left\{ \int_0^T \left[J_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t)\delta p + J_{\partial p}(\mathbf{x}(t), p)\delta p \right] dt \right\} \\ &= \lim_{T \to \infty} \frac{1}{T} \left\{ \int_0^T \left[J_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) + J_{\partial p}(\mathbf{x}(t), p) \right] dt \right\} \end{aligned} \tag{4.9}$$

Performing the differentiation of (4.6) with respect to the time, we get the equation for the evolution of $\mathbf{y}(t)$

$$\begin{aligned} \delta \dot{\mathbf{x}} &= \dot{\mathbf{x}}'(t) - \dot{\mathbf{x}}(t) \\ &\underset{(4.1)}{\overset{(4.5)}{=}} \mathbf{f}(\mathbf{x}'(t), p') - \mathbf{f}(\mathbf{x}(t), p) \\ &= \dot{\mathbf{y}}(t)\delta p \end{aligned} \tag{4.10}$$

Linearising the vector field around the reference trajectory leads to

$$\begin{aligned} \mathbf{f}(\mathbf{x}'(t), p') &= \mathbf{f}(\mathbf{x}(t), p) + \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p)\delta \mathbf{x}(t) + \mathbf{f}_{\partial p}(\mathbf{x}(t), p)\delta p + \mathcal{O}(\delta p^2) \\ &\overset{(4.7)}{=} \mathbf{f}(\mathbf{x}(t), p) + \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t)\delta p + \mathbf{f}_{\partial p}(\mathbf{x}(t), p)\delta p + \mathcal{O}(\delta p^2) \end{aligned} \tag{4.11}$$

and substituting it in (4.10) we obtain

$$\dot{\mathbf{y}}(t)\delta p = \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t)\delta p + \mathbf{f}_{\partial p}(\mathbf{x}(t), p)\delta p \tag{4.12}$$

where $\mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \in \mathbb{R}^{N \times N}$ is the system Jacobian containing the partial derivatives of the vector field with respect to the state space coordinates whilst $\mathbf{f}_{\partial p}(\mathbf{x}(t), p) \in \mathbb{R}^N$ is a vector containing the partial derivatives of the vector field with respect to the parameters. Coupling (4.12) with the initial condition $\mathbf{y}(0) = \mathbf{0}$, which is the linearisation of $\mathbf{x}_0' = \mathbf{x}_0$, we obtain the initial value problem for the conventional tangent approach

$$\begin{cases} \dot{\mathbf{y}}(t) = \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) + \mathbf{f}_{\partial p}(\mathbf{x}(t), p), & t \in [0, T] \quad \text{(4.13a)} \\ \mathbf{y}(0) = 0 & \text{(4.13b)} \end{cases}$$

This choice arises from optimal control theory ideas, where the focus is typically on the effects of parameter perturbations on the future evolution of the system, starting from the same initial condition.

4.1.2 Adjoint Method

When the sensitivity with respect to several parameters is required, we rewrite the initial value problem in such a way we do not have to explicitly compute $\mathbf{y}(t)$. For this reason we introduce the adjoint variable $\mathbf{q}(t) \in \mathcal{X} \equiv \mathbb{R}^N$ and we combine it with the sensitivity equation (4.12), obtaining the identity

$$\frac{1}{T} \int_0^T \mathbf{q}^\intercal(t) \left[\dot{\mathbf{y}}(t) - \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) - \mathbf{f}_{\partial p}(\mathbf{x}(t), p) \right] dt = 0 \tag{4.14}$$

Expanding the multiplication we obtain

$$\frac{1}{T} \int_0^T \left[\mathbf{q}^\intercal(t) \dot{\mathbf{y}}(t) - \mathbf{q}^\intercal(t) \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) - \mathbf{q}^\intercal(t) \mathbf{f}_{\partial p}(\mathbf{x}(t), p) \right] dt = 0 \tag{4.15}$$

and integrating by parts the first term leads to

$$\mathbf{q}^\intercal(t) \dot{\mathbf{y}}(t) = \left[\mathbf{q}^\intercal(t) \mathbf{y}(t) \right] \Big|_0^T - \int_0^T \dot{\mathbf{q}}^\intercal(t) \mathbf{y}(t) dt \tag{4.16}$$

Substituting in (4.15) leads to

$$\frac{1}{T} \int_0^T \Big\{ \left[-\dot{\mathbf{q}}^\intercal(t) - \mathbf{q}^\intercal(t) \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \right] \cdot \mathbf{y}(t) - \left[\mathbf{q}^\intercal(t) \mathbf{f}_{\partial p}(\mathbf{x}(t), p) \right] \Big\} dt + \\ + \frac{1}{T} \left[\mathbf{q}^\intercal(t) \cdot \mathbf{y}(t) \right] \Big|_0^T = 0 \tag{4.17}$$

Since the identity to the null value is verified, adding this term to (4.9) and considering finite-time integration

$$\begin{aligned}\mathcal{J}_{dp}^{T}(p) = \frac{1}{T}\int_0^T \left[J_{\partial \mathbf{x}}(\mathbf{x}(t),p)\cdot \mathbf{y}(t) + J_{\partial p}(\mathbf{x}(t),p)\right]dt + \\ \frac{1}{T}\int_0^T \left\{\left[-\dot{\mathbf{q}}^{\intercal}(t) - \mathbf{q}^{\intercal}(t)\mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t),p)\right]\cdot \mathbf{y}(t) - \left[\mathbf{q}^{\intercal}(t)\mathbf{f}_{\partial p}(\mathbf{x}(t),p)\right]\right\}dt + \\ + \frac{1}{T}\left[\mathbf{q}^{\intercal}(t)\cdot \mathbf{y}(t)\right]\Big|_0^T\end{aligned} \tag{4.18}$$

Factoring out $\mathbf{y}(t)$

$$\begin{aligned}\mathcal{J}_{dp}^{T}(p) = \frac{1}{T}\Bigg\{\int_0^T \left[-\dot{\mathbf{q}}^{\intercal}(t) - \mathbf{q}^{\intercal}(t)\mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t),p) + J_{\partial \mathbf{x}}(\mathbf{x}(t),p)\right]\cdot \mathbf{y}(t)dt + \\ + \int_0^T \left[-\mathbf{q}^{\intercal}(t)\mathbf{f}_{\partial p}(\mathbf{x}(t),p) + J_{\partial p}(\mathbf{x}(t),p)\right]dt\Bigg\} + \\ + \frac{1}{T}\left[\mathbf{q}^{\intercal}(t)\cdot \mathbf{y}(t)\right]\Big|_0^T\end{aligned} \tag{4.19}$$

To avoid the explicit computation of $\mathbf{y}(t)$, contained in the first and last term of (4.19), for each design variable we need to select a particularly adjoint variable $\mathbf{q}(t)$ which allows obtaining

$$\dot{\mathbf{q}}^{\intercal}(t) + \mathbf{q}^{\intercal}(t)\mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t),p) = J_{\partial \mathbf{x}}(\mathbf{x}(t),p) \tag{4.20}$$

and

$$\begin{aligned}\frac{1}{T}\left[\mathbf{q}^{\intercal}(t)\cdot \mathbf{y}(t)\right]\Big|_0^T &= \frac{1}{T}\left[\mathbf{q}^{\intercal}(T)\mathbf{y}(T) - \mathbf{q}^{\intercal}(0)\mathbf{y}(0)\right] \\ &\overset{(4.13b)}{=} \frac{1}{T}\mathbf{q}^{\intercal}(T)\mathbf{y}(T)\end{aligned} \tag{4.21}$$

Thus, choosing $\mathbf{q}^{\intercal}(T) = 0$ for the adjoint variables, we obtain $\mathbf{q}^{\intercal}(t)$ solving backward in time the following system

$$\begin{cases}\dot{\mathbf{q}}^{\intercal}(t) + \mathbf{q}^{\intercal}(t)\mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t),p) = J_{\partial \mathbf{x}}(\mathbf{x}(t),p) \\ \mathbf{q}^{\intercal}(T) = 0\end{cases} \tag{4.22}$$

Thus, the sensitivity is

$$\mathcal{J}_{dp}^{T} = \frac{1}{T}\int_0^T \left[-\mathbf{q}^{\intercal}(t)\mathbf{f}_{\partial p}(\mathbf{x}(t),p) + J_{\partial p}(\mathbf{x}(t),p)\right]dt \tag{4.23}$$

4.1.3 Limitations of Conventional Methods

For a chaotic system with unstable dynamic, the initial values problems (4.13) and (4.22) are ill-conditioned. Thus, the computed sensitivity derivative (4.12),(4.23) diverges to infinity rather than converging to the sensitivity derivative of the statistical quantity [4]. In particular, these approaches fail when the observable of interest is averaged over a time span T long relative to the time scale of the problem. This occurs because a chaotic dynamical system is extremely sensitive to initial conditions perturbations causing an exponential growth of the linearised initial value problem, as we can observe in figure 4.1. In figure 4.1 we show the time evolution of the distance between

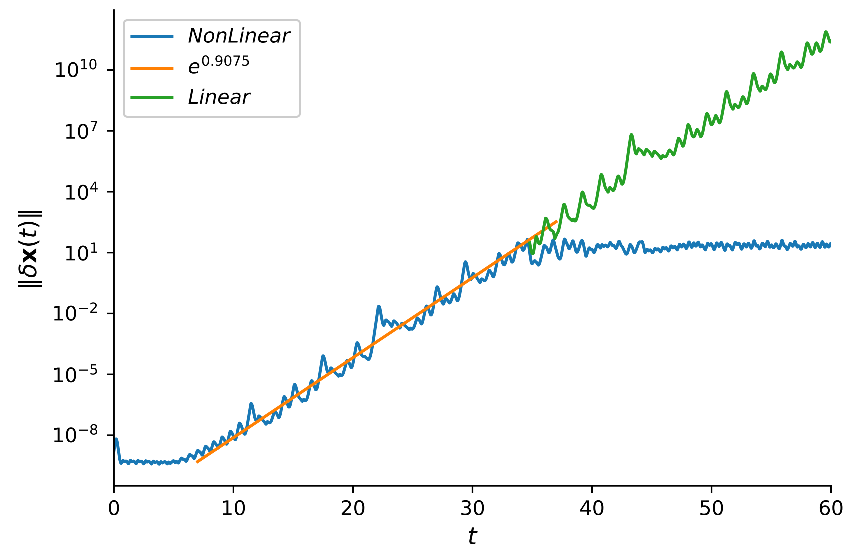

Figure 4.1: Time evolution of the distance between trajectories of Lorenz system initially spaced by 10^{-9}. The blue and green lines represent the nonlinear and linearised simulation, respectively. The regression line corresponds to the maximal Lyapunov exponent of the system.

the trajectories considering nonlinear (6.1) and linearised system. We observe that the positive slope for the nonlinear simulation only holds up for the first 35 time units and, after that, the curve levels off. This occurs because all trajectories of the Lorenz system wind up in its attractor. On the other hand, linearised simulation grows until that threshold with the same slope of the nonlinear simulation, but after that, they cannot describe the phenomenon related to the saturation of the distance. In fact, after that threshold, there is an endless increase in distance $\|\mathbf{y}(t)\|\delta p$ as $T \to \infty$. Thus, the sensitivity obtained is too large and does not represent the real behaviour of the system.

Role of Lyapunov Exponent

The growth rate of distance between the reference and perturbed trajectories is $e^{\lambda t}$, where λ is the maximal Lyapunov exponent of the system under consideration. Lyapunov exponents are characteristic quantities of dynamical systems which provide a measure of the rate of separation of infinitesimally close trajectories in phase space. Consider two points $\mathbf{x}(t)$ and $\mathbf{x}(t) + \delta\mathbf{x}(t)$ of N-dimensional chaotic system spaced by infinitesimal $\delta\mathbf{x}(t)$ at time t. Figure 4.2 shows a snippet of an evolution in the state spaces of two generic trajectories.

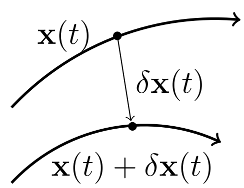

Figure 4.2: Evolution of two different trajectories initially spaced $\delta\mathbf{x}_0$

The rate of change of $\delta\mathbf{x}(t)$ over time is provided by the maximal Lyapunov exponent of the system, defined as

$$\lambda = \lim_{t\to\infty} \frac{1}{t} ln \frac{\|\delta\mathbf{x}(t)\|}{\|\delta\mathbf{x}(0)\|} \tag{4.24}$$

where $\delta\mathbf{x}(0)$ is the initial distance between the trajectories. Dynamical systems has a spectrum of Lyapunov exponents, one for each dimension of its phase space and the maximal Lyapunov exponent is responsible for the dominant behaviour of a system. Using (4.24) we can obtain the maximal Lyapunov exponent as the gradient of the distance between the trajectories plotted as a function of time, as depicted in figure 4.1. The maximal Lyapunov exponent evaluated is ≈ 0.9075, but if we average over many trajectories we can obtain the more accurate value of ≈ 0.906 [24].

4.2 Tangent Periodic Shadowing

Our aim is to obtain the gradient (4.27) using a linear method, with a tangent or adjoint formulation. Consider again the dynamic system presented in equation (4.1). As stated, the observable of interest depends on the initial condition and parameters value, and it is represented by the scalar-valued function $J(\mathbf{x}(t), p)$, which its finite-time average is

$$\mathcal{J}^T(\mathbf{x_0}, p) = \frac{1}{T} \int_0^T J(\mathbf{x}(t; \mathbf{x_0}, p), p)dt \tag{4.25}$$

The dependence on the initial condition can be avoided assuming ergodicity, so

$$\mathcal{J}^{\infty}(p) = \lim_{T\to\infty} \mathcal{J}^{T}(\mathbf{x_0}, p) \tag{4.26}$$

To understand how the observable of interest modifies due to a small parameter perturbation $\delta p = p' - p$, we need to evaluate the gradient $\mathcal{J}^{\infty}_{dp}(p)$, formally defined as

$$\begin{aligned} \mathcal{J}^{\infty}_{dp}(p) &= \lim_{\delta p\to 0} \frac{1}{\delta p} \left[\mathcal{J}^{\infty}(p') - \mathcal{J}^{\infty}(p)\right] \\ &= \lim_{\delta p\to 0} \frac{1}{\delta p} \left[\lim_{T'\to\infty} \mathcal{J}^{T'}(\mathbf{x}'_{\mathbf{0}}, p') - \lim_{T\to\infty} \mathcal{J}^{T}(\mathbf{x_0}, p)\right] \end{aligned} \tag{4.27}$$

Expression (4.27) is the basis of the familiar finite-difference gradient approximation. However, the existence of this limit, i.e. the differentiability of the infinite-time averages of a dynamical system, is a long-standing question in dynamical systems theory, but, as stated in Ref. [25] it exists if uniformly hyperbolic systems are considered.
The dynamic system characterized by the perturbed parameter is formally expressed as

$$\dot{\mathbf{x}}'(t) = \mathbf{f}(\mathbf{x}'(t), p') \tag{4.28}$$

and defined over a time span $[0, T']$ where $T' = T + \delta T$. Using different time span does not affect (4.27) and it can be advantageous for the approach we will explain in the following. For this reason, we start defining the distance between the reference and perturbed trajectories, leaning on the Linstedt-Poincaré technique by which, introducing $\omega = T'/T$ term representing how the time scales of the trajectories are related, it is possible to define

$$\delta\mathbf{x}(t) = \mathbf{x}'(t\omega) - \mathbf{x}(t) \tag{4.29}$$

This technique (see e.g. Refs. [26, 27]), allows to define equation (4.29) over $t \in [0, T]$ but actually extending the time to the entire interval $[0, T']$ on the perturbed trajectory. This means that $t \in [0, T]$ represents now the independent variable parameterising points of the perturbed system. Using different time span, relation (4.29) would not be T-periodic due to the algebraically growing modes known as secular terms introduced performing the linearisation [28]. It can be therefore understood that the choice of ω is not completely arbitrary, but should reflect the change in the relevant time scale of the system under parameter perturbations.
Considering a small δp, we can write the linearisation of the distance between the trajectories as

$$\delta\mathbf{x}(t) = \mathbf{y}(t)\delta p + \mathcal{O}(\delta p^2) \tag{4.30}$$

where $\mathbf{y}(t)$ is the sensitivity of the system.
Using this term to linearise the observable of interest around the reference trajectory leads to

$$J(\mathbf{x}'(t\omega), p') = J(\mathbf{x}(t), p) + J_{\partial\mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t)\delta p + J_{\partial p}(\mathbf{x}(t), p)\delta p + \mathcal{O}(\delta p^2) \tag{4.31}$$

with $t \in [0, T]$ as discussed in (4.29). Now it is possible to rewrite (4.27) as

$$\begin{aligned}
\mathcal{J}^{\infty}_{dp}(p) &= \lim_{\delta p \to 0} \lim_{T,T' \to \infty} \frac{1}{\delta p} \left[\frac{1}{T'} \int_0^{T'} J(\mathbf{x}'(t; \mathbf{x}'_\mathbf{0}, p'), p')dt - \frac{1}{T} \int_0^T J(\mathbf{x}(t; \mathbf{x_0}, p), p)dt \right] \\
&= \lim_{\delta p \to 0} \lim_{T \to \infty} \frac{1}{\delta p} \left[\frac{1}{T'} \int_0^{T} J(\mathbf{x}'(t\omega; \mathbf{x}'_\mathbf{0}, p'), p')dt\omega - \frac{1}{T} \int_0^T J(\mathbf{x}(t; \mathbf{x_0}, p), p)dt \right] \\
&\overset{(4.31)}{=} \lim_{\delta p \to 0} \lim_{T \to \infty} \frac{1}{\delta p} \frac{1}{T'} \omega \int_0^T \left[J_{\partial\mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t)\delta p + J_{\partial p}(\mathbf{x}(t), p)\delta p \right] dt \\
&= \lim_{T \to \infty} \frac{1}{T} \int_0^T \left[J_{\partial\mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) + J_{\partial p}(\mathbf{x}(t), p) \right] dt
\end{aligned} \tag{4.32}$$

Performing the differentiation of (4.29) with respect to the time t we get the equation for the evolution of $\mathbf{y}(t)$

$$\begin{aligned}
\delta\dot{\mathbf{x}} &= \omega\dot{\mathbf{x}}'(t\omega) - \dot{\mathbf{x}}(t) \\
&\underset{(4.28)}{\overset{(4.1)}{=}} \omega\mathbf{f}(\mathbf{x}'(t\omega), p') - \mathbf{f}(\mathbf{x}(t), p) \\
&= \dot{\mathbf{y}}(t)\delta p
\end{aligned} \tag{4.33}$$

Linearising the vector field around the reference trajectory $\mathbf{x}(t)$ like in (4.31)

$$\mathbf{f}(\mathbf{x}'(t\omega), p') = \mathbf{f}(\mathbf{x}(t), p) + \mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t)\delta p + \mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t), p)\delta p + \mathcal{O}(\delta p^2) \tag{4.34}$$

and substituting it in the evolution equation (4.33)

$$\begin{aligned}
\dot{\mathbf{y}}(t)\delta p &= \omega[\mathbf{f}(\mathbf{x}(t), p) + \mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t)\delta p + \mathbf{f}_{\partial p}(\mathbf{x}(t), p)\delta p] - \mathbf{f}(\mathbf{x}(t), p) \\
&= \omega\mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t)\delta p + \omega\mathbf{f}_{\partial p}(\mathbf{x}(t), p)\delta p + (\omega - 1)\mathbf{f}(\mathbf{x}(t), p)
\end{aligned}$$

Since to the first order $\omega = 1 + \dfrac{T_{dp}}{T}\delta p$ and dividing by δp

$$\dot{\mathbf{y}}(t) = \left(1 + \frac{T_{dp}}{T}\delta p\right) \mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) + \left(1 + \frac{T_{dp}}{T}\delta p\right) \mathbf{f}_{\partial p}(\mathbf{x}(t), p) + \frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t), p)$$

Performing the $\delta p \to 0$ limit we obtain the sensitivity equations

$$\dot{\mathbf{y}}(t) = \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) + \mathbf{f}_{\partial p}(\mathbf{x}(t), p) + \frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t), \mathbf{p}) \tag{4.35}$$

where $\mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \in \mathbb{R}^{N \times N}$ is the system Jacobian containing the partial derivatives of the vector field with respect to the state space coordinates whilst $\mathbf{f}_{\partial p}(\mathbf{x}(t), p) \in \mathbb{R}^{N}$ is a vector containing the partial derivatives of the vector field with respect to the parameters.

Note that the gradient T_{dp} is still an unknown and "arbitrary" quantity. Once obtained the equation of sensitivity we need to introduce the initial condition for it. Recalling the ergodicity assumption performed on (4.26), we are free to choose any initial condition for the perturbed system (4.28), i.e. in a linearised formulation it corresponds to select an arbitrary $\mathbf{y}(0)$. Classical sensitivity analysis methods (4.13) select the initial condition $\mathbf{y}(0) = 0$, the linearisation of $\mathbf{x}'_0 = \mathbf{x}_0$, as used for instance in (4.13b). However, it was proven that two trajectories originating at the same point separate initially at an exponential rate, and the difference saturates in a finite-time around a finite value due to global boundedness, as seen in figure 4.1. The problem that arises from the linearised formulation is that (4.12), by definition, does not model the nonlinear effects, and thus $\mathbf{y}(t)$ continues growing at an average exponential rate for all the considered time span [4]. In other words, for any finite δp, there is a finite T at which the linearisation fails and higher order terms neglected in (4.7) and (4.11) become important [17]. This growth is reflected in a non-physical exponential increase of the gradient (4.32) as T is increased [4].

In order for the linearisation to remain valid, and thus for the gradient (4.32) to converge while integrated on $T \to \infty$, the sensitivity should remain bounded. Regarding this, the shadowing lemma gives us the necessary theoretical aid, although it does not formulate algorithms to provide initial or boundary conditions that can be used to solve (4.35). In the following the required trajectory is called shadow trajectory for a finite perturbation and shadowing direction for an infinitesimal one. The peculiarity of the method developed and here proposed as Periodic Shadowing is precisely on the choice of the periodic boundary conditions. The key idea is that it is imposed the condition that the end points of the perturbed trajectory move in the same unspecified direction by the same unspecified amount, as presented by the arrows in panel (a) and (b) of figure 4.3 for aperiodic and periodic trajectory, respectively.

From a mathematical standpoint this feature is expressed as

$$\mathbf{x}'_{\mathbf{0}} - \mathbf{x}_{\mathbf{0}} = \mathbf{x}'(T'; \mathbf{x}'_{\mathbf{0}}, p') - \mathbf{x}(T; \mathbf{x}_{\mathbf{0}}, p)$$

whose linearisation with respect to the reference trajectory, obtained by dividing both

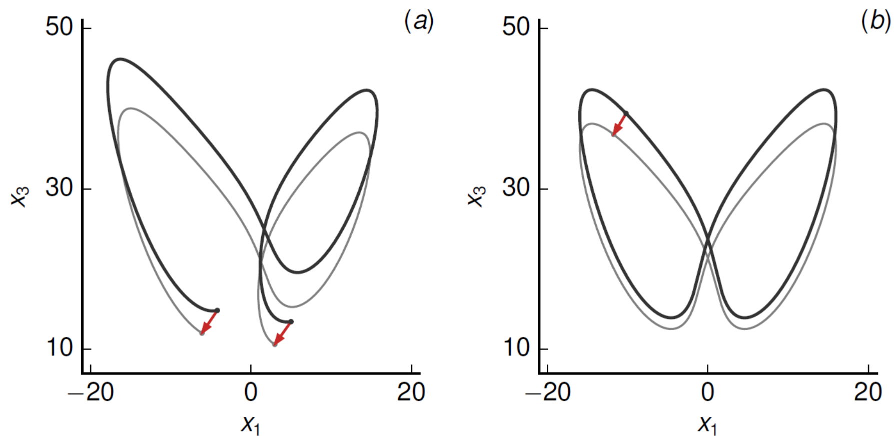

Figure 4.3: Geometry of the periodicity condition 4.36, for aperiodic (a) and periodic (b) trajectories of the Lorenz equations [21].

sides by δp and taking the $\delta p \to 0$ limit, leads to

$$\mathbf{y}(0) = \mathbf{y}(T) \tag{4.36}$$

Combining the relationship (4.35) with (4.36) the sensitivity problem is only partially defined because, unlike the conventional tangent method, there is an additional term related to the time gradient. In fact, considering only the periodic boundary conditions, there are no constraints which allow us to determine the gradient T_{dp}, which remains arbitrary. If we hypothesize to forcibly set $T_{dp} = 0$ we do not get an accurate solution because neglecting the growth of algebraic modes produces a spurious sensitivity error that does not vanish as $T \to \infty$. In the following, an approach that is frequently employed in bifurcation analysis for periodic systems [29] are presented. Rather than fixing the gradient T_{dp} a priori, we impose that the solution of the sensitivity equations (4.35) satisfies the additional orthogonality condition

$$\mathbf{f}(\mathbf{x}(0), p)^{\intercal} \cdot \mathbf{y}(0) = 0 \tag{4.37}$$

at the trajectory initial point.
Finally, combining the sensitivity equations (4.35), the periodic boundary conditions (4.36) and the orthogonality constraint (4.37) we obtain the tangent Periodic Shadowing problem as

$$\begin{cases} \dot{\mathbf{y}}(t) = \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) + \mathbf{f}_{\partial p}(\mathbf{x}(t), p) + \dfrac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t), p), & t \in [0, T] \\ \mathbf{y}(0) = \mathbf{y}(T) \\ \mathbf{f}(\mathbf{x}(0), p)^{\intercal} \cdot \mathbf{y}(0) = 0 \end{cases} \tag{4.38}$$

From a mathematical viewpoint, this set of equations represents a boundary-value problem. The periodic boundary conditions ensure that the solution will remain bounded for all the time span, regardless the choice of T.
Having a problem of this type involves to use of a dedicated numerical method for boundary value problems. An efficient numerical methods to solve such problems is based on shooting technique [30], presented in the next chapter.

4.3 Adjoint Periodic Shadowing

The tangent method presented in the previous section allows calculating the sensitivity of the observable of interest to a variation in one of design parameters. If the sensitivity with respect to several parameters is required, the algorithm must be repeated for each perturbations applied. In such a context, the tangent method becomes computationally expensive, and therefore, as mentioned in the introductory chapter, adjoint methods are preferable.
To obtain the adjoint Periodic Shadowing method we introduce again the adjoint variables $\mathbf{q}(t) \in \mathcal{X} \equiv \mathbb{R}^N$ defined over $[0, T]$ and we combine it with the sensitivity equation (4.35) leads to the identity

$$\frac{1}{T}\int_0^T \mathbf{q}^\intercal(t)\left[\dot{\mathbf{y}}(t) - \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p)\cdot \mathbf{y}(t) - \mathbf{f}_{\partial p}(\mathbf{x}(t), p) - \frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t), p)\right] dt = 0 \tag{4.39}$$

Expanding the multiplication

$$\frac{1}{T}\int_0^T \left[\mathbf{q}^\intercal(t)\dot{\mathbf{y}}(t) - \mathbf{q}^\intercal(t)\mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p)\cdot \mathbf{y}(t) - \mathbf{q}^\intercal(t)\mathbf{f}_{\partial p}(\mathbf{x}(t), p) - \mathbf{q}^\intercal(t)\frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t), p)\right] dt = 0 \tag{4.40}$$

and integrating by parts the first term

$$\mathbf{q}^\intercal(t)\dot{\mathbf{y}}(t) = [\mathbf{q}^\intercal(t)\mathbf{y}(t)]\Big|_0^T - \int_0^T \dot{\mathbf{q}}^\intercal(t)\mathbf{y}(t)dt \tag{4.41}$$

Replacing it in (4.40) and rearranging

$$\begin{aligned}\frac{1}{T}\int_0^T &[-\dot{\mathbf{q}}^\intercal(t) - \mathbf{q}^\intercal(t)\mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p)]\cdot \mathbf{y}(t)dt + \\ &+ \frac{1}{T}\int_0^T \left[-\mathbf{q}^\intercal(t)\mathbf{f}_{\partial p}(\mathbf{x}(t), p) - \frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t), p)\right] dt + \frac{1}{T}\left[\mathbf{q}^\intercal(t)\cdot \mathbf{y}(t)\right]\Big|_0^T = 0\end{aligned} \tag{4.42}$$

Now it is possible to add this term (since the identity to null value is verify) to (4.32), leading to

$$\begin{aligned}\mathcal{J}_{dp}^{\infty}(p) = \lim_{T\to\infty}\frac{1}{T}\Bigg\{\int_0^T &\left[J_{\partial\mathbf{x}}(\mathbf{x}(t),p)\cdot\mathbf{y}(t) + J_{\partial p}(\mathbf{x}(t),p)\right]dt+\\ &+\int_0^T\left[-\dot{\mathbf{q}}^\intercal(t) - \mathbf{q}^\intercal(t)\mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t),p)\right]\cdot\mathbf{y}(t)dt+\\ +\int_0^T\left[-\mathbf{q}^\intercal(t)\mathbf{f}_{\partial p}(\mathbf{x}(t),p) - \mathbf{q}^\intercal(t)\frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t),p)\right]dt &+ \left[\mathbf{q}^\intercal(t)\cdot\mathbf{y}(t)\right]\Bigg|_0^T\Bigg\}\end{aligned} \tag{4.43}$$

Factoring out $\mathbf{y}(t)$

$$\begin{aligned}\mathcal{J}_{dp}^{\infty}(p) = \lim_{T\to\infty}\frac{1}{T}\Bigg\{\int_0^T &\left[-\dot{\mathbf{q}}^\intercal(t) - \mathbf{q}^\intercal(t)\mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t),p) + J_{\partial\mathbf{x}}(\mathbf{x}(t),p)\right]\cdot\mathbf{y}(t)+\\ +\int_0^T&\left[-\mathbf{q}^\intercal(t)\mathbf{f}_{\partial p}(\mathbf{x}(t),p) - \mathbf{q}^\intercal(t)\frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t),p) + J_{\partial p}(\mathbf{x}(t),p)\right]dt+\\ &+\left[\mathbf{q}^\intercal(t)\cdot\mathbf{y}(t)\right]\Bigg|_0^T\Bigg\}\end{aligned} \tag{4.44}$$

The key idea that makes the Periodic Shadowing adjoint method so powerful is to select, as done for the conventional approach, a particularly adjoint variables to avoid the explicit computation of $\mathbf{y}(t)$ for every parameter of interest. To achieve this goal we need to select a particularly adjoint variable $\mathbf{q}(t)$ which allows to obtain

$$\dot{\mathbf{q}}^\intercal(t) + \mathbf{q}^\intercal(t)\mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t),p) = J_{\partial\mathbf{x}}(\mathbf{x}(t),p) \tag{4.45}$$

Secondly, the last term can be treated as

$$\begin{aligned}\frac{1}{T}\left[\mathbf{q}^\intercal(t)\cdot\mathbf{y}(t)\right]\Bigg|_0^T &= \frac{1}{T}\left[\mathbf{q}^\intercal(T)\mathbf{y}(T) - \mathbf{q}^\intercal(0)\mathbf{y}(0)\right]\\ &\overset{(4.36)}{=} \frac{1}{T}\mathbf{y}(0)\left[\mathbf{q}^\intercal(T) - \mathbf{q}^\intercal(0)\right]\end{aligned} \tag{4.46}$$

Thus, choosing the periodic boundary condition $\mathbf{q}^\intercal(T) = \mathbf{q}^\intercal(0)$ for the adjoint variables as well, allow us to obtain $\mathbf{q}^\intercal(t)$ from

$$\begin{cases}\dot{\mathbf{q}}^\intercal(t) + \mathbf{q}^\intercal(t)\mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t),p) = J_{\partial\mathbf{x}}(\mathbf{x}(t),p)\\ \mathbf{q}^\intercal(T) = \mathbf{q}^\intercal(0)\end{cases} \tag{4.47}$$

in order to evaluate $\mathcal{J}_{dp}^{\infty}$ without compute $\mathbf{y}(t)$ explicitly for each design variables. Thus, the sensitivity is

$$\mathcal{J}_{dp}^{\infty} = \lim_{T\to\infty}\frac{1}{T}\int_0^T\left[-\mathbf{q}^\intercal(t)\mathbf{f}_{\partial p}(\mathbf{x}(t),p) - \frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t),p) + J_{\partial p}(\mathbf{x}(t),p)\right]dt \tag{4.48}$$

The periodic boundary conditions guarantee that the adjoint solution remains bounded, i.e. it does not exhibit the typical exponential growth. This method is extremely powerful if we were to evaluate the influence of more perturbations.
The gradient T_{dp} is obtained from the resolution of an additional adjoint problem. Mathematical details about how obtain this term are presented in appendix B of Ref. [21].

Chapter 5

Numerical Methods

Once the boundary value problems (4.38) and (4.47) have been formulated, it is necessary to use appropriate numerical methods for the solution. In fact, as known, the success of a numerical process depends on a combination of two factor: a well-conditioned problem and a stable algorithm to solve it. Generally, finding the solution of boundary value problems is more diffult than finding the solution of initial value problems due to the more complex structure. Thus it makes sense to use a numerical method for a given boundary value problem by relating the problem to corresponding (easier) initial value problem and solving the latter numerically [30]. To perform this, shooting technique are used. This chapter will provide an overview of the shooting methods: starting from the single shooting and its limitations, we will reach the multiple shooting formulation for the tangent approach of Periodic Shadowing.

5.1 Shooting Methods

The key idea of shooting method is to reduce the given boundary value problem to several initial value problems. The name of these methods perfectly represents the way they work: in fact, the goal is to shoot out trajectories in different directions until we find a trajectory that has the desired boundary value. This techniques are widely used because, in addition to the different structure of the problem, for the latter, good numerical methods are well understood and a wide range of software own simple and flexible mathematical library. They are applicable to both linear and non-linear boundary value problems and, once the "conversion" into an initial value problem is done, the resulting system is solved iteratively.

For both of following descriptions tangent formulation of Periodic Shadowing (4.38) will be considered.

5.1.1 Single Shooting Method

The simplest shooting technique is the single shooting method. Its easy understanding and implementation are balanced by the possible presence (as in our case) of instability drawbacks. Consider the problem of tangent Periodic Shadowing (4.38)

$$\begin{cases} \dot{\mathbf{y}}(t) = \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) + \mathbf{f}_{\partial p}(\mathbf{x}(t), p) + \dfrac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t), p), \qquad t \in [0, T] & (5.1a) \\ \mathbf{y}(0) = \mathbf{y}(T) & (5.1b) \\ \mathbf{f}(\mathbf{x}(0), p)^{\intercal} \cdot \mathbf{y}(0) = 0 & (5.1c) \end{cases}$$

Based on the principle of superposition discussed in detail in Refs. [31, 32], any solution of (5.1a) can be written as a linear combination of N linearly independent solution, where N is the dimension of the square matrix $\mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p)$. Thus, the general solution can be represented by

$$\mathbf{y}(t) = \mathbf{Y}(t) \cdot \mathbf{y}_0 + \mathbf{h}(t) + \frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t), p), \qquad t \in [0, T] \tag{5.2}$$

where $\mathbf{Y}(t)$ is the fundamental matrix solution, $\mathbf{y}(t) \in \mathbb{R}^N$ is the parameter vector and $\mathbf{h}(t)$ is a particular solution. For definiteness, the fundamental matrix solution is an $N \times N$ matrix such that $\mathtt{det}(\mathbf{Y}(t)) \neq 0 \ \forall t$, i.e. the N columns of $\mathbf{Y}(t)$ are linearly independent. The fundamental matrix solution satisfies

$$\begin{cases} \dot{\mathbf{Y}}(t) = \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{Y}(t), \qquad t \in [0, T] \\ \mathbf{Y}(0) = \mathbf{I} \end{cases} \tag{5.3}$$

where $\mathbf{I}$ is the identity matrix. The particular solution is defined by the initial value problem

$$\begin{cases} \dot{\mathbf{h}} = \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{h} + \mathbf{f}_{\partial p}(\mathbf{x}(t), p), \qquad t \in [0, T] \\ \mathbf{h}(0) = \mathbf{0} \end{cases} \tag{5.4}$$

Thus, the N columns of $\mathbf{Y}(t)$ and the vector $\mathbf{h}(t)$ can be computed as solution of $N+1$ initial value problems. The boundary value problem solution $\mathbf{y}(t)$ is given as their superposition in equation (5.2). Exploiting (5.1b) it is possible to write

$$\mathbf{Y}(T) \cdot \mathbf{y}_0 + \mathbf{h}(T) \overset{(5.2)}{=} \mathbf{y}(T) = \cdot\mathbf{y}(0) \tag{5.5}$$

which is equivalent to

$$(\mathbf{Y}(T) - \mathbf{I})\, \mathbf{y}_0 = -\mathbf{h}(T) \tag{5.6}$$

The boundary value problem (5.1) has a unique solution iff (if and only if) the $(\mathbf{Y}(T) - \mathbf{I})$ matrix is nonsingular. Solving (5.6) and then (5.2) we obtain obtain the solution. However, as mentioned above, the single shooting method cannot always be applied due to its stability drawbacks. In fact, the initial value problems integrated in the process could be unstable even if the boundary value problem is well posed. The major trouble is connected with the roundoff error accumulation as described in Ref. [30]. In figure 5.1 we present examples of the left-hand side matrix $(\mathbf{Y}(T) - \mathbf{I})$ condition numbers obtained with single shooting method fixing the initial conditions and time step.

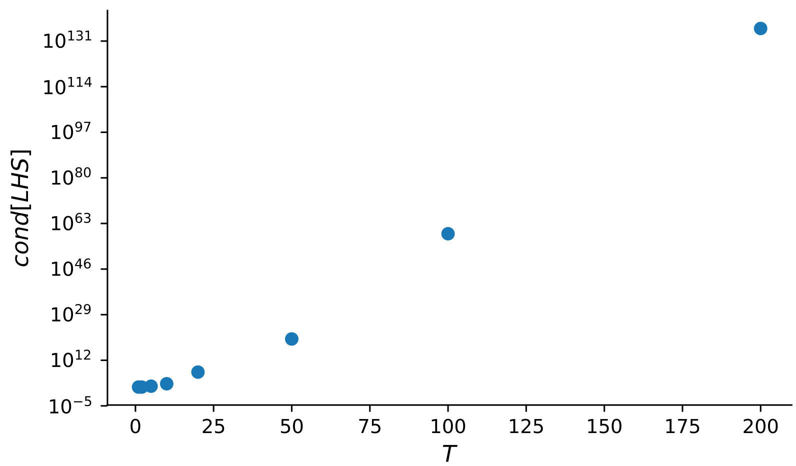

Figure 5.1: Condition number of the left-hand side matrix for the single shooting algorithm. Fixed the initial condition and the time step ($\Delta t = 10^{-2}$) the condition number is computer for increasing time span, $T = [1, 2, 5, 10, 20, 50, 100, 200]$.

We observe that in for small time intervals the condition number remains bounded, but considering integration time long relative to the time scale of the system, the condition number explode numerically. To reduce these spurious effects and make the results more trustworthy, methods that work by reducing the interval over which the initial values problems are solved have been developed. One of them is the multiple shooting method.

5.1.2 Multiple Shooting Method

Having to keep the roundoff error limited, the logical expedient adopted is to consider smaller time interval over which integrate the initial value problems. For this reason the whole time span of integration $[0, T]$ is subdivided into a mesh of N shooting points evenly spaced at

$$t_i = i\frac{T}{N}, \qquad i = 0, \dots, N-1$$

The number of fragments in which to subdivide the whole integration time is an important input parameter that must be chosen appropriately. In fact, the T/N ratio is strictly connected to the condition number of the associated matrix. Thus, it is therefore clear that the correct number of shooting points must be chosen in order to maximize the efficiency of the algorithm in terms of computational cost and to not undermine the solution of the problem. The key point is, therefore, to perform the steps presented in the previous section on each subinterval $i \in [t_i, t_{i+1}]$ created.

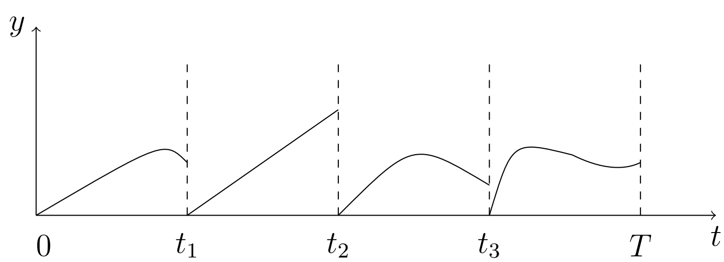

Figure 5.2: Graphical representation of multiple shooting technique.

Using the previous notation, it is possible to express the solution $\mathbf{y}_i(t)$ over the i-th subinterval originating from a particular initial condition $\mathbf{y}_i^0$ as

$$\begin{aligned} \mathbf{y}_i(t) &= \mathbf{Y}_i(t, \mathbf{x}_i) \cdot \left\{ \mathbf{y}_i^0 + \int_{t_i}^{t} \mathbf{Y}_i^{-1}(s, \mathbf{x}_i) \cdot \left[\mathbf{f}_{\partial p}(\mathbf{x}(s), p) + \frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(s), p) \right] ds \right\} \\ &= \mathbf{Y}_i(t, \mathbf{x}_i) \cdot \mathbf{y}_i^0 + \mathbf{h}_i(t) + (t - t_i)\frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t)), \qquad i = 0, \dots, N-1 \end{aligned} \tag{5.7}$$

The fundamental matrix solution and the particular solution are still obtained solving the initial values problems (5.3) and (5.4) but now referred to the i-th segment. Now it is possible to solve the problem looking for the initial condition $\mathbf{y}_i^0, i = 0, \dots, N-1$, such that the overall solution is continuous at the shooting point [21]. The system obtained, considering also the condition of orthogonality, turns out to be

$$\mathbf{y}_{i+1}^0 = \mathbf{Y}_i(t_{i+1}, \mathbf{x}_i) \cdot \mathbf{y}_i^0 + \mathbf{h}_i(t_{i+1}) + (t_{i+1} - t_i)\frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t_{i+1}), p), \qquad i = 0, \dots, N-1 \tag{5.8}$$

which can be written in matrix form as

$$\left[\begin{array}{cccc|c} \mathbf{Y}_0(t_1, x_0) & -\mathbf{I} & \cdots & & \mathbf{f}(t_1) \\ \vdots & \ddots & \ddots & & \vdots \\ & & \ddots & -\mathbf{I} & \\ -\mathbf{I} & & & \mathbf{Y}_{N-1}(t_N, \mathbf{x}_{N-1}) & \mathbf{f}(t_N) \\ \hline \mathbf{f}^\intercal(0) & \mathbf{0}^\intercal & \cdots & \mathbf{0}^\intercal & 0 \end{array}\right] \cdot \left[\begin{array}{c} \mathbf{y}_0^0 \\ \vdots \\ \\ \mathbf{y}_{N-1}^0 \\ \hline T_{dp}^0/N \end{array}\right] = - \left[\begin{array}{c} \mathbf{h}_0(t_1) \\ \vdots \\ \\ \mathbf{h}_{N-1}(t_N) \\ \hline 0 \end{array}\right] \tag{5.9}$$

The bordering vector represents the orthogonality constraint included in system (4.38). The system is solved using L, U factorisation. The left-hand side matrix is a sparse matrix which takes, for instance, the configuration presented in figure 5.3 for the Lorenz system (6.6) with twenty shooting points.

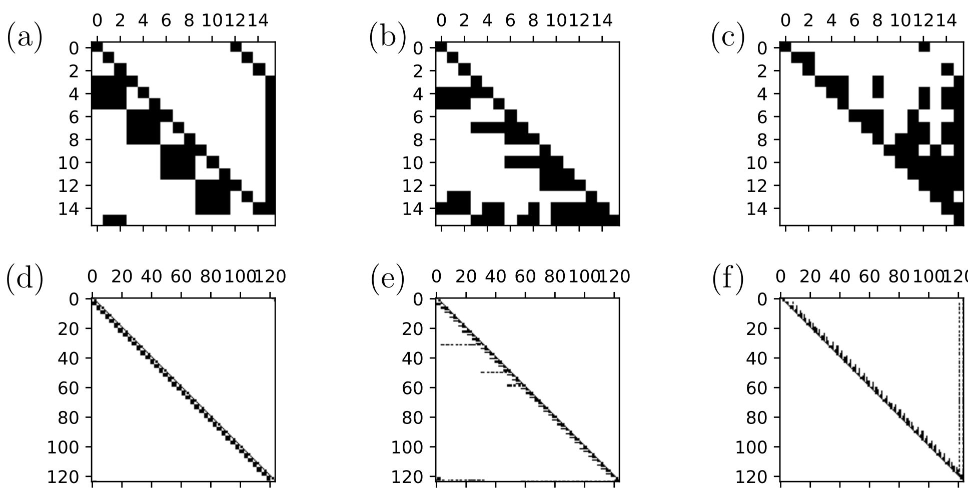

Figure 5.3: Top: configuration of left-hand side matrix of (5.9) computed for four shooting points (a) and the LU factorisation in (b) and (c), respectively. Bottom: configuration of left-hand side matrix of (5.9) computed for forty shooting points (d) and the LU factorisation in (e) and (f), respectively. The black squares represent the non-zero element.

The spy plot of figure 5.3a represents the configuration of the left-hand side matrix considering four shooting points, i.e. a time span of $T = 20$ time units, for the multiple shooting formulation of the tangent Periodic Shadowing approach applied to Lorenz equations (a) whilst in panels (b) and (c) the L, U factors is presented. The features of the matrix in panel (a) depends on two factors: number of shooting points and system dimensions. Primarily, the sparsification of the matrix increases with increasing time interval considered, i.e. the number of shooting points. This, also dependent on dimensions of system analysed. In fact, hypothetically considering high-dimensional

systems, e.g. Navier-Stokes equations, coupled with a long integration time, the matrix would become so large that a block elimination method could not be applied. From what is evident, the considerable increase in storage and/or computation requirements are nowadays the bottleneck of this application.
The choice in how many shooting points to subdivide the whole trajectory is performed following the verification of the condition number of the matrix. In fact, choosing a few shooting points increases the speed of the code but we stumble in the problems related to single-shooting. On the other hand, if we use too many shooting points the dimensions of the matrix increases more and more spoiling the condition number. In figure 5.4 we plot the condition number of the left-hand side matrix for the Lorenz equation as a function of T/N ratio fixed the initial condition and time step. We

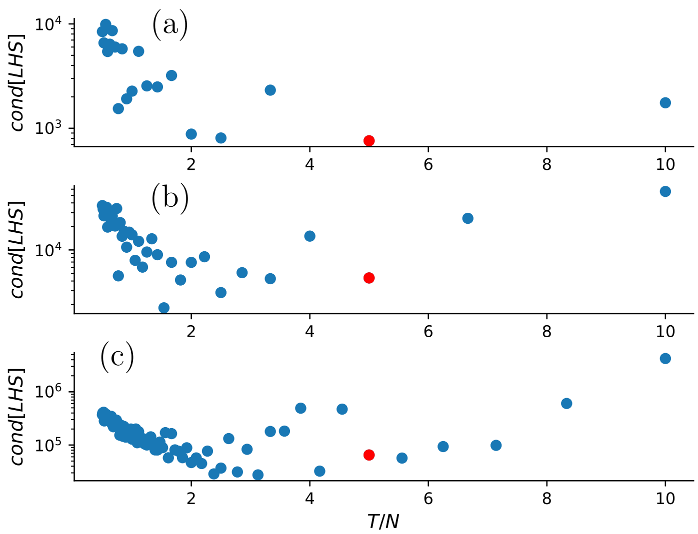

Figure 5.4: Condition number of the left-hand side sparse matrix as a function of T/N ratio for $T = 10$ (a) $T = 20$ (b) $T = 100$ (c) time units. The red points corresponds to $T/N = 5$.

observe that the red markers, at $T/N = 5$, correspond (on average) to the ratio that provides the minimum condition number. For this reason, for both systems we select $T/N = 5$ as input parameter.

5.2 Numerical Solution of Linearised Equations

The numerical integration of nonlinear and linearised equation is obtained trough a classical fourth-order Runge-Kutta scheme with step size $dt = 10^{-3}$. This value is chosen as step size in order to obtain accurate solution will be discussed in more detail in the appropriate subsection. The linearised equations are solved in a coupled manner with the nonlinear equations by propagating forward in time the augmented system [21].

$$\begin{cases} \dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) \\ \dot{\mathbf{y}}(t) = \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) + \mathbf{f}_{\partial p}(\mathbf{x}(t), p) + \frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t), \mathbf{p}) \end{cases} \tag{5.10}$$

The shooting points are then created starting from a position on the trajectory obtained by propagating the initial condition for a number of temporal units sufficiently high to eliminate the spurious effects of the initial transient, as represented in figure 5.5.

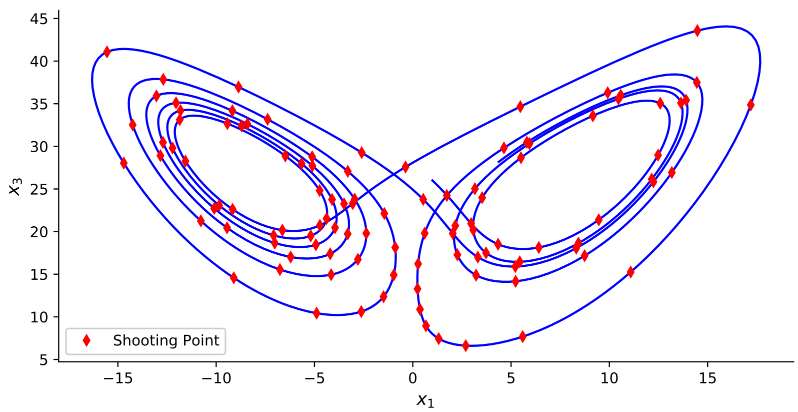

Figure 5.5: Representation of the shooting points in the Lorenz attractor. The ratio between the time span and the number of shooting points must be, for the chaotic system arising from Lorenz equations, not greater than 5.

In figure 5.5 we present a classical trajectory of the Lorenz attractor and the shooting points used to construct the matrix. It must be underlined that, for the Lorenz system, the relationship between the integration time T and the number of shooting points N must never be greater than 5. To obtain the sensitivity of the observable of interest (4.32) with respect a certain design variable p, the system 5.10 is augmented coupling a quadrature equation, leading to

$$\begin{cases} \dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) \\ \dot{\mathbf{y}}(t) = \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{y}(t) + \mathbf{f}_{\partial p}(\mathbf{x}(t), p) + \frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t), \mathbf{p}) \\ \dot{q}(t) = \mathbf{g}(\mathbf{x}(t), \mathbf{y}(t)) \end{cases} \tag{5.11}$$

and when integrated, $\mathbf{q}(T)$ is equal to the integral

$$q(T) = \int_0^T \mathbf{g}(\mathbf{x}(t), \mathbf{y}(t)) dt \tag{5.12}$$

with high accuracy.

5.3 Convergence Analysis

In this section, the mutual interaction between the time step used in the integration process and the number of shooting points considered to subdivide the whole time span will be analysed. Indeed it is necessary to verify that the time step is enough small to analyse the system correctly. In order to observe this, we present in figure 5.6 the time evolutions of the sixth component of Reiterer system obtained by integrating the equations (6.9) and using $\Delta t = [0.01, 0.02, 0.04]$, in panel (a),(b) and (c), respectively. We fix the integration time $T = 20$ and we modify the number of shooting points in order to obtain three different T/N ratio: $N = [4, 8, 16]$. We observe that if we select

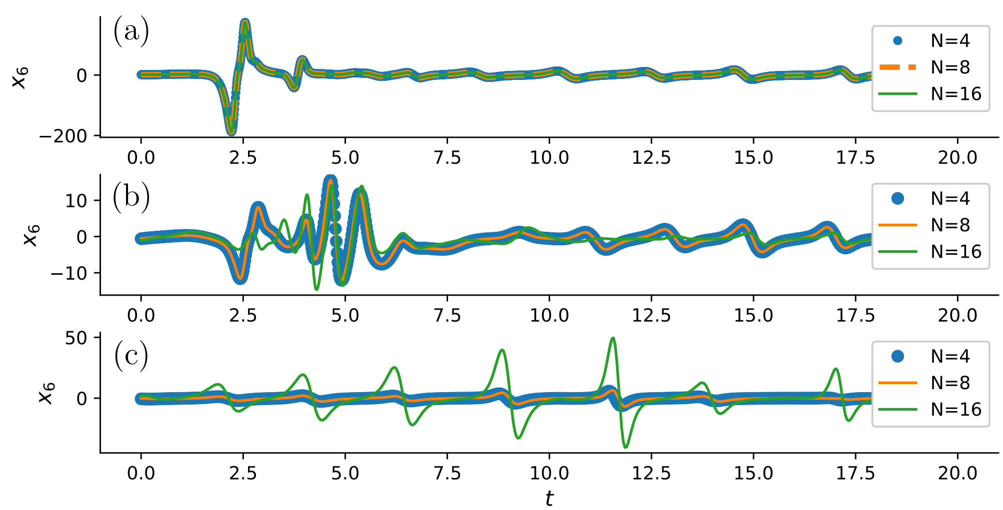

Figure 5.6: Time evolution of the sixth component using an increasing number of shooting points for $\Delta t = 0.01$ (a), $\Delta t = 0.02$ (b) $\Delta t = 0.04$.

a time step too wide, the number of shooting points in which we subdivide the whole time span affects the results. This phenomenon should be avoided because the dynamic, and thus the sensitivity, of the system are corrupted by the mathematical model. By contrast, if $\Delta t = 0.01$ the N value does not affect the results and thus will be the input for all the following analysis. In what follows, we have used $\Delta t = 0.01$ for both Lorenz and Reiterer systems.

Chapter 6

Mathematical Models

This chapter describes the mathematical and physical models of the low-dimensional chaotic systems used as application models for the sensitivity analysis presented in this project. Both systems arise from natural convection problem in fluid cells bounded by two plates and heated from below. To investigate and validate the response of the method, it was necessary initially to use a chaotic system with low computational cost. For this reason, our choice fell on from Lorenz equations [5], whose low-dimensionality and the vast amount of literature sources make it particularly suitable for testing and validation. Subsequently, we apply it to a system of nine nonlinear ordinary differential equations describing the three-dimensional cells in Rayleigh-Bénard convection problem [23]. This chapter contains the physical model of the chaotic systems and describes the influence on the chaotic behaviour of the temperature difference between the plates. For both systems the most important views of the arising attractors will be presented whilst in the last sections, the coordinate transformation and the linearisation will be introduced.

6.1 The Lorenz System

"Does the Flap of a Butterfly's wings in Brazil Set Off a tornado in Texas?"

Surely this sentence has already been heard, probably mistakenly associated with the philosophical or poetic field. Actually, it is the title [33] given by Edward Lorenz, following suggestions from colleagues, for a talk held at the $139th$ meeting of the *American Association for the Advancement of Science* in 1972. To appropriately contextualize the above quotation it is necessary to know that Lorenz was an American meteorologist whose goals ware to construct a mathematical model that could represent accurately

the phenomenology related to long-term weather forecasts, due to his scepticism about the linear statistical models used in the '50. This study led him to be the first to recognize what is now called *chaotic behaviour* in the mathematical modelling, and he realized that small differences in the initial condition of a dynamic system characterized by chaotic nature, could trigger vast and often unsuspected results.
From the late seventies to today, literature has grown a lot with regard to the mutual relationship between fluid dynamic and chaos theory, in particular for interpreting the increasing complexity of the dynamics all along the cascade of instabilities leading to turbulence [34, 35]. Lorenz developed a simplified mathematical model of the two-dimensional convection of Rayleigh-Bénard, i.e. the flow between two infinite parallel plates spaced h where the lower one is heated like represented in figure 6.1. The change in density due to temperature variations gives rise to a flow generated by buoyancy, which opposes to the viscous forces in the fluid. The equilibrium between these forces determines whether the flow is stable or not: such system has a steady state solution, in which there is no movement and the temperature varies linearly with depth, and an unstable one, in which the phenomenon of convection arisen. This simple convection problem can be considered a drastically simplified model of the Earth's atmosphere.

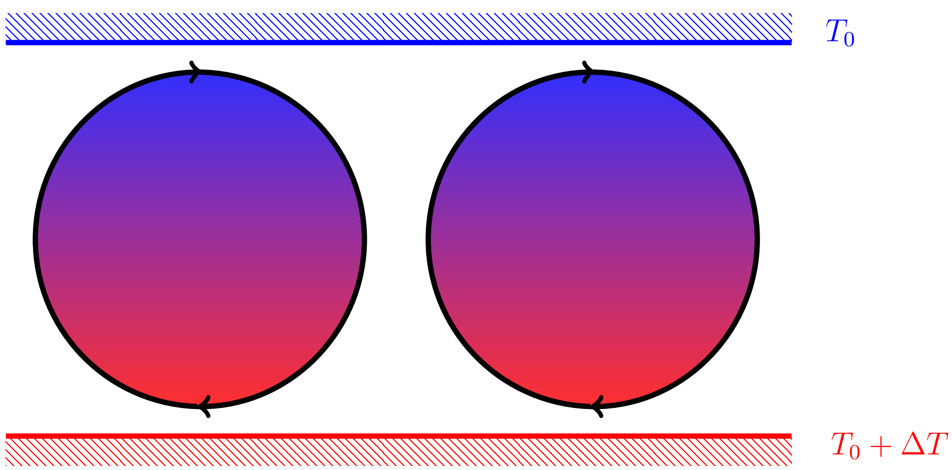

Figure 6.1: Graphical representation of fluid cells bounded by two plates and heated from below. This illustrationn represents Rayleigh-Bénard convection problem.

The upper and lower plates have a temperature of $T_u = T_0$ and $T_L = T_0 + \Delta T$, respectively, and the fluid contained between these has density ρ, kinematic viscosity ν, thermal diffusivity κ and thermal diffusivity α. The starting point of the Lorenz digression is the system constituted by the Navier-Stokes equations coupled to the thermal energy equation upon which the Boussinesq approximation is applied, i.e. the variations in fluid properties were ignored except for the density $\rho = \rho(p, T)$ multiplied by the gravitational acceleration g. The boundary conditions of this problem are the constant temperature of both plates and no-slip condition. From this system, Lorenz succeeded in obtaining, through dimensionless variable, Fourier expansion and Galerkin truncation as presented in Ref. [5], the system of ordinary differential equations written in the canonical form

$$\begin{cases} \dot{x_1} = \sigma(x_2 - x_1) \\ \dot{x_2} = x_1(\rho - x_3) - x_2 \\ \dot{x_3} = x_1 x_2 - \beta x_3 \end{cases} \tag{6.1}$$

where x_1 is proportional to the intensity of the convective motion, x_2 is related to the temperature difference between the ascending and descending current (same signs of x_1 and x_2 mean that warmer fluid is rising) while x_3 is proportional to the distortion of the vertical temperature profile from linearity (positive value means strong gradients occur near the boundaries). The constants σ, ρ and β are system parameters and correspond to the Prandtl[1] number, a value proportional to Rayleigh[2] number and characteristic dimension of the domain, respectively. If the σ and β parameters are generally considered constant and equal to $\sigma = 10$ and $\beta = 8/3$, become important to pay attention to the parameter ρ, which is the so-called reduced Rayleigh number. The two parameters σ and β are kept constant unless otherwise specified whilst the influence of this term will be resumed later. The value of the system at the generic time t can be illustrated by a single point with components x_1, x_2, x_3. Over time the state of the system changes and the correspondent displacement of the point in the phase space along a curve is called trajectory which, under particular constraints, takes the well-known form of the Lorenz attractor presented in figure 6.2.

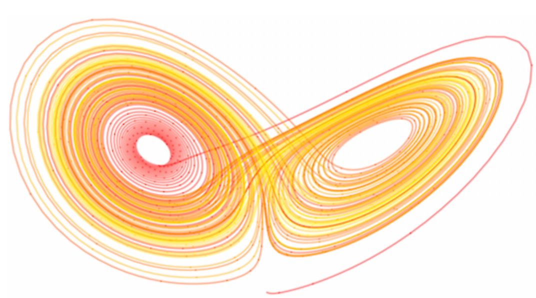

Figure 6.2: Lorenz attractor obtained for standard parameters values.

[1] $Pr = \dfrac{\text{kinematic viscosity}}{\text{thermal diffusivity}} = \dfrac{c_p \mu}{\kappa}$

[2] The Rayleigh number is a dimensionless number that is associated with buoyancy-driven airflow and can be regarded as a measure of the driving forces of natural convection: $Ra \propto \Delta T h^3$

6.1.1 Mathematical Properties of the Lorenz System

Before entering into details of the chaotic behaviour it is necessary to describe the properties of this system:

- nonlinearity: the nonlinearity behaviour of the system is given by the x_1x_3 and x_1x_2 terms in the second and third equation, respectively;

- symmetry: equations are invariant under $(x_1, x_2) \to (-x_1, -x_2)$. Hence, if $(x_1(t); x_2(t); x_3(t))$ is a solution, $(-x_1(t); -x_2(t); x_3(t))$ is also one. This can be seen by simple substitution into the Lorenz equations (6.1);

- dissipative: a system is termed dissipative when the divergence of f is negative. Considering the system

$$\begin{cases} \mathrm{div}(f) = \dfrac{\partial(\sigma(x_2 - x_1))}{\partial x_1} + \dfrac{\partial(x_1(\rho - x_3) - x_2))}{\partial x_2} + \dfrac{\partial(x_1x_2 - \beta x_3)}{\partial x_3} = -(1 + \sigma + \beta) \\ \sigma = 10 \\ \beta = 8/3 \end{cases}$$

  it turns out that $\mathrm{div}(f) < 0$ so, volumes shrink exponentially fast under the function $V(t) = V(0)e^{-(1+\sigma+\beta)}$ to limiting of zero volume;

- fixed points: the origin $(x_1, x_2, x_3) = (0, 0, 0)$ is a fixed point for all values of the parameters. If $\rho = 1$ a symmetric pair of fixed points appears in addition to the previous one

$$(x_1, x_2, x_3) = (\pm\sqrt{\beta(\rho - 1)}, \pm\sqrt{\beta(\rho - 1)}, \rho - 1) \tag{6.2}$$

  Lorenz called them C^+ and C^-, and represent the left/right turning convection rolls. So, if $\rho = 1^+$ a pitchfork bifurcation occurs, i.e. the system evolves from one to three fixed points.

6.1.2 Role of the Parameter ρ

The ρ value, proportional to the temperature gradient between the two plates, represents the reduced Rayleigh number (Rayleigh number reduced to the critical value $R_c = 27\pi^4/4$) and plays a fundamental role in the analysis of the trajectory configuration. For $\rho < 1$, the origin is globally stable, while, as stated previously, for $1 < \rho < \rho_a \approx 14$ there are two stable fixed points: C^+ and C^-.

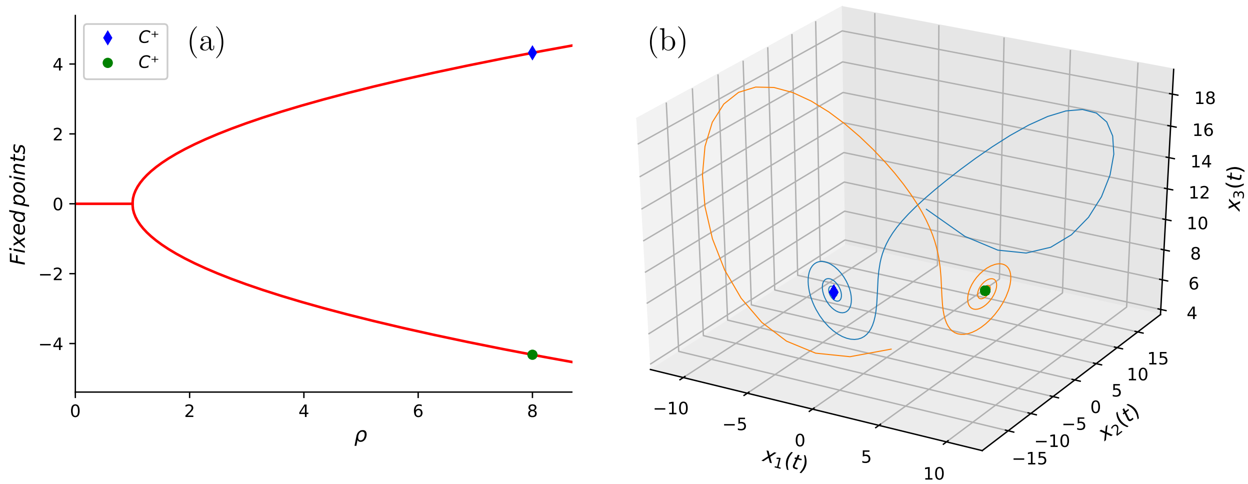

Figure 6.3: Stable fixed points values for $0 < \rho < \rho_a$ (a) and trajectories obtained with $(\sigma, \rho, \beta) = (10, 8, 8/3)$ with initial condition $x_0 = (\mp 5, \pm 18, 8)$ and integration time $T = 40$ time units (b).

From the images of figure 6.3 the continuity of the results can be verified, e.g. if $\rho = 8$, through the equation (6.2) we obtain $(\pm 4.32, \pm 4.32, 7)$, corresponding to C^+ and C^- in figure 6.3a, respectively. Continuing to increase the ρ value up to

$$\rho_{hb} = \frac{\sigma(\sigma + \beta + 3)}{\sigma - \beta - 1} = 24.73 \tag{6.3}$$

we enter into the “transient chaos” or “semi-chaotic zone”. At $\rho = \rho_{hb}$ a subcritical Hopf bifurcation occurs[3] by which the stable points for $C^\pm$ lose their stability creating an unstable periodic orbit. Once this value is exceeded, the chaotic behaviour is manifested through the presence of the strange attractor, also named Lorenz attractor. The characteristic of a strange attractor is hardly to foresee due to the switching between attraction and repulsion effects.

With the configuration set in figure 6.4, the macroscopic shape of the Lorenz attractor (sometimes so-called the Lorenz butterfly due to its shape) doesn’t change significantly in time and the trajectory never repeats itself. The graph that represents the temporal trend of the components of the system is decidedly more usable, in fact, starting from the same initial condition, it is possible to observe in figure 6.5 the two different time-evolutions for a non-chaotic and chaotic configuration. As it can be seen, after an initial transient, the solution settles into an irregular oscillation that persists in time but which never repeats exactly.

[3] In Hopf bifurcation a complex conjugate pair of eigenvalues becomes purely imaginary

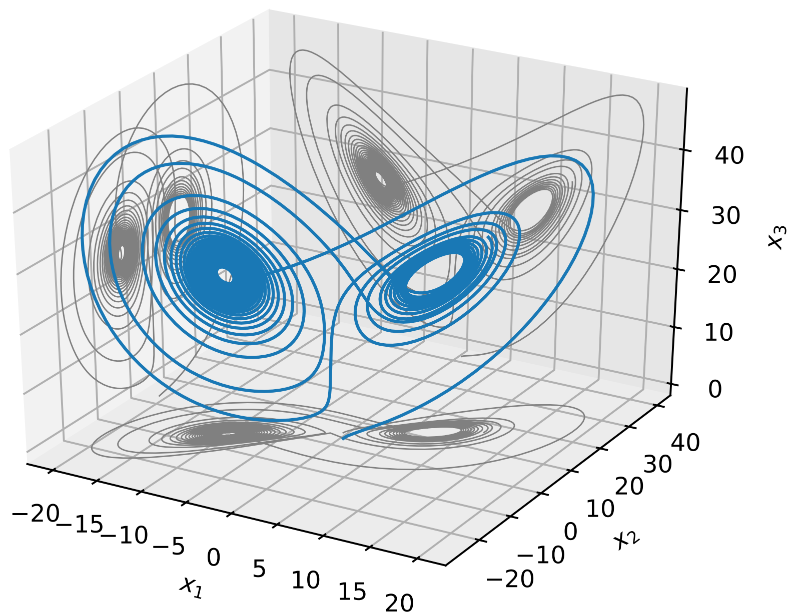

Figure 6.4: Lorenz attractor simulation with $(\sigma, \rho, \beta) = (10, 470/19, 8/3)$ with integration time of $T = 50$ time units.

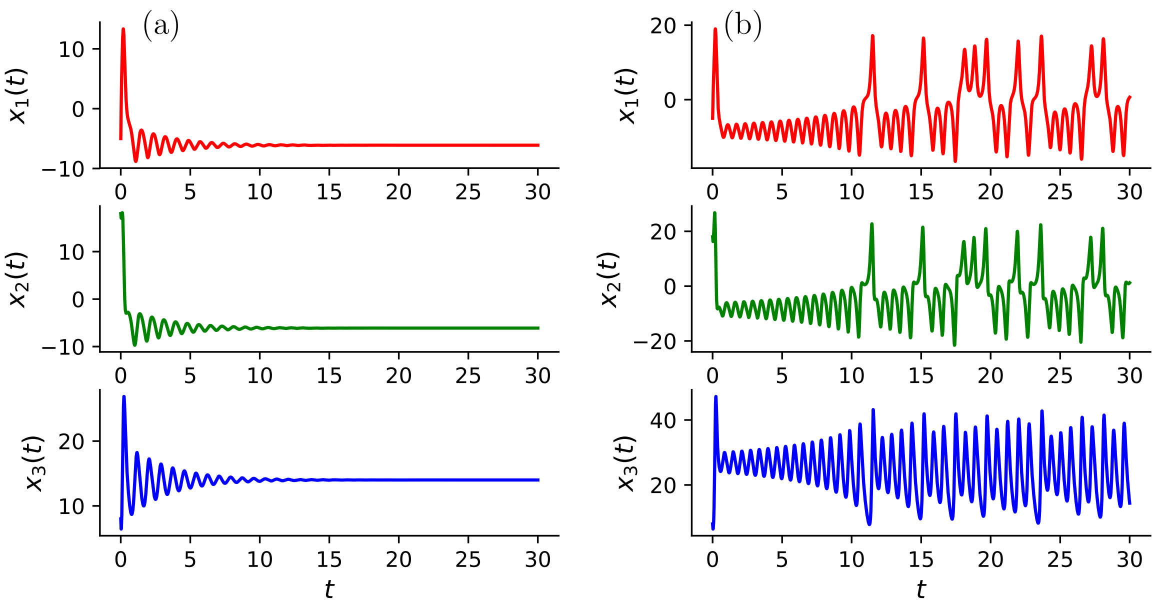

Figure 6.5: Time evolution of the components obtained with initial condition $x_0 = (-5, 18, 8)$ for a non-chaotic $(\sigma, \rho, \beta) = (10, 15, 8/3)$ (a) and chaotic $(\sigma, \rho, \beta) = (10, 28, 8/3)$ (b) behaviour.

The peculiarity of this system, or better, of all the chaotic systems in general, is the high sensitivity on initial conditions. The distance between two arbitrarily selected and almost coincident points diverges exponentially when the integration interval starts to become relatively long. This, coupled with the unstable periodic dynamics that ensures that the trajectory will never escape from the attractor, creates a complex structure on which it is difficult to perform sensitivity analysis. In figure 6.6a the two trajectories

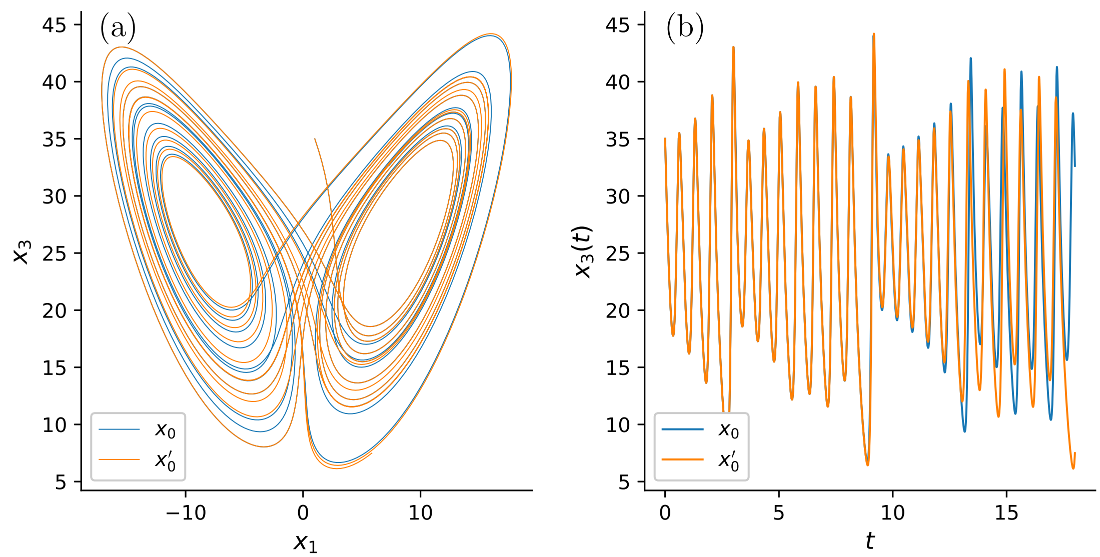

Figure 6.6: Lorenz attractor (a) and time evolution of its third component (b) for two trajectories initially spaced out by 10^{-4} with $(\sigma, \beta, \rho) = (10, 8/3, 28)$.

initially divided by an infinitesimal perturbation overlap until a certain integration time ($T \approx 10$ time units) beyond which they do not contain common elements. This feature can be noticed more easily in the time evolution plot of the x_3-component, as figure 6.6b reports. We observe that until ≈ 25 time units, the initial perturbation does not manifest itself clearly in the time evolution.

If we increase the ρ value to approximately $\rho \approx 145$, we pass from the chaotic regime to one called Double Periodic Orbits. Further increasing the value of $\rho > 165$ entails entering into an intermittent chaotic regime. The regimes for these values are not subject of discussion in this project but more information are present in the cited references inherent to Lorenz equations. Everything we have discussed in this section represents a small part of the physical characteristics of Lorenz equations. For our sensitivity analyses, we will focus on parameter values for which chaotic behaviour occurs.

6.2 Nine-Dimensional System

The chaotic system that will be analysed in this section has the same physical roots as the one presented in the previous section (see figure 6.1). The difference between the two formulations lies in the fact that, once applied the Boussinesq-Oberbeck approximation and boundary conditions, Lorenz equations have been expanded with a double Fourier series [5], whilst Reiterer et al. in Ref. [23], in order to obtain an higher-dimensional model, applied a triple Fourier series ansatz up to second order. From a physical standpoint Lorenz approach corresponding to the description of two-dimensional "fluid rolls" whereas in this case the analysis has as its object three-dimensional cells.
From that point, operating according to Ref. [23], the system of nine nonlinear ODEs is presented.

$$\begin{cases} \dot{x}_1(t) = -\sigma\beta_1 x_1(t) - x_2(t)x_4(t) + \beta_4 x_4^2(t) + \beta_3 x_3(t)x_5(t) - \sigma\beta_2 x_7(t) \\ \dot{x}_2(t) = -\sigma x_2(t) + x_1(t)x_4(t) - x_2(t)x_5(t) + x_4(t)x_5(t) - \sigma x_9(t)/2 \\ \dot{x}_3(t) = -\sigma\beta_1 x_3(t) + x_2(t)x_4(t) - \beta_4 x_2^2(t) - \beta_3 x_1(t)x_5(t) + \sigma\beta_2 x_8(t) \\ \dot{x}_4(t) = -\sigma x_4(t) - x_2(t)x_3(t) - x_2(t)x_5(t) + x_4(t)x_5(t) + \sigma x_9(t)/2 \\ \dot{x}_5(t) = -\sigma\beta_5 x_5(t) + x_2^2(t)/2 - x_4^2(t)/2 \\ \dot{x}_6(t) = -\beta_6 x_6(t) + x_2(t)x_9(t) - x_4(t)x_9(t) \\ \dot{x}_7(t) = -\beta_1 x_7(t) - \rho x_1(t) + 2x_5(t)x_8(t) - x_4(t)x_9(t) \\ \dot{x}_8(t) = -\beta_1 x_8(t) + \rho x_3(t) - 2x_5(t)x_7(t) + x_2(t)x_9(t) \\ \dot{x}_9(t) = \rho x_4(t) - x_9(t) - \rho x_2(t) - 2x_2(t)x_6(t) + 2x_4(t)x_6(t) + x_4(t)x_7(t) - x_2(t)x_8(t) \end{cases} \tag{6.4}$$

where σ, ρ, β_i refer to the Prandtl number, reduced Rayleigh number and particular physical dimensions of the fluid layer,respectively. This nine-dimensional system is a very interesting model for investigating chaotic dynamics in a phase space with dimensions greater than 3.

6.2.1 Role of the Parameter ρ

In order to analyse the effects of the temperature difference between the plates on chaotic behaviour, the Prandtl number σ and the geometric characteristics β_i of the system will be fixed while the reduced Rayleigh number ρ will be varied, exactly as performed for the Lorenz equations. To aid the comparison with the results presented in Ref. [23], we select $\sigma = 0.5$. For $\rho < 1$ the trivial fixed point is stable and the fluid is at rest, while for $\rho = 1$ that point become unstable and a new equilibrium, characterized by a stationary convective flow, is established. Continuing to increase the temperature difference, we observe that for $\rho \approx 13.07$ a limit cycle is generated thanks to Hopf bifurcation. This limit cycle is clearly visible in figure 6.7, where the nonlinear solution is plotted on $x_7(x_6)$ and $x_9(x_6)$ graphs, respectively.

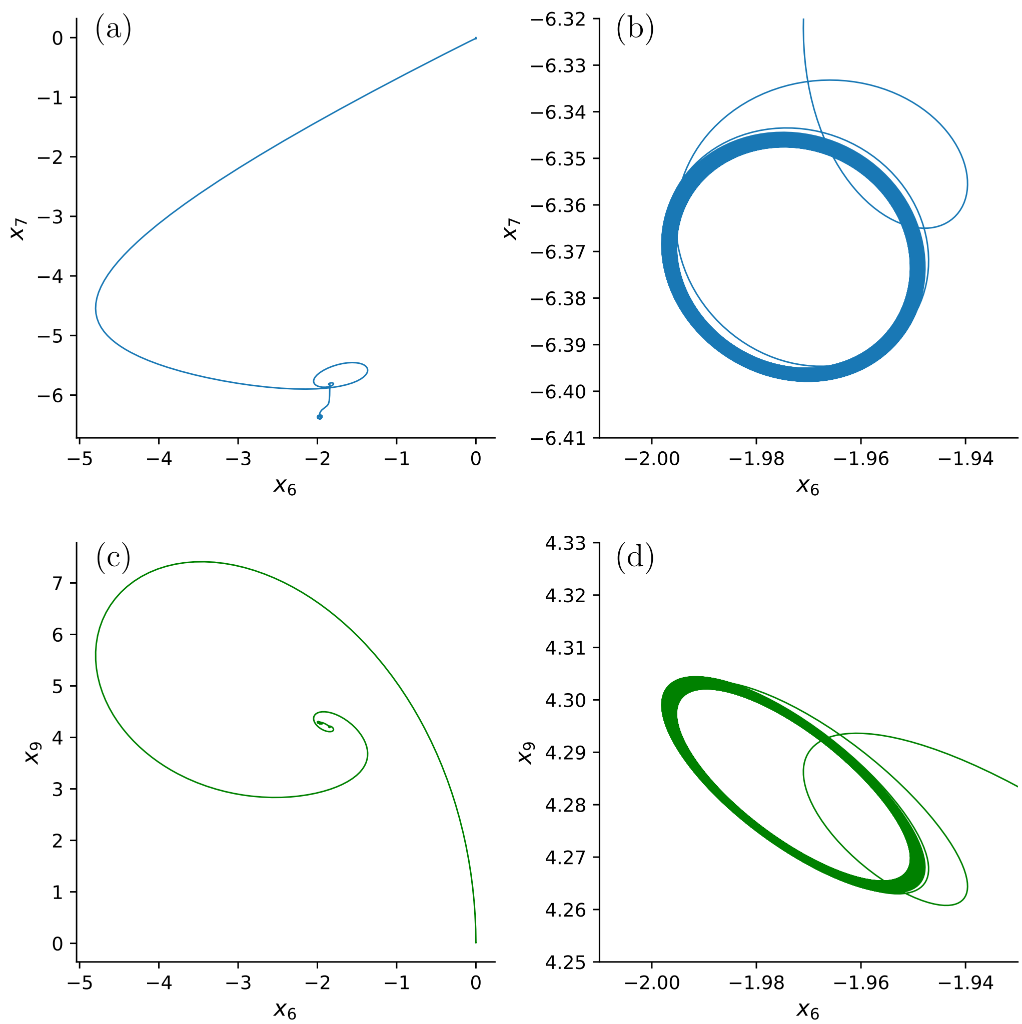

Figure 6.7: Phase space x_6-x_7 (a),(b) and x_6-x_9 (c),(d) of the nine-dimensional chaotic system for $(\sigma, \rho) = (0.5, 13.07)$ and integration time $T = 1000$ time units. In panel (b) and (d) a close-up of the respective limit cycle is presented.

The bifurcation which leads to the limit cycle can be supercritical or subcritical, resulting in stable or unstable limit cycle. The choice of using these two projection planes will be clearer later, when, it will be necessary to use particular representations in order to observe the attractors for the chaotic behaviour.

To achieve the chaotic behaviour it is necessary to exceed the value of $\rho \approx 14.22$. In fact, in the range $\rho \in [14.10, 14.22]$ the period-doubling cascade is easily identified and does not seem to depend on the plane projections used. Once this threshold is exceeded, differences in the representations of the dynamics are clearly visible and, to have continuity of results with those of Ref. [23], the attractor will be projected on $x_7(x_6)$ and $x_9(x_6)$ planes.

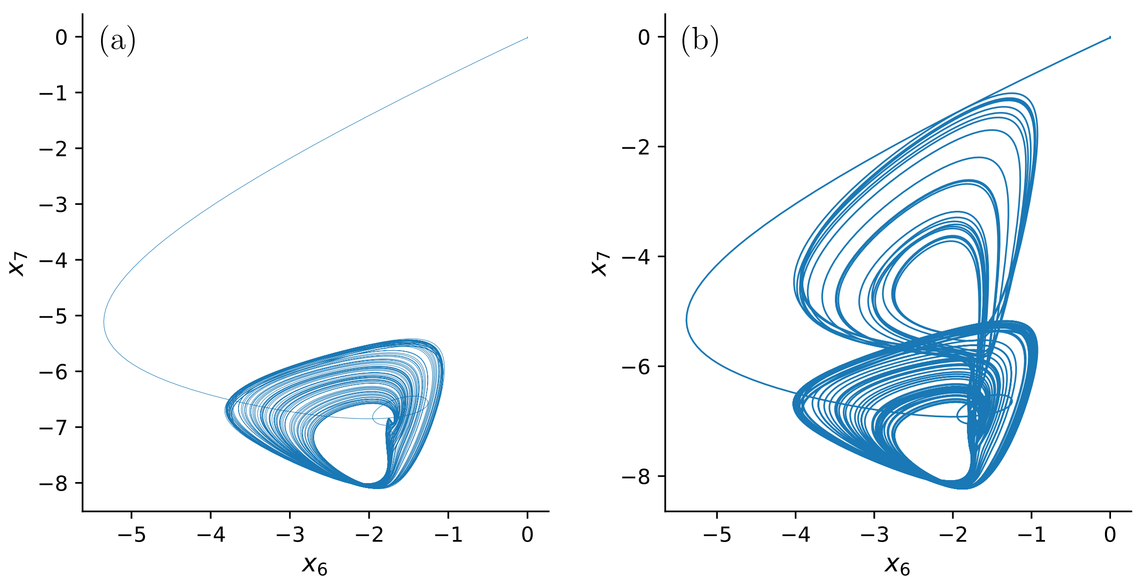

Figure 6.8: Trajectories projected on x_6-x_7 for $(\sigma, \rho) = (0.5, 14.22)$ (a) and $(\sigma, \rho) = (0.5, 14.3)$ (b).

In figure 6.8 is reported the effect of parameter ρ in the trajectory in $x_7(x_6)$ projection. This system response is also observable, albeit with some differences, in the projection plane $x_9(x_6)$ of the attractor, presented in figure 6.9. A different element between the two figures appears, for instance, considering $\rho = 14.30$: a double attractor is observed in the $x_7(x_6)$ plane projection (figure 6.8) whilst a simple attractor is found in the $x_9(x_6)$ plane projection (figure 6.9).
To observe the chaotic behaviour, it is more representative to highlight the strong dependence on initial conditions. For this reason in figure 6.10 we evaluate the time evolution of two trajectories originating from identical initial condition except for $\Delta x_6 = 10^{-7}$. We observe in figure 6.10(a) that the effect of the perturbation does not manifest clearly into the time evolutions, however, exceeded the value of 500 time units (this value is not absolute and depends on the initial conditions and parameters value) the two trajectories start to be completely unrelated to each other, symptom of the strong dependence on initial conditions characteristic of chaotic system.

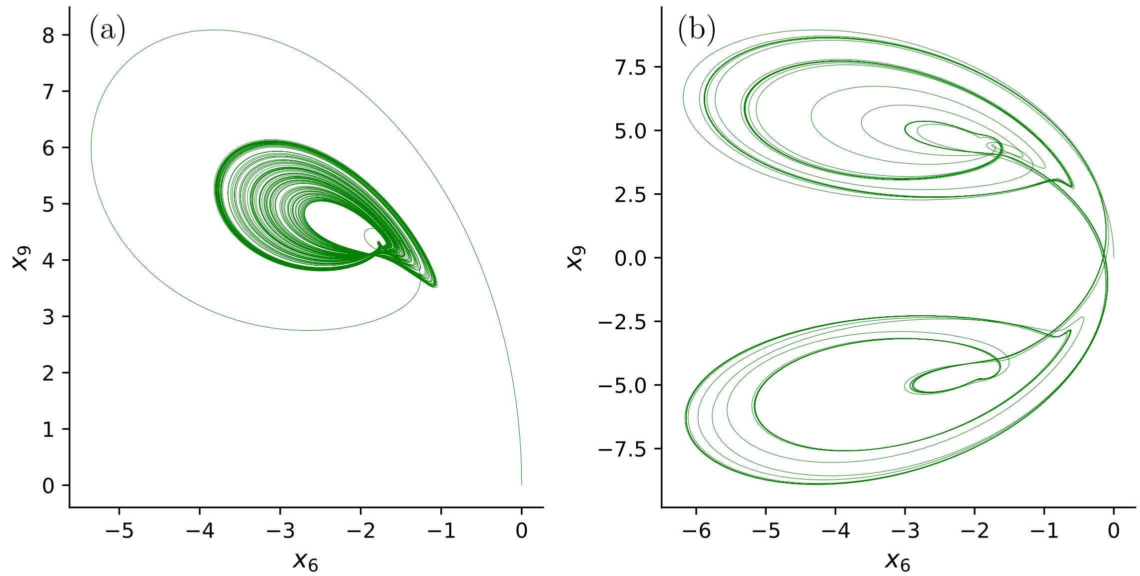

Figure 6.9: Trajectories projected on x_6-x_9 for $(\sigma, \rho) = (0.5, 14.22)$ (a) and $(\sigma, \rho) = (0.5, 15)$ (b).

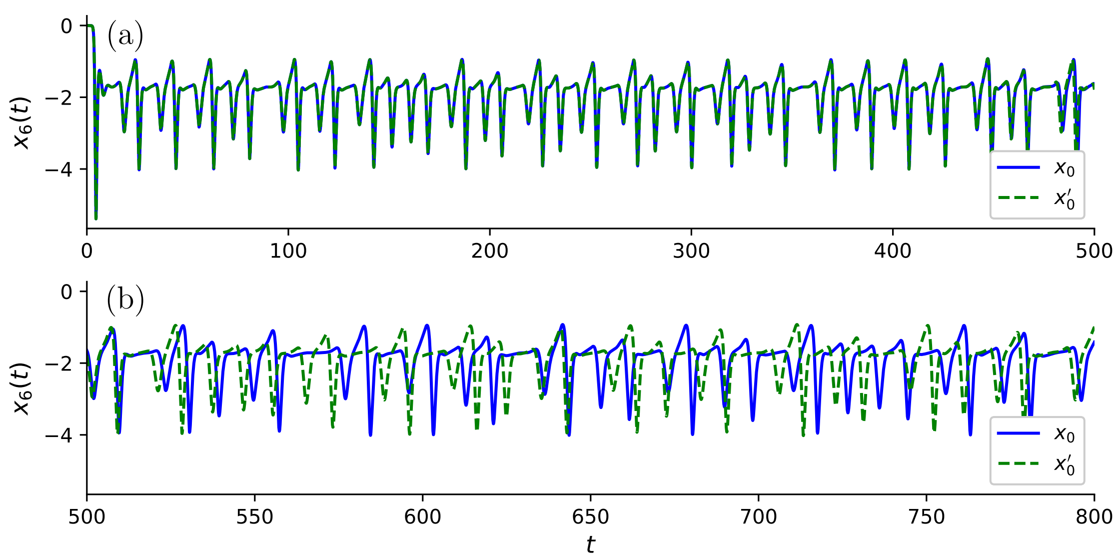

Figure 6.10: Time evolutions of two x_6-components over $T \in [0, 500]$ (a) and $T \in [500, 800]$ (b) obtained for $(\sigma, \rho) = (0.5, 14.3)$. The initial condition are spaced out by 10^{-7} in x_6 state.

6.3 Coordinate Transformation

For our project, it was necessary to include an additional parameter within the equations of both models given the fact that, a perturbation in a design variable representing a coordinate transformation does not lead the problem into bifurcation. In other words, it means that the attractor characterizing the system remains unchanged from a topological standpoint. The solutions are analogous to the case of hyperbolic systems where the perturbed system is topologically conjugate to the original one [21, 36]. To obtain the system containing this parameter it was necessary to consider the equations with the transformed coordinate $\mathbf{x}' = \mathbf{h}_\gamma(\mathbf{x}(t))$. The coordinate transformation $\mathbf{h}_\gamma$ represents a stretching or a compression of the evolution of the state spaces in a particular direction for $\gamma > 1$ or $\gamma < 1$, respectively. Differentiating this expression with respect to time and then using the chain rule and the inverse $\mathbf{h}_\gamma^{-1}$ we obtain the new set of equations containing γ. The most important justification for considering this new formulation lies in the fact that the shadowing direction could be known explicitly, enabling a verification of the theoretical prediction. Furthermore, the stretching (or the compression) does not affect the temporal dynamics of the problem. More details about this topic can be found in Ref. [21].

6.3.1 Lorenz System

For the Lorenz system, the coordinate transformation will be applied to the third state. From a practical point of view, to obtain the new set of equations, it is necessary to divide the interested state, $x_3(t)$, and multiply the third equation by γ.

$$\mathbf{h}_\gamma : (x_1, x_2, x_3) \to (x_1, x_2, \gamma x_3) \tag{6.5}$$

Thus, the new set of equations are

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \dot{x_1}(t) \\ \dot{x_2}(t) \\ \dot{x_3}(t) \end{bmatrix} = \begin{bmatrix} \sigma(x_2(t) - x_1(t)) \\ \rho x_1(t) - x_2(t) - x_1(t)x_3(t)/\gamma \\ \gamma x_1(t)x_2(t) - \beta x_3(t) \end{bmatrix} = \mathbf{f}(\mathbf{x}(t), \mathbf{p}) \tag{6.6}$$

For a trajectory defined by (6.6), the shadowing direction $\mathbf{y}_S(t) = [0, 0, z(t)]^\intercal$ is the solution of the sensitivity equations at $\gamma = 1$. Thus, the sensitivity can thus be found analytically as

$$\mathcal{J}^T_{\partial\gamma}(\mathbf{x}_0, \gamma) = \frac{1}{T}\int_0^T J_{\partial\mathbf{x}}(t)^\intercal \cdot \mathbf{y}(t)dt = \int_0^T x_3(t)dt = \mathcal{J}^T(\mathbf{x}_0, \gamma) \tag{6.7}$$

In panel (a) of figure 6.11 we present a common ($\gamma = 1$) and a compressed ($\gamma = 0.8$) in x_3-direction attractors. The panel (b) shows the time evolution of the third state for both systems.

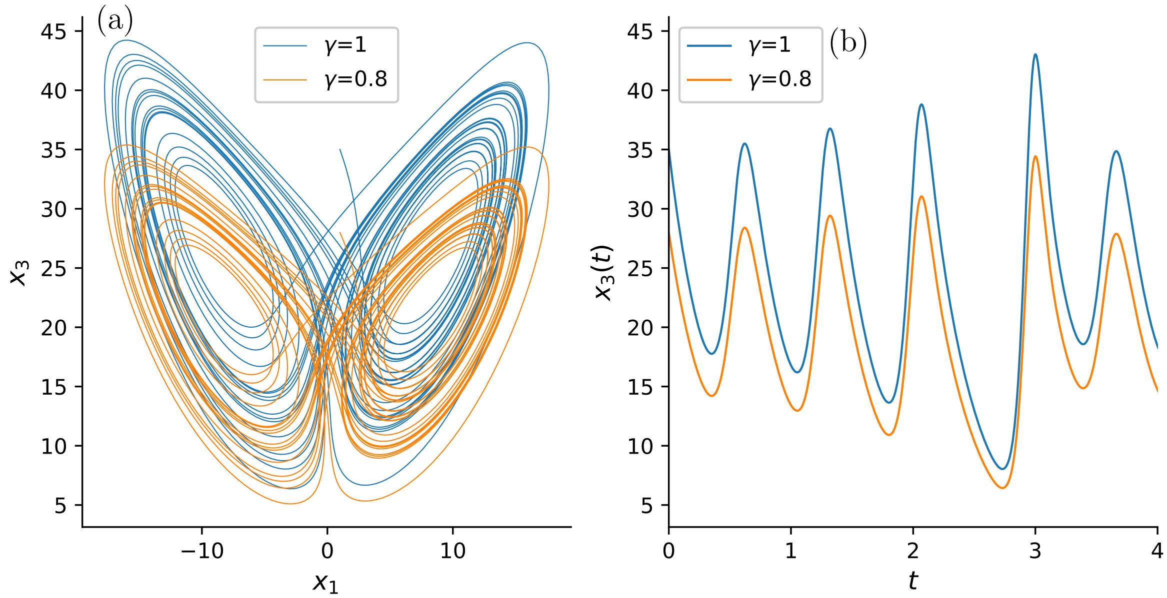

Figure 6.11: Lorenz attractor in plane projections x_1-x_3(a) with the non-perturbed and compressed orbit in x_3-direction. Panel (b) represents the time-evolution of their third components (b). The initial condition of the trajectory for $\gamma = 0.8$ is obtained by applying the transformation (6.5) to the initial condition $[1, 4, 35]$ of the trajectory for $\gamma = 1$.

We observe in figure 6.11 that the stretching affects only the third component of the attractor. The applied transformation also does not change the time scale of the system. In fact, as will be clearer in the chapter 7 concerning results, considering longer and longer trajectories, the time gradient $T_{dp}/T \to 0$.

6.3.2 Nine-Dimensional System

For the latter chaotic system, the coordinate transformation representing a stretching (or a compression) will be performed along the x_6-direction. This choice will be justified in detail in the following chapter. Temporarily neglecting motivation, the coordinate transformation $\mathbf{h}_\gamma$ for this system is defined by

$$\mathbf{h}_\gamma : (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9) \to (x_1, x_2, x_3, x_4, x_5, \gamma x_6, x_7, x_8, x_9) \tag{6.8}$$

which, as yet stated, represents a stretching or a compression of the evolution of the state spaces in x_6-direction for $\gamma > 1$ or $\gamma < 1$, respectively. To introduce this additional parameter into the equations, the analogous procedure described for the Lorenz equations is followed. The resulting system is

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -\sigma\beta_1 x_1(t) - x_2(t)x_4(t) + \beta_4 x_4^2(t) + \beta_3 x_3(t)x_5(t) - \sigma\beta_2 x_7(t) \\ -\sigma x_2(t) + x_1(t)x_4(t) - x_2(t)x_5(t) + x_4(t)x_5(t) - \sigma x_9(t)/2 \\ -\sigma\beta_1 x_3(t) + x_2(t)x_4(t) - \beta_4 x_2^2(t) - \beta_3 x_1(t)x_5(t) + \sigma\beta_2 x_8(t) \\ -\sigma x_4(t) - x_2(t)x_3(t) - x_2(t)x_5(t) + x_4(t)x_5(t) + \sigma x_9(t)/2 \\ -\sigma\beta_5 x_5(t) + x_2^2(t)/2 - x_4^2(t)/2 \\ -\beta_6 x_6(t) + x_2(t)x_9(t)\gamma - x_4(t)x_9(t)\gamma \\ -\beta_1 x_7(t) - \rho x_1(t) + 2x_5(t)x_8(t) - x_4(t)x_9(t) \\ -\beta_1 x_8(t) + \rho x_3(t) - 2x_5(t)x_7(t) + x_2(t)x_9(t) \\ \rho x_4(t) - x_9(t) - \rho x_2(t) - \dfrac{2x_2(t)x_6(t)}{\gamma} + \dfrac{2x_4(t)x_6(t)}{\gamma} + x_4(t)x_7(t) - x_2(t)x_8(t) \end{bmatrix} \tag{6.9}$$

The applied transformation is reflected in an effective compression of the attractor in x_6-direction, as figure 6.12a shows.

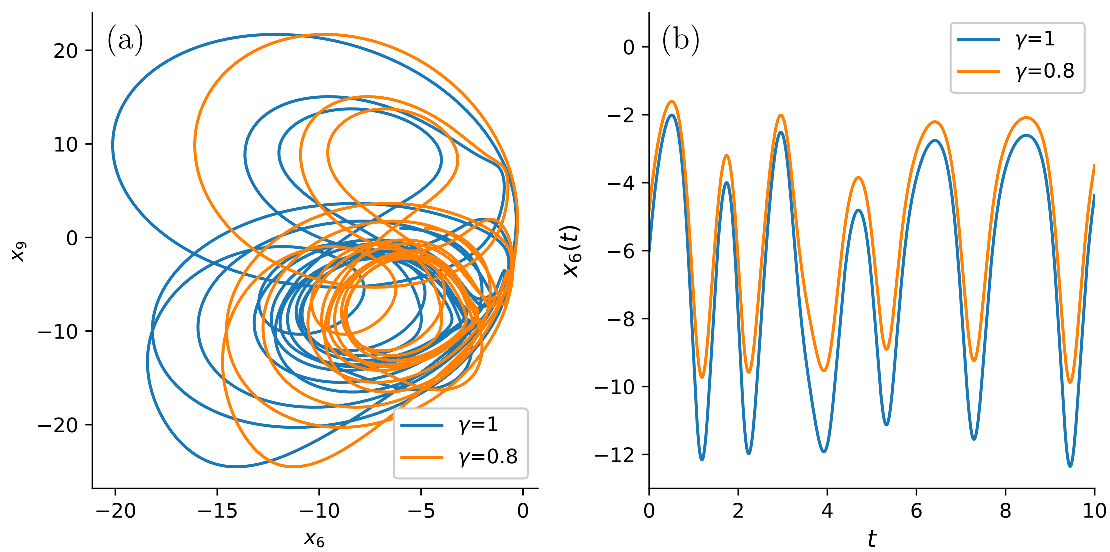

Figure 6.12: Phase space projections of the 9D attractor on the x_6-x_9 plane (a) with non-perturbed and compressed orbits in x_6-direction and time-evolution of their sixth components (b). The initial condition of the trajectory for $\gamma = 0.8$ is obtained by applying the transformation (6.8) to the initial condition of the trajectory for $\gamma = 1$.

It is necessary to underline that in the plot of figure 6.12b only the initial fragment of the time evolutions of the sixth states of the simulation presented in figure 6.12a are shown. This is justified because to observe the attractor of this chaotic system (panel (a)) it is necessary to march in time up to a value such that the actual dynamic of the shadow (panel (b)) could not have been clearly observed. The sensitivity of the system can, therefore, be calculated analytically as

$$\mathcal{J}^T_{\partial\gamma}(\mathbf{x}_0,\gamma) = \frac{1}{T}\int_0^T J_{\partial\mathbf{x}}(t)^\intercal \cdot \mathbf{y}(t)dt = \int_0^T x_6(t)dt = \mathcal{J}^T(\mathbf{x}_0,\gamma) \tag{6.10}$$

6.4 Linearisation

The formulated problems (4.38) and (4.47) still need the implementation of the Jacobian matrix containing the derivatives with respect to the state space and parameter considered. In the following subsections, the linearisation required for the Periodic Shadowing application will be presented for both systems.

6.4.1 Lorenz System

Considering (6.6), the linearisation with respect to the state space is

$$\dot{\mathbf{y}}(t) = \begin{bmatrix} \dot{y_1}(t) \\ \dot{y_2}(t) \\ \dot{y_3}(t) \end{bmatrix} = \begin{bmatrix} -\sigma & \sigma & 0 \\ \rho - x_3(t)/\gamma & -1 & -x_1(t)/\gamma \\ \gamma x_2(t) & \gamma x_1(t) & -\beta \end{bmatrix} \cdot \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} = \mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t),\mathbf{p}) \cdot \mathbf{y}(t) \tag{6.11}$$

The Jacobian containing the derivatives with respect to the parameter considered, representing the forcing functions for the non-homogeneous sensitivity equations, depends on which sensitivity analyses we will perform. In particular, we compute the sensitivity with respect to parameters γ and ρ. Thus, the forcing terms are respectively

$$\frac{\partial\mathbf{f}(\mathbf{x}(t))}{\partial\gamma} = \begin{bmatrix} 0 \\ \dfrac{x_1(t)x_3(t)}{\gamma^2} \\ x_1(t)x_2(t) \end{bmatrix} = \mathbf{f}_{\partial\gamma}(x(t)) \tag{6.12}$$

$$\frac{\partial\mathbf{f}(\mathbf{x}(t))}{\partial\rho} = \begin{bmatrix} 0 \\ x_1(t) \\ 0 \end{bmatrix} = \mathbf{f}_{\partial\rho}(\mathbf{x}(t)) \tag{6.13}$$

From a computational standpoint, once the nonlinear and linearised equations were implemented, the forcing term related to the problem that you want to solve is passed as input for the latter.

6.4.2 Nine-Dimensional System

Considering (6.9), the Jacobian matrix containing the linearisation with respect to the states becomes

$$\begin{bmatrix} -\sigma\beta_1 & -x_4 & \beta_3 x_5 & -x_2+2\beta_4 x_4 & \beta_3 x_3 & 0 & -\sigma\beta_2 & 0 & 0 \\ x_4 & -\sigma-x_5 & 0 & x_1+x_5 & -x_2+x_4 & 0 & 0 & 0 & -\sigma/2 \\ -\beta_3 x_5 & x_4-2\beta_4 x_2 & -\sigma\beta_1 & x_2 & -\beta_3 x_1 & 0 & 0 & \sigma\beta_2 & 0 \\ 0 & -x_3-x_5 & -x_2 & -\sigma+x_5 & -x_2+x_4 & 0 & 0 & 0 & \sigma/2 \\ 0 & x_2 & 0 & -x_4 & -\sigma\beta_5 & 0 & 0 & 0 & 0 \\ 0 & x_9\gamma & 0 & -x_9\gamma & 0 & \beta_6 & 0 & 0 & x_2-x_4\gamma \\ -\rho & 0 & 0 & -x_9 & 2x_8 & 0 & -\beta_1 & 2x_5 & -x_4 \\ 0 & x_9 & \rho & 0 & -2x_7 & 0 & -2-x_5 & -\beta_1 & x_2 \\ 0 & -\rho-\dfrac{2x_6}{\gamma}-x_8 & 0 & \rho+\dfrac{2x_6}{\gamma}+x_7 & 0 & \dfrac{-2x_2+2x_4}{\gamma} & x_4 & -x_2 & -1 \end{bmatrix} \tag{6.14}$$

The forcing functions for the non-homogeneous sensitivity equations for this system can be written as

$$\mathbf{f}_{\partial\gamma}(x(t)) = \left[0,0,0,0,0,x_2(t)x_6(t)-x_1(t)x_6(t),0,0,\frac{2x_2(t)x_3(t)}{\gamma^2}-\frac{2x_1(t)x_3(t)}{\gamma^2}\right] \tag{6.15}$$

$$\mathbf{f}_{\partial\gamma}(x(t)) = [0,0,0,0,0,0,-x_1,x_3,-x_2+x_4] \tag{6.16}$$

At this point it is possible to proceed with the calculation of the sensitivity with respect to γ and ρ for both chaotic systems. In the following chapter we present the sensitivity analysis results obtained by application of the Periodic Shadowing.

Chapter 7

Results

This section will present the most relevant results concerning the application of the Periodic Shadowing method on the chaotic systems described in the previous chapter. In particular, the chapter is organized as follows: the Lorenz and the Reiterer model will be described separately, explaining the results of the sensitivity analysis with respect to the parameters γ and ρ. In the last section, the convergence analysis performed will be detailed. All the results treated in this chapter refer to the tangent approach (4.38) of Periodic Shadowing.

7.1 Periodic Shadowing and Lorenz System

The application of the Periodic Shadowing allows obtaining solutions of a linearised problem that remain bounded as the time span is increased. In fact, recalling figure 4.1, the linearised problem arisen from conventional methods is not able to grasp the saturation of the system. The solution remains bounded over time and does not show the classical exponential growth. In figure 7.1 we compare the time evolution of the distance between the trajectories computed with conventional linearised methods and Periodic Shadowing. In particular, we select two trajectories initially spaced out 10^{-9} and we perform the integration with $\Delta t = 0.01$. We observe that even in the distance between the trajectory is comparable for small time span, the conventional sensitivity methods show an evolution of the distance between the trajectories which is high unrepresentative of the system behaviour. On the other hand, Periodic Shadowing allows to obtain a bounded solution for the time evolution of the distance.

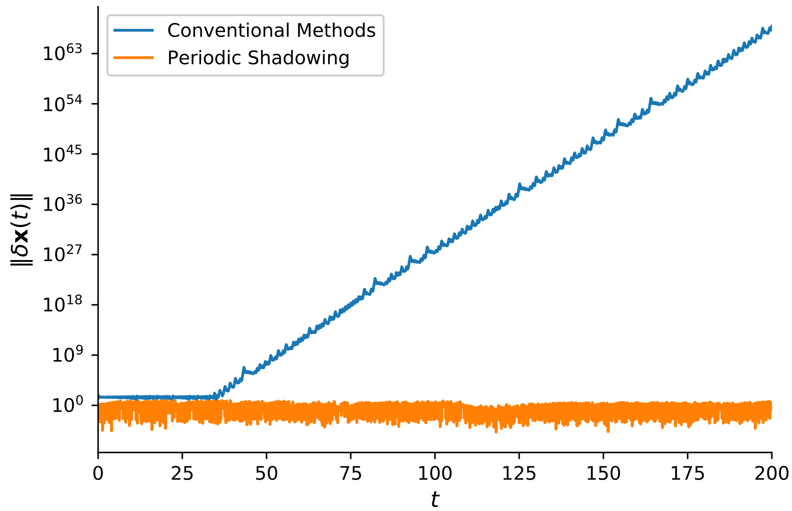

Figure 7.1: Time evolution of the distance between trajectories obtained with conventional sensitivity analysis methods and Periodic Shadowing.

7.1.1 Sensitivity Analysis with Respect to γ

Since the distance between the trajectories remains bounded, linearisation holds and therefore time evolutions of the solution obtained remain bounded as well. Panel (a) of

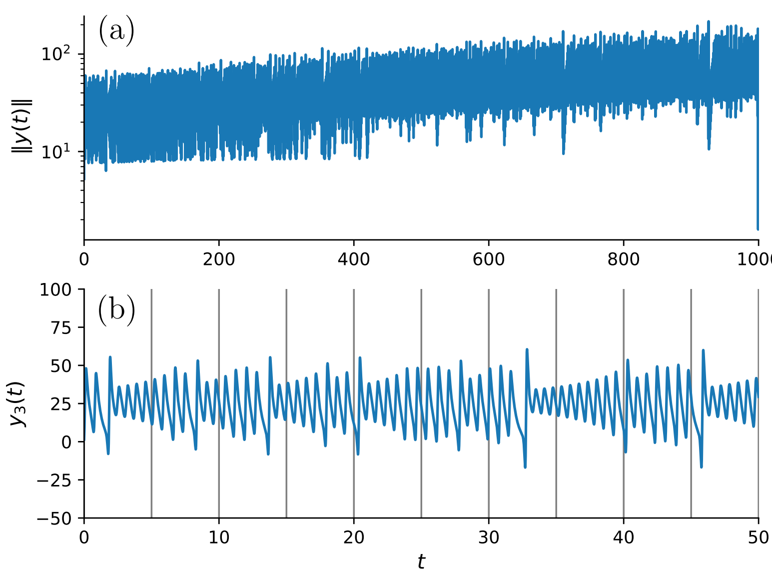

Figure 7.2: Time evolution of the normed solution $t \in [0, 1000]$ (a) and its third state over the fragment $t \in [0, 50]$ (b). Vertical lines in panel (b) represent the shooting points which subdivide the whole time interval.

figure 7.2 shows that the integration over a time span very long relative to the time scale of the system produces a bounded solution, whilst in panel (b) we focus the attention on the first fragment of the time evolution of the third component of $\mathbf{x}(t)$.
As hinted above, despite such systems being deterministic, due to their chaotic nature it was necessary to hold a statistical approach in order to actually validate the method. In fact, attention has been placed not only on the mean (median) values but also on the relative error bars (confidence intervals). For each time interval considered, the Periodic Shadowing method is applied to consecutive trajectory segments lying on the attractor thanks to the propagator function already used and defined in chapter 5 in order to analyse a different initial condition for each iteration. Results of the statistical distribution of $\mathcal{J}^T_{d\gamma}$ are reported in figure 7.3. Panel (a) shows the normalized distribution of $\mathcal{J}^T_{d\gamma}$, panel (b) depicts a close-up of the right tail of the distribution. The value C, subtracted to $\mathcal{J}^T_{d\gamma}$, is the sensitivity obtained for long integration time. The samples are obtained by repeating the sensitivity calculation over 10^6 different initial conditions.

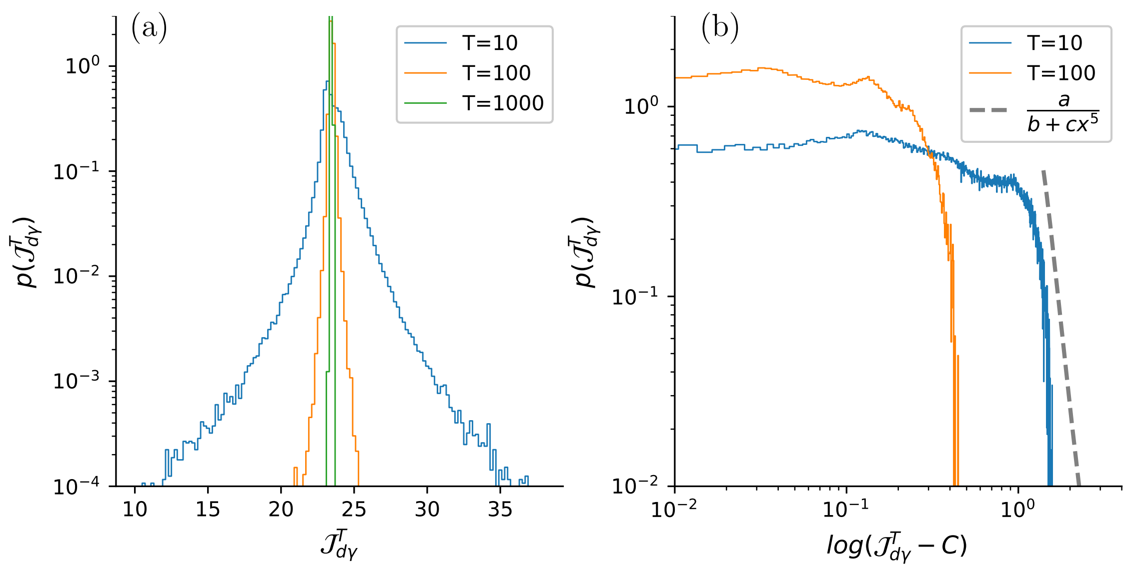

Figure 7.3: Normalized distribution of the sensitivity $\mathcal{J}^T_{d\gamma}$ for increasing time spans (a) and a close-up of the right tail of the full distribution (b).

The increase of time interval causes a reduction in the dispersion of results. The fast drop of the right tail of the full probability distributions, presented in a log-log plot to highlight the asymptotic trend, means that both mean and standard deviation converge as the number of samples is increased. In fact, analysing the law of decay of tails, it is possible to state that it is not a heavy-tailed distribution (with, for instance a decay

law $\sim a/(b + cx^2)$). If the distribution were heavy-tailed, as it will happen for the 9-state system, it is necessary to select the median and interquartile range as statistical properties.
The results of the sensitivity calculations are compared with the approximation of the gradients obtained with centred second-order finite difference (FD) scheme

$$\mathcal{J}_{d\gamma}^{T,FD}(\mathbf{x}_0) = \frac{\mathcal{J}^T(\mathbf{x}_0, \gamma + \Delta\gamma) - \mathcal{J}^T(\mathbf{x}_0, \gamma - \Delta\gamma)}{2\Delta\gamma} \tag{7.1}$$

with $\Delta\gamma = 0.5$ and results are shown in figure 7.4. The arithmetic averages of the sensitivity obtained from multiple repetitions of both algorithms are reported in panel (a), while panel (b) shows the sample standard deviation. The number of samples N

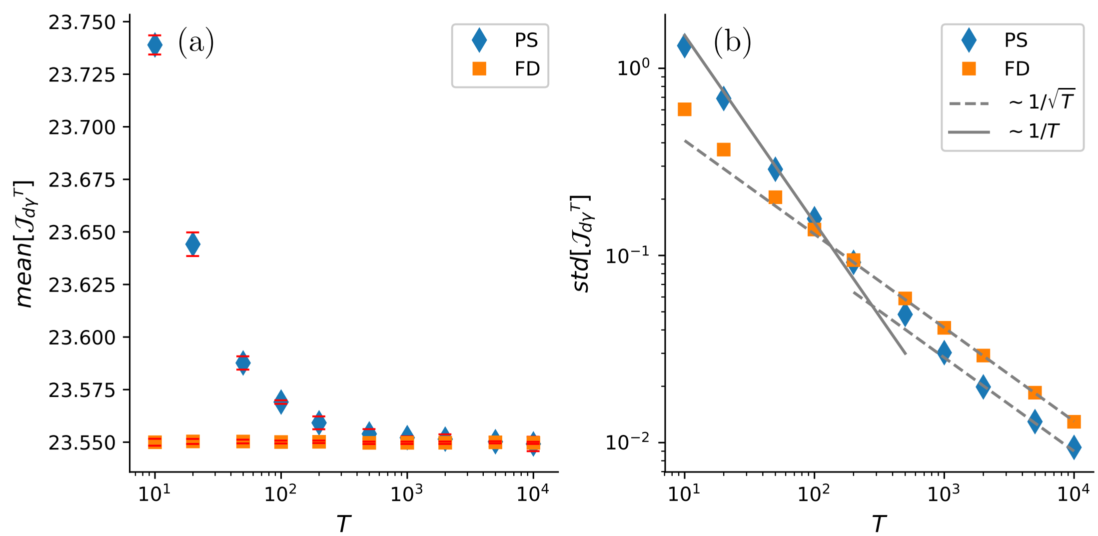

Figure 7.4: Mean (a) and standard deviation (b) of the sensitivity $\mathcal{J}_{d\gamma}^T$ computed for the Lorenz system using Periodic Shadowing. The line in (b) represents the predicted standard deviation scaling of finite-difference and Periodic Shadowing algorithms.

used is sufficiently high ($\geq 10^6$) to decrease opportunely the error bars related to the confidence level of 99.73%. In particular the error bars shown in figure 7.4a (and in all the following figures inherent the sample mean) enclose the values

$$\mathrm{mean}[\mathcal{J}_{d\gamma}^T] \pm 3\frac{\mathrm{std}[\mathcal{J}_{d\gamma}^T]}{\sqrt{N}}$$

where we use mean[$\star$] and std[$\star$] to indicate sample mean and standard deviation. As can be seen in figure 7.4a, the averaged sensitivity value obtained through Periodic

Shadowing converges to the gradient obtained from the finite differences approximation as the integration time increases, as predicted in (6.7). For short trajectories the average effect of the initial condition is reflected in a gap reducing by considering longer integration intervals. Regarding panel (b) of figure 7.4, the trend of the sample standard deviation is in line with what has been stated in the error analysis section of Ref. [21]. In that work it is proven that the error is composed by the sum of two different terms. The former is the random error of the finite-time sensitivity $\mathcal{J}_{dp}^{T,S}$ of the shadowing trajectory while the latter, is the error of the same quantity with respect to the sensitivity $\mathcal{J}_{dp}^{T,P}$ computed using the periodic solution of (4.38). Thus, the global error

$$\mathcal{E}^T(\mathbf{x}_0, p) = \mathcal{E}_0^T(\mathbf{x}_0, p) + \mathcal{E}_1^T(\mathbf{x}_0, p) = \frac{C_0^T(\mathbf{x}_0, p)}{\sqrt{T}} + \frac{C_1^T(\mathbf{x}_0, p)}{T} \tag{7.2}$$

where $C_0^T(\mathbf{x}_0, p)$ and $C_1^T(\mathbf{x}_0, p)$ are statistically distributed according to a certain probability density function (PDF) that is independent of T, but only depends on the dynamics (4.1) and the choice of the observable. The rapid $1/T$ decay implies that for some sufficiently large T the shadowing error $\mathcal{E}_1^T(\mathbf{x}_0, p)$ will be, on average, smaller than that of the random error $\mathcal{E}_0^T(\mathbf{x}_0, p)$ and the global error is mostly dominated by the first term.

We now consider the gradient $T_{d\gamma}/T$, related to how the parameter perturbation affects the time scale of the system. Panel (a) of figure 7.5 shows the mean of $T_{d\gamma}/T$ computed for the same amount of different initial conditions whilst panel (b) depicts its standard deviation and its decay rate.

It can be observed that the ratio $T_{d\gamma}/T$ decays towards zero. In fact, the parameter γ is a simply coordinate transformation, so it does not change the time scales of the problem. In other words, the most important result of the figure 7.5 is that the stretched system has the same time scale of the non-perturbed one when $T \to \infty$.

To understand the behaviour of Periodic Shadowing, we represent a quantity that would allow us to verify the real influence of considering longer and longer integration time. Our choice fell on

$$\boldsymbol{\chi}(\mathbf{x}_0, T) = \frac{1}{N} \sum_{i=1}^{N} \|\mathbf{y}(t_i)\|^2 \tag{7.3}$$

where N is the number of samples (different initial conditions) and $\mathbf{y}(t_i)$ is the solution evaluated in the i-shooting point. In other words, this quantity represents the mean of the squared norm solutions evaluated at the shooting points. We consider 10^6 samples and the results are reported in figure 7.6. The convergence of $\boldsymbol{\chi}(\mathbf{x}_0, T)$ as $T \to \infty$

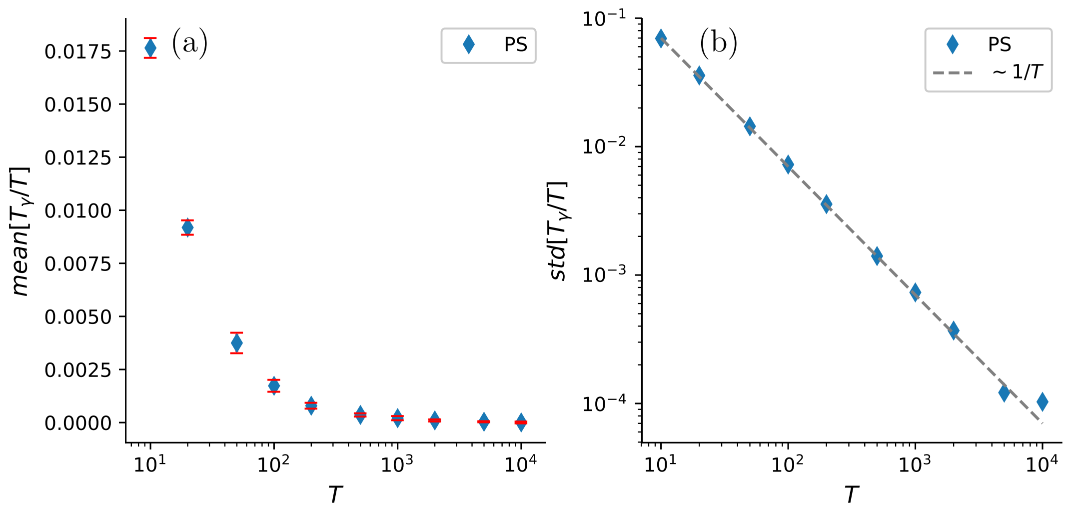

Figure 7.5: Arithmetic average (a) and standard deviation (b) of the gradient $T_{d\gamma}/T$ as a function of time span computed for 10^6 iteration of the Periodic Shadowing.

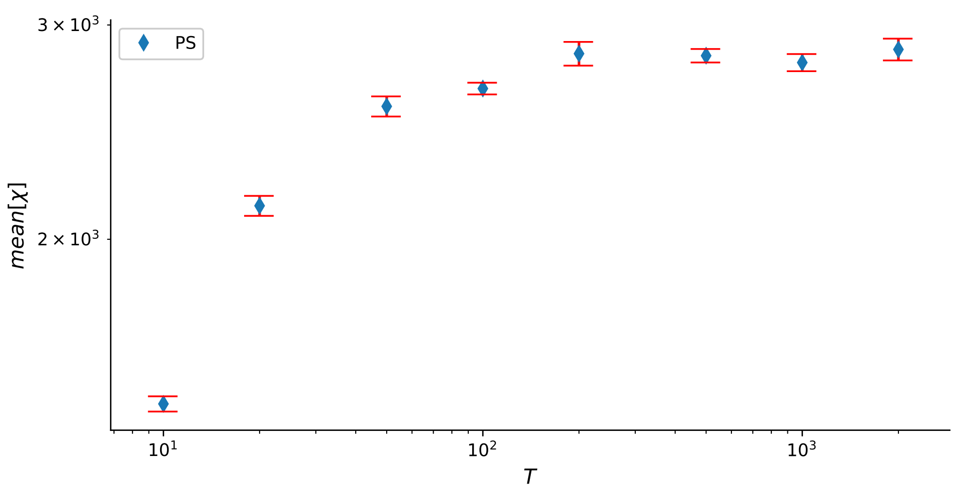

Figure 7.6: Averaged $\boldsymbol{\chi}(\mathbf{x}_0, T)$ as a function of time span obtained by repeating the algorithm for 10^6 different initial conditions.

depicted in the figure 7.6 is supported by the fact that a perturbation of the parameter γ does not violate the assumption of Periodic Shadowing.

7.1.2 Sensitivity analysis with Respect to ρ

The application of the Periodic Shadowing to evaluate the sensitivity of the observable of interest with respect parameter ρ leads to a solution which remains bounded in time as obtained in figure 7.2 for a perturbation with respect to γ. In figure 7.7 we present the time evolution of the normedsolution and the time evolution of its third component, in panel (a) and (b), respectively.

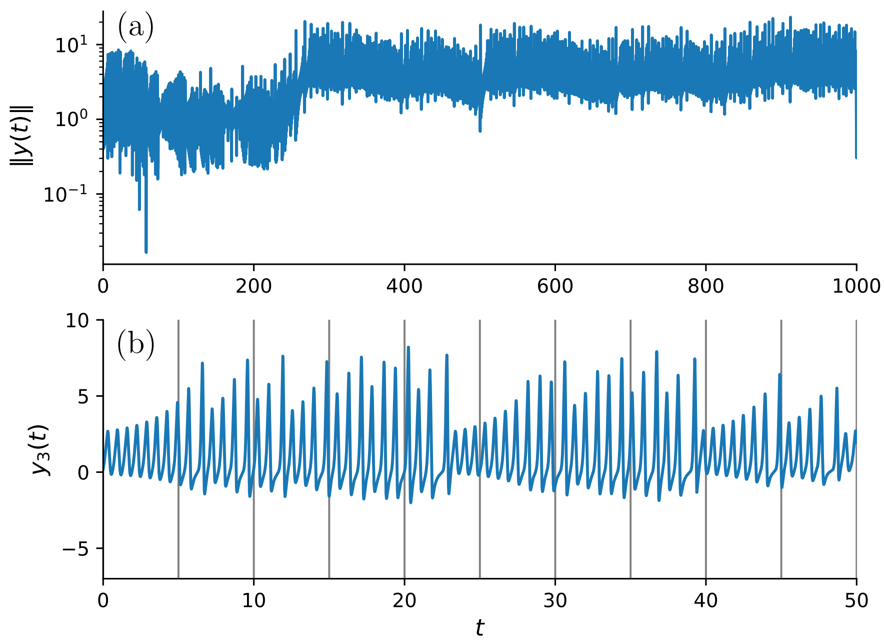

Figure 7.7: Time evolution of the normed solution norm for $t \in [0, 1000]$ (a) and its third state over the fragment $t \in [0, 50]$ (b). Vertical lines in panel (b) represent the shooting points which subdivide the whole time interval.

The results obtained in figure 7.7 show a bounded behaviour of the solution instead of presenting characteristic numerical explosion underlined in Ref. [4]. In panel (b) we depict only the first 50 time units of the evolution in order to make clearly visible the matching boundary condition in the shooting points, represented by the vertical gray lines.

Distributions of $\mathcal{J}^T_{d\rho}$ for increasing time span are report in figure 7.8. Panel (a) shows the normalized distribution whilst panel (b) depicts a close-up of the right tail. We notice that the increase of time interval causes a reduction in the dispersion of results. In panel (b) are presented the right tail of the full distribution for two different time spans: the way in which the decay of the tails takes place, therefore, admits the convergence of sample mean and standard deviation.

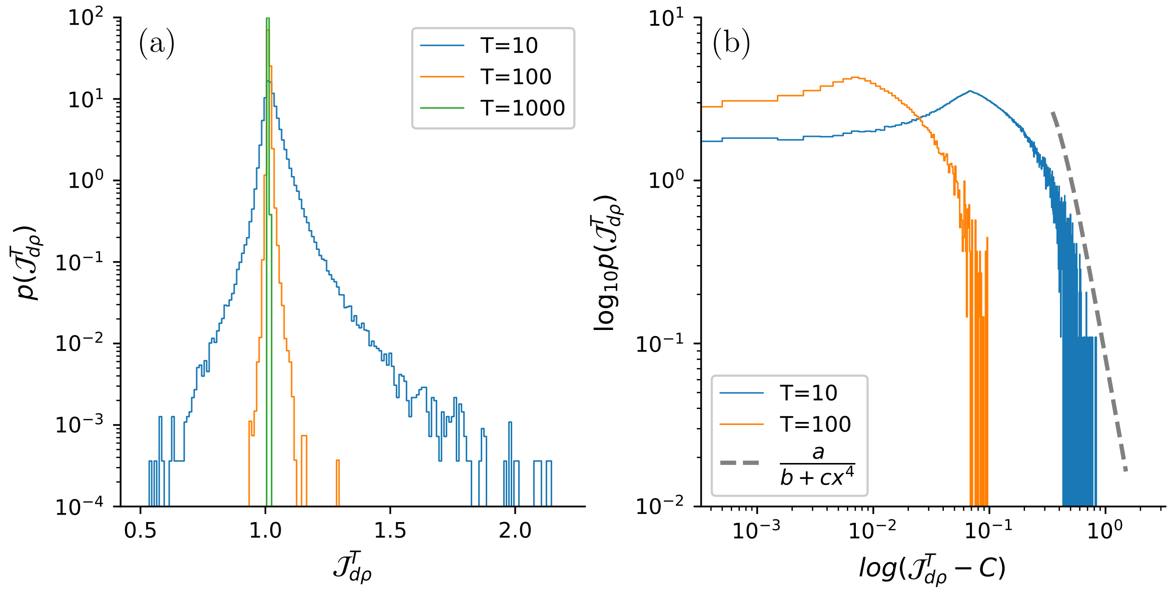

Figure 7.8: Normalized probability distribution of $\mathcal{J}^T_{d\rho}$ for increasing time span (a) and a close-up of the right tail of the full distribution (b).

In this analysis the results will not be compared exclusively with those computed from the finite-difference approximation (7.1) with $\Delta\rho = 0.5$, but also with a digitization of the data points obtained from the Least Square Shadowing (LSS) method reported in figure 6 of Ref. [7]. However, we underline that in Ref. [7] the algorithm was applied to only 10 different initial conditions, while for the method presented here the iterations performed for 10^6 samples. Results are reported in figure 7.9. The arithmetic average of the sensitivity obtained from multiple repetitions of the various algorithms is reported in panel (a) of figure 7.9, whilst panel (b) shows the sample standard deviation. The values obtained through the present algorithm underline that the arithmetic average of the sensitivity converges to a value around $\mathcal{J}^T_{d\rho} \simeq 1.017$. Similar values of the gradients can be found also in the digitalization of the Least Square Shadowing method in Ref. [7] and in the results section of Refs. [8, 21, 37]. It is immediately evident the presence of a gap in figure 7.9a between the sensitivity computed through shadowing-based methods and the finite difference approximation of the gradient. It was carefully checked that this difference is not led by wrong choice of the step size of the numerical integration or in the $\Delta\rho$ applied to obtain the approximation. This feature represents a crucial point of this analysis but, at the same time, provides valuable indications for future developments. Further details are provided in the final chapter. Concerning how the standard deviation of the sensitivity evolves as a function of the time span, figure 7.9b

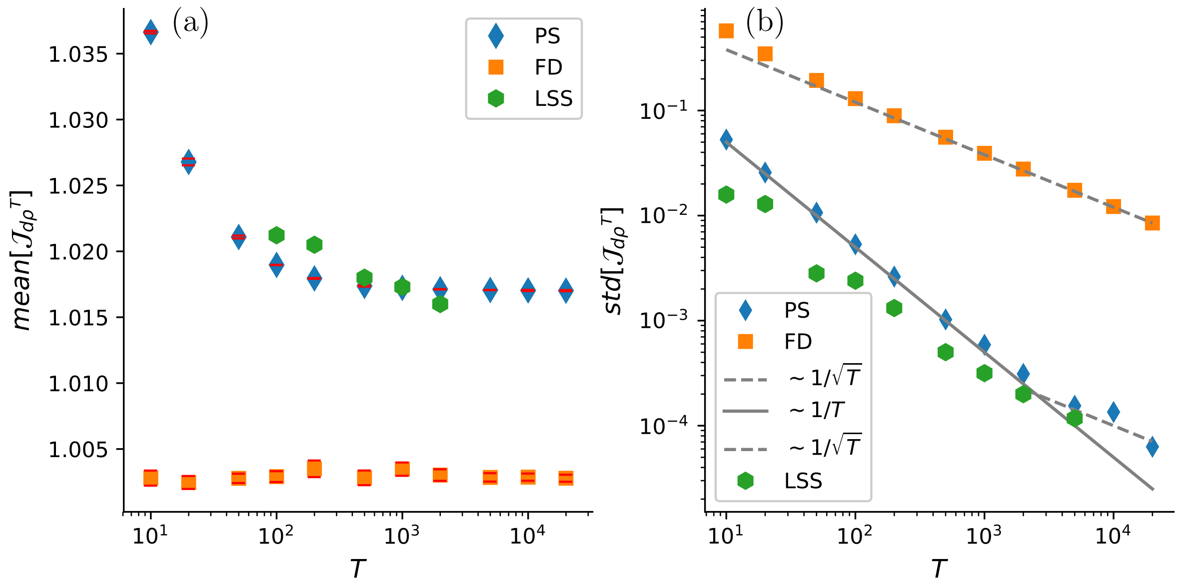

Figure 7.9: Arithmetic average (a) and standard deviation (b) of the sensitivity $\mathcal{J}_{d\rho}^{T}$ computed for the Lorenz system using Periodic Shadowing (PS), finite-difference (FD) and Least Square Shadow (LSS). The lines in (b) represent the predicted standard deviation scaling.

shows that it initially decays as $1/T$ and asymptotically as $1/\sqrt{T}$. The threshold at which the decay rate changes is around 5000 time units, in agreement with the results obtained from the digitization of results of the Least Square Shadowing. This threshold is higher than what observed in figure 7.4b for the sensitivity with respect to γ.
As already performed for the sensitivity analysis with respect to γ, the averaged gradient $T_{d\rho}/T$, related to the time scales of the two trajectory, is evaluated and expounded. In panel (a) of figure 7.10 the time gradient $T_{d\rho}/T$ is reported, whilst its standard deviation is reported in panel (b) of the same figure. The key result of figure 7.10a is that the gradient $T_{d\rho}/T$ converges, if $T \to \infty$, to the well defined value $-2.42 \cdot 10^{-2}$, which is the same obtained from periodic orbits in Ref. [8], characterized by no shadowing error. As regards the standard deviation, panel (b) shows that the convergence to the asymptotic value is achieved at a $1/T$ rate initially and then at a $1/\sqrt{T}$ rate for time spans longer than 200 time units.
We now consider the quantity $\boldsymbol{\chi}(\mathbf{x}_0, T)$. We select two random initial condition to compute $\boldsymbol{\chi}(\mathbf{x}_0, T)$ for increasing time span and the results are reported in figure 7.11. We observed in panel (a) that $\chi(\mathbf{x}_0, T)$ grows weakly ($\sim \sqrt{T}$) when considering longer time intervals while panels (b) and (c) provide a detail of the same quantity for the time interval $t \in [520, 740]$,$t \in [673, 683]$, respectively. The importance of panel (b) and

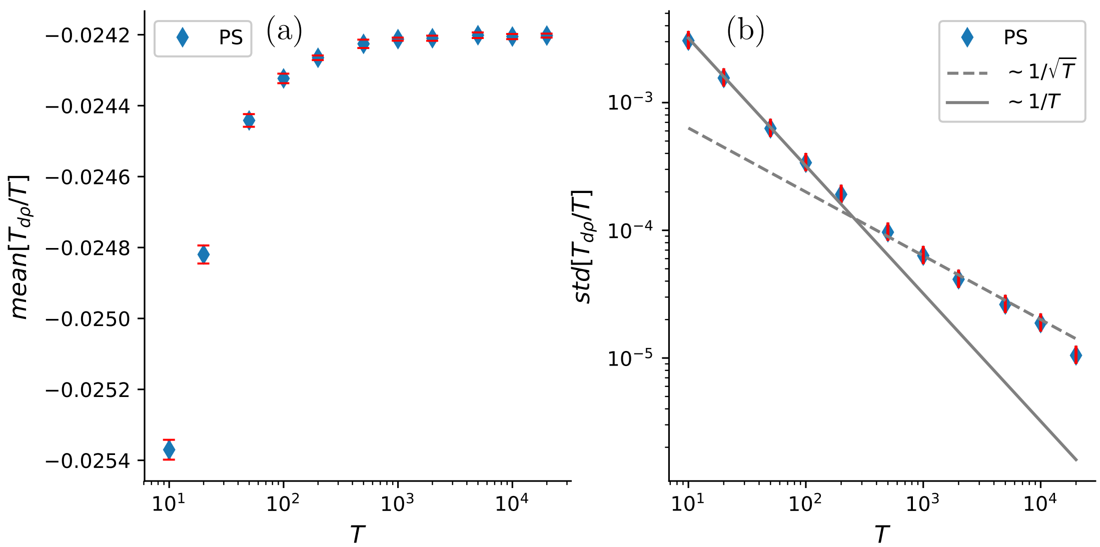

Figure 7.10: Arithmetic average (a) and standard deviation (b) of the period gradient $T_{d\rho}/T$ computed for the Lorenz system using Periodic Shadowing. The lines in (b) represent the standard deviation decay ratio for this term.

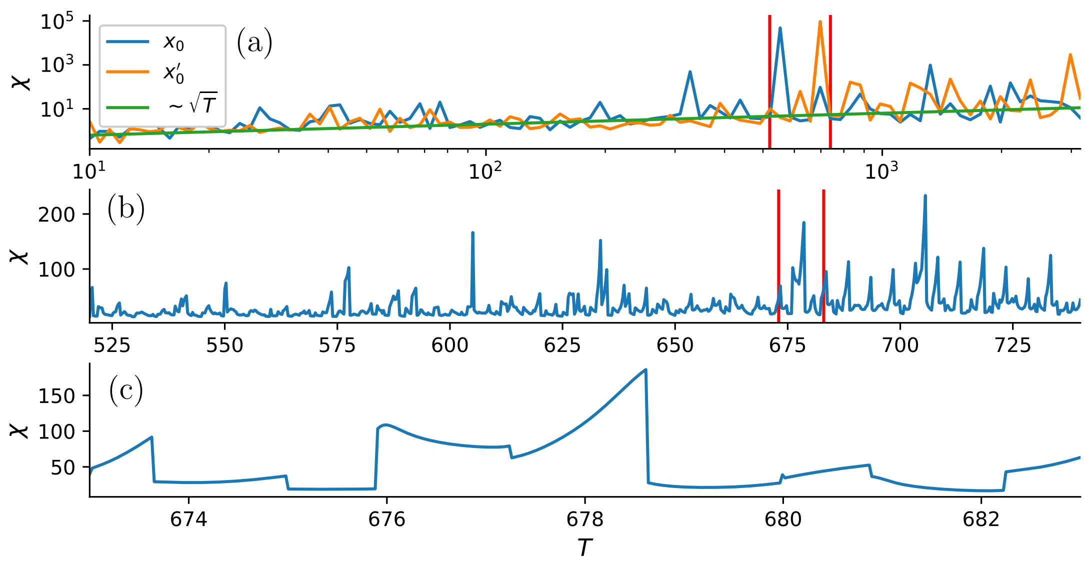

Figure 7.11: Time evolution of $\chi(\mathbf{x}_0, T)$ for two different initial condition (a) and a particular fragments over $t \in [520, 740]$ (b) and $t \in [673, 683]$ time unit. The green line in panel (a) represents how mean squared norms increase considering different time span.

(c) of figure 7.11 is to check that the solution does not present asymptotes considering increasing time span. To underline that this growth is actually characteristic of the interaction chaotic system/Periodic Shadowing and not a trend brought by the initial conditions considered, we repeat the algorithm for several (10^6) different initial conditions. Results of the averaged $\boldsymbol{\chi}(\mathbf{x}_0, T)$ as function of time span are reported in figure 7.12. We observe a weak growth ($\sim \sqrt{T}$) in averaged quantity unlike what presented

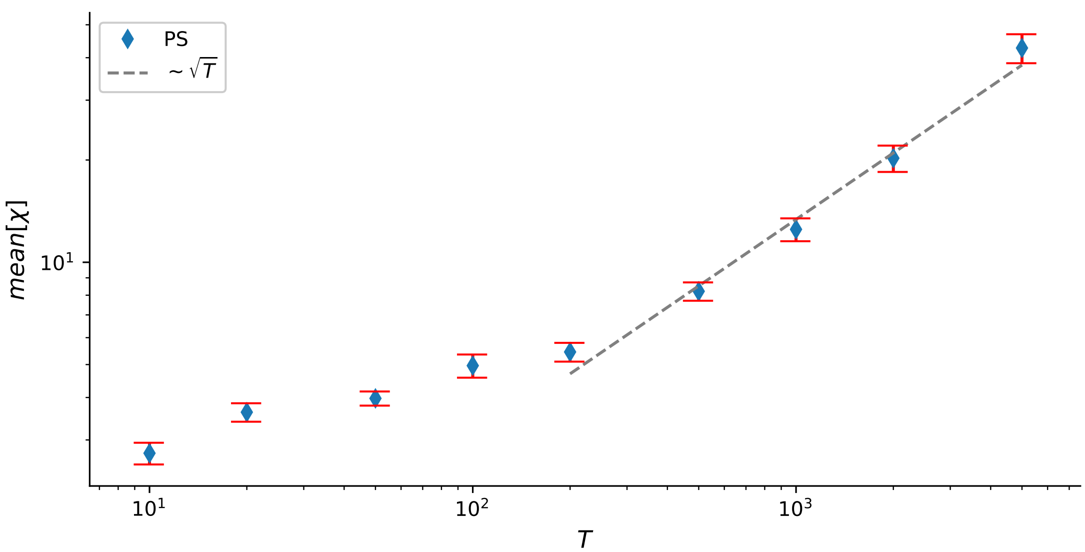

Figure 7.12: Averaged $\boldsymbol{\chi}(\mathbf{x}_0, T)$ as a function of time span obtained by repeating the algorithm for 10^6 different initial conditions.

for the sensitivity with respect to γ. This trend is very low if compared to the numerical explosion of the solution obtained using conventional methods (remember figure 2.2). The results obtained by the application of the Periodic Shadowing on chaotic system arising by the Lorenz equations, and summarized in this section, show a good response of the method for both sensitivity evaluated with respect to γ and ρ. In fact, if on the one side the sensitivity evaluated with respect to the coordinate transformation matches the analytical solution for $T \to \infty$, on the other hand, the sensitivity with respect to a ρ perturbation is supported by other methods already present in the bibliography.

7.2 Periodic Shadowing and Reiterer Equations

This application allows us to go further and test our method on a chaotic system of greater complexity. Similarly to what has been done with the Lorenz system, the sensitivity with respect to γ and ρ perturbations will be computed. However, unlike before, a brief premise about non-linear analyses of the system is necessary.

7.2.1 Nonlinear Analysis

The choice to apply the coordinate transformation $\mathbf{h}_\gamma$ on the third component was a guided operation due to the double symmetry in x_1 and x_2 of the Lorenz system. In this case it was necessary to carry out a more complete analysis in order to understand on which state the transformation could applied and so, on which one evaluate the sensitivity. In particular, we perform an analysis concerning the relationship between the long time average of the states and the parameter ρ, and results are reported in figure 7.13. Remember that in this case ρ is defined differently with respect to the previous problem from an analytical standpoint, however, it always refers to the reduced Rayleigh number which is proportional to the temperature difference between the plates of the fluid cell.

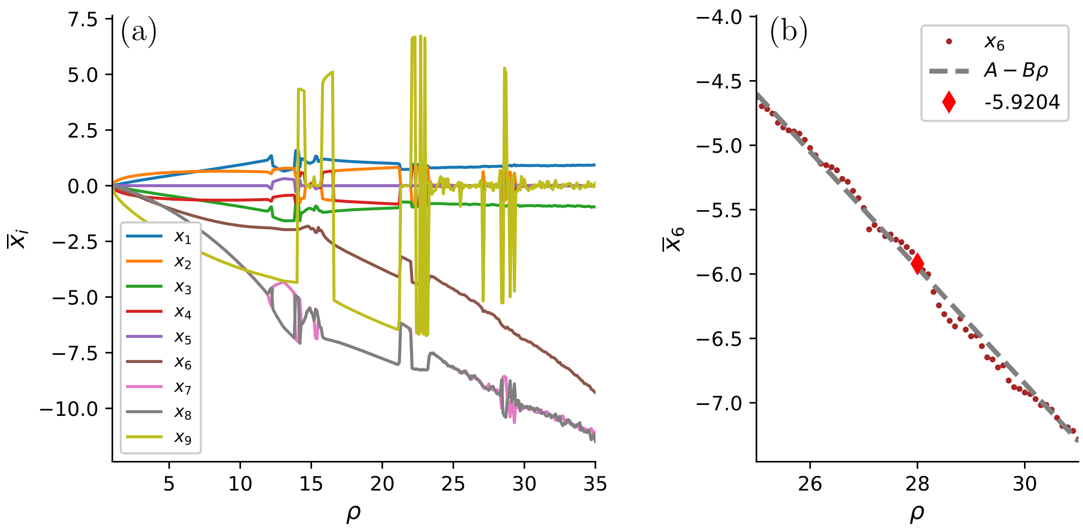

Figure 7.13: Evolution as a function of ρ of the time averaged nine states x_i of Reiterer equations (a) and a close-up of the interval $\rho = [25, 31]$ (b) for $T = 20000$ time units. The red marker in panel (b) shows the value of the average for $\rho = 28$.

Panel (a) shows the trend of the long time average states whilst panel (b) focuses on sixth state and on the range containing the value $\rho = 28$ considered for the further analysis. For the results presented in figure 7.13, ρ is varied with a step size $\Delta\rho = 0.1$ and the finite time average is evaluated over $T = 20000$ time units. We observe that the trends of x_1, x_2, x_3, x_4, x_5 and x_9 present an asymptotically behaviour characterized by a constant value of the average. In fact, neglecting some ρ values which induce strong perturbations in the finite time average (as, for instance, it happens for the x_9 state considering $\rho = [22, 23]$), evaluating the effect of a variation of parameter ρ for these states would not be a quantity of particular interest since they are not particularly influenced by it. Among the remaining states, the choice is made according to the following principles. All three remaining states (x_6, x_7, x_8) show a strong influence with respect to the parameter ρ. However, since in chapter 6 all the representations of the attractor had been performed on projection plans containing x_6, we decided to use the sixth as state on which to calculate the sensitivity. The choice to use this state instead of the others is certainly questionable but, thusly, we can also provide a continuity of analysis with what is done in Ref. [23]. Furthermore, our interest is on the response of the algorithm, not the physics of the problem. For this reason, from now, the sixth state will be selected for both coordinate transformation and sensitivity analysis calculation.

7.2.2 Sensitivity Analysis with Respect to γ

The time evolution of the sixth state remains bounded when Periodic Shadowing method is applied to the 9-states system to evaluate the sensitivity of the observable of interest with respect γ. In figure 7.14 we present the time evolution of the solution and the time evolution of its sixth component, in panel (a) and (b), respectively. For an example we select trajectory of length $T = 1000$ time units and originating from the origin. This result allows us to verify that Periodic Shadowing method responds correctly also for a greater complexity chaotic system. Since the trajectories remain close for whole integration time, the correct sensitivity can be computed.
In figure 7.15 we show the probability distribution of $\mathcal{J}^T_{d\gamma}$ obtained by applying the method for 10^6 different initial conditions and for different time spans. Looking at the figure 7.15a, we observe that the dispersion of $\mathcal{J}^T_{d\gamma}$ decreases with the increase of the time span. However, the close-up of the right tail depicted in panel (b) shows a decay law characteristic of the heavy-tailed distribution. For this reason, it is necessary to underline that the statistical quantities taken into consideration in this analysis are the median and the interquartile range instead sample mean and standard deviation. The median is the value separating the higher half from the lower half of a data sample whilst the interquartile range (indicated as *iqr* in the following), is a measure of statistical dispersion, being equal to the difference between 75^{th} and 25^{th} percentiles. In fact, considering this particular type of distribution, the *Central Limit Theorem* (CLT) does not work [38] and so the computed sample mean and standard deviation would be very misleading. Using the median instead of the mean also implies a new definition of

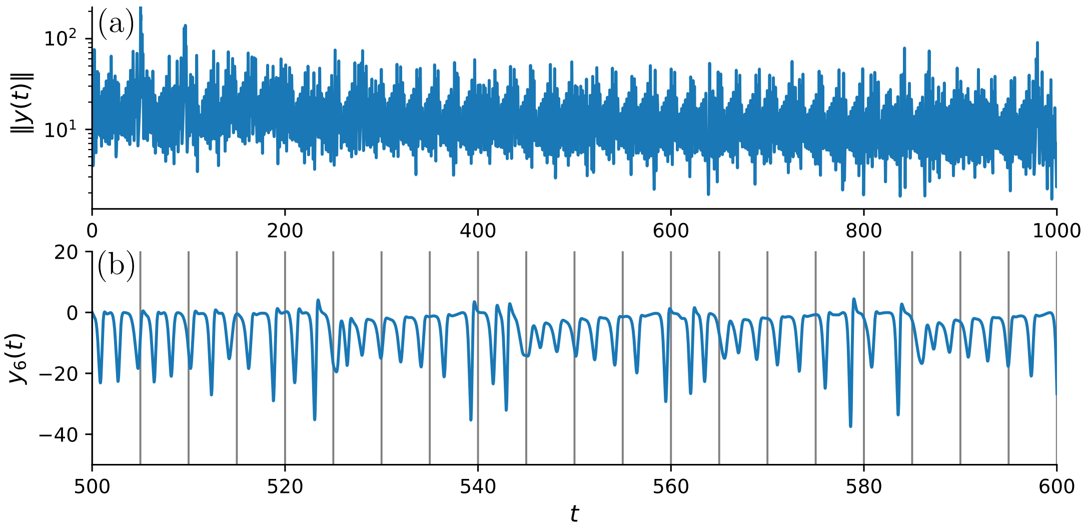

Figure 7.14: Time evolution of the normed solution $t \in [0, 1000]$ (a) and its sixth state over the fragment $t \in [500, 700]$ (b). Vertical lines in panel (b) represent the shooting points which subdivide the whole time interval.

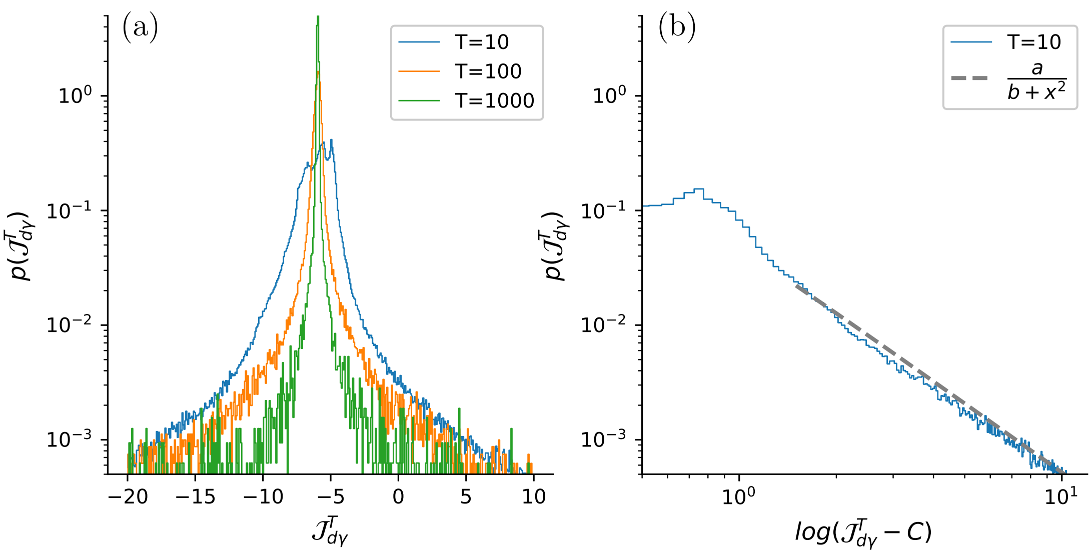

Figure 7.15: Normalized probability distribution of $\mathcal{J}^T_{d\gamma}$ for increasing time span (a) and a close-up of the right tail of the full distribution (b).

the error bar. In fact in this case the confidence interval is 95%, which can be written as

$$\mathcal{J}^T_{d\gamma}[\text{lower}] < \mathcal{J}^T_{d\gamma} < \mathcal{J}^T_{d\gamma}[\text{higher}] \tag{7.4}$$

where *lower* and *higher* represent the

$$\begin{cases} \text{lower} = \dfrac{N}{2} - \dfrac{1.96\sqrt{N}}{2}\text{th element} \\ \text{higher} = \dfrac{N}{2} + \dfrac{1.96\sqrt{N}}{2}\text{th element} \end{cases} \tag{7.5}$$

of the sorted vector containing N-$\mathcal{J}_{d\gamma}^{T}$ elements [39].
At this point it is possible to compare the median of the sensitivity computed for different initial conditions and the approximate value of the gradient obtained with centred second-order finite difference scheme (7.1) with $\Delta\gamma = 0.5$. Results are reported in figure 7.16. The plots show the median of $\mathcal{J}_{d\gamma}^{T}$ and the interquartile range of $\mathcal{J}_{d\gamma}^{T}$ as a function of time span, panel (a) and (b) respectively. As noticeable in figure 7.16a the

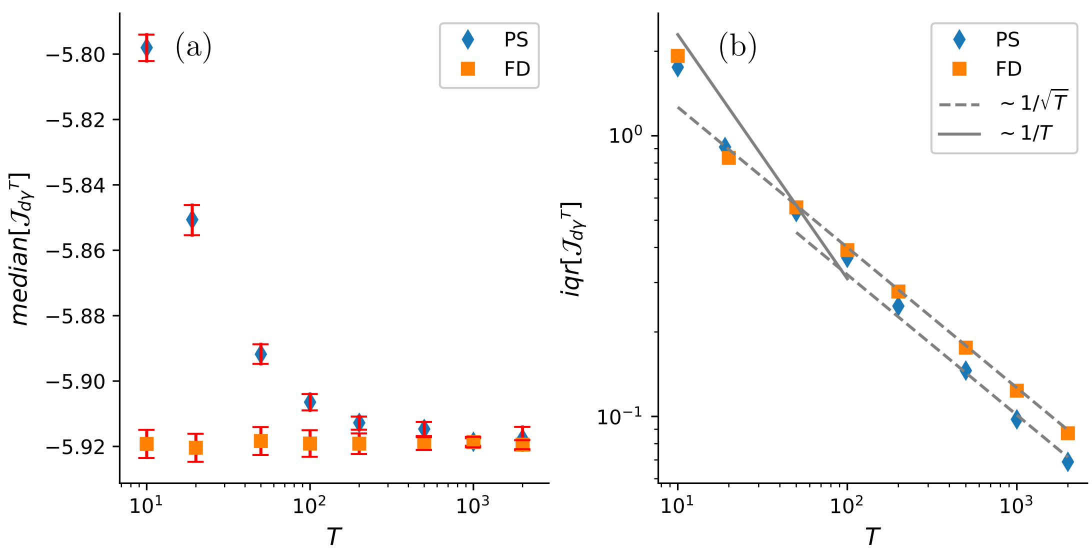

Figure 7.16: Median (a) and interquartile range (b) of the sensitivity $\mathcal{J}_{d\gamma}^{T}$ computed for the three-dimensional cells in natural convection using Periodic Shadowing. The line in (b) represents the decay of the interquartile range of the finite-difference and Periodic Shadowing algorithms.

statistical (10^6 different initial conditions on the trajectory) convergence of the time averaged sensitivity with respect to parameter γ is therefore also verified for a system of greater complexity compared to Lorenz equations. For increasing T the sensitivity converges to ≈ -5.92, value already presented in figure 7.13b and also obtained with the second-order finite-difference approximation.
Regarding how the time scales of the systems are related to each other as a function of the integration time, it is necessary to plot $T_{d\gamma}/T$ quantity for different time span. Results are reported in figure 7.17.

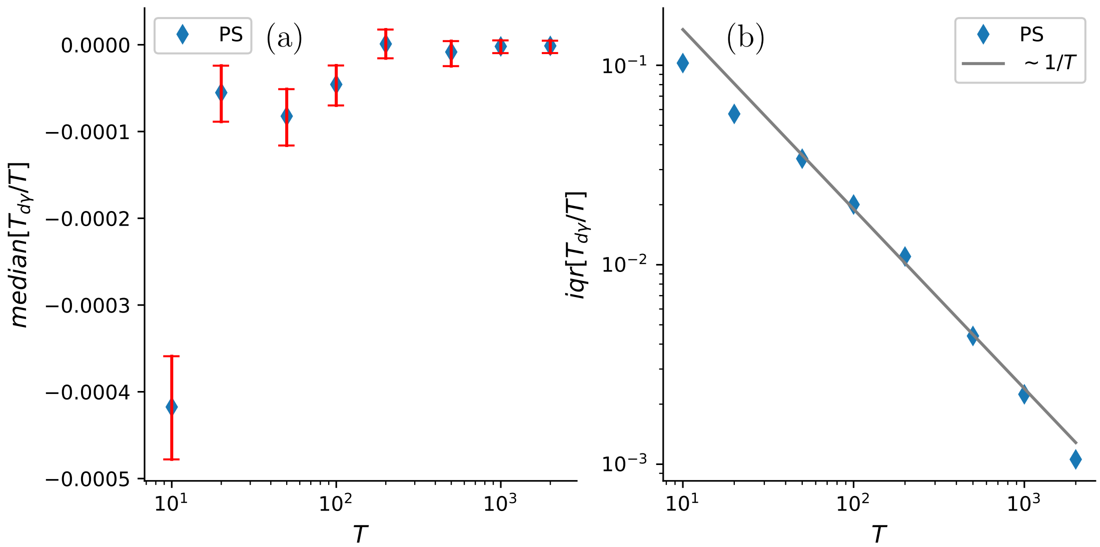

Figure 7.17: Median (a) and interquartile range (b) of the time gradient $T_{d\gamma}/T$ as a function of the time span. The line in (b) represents the decay of the interquartile range of this gradient.

As obtained for the sensitivity with respect to γ for the Lorenz system, the statistical quantity converges to zero if $T \to \infty$, because, as previously stated, a perturbation in this parameter does not affect the time scale of the problem.
To observe how the solution behaves, the quantity $\boldsymbol{\chi}(\mathbf{x}_0, T)$ for different time span is presented. To construct this result, plotted in figure 7.18, two random initial conditions are selected and the time span is increased for logarithmic spaced time in the range $T = [10, 1000]$. Panel (b) and (c) of figure 7.18 show the presence of asymptotes in $\boldsymbol{\chi}(\mathbf{x}_0, T)$ as a function of time span. In particular we depict in panel (c) the asymptotic value found approximatively at $T = 156.305$ for that initial condition. To justify the trend of panel (a), we now consider the median value of $\boldsymbol{\chi}(\mathbf{x}_0, T)$ obtained by repeating the algorithm for 10^3 different initial conditions. Results are reported in figure 7.19. The error bar are represent the 95% confidence interval and they are constructed using (7.5). We observe in figure 7.19 a convergence of $\boldsymbol{\chi}(\mathbf{x}_0, T)$ as $T \to \infty$. This behaviour is analogous to what found for the sensitivity evaluated for Lorenz system.

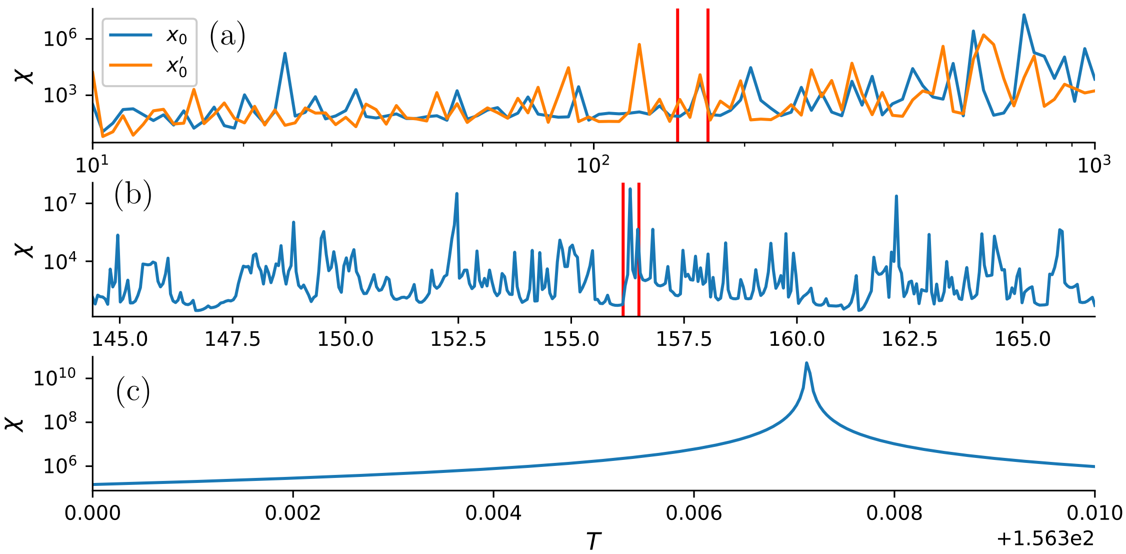

Figure 7.18: Value of χ evaluated in the shooting points for two different initial condition (a) and a particular fragments over $T \in [144, 166]$ (b) and $T \in [156.3, 156.31]$ time units.

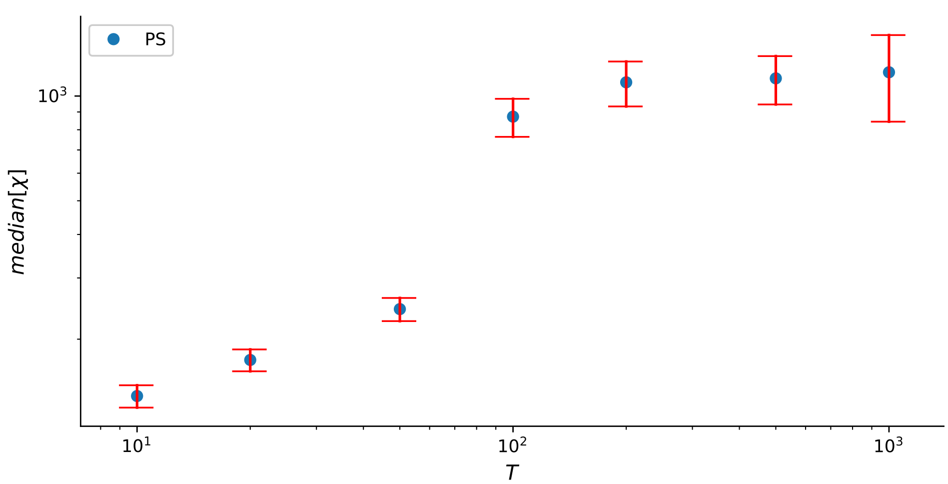

Figure 7.19: Averaged $\boldsymbol{\chi}$ as a function of time span obtained by repeating the algorithm for 10^6 different initial conditions.

7.2.3 Sensitivity Analysis with Respect to ρ

We now discuss sensitivity analysis with respect to the parameter ρ. The time evolution of the solution obtained with Periodic Shadowing and the sixth state are presented in figure 7.20. We select a trajectory of length $T = 10000$ time units and originating in

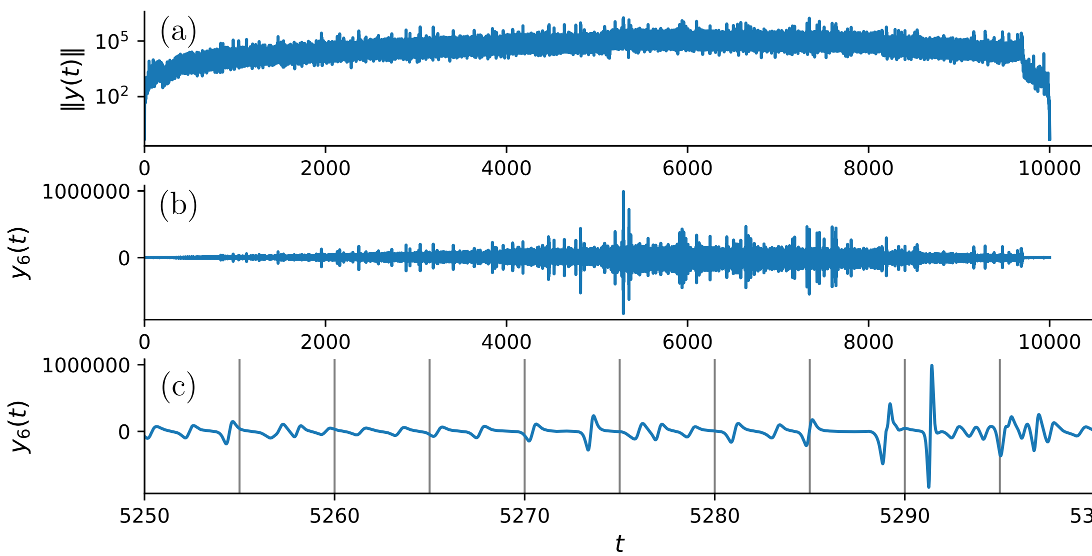

Figure 7.20: Time evolution of the complete solution (a) of its sixth component (b) and a detail of y_6 over $t = [5250, 5300]$ time units. The vertical grey lines in panel (c) represent the shooting points.

the origin. We observe in panels (a) and (b) that the solution and the sixth component remain bounded for the whole time span. However, we note in panel (b) that the sixth component is characterized by the presence of peak values significantly high although the periodic boundary condition $\mathbf{y}(0) = \mathbf{y}(T)$ and the continuity in the shooting points (observable in panel (c)) are satisfied. This behaviour can be arisen by the lack of hyperbolicity of the system led by perturbation in parameter ρ.
As performed in previous section, the whole discussion is subordinated to the analysis of the probability distribution of $\mathcal{J}^T_{d\rho}$. The normalized probability distribution of $\mathcal{J}^T_{d\rho}$ for increasing time span is presented in panel (a) of figure 7.21, whilst panel (b) shows a close-up of the right tail in a log-log plot. In particular, in panel (a) we observe a behaviour contrary to what has been observed so far because the dispersion does not reduce when the time span is increased. This phenomenon is also coupled to the fact that the close-up in panel (b) shows a decay law characteristic of heavy-tailed distribution with power-law $a/(b + cx^2)$. This kind of distributions goes to zero slower than the previous presented, which means there are many outliers with very high values. For

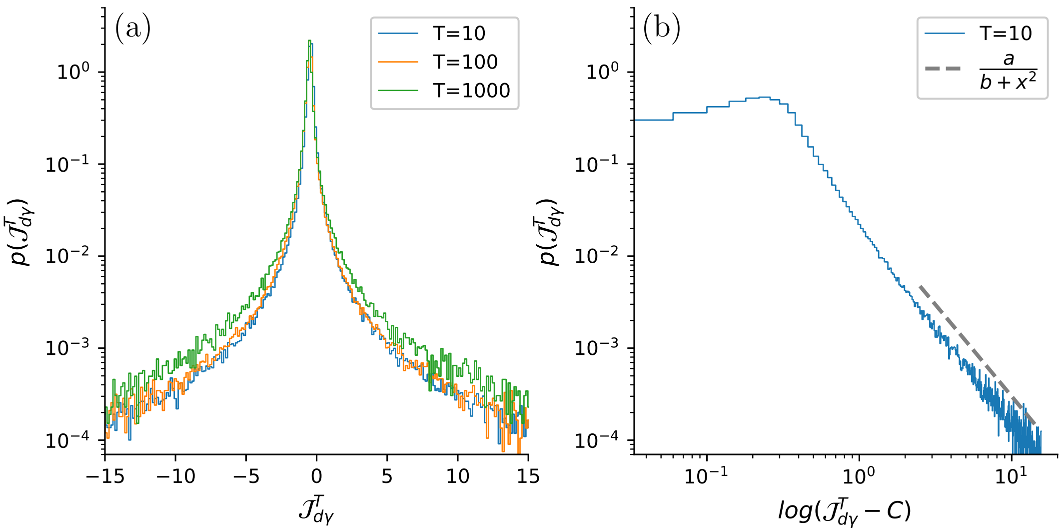

Figure 7.21: Normalized probability distribution of $\mathcal{J}^T_{d\gamma}$ for increasing time span (a) and a close-up of the right tail of the full distribution (b).

this reason, median and interquartile range will be chosen as statistical quantity to be studied instead sample mean and standard deviation.
The sensitivity of the observable of interest with respect ρ computed for different initial condition is presented as a function of time span considered in figure 7.22. The approximation of the gradient obtained with a centred second-order finite-difference scheme ($\Delta\rho = 0.5$) is also presented. The red errors bars represent the 95% confidence interval, obtained with (7.5) for 10^6 different initial condition. Panel (a) of figure 7.22 shows that the sensitivity converges to the value approximately equals to ≈ -0.471. On the other hand panel (b), depicts the noticeable increase of the interquartile range of $\mathcal{J}^T_{d\rho}$ computed with Periodic Shadowing. In particular, we observe that before $T = 100$ time units the iqr$\mathcal{J}^T_{d\rho}$ grows very weakly ($\sim T^{0.05}$), but after this threshold, it grows more rapidly as $\sim T^{0.7}$. This results is coherent with the non-decreasing dispersion of the full distribution shown in panel (a). As occurred for the Lorenz system, the sensitivity solutions obtained with the Periodic Shadowing method are affected by a bias error with respect to the value obtained through finite differences.
Regarding how the time scales of the systems are related to each other as a function of the integration time, it is necessary to plot $T_{d\rho}/T$ quantity for different time span. Results are reported in figure 7.23. We observe that the time gradient between the reference and perturbed trajectories has a weaker convergence to the value -0.0625. As observed in panel (b) of figure 7.22, the interquartile range grows considering longer

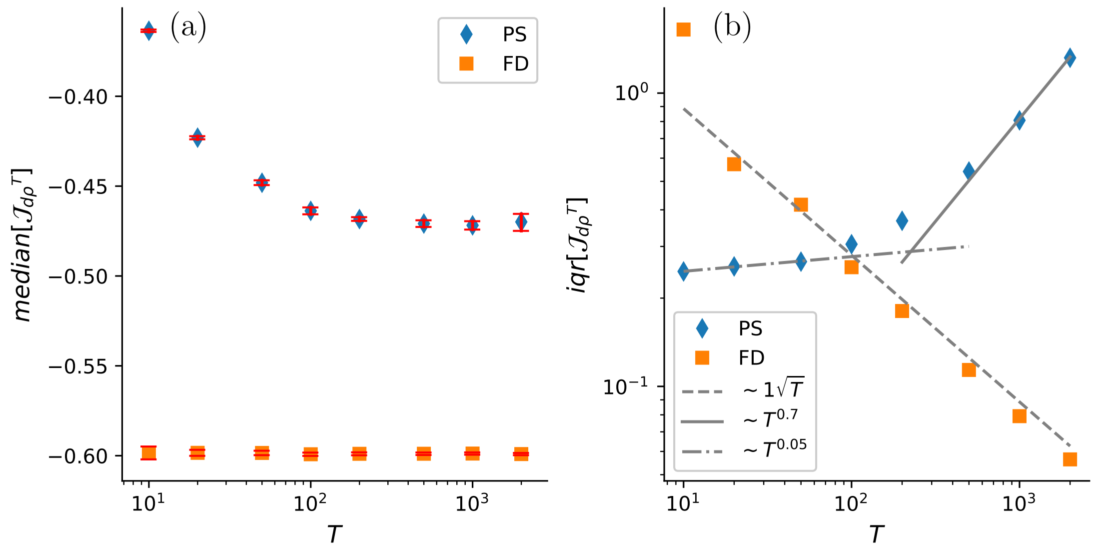

Figure 7.22: Median (a) and interquartile range (b) of the sensitivity $\mathcal{J}_{d\rho}^T$ computed for the three-dimensional cells in natural convection using Periodic Shadowing. The gray line in panel (b) shows the decay/growth of the interquartile range.

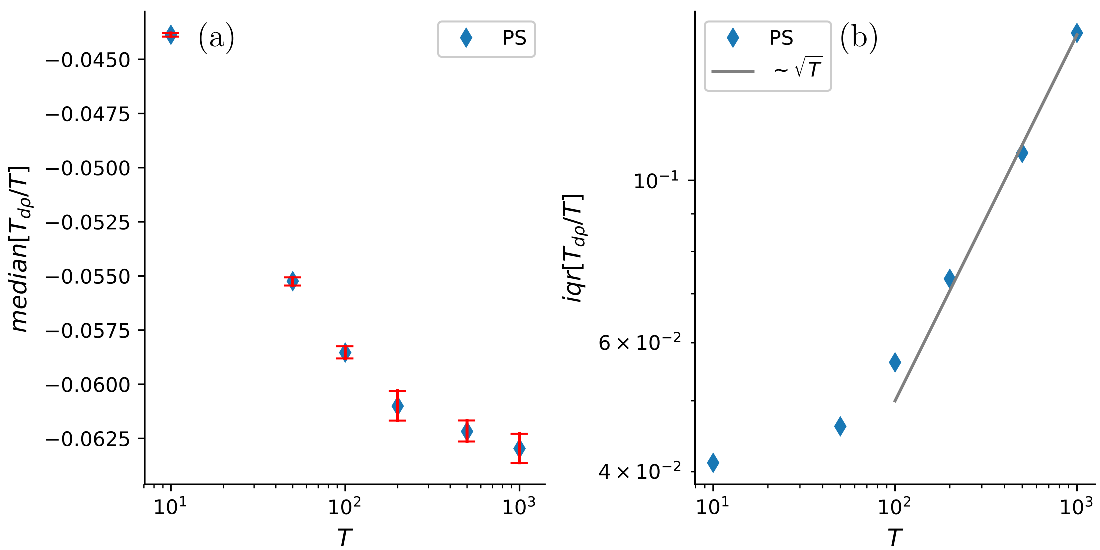

Figure 7.23: Median (a) and interquartile range (b) of the time gradient $T_{d\rho}/T$ as a function of the time span obtained by repeating the algorithm for 10^6 different initial conditions. The gray lines in panel (b) represent the growth of the interquartile range of $T_{d\rho}/T$.

time span. In particular, we observe that after 100 time units the interquartile range grows as $\sqrt{T}$. This occurs because the normalized distributions of the sensitivity present

an increasing dispersion if longer intervals are considered.
We now consider $\boldsymbol{\chi}(\mathbf{x}_0, T)$. In figure 7.24 we present the values assumed by $\boldsymbol{\chi}(\mathbf{x}_0, T)$ (for two different initial condition in panel (a)) considering increasing time span up to 1000 time units. Panels (b) and (c) represent a close-up of the range $T \in [48, 78]$ and $T \in [57.288, 57.290]$ time units, respectively.

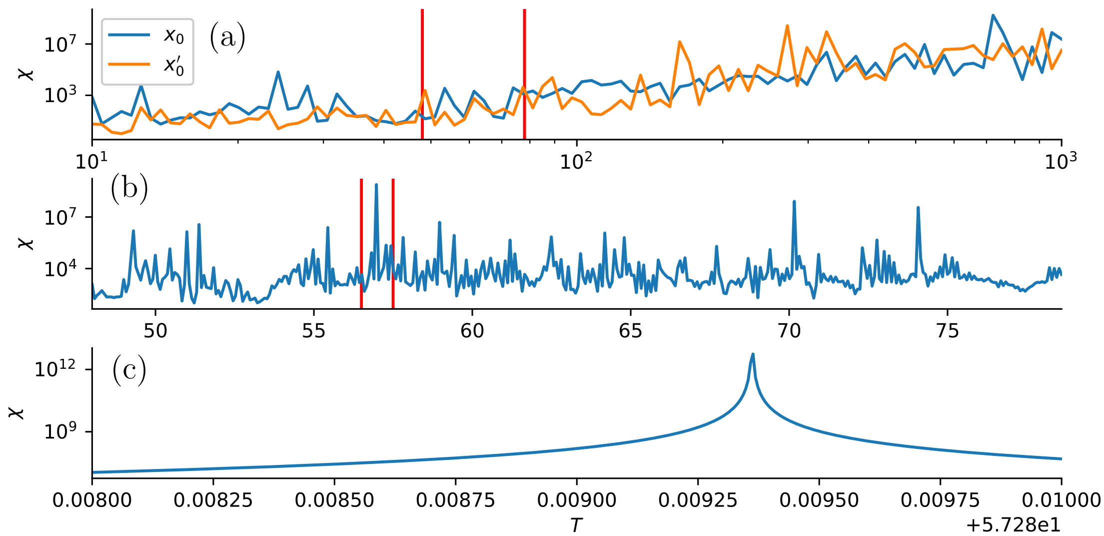

Figure 7.24: Value of $\boldsymbol{\chi}(\mathbf{x}_0, T)$ for two different initial conditions (a) and a particular fragments over $T \in [48, 78]$ (b) and $T \in [57.288, 57.290]$ time unit. The green line in panel (a) represents how mean squared norms increase considering different time span.

The importance of panel (b) and (c) of figure 7.24 is to check if $\boldsymbol{\chi}(\mathbf{x}_0, T)$ presents asymptotes in the solution. In particular we depict in panel (c) the asymptotic value found approximatively at $T = 57.2893$ for that initial condition. To underline that this growth is actually characteristic of the interaction chaotic system/Periodic Shadowing and not a trend brought by the initial conditions considered, we repeat the algorithm for several (10^5) different initial conditions. Results of the averaged $\boldsymbol{\chi}(\mathbf{x}_0, T)$ as function of time span are reported in figure 7.25. We observe that the quantity $\boldsymbol{\chi}(\mathbf{x}_0, T)$ grows rapidly ($\sim T^{2.7}$) as a function of time span. In particular, also in this case the threshold after which it is possible to recognize the growth rate is near $T = 100$ time units.
This application of Periodic Shadowing to a chaotic system characterised by greater complexity than the one arising from the Lorenz equations manifests itself in a complete solution (figure 7.20) that results with higher values than those presented in figure 7.7, but at the same time it remains bounded between values significantly smaller than those

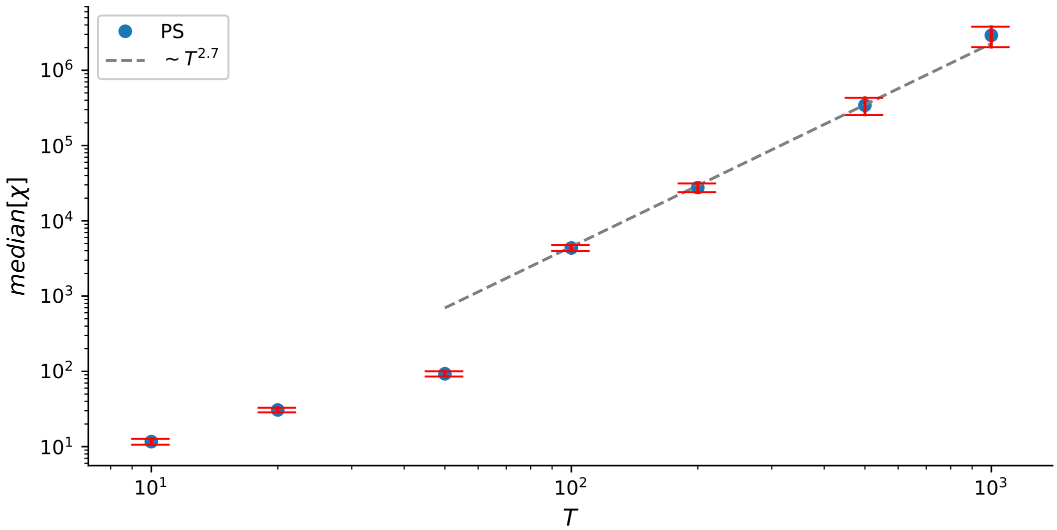

Figure 7.25: Averaged $\boldsymbol{\chi}(\mathbf{x}_0, T)$ as a function of time span obtained by repeating the algorithm for 10^5 different initial conditions.

obtained with the application of conventional methods.

Chapter 8

Conclusions

In this project a new shadowing-based sensitivity analysis method is presented and applied to two low-dimensional chaotic systems arising from natural convection problems. Sensitivity analysis is still an important tool when coupled with other numerical simulations, e.g. the pairing with Reynolds-Averaged Navier-Stokes equations in the aerodynamic shape design of wing-fuselage bodies or in turbomachinery applications. The need to develop a new approach lies in the fact that when chaotic systems are considered, as during unsteady turbulent flow simulations, the extremely sensitivity on the initial conditions leads the current generation of sensitivity analysis methods to be completely ineffective. The major contribution of this thesis is to provide an alternative shadowing-based sensitivity algorithm to evaluate how a small perturbations in design and control variables of a chaotic systems affect certain observables of interest. Periodic Shadowing algorithm is applied to the chaotic systems arising from Lorenz equations [5] and from a greater complexity system characterised by six additional states [23]. The sensitivity of the time average of a quantity related to the heat transfer in the fluid cells with respect a coordinate transformation γ and to a physical parameter ρ proportional to the temperature difference between the plates is evaluated. In the present chapter we outline the most important results obtained from the application of the Period Shadowing to the considered chaotic system and then, we express our intentions for future development.

8.1 Discussion

The obtained results show an excellent response of the method when perturbations in parameter γ are considered. In fact, for both systems, the averaged sensitivity computed for repeated samples converges to the approximation of the gradient obtained through a second-order finite difference scheme as $T \to \infty$. In parallel, the standard deviation of the computed sensitivity decays with the same rate of the sensitivity error presented in section 3 of Ref. [21]. The rapid $1/T$ decay implies that for some sufficiently large time span the shadowing error $\mathcal{E}_1^T(\mathbf{x}_0, p)$ will be, on average, smaller than that of the random error $\mathcal{E}_0^T(\mathbf{x}_0, p)$ and the global error is mostly dominated by the first term.

As well underlined in Ref. [8], using a periodic trajectory for sensitivity analysis, the error $\mathcal{E}_1^T(\mathbf{x}_0, p)$ is identically zero and only the random error $\mathcal{E}_0^T(\mathbf{x}_0, p)$ affects the sensitivity results.

The time gradient $T_{d\gamma}/T$ between the reference and perturbed trajectories tends to 0 as $T \to \infty$, which means that stretched (or compressed) state space evolution has the same time scale of the reference trajectory if integration time is sufficiently long. Furthermore we observe that the mean of the squared norm of the linearised solution evaluated in shooting points tends to a constant value for both Lorenz and Reiterer systems, respectively. These results can all be justified by the fact that the introduction of a coordinate transformation like the one proposed in the chapter 6, does not bring the system into bifurcations.

Concerning the average sensitivity of the observable with respect to parameter ρ, we observe converging values if longer and longer time span are considered. However, they present a consistent, reproducible bias between the finite difference gradient approximation. The gap, shown in figure 7.9 and figure 7.22, is imputable due to the fact that the chaotic systems considered and described by (6.6),(6.9) are not completly hyperbolic and thus not structurally stable, as stated in Refs. [40, 41, 42]. On the other hand, it may seem that dynamical behaviour of Lorenz attractor is very robust, that is that nonlinear simulation which provides, for instance, the classical "butterfly" of Lorenz attractor looks like pretty much the same (obviously considering the chaotic field) if we take slightly different values for the design parameters σ, β, ρ. The first question that comes to mind after these statements should be: how can this be true, if this system is unstable? The answer to this problem is the key presence of a weak form of hyperbolicity that in Ref. [41] is called singular hyperbolicity. The authors of Ref. [43] prove that any robust attractor containing an equilibrium point is singular hyperbolic. The fact of not being a complete hyperbolic system, leads the attractive set always infinitely close to the bifurcations [44]. These bifurcations cause a discontinuity in the statistical quantities such as (4.26), and hence not differentiable with the parameters. In such situations, the limit (4.27) does not itself formally exist, as the infinite time average is not a continuous function of the parameters. However, the empirical observation on this set of equations suggest that statistics appear as if they were smooth functions

of ρ. The details on the hyperbolic characteristics of chaotic systems are not object of discussion in this project but more information are present in the cited references, especially in Ref. [45] concerning the Lorenz equations.
This feature, known for low dimensional systems [46], is expected to be a general property of many physical systems characterized by a non complete form of hyperbolicity. Logically, therefore, the use of an approach based on the shadowing lemma such as Periodic Shadowing or Least Square Shadowing can be questionable. The perturbed trajectory obtained from the linear problem may not belong to the attractor of the perturbed system, although it may lie close to it. Hence, in the calculation of a statistical convergence of the sensitivity, a spurious contribution is introduced (we underline that the gap between the two results is $\sim 1.5\%$ for the Lorenz system), resulting in a sensitivity error that does not vanish as $T \to \infty$. This outcome does not preclude the effective validity of the method because it is supported by the results of the application of other algorithms such as those presented in Refs. [7, 8, 37]. The presence of the bias can also be connected with the so-called "chaotic hypothesis" [47]. The hypothesis asserts that if the dimension of the system is large enough, perturbations in design parameter which can induce bifurcations, such as ρ, do not lead to such a noticeable evidence in the calculation of the average statistic of the sensitivity with respect to it. In other words, the greater the number of states, the greater is the stability of the systems with respect to small parameter perturbations, and so they behave as they were hyperbolic [48]. In this way the bias error should be reduced if high-dimensional systems are considered. To date, this turns out to be a conjecture but research in this direction is currently under way to provide additional support to shadowing-based algorithms.

8.2 Future Development

There are several important aspects requiring further research. Firstly, develop a fundamental understanding of how the proposed method performs in high-dimensional systems is warranted. In other words, it is necessary to understand how the number of degrees of freedom affects the aforementioned bias error, verifying the validity of the conjecture. The achievement is accomplished by analysing the effects of the lack of hyperbolicity of systems on the shadowing error. A better understanding and characterization of the spectral properties of the multiple-shooting system resulting from the Periodic Shadowing approach is also needed. A further research problem is to implement the Least Square Shadowing method to observe how sensitivity distributions. In this way, it will also be possible to compare the computational cost and accuracy of the methods and understand how Least Square Shadowing behaves for chaotic systems of greater complexity than the one arising from Lorenz equations. Finally, find alternative strategies to set the gradient T_{dp} are required. The aim would be to prevent that the boundary value problem (4.38) and the multiple-shooting system (5.9) become singular. This might in turn improve the conditioning of the problem and result in a more favourable probability distribution of the sensitivity. We wish to address these aspects in future work.

Appendix A

Lagrangian Approach for Adjoint Sensitivity Method

A classical Lagrangian approach [49][21] starts by constructing the finite-time Lagrangian function

$$\mathcal{L}^T = \mathcal{J}^T(p) + \frac{1}{T}\int_0^T \mathbf{q}^\intercal(t)\cdot\left[\dot{\mathbf{y}}(t) - \mathbf{f}_{\partial\mathbf{x}}(\mathbf{x}(t),p)\cdot\mathbf{y}(t) - \mathbf{f}_{\partial p}(\mathbf{x}(t),p) - \frac{T_{dp}}{T}\mathbf{f}(\mathbf{x}(t),\mathbf{p})\right]dt \tag{A.1}$$

$\mathcal{L}^T : \mathbb{R} \to \mathbb{R}$, by adding the sensitivity equation 4.35 (as second term in right-hand side) to the cost function 4.25, with the adjoint variables $\mathbf{q}(t) \in \mathcal{X} \equiv \mathbb{R}^N$. Since the sensitivity equation just added is satisfied for all p and for all $t \in [0,T]$ along the trajectory, $\mathcal{L}^T(p) = \mathcal{J}^T(p)$ for every p and thus $\mathcal{L}^T_{dp} = \mathcal{J}^T_{dp}$. This identity is exploited to obtain the gradient $\mathcal{J}^T_{dp}$ from $\mathcal{L}^T_{dp}$ at a much reduced computational cost when the sensitivity with respect to many parameters is required. As foreseeable, the derivative of the finite-time Lagrangian with respect to the parameter is defined as

$$\mathcal{L}^T_{dp}(p) = \lim_{\delta p\to 0}\frac{1}{\delta p}\left[\mathcal{L}^{T'}(p') - \mathcal{L}^T(p)\right] \tag{A.2}$$

where $\mathcal{L}^{T'}(p')$ and $\mathcal{L}^T(p)$ refer to the perturbed and reference trajectories, respectively. Now, replacing into A.1 the 4.32 and expanding the multiplication on the additional term leads to

$$\begin{aligned}\mathcal{L}^T_{dp}(p) = \frac{1}{T}\int_0^T \left[\mathcal{J}_{\partial\mathbf{x}}(t)\cdot\mathbf{y}(t) + \mathcal{J}_{\partial p}(t)\right]dt + \frac{1}{T}\int_0^T\Big[\mathbf{q}^\intercal(t)\cdot\dot{\mathbf{y}}(t) - \\ + \mathbf{q}^\intercal(t)\mathbf{f}_{\partial\mathbf{x}}(t)\cdot\mathbf{y}(t) - \mathbf{q}^\intercal(t)\mathbf{f}_{\partial p}(t) - \mathbf{q}^\intercal(t)\frac{T_{dp}}{T}\mathbf{f}(t)\Big]dt\end{aligned} \tag{A.3}$$

Integrating by parts the $\mathbf{q}^{\intercal}(t) \cdot \dot{\mathbf{y}}(t)$ term

$$\mathbf{q}^{\intercal}(t) \cdot \dot{\mathbf{y}}(t) = [\mathbf{q}^{\intercal} \cdot \mathbf{y}(t)]\Big|_0^T - \int_0^T \dot{\mathbf{q}}^{\intercal} \cdot \mathbf{y}(t) dt$$

and rearranging correctly

$$\mathcal{L}_{dp}^{T}(p) = \frac{1}{T}\int_0^T \left\{ \underbrace{[\mathcal{J}_{\partial \mathbf{x}} - \dot{\mathbf{q}}^{\intercal}(t) - \mathbf{q}^{\intercal}(t) \cdot \mathbf{f}_{\partial \mathbf{x}}]}_{A} \cdot \mathbf{y}(t) + \mathcal{J}_{\partial p} - \mathbf{q}^{\intercal}(t) \cdot \left[\mathbf{f}_{\partial p}(t) + \frac{T_{dp}}{T} \mathbf{f}(t) \right] \right\} dt - \\ - \underbrace{\frac{1}{T}[\mathbf{q}^{\intercal} \cdot \mathbf{y}(t)]\Big|_0^T}_{B} \tag{A.4}$$

The key idea of this approach to obtain the adjoint formulation is to select $\mathbf{q}(t)$ variables such that both A and B vanish identically, to avoid the explicit computation of $\mathbf{y}(t)$ for every parameter of interest.
Requiring A to vanish leads to an adjoint equation whilst $B = 0$ is obtained imposing periodic boundary condition in time on the adjoint problem. This leads to the adjoint Periodic Shadowing problem

$$\begin{cases} \dot{\mathbf{q}}(t) = \mathcal{J}_{\partial \mathbf{x}}(\mathbf{x}(t), p) \cdot \mathbf{q}(t) - \mathbf{f}_{\partial \mathbf{x}}(\mathbf{x}(t), p), & t \in [0, T] \\ \mathbf{q}(0) = \mathbf{q}(T) \end{cases} \tag{A.5}$$

Upon the solution of A.5 we obtain the sensitivity of the time average as

$$\mathcal{J}_{dp}^{T} = \mathcal{L}_{dp}^{T} = \frac{1}{T}\int_0^T \mathcal{J}_{\partial p}(\mathbf{x}(t), p) - \mathbf{q}^{\intercal}(t) \cdot \left[\mathbf{f}_{\partial p}(\mathbf{x}(t), p) + \frac{T_{dp}}{T} \mathbf{f}(\mathbf{x}(t), p) \right] dt \tag{A.6}$$

Bibliography

[1] Razvan Florea and Kenneth C. Hall. "Sensitivity Analysis of Unsteady Inviscid Flow Through Turbomachinery Cascades". *AIAA Journal* vol. 39(6) (June 2001), pp. 1047–1056.

[2] Joaquim R. R. A. Martins and Juan J. Alonso. "Aero-Structural Wing Design Optimization Using High-Fidelity Sensitivity Analysis" (August 2001).

[3] Zhoujie Lyu, Gaetan K. W. Kenway, and Joaquim R. R. A. Martins. "Aerodynamic Shape Optimization Investigations of the Common Research Model Wing Benchmark". *AIAA Journal* vol. 53 (April 2015).

[4] Daniel Lea, Allen Myles, and Thomas Haine. "Sensitivity analysis of the climate of a chaotic system". *Tellus Series a-Dynamic Meteorology and Oceanography* vol. 52 (March 2000), pp. 523–532.

[5] Edward N. Lorenz. "Deterministic Nonperiodic Flow". *Journal of the Atmospheric Sciences* vol. 20 (January 1963), pp. 130–141.

[6] Rufus Bowen. "ω-Limit Sets for Axiom A Diffeomorphisms". *Journal of Differential Equations* vol. 18 (1975), pp. 333–339.

[7] Qiqi Wang and Rui Hu. "Least Square Shadowing sensitivity analysis of chaotic limit cycle oscillations". *Journal of Computational Physics* vol. 267 (June 2014), pp. 210–224.

[8] Davide Lasagna. "Sensitivity Analysis of Chaotic Systems Using Unstable Periodic Orbits". *SIAM, Journal on Applied Dynamical Systems* vol. 17(1) (August 2017), pp. 547–580.

[9] Paolo Luchini and Alessandro Bottaro. "Supplemental Appendix to: An Introduction to Adjoint Problems". *Annual Review of Fluid Mechanics* vol. 46 (2014), 493–517.

[10] E. D. Pendlebury. "General Perturbation Theory in Neutronics". *Proceedings of the Physical Society, Section A* vol. 68 (June 1955), pp. 474–481.

[11] Chun Hsiung Chen and Chuen-Der Lien. "The Variational Principle for Non-Self-Adjoint Electromagnetic Problems". *IEEE Transactions on Microwave Theory and Techniques* vol. 28 (August 1980), pp. 878–886.

[12] G.L. Eyink et al. "Ruelle's linear response formula, ensemble adjoint schemes and Lévy flights". *Nonlinearity* vol. 17 (2004), pp. 1867–1889.

[13] Thomas R. Bewley, Parviz Moin, and Roger Temam. "DNS-based predictive control of turbulence: an optimal benchmark for feedback algorithms". *Journal of Fluid Mechanics* vol. 447 (2001), pp. 179–225.

[14] S. Corti and T. N. Palmer. "Sensitivity analysis of atmospheric low-frequency variability". *Quarterly Journal of the Royal Meteorological Society* vol. 123 (1997), pp. 2425–2447.

[15] Daniel J. LEA et al. "Sensitivity analysis of the climate of a chaotic ocean circulation model". *Quarterly Journal of the Royal Meteorological Society* vol. 128 (2002), pp. 2587–2605.

[16] Patrick J. Blonigan and Qiqi Wang. "Probability density adjoint for sensitivity analysis of the Mean of Chaos". *Journal of Computational Physics* vol. 270 (August 2014), pp. 660–686.

[17] John Thuburn. "Climate sensitivities via a Fokker–Planck adjoint approach". *Q. J. R. Meteorol. Soc* vol. 3 (2005), pp. 73–92.

[18] Qiqi Wang. "Forward and adjoint sensitivity computation of chaotic dynamical system". *Journal of Computational Physics* vol. 235 (February 2013), pp. 1–13.

[19] Patrick J. Blonigan and Qiqi Wang. "Multiple shooting shadowing for sensitivity analysis of chaotic dynamical systems". *Journal of Computational Physics* vol. 354 (February 2018), pp. 447–475.

[20] T. Sauer and J. A. Yorke. "Shadowing trajectories of dynamical systems". *Computer-Aided Proofs in Analysis*. Ed. by K. Meyer and D. Schmidt. 1990, pp. 229–234.

[21] Davide Lasagna, Ati Sharma, and Johan Meyers. "Periodic Shadowing Sensitivity Analysis of Chaotic Systems". *arXiv* (June 2018).

[22] David Ruelle. "Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics". *Journal of Statistical Physics* vol. 95 (April 1999), pp. 393–468.

[23] Peter Reiterer et al. "A nine-dimensional Lorenz system to study high-dimensional chaos". *Journal of Physics A: Mathematical and General* vol. 31 (August 1998).

[24] Ilia Grigorenko and Elena Grigorenko. “Chaotic Dynamic of the Fractional Lorenz System”. *Physical Review Letter* vol. 91 (July 2003).

[25] David Ruelle. “A Review of Linear Response theory for General Diffentiable Dynamical Systems”. *Nonlinearity* vol. 22 (March 2009), pp. 855–870.

[26] Divakar Viswanath. “The Lindstedt Poincaré Technique as an Algorithm for Computing Periodic Orbits”. *Society for Industrial and Applied Mathematics Journal* vol. 43(3) (2001), pp. 478–495.

[27] Ali Hasan Nayfeh. *Introduction to Perturbation Techniques*. Ed. by Wiley-Interscience. 1993.

[28] Ralma Larter. “Sensitivity Analysis of Autonomous Oscillators. Separation of Secular Terms and Determination of Structural Stability”. *The Journal of Physical Chemistry,* vol. 87(3) (1983), pp. 3114–3121.

[29] Henk A. Dijkstra et al. “Numerical BifurcationMethods and their Application to Fluid Dynamics: Analysis beyond Simulation”. *Communications in Computational Physics* vol. 15 (January 2014), pp. 1–45.

[30] Uri M. Ascher, Robert M.M. Mattheij, and Robert D. Russel. *Numerical solution of boundary value problems for ordinary differential equations*. Ed. by Society for Industrial and Applied Mathematics Journal. 1995.

[31] Melvin R. Scott and Herman A. Watts. “Computational solution of Linear Two-Point Boundary Value Problems via Orthonormalization”. *SIAM Journal on Numerical Analysis* vol. 14 (1977), pp. 40–70.

[32] Jack K. Hale. *Ordinary Differential Equations*. Ed. by Wiley-Interscience. Vol. 14.

[33] Edaward N. Lorenz. *“Does the Flap of a Butterfly’s wings in Brazil Set Off a tornado in Texas?”* 139th Meeting of American Association for the Advancement of Science, December 1972.

[34] Paul Manneville. *Instabilities, Chaos and Turbolence*. Ed. by Imperial College Press. 2004.

[35] Charles Y. Li. “Chaos in Partial Differential Equations, Navier-Stokes Equations and Turbulence”. *International Congress of Chinese Mathematician* vol. 3 (2007), pp. 110–122.

[36] John Guckenheimer and Philip Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Ed. by Springer-Verlag. 1983.

[37] Haitao Liao. “Efficient sensitivity analysis method for chaotic dynamical systems”. *Journal of Computational Physics* vol. 313 (May 2016), pp. 57–75.

[38] Maurice C. Bryson. "Heavy Tailed Distributions: Properties and Tests". *Technometrics* vol. 16 (February 1974), pp. 61–68.

[39] John Martin Bland. *An Introduction to Medical Statistics*. Ed. by Oxford University Press. Vol. 3. 2000.

[40] Marcelo Viana. "What's new on Lorenz Strange attractors". *The Mathematical Intelligencer* vol. 22 (2000), pp. 6–19.

[41] Carlos Arnoldo Morales, Maria José Pacifico, and Enrique Ramiro Pujals. "Singular Hyperbolic System". *Proceeding of the American Mathematical Society* vol. 127 (May 1999), pp. 3393–3401.

[42] John Guckenheimer. *A Strange, Strange Attractor. In: The Hopf Bifurcation and Its Applications. Applied Mathematical Sciences*. Ed. by Springer-Verlag. Vol. 19. 1976.

[43] Carlos Arnoldo Morales, Maria José Pacifico, and Enrique Ramiro Pujals. "On C^1 robust singular transitive sets for three-dimensional flows". *Comptes Rendus de l'Académie des Sciences* vol. 326 (1998), pp. 81–86.

[44] Christian H. Reick. "Linear response of the Lorenz system". *Physical Review E* vol. 66 (2002).

[45] Colin Sparrow. *The Lorenz equations: bifurcations, chaos and strange attractors*. Ed. by Springer-Verlag. 1982.

[46] Doyne J. Farmer. "Sensitive Dependence on Parameters in Nonlinear Dynamics". *Physical Review Letters* vol. 55 (22 July 1985).

[47] Giovanni Gallavotti and Ezechiel Godert David Cohen. "Dynamical ensembles in stationary states". *Journal of Statistical Physics* vol. 80 (September 1995), pp. 931–970.

[48] D. J. Albers1 and J. C. Sprott. "Structural Stability and Hyperbolicity Violation in High-Dimensional Dynamical Systems". *Institute of Physics Publishing. Nonlinearity* vol. 19 (August 2006), pp. 1801–1847.

[49] Alfio Borzì and Volker Schulz. *Computational Optimization of Systems Governed by Partial Differential Equations*. Ed. by Society for Industrial and Applied Mathematics Journal. May 2012.

[50] Robert P.Dickinson and Robert J. Gelinas. "Sensitivity Analysis of Ordinary Differential Equation Systems-A Direct Method". *Journal of Computational Physics* vol. 21 (June 1976), pp. 123–143.

[51] Dan G. Cacuci and Michael E.Schlesinger. "On the application of the adjoint method of sensitivity analysis to problem in the atmospheric sciences". *Atmósfera* vol. 7 (1994), pp. 47–59.

[52] Stephen M Hammel, James A. Yorke, and Celso Grebogi. "Do numerical orbits of chaotic dynamical processes represent true orbits?" *Journal of Complexity* vol. 3 (January 1987), pp. 136–145.

[53] Vasile Marinca and Nicolae Herisanu. *Nonlinear Dynamical Systems in Engineering*. Ed. by Springer. 2012.

[54] Helmut Sobieczky. "Parametric Airfoils and Wings". *Notes on Numerical Fluid Mechanics* vol. 68 (1998), pp. 71–88.

[55] Paolo Luchini and Alessandro Bottaro. "Adjoint Equations in Stability Analysis". *Annual Review of Fluid Mechanics* vol. 46 (January 2014), 493–517.

[56] Ronald M. Errico. "What Is an Adjoint Model?" *Bulletin of the American Meteorological Society* vol. 78 (1997), pp. 2577–2591.

[57] Lloyd N. Trefethen and David Baus. *Numerical Linear Algebra*. Ed. by Society for Industrial and Applied Mathematics. 1997.

[58] Warwick Tucker. "The Lorenz attractor exists". *Comptes Rendus de l'Académie des Sciences* vol. 328 (1999), pp. 1197–1202.

[59] Syukuro Manabe and Ronald J. Stouffer. "Sensitivity of a global climate model to an increase of CO_2 concentration in the atmosphere". *Journal of Geophysical Research: Oceans* vol. 85 (20 October 1980).

[60] Ronald M. Errico and Tomislava Vukicevic. "Sensitivity Analysis Using an Adjoint of the PSU-NCAR Mesoscale Model". *Monthly Weather Review* vol. 120 (November 1991), pp. 1644–1660.

[61] Alan George. "On Block Elimination for Sparse Linear Systems". *SIAM Journal on Numerical Analysis* vol. 3 (June 1974), pp. 585–603.