MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21d/2023/08.22.19.06-TDI

GRADIENT PATTERN ANALYSIS: ENHANCEMENTS AND APPLICATIONS INCLUDING THE INFLUENCE OF NOISE ON PATTERN FORMATION

Rubens Andreas Sautter

Doctorate Thesis of the Graduate Course in Applied Computing, guided by Dr. Reinaldo Roberto Rosa, approved in August 11, 2023.

URL of the original document:
<http://urlib.net/8JMKD3MGP3W34T/49M7R4L>

INPE
São José dos Campos
2023

PUBLISHED BY:

Instituto Nacional de Pesquisas Espaciais - INPE
Coordenação de Ensino, Pesquisa e Extensão (COEPE)
Divisão de Biblioteca (DIBIB)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21d/2023/08.22.19.06-TDI

GRADIENT PATTERN ANALYSIS: ENHANCEMENTS AND APPLICATIONS INCLUDING THE INFLUENCE OF NOISE ON PATTERN FORMATION

Rubens Andreas Sautter

Doctorate Thesis of the Graduate Course in Applied Computing, guided by Dr. Reinaldo Roberto Rosa, approved in August 11, 2023.

URL of the original document:
<http://urlib.net/8JMKD3MGP3W34T/49M7R4L>

INPE
São José dos Campos
2023

Cataloging in Publication Data

Sautter, Rubens Andreas.

Sa88g Gradient pattern analysis: enhancements and applications including the influence of noise on pattern formation / Rubens Andreas Sautter. – São José dos Campos : INPE, 2023.

xvi + 96 p. ; (sid.inpe.br/mtc-m21d/2023/08.22.19.06-TDI)

Thesis (Doctorate in Applied Computing) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2023.

Guiding : Dr. Reinaldo Roberto Rosa.

1. Gradient pattern analysis. 2. Spatiotemporal pattern formation. 3. Nonlinear fluctuations. 4. Stochastic complex ginzburg-landau. 5. Structures in space physics. I.Title.

CDU 519.246.8

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

DEFESA FINAL DE TESE RUBENS ANDREAS SAUTTER
BANCA Nº 201/2023, REGISTRO 133136/2018

No dia 11 de agosto de 2023, por teleconferência, o(a) aluno(a) mencionado(a) acima defendeu seu trabalho final (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. O(A) aluno(a) foi APROVADO(A) pela Banca Examinadora, por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Doutor em Computação Aplicada, com a exigência de que o trabalho final a ser publicado deverá incorporar as correções sugeridas pela Banca Examinadora, com revisão pelo(s) orientador(es).

Título: "GRADIENT PATTERN ANALYSIS: ENHANCEMENTS AND APPLICATIONS INCLUDING THE INFLUENCE OF NOISE ON PATTERN FORMATION".

Membros da Banca:

Dr. Stephan Stephany – Presidente – INPE
Dr. Reinaldo Roberto Rosa – Orientador (a) – INPE
Dr. Leonardo Bacelar Lima Santos – Membro Interno – INPE
Dr. Flavio Hernández Fenton – Membro Externo – GATECH
Dr. Érico Luiz Rempel – Membro Externo – ITA

Declaração de aprovação do Membro Externo Flavio Hernández Fenton anexa ao processo.

Documento assinado eletronicamente por **Leonardo Bacelar Lima Santos**, **Pesquisador**, em 14/08/2023, às 08:52 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Erico Luiz rempel (E)**, **Usuário Externo**, em 14/08/2023, às 10:28 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Stephan Stephany**, **Pesquisador Titular**, em 14/08/2023, às 10:40 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Reinaldo Roberto Rosa**, **Pesquisador Titular**, em 17/08/2023, às 16:13 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.mcti.gov.br/verifica.html, informando o código verificador **11280252** e o código CRC **9C1B6C3E**.

Referência: Processo nº 01340.006588/2023-11 SEI nº 11280252

ACKNOWLEDGEMENTS

I express my gratitude to Dr. Reinaldo Roberto Rosa for his guidance in this project. To my parents, Frank and Claudia, my dear little sister Cecilia, and my beloved Nathana, your unwavering support has been my foundation and inspiration.

I also wish to express my sincere thanks to my professors and colleagues at CAP for their invaluable contributions to the development of this work. Your insights and collaborative efforts have played an integral role in enriching the depth of this research.

I am also grateful to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for their financial support.

ABSTRACT

This doctoral project in applied computing enhances and evaluates the performance of the Gradient Pattern Analysis (GPA) technique in characterizing the spatiotemporal pattern formation processes generated from the multidimensional Stochastic Complex Ginzburg-Landau (SCGL) model. It is verified that the inclusion of noise, forming a stochastic extended system, leads to vortices and spiral defects that are structures resulting from the breaking of localized gradient symmetries which is compatible with the main GPA measurement criteria. The influence of the colored noise model (additive and multiplicative) on the formation of such patterns is investigated with an improved version of the GPA. For this, we present a refinement of the GPA considering two new increments: (i) A more precise metric to measure the phase fluctuation of the gradient lattice corresponding to each snapshot; and (ii) a new Lyapunov stability analysis of the reactive part is considered to evaluate which of the colored noises results in the highest rate of asymmetric fluctuations. Considering the stochastic model studied in this work, the stability analysis indicates that white and pink noise do not contribute to the formation of asymmetric fluctuations in the reactive phase space, while red noise is effective. Furthermore, based on the GPA measurements for modulus (G_2) and phase (G_3), we show that the SCGL multiplicative red noise model presents greater structural coherence in the generation of spiral defects than the additive model. Based on these results, applications of incremental GPA in a future machine learning context are presented, taking as examples data frames related to pattern formation processes in solar physics, cosmology, and medicine.

Palavras-chave: Gradient Pattern Analysis. Spatiotemporal Pattern Formation. Nonlinear fluctuations. Stochastic Complex Ginzburg-Landau. Structures in Space Physics.

ANÁLISE DE PADRÕES GRADIENTE: APRIMORAMENTOS E APLICAÇÕES INCLUINDO A INFLUÊNCIA DO RUÍDO NA FORMAÇÃO DE PADRÕES

RESUMO

Este projeto de doutorado em computação aplicada aprimora e avalia o desempenho da técnica Gradient Pattern Analysis (GPA) na caracterização dos processos de formação de padrões espaço-temporais gerados a partir do modelo multidimensional, complexo e estocástico, de Ginzburg-Landau (SCGL). Verifica-se que a inclusão de ruído, formando um sistema estendido estocástico, leva a vórtices e defeitos espirais que são estruturas coerentes resultantes da quebra de simetrias localizadas compatíveis com os principais critérios de medidas do GPA. A influência do modelo de ruído colorido (aditivo e multiplicativo) na formação de tais padrões é investigada com uma versão melhorada do GPA. Para isso, apresentamos um refinamento do GPA considerando dois novos incrementos: (i) Uma métrica mais precisa para medir a flutuação de fase da grade gradiente correspondente a cada instantâneo; e (ii) uma nova análise de estabilidade de Lyapunov da parte reativa é considerada para avaliar qual dos ruídos coloridos resulta na maior taxa de flutuações assimétricas. Considerando o modelo estocástico estudado neste trabalho, a análise de estabilidade indica que os ruídos branco e rosa não contribuem para a formação de flutuações assimétricas no espaço de fase reativa, enquanto o ruído vermelho é efetivo. Além disso, com base nas medidas de GPA para módulo (G_2) e fase (G_3), mostramos que o modelo SCGL com ruído vermelho multiplicativo apresenta maior coerência estrutural na geração de defeitos espirais do que o modelo aditivo. Com base nesses resultados, são apresentadas aplicações de GPA incremental em um futuro contexto de aprendizado de máquina, tomando como exemplos dataframes relacionados a processos de formação de padrões em física solar, cosmologia e medicina.

Palavras-chave: Análise de Padrões Gradiente. Formação de Padrões espaço-temporais. Flutuações não-lineares. Equação Complexa Estocástica de Ginzburg-Landau. Estruturas em física espacial.

LIST OF FIGURES

Page

LIST OF TABLES

Page

CONTENTS

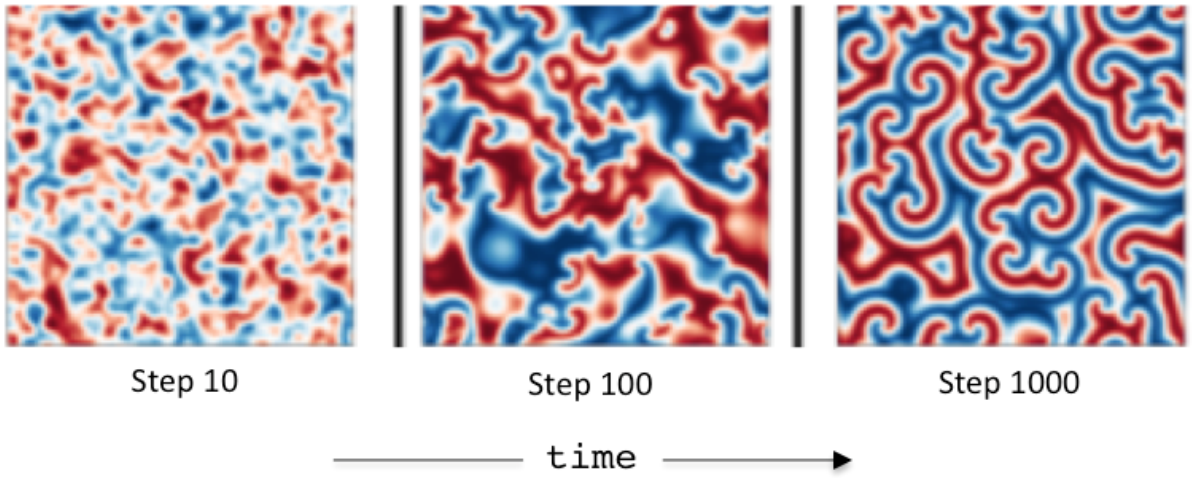
1 INTRODUCTION

This doctoral thesis project in Applied Computing has as main objective to continue an advanced study on the technique known as Gradient Pattern Analysis (GPA) (ROSA et al., 2000; RAMOS et al., 2000; BARONI; ROSA, 2011; ROSA et al., 2018). The advanced study aims to improve the performance of the technique based on a critical study of its performance, always trying to find characteristics that can be improved. In my master's project (SAUTTER, 2018) we performed a critical study on the performance of GPA applied to so-called Coupled Map Lattices (CMLs) in their 2D spatiotemporal configuration (see Appendix A).

CMLs are canonical nonlinear systems that allow studying the formation and evolution of spatiotemporal patterns whose characteristics of structural complexity are sensitive to initial conditions, boundary conditions and level of chaotic behavior imposed on the model. In this doctoral project, we generalize the application of GPA to a system that, in addition to involving chaos, also incorporates noise and can be simulated and analyzed in multiple dimensions: ordinary time series (without spatial domains), 2D spatiotemporal series (equivalent to a sequence of images based on pixels) and 3D (equivalent to a sequence of hypercube based on voxels). Figure 1.1 shows a spatiotemporal pattern formation event (2D+1) visualized as a sequence of some snapshots that represent the temporal evolution of the Ginzburg-Landau reactive-diffusive process. We are using the notation of the theory of amplitude equations where nD+1 denotes a sequence of snapshots where n is the spatial dimension and "1" means the measure of amplitude over time. Thus, the most common are 2D+1 (for matrices or 2D images) and 3D+1 for hypercubes or 3D images.

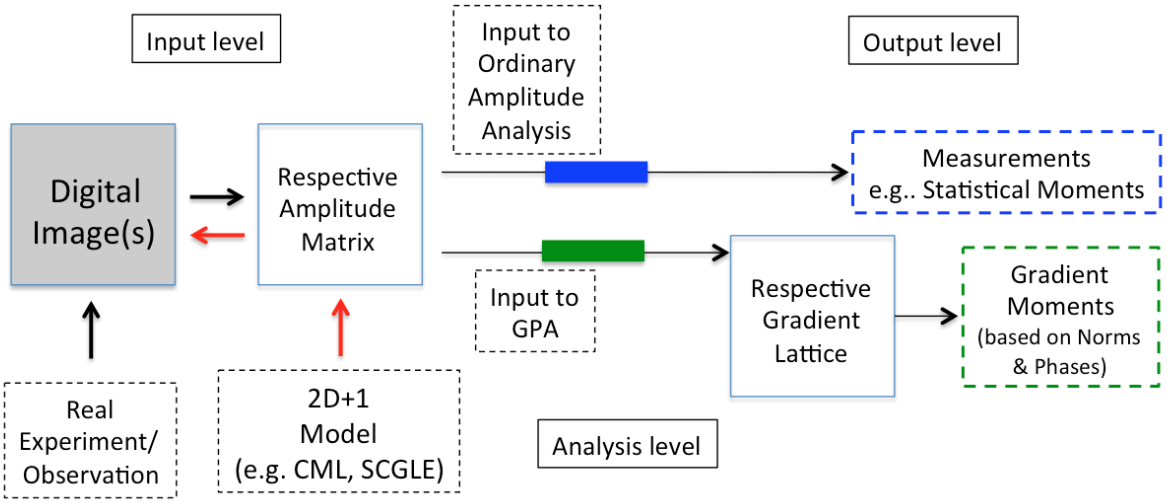
Figure 1.2 illustrates the scenario that relates the study of GPA as a technique for analyzing simulated and/or observed spatiotemporal patterns from various mathematical, physical, chemical and biological observed processes (or models). In general terms, GPA, instead of extracting measurements from numerical values of a matrix, first operates on it to generate the respective gradient lattice and then extracts geometric measurements from the gradient pattern (first gradient moment) and also from the norm's set (second gradient moment) and corresponding phases (third gradient moment), and may also compose a fourth measure based on the complex entropy written from the norms and phases (fourth gradient moment) (RAMOS et al., 2000; ROSA et al., 2003).

Figure 1.1 - Example of a typical 2D+1 Spatiotemporal pattern formation.



Three time steps are shown in the simulation of spiral pattern formation, from the perspective of Amplitude Equations Theory.

Figure 1.2 - The GPA 2D+1 analytical paradigm.



The GPA paradigm, which explores more information about a complex 2D+1 structure that may be contained in the norms and phases of the gradient lattice, which is obtained from the corresponding amplitude matrix.

The details of the research problem and the structure of the thesis are described in the sections below.

1.1 Motivations and hypothesis

The study of SpatioTemporal Pattern Formation (STPF) began in 1900s with the study of convection(RAYLEIGH, 1916). Subsequent studies introduced new terms to the Partial Differential Equations (PDEs), specifically associated with population dynamics, which are now recognized as reaction-diffusion systems (FISHER, 1999; TURING, 1990; DÚZS et al., 2023)[1]. Within this context, the so-called "coherent structures" that emerge in many physical, chemical and biological systems have been observed and simulated from a multitude of methods and techniques (CLADIS; PALFFY-MUHORAY, 1995; WALGRAEF, 1997). With the advent of scientific computing approaches, this area has become very broad, both in terms of improving models and techniques for numerical simulations, as well as advancing the quality of observations through digital images with high resolution and multifrequency samples. Therefore, from the point of view of data analysis, new challenges have arisen in relation to the study of techniques capable of quantifying and identifying the formation of spatial structures over time (mainly in 2D, in the form of digital images and their respective matrices).

Unlike a large amount of techniques available for the analysis of ordinary time series (without spatial information), there are few techniques involving measurements that are effective in characterizing the formation of spatiotemporal patterns (DIMITROPOULOS et al., 2014; HAO et al., 2019; XU et al., 2015), mainly when underlying nonlinear processes are addressed as the cause of coherent structures and defects as observed in reactive-diffusive systems, chaotic and turbulent patterns in fluids, structure formation in astrophysics and cosmology, propagation and growth of structures in biological systems, complex materials, etc) (CLADIS; PALFFY-MUHORAY, 1995; WALGRAEF, 1997; CROSS; HOHENBERG, 1993; DEKEL; OSTRIKER, 1999). Added to this context is the need for measures that act as attributes and labels in machine learning, making it possible to incorporate Artificial Intelligence in the monitoring and recognition of complex spatiotemporal patterns[2].

Glimpsing the above scenario, spatiotemporal data analysis is an emerging research

[1]The seminal paper of Turing (1990) is a reprinted version of a paper presented in the *Philosophical Transactions of the Royal Society*, 1953

[2]Complex systems here are those that present at least three properties: non-linearity, collective transitive behavior and/or self-organization and entropy exchange with the environment(CLADIS; PALFFY-MUHORAY, 1995).

area due to the development and application of new computational techniques that allow the analysis of complex spatiotemporal systems considering even large datasets. Spatiotemporal models arise when data are collected in time and space and have at least one spatial and one temporal property. In our context, we consider that a spatiotemporal dataset describes a spatial and temporal phenomenon that exists at a given time t and location (x,y) or (x,y,z). It is in this context that the GPA technique emerged in the 1990s, originally conceived as a new tool to analyze the asymmetric fragmentation of an active solar region observed in X-rays by the Yohkoh mission(ROSA et al., 1999). Since there, the technique has been tested and improved in several applications(DANTAS, 2014; BARONI, 2009; GUERRA, 2008; SANTOS, 2009; COSTA JÚNIOR, 2004; FREITAS, 2012; SAUTTER, 2018). Taking into account all those masters and doctoral projects, GPA applications include cases in the domains of physics, astrophysics, cosmology, chemistry, biology, meteorology, and medicine.

In some of these works GPA was compared with other techniques and proved to be a unique method to detect patterns via bilateral symmetry breaking measurements on the gradient lattice. For instance, some techniques that were compared to GPA was: texture attributes (SOUZA; ASSIREU, 2016), Bidimensional Detrended Fluctuation Analysis (DFA-2D) (SOUZA; ASSIREU, 2016) and Minkowski functionals (ROSA et al., 2007; BARONI, 2009).

In this project, motivated by the challenges described below, we elaborated two hypotheses to be investigated in this applied computing thesis:

(i) The Stochastic Complex Ginzburg-Landau Equation (SCGLE) is a popular PDE in physics to study phase-transitions(ARANSON; KRAMER, 2002). The inclusion of noise terms in this equation can change the formation patterns. Furthermore, this type of model allows studying the intriguing pattern formation mechanism in systems as the largest forest ecosystems(NOBRE; FABRICIO-NETO, 2021; ASSAD et al., 2022). The presence and characterization of noise, in this context, is still a relevant research topic that has never been addressed by the GPA technique. In this sense, the hypothesis that the GPA technique can be effective in the characterization of patterns generated from the SCGLE will be carefully investigated, admitting the need of new technical GPA increments and the implementation of an effective numerical model that allows to apply the GPA in the ordinary time series, 2D and 3D solutions from SCGLE, which

lead to the formation of patterns on the influence of noise.

(ii) The GPA technique, when seen within a Machine Learning context, demands testing its usefulness within a new scope of analysis, incorporating Dashboard solutions and perspectives of using GPA parameters as attributes that works as training labels for machine learning models primarily for pattern detection and classification. This perspective extends the applications of GPA to areas such as monitoring solar activity and also to the use of artificial intelligence in health sciences. Thus, the second hypothesis of this thesis consists of evaluating GPA as a general-purpose machine learning technique(ROSA et al., 2018; BARCHI et al., 2020).

1.2 Main goals

Based on the motivations and hypotheses described above, the main objectives of this research are as follows:

a) Implement simulations based on the SCGLE numerical solution to test, improve and validate the GPA technique in the characterization of STPF and regimes associated to them under the presence of noise;

b) Start applying the technique within a broader context involving Machine Learning challenges. This objective does not intend to exhaust the subject in this thesis, but only to discuss in an incipient way the hypothesis that a GPA- Dashboard Analytics solution can be useful, for example, in applications in solar physics, cosmology and medicine. The case study in medicine was incorporated into the thesis project for the following reasons: The application of the GPA in a FAPESP PIPE-Phase 2 project that is in progress [3] and will offer a postdoctoral scholarship recommended by FAPESP. In addition, the same project foresees developments to provide solutions within NASA's Artemis Mission[4] regarding the use of computer vision to assess practical skills in aerospace medicine.

1.3 Document outline

We have organized the content as follows: Chapter 2 provides a comprehensive review of the GPA formalism and introduces the Stochastic Complex Ginzburg-Landau

[3] <https://bv.fapesp.br/pt/pesquisa/buscador/?q2=(id_pesquisador_exact%3A2682)%20AND%20(auxilio:*%20AND%20situacao:%22Em%20andamento%22)>

[4] <https://www.nasa.gov/specials/artemis/>

Equation (SCGLE). Chapter 3 presents, in Section 3.1., the numerical method used in the SCGLE solution. In Section 3.2., the mathematical increments of GPA are investigated. The data that will be analyzed by GPA are presented in Chapter 4. Results and Interpretations are in Chapter 5. An overview of GPA in a Visual Analytics context is presented in Chapter 6. The overall contributions of this thesis can be found in Appendix B.

2 THEORETICAL BACKGROUND

The exploration of spatiotemporal pattern formation (WALGRAEF, 1997; SWINNEY; KRINSKY, 1992; COELHO et al., 2021) plays a central role in understanding nonequilibrium regimes within reactive-diffusive systems (CROSS; HOHENBERG, 1993; ARIS et al., 1991), such as fluid dynamic instabilities in complex media (GALLAIRE; BRUN, 2017), Belouzov-Zhabotinsky reaction (BARZYKINA, 2020), and electrophysiological propagation in biological systems (WINFREE, 1997). These systems typically exhibit long-term relaxation, which arises from the coupling of slow modes in their amplitude dynamics (MANNEVILLE, 2005) and is statistically represented by universality classes (COLLET; ECKMANN, 2014; JABEEN; GUPTE, 2010). Therefore, comprehending the role of instabilities in pattern formation is vital for formulating thermodynamics (BUDRONI; WIT, 2017) and developing wave theories (HUPKES et al., 2020) applicable to nonlinear, out-of-equilibrium extended systems. Among the extensively studied systems, the complex Ginzburg-Landau equation stands out, revealing fascinating spiral patterns, double spirals, and cellular patterns (ARANSON; KRAMER, 2002). These intricate patterns offer valuable insights into the underlying dynamics and complexity of many biological and physical systems.

2.1 Complex Ginzburg-Landau

The Complex Ginzburg–Landau (CGL) equation is one of the most remarkable nonlinear equations in physics. It describes superconductors (YANG; HOU, 2008; ARANSON; KRAMER, 2002), being an important model for the study of phase transitions (ARANSON; KRAMER, 2002). It is also an important model to study pattern formation since it is able to capture the emergence of complex spatiotemporal patterns from the interaction of a small number of parameters. The system is governed by the Equation 2.1, where A is a complex amplitude envelope, related to a spatial variable at time t, b is a spatial parameter related to diffusion term $\nabla^2 A$, c is a parameter related to the reactive part of the equation, and $i = \sqrt{-1}$.

$$\partial_t A = (1 + ib)\nabla^2 A + A - (1 + ic)|A|^2 A \tag{2.1}$$

In practical applications, it is commonly observed noise interference in Partial Differential Equations. Therefore, recent studies include noise terms in the CGL equation, which is formally known as Stochastic Complex Ginzburg-Landau (SCGL) equation (SHEN et al., 2018; YANG; HOU, 2008). The main approaches include the additive

or multiplicative noise terms, according to Equations 2.2 and 2.3 respectively.

$$\partial_t A = (1+ib)\nabla^2 A + A - (1+ic)|A|^2 A + \sigma\partial_t(\eta_\beta) \tag{2.2}$$

$$\partial_t A = (1+ib)\nabla^2 A + A - (1+ic)|A|^2 A + \sigma A\partial_t(\eta_\beta) \tag{2.3}$$

In those equations, η_β is a colored noise, which typically is a white noise term ($\beta = 0$), in Chapter 3 we discuss the algorithm to generate this signal.

The main motivation to study this stochastic reactive system is to mimic rainforest dynamics (LANGNER et al., 2018; BOLZAN et al., 2002; RODRIGUES NETO et al., 2001). In Amazon rainforest, for example, the dynamics of air currents are directly influenced by the heterogeneous patterns created by the canopy layer (CAMPANHARO et al., 2008; CAMPOS VELHO et al., 2001). The formation of short and long-term biophysical patterns in this ecosystem can also be influenced by stochastic reactive-diffusion processes, where weak turbulence plays a significant role in promoting self-organization under non-equilibrium conditions, such as wildfires, deforestation, and the precipitation of pollutants. These conditions often give rise to dissipative structures, including chaotic heat ramps and spiral defects, which involve the presence of diffusive noise (CAMPANHARO et al., 2008; CAMPOS VELHO et al., 2001; LEJEUNE et al., 1999; LANGNER et al., 2018).

An important feature apparent in these systems is their symmetry, which serves as an indicator for various physical processes, including turbulence, chaos, and boundary-induced instabilities. Within this context, GPA emerges as a promising technique for pattern characterization. This approach stands out by not solely relying on symmetry but also incorporanting the asymmetries in the gradient lattices which are a signature of nonlinear local fluctuations as well as the emergence of coherent structures as spatiotemporal patterns. Consequently, in the subsequent section we will discuss the theoretical bases of GPA.

2.2 Gradient Pattern Analysis

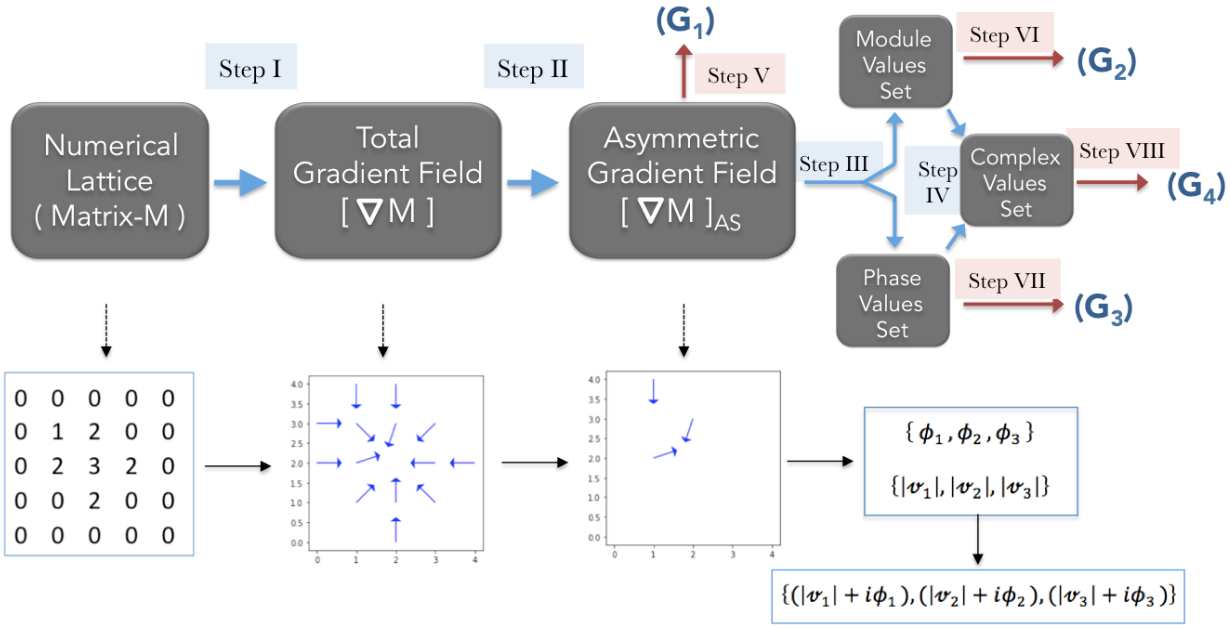
In complex systems, the characterization of spatiotemporal patterns is one of the main challenges, involving the definition and interpretation of techniques that are capable of measuring the emergence of complex structures over time. Addressing this challenge, the GPA formalism offers a well-defined set of criteria that enable the comprehensive characterization of complex spatial patterns by considering the

bilateral symmetry breaking from the corresponding matrix gradient lattice. The initial application of this technique was primarily associated with solar bursts coming from asymmetric magnetic sources in the solar atmosphere(ROSA et al., 1999), and subsequent research expanded its application to time-series (0D+1) domain (ASSIREU et al., 2002), introduced new metrics (ROSA et al., 2003; RAMOS et al., 2000; ROSA et al., 2018), and incorporated preprocessing enhancements related to pattern symmetry (ROSA et al., 2018; BARCHI et al., 2020). In this section we revisit the GPA technique.

2.2.1 The GPA operator

As a measurement technique, GPA is a computational operation that takes place on a lattice as shown in Figure 2.1.

Figure 2.1 - GPA operations.



Scheme representing the GPA operations and respective Gradient Moments (GM). The set of operations are denoted by I-VIII, where the features V - VIII are the Gradient Moments measures.

In this formalism, the lattice preprocessing involves four key operations, denoted as I-IV. Operation I entails measuring the gradient, which involves computing local

differences in each direction. The literature commonly adopts the Central Finite Differences (CFD) scheme for its simplicity and effectiveness (ROSA et al., 1999; BARCHI et al., 2020). In specific cases where boundary points are involved, either Forward Finite Differences (FFD) or Backward Finite Differences (BFD) are employed. A summary of this operation is presented in Table 2.2.1.

Table 2.1 - Finite difference scheme for computing gradient lattice.

Numerical scheme	$\partial_x(A_{i,j})$	$\partial_y(A_{i,j})$
CFD	$(A_{i+1,j} - A_{i-1,y})/2$	$(A_{i,j+1} - A_{i,j-1})/2$
FFD	$A_{i+1,j} - A_{i,j}$	$A_{i,j+1} - A_{i,j}$
BFD	$A_{i,j} - A_{i-1,j}$	$A_{i,j} - A_{i,j-1}$

Central Finite Difference (CFD) scheme, Forward Finite Difference (FFD), and Backward Finite Difference (BFD). In a numerical lattice with amplitudes $A_{i,j}$, the vectors are computed using CFD, and FFD or BFD for the boundaries. $\partial_x(A_{i,j})$ is the spatial derivative in x-direction and $\partial_y(A_{i,j})$ is the spatial derivative in y-direction.

Operation II involves the removal of symmetric vectors from the lattice, as detailed in Section 2.2.2. Operations III and IV are focused on analyzing specific properties of each vector within the lattice.

In Operation III, the modulus and phases of each vector are measured to understand its characteristics. The modulus, denoted as $||v||$, represents the vector's length and is computed as the Euclidean distance: $||v|| = \sqrt{[\partial_x(A_{i,j})]^2 + [\partial_y(A_{i,j})]^2}$. This provides valuable information about the intensity of amplitude change in the vector's neighborhood, a measure that defines GM2.

On the other hand, the phase $\phi_{i,j}$ indicates the direction in which the amplitude of the lattice is locally increasing and is expressed as an angle, forming GM3. The phases are calculated using the function $\phi_{i,j} = \text{atan2}(\partial_y(A_{i,j}), \partial_x(A_{i,j}))$, where atan2 is the arc tangent function that solves the ambiguity of measuring an angle.

Finally, Operation IV involves the recombination of moduli and phases in a complex form: $z_{i,j} = |v_{i,j}|\exp(\sqrt{-1}\theta_{i,j})$, generating the GM4.

The set of Gradient Moments (GM) obtained from these operations allows the calculation of a comprehensive metric, as outlined in Section 2.3.

2.2.2 Breaking bilateral symmetry

The literature offers numerous definitions of symmetry (LIU et al., 2010). These definitions revolve around the concept of elements remaining invariant under specific operations, such as translation, rotation, or reflection. In the literature of GPA (BARCHI et al., 2020; ROSA et al., 2018), the criterion for spatial bilateral symmetry has been established. According to this criterion, elements that are equidistant from a separation axis and perpendicular to it are considered potential candidates for symmetry. In the gradient field, each element is characterized by a modulus and phase, which determine whether the candidates exhibit actual symmetry. If two candidate vectors possess equal modulus but opposite phases, they are deemed symmetrical.

More specifically, any pair of vectors that are equidistant from a designated center of symmetry can be considered candidates for symmetry. The axis of symmetry is determined by the midpoint between two equally distant points and the center of symmetry. Additionally, vectors with the same modulus but opposite phases result in a vectorial sum of zero.

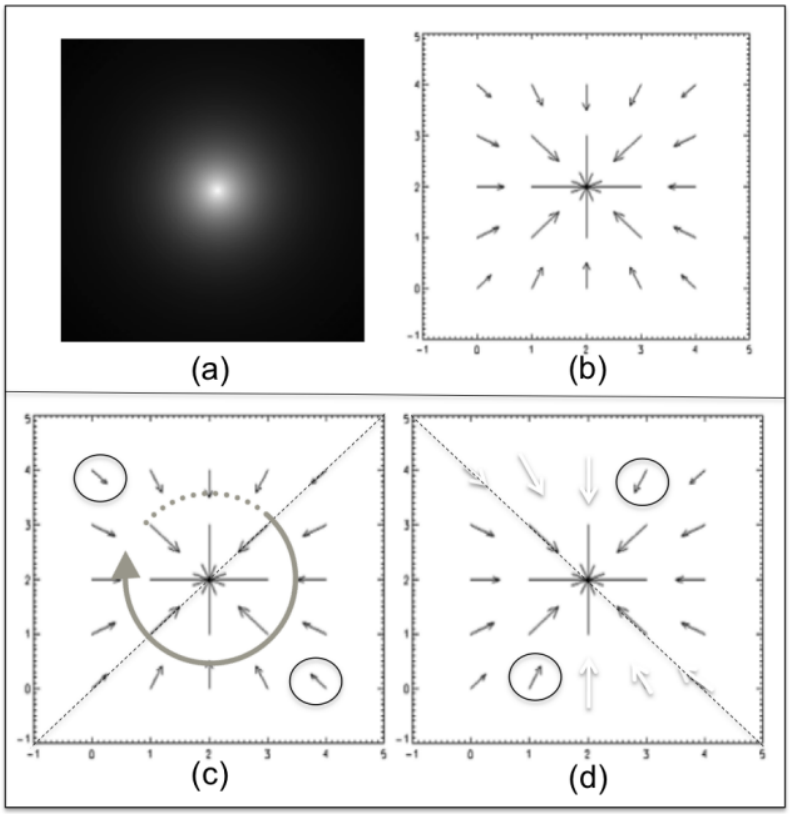
Hence, an algorithm specifically crafted to detect symmetrical pairs operates by first identifying vectors equidistant from the center of symmetry. It proceeds to check whether each vector has a modulus greater than a defined tolerance parameter. Vectors with a modulus below the tolerance are labeled as 'unknown symmetry.' For vectors surpassing the tolerance, the algorithm further examines if the modulus of their sum falls below a specified threshold. Figure 2.2 provides a visual representation of this procedure.

2.3 Gradient moments

To characterize an asymmetrical vector lattice, many measures can be taken to highlight or diminish certain features of the pattern. It is important to have a variety of measures available, in order to represent all the aspects that a given pattern may provide. To this end, Rosa et al. (2003) proposed a measure classification based on the representation of the gradient lattice, which was named the Gradient Moment (GM) due to its similarity to statistical moments. This classification is based on the steps to generate every gradient lattice representation, as shown in Figure 2.1.

With GMs, it is possible to track the trends of the gradient lattice, as well as determine the presence of anomalies. Ultimately, these measures can be used to

Figure 2.2 - Algorithm for symmetry detection.



Detection of symmetry based on the concentric criteria. Vectors equidistant from the center of symmetry are summed to determine their symmetry classification.

SOURCE: Rosa et al. (2018).

gain a deeper understanding of a given vector lattice and help discover patterns that may otherwise be hidden.

The first Gradient Moment (GM1) deals with the characteristics of vectors, including their phases, modules, and positions. In order to aggregate all of these characteristics, Rosa et al. (1999) proposed an intermediate operation: for each vector in the lattice, the position is added to the vector, and a triangulation is performed. Of the various triangulation methods, the Delaunay triangulation was chosen as the standard due to its maximization of interior angles. Finally, the ratio between the number of connections in the Delaunay triangulation, N_C, and the number of vertices, N_V, is used to compose the measure, as described in Equation 2.4.

$$G_1^C = \frac{N_C - N_V}{N_V} \tag{2.4}$$

This metric is an important feature in terms of local perturbations. However, there is a drawback in applications with large matrices, as in the following section. In order to achieve a more robust measure, it is essential to utilize more refined data. In this context, additional moments of the gradient field have been developed, which consider a reduced number of gradient properties.

The GM2 metrics incorporate information related to the vectorial modulus, obtained by operation (ii) in Figure 2.1. From the moduli set, we can compute the metric from Equation 2.5.

$$G_2 = \frac{V_A}{V}\left(1 - \frac{\left|\sum_{i=0}^{V_A} v_i^A\right|}{2\sum_{i=0}^{V_A}|v_i^A|}\right) \tag{2.5}$$

In this Equation, V_A is the number of asymmetrical vectors, V is the total number of vectors in the lattice, and v_i^A is an asymmetrical vector. Equation 2.5 is a normalized version of the classical measure G_2^C, proposed by Rosa et al. (2018), which is bounded between $[0, 2]$ to match the same range as G_1^C. Notably, $G_2 = G_2^C/2$, preserving the constant values in a new scale. For example, previous studies have reported that $G_2^C(L_P) = 1$ for a laminar pattern L_P, while the metric for the random pattern (U) for both versions are $G_2^C(U) = 2$ and $G_2(U) = 1$, respectively.

Unlike the other GMs, no metric for GM-3 has been previously defined in the literature until a recent study (SAUTTER et al., 2023). Arithmetic operations involving phases exhibit substantial distinctions from conventional numerical sets. For example, the summation operation of an angle encounters a challenge near $\theta = 0 = 2\pi$ radians. This issue also arises in defining the distance between phases; an angle of $\phi_1 = \pi/2$ radians is considered more distant to θ than a phase angle of $\phi_2 = 5\pi/3$ radians. To address this issue, a simpler approach involves operations of unitary vectors within a polar coordinate system.

This representation neither incorporates new information regarding the original vector nor excludes any phase-related information. When computing the phases of a vector using the atan2 function, the reference point is the x-axis. However, as the

Figure 2.3 - Example of G_3 angles.

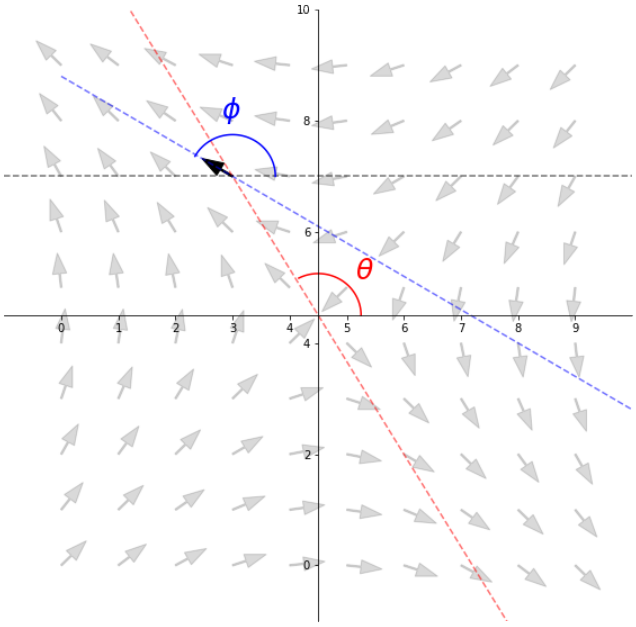


Illustration of angles used in the calculation of G_3.

interpretation of the x-axis may vary in different applications, it becomes crucial to adjust the reference basis of the phases accordingly. One approach to achieve this adjustment is by taking into account the vector's position relative to the lattice center and transforming it using the dot product, as illustrated in Figure 2.3. This transformation is formulated in Equation 2.6, where $u_i = (cos(\phi_i), sin(\phi_i))$ denotes a unitary vector related to the phase, and $v_i = (cos(\theta_i), sin(\theta_i))$ represents a unitary vector representing the position relative to the lattice center.

$$G_3 = \frac{1}{2}\left(\frac{V_A}{V} + \frac{1}{2V_A}\sum_{i=0}^{i=V_A} u_i.v_i + 1\right) \tag{2.6}$$

In this manner, G_3 expresses the collective direction of all vectors with respect to the center of the lattice. Individually, the phase and modulus provide detailed information about the variation in location and intensity. However, in certain applications, the combined information becomes crucial. In this regard, (RAMOS et al., 2000) introduces an entropic metric based on the complex notation of each vector. Equation 2.7 presents the entropy of the vector set $z_i = |v_i|\exp(i\theta_i)$, where $|v_i|$ represents the modulus and θ_i represents the phase. In order to normalize this metric, Rosa et

al. (2003) proposes a normalization considering the sum of all vectors, which can be expressed as $z = \sum_i^{V_A} z_i$.

$$G_4 = -\sum_{i=0}^{V_A} \frac{z_i}{z} \ln\left(\frac{z_i}{z}\right) \tag{2.7}$$

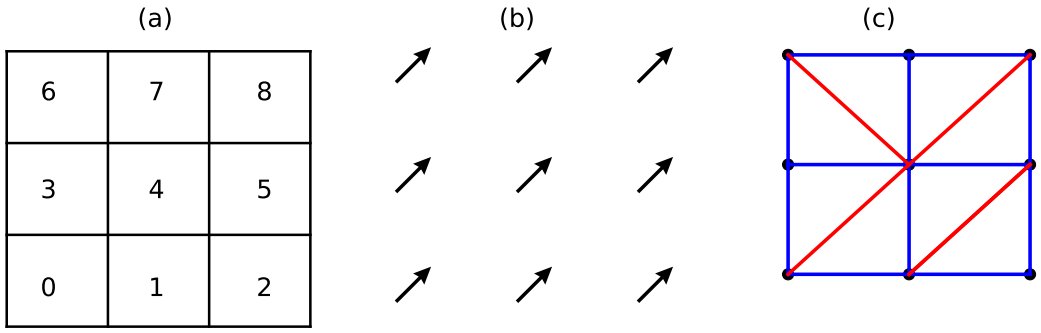
Ramos et al. (2000) also proposes a Stirling approximation to the logarithm function for complex numbers. However, recent implementations of Python provide an optimized implementation for the complex logarithm operation.

From the definition of GM, we can perform the measurement with the GPA technique. However, the interpretation of the values produced is extremely important for data analysis. In this regard, the patterns presented in the following sections encompass controlled experimental scenarios, including laminar patterns, random patterns, and colored noise patterns, which serve as informative benchmarks for understanding the capabilities and outcomes of GPA analysis.

2.3.1 Laminar pattern

One of the most important canonical pattern is the laminar flow (L_P). This pattern is a matrix with constant increments in the x and y directions and has a fully asymmetrical gradient field with constant phases and moduli. The minimal example of this pattern and the corresponding gradient field and triangulation can be seen in Figure 2.4. In this section, we will present analytical and empirical solutions for all GMs in the case of a laminar flow of a square grid. We will first provide an in-depth analysis of both GM1 and GM2 to illustrate their corresponding solutions for the first time. Then move on to the empirical solution for all GMs.

Figure 2.4 - Laminar Pattern (L_P).



A laminar pattern is defined as a matrix with constant increments in every direction, such as the matrix (a). Thus, the gradient field (b) contain constant size vectors pointing to the same direction. The corresponding triangulation field (c) is build from the point of the vectors. In the case of laminar pattern, there are two kinds of connections in triangulation field: regular grid connections in blue and diagonal connections in red.

Let's consider the minimal example of L_P Delaunay triangulation (see Figure 2.4), where there are diagonal connections and regular grid connections. The total number of triangulation connections N_C in a square matrix, given the matrix order N, is expressed by Equation 2.8. The term $2N(N-1)$ is relative to the regular grid of the triangulation and $(N-1)^2$ is relative to the diagonal connections.

$$N_C = 2N(N-1) + (N-1)^2 \tag{2.8}$$

This equation can be expanded to yield Equation 2.9.

$$N_C = 3N^2 - 4N + 1 \tag{2.9}$$

Considering $N_V = N^2$ and Equation 2.9, the value of $G1^C(L_P)$ is found to be inversely proportional to the matrix size, which can be described as Equation 2.10.

$$G_1^C(L_P) = 2 - \frac{4}{N} + \frac{1}{N^2} \tag{2.10}$$

This indicates that for large enough matrices, the laminar pattern becomes indistinguishable from the random pattern since $\lim_{N\to\infty} G_1^C(L_P) = 2$.

Referring to Equation 2.5 introduced in the previous section, it is evident that the metric G_2 is directly related to the alignment of vectors and the proportion

of asymmetrical vectors. The laminar pattern's vectorial summation, denoted as $|\sum v_i| = V_A \sum |v_i|$, leads to the expression $G_2 = V_A/(2V)$. In the case of a laminar pattern where all vectors are inherently asymmetrical, the ratio V_A/V equals 1. Consequently, the metric of the laminar pattern is the constant $G_2(L_P) = 1/2$.

The metric G_3 contain information related to the phases of the system. Consequently, this metrics is not constant with respect to the matrix size for the laminar pattern. However, when a constant phase is considered, these metrics assume constant values. Therefore, an analytical test can be employed in this scenario. An analysis with all GMs is presented in Table 2.3.1. This example serves to demonstrate the application and calculation of metrics in the context of laminar patterns generated with the unitary differences between the elements in the x-axis.

Table 2.2 - Gradient Moments for Laminar Pattern.

Matrix size	G_1^C	G_2	G_3
3x3	0.778	0.500	0.759
5x5	1.240	0.500	0.752
9x9	1.568	0.500	0.750
17x17	1.768	0.500	0.750
33x33	1.880	0.500	0.750
65x65	1.939	0.500	0.750
129x129	1.969	0.500	0.750

2.3.2 Random patterns

Random matrices serve as another crucial example of canonical patterns in GPA context. The literature on GPA includes at least two types of random matrices (ANDRADE et al., 2006; ROSA et al., 1999). The first type involves random matrices generated from a uniform distribution, while the second type entails matrices created by shuffling laminar matrix elements. Overall, there is no statistical difference between the GMs of both random matrices. The corresponding average and standard deviation for the random matrices can be found in Table 2.3.2. A set of 100 random matrices, of size $2^N + 1$ elements, has been applied to this test.

Table 2.3 - Gradient Moments for Shuffled Laminar Pattern.

Matrix size	G_1^C	G_2	G_3
3x3	1.081 ± 0.154295	0.792 ± 0.132010	0.722 ± 0.087462
5x5	1.613 ± 0.053446	0.889 ± 0.058783	0.744 ± 0.054518
9x9	1.867 ± 0.017766	0.950 ± 0.024025	0.749 ± 0.021603
17x17	1.960 ± 0.005387	0.979 ± 0.011278	0.750 ± 0.007761
33x33	1.988 ± 0.001477	0.991 ± 0.004227	0.750 ± 0.002983
65x65	1.997 ± 0.000316	0.997 ± 0.001397	0.750 ± 0.001474
129x129	1.999 ± 0.000119	0.999 ± 0.000529	0.750 ± 0.000718

It is worth noting that G_3 exhibits the same constant value for both types of random matrices and the laminar pattern, indicating a limitation of this metric. Conversely, both G_1 converge to a value of 1 as the matrix size increases. Additionally, G_1^C tends to approach a value of 2, a finding that has been previously observed by Rosa et al. (1999).

The randomly generated pattern is commonly referred to as white noise, a term that holds significance in comprehending the underlying physics of the system. This classification of noise types serves as an additional benchmark for GPA. In the subsequent section, we delve into the theory of colored noise and present an algorithm for generating colored noise, enabling a deeper exploration of noise patterns within the GPA framework.

2.3.3 Colored noise

Understanding noise is essential since it is ubiquitous. Furthermore, the noise signal may indicate important characteristics of the physics involved in the observed phenomenon and flaws in the data acquisition process. Thus, generating synthetic noise is important to understand the data and test analysis tools. In this regard, different techniques are applied depending on system stability requirements, the spatiotemporal domain, observational signal-to-noise ratio (SNR), and computational resources. There are two main families of noise signals established in the literature; the colored noise signal and the procedural noise.

Colored noise signals are typically described by a log-log relationship between frequency and energy content at the given frequency, which is known as Power Spectrum Density (PSD). The PSD of most noise signals follows a power law of the form $PSD(f) = 1/f^{\beta}$, where β is the power law exponent. Replicating noises with specific

PSD decays (β) is a challenging task, particularly for $\beta \geq 3$ (XU, 2019). Various techniques have been proposed to generate colored noise signals, including filtering of white noise (LANGBEIN, 2004), generation and application of a function that approaches $1/f^{\beta}$ (TIMMER; KOENIG, 1995), or a combination of different techniques (XU, 2019).

The family of Procedural Noise is popular in applications that require spatial noise modeling, which include: cloud mimic, textures generation, generation of geographical formation, and visualization of galaxy gas distribution. These techniques generate noise from a regular grid (simplex grid or a traditional grid), which are then interpolated to generate the noise (LAGAE et al., 2010). The produced noise is visually similar to natural phenomena; however, it has a high computational cost.

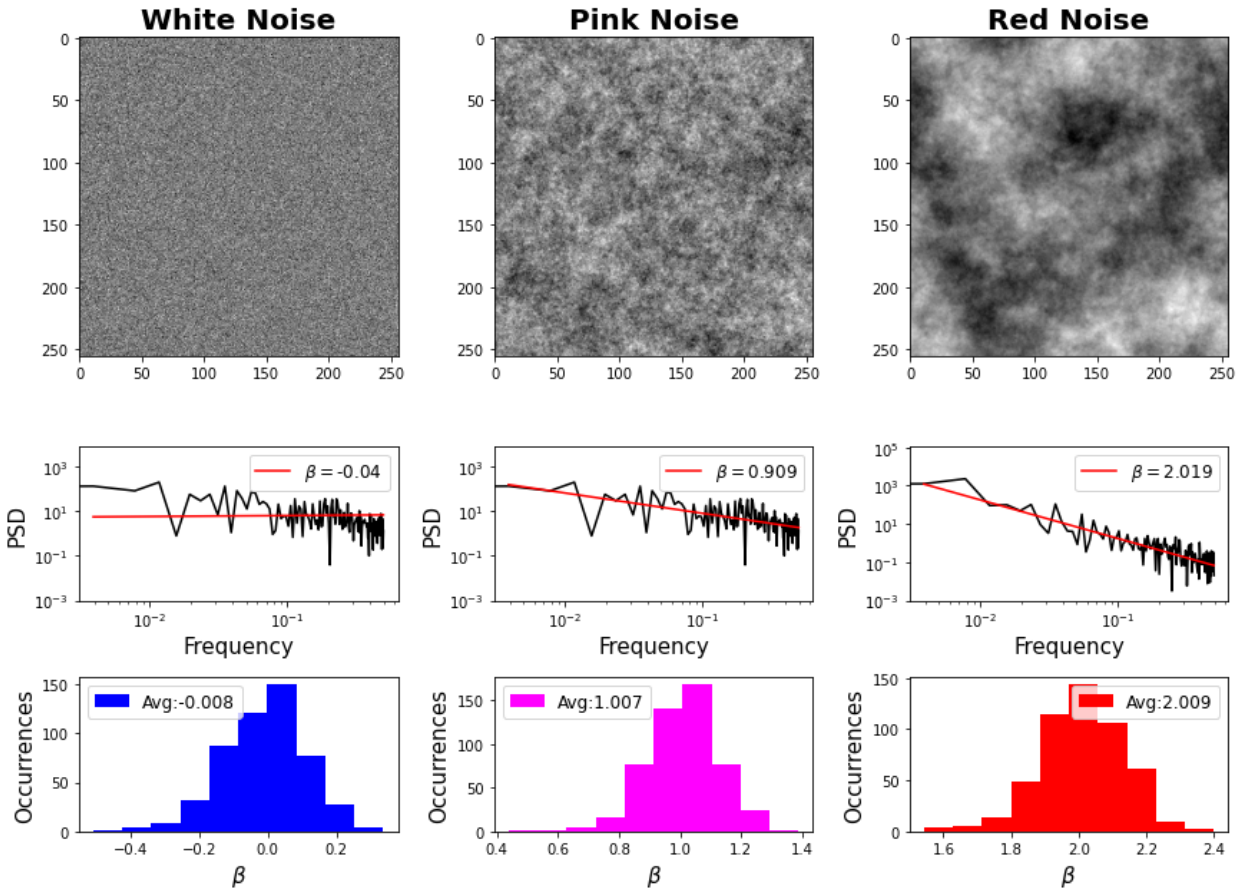
In this regard, we introduce a new technique which is a generalization of the squared-root algorithm (TIMMER; KOENIG, 1995), that is similar to Procedural Noises but in a multidimensional sense. As the squared-root noise algorithm, we sample a random Gaussian set with zero average and standard deviation (σ). A new term s is introduced, which is fitted to ensure that every 0D+1 slice of the noise has the β parameter as a log-log linear rule. The following procedure is applied to fit the parameter and generate the noise.

Algorithm 1 Multidimensional Noise Generator

Require: β, ξ, υ
Ensure: f(t)
  $s \leftarrow +\sqrt{2}^{D}$
  $a \leftarrow +\infty$
  while $a > \xi$ **do**
    $\hat{f}(\omega) \leftarrow \eta(0,1)\omega^{\frac{-s\beta}{2}}$
    $f(t) = \Re(\text{IFFT } \hat{f}(\omega))$
    $b \leftarrow [\,]$
    for all $\rho, \omega \in$ slice of FFT(f(t)) **do**
      $\alpha \leftarrow$ coefficient of log-log linear fit(ρ, ω)
      $b \leftarrow [b, \alpha]$
    end for
    $a \leftarrow average(b - \beta)$
    $s \leftarrow s - \upsilon a$
  end while
  $f(t) \leftarrow f(t) - f(t)/max(|f(t)|)$ ▷ Normalize the output

In this procedure, β is the desired PSD rule, ξ is the tolerance of the generated noise average β, and υ is the step size of the optimization.

Figure 2.5 - Sample of colored noise.



Examples of outputs from the proposed colored noise algorithm. The first row shows a $2D + 1$ sample for each type of colored noise; the second row are the middle slice PSD fitted and third row are the distribution of fitted PSD. This is an example from a set of 200 samples for each colored noise.

The analysis of red noise ($\beta = 2$) is presented in Table 2.3.3, where one can see the same tendency in noise terms.

Table 2.4 - Gradient Moments for Red Noise.

Matrix size	G_1^C	G_2	G_3
3x3	0.970 ± 0.170718	0.684 ± 0.173975	0.754 ± 0.112515
5x5	1.590 ± 0.075505	0.933 ± 0.054599	0.744 ± 0.078924
9x9	1.837 ± 0.029571	0.976 ± 0.021875	0.752 ± 0.053197
17x17	1.943 ± 0.010992	0.992 ± 0.005938	0.751 ± 0.035004
33x33	1.982 ± 0.003310	0.997 ± 0.002475	0.748 ± 0.019820
65x65	1.995 ± 0.000853	0.999 ± 0.000721	0.750 ± 0.013090
129x129	1.999 ± 0.000205	1.000 ± 0.000286	0.750 ± 0.010874

The context presented in this chapter encompasses crucial aspects of the GPA technique, including its interpretation and limitations. Building upon this foundation, the following chapter delves into the exploration of GPA within methodology of this work, which includes improvements to GPA technique and the details of the numerical methods.

3 METHODOLOGY

This chapter is dedicated to the discussion of the methodology employed in generating synthetic data and the analysis technique's generalization for studying ordinary time series and spatiotemporal data. While some numerical methods presented here have been previously applied to various problems, such as the pseudospectral method's application to the Ginzburg-Landau equation(ZENG et al., 2019; FARAG et al., 2022), the innovation lies in the novel combination of methods and the introduction of noise into this approach. As a result, this chapter highlights the principal innovations introduced by this thesis.

3.1 Complex Ginzburg-Landau numerical methods

Recently, a pseudospectral method has emerged as a promising solution for solving the CGL equation. This method offers significant advantages over other numerical approaches in terms of accuracy (ZENG et al., 2019; FARAG et al., 2022). The pseudospectral method employs a series of polynomial or non-linear functions to approximate the signal. In the context of spatial derivatives, the Fourier Transform is commonly used, where the signal is represented as a series of sines and cosines. This choice allows for the utilization of the efficient Fast Fourier Transform (FFT) algorithm, albeit with the requirement of periodic boundary conditions.

Applying Euler's formula, the sines and cosines can be expressed in complex form, as depicted in Equation 3.1, enabling the representation of the signal in terms of frequency. Specifically, if we consider a sampled signal denoted by $f(t)$, we can approximate it using a set of complex sines and cosines expressed as $\hat{f}(\omega)e^{-i2\pi\omega t}$, where $\hat{f}(\omega)$ is the transformed signal, ω represents the frequency, and t denotes the normalized time.

$$e^{ix} = sin(x) + icos(x) \tag{3.1}$$

Using the Gauss method, we minimize the quadratic error as follows:

$$\frac{d}{d\hat{f}(\omega)} \sum_{\omega} \sum_{t} \left(f(t) - \hat{f}(\omega)e^{-i2\pi\omega t} \right)^2 = 0. \tag{3.2}$$

Considering the orthogonality property of the series, the minimization results in the

Fourier transform:

$$\hat{f}(\omega) = \int_T f(\omega) e^{-i2\pi\omega t} dt, \tag{3.3}$$

where ω is a frequency parameter, t is the time, $f(t)$ is the original signal and $\hat{f}(\omega)$ is the transformed signal. The inverse Fourier transform is given by:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{i2\pi\omega t} d\omega, \tag{3.4}$$

that can be derived with respect to the time on both hand sides, producing:

$$\frac{df(t)}{dt} = \frac{1}{2\pi} \int_{-\infty}^{\infty} (i2\pi\omega) \hat{f}(\omega) e^{i2\pi\omega t} d\omega. \tag{3.5}$$

This approach can be applied iteratively to obtain higher-order derivatives. In the case of CGLE, the second-order derivative is necessary to describe the diffusion process ($\nabla^2 A$), which can be approached with the pseudospectral method as:

$$\frac{d^2 f(t)}{dt^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} (-4\pi^2\omega^2) \hat{f}(\omega) e^{i2\pi\omega t} d\omega. \tag{3.6}$$

It is important to highlight that the chosen approach can be susceptible to numerical instability in the vicinity of discontinuities, which is known as the Gibbs effect (LIU, 1999; DJEDDI; EKICI, 2016). Although limiting high-frequency terms in the derivative calculation can mitigate this issue, it comes at the cost of reduced precision. Instead, we propose employing the Runge-Kutta-Fehlberg 4(5) integration method (SHAMPINE, 1986) to address the problem of instability. This method extends the traditional Runge-Kutta method and dynamically adjusts the integration time-step based on the discrepancy between an $O(h^4)$ approximation and an $O(h^5)$ approximation. The integration steps used in this method are given by Equation 3.7.

$$
\begin{aligned}
k_1 &= \Delta t f(t_k,\ y_k) \\
k_2 &= \Delta t f\left(t_k + \tfrac{1}{4}\Delta t,\ y_k + \tfrac{1}{4}k_1\right) \\
k_3 &= \Delta t f\left(t_k + \tfrac{3}{8}\Delta t,\ y_k + \tfrac{3}{32}k_1 + \tfrac{9}{32}k_2\right) \\
k_4 &= \Delta t f\left(t_k + \tfrac{12}{13}\Delta t,\ y_k + \tfrac{1932}{2197}k_1 - \tfrac{7200}{2197}k_2 + \tfrac{7296}{2197}k_3\right) \\
k_5 &= \Delta t f\left(t_k + \Delta t,\ y_k + \tfrac{439}{216}k_1 - 8k_2 + \tfrac{3680}{513}k_3 - \tfrac{845}{4104}k_4\right) \\
k_6 &= \Delta t f\left(t_k + \tfrac{\Delta t}{2},\ y_k - \tfrac{8}{27}k_1 + 2k_2 - \tfrac{3544}{2565}k_3 + \tfrac{1859}{4104}k_4 - \tfrac{11}{40}k_5\right)
\end{aligned}
\tag{3.7}
$$

With the pre-steps, the O(h^4) approach and the O(h^5) approach are given by Equation 3.8.

$$
\begin{aligned}
y^4_{t+\Delta t} &= y^t + \tfrac{25}{216}k_1 + \tfrac{1408}{2565}k_3 + \tfrac{2197}{4101}k_4 + \tfrac{1}{5}k_5 \\
y^5_{t+\Delta t} &= y_t + \tfrac{16}{135}k_1 + \tfrac{6656}{12825}k_3 + \tfrac{28561}{56430}k_4 - \tfrac{9}{50}k_5 + \tfrac{2}{55}k_6
\end{aligned}
\tag{3.8}
$$

If the largest difference between both estimations is greater than a tolerance parameter τ, then the new step can be updated according to Equation 3.9.

$$
\Delta t' = \Delta t \left(\frac{\tau}{max\{|y^4_{t+\Delta t} - y^5_{t+\Delta t}|\}} \right)^{\frac{1}{4}}, \tag{3.9}
$$

Once the numerical methods have been established, it becomes necessary to define the initial conditions for the system. In the subsequent section, we present a stability analysis that not only enhances the determination of optimal initial conditions but also complements the GPA technique.

3.1.1 Initial conditions

The analysis of the reactive component of equations plays a pivotal role in comprehending the stability of system regimes, identifying optimal initial conditions, and detecting fixed points. While this analysis has been presented in various forms, it does not consider any spatial transport mechanism. In the case of the classical Com-

plex Ginzburg-Landau (CGL) equation, we can express it as Equation 3.10, where A_r represents a complex number and c denotes the reactive dispersion parameter.

$$\partial A_r = A_r - (1 + ic)|A_r|^2 A_r \tag{3.10}$$

To find equilibrium points, one should consider the derivative zero, as shown in Equation 3.11, that solution is the origin.

$$A_r - (1 + ic)|A_r|^2 A_r = 0 \tag{3.11}$$

To determine whether a equilibrium point is a source or a sink, one should consider the derivative of the reaction, which is given by Equation 3.12.

$$\partial^2 A_r = 1 - 3(1 + ic)A_r^2 \tag{3.12}$$

As the derivative of the reaction at the origin is positive, it serves as a source-fixed point. However, in the presence of a single source-type fixed point, it is necessary to verify the existence of limit cycles, as the absence of such cycles would result in a divergent solution. To determine the limit cycle of the reactive part of the CGL, we can use the complex notation as shown in Equation 3.13, where $|a|$ is the complex number's absolute value, and θ is its phase.

$$A_r = |a|e^{i\theta} \tag{3.13}$$

Thus, the reactive term given in Equation 3.10 can be expressed as Equation 3.14.

$$dA_r = |a|e^{i\theta} - (1 + ic)|a|^3 e^{i\theta}. \tag{3.14}$$

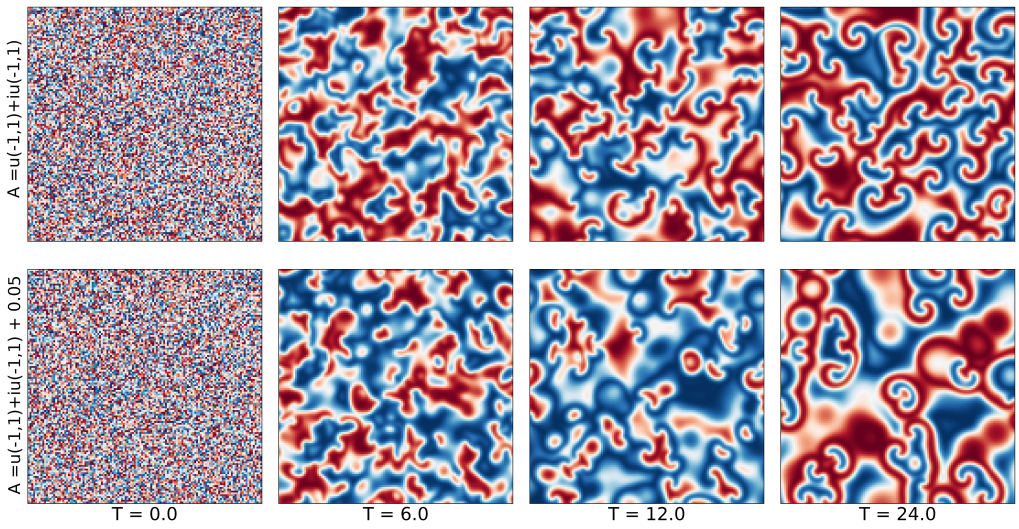
Using the Euler's formula (Equation 3.1), the reaction in Equation 3.14 can be written as Equation 3.15, where the term $|a| - |a|^3$ is zero for $|a| = 0$ and $|a| = 1$.

$$dA_r = c\cos\theta + (|a| - |a|^3)\sin\theta + i(-c\sin\theta + (|a| - |a|^3)\cos\theta) \tag{3.15}$$

$$dA_r = c(\sin\theta - i\cos\theta) \tag{3.16}$$

Hence, the reaction at a unitary distance can be described by Equation 3.16, which exhibits a clockwise orthogonal rotation with respect to the position. Consequently, the circumference $|A_r| = 1$ forms a closed path, and any set of points in proximity to $|A_r| \approx 1$ is conducive to rapid convergence towards the pattern. To illustrate the speed of pattern formation, we present an example of system iteration in Figure 3.1. In this example, the simulation starts with matrices with random initial conditions of $A = u(-1,1) + iu(-1,1)$ and $A = u(-1,1) + iu(-1,1) + 0.05$, where $u(-1,1)$ represents a uniform white noise ranging between -1 and 1.

Figure 3.1 - Convergence time comparison of Complex Ginzburg-Landau.

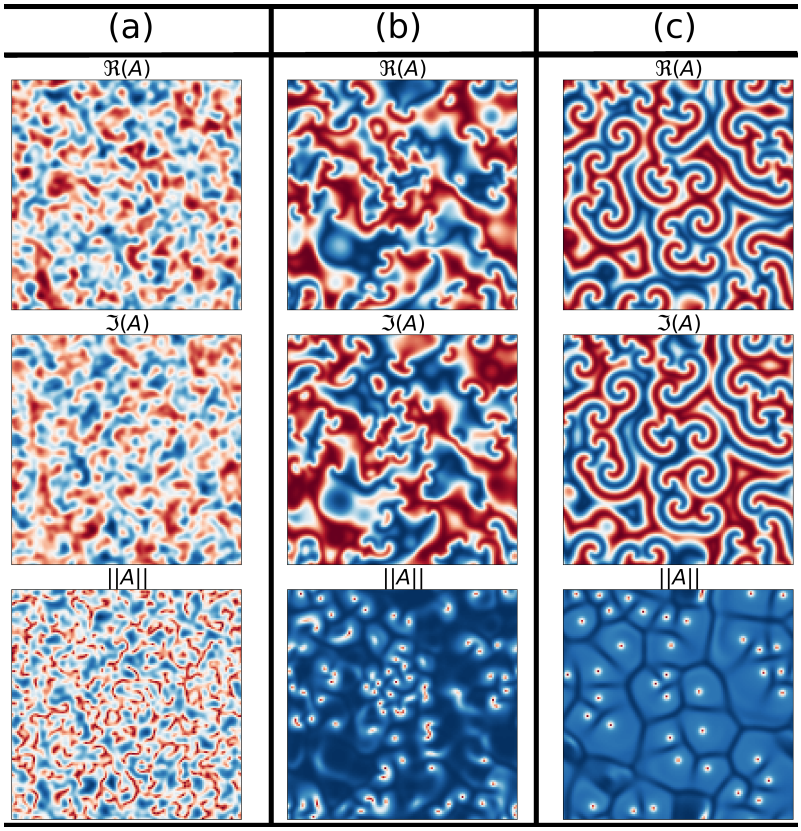


Sequences of snapshots of CGL real part without noise terms, starting from random initial conditions, where $u(-1,1)$ represents a uniform noise that ranges between -1 and 1. In this example, we computed 2,000 iterations using the parameters $c = 1.5$ and $b = 0.0$. The system with an initial condition of $A = u(-1,1) + iu(-1,1)$ (sequence at the top) forms the pattern more quickly than the system with an initial condition of $A = u(-1,1) + iu(-1,1) + 0.05$ (sequence at the bottom).

It is important to note that for longer-period interactions, using the appropriate initial conditions, at least three pattern formation regimes emerge. These are known as: (a) accretion, (b) spiral formation and (c) spiral relaxation, as shown in Figure 3.2. Note that the inclusion of noise terms can affect the spatiotemporal stability

of these regimes by accelerating transitions and/or changing the structure of the pattern.

Figure 3.2 - Complex Ginzburg-Landau pattern formation regimes.



Complex Ginzburg-Landau pattern formation regimes: (a) Accretion ($T = 2.00$), (b) Spiral formation ($T = 12.00$), and (c) Spiral relaxation ($T = 57.00$). In this example, the system has been iterated in absence of noise term. Note that the three regimes occur in all representation domains: real (top), imaginary part of the complex (middle), and modulus A (bottom).

However, performing the same analysis for the SCGL equation poses a challenge as the noise terms, $\sigma\delta_t(\eta_\beta)$ and $\sigma A\delta_t(\eta_\beta)$, exhibit location and time dependencies. To address this issue, The noise term used in this application normalized between -1 and 1, according to the noise algorithm shown in Section 2.3.3. Overall, the parameters applied to this study are shown in Table 3.1, which describes the formation of double-spiral patterns (ARANSON; KRAMER, 2002).

Table 3.1 - Input parameters for the SCGL model.

Parameter	**Symbol**	**Value**
Spatial dispersion parameter	b	0.00
Reaction dispersion parameter	c	1.50
Stochastic strength	σ	100.00
Power-law	β	2.00
Noise tolerance	ξ	0.01
Noise optimization step	ν	0.60
Time-step	Δt	0.05
Spatial-step	Δh	1.00
Integration tolerance	τ	10^{-4}
Grid-size	L	128x128
Initial condition	A_0	$\eta + i\eta$

3.1.2 Stability analysis

Understanding the influence of colored noise on system dynamic stability is a significant concern. The reactive part of the SCGL model includes a noise term that poses challenges in terms of stability modeling. To investigate the noise's impact on system stability, we propose a Lyapunov analysis focused on the reactive component.

The reactive part of the SCGL equation, incorporating both additive and multiplicative noise terms, is expressed by Equations 3.17 and 3.18, where A_r is a complex number, η_β represents 0D+1 noise, c is the reactive dispersion parameter, and σ is the stochastic strength parameter. The analysis is conducted in the reactive phase space, as depicted in Figure 3.3. A similar representation with noiseless high-order reactive terms has been previously presented in the literature (BYRNE et al., 2015).

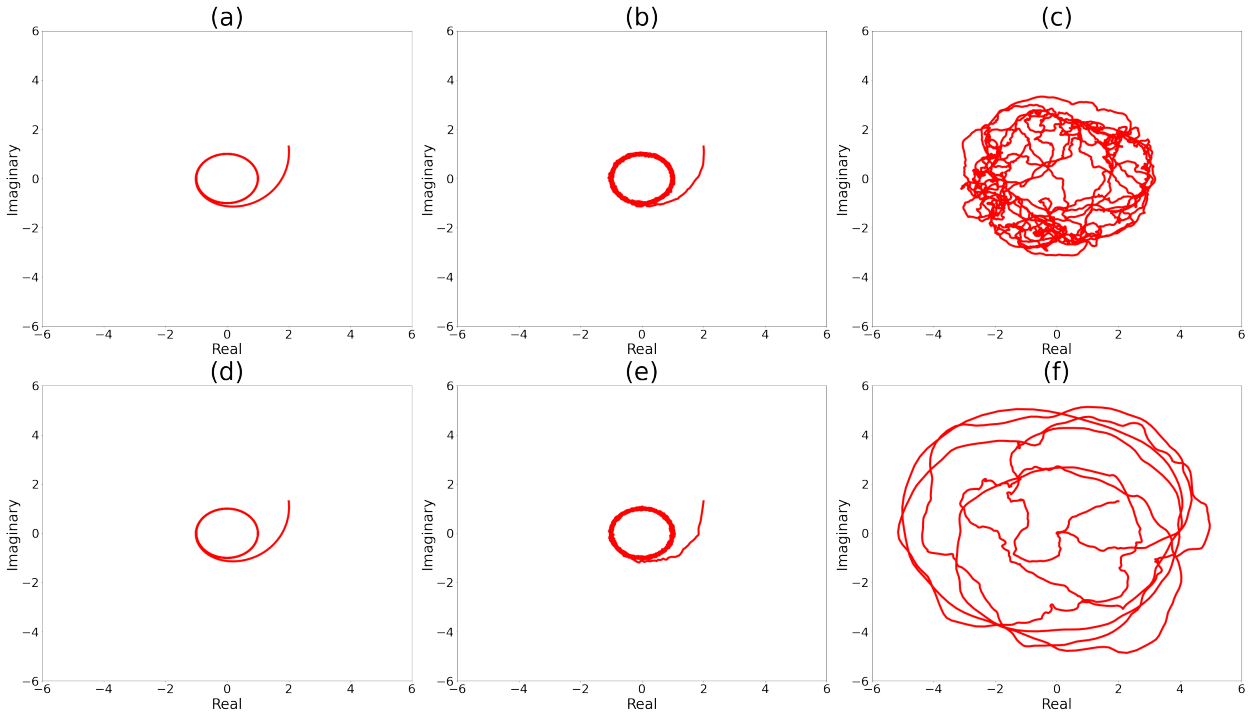
$$\partial A_r = A_r - (1 + ic)|A_r|^2 A_r + \sigma d(\eta_\beta) \tag{3.17}$$

$$\partial A_r = A_r - (1 + ic)|A_r|^2 A_r + \sigma A_r d(\eta_\beta) \tag{3.18}$$

In Figure 3.3, it is presented three examples of iterations with colored noises. The observations clearly demonstrate that significant instabilities only arise when red noise is present, while the system remains bounded in all scenarios. Additionally, an interesting distinction emerges between additive and multiplicative red noise. Specifically, iterations involving multiplicative noise exhibit fewer crossings near

the stable point at $A_r = 0 + 0i$ compared to iterations involving additive noise. This finding suggests that the presence of multiplicative noise does not hinder the formation of patterns.

Figure 3.3 - Lyapunov paths for Reactive SCGL.



The Lyapunov paths are presented in red for Reactive Noisy CGL in the complex plane. (a), (b), and (c) depict solutions with respective white, pink, and red additive noise terms, while (d), (e), and (f) represent solutions with respective white, pink, and red multiplicative noise terms. The red noise increases the path asymmetric fluctuations as well as its ergodicity area.

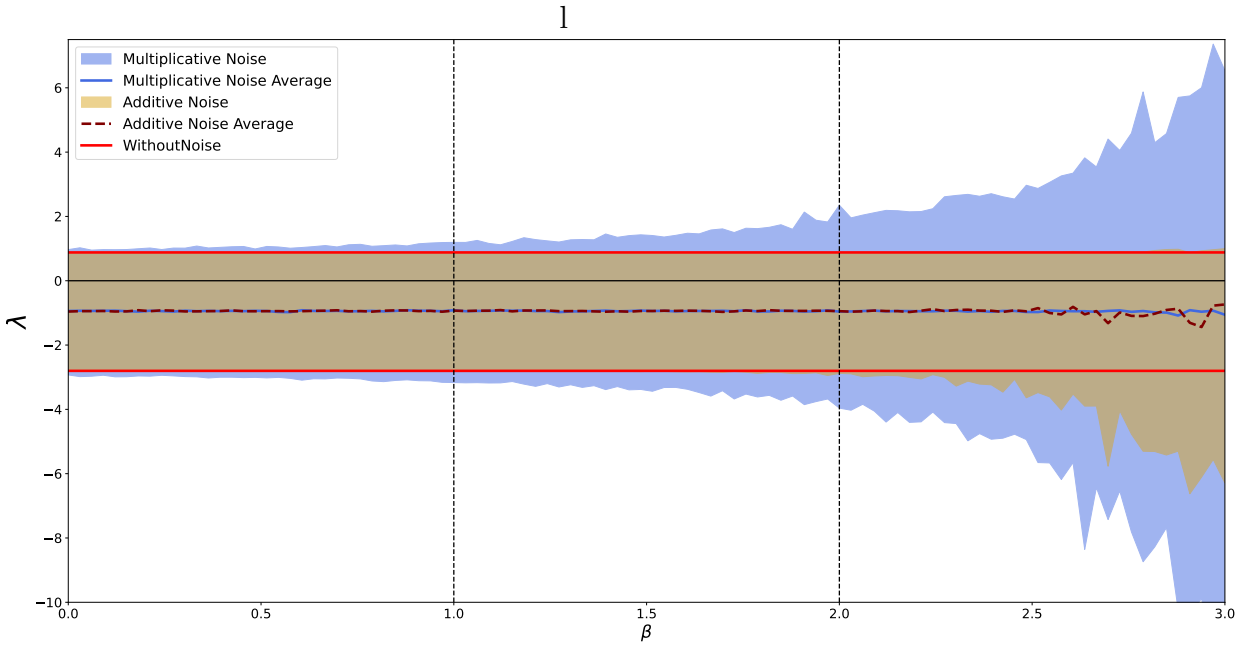
To characterize the influence of colored noise on dynamic stability, we employ a Lyapunov analysis based on the reactive part of the system. The phase space considers the real and imaginary parts of the system as free variables. By introducing noise as an additional free variable, we can test the hypothesis of chaos in the reaction terms using the Lyapunov exponent (λ). Due to the complexity of measuring the Jacobian matrix, particularly for the stochastic term, we adopt an empirical approach to estimate the Lyapunov exponent. At each time step, the system is integrated around a randomly selected point within a distance of δ_t. The Lyapunov exponent

is then computed as the natural logarithm of the ratio between the distances of the integrated points $\delta_{t+\Delta_t}$ and the initial distance, as given by Equation 3.19, where Δt represents the time step. The Lyapunov exponent indicates the average rate of exponential growth or decay of infinitesimally small perturbations in various directions within the phase space. In this study, we consider the Lyapunov exponent to be omnidirectional, meaning that we compute exponents in all directions. A positive exponent indicates local instabilities, signifying that small perturbations grow rapidly and the system is generally sensitive to initial conditions.

$$\lambda = \ln\left(\frac{\delta_{t+\Delta_t}}{\delta_t}\right), \tag{3.19}$$

Figure 3.4 provides a visualization of the Lyapunov path in the complex plane for Reactive SCGL. The solutions with white, pink, and red additive noise terms are depicted in (a), (b), and (c) respectively, while (d), (e), and (f) represent solutions with white, pink, and red multiplicative noise terms. Red noise amplifies asymmetric fluctuations in the path and expands its ergodicity area as denoted by the local transformation value.

Figure 3.4 - Lyapunov exponent boundaries and averages as a function of noise color β.



The distribution of Lyapunov exponents with additive white noise and multiplicative white noise terms is similar to that of the noiseless system. As the noise intensity increases, smaller exponents become more common in both cases. However, the presence of multiplicative noise leads to a higher occurrence of positive Lyapunov exponents, indicating instability.

By defining initial conditions, employing numerical integration methods, and conducting stability analysis, we can successfully simulate the dynamic behavior of the Complex Ginzburg-Landau equation in both 2D and 3D settings. Building upon this foundation, the subsequent section takes a step further in GPA formalism.

3.2 Improved Gradient Pattern Analysis

Several improvements have been implemented in the GPA technique. This section will highlight the key modifications, including the introduction of a novel measure for GM1 and the technique's generalization to time series and hypercubes. To begin, we present a new metric for GM1, recently presented in a conference (SAUTTER et al., 2023).

3.2.1 New G_1

We proposed a new measure for GM1 that is proportional to the inequality of the set of Delaunay connections length L. We sort the L set and measure the coefficient from Equation 3.20.

$$G_1 = \frac{\overline{L}_{\text{sup}} - \overline{L}_{\text{inf}}}{\max(L)}, \tag{3.20}$$

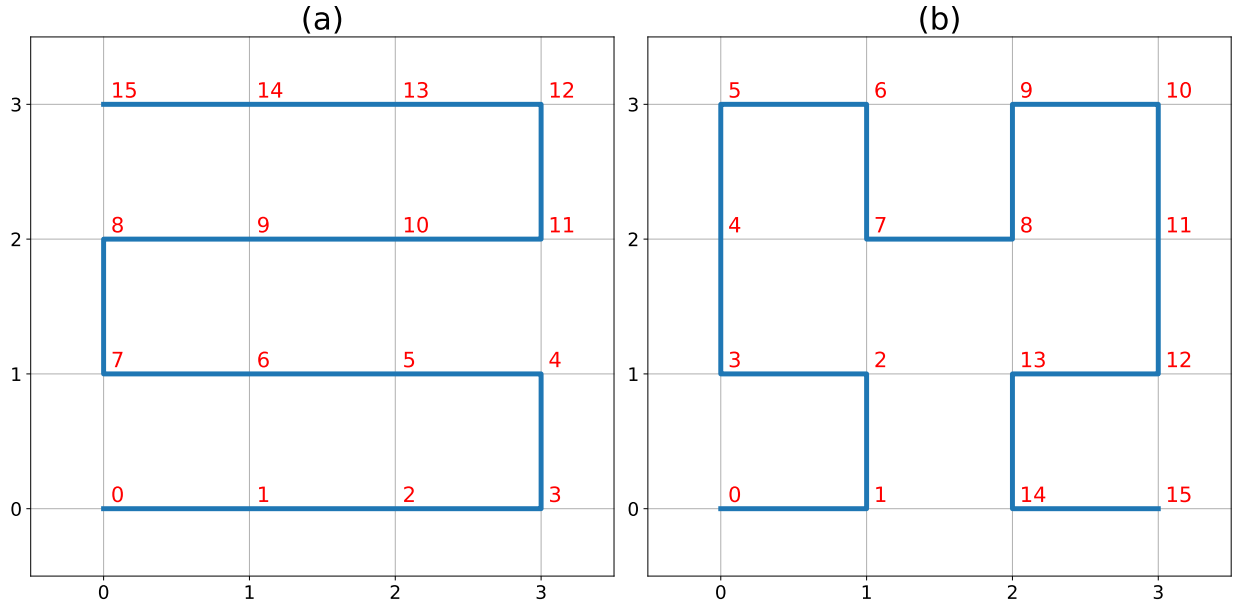
In Equation 3.20, $\overline{L}_{\text{inf}}$ is the set of connections smaller than the median, $\overline{L}_{\text{inf}}$ is the set of connections larger than the median, and $max(L)$ is the largest connection. In practice, the metric has an important application in the detection of dynamical patterns that partially breaks the symmetry.

3.2.2 Unidimensional generalization

Numerous applications often contain information in a 0D+1 format. To adapt the technique for such cases, an approach involves transforming the time series into a collection of matrices, as described by Assireu et al. (2002). There exist several mathematical models that can be utilized to perform this transformation interchangeably, as shown in Figure 3.5. These models fall under the category of space-filling curves, some examples of which are presented in Figure 3.8. One prominent algorithm in this domain is the Hilbert-space filling curve, which has found extensive applications, including data representation in databases (JENKINS et al., 1998) and encoding spatial-related data for neural network analysis (YIN et al., 2018; WANG et al., 2021). One of the main advantages of this technique is its ability to achieve spatial convergence (BUTZ, 1969); for different scaling measures, a given point in the series converges into a specific location as the matrix increases.

The differences between space-filling curves play a critical role as they have the potential to significantly impact the resulting patterns generated. Recognizing this, Yin et al. (2018) proposed a test in which a vector comprising 16 elements is transformed into a 4x4 matrix. Subsequently, the distances between neighboring elements in the transformed matrix are computed. In an ideal scenario, a smaller distance indicates a superior space-filling curve, as it implies that adjacent elements exhibit temporal similarities. Through this test, it was determined that the Snake curve and Hilbert curve achieve the minimal distance between space-filling curves. Therefore, we adopt both curves in our testings.

Figure 3.5 - Techniques to convert vectors into matrices.



The image shows different approaches to convert size 16 vectors into 4x4 matrices. The numbers in red indicate the position of a value in the vector, while the numbers in black indicate the equivalent position in the matrix. (a) is the Snake curve and (b) is the Hilbert curve.

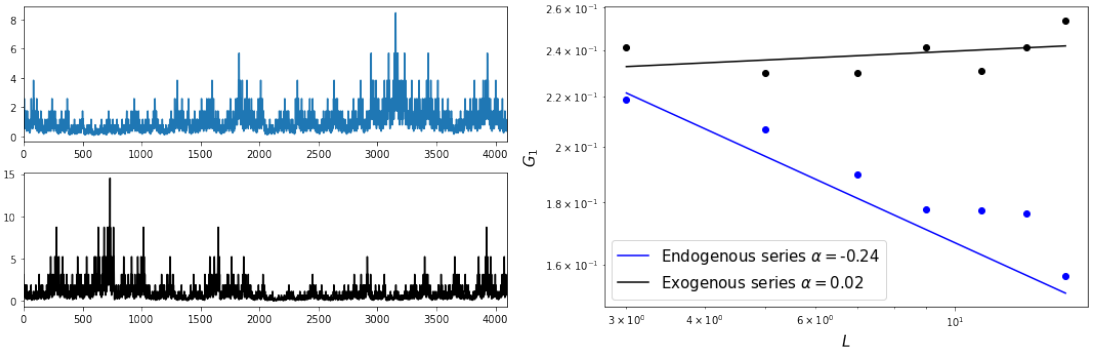
This approach produces a set of matrices, which can have various GMs metrics. Assireu et al. (2002) proposes the average of GMs metrics as a measure to characterize the time series. This can be an issue in a extreme event context since the extreme fluctuations can represent a small portion of the overall set. Another approach is to define a scaling law, for matrices of various sizes, we look for the power-law from Equation 3.21, where L is the matrix size, $G_N(x, L)$ is the Gradient Moment measure N, and $\alpha(x)$ is the scaling law.

$$G_N(x, L) \approx L^{\alpha(x)} \tag{3.21}$$

A similar approach has been proposed by Rosa et al. (2007), where a relation between the frequency and the GMs in log-log scale has been established in the context of wavelets. In this work, we look for the coefficients in a set of scales and look for the log-log rule. To find the scaling coefficient $\alpha(x)$, a linear fit between the logarithm of

$G_N(x, L)$ and the logarithm of the matrix size L is measured using the least squares regression. Two different examples of time series and scaling law fit are presented in Figure 3.6. Note that the more extreme the fluctuation, the greater the slope.

Figure 3.6 - GPA scaling law fitting example.



3.2.3 Tridimensional generalization

The analysis of tridimensional data introduces distinct challenges that cannot be effectively addressed using bidimensional tools alone. In this section, we present a novel extension of the GPA specifically designed for tridimensional data. The proposed extension of the GPA aims to enhance the interpretability of metrics associated with tridimensional data, even with a potential increase in computational cost.

The tridimensional lattice, also known as a hypercube, can be formally represented by a collection of amplitudes $A_{i,j,k}$ arranged in a regular grid, these amplitude elements are also known as voxels (volume elements). Here, i, j, and k serve as positional indices within the lattice. The majority of techniques employed in this context do not necessitate any conceptual alterations in this representation; only minor modifications in the code are required. For instance, computing a tridimensional gradient follows the same principles as the bidimensional case, with an additional consideration for the third dimension. The gradient can be calculated in a similar manner, incorporating the rules previously defined. Table 3.2.3 outlines the inclusion of the third case using the established rules.

Similarly, the algorithm for symmetry detection has been formulated based on the

Table 3.2 - Finite difference scheme for computing gradient lattice 3D.

Numerical scheme	$\partial_x(A_{i,j,k})$	$\partial_y(A_{i,j,k})$	$\partial_z(A_{i,j,k})$
CFD	$\frac{(A_{i+1,j,k}-A_{i-1,j,k})}{2}$	$\frac{(A_{i,j+1,k}-A_{i,j-1,k})}{2}$	$\frac{(A_{i,j,k+1}-A_{i,j,k-1})}{2}$
FFD	$A_{i+1,j,k} - A_{i,j,k}$	$A_{i,j+1,k} - A_{i,j,k}$	$A_{i,j,k+1} - A_{i,j,k}$
BFD	$A_{i,j,k} - A_{i-1,j,k}$	$A_{i,j,k} - A_{i,j-1,k}$	$A_{i,j,k} - A_{i,j,k-1}$

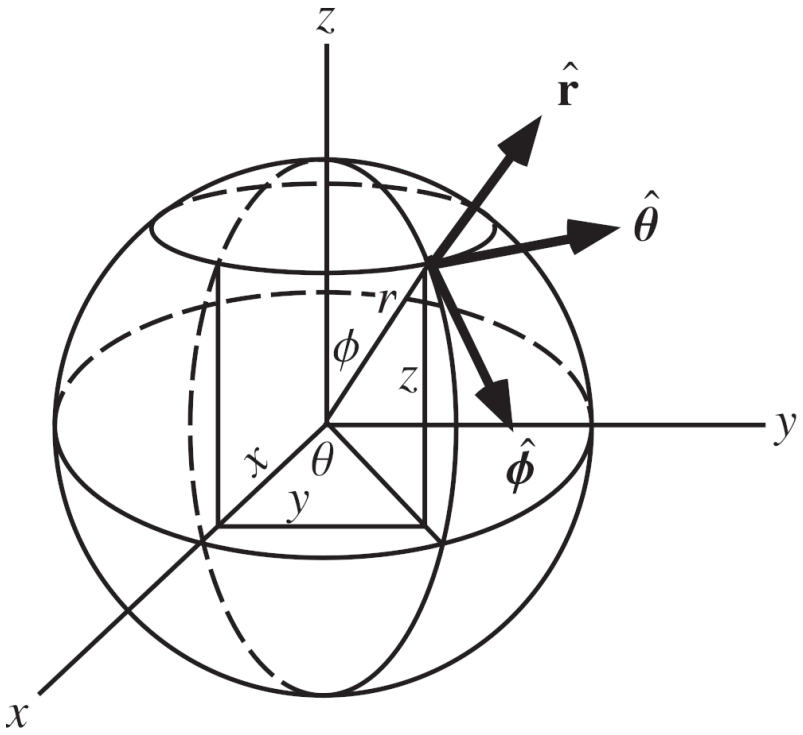
Central Finite Difference (CFD) scheme, Forward Finite Difference (FFD), and Backward Finite Difference (BFD). In a numerical lattice with amplitudes $A_{i,j}$the vectors are computed using CFD, and FFD or BFD for the boundaries. $\partial_x(A_{i,j,k})$ is the spatial derivative in x-direction, $\partial_y(A_{i,j,k})$ is the spatial derivative in y-direction, and $\partial_z(A_{i,j,k})$ is the spatial derivative in z-direction.

Euclidean distance from the center of symmetry. Consequently, the same principles can be extended to a tridimensional space without significant modifications. However, the measurements of GM1 and GM3 require some important adaptations. When extending the application to a 3D domain, an additional phase is needed to represent the vector. In this context, the unit vector used in the measure G_3 can be denoted in spherical coordinates according to the Equation 3.22, where ϕ is the phase relative to the z axis and the phase θ is relative to the x-y plane.

$$\begin{aligned} x &= \cos(\theta)\sin(\phi) \\ y &= \sin(\theta)\sin(\phi) \\ z &= \cos(\phi) \end{aligned} \tag{3.22}$$

A schematic of the spherical coordinate system is shown in Figure 3.7, where in a context of unitary vectors $r = 1$.

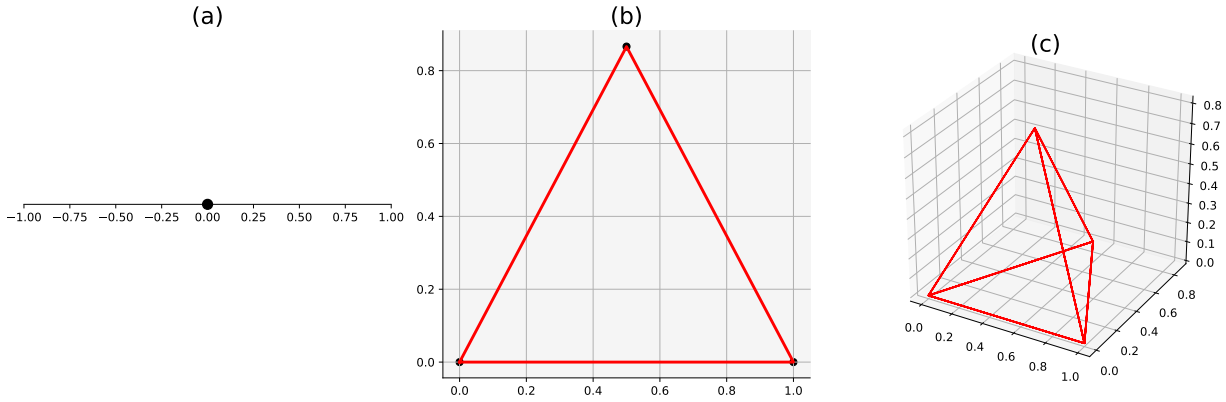
Figure 3.7 - Spherical Coordinates.



SOURCE: Weisstein (2005).

In GM1, a set of triangles are applied to incorporate the vectorial information of position, phases, and moduli. In a tridimensional context, this triangulation can be generalized to a set of simplices. This geometrical structure is a generalization of a triangle for a multidimensional scenario. Some examples of simplices are shown in Figure 3.8.

Figure 3.8 - Simplices and dimension.



Simplex is a structure that generalizes triangles (b) for multidimensions. For instance a zero-dimension simplex is a point (a), and a 3-dimension simplex is a tetrahedron (c).

In the specific case of 3D, the simplex is represented by a tetrahedron, which composition structure can be measured with the flipping algorithm from the QHull library (VIRTANEN et al., 2020). We also observed a change in the metric range to values higher than dimension of the system. In this sense we review in the next section some properties of GPA in a tridimensional context.

3.2.4 Simplex generalization and GM1 properties changes

In order to understand the measure of GPA, Rosa et al. (1999) proposes a test set that encompasses canonical matrices, symmetric matrices, quasi-symmetric matrices, laminar matrices, and random matrices. The hypothesis of G_1^C of random matrices as a natural measure of the system dimension has been proposed, due to the trend towards a maximum value of 2. However, the assumption is not valid with a simplex generalization. To exhibit this, consider the Algorithm 3.1.

Algorithm 3.1 - Function to generate the number of Delaunay connections.

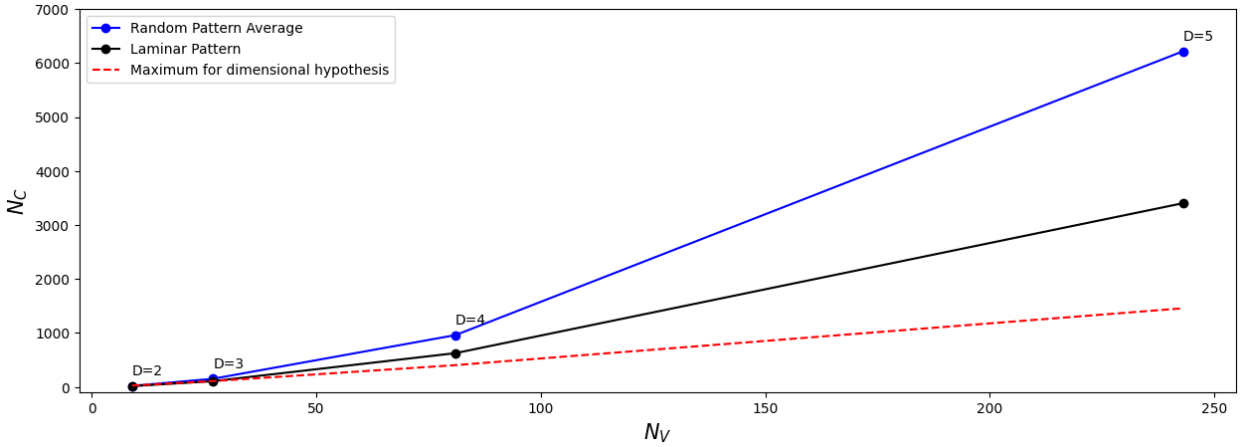
```
import numpy as np
from scipy.spatial import Delaunay
def getNDelaunayConn(d, n=3,nRandSample=100):
  x = np.linspace(0,1,n)
  grid = np.meshgrid(*tuple(np.repeat([x],d,axis=0)))
  pts = tuple(map(np.ravel, tuple(grid)))
  pts = np.column_stack(pts)
  _,conn2 = Delaunay(pts).vertex_neighbor_vertices
  randValues = []
  for i in range(nRandSample):
    randomSet = np.random.rand(*pts.shape)
    _,conn = Delaunay(randomSet).vertex_neighbor_vertices
    randValues.append(len(conn)//2)
  NCh = (dim+1)*pts.shape[0]
  NV = pts.shape[0]
  return np.max(randValues), len(conn2)//2, NCh, NV
```

The code provides the number of connections of the Delaunay triangulation of the laminar flux and random patterns in the lattice of N^D points, where N is the number of points in every dimension D. It also computes the maximal hypothetical number of connections, which can be calculated from G_1^C, shown in the introduction, according to Equation 3.23. This Equation can be deduced by considering $G_1^C = D$ and $N_V = N^D$.

$$N_C(N, D) = DN^D + D \tag{3.23}$$

From the code 3.1, Figure 3.9 has been computed. This Figure also shows an important property of the metric, which is the boundless asymptotic behavior.

Figure 3.9 - Dimensional Hypothesis Test.



The graph shows the relation between the number of vectors N_V and the number of connections N_C. This test was applied to a lattice with dimensions D, with $N = 3$ points in every dimension, resulting in N^D points. In order for the dimensional hypothesis to be valid for simplices, no measure can be found above the red line.

Instead of the previous hypothesis, the minimal structure of the system can determine its dimension. To illustrate this concept, we will examine the simplest case using the minimal structure denoted as $S_{\min}$. In a given dimension D, the minimal structure consists of $N_V = D + 1$ elements. The Delaunay connections of this minimal structure can be calculated using Equation 3.24, which leverages the relationship between the simplex cofaces and Pascal's triangle (CONWAY et al., 2016).

$$N_C = \frac{D(D+1)}{2} \tag{3.24}$$

Applying the classical GM1 metric for this minimal structure, yields Equation 3.25.

$$G_1^C(S_{min}) = \frac{D}{2} - 1 \tag{3.25}$$

In fact, one can rewrite Equation 3.25 as Equation 3.26.

$$D = 2(G_1^C(S_{min}) + 1) \tag{3.26}$$

For instance, in the bidimensional example, when a single triangle is computed,

$N_C = N_V = 3$, and therefore $G_1^C = 0$. Thus, the dimension $D = 2$ of the system can be calculated according to Equation 3.26. With that generalization defined, we conclude the extension of GPA for a multidimensional scenario. In the following chapter we discuss the data.

4 DATA

This chapter serves as an introduction to the data that will be analyzed in this work. To facilitate understanding, we have organized the Table 4.1 that includes the dimensionality of the data, as described in the theory of Amplitude Equations, along with the main reference for each dataset. In addition, this Table includes information about the nature of the data source, where 's' denotes synthetic data, and 'o' denotes observational data.

Table 4.1 - Datasets and dimensions.

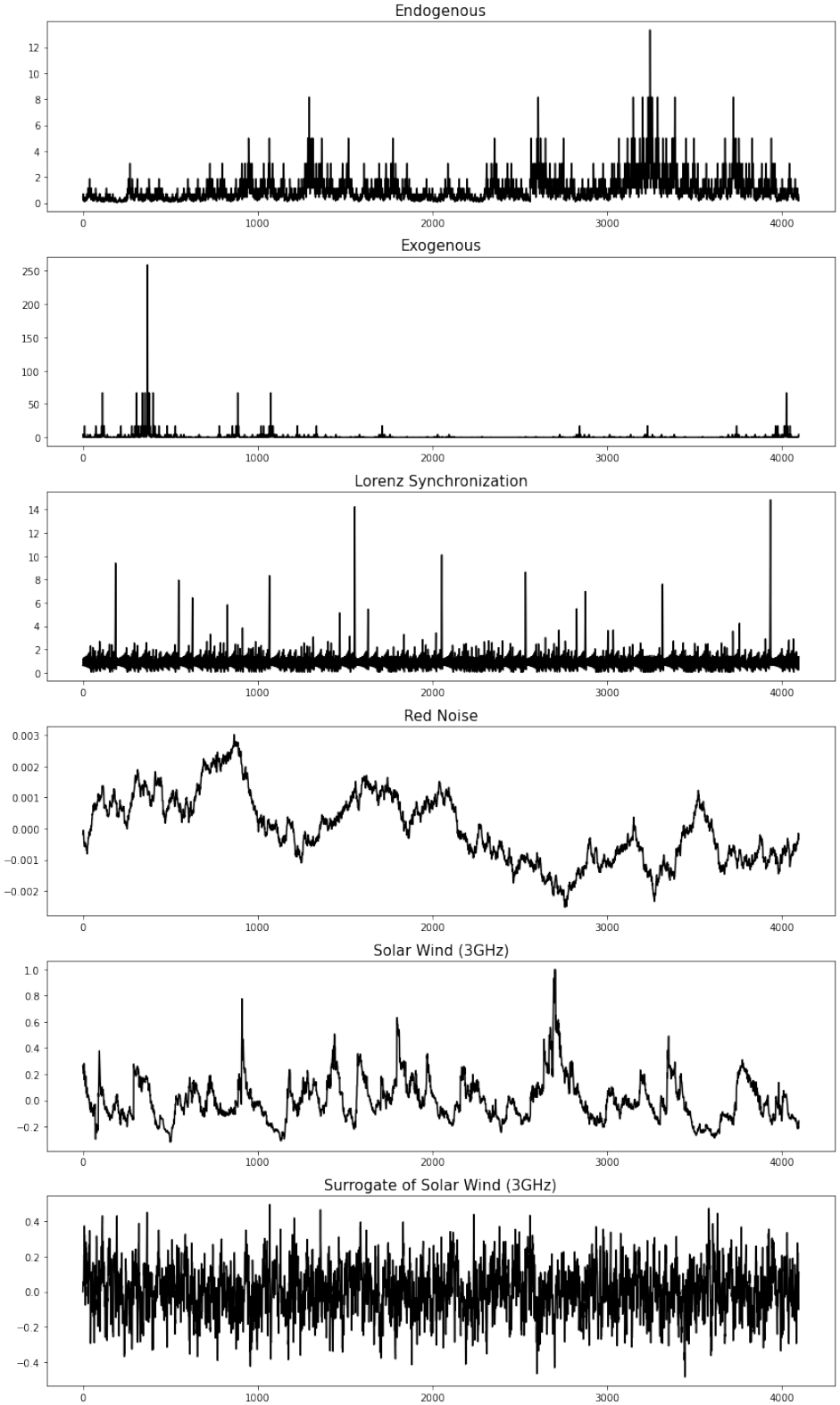
	Nature	**Dimensionality**	**Main Reference**
P-Model	S	0D+1	Meneveau and Sreenivasan (1987)
Lorenz	S	0D+1	Krawiecki and Matyjaśkiewicz (2001)
Solar burst 3 GHz	O	0D+1	Rosa et al. (2008)
NOAA 11131	O	2D+1	Sych et al. (2012)
Milli-Millennium	S	3D+1	Guo and White (2008)
Colored Noise	S	nD+1	Timmer and Koenig (1995)
Ginzburg-Landau	S	nD+1	Aranson and Kramer (2002)

This chapter is structured in alignment with the dimensionality of the data, allowing us to start with the presentation of the time series dataset.

4.1 Extreme Events

Three canonical systems have been selected as illustrative examples to explore different patterns of extreme fluctuations: the multifractal extreme dissipation of the p-model, the coupling between two Lorenz chaotic models, and the noise-like dynamics. Additionally, real-world datasets are compared with observations of Solar wind at 3GHz, which have been previously studied by Rosa et al. (2008). To understand the impact of the time series structure, we apply the surrogate operation (Appendix A) on the observational data and compare the result. Figure 4.1 shows some examples of the model, the Solar wind data, and the corresponding surrogate. In the following section, we start with one of the most popular theoretical model to generate Extreme time series, followed by the chaotic desynchronization model.

Figure 4.1 - Extreme time series.



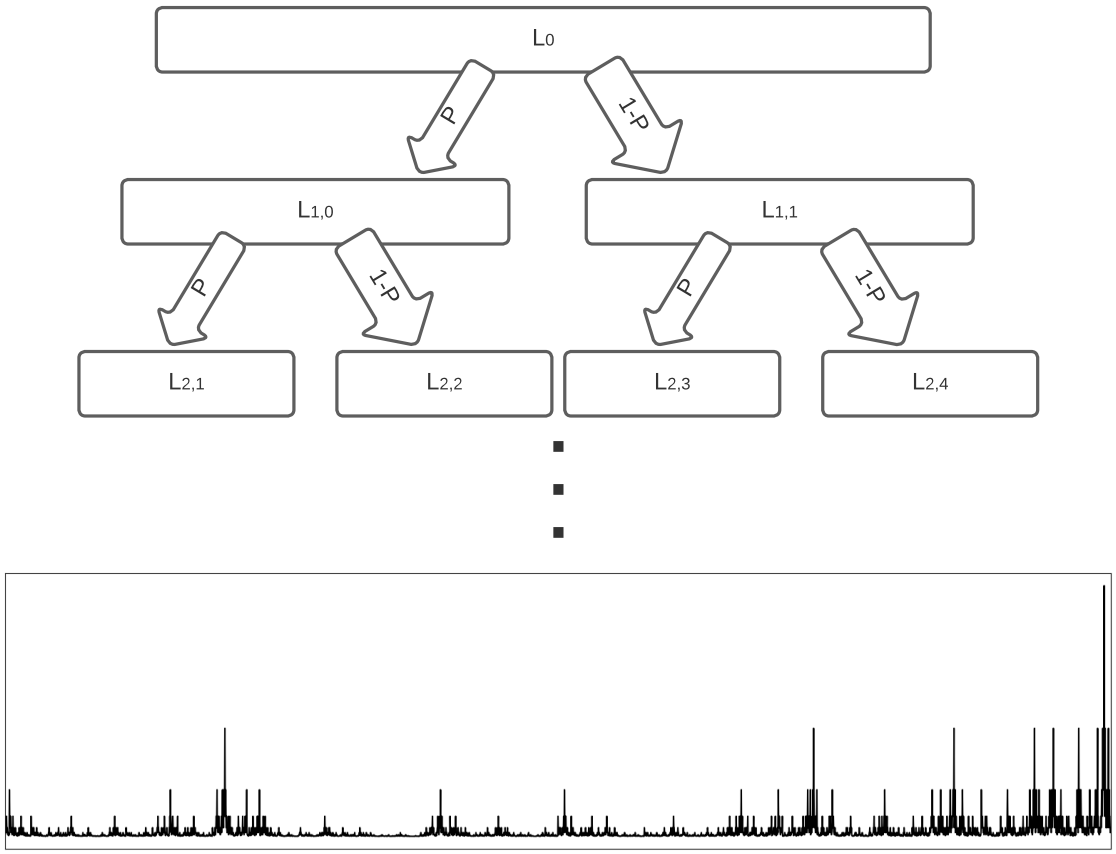
Sample of canonical models, the observation of a solar burst in radio, and the correponding surrogate of the solar wind time series.

The study of extreme events (*X-events*) has attained increasing interest, particularly in fields like space and environmental physics, where such occurrences are prevalent. Three canonical systems have been selected as illustrative examples to explore different patterns of extreme fluctuations: the multifractal extreme dissipation of the p-model, the coupling between two Lorenz chaotic models, and nonlinear noise-like dynamics.

4.1.1 P-model

The P-model is model for nonhomogenous turbulent-like cascade proposed by Meneveau and Sreenivasan (1991). It is a generalization form of two-scale cantor set with a balanced distribution of length which shows multifractal properties of one-dimensional sections of the dissipation field. The generalized form starts from the classical view of eddy cascade before the inertial range of fully developed turbulence, where a flux of energy (E_K) of size L_K cascades down along the inertial range and eventually dissipates, when it reaches the Kolmogorov length scale β, In the process of cascading, the energy is recursively divided into two part, as shown in Figure

Figure 4.2 - P-Model nonhomogenous energy cascading process.



P-Model cascading of inertial scale, generated using binomial distribution. In this example, an exogenous time-series has been generated with $P = 0.75$.

Hence, in a cascade step, each eddy of size L_K is divided into two equal parts expressed as L_{K_1} and L_{K_2} each. However, in each cascade step, the flux of energy is distributed, as a probability, unequally in fractions of p_1 and $p_2 = 1 - p_1$ where $p_1 + p_2 = 1$. This process is iterated over fixed p_1 until eddies reach the Kolmogorov scale β.

While the P-Model is widely recognized in the field of extreme events, it is important to note that it represents just one of several potential mechanisms for generating such events. In the following section, we explore an alternative approach to producing extreme events from the partial synchronization process.

4.1.2 Chaotic extreme patterns

The study of coupled chaotic systems has been a significant focus in dynamical systems, particularly in extending dynamical systems to spatial domains and investigating synchronization phenomena. In specific scenarios, synchronization processes occur temporarily, followed by sudden desynchronization and subsequent resynchronization. This intriguing phenomenon is commonly referred to as a blowout bifurcation. Noteworthy practical applications that exhibit such phenomena include laser synchronization (ASHWIN et al., 1998) and neuronal networks (KANAMARU, 2006). The following desynchronization time series were provided by Professor Juan Alejandro Valdivia in collaboration with the University of Chile.

In this work the synchronization of a dual Lorenz system is investigated, a similar system presented by Krawiecki and Matyjaśkiewicz (2001). The Lorenz system is a set of three nonlinear differential equations, shown in Equation 4.2, that describes a simplified model of atmospheric convection. The variables x, y, and z correspond to spatial coordinates, while the parameters σ, β, and ρ play important roles in the system's behavior.

$$\begin{aligned} f_x(x,y,z,\sigma,\beta,\rho) &= \sigma(y-x) \\ f_y(x,y,z,\sigma,\beta,\rho) &= x(\rho-z)-y \\ f_z(x,y,z,\sigma,\beta,\rho) &= xy-\beta z \end{aligned} \tag{4.1}$$

In order to induce extreme events using this system, we introduce a second Lorenz system that is driven by the dynamics of the former system. Specifically, the latter system is influenced by the derivatives $f_x^{(0)}$, $f_y^{(0)}$ and $f_z^{(0)}$ of the former system. Equation 4.3 describes the latter system, where the parameter μ regulates the strength of the driving force.

$$\begin{aligned} f_x(x,y,z,\sigma,\beta,\rho,\mu,f_1) &= \sigma(y-x)-\mu(x-f_x^{(0)}) \\ f_y(x,y,z,\sigma,\beta,\rho,\mu,f_1) &= x(\rho-z)-y-\mu(y-f_y^{(0)}) \\ f_z(x,y,z,\sigma,\beta,\rho,\mu,f_1) &= xy-\beta z-\mu(z-f_z^{(0)}) \end{aligned} \tag{4.2}$$

Finally, the time series of desynchronization can be measured by the distance between both system, using the Euclidean distance shown in Equation 4.3.

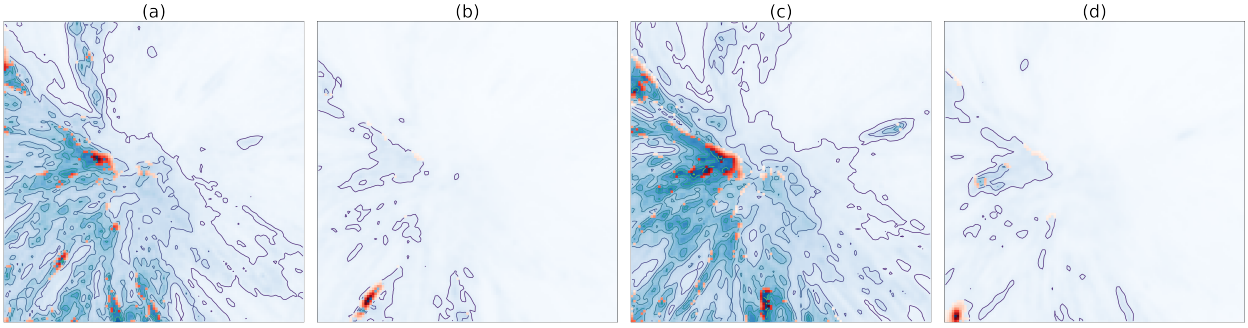
$$e(t) = \sqrt{(x_1(t) - x_2(t))^2 + (y_1(t) - y_2(t))^2 + (z_1(t) - z_2(t))^2} \tag{4.3}$$

From this error function we conclude the introduction of 1D data. The next section presents an example of 2D data.

4.2 Sunspot NOAA 11131

Solar corona seismology plays a crucial role in unraveling the formation and evolution of sunspots. Recent studies have focused on investigating the oscillations of magneto-acoustic waves with periods of 5 minutes and 3 minutes. Various candidate models have been proposed to explain these dynamics (CHO et al., 2019; FELIPE, 2019), including hypotheses related to small-scale plasma convection or the cut-off frequency of a p-mode. In this study, we specifically examine the NOAA 11131 region, which serves as a standard observation for 3-minute sunspot oscillatory dynamics (DERES; ANFINOGENTOV, 2015; SYCH et al., 2021; SYCH et al., 2012). The observations of this region were obtained on December 10, 2010, using the Solar Digital Observatory (SDO). The umbra at 304 Å was tracked over a duration of 6 hours with a time resolution of 12 seconds, employing the methodology proposed by Sych et al. (2012).

Figure 4.3 - Snapshots of NOAA 11131.



Contour line and heatmap of NOAA 11131, observed at 304 Å. Blue regions contain symmetrical vectors, whereas red regions contain asymmetrical vectors. (a) and (c) are snapshots of the oscillatory regime, while (b) and (d) are snapshots of the critical regime.

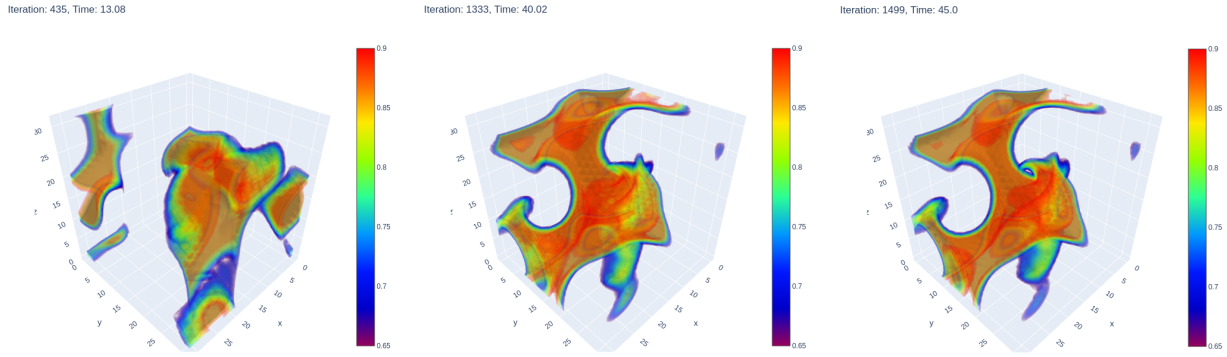
SOURCE: Sautter et al. (2023).

To complement the existing studies on NOAA 11131, we introduce the GPA technique. By analyzing the temporal evolution of GPA metrics and conducting subsequent time-series analysis, we aim to characterize the transient stages and oscillatory characteristics induced by the underlying solar physics. Figure 4.3 shows some regimes of the event, which are the oscillatory regime and the critical regime. In an oscillatory regime, a diversity of structures is observed, while in a critical regime, a single region presents higher amplitude values. Our aim is to distinguish these moments and analyze the oscillatory behavior.

4.3 Complex Ginzburg-Landau 2D+1 and 3D+1

As described in methodology chapter, the parameters, initial conditions and boundary conditions follows the stability presented previously. Also, the technique discussed in Chapter 2.1 can be extended to multiple dimensions. The pseudospectral method establishes a relationship between frequencies and the Fourier spectrum. In a Multidimensional Fast Fourier Transform, the frequencies in each dimension are linearly independent. Consequently, by applying the same shift to each direction of the spectrum, we can effectively approximate the diffusion term of the 3D-CGL. Some snapshots are depicted in Figure 4.4, where the selected snapshots has been selected from the critical points of the GPA.

Figure 4.4 - Example of CGL 3D Snapshots.



4.4 Milli-Millennium

The Millennium Simulation, conducted by the Virgo Consortium (JENKINS et al., 1998), is a pioneering project in cosmology. Its primary aim was to simulate the evolution of dark matter distribution on a massive scale and observe the formation of large-scale structures in the universe. The simulation involved more than 10 billion particles within a cosmological box measuring approximately 500 h^{-1}Mpc. To carry out the simulation, the state-of-the-art GADGET2 code (SPRINGEL, 2005) was employed.

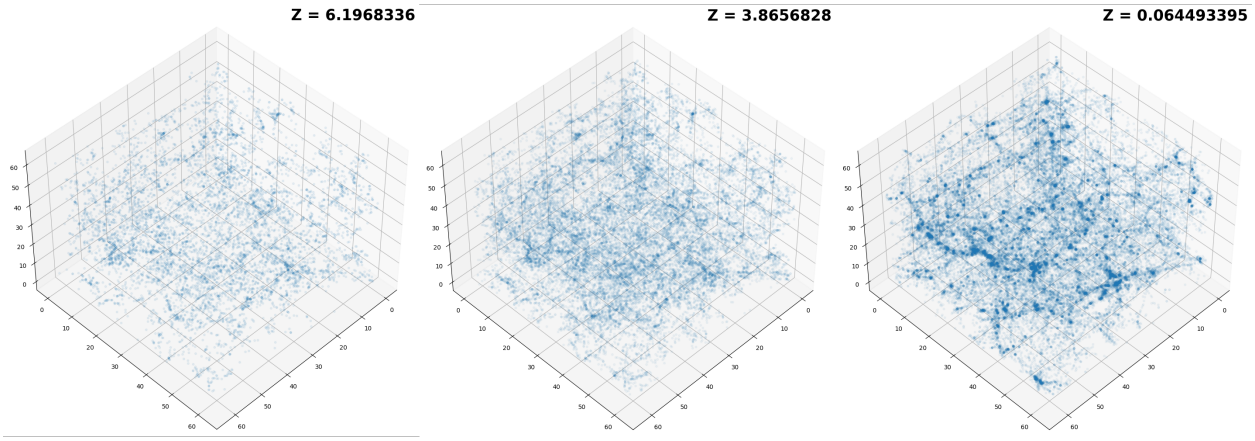
In 2008, a second simulation was completed using the same cosmology, output structure, and particle count as the original Millennium Simulation. This time, the simulation was performed in a box five times smaller, resulting in an 125 times increase in mass resolution. This enabled a more detailed examination of the underlying dynamics and processes shaping the formation and evolution of cosmic structures.

The Milli-Millennium dataset is a post-processed compilation derived from the Subhalo tree of the original Millennium simulation. The Subhalo tree is a hierarchical structure that summarizes the mass distribution within a specific region, effectively reducing the number of data points required to describe the structures by a factor of 1/512 in the Milli-Millennium. The pioneering work of Lucia et al. (2006) introduced the concept of post-processing the first simulation, while Guo and White (2008) conducted an in-depth analysis of the simulation's structure in the context of

the second simulation. Figure 4.5 illustrates the outcome of the analysis performed by Guo and White (2008).

To facilitate access to the Milli-Millennium dataset, the code it has been made publicly available[1]. The dataset itself can be obtained through SQL queries. In addition, to simplify the process of data retrieval and loading, we have developed a code[2] that enables automated data downloading and seamless integration with the Pandas and Mechanize libraries.

Figure 4.5 - Example of Milli-Millenium snapshot.



Sample of Mili-Millenium snapshot. The variable Z indicates the redshift of the snapshot.

To investigate the formation of structures using GPA, the dataset was discretized into a hypercube, as illustrated in Figure 4.5. This involved computing the logarithm of subhalo density within each voxel, which corresponded to a spatial scale of 1^3 Mpc. As a result, the hypercube consisted of 64x64x64 voxels, which can be analyzed with 3DGPA. From this point, we conclude the presentation of all the data that will be analyzed in the next chapter.

[1] http://gavo.mpa-garching.mpg.de/Millennium/
[2] https://github.com/rsautter/MillenniumQuery

5 RESULTS AND INTERPRETATION

This chapter encompasses the presentation of all the analyses conducted and their corresponding results. To enhance clarity, we have organized the order of presentation according to the dimensionality of the data. We present the results of the 2D and 3D GPA in two types of outputs. The first type of output presents an individual analysis of each measure, while the second type of output presents an analysis of the combination of measures. The objective of this analysis is to detect, whenever possible, regimes, periodicities and trends. These characteristics are important to interpret the observed processes and validate the metrics.

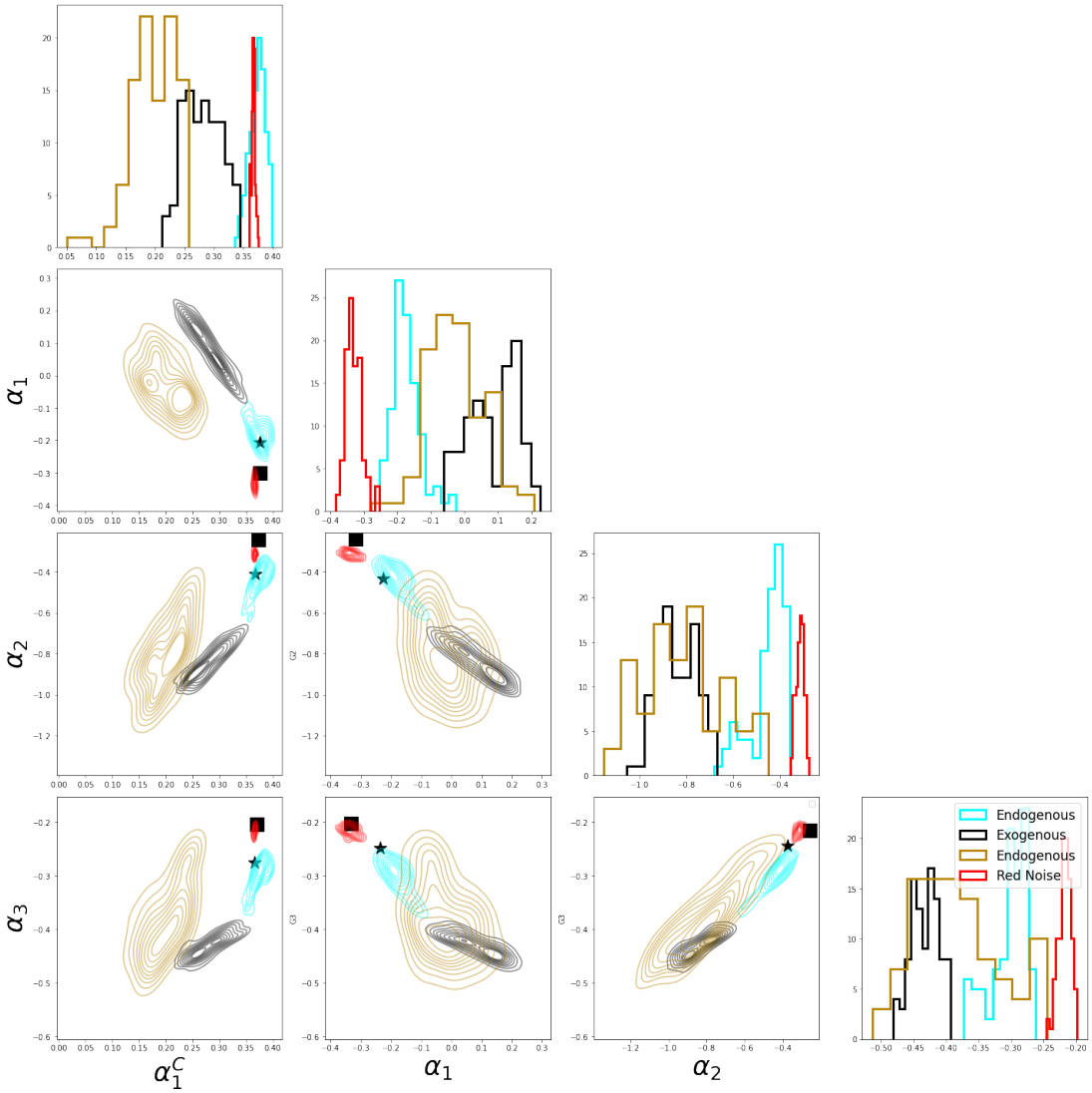
In the first type of output, the colors at the background of the time series represent the regimes. Each segment is detrended to calculate its periodogram. The trend calculation is carried out by fitting a linear function, whose coefficient is μ. After detrending, the segment is normalized by the respective standard deviation. Finally the periodogram is calculated using the Bartlett kernel (BARTLETT, 1950).

The second type of output presents an analysis composed by the combination of all GMs, where the colors represent the time of the series. In this graph we look for clusters that represent the regimes of the series. Furthermore, the distribution of each GM on the main diagonal is shown in this graph. It is worth noting that due to the unique nature of analyzing 1D data, we have adopted a different method specifically for these datasets. Also, we do not present the analysis with GM4 in this study, as it is a measure in complex numbers that makes visualization and interpretation difficult.

5.1 GPA characterization of extreme events in ordinary time series

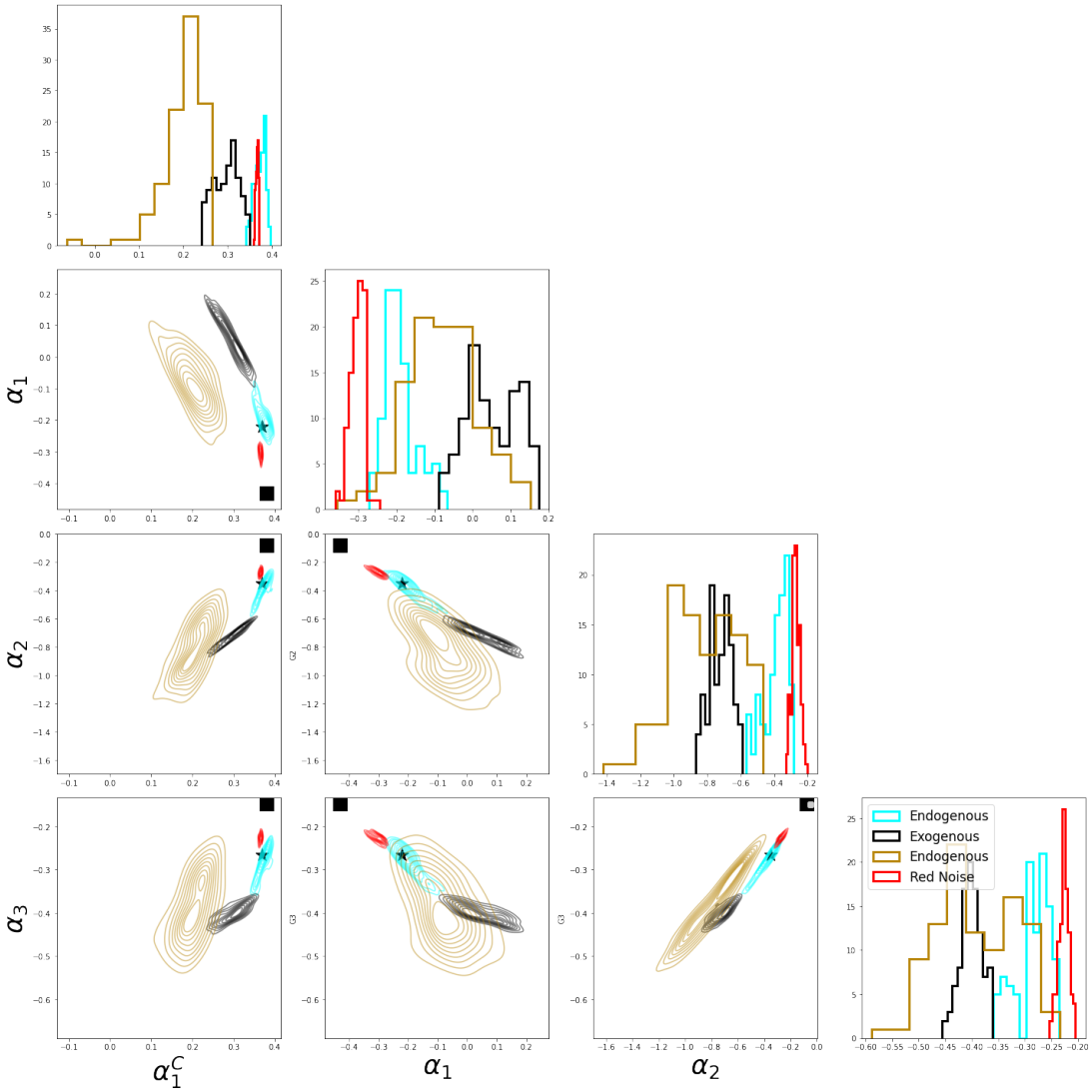
Applying this technique to the synthetic extreme fluctuation dataset, we obtain the Coefficients distributions from Figures 5.1 and 5.2.

Figure 5.1 - GPA scaling coefficient of extreme events using Snake curve.



Scaling coefficient of GPA 1D generalization using for synthetic models, solar wind observational data, and corresponding surrogate. Notice that surrogate series of solar wind is near the original series and the red noise region.

Figure 5.2 - GPA scaling coefficient of extreme events using Hilbert's curve.



Scaling coefficient of GPA 1D generalization using for synthetic models, solar wind observational data, and corresponding surrogate. Note that the surrogate series is always in an empty region. This result is expected, since the surrogate operation breaks the temporal relation of the series.

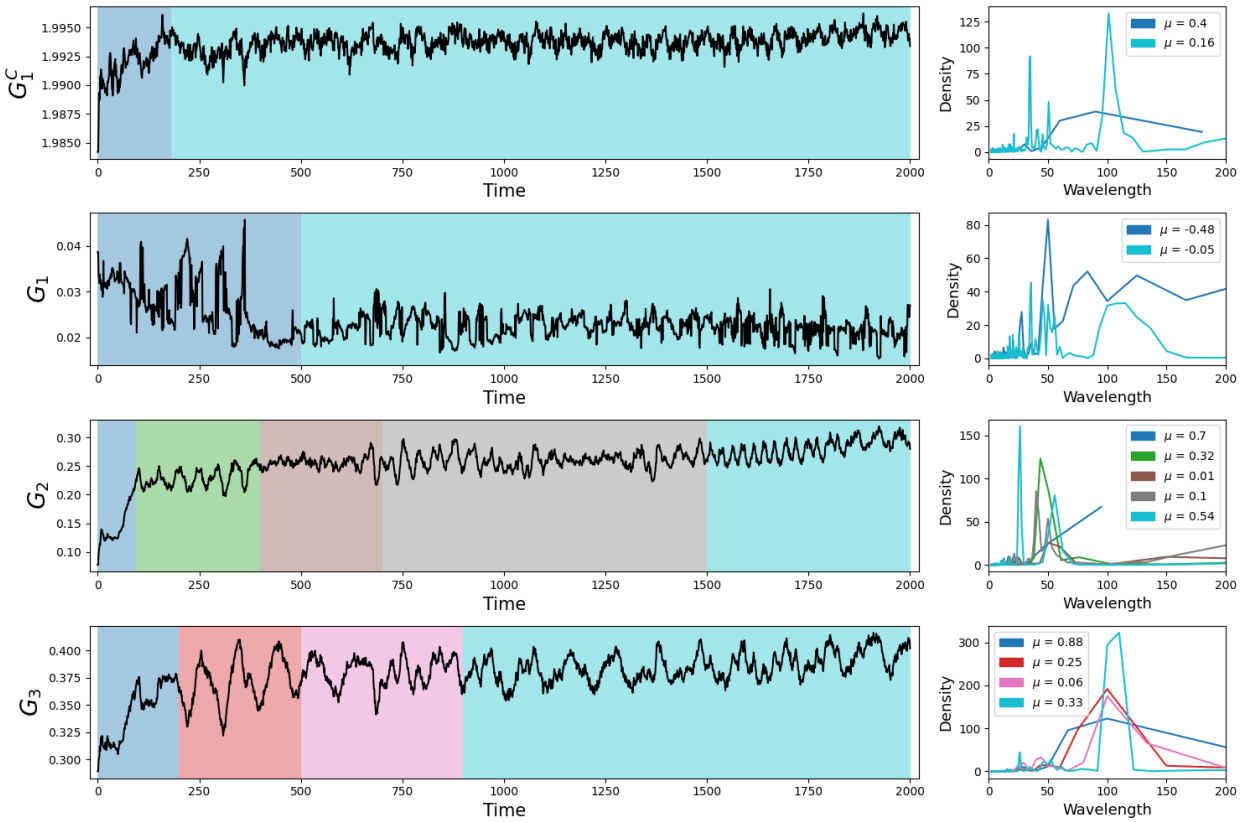
Previous works utilizing GPA in the 1D context have primarily focused on the Reshape curve (ASSIREU et al., 2002) along with a statistical measure. However, in this study, we have extended the investigation by replacing the reshape curve with the Snake curve and Hilbert curve, due to the analyzes presented in the literature. By examining the results depicted in Figures 5.1 and 5.2, we can conclude that the Hilbert's curve is particularly effective at characterizing extreme events. It exhibits

a clear ability to distinguish between the four classes of synthetic extreme events, and the ability to distinguish the original series from the surrogate series. It is also expected a blank region for the surrogated series, since this operation breaks the time structural relation of the series. With this analysis we can conclude that the solar wind time series studied is similar to an endogenous event in terms of symmetry and scaling law. In the subsequent section, we further explore a different generalization, this time focusing on tridimensional datasets.

5.2 GPA of complex Ginzburg-Landau 2D+1

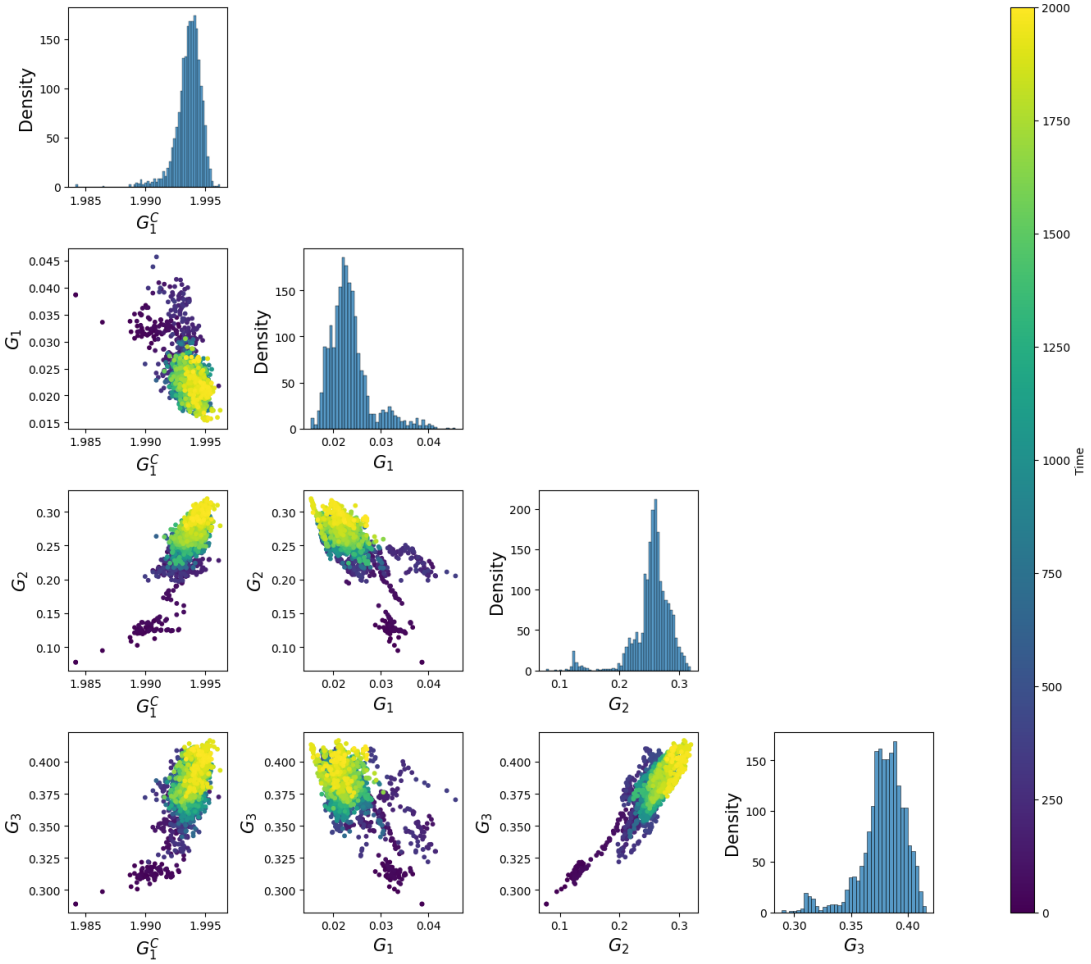
We conducted a GPA analysis on the real part of the system at each snapshot to investigate the formation of spirals and their dynamic defects.

Figure 5.3 - Analysis of 2D Complex Ginzburg-Landau pattern formation without noise.



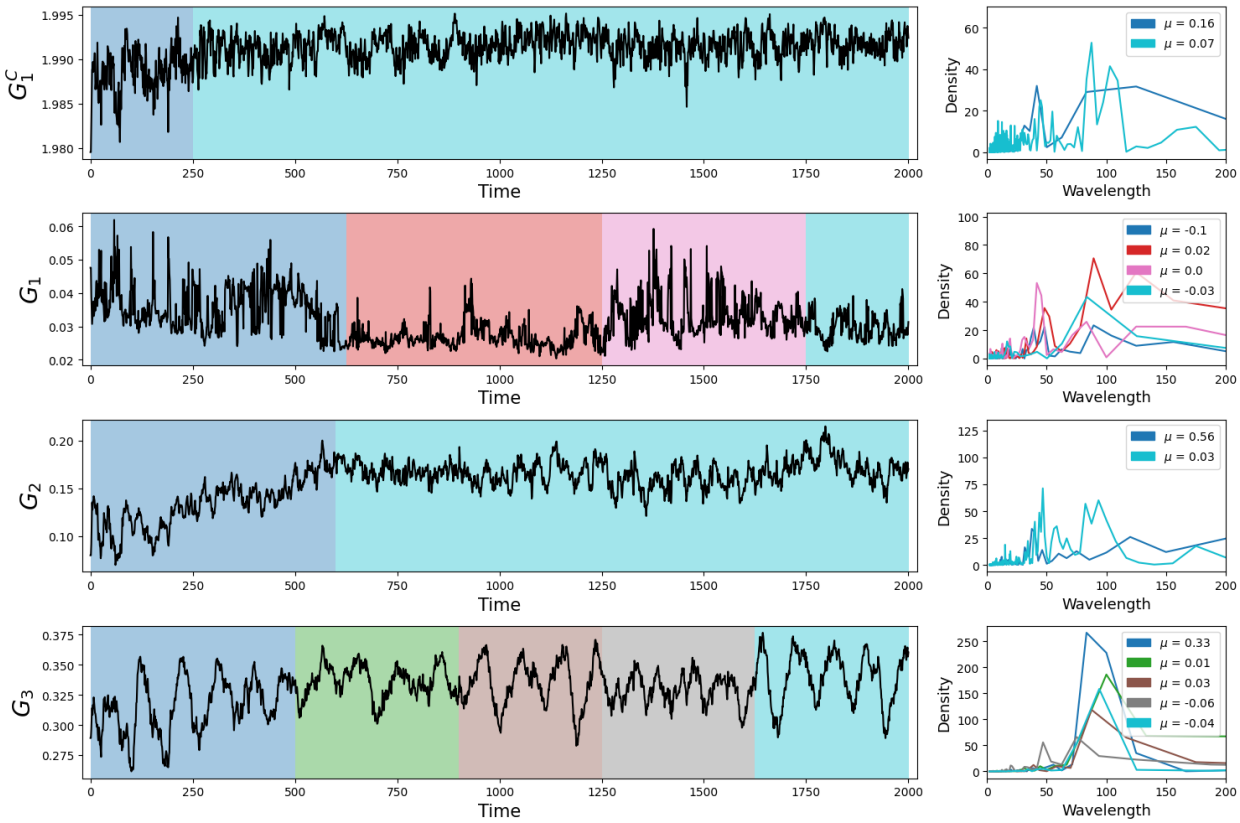
The number of observed regimes is different for each gradient moment. GM1 presents the initial and secondary stages that indicates the accretion and spiral formation. GM2 and GM3 have different oscillatory characteristics that indicate the presence of more regimes.

Figure 5.4 - Combined analysis of 2D Complex Ginzburg-Landau pattern formation without noise.



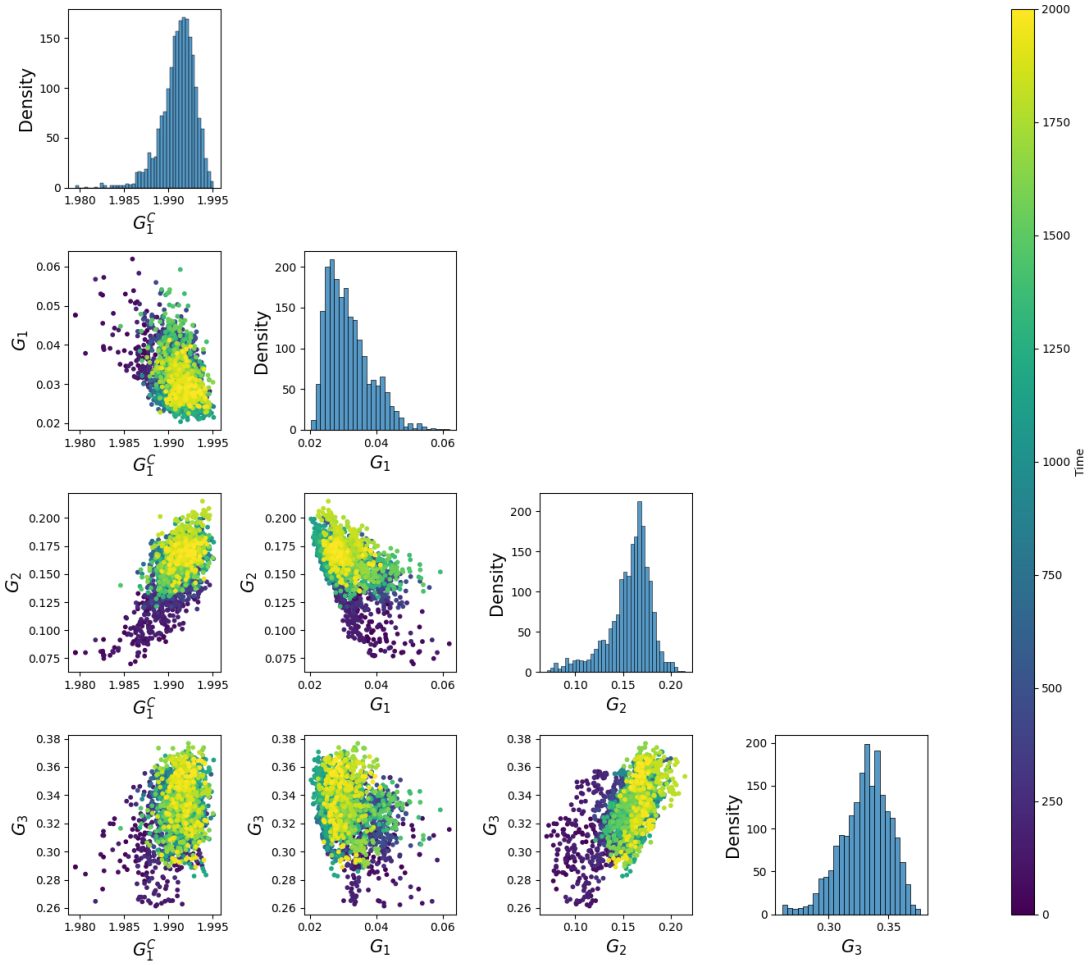
Combination of all metrics, colors indicate the time.

Figure 5.5 - Analysis of 2D Complex Ginzburg-Landau pattern formation with additive noise.



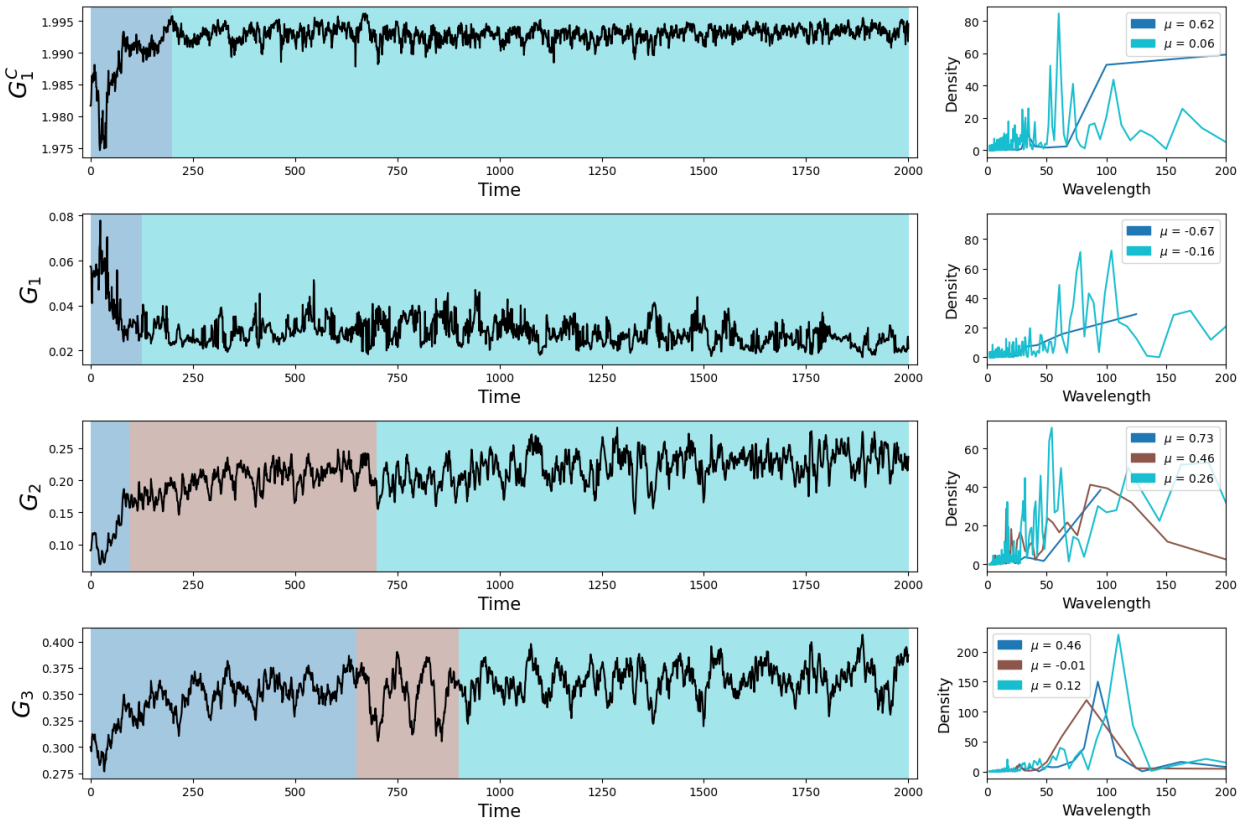
The number of observed regimes is different for each gradient moment. GM1 presents the initial and secondary stages that indicates the accretion and spiral formation. GM2 and GM3 have different oscillatory characteristics that indicate the presence of more regimes.

Figure 5.6 - Combined analysis of 2D Complex Ginzburg-Landau pattern formation with additive noise.



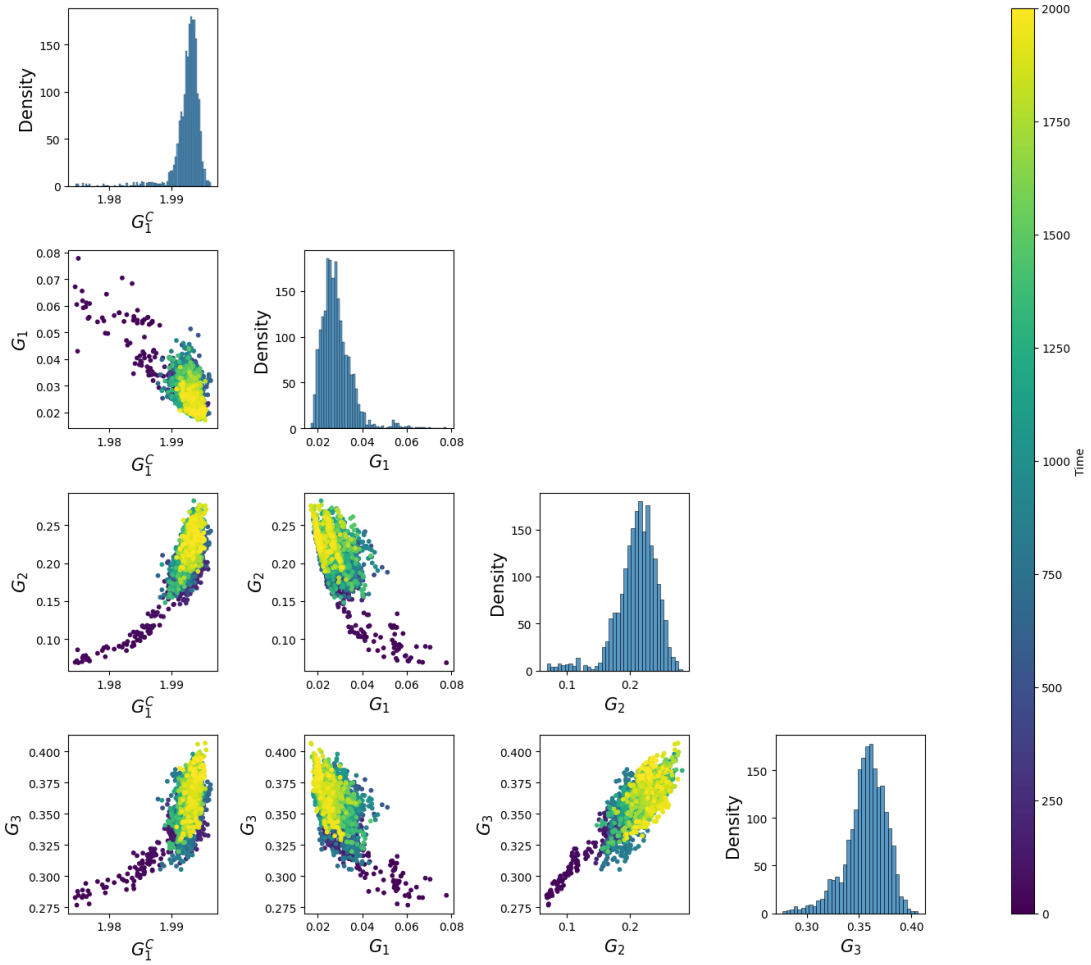
Combination of all metrics, colors indicate the time.

Figure 5.7 - Analysis of 2D Complex Ginzburg-Landau pattern formation with multiplicative noise.



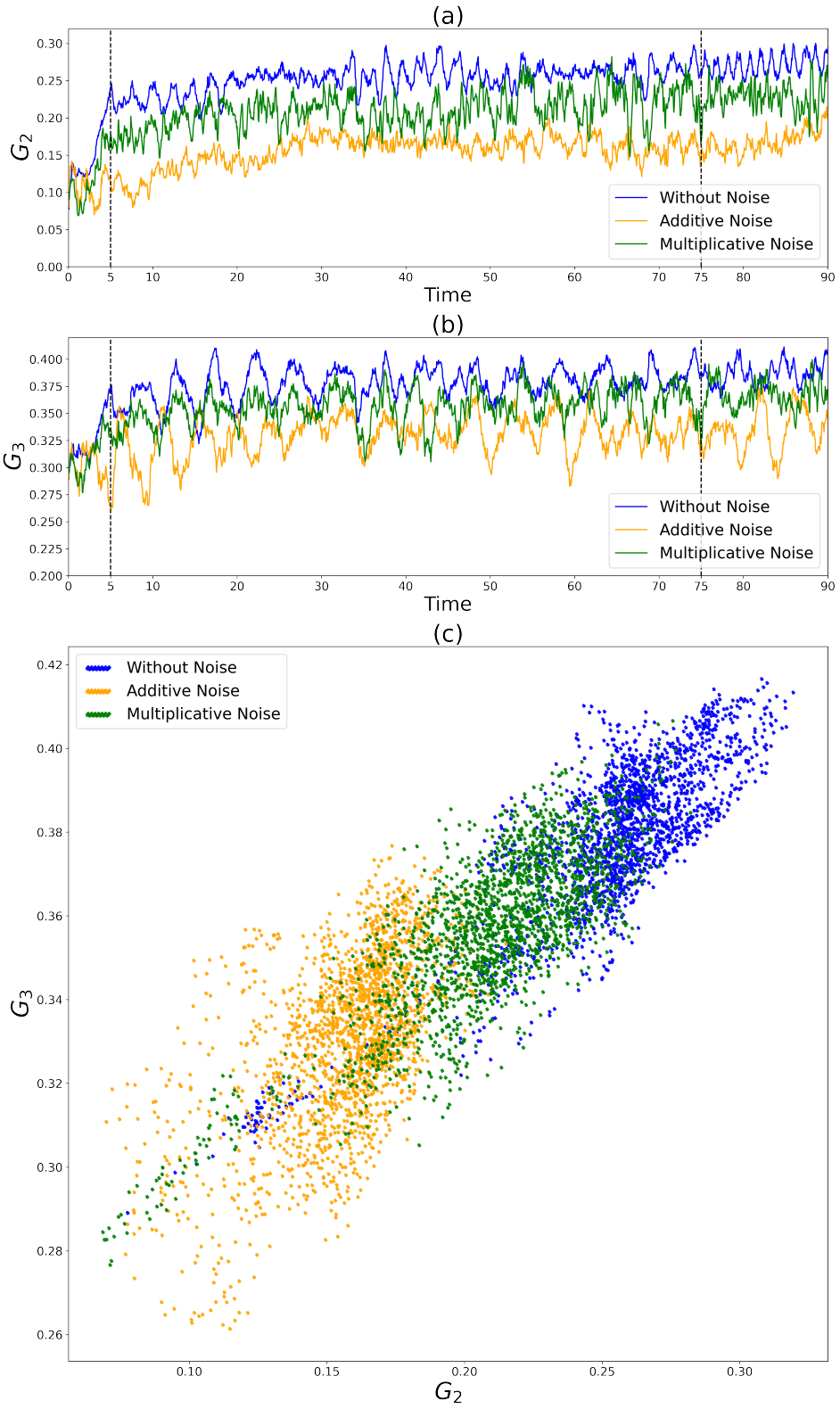
The number of observed regimes is different for each gradient moment. GM1 presents the initial and secondary stages that indicates the accretion and spiral formation. GM2 and GM3 have different oscillatory characteristics that indicate the presence of more regimes.

Figure 5.8 - Combined analysis of 2D Complex Ginzburg-Landau pattern formation with multiplicative noise.



Combination of all metrics, colors indicate the time.

Figure 5.9 - Phase space of Complex Ginzburg-Landau 2D and its Parametric space.



Every color represents a type of stochastic/non-stochastic CGL. Notice that, every type of stochastic and the non-stochastic systems occupy different regions.

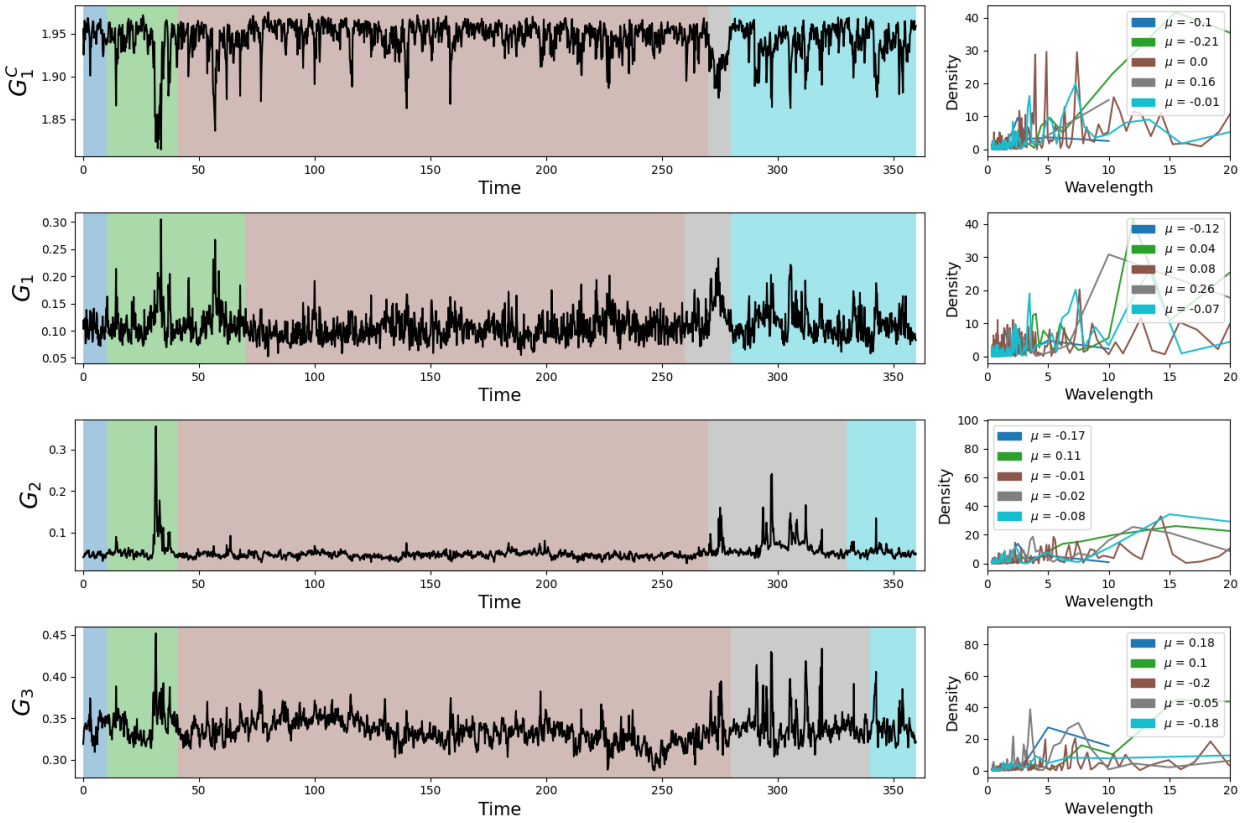
The introduction of stochastic terms altered the pattern formation pathway, leading to a non-relaxing system, as shown in Figures 5.3, 5.5, and 5.7. The additive noise term often inhibited the formation of spiral defects, preventing metric relaxation. In the presence of multiplicative noise, large structures were preserved, but local fluctuations disturbed the system, resulting in non-periodic variations within a narrow dynamical range. However, the technique still provides evidence of the accretion regime and spiral formation, even for the case of additive noise, as shown in the measure G_2. Furthermore, the intermittent periodicity of G_3 indicates periods where noise has less influence. It is important to note that, depending on the noise term, both G_2 and G_3 occupy unique regions, as illustrated in Figure 5.9. These findings underscore the intricate interplay between stochastic elements and pattern formation dynamics.

Figures 5.4, 5.6, and 5.8 present the combined analysis that unveiled essential aspects related to pattern formation. In the absence of noise, three clusters are visible in $G1$x G_2, which indicates distinct stages. In the additive noise scenario, the parametric space G_2xG_3 exhibited a sparse set of points due to the higher diversity of asymmetric patterns. Conversely, the system with multiplicative noise displayed a similar region in G_2xG_3 space to the noise-free system, aligning with the formation of large structures. The distinction between the system with multiplicative noise and the noise-free system lay in the amplitude of G_3 and G_2, indicating local fluctuations generated by the multiplicative noise term.

5.3 GPA of solar active region (NOAA 11131)

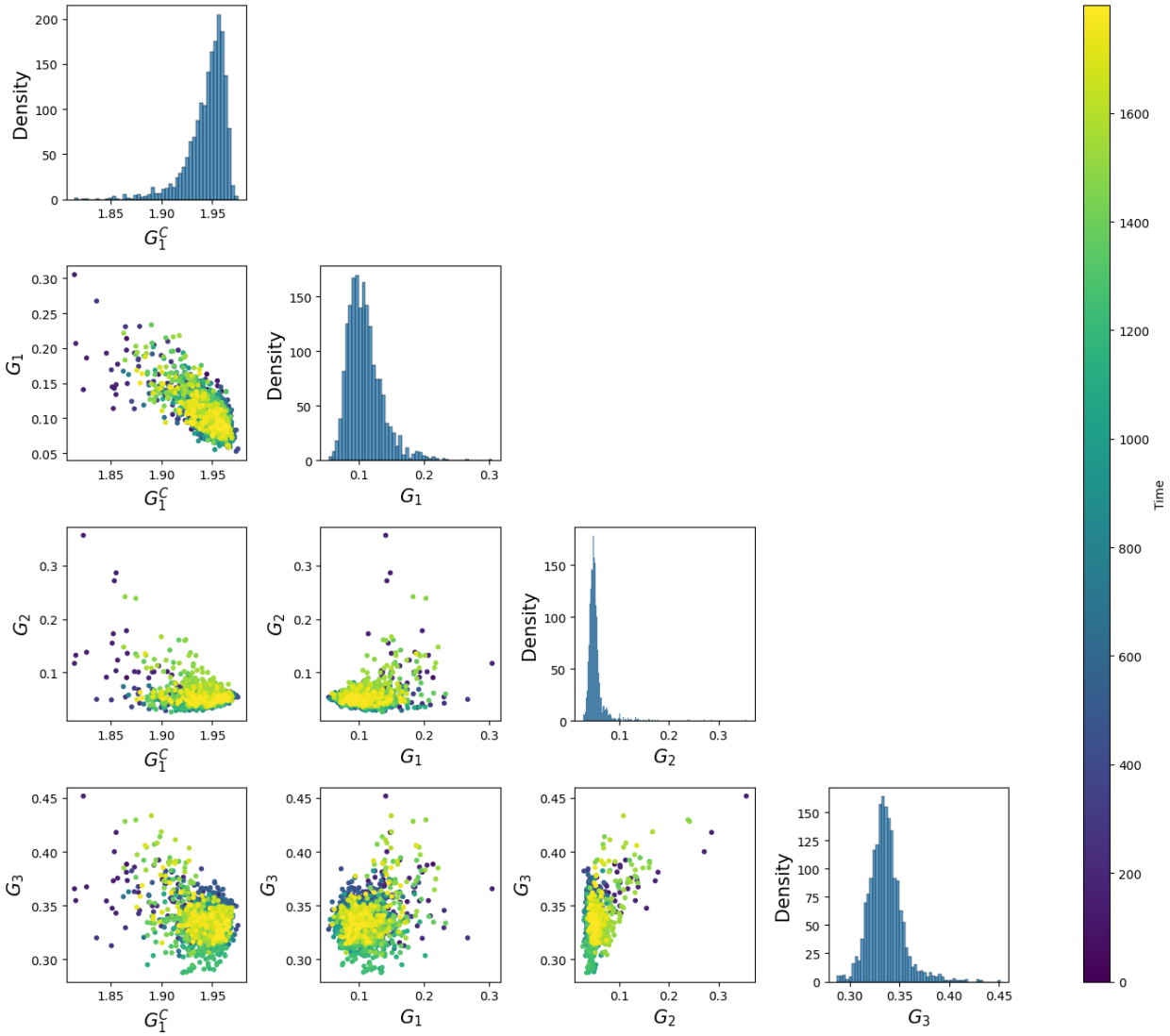
The analysis of sunspots reveals two distinct critical regimes characterized by maximum pattern asymmetry, as illustrated in Figure 5.10. Those regimes exhibit distinct durantion and amplitudes of Gradient Moments (GMs). The first stage is relatively shorter but exhibits higher GMs compared to the second stages at every GM. Among the measures applied to the problem, G_1 is the only one to present a result consistent with the literature in terms of oscillation. Three peaks are observed in the regime between 41 minutes and 270 minutes, which correspond to oscillations of 3, 4.5 and 7 minutes.

Figure 5.10 - Gradient Moments of NOAA 11131.



The same result can be seen in the parametric spaces, as shown in Figure 5.11. In this Figure, there is a region of greater concentration of points and some outliers. The region with the highest concentration of points represents the oscillatory regime, while the outliers represent the critical regime. Note that the outliers are two-colored, as extreme fluctuations are observed at different times.

Figure 5.11 - Combined analysis of NOAA 11131.



Time is displayed according to the colors. Note that most points are concentrated in a given region of parametric space, regardless of the moment of the gradient. The points in this region belong to the oscillatory regime, while the outliers belong to the critical regime.

5.4 GPA of complex Ginzburg-Landau 3D+1

Upon observing the dynamics of the system, we can identify the presence of 3D spirals. Consequently, the metrics exhibit periodic behavior starting from time 600, as depicted in Figure 5.12. The extension of the technique to the three-dimensional

domain brought advantages to the analysis of the system. Periodic dynamics are more pronounced in the 3D case than in the 2D case. This periodicity is also reflected in the parametric space (Figure 5.13), especially in the G_2xG_3 space.

Figure 5.12 - Analysis of 3D Complex Ginzburg-Landau pattern formation without noise.

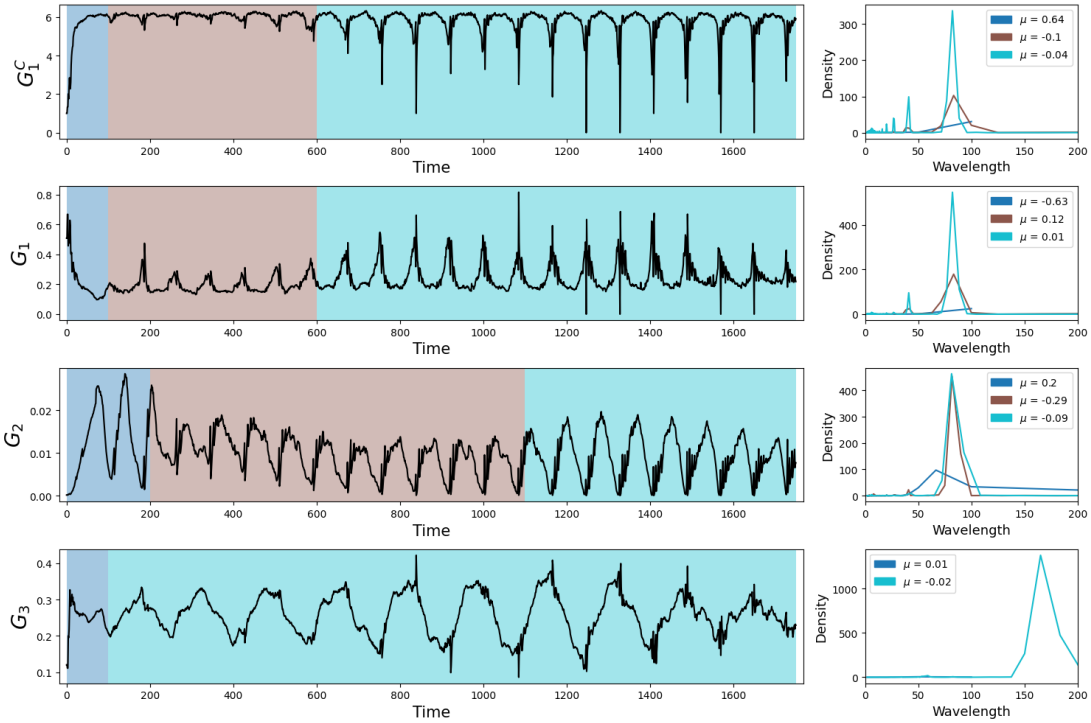
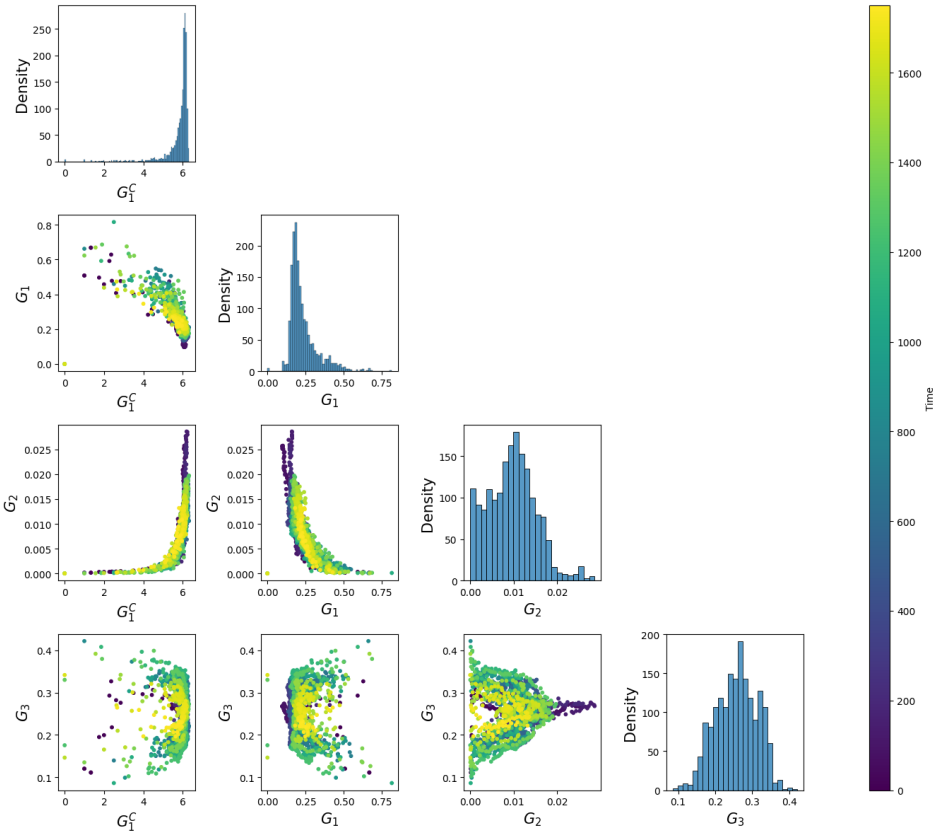


Figure 5.13 - Combined analysis of 3D Complex Ginzburg-Landau pattern formation without noise.



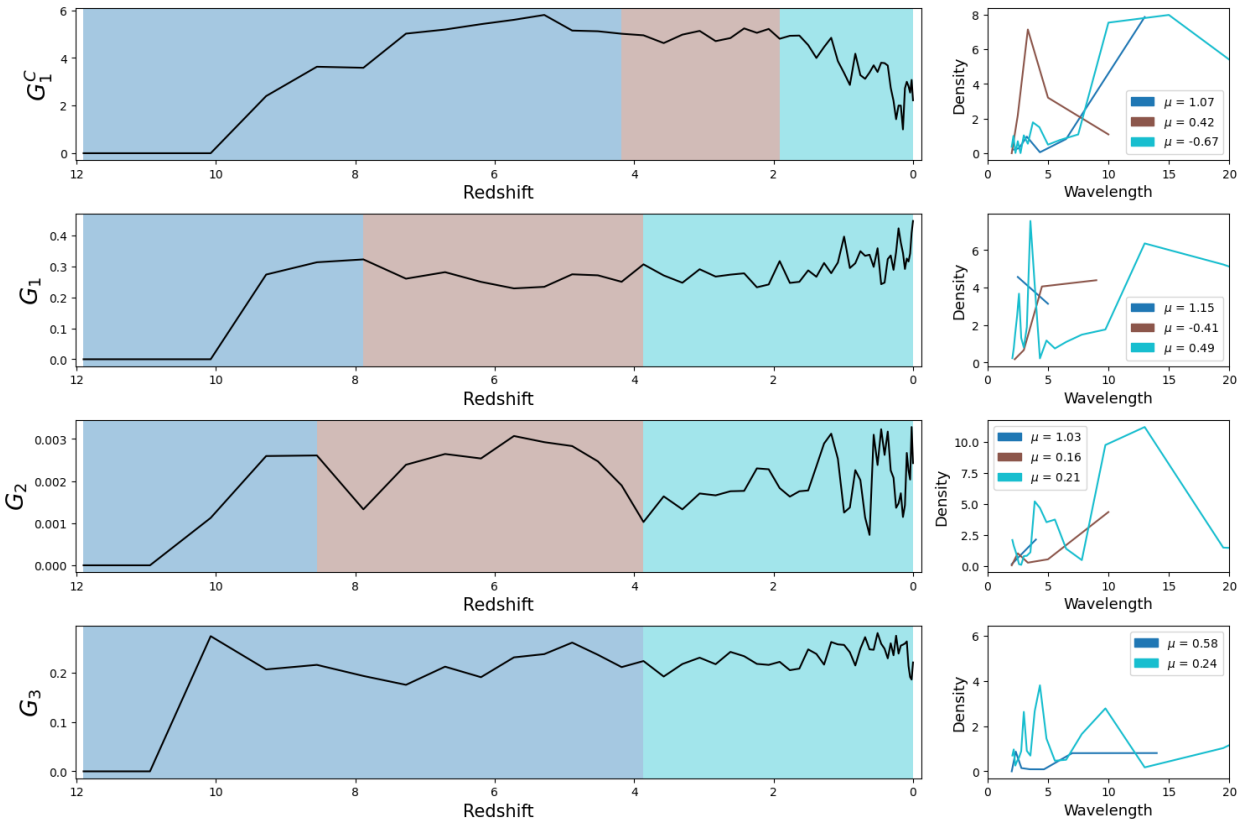
The analysis presented in this section demonstrates the efficacy of 3D GPA in describing a controlled experiment. In the subsequent section, we delve into a more challenging dataset that pertains to the cosmological formation of structures. This dataset has previously been examined using earlier versions of GPA (ANDRADE et al., 2006), making it an intriguing and engaging test for the proposed technique.

5.5 GPA of Milli-Millennium

The 3D GPA analysis, depicted in Figure 5.14, encompassed the study of 63 hypercubes representing the redshift range $12 < z < 0$. Notably, the initial snapshots contained no density values exceeding the minimum subhalo size determined by Guo and White (2008). Consequently, the associated metrics were all zero within the redshift interval $12 < z < 11$. The combined analysis is more complicated in the case of few elements, as it is difficult to determine the region of the clusters. Even so,

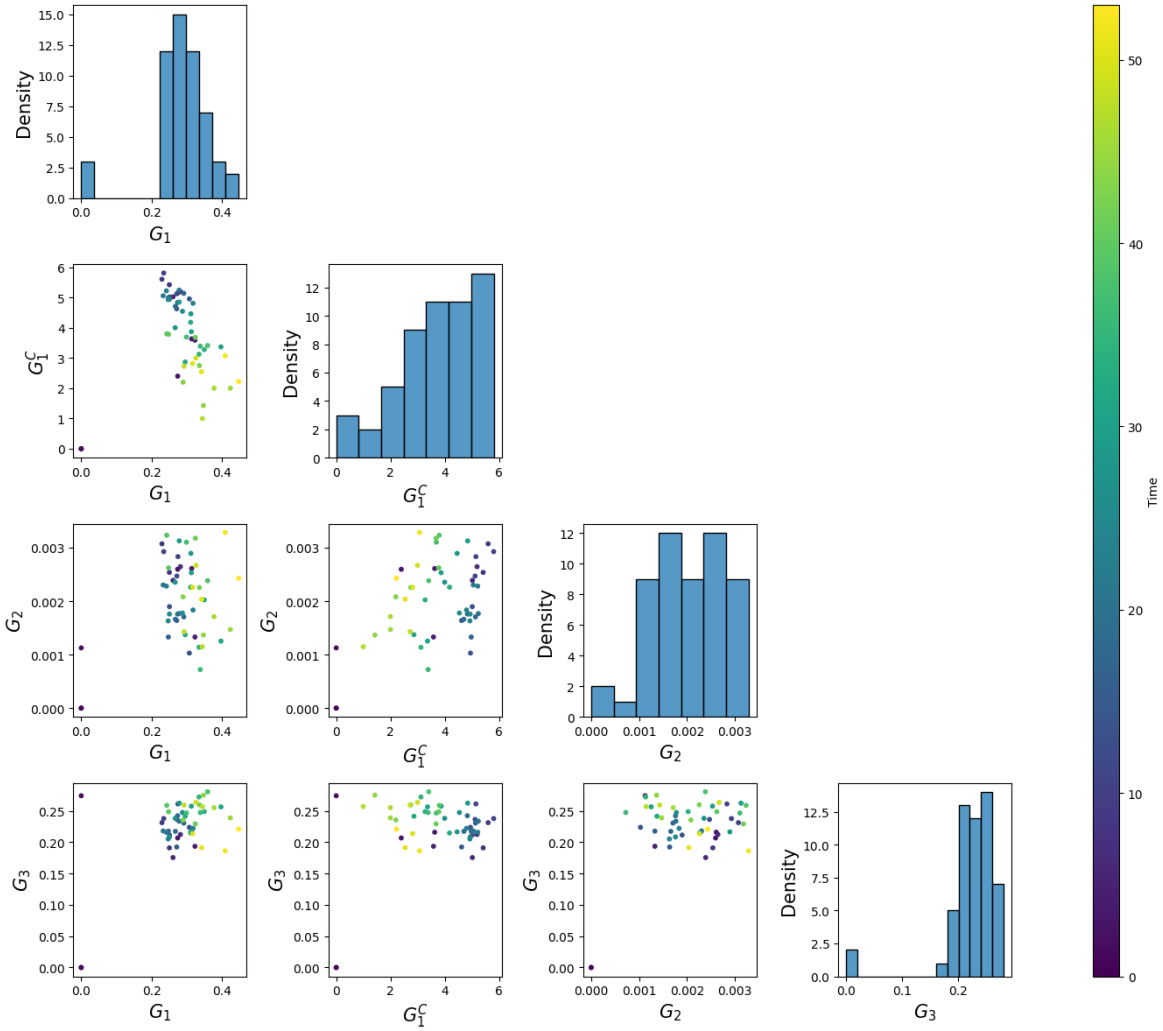
it is possible to verify some outliers in this analysis in Figure 5.15, which represent the initial stages of the simulation. Periodogram analysis is also impaired in this scenario. However, trend analysis is more robust in this scenario, as shown in Figure 5.14.

Figure 5.14 - Analysis of the Milli-Millennium.



The redshift scale is reversed, as the simulation started at $z \approx 12$, until the present $z = 0$.

Figure 5.15 - Combined analysis o of the Milli-Millennium.



The results obtained from the 3D GPA analysis for G_1^C align closely with the previous findings reported by Andrade et al. (2006), Caretta et al. (2008). The G_1^C curve provides insights into the accretion history of matter in the universe. Initially, the matter is uniformly distributed throughout space. However, local high-density regions asymmetrically distributed start to form gradually, as indicated by the increasing values of G_1^C, until approximately $z \approx 5$. Subsequently, the local clumps of matter merge and aggregate into larger structures, resulting in a decrease in the G_1^C values.

The metrics G_1, G_2, and G_3 consistently demonstrate that the matter and structures exhibit isotropy in terms of density variation from approximately $z \approx 9$ to $z \approx 1$. As a consequence, these metrics maintain relatively constant values within this redshift range. Conversely, the newly introduced metric G_1 exhibits an increase in recent times ($z < 2$). As this metric is associated with the diversity of Delaunay connections, this upward trend suggests the formation of structures at different scales during more recent cosmic epochs.

6 A VISUAL ANALYTICS TOOL FOR GPA APPLICATIONS

Numerous important works have focused on addressing the visualization of datasets (FRIENDLY, 2008; SCHWENDIMANN et al., 2016; JANES et al., 2013). By employing the tools presented in this chapter, analysts can readily understand the results derived from the GPA technique, enabling them to make more informed decisions. Analysis visualization plays a crucial role in interpreting and understanding complex data patterns, especially as datasets continue to grow in size and complexity. The primary objective of analysis visualization is to present data in a visual form that is easily comprehensible, enabling users to perceive patterns, trends, and relationships that may not be apparent through raw data alone. By leveraging tools like libraries such as Matplotlib and Plotly, analysts have access to flexible and interactive ways of visualizing data. These visualization libraries offer a wide range of options to represent data, allowing for dynamic exploration and manipulation of visualizations.

There is a wide range of definitions for dashboards found in the literature, and their interpretation may vary depending on the specific context. One commonly cited definition, provided by Janes et al. (2013), states that within an organization, a "dashboard" refers to a system that visually represents data in a manner that aids decision-making. Similar to the dashboard in a car, the primary objective of organizational dashboards is to inform users without distracting them from their main tasks. To achieve this, dashboards employ various visualization techniques such as charts, tables, gauges, and other concise summarizations of data.

Furthermore, in the context of education, Schwendimann et al. (2016) offer another significant definition that pertains specifically to learning dashboards. According to Schwendimann et al. (2016), a learning dashboard is a unified display that brings together diverse indicators related to learners, the learning process, and the learning context, presenting them through one or multiple visualizations. The purpose of a learning dashboard is to provide a comprehensive overview of relevant information to facilitate informed decision-making in educational settings.

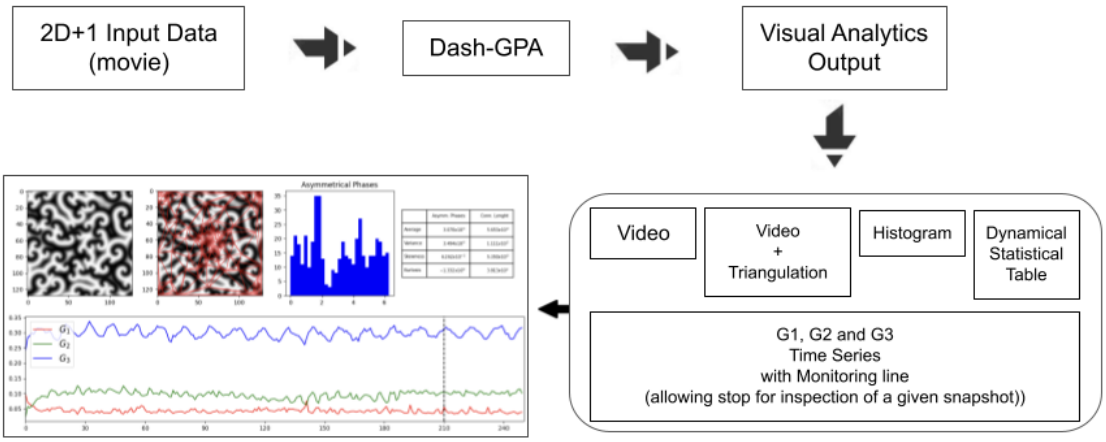
In general, when designing dashboards, simplicity and flexibility are sought-after attributes. The visualizations used should be straightforward and easy to understand, allowing users to grasp the information at a glance. Additionally, flexibility is crucial to accommodate different types of data and the evolving needs of users. Schwendimann et al. (2016) also highlights some popular graphics commonly employed in dashboards, including bar plots, line plots, and tables. These visual representations

effectively present data in a clear and concise manner, enabling users to interpret the information efficiently. In this chapter we approach GPA in a context of visual analytic that aims to meet at least three contemporary demands in the context of data science:

- Introduce a GPA-Dashboard for monitoring and detecting spatiotemporal nonlinear regimes in real time.
- Evaluating which are the best metrics to work together in a classification model.
- Using such metrics as labels for training supervised data mining models, machine and deep learning.

Therefore, considering GPA as a tool for monitoring the 2D+1 pattern formation case, a dashboard was built to work as a visual analytic interface in general applications. The system, developed in matplotlib, is exemplified in Figure 6.1. The Dash output includes the movie of the pattern formation (left above) and the same (above center) with the triangulation pattern among the vectors of the gradient lattice. The respective evolution of the distribution of geometric connection values along the time evolution is at the top right. The update of the respective statistical moments is also included as a complementary measure of the geometric complexity in each snapshot. The bottom part shows the values of the gradient moments (G_1, G_2 and G_3) along the spatiotemporal pattern evolution. Alternatively, norms (G_2) and phase (G_3) statistics can also be pre-selected to replace triangulation monitoring (G_1). This output corresponds to the application in SCGLE pattern formation.

Figure 6.1 - The GPA-Dash output scheme with its multiple features.

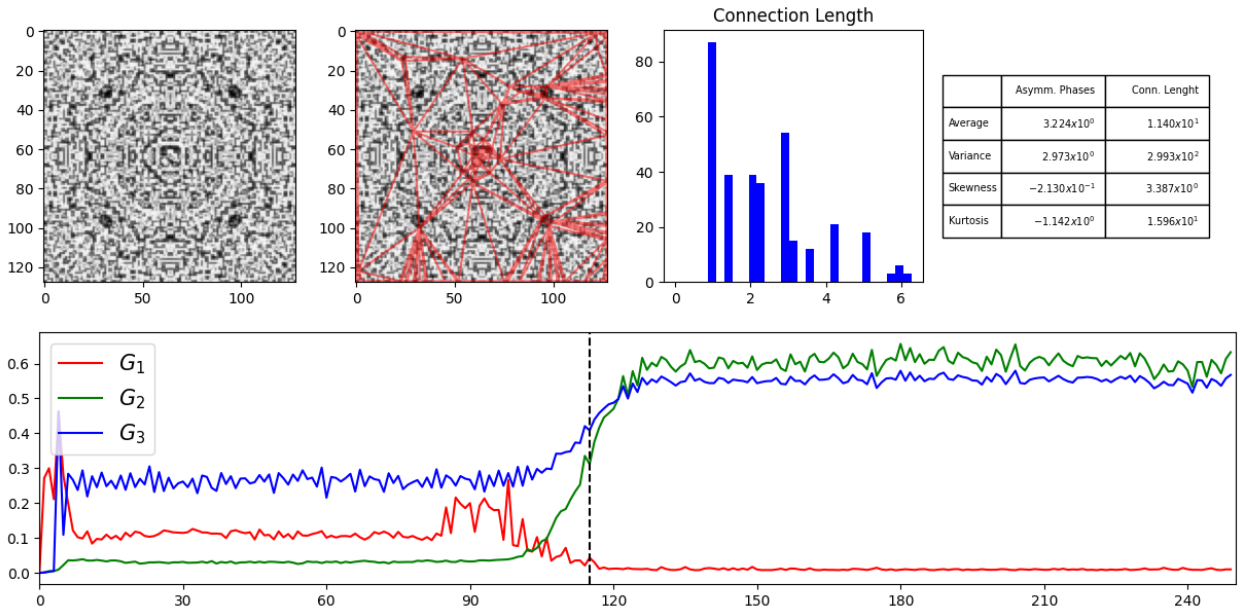


Norms (G_2) and phase (G_3) statistics can also be pre-selected to replace triangulation monitoring (G_1).

To demonstrate the capabilities of this tool, we have chosen three application examples: Coupled Map Lattice (CML), Solar Activity(LARA et al., 2008; SAWANT et al., 2011), and medical ultrasound(SAUTTER et al., 2023). One of these examples is the monitoring output of chaotic pattern formation from a CML, a system that extends classical chaotic maps and exhibits diverse dynamical behaviors. These behaviors include spatial synchronization, frozen chaos, and spatially driven chaos, depending on the chosen parameter set and initial conditions.

Figure 6.2 illustrates the analytical output of the CML monitoring process.

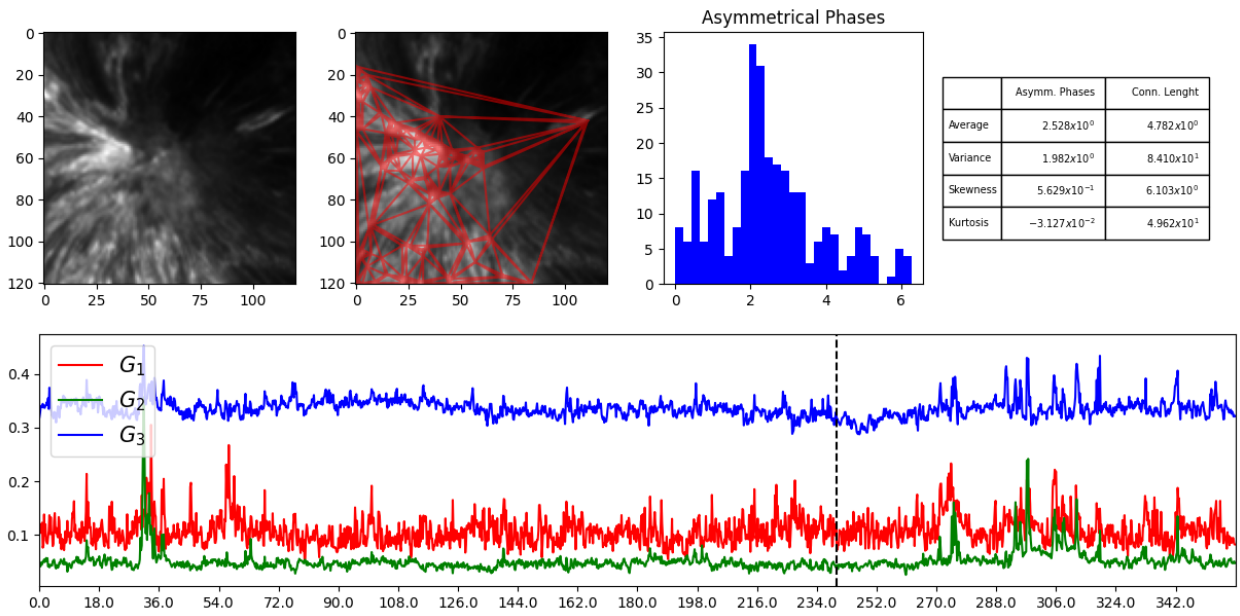
Figure 6.2 - Coupled Map Lattice Dashboard.



The line indicates the snapshot being analyzed.

Figure 6.3 shows the monitoring output of a solar active region NOAA 11131 observed by the SDO, as previously presented in Chapter 4.2, which has been presented in (SAUTTER et al., 2023).

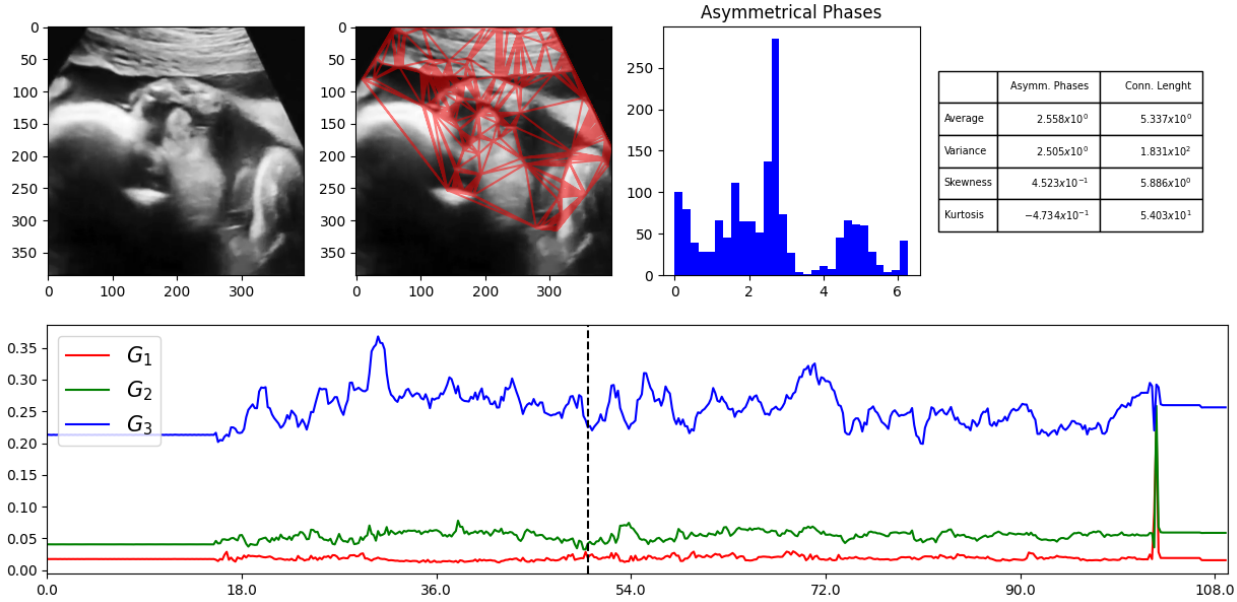
Figure 6.3 - Sunspot Dashboard.



The line indicates the snapshot being analyzed.

Figure 6.4 illustrates the results of an ultrasound scan used to detect and monitor the health and growth status of an unborn baby during pregnancy. The primary focus of this analysis is to examine the feasibility and effectiveness of using GPA in diagnostic medical imaging for identifying the baby through 2D ultrasound. This case study serves as a valuable exploration to evaluate the potential of GPA in enhancing diagnostic capabilities. The findings of this analysis have been presented in a paper (SAUTTER et al., 2023). Moreover, this particular application example has been recently integrated into the PIPE FAPESP Project 2021/15114-8 as an essential topic in applied research.

Figure 6.4 - Ultrasonography Dashboard.



The line indicates the snapshot being analyzed.

With this latest improvement over the GPA technique, we achieved the main objectives outlined in the original proposal. Conclusions and future work from this doctoral project in applied computing are discussed in the last chapter.

7 CONCLUDING REMARKS

In this doctoral work, we performed an unprecedented and improved application study of the Gradient Pattern Analysis (GPA) technique in the presence of noise and explored its applicability in different scenarios of dimensional data. Initially, we discuss the interpretation of GPA from a reference perspective, as described in the existing literature, and introduce new techniques for data visualization that had not been explored in the context of GPA until this work. The introduction of our proposed Dashboard structure allowed an efficient and flexible visualization of the analysis results.

Focusing on the pattern formation from the SCGL model, the hypothesis that the GPA is effective in the characterization of coherent structures obtained from Amplitude Equations was confirmed. For this, a new numerical methodology to calculate the SCGLE was developed and, in this context, new metrics for G_1 and G_3 were improved and implemented. In particular, G_1 proved to be effective in the dynamic analysis of an active solar region, strengthening its application in the context of space weather monitoring. It's also important to address the comparison between SCGL spatiotemporal patterns with emerging methodologies, such as the point vortex dynamics(KRIEGER et al., 2020).

For the first time, the GPA technique was applied in an integrated way to analyze the same reactive-diffusive phenomenon purely in time and in the 2D and 3D spatiotemporal domains. Still in the 3D+1 context, in addition to the SCGLE, the GPA was applied to large-scale structure formation data produced by the Millennium Consortium for Cosmological Simulation. The obtained results confirm and improve a previous analysis on the 2D data obtained from the Virgo (ANDRADE et al., 2006) consortium, suggesting new applications of this approach as a future work in computational cosmology.

Furthermore, we investigated the application of GPA in the analysis of Extreme Events (EE) in ordinary time series. Notably, the technique successfully distinguished different fluctuation patterns in time series associated with EE. In summary, this work added new contributions to the GPA technique, improving the interpretation methods in new areas of application. Based on these achievements, we propose an adaptation of the GPA technique to the context of Deep Learning (see Appendix D).

Regarding performance benchmarking with other measures in the context of this

thesis, we must clarify why they were not performed. Considering previous works that addressed the performance of GPA compared to other measures (e.g. DFA-2D(SOUZA; ASSIREU, 2016) and Minkowski functionals (ROSA et al., 2007; BARONI, 2009)) we ruled out the need for an equivalent benchmark in this project, since in previous works it became evident that the characterization of asymmetries in gradient lattices are not detected by any other technique or measurement published in the literature. Therefore, it is worth mentioning that the main objective of this project its just exploiting the thesis that GPA, with some increments, is effective to characterize the formation of 2D+1 patterns in the presence of noise. Detecting, for example, in the SCGLE dynamics, local oscillatory regimes caused by the presence of coherent structures like spiral defects (which are structures with bilateral asymmetry that affect the norms and mainly the phases of the gradient lattice vectors). However, we can also consider improvements to other techniques for future benchmark applications, following the same criteria presented for morphological classification of Galaxies by using GPA (ROSA et al., 2018; BARCHI et al., 2020).

REFERENCES

ANDRADE, A.; RIBEIRO, A. L. B.; ROSA, R. R. Gradient pattern analysis of cosmic structure formation: norm and phase statistics. **Physica D: Nonlinear Phenomena**, v. 223, n. 2, p. 139–145, 2006. 17, 67, 69, 77

ARANSON, I. S.; KRAMER, L. The world of the complex ginzburg-landau equation. **Reviews of Modern Physics**, v. 74, n. 1, p. 99, 2002. 4, 7, 28, 43

ARIS, R.; ARONSON, D. G.; SWINNEY, H. L. **Patterns and dynamics in reactive media**. [S.l.]: Springer Science & Business Media, 1991. 7

ASHWIN, P.; TERRY, J. R.; THORNBURG JUNIOR, K. S.; ROY, R. Blowout bifurcation in a system of coupled chaotic lasers. **Physical Review E**, v. 58, n. 6, p. 7186, 1998. 47

ASSAD, E. D.; CALMON, M.; LOPES-ASSAD, M. L.; FELTRAN-BARBIERI, R.; POMPEU, J.; DOMINGUES, L. M.; NOBRE, C. A. Adaptation and resilience of agricultural systems to local climate change and extreme events: an integrative review. **Pesquisa Agropecuária Tropical**, v. 52, p. e72899, 2022. ISSN 1983-4063. Available from: <<https://doi.org/10.1590/1983-40632022v5272899>>. 4

ASSIREU, A.; ROSA, R.; VIJAYKUMAR, N.; LORENZZETTI, J.; REMPEL, E.; RAMOS, F.; ABREU SÁ, L.; BOLZAN, M.; ZANANDREA, A. Gradient pattern analysis of short nonstationary time series: an application to lagrangian data from satellite tracked drifters. **Physica D: Nonlinear Phenomena**, v. 168-169, p. 397–403, 2002. ISSN 0167-2789. VII Latin American Workshop on Nonlinear Phenomena. Available from: <<https://www.sciencedirect.com/science/article/pii/S0167278902005274>>. 9, 33, 34, 55

BARCHI, P. H.; CARVALHO, R. de; ROSA, R. R.; SAUTTER, R.; SOARES-SANTOS, M.; MARQUES, B. A.; CLUA, E.; GONÇALVES, T.; SÁ-FREITAS, C. de; MOURA, T. Machine and deep learning applied to galaxy morphology-a comparative study. **Astronomy and Computing**, v. 30, p. 100334, 2020. 5, 9, 10, 11, 78

BARONI, M.; ROSA, R. Detrended fluctuation analysis of numerical density and viscous fingering patterns. **Europhysics Letters**, v. 92, n. 6, p. 64002, 2011. 1

BARONI, M. P. M. A. **Simulação, análise e caracterização computacional de padrões de fingering em deslocamentos de fluidos em meios porosos**. 123 p. Tese (Doutorado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2009. 4, 78

BARTLETT, M. S. Periodogram analysis and continuous spectra. **Biometrika**, v. 37, n. 1/2, p. 1–16, 1950. 53

BARZYKINA, I. Chemistry and mathematics of the belousov–zhabotinsky reaction in a school laboratory. **Journal of Chemical Education**, v. 97, n. 7, p. 1895–1902, 2020. 7

BOLZAN, M. J.; RAMOS, F. M.; SÁ, L. D.; NETO, C. R.; ROSA, R. R. Analysis of fine-scale canopy turbulence within and above an amazon forest using tsallis' generalized thermostatistics. **Journal of Geophysical Research: Atmospheres**, v. 107, n. D20, p. LBA–30, 2002. 8

BUDRONI, M.; WIT, A. D. Dissipative structures: from reaction-diffusion to chemo-hydrodynamic patterns. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, v. 27, n. 10, p. 104617, 2017. 7

BUTZ, A. R. Convergence with hilbert's space filling curve. **Journal of Computer and System Sciences**, v. 3, n. 2, p. 128–146, 1969. 33

BYRNE, G.; MARCOTTE, C. D.; GRIGORIEV, R. O. Exact coherent structures and chaotic dynamics in a model of cardiac tissue. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, v. 25, n. 3, p. 033108, 2015. 29

CAMPANHARO, A.; RAMOS, F.; MACAU, E.; ROSA, R.; BOLZAN, M.; SÁ, L. Searching chaos and coherent structures in the atmospheric turbulence above the amazon forest. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 366, n. 1865, p. 579–589, 2008. 8

CAMPOS VELHO, H. F.; ROSA, R. R.; RAMOS, F. M.; PIELKE, R. A.; DEGRAZIA, G. A.; RODRIGUES NETO, C.; ZANANDREA, A. Multifractal model for eddy diffusivity and counter-gradient term in atmospheric turbulence. **Physica A: Statistical Mechanics and its Applications**, v. 295, n. 1, p. 219–223, 2001. ISSN 0378-4371. Available from: <<https://www.sciencedirect.com/science/article/pii/S0378437101000772>>. 8

CARETTA, C. A.; ROSA, R. R.; VELHO, H. de C.; RAMOS, F. M.; MAKLER, M. Evidence of turbulence-like universality in the formation of galaxy-sized dark matter haloes. **Astronomy & Astrophysics**, v. 487, n. 2, p. 445–451, 2008. 69

CHO, K.; CHAE, J.; LIM, E.-k.; YANG, H. The observational evidence for the internal excitation of sunspot oscillations inferred from the fe i 5435 å line. **The Astrophysical Journal**, v. 879, n. 2, p. 67, 2019. 48

CLADIS, P.; PALFFY-MUHORAY, P. **Spatio-temporal patterns in nonequilibrium complex systems**. 1. ed. Massachusetts: Addison Wesley Longman, 1995. 704 p. ISBN 978-0201409871. 3

COELHO, D. L.; VITRAL, E.; PONTES, J.; MANGIAVACCHI, N. Stripe patterns orientation resulting from nonuniform forcings and other competitive effects in the swift–hohenberg dynamics. **Physica D: Nonlinear Phenomena**, Elsevier, v. 427, p. 133000, 2021. 7

COLLET, P.; ECKMANN, J.-P. **Instabilities and fronts in extended systems**. [S.l.]: Princeton University Press, 2014. 7

CONWAY, J. H.; BURGIEL, H.; GOODMAN-STRAUSS, C. **The symmetries of things**. [S.l.]: CRC Press, 2016. 40

COSTA JÚNIOR, R. A. **Desenvolvimento e aplicações de um ambiente computacional para a análise de padrões-gradientes**. 202 p. Tese (Doutorado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2004. 4

CROSS, M. C.; HOHENBERG, P. C. Pattern formation outside of equilibrium. **Reviews of Modern Physics**, v. 65, n. 3, p. 851, 1993. 3, 7

DANTAS, M. S. **Um ambiente virtual colaborativo de computação científica para análise avançada de séries temporais**. 175 p. Tese (Doutorado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2014. 4

DEKEL, A.; OSTRIKER, J. P. **Formation of structure in the Universe**. [S.l.]: Cambridge University Press, 1999. 3

DERES, A.; ANFINOGENTOV, S. Measurement of the formation heights of uv and euv emission above sunspot umbrae from observations of three-minute oscillations. **Astronomy Reports**, v. 59, p. 959–967, 2015. 48

DIMITROPOULOS, K.; BARMPOUTIS, P.; GRAMMALIDIS, N. Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 25, n. 2, p. 339–351, 2014. 3

DJEDDI, R.; EKICI, K. Resolution of gibbs phenomenon using a modified pseudo-spectral operator in harmonic balance cfd solvers. **International Journal of Computational Fluid Dynamics**, v. 30, n. 7-10, p. 495–515, 2016. 24

DÚZS, B.; HOLLÓ, G.; KITAHATA, H.; GINDER, E.; SUEMATSU, N. J.; LAGZI, I.; SZALAI, I. Appearance and suppression of turing patterns under a periodically forced feed. **Communications Chemistry**, v. 6, n. 1, p. 3, 2023. 3

FARAG, N. G.; ELTANBOLY, A. H.; EL-AZAB, M.; OBAYYA, S. Pseudo-spectral approach for extracting optical solitons of the complex ginzburg landau equation with six nonlinearity forms. **Optik**, v. 254, p. 168662, 2022. 23

FELIPE, T. Origin of the chromospheric three-minute oscillations in sunspot umbrae. **Astronomy & Astrophysics**, v. 627, p. A169, 2019. 48

FISHER, R. A. **The genetical theory of natural selection: a complete variorum edition**. [S.l.]: Oxford University Press, 1999. 3

FREITAS, R. M. **Laboratório virtual para visualização e caracterização do uso e cobertura da terra utilizando imagens de sensoriamento remoto**. 202 p. Tese (Doutorado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2012. 4

FRIENDLY, M. A brief history of data visualization. In: CHEN, C. H. (Ed.). **Handbook of data visualization**. Berlin: Springer, 2008. p. 15–56. 71

GALLAIRE, F.; BRUN, P.-T. Fluid dynamic instabilities: theory and application to pattern forming in complex media. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 375, n. 2093, p. 20160155, 2017. 7

GUERRA, J. M. **Caracterização fina dos padrões de variabilidade do ECG para validação de modelos e aplicações em microgravidade**. 141 p. Dissertação (Mestrado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2008. 4

GUO, Q.; WHITE, S. D. Galaxy growth in the concordance λ cdm cosmology. **Monthly Notices of the Royal Astronomical Society**, v. 384, n. 1, p. 2–10, 2008. 43, 50, 51, 67

HAO, Y.; XU, Z.-J.; LIU, Y.; WANG, J.; FAN, J.-L. Effective crowd anomaly detection through spatio-temporal texture analysis. **International Journal of Automation and Computing**, v. 16, n. 1, p. 27–39, 2019. 3

HUPKES, H. J.; MORELLI, L.; SCHOUTEN-STRAATMAN, W. M.; VLECK, E. S. V. Traveling waves and pattern formation for spatially discrete bistable reaction-diffusion equations. In: BOHNER, M.; SIEGMUND, S.; HILSCHER, R. Š.; STEHLÍK, P. (Ed.). **Difference Equations and Discrete Dynamical Systems with Applications**. Berlim: Springer International Publishing, 2020. p. 55–112. ISBN 978-3-030-35502-9. 7

JABEEN, Z.; GUPTE, N. The dynamical origin of the universality classes of spatiotemporal intermittency. **Physics Letters A**, v. 374, n. 44, p. 4488–4495, 2010. 7

JANES, A.; SILLITTI, A.; SUCCI, G. Effective dashboard design. **Cutter IT Journal**, v. 26, n. 1, p. 17–24, 2013. 71

JENKINS, A.; FRENK, C.; PEARCE, F.; THOMAS, P.; COLBERG, J.; WHITE, S. D.; COUCHMAN, H.; PEACOCK, J.; EFSTATHIOU, G.; NELSON, A. Evolution of structure in cold dark matter universes. **The Astrophysical Journal**, v. 499, n. 1, p. 20, 1998. 33, 50

KANAMARU, T. Blowout bifurcation and on–off intermittency in pulse neural networks with multiplec modules. **International Journal of Bifurcation and Chaos**, v. 16, n. 11, p. 3309–3321, 2006. 47

KRAWIECKI, A.; MATYJAŚKIEWICZ, S. Blowout bifurcation and stability of marginal synchronization of chaos. **Physical Review E**, v. 64, n. 3, p. 036216, 2001. 43, 47

KRIEGER, M. S.; SINAI, S.; NOWAK, M. A. Turbulent coherent structures and early life below the kolmogorov scale. **Nature Communications**, v. 11, n. 1, p. 2192, 2020. 77

LAGAE, A.; LEFEBVRE, S.; COOK, R.; DEROSE, T.; DRETTAKIS, G.; EBERT, D. S.; LEWIS, J. P.; PERLIN, K.; ZWICKER, M. A survey of procedural noise functions. v. 29, n. 8, p. 2579–2600, 2010. 19

LANGBEIN, J. Noise in two-color electronic distance meter measurements revisited. **Journal of Geophysical Research: Solid Earth**, v. 109, n. B4, 2004. 19

LANGNER, A.; MIETTINEN, J.; KUKKONEN, M.; VANCUTSEM, C.; SIMONETTI, D.; VIEILLEDENT, G.; VERHEGGHEN, A.; GALLEGO, J.; STIBIG, H.-J. Towards operational monitoring of forest canopy disturbance in evergreen rain forests: A test case in continental southeast asia. **Remote Sensing**, v. 10, n. 4, p. 544, 2018. 8

LARA, A.; BORGAZZI, A.; MENDES, O. J.; ROSA, R. R.; DOMINGUES, M. O. Short-period fluctuations in coronal mass ejection activity during solar cycle 23. **Solar Physics**, v. 248, p. 155–166, 2008. 73

LEJEUNE, O.; COUTERON, P.; LEFEVER, R. Short range co-operativity competing with long range inhibition explains vegetation patterns. **Acta Oecologica**, v. 20, n. 3, p. 171–183, 1999. ISSN 1146-609X. Available from: <<https://www.sciencedirect.com/science/article/pii/S1146609X99800307>>. 8

LI, S. et al. Predicting lung nodule malignancies by combining deep convolutional neural network and handcrafted features. **Physics in Medicine & Biology**, v. 64, n. 17, p. 175012, 2019. 95

LIU, Q. H. Large-scale simulations of electromagnetic and acoustic measurements using the pseudospectral time-domain (pstd) algorithm. **IEEE Transactions on Geoscience and Remote Sensing**, v. 37, n. 2, p. 917–926, 1999. 24

LIU, Y. et al. Computational symmetry in computer vision and computer graphics. **Foundations and Trends® in Computer Graphics and Vision**, v. 5, n. 1–2, p. 1–195, 2010. 11

LUCIA, G. D.; SPRINGEL, V.; WHITE, S. D.; CROTON, D.; KAUFFMANN, G. The formation history of elliptical galaxies. **Monthly Notices of the Royal Astronomical Society**, v. 366, n. 2, p. 499–509, 2006. 50

MANNEVILLE, P. Dissipative structures and weak turbulence. In: WERON M. W. A.; GARBACZEWSKI, P. (Ed.). **Chaos—the Iinterplay Bbetween stochastic and deterministic behaviour: Proceedings of the XXXIst**

Winter School of Theoretical Physics Held in Karpacz, Poland 13–24 February 1995. [S.l.]: Springer, 2005. p. 257–272. 7

MECHETER, I.; ABBOD, M.; AMIRA, A.; ZAIDI, H. Deep learning with multiresolution handcrafted features for brain mri segmentation. **Artificial Intelligence in Medicine**, v. 131, p. 102365, 2022. 95

MENEVEAU, C.; SREENIVASAN, K. Simple multifractal cascade model for fully developed turbulence. **Physical Review Letters**, v. 59, n. 13, p. 1424, 1987. 43

_____. The multifractal nature of turbulent energy dissipation. **Journal of Fluid Mechanics**, v. 224, p. 429–484, 1991. 45

NANNI, L.; LUCA, E. D.; FACIN, M. L.; MAGUOLO, G. Deep learning and handcrafted features for virus image classification. **Journal of Imaging**, v. 6, n. 12, p. 143, 2020. 95

NOBRE, C. A.; FABRICIO-NETO, A. The amazon forest and climate change: A sustainable pathway to avoid a tipping point. In: M, P.; C., R.; M., D. (Ed.). **ur Warming Planet: climate change impacts and adaptation**. Singapura: World Scientific, 2021. p. 412–415. 4

RAMOS, F. M.; ROSA, R. R.; RODRIGUES NETO, C.; ZANANDREA, A. Generalized complex entropic form for gradient pattern analysis of spatio-temporal dynamics. **Physica A: Statistical Mechanics and its Applications**, v. 283, n. 1-2, p. 171–174, 2000. 1, 9, 14, 15, 89

RAYLEIGH, L. Lix. on convection currents in a horizontal layer of fluid, when the higher temperature is on the under side. **The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science**, v. 32, n. 192, p. 529–546, 1916. 3

RODRIGUES NETO, C.; ZANANDREA, A.; RAMOS, F.; ROSA, R.; BOLZAN, M.; SÁ, L. Multiscale analysis from turbulent time series with wavelet transform. **Physica A: Statistical Mechanics and its Applications**, v. 295, n. 1-2, p. 215–218, 2001. 8

ROSA, R.; BARONI, M.; ZANIBONI, G.; FERREIRA DA SILVA, A.; ROMAN, L.; PONTES, J.; BOLZAN, M. Structural complexity of disordered surfaces: analyzing the porous silicon sfm patterns. **Physica A: Statistical Mechanics and its Applications**, v. 386, n. 2, p. 666–673, 2007. ISSN 0378-4371. Disorder and Complexity. Available from: <<https://www.sciencedirect.com/science/article/pii/S0378437107009557>>. 4, 78

ROSA, R.; CARVALHO, R. D.; SAUTTER, R.; BARCHI, P.; STALDER, D.; MOURA, T.; REMBOLD, S.; MORELL, D.; FERREIRA, N. Gradient pattern analysis applied to galaxy morphology. **Monthly Notices of the Royal Astronomical Society: Letters**, v. 477, n. 1, p. L101–L105, 2018. 1, 5, 9, 11, 12, 13, 78

ROSA, R.; KARLICKỲ, M.; VERONESE, T.; VIJAYKUMAR, N.; SAWANT, H.; BORGAZZI, A.; DANTAS, M.; BARBOSA, E.; SYCH, R.; MENDES, O. Gradient pattern analysis of short solar radio bursts. **Advances in Space Research**, v. 42, n. 5, p. 844–851, 2008. 43

ROSA, R.; PONTÈS, J.; CHRISTOV, C.; RAMOS, F. M.; NETO, C. R.; REMPEL, E. L.; WALGRAEF, D. Gradient pattern analysis of swift–hohenberg dynamics: phase disorder characterization. **Physica A: Statistical Mechanics and its Applications**, v. 283, n. 1-2, p. 156–159, 2000. 1

ROSA, R.; R, C.; RAMOS, F.; L, V.; FUJIWARA, S.; SATO, T. Gradient pattern analysis of structural dynamics: application to molecular system relaxation. **Brazilian Journal of Physics**, v. 33, 09 2003. 1, 9, 11, 15

ROSA, R. R. Data science strategies for multimessenger astronomy. **Anais da Academia Brasileira de Ciências**, v. 93, p. e20200861, 2020. 95

ROSA, R. R.; DANTAS, M. S.; RANGEL, R.; VIJAYKUMAR, N. L.; VERONESE T. B., B. Gradient pattern analysis of asymmetric fluctuations: applications for short time series and disordered spatial structures. In: ISI INTERNATIONAL STATISTICAL INSTITUTE, 56. **Proceedings...** [S.l.], 2007. 34

ROSA, R. R.; SHARMA, A.; VALDIVIA, J. Characterization of asymmetric fragmentation patterns in spatially extended systems. **International Journal of Modern Physics C**, v. 10, n. 01, p. 147–163, 1999. 4, 9, 10, 12, 17, 18, 38

SANTOS, L. **Caracterização computacional de padrões estruturais em seqüências de DNA relacionadas a processos em redes metabólicas**. 126 p. Dissertação (Mestrado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2009. 4

SAUTTER, R.; RR, R.; SYCH, R.; SAWANT, H.; BISOI, S. Gradient pattern analysis of the solar active region noaa 11131. **Boletim da Sociedade Astronômica Brasileira**, v. 34, p. 231–232, 2023. 13, 32, 49, 73, 74, 91

SAUTTER, R. A. **Gradient Pattern Analysis: new methodological and computational features with applications**. 97 p. Dissertação (Mestrado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2018. 1, 4, 89

SAUTTER, R. A.; ROSA, R. R.; ALAVARCE, D. C.; SILVA, D. G. Gradient pattern analysis applied for computer vision in medical ultrasound diagnosis. **international journal on computers science and information systems**, 2023. ISSN 1646-3692. 75

SAWANT, H. S.; GOPALSWAMY, N.; ROSA, R. R.; SYCH, R. A.; ANFINOGENTOV, S. A.; FERNANDES, F. C.; CECATTO, J. R.; COSTA, J. E. The brazilian decimetric array and space weather. **Journal of Atmospheric and Solar-Terrestrial Physics**, v. 73, n. 11-12, p. 1300–1310, 2011. 73

SCHREIBER, T.; SCHMITZ, A. Surrogate time series. **Physica D: Nonlinear Phenomena**, v. 142, n. 3-4, p. 346–382, 2000. 93

SCHWENDIMANN, B. A.; RODRIGUEZ-TRIANA, M. J.; VOZNIUK, A.; PRIETO, L. P.; BOROUJENI, M. S.; HOLZER, A.; GILLET, D.; DILLENBOURG, P. Perceiving learning at a glance: A systematic literature review of learning dashboard research. **IEEE Transactions on Learning Technologies**, v. 10, n. 1, p. 30–41, 2016. 71

SHAMPINE, L. F. Some practical runge-kutta formulas. **Mathematics of Computation**, v. 46, n. 173, p. 135–150, 1986. 24

SHEN, T.; XIN, J.; HUANG, J. Time-space fractional stochastic ginzburg-landau equation driven by gaussian white noise. **Stochastic Analysis and Applications**, v. 36, n. 1, p. 103–113, 2018. 7

SOUZA, V.; ASSIREU, A. Detrended fluctuation analysis of spatially extended digital surfaces: the classification process of 1/f noise and computacional performance. **Journal of Computational Interdisciplinary Sciences**, v. 24, p. 25, 2016. 4, 78

SPRINGEL, V. The cosmological simulation code gadget-2. **Monthly Notices of the Royal Astronomical Society**, v. 364, n. 4, p. 1105–1134, 2005. 50

SWINNEY, H.; KRINSKY, V. **Waves and patterns in chemical and biological media**. [S.l.]: MIT Press, Cambridge, MA and Elsevier, Amsterdam, 1992. 7

SYCH, R.; JESS, D. B.; SU, J. The dynamics of 3-min wavefronts and their relation to sunspot magnetic fields. **Philosophical Transactions of the Royal Society A**, v. 379, n. 2190, p. 20200180, 2021. 48

SYCH, R.; ZAQARASHVILI, T.; NAKARIAKOV, V.; ANFINOGENTOV, S.; SHIBASAKI, K.; YAN, Y. Frequency drifts of 3-min oscillations in microwave and euv emission above sunspots. **Astronomy & Astrophysics**, v. 539, p. A23, 2012. 43, 48

THEILER, J.; EUBANK, S.; LONGTIN, A.; GALDRIKIAN, B.; FARMER, J. D. Testing for nonlinearity in time series: the method of surrogate data. **Physica D: Nonlinear Phenomena**, v. 58, n. 1-4, p. 77–94, 1992. 93

TIMMER, J.; KOENIG, M. On generating power law noise. **Astronomy and Astrophysics**, v. 300, p. 707, 1995. 19, 43

TURING, A. M. The chemical basis of morphogenesis. **Bulletin of Mathematical Biology**, v. 52, p. 153–197, 1990. 3

VIRTANEN, P. et al. Scipy 1.0: fundamental algorithms for scientific computing in python. **Nature Methods**, v. 17, n. 3, p. 261–272, 2020. 38

WALGRAEF, D. The hopf bifurcation and related spatio-temporal patterns. In: WALGRAEF, D. (Ed.). **Spatio-Temporal Pattern Formation: with examples from physics, chemistry, and materials science**. Berlin: Springer, 1997. p. 65–85. 3, 7

WANG, L.; XU, N.; SONG, J. Decoding intra-tumoral spatial heterogeneity on radiological images using the hilbert curve. **Insights into Imaging**, v. 12, p. 1–10, 2021. 33

WEISSTEIN, E. W. Spherical coordinates. In: . [S.l.]: Wolfram Research, Inc., 2005. 37

WINFREE, A. Heart muscle as a reaction–diffusion medium: the roles of electric potential diffusion, activation front curvature, and anisotropy. **International Journal of Bifurcation and Chaos**, v. 7, n. 03, p. 487–526, 1997. 7

WU, J. Introduction to convolutional neural networks. **China: Nanjing**, 2017. 95

XU, C. An easy algorithm to generate colored noise sequences. **The Astronomical Journal**, v. 157, n. 3, p. 127, 2019. 19

XU, L.; ZHANG, G.; HAN, B.; ZHANG, L.; LI, M.; HAN, Y. Turing instability for a two-dimensional logistic coupled map lattice. **Physics Letters A**, v. 374, n. 34, p. 3447–3450, 2010. 89

XU, Y.; QUAN, Y.; ZHANG, Z.; LING, H.; JI, H. Classifying dynamic textures via spatiotemporal fractal analysis. **Pattern Recognition**, v. 48, n. 10, p. 3239–3248, 2015. 3

YANG, D.; HOU, Z. Large deviations for the stochastic derivative ginzburg–landau equation with multiplicative noise. **Physica D: Nonlinear Phenomena**, v. 237, n. 1, p. 82–91, 2008. 7

YIN, B.; BALVERT, M.; ZAMBRANO, D.; SCHÖNHUTH, A.; BOHTE, S. An image representation based convolutional network for dna classification. **arXiv preprint arXiv:1806.04931**, 2018. 33

ZENG, W.; XIAO, A.; LI, X. Error estimate of fourier pseudo-spectral method for multidimensional nonlinear complex fractional ginzburg–landau equations. **Applied Mathematics Letters**, v. 93, p. 40–45, 2019. 23

APPENDIX A - COUPLED MAP LATTICES

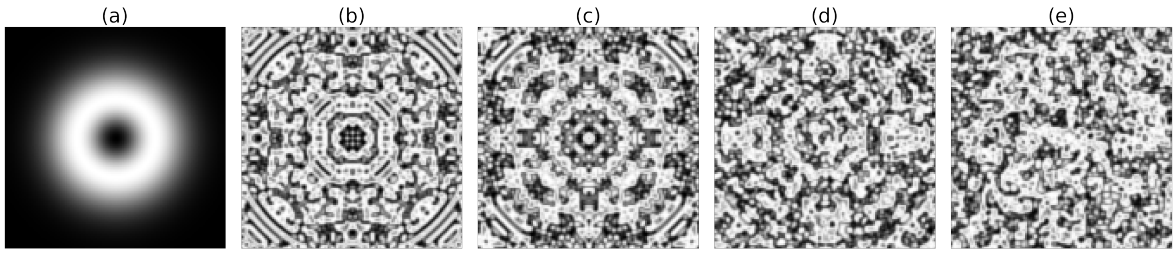
Here we present the Coupled Map Lattice (CML)[1], a dynamical system that has been proposed as an extension to classical chaotic maps.

The CML is composed of oscillators $f(u^t_{x,y})$ in a regular grid, in which spatial interaction is controlled by a coupling parameter ϵ. The system state $u^t_{x,y}$ is updated at every time step according to Equation A.1.

$$u^{t+1}_{x,y} = (1-\epsilon)f(u^t_{x,y}) + \frac{\epsilon}{4}\left(f(u^t_{x+1,y}) + f(u^t_{x-1,y}) + f(u^t_{x,y+1} + f(u^t_{x,y-1})\right). \quad \text{(A.1)}$$

In CML studies, the logistic map has become a standard case (RAMOS et al., 2000; XU et al., 2010), which has the form: $f(u) = \lambda u(1-u)$. The term λ is a parameter that controls the local regime of the series. Some snapshots of this system's iteration are presented in Figure A.1

Figure A.1 - Example of Coupled Map Lattice snapshots.



The simulation was performed in a 128x128 lattice. The parameter $\lambda = 3.9$ and $\epsilon = 0.5$. Snapshots (a) - (e) are respectively the iterations: 1, 50, 100, 150, and 200.

In a range $\lambda = [3, 4]$, the map presents all 2^N periodic orbits and chaotic regimes. We have presented a more in-depth study of this dynamical system in a previous work (SAUTTER, 2018).

[1]The source code is public-available at: https://github.com/rsautter/CML-CoupledMapLattice

APPENDIX B - CONTRIBUTIONS AND PUBLICATIONS

This work has made significant contributions both in terms of scientific publications and software development. In terms of scientific publications, two papers have been submitted. The first paper explores the analysis of GPA within the context of the Stochastic Complex Ginzburg-Landau and has been submitted to Physical Review E. The second paper, submitted to EuroPhysics Letters, includes the GPA analysis of Coupled Map Lattices, a previous result from our work, with adaptations in terms of data visualization.

Additionally, this work has actively contributed to the scientific community through participation in RASAB (SAUTTER et al., 2023), where a case study of NOAA 11131 has been presented. In addition, a paper analyzing the milli-millennium set was accepted at the Complex Systems Conference 2023.

In terms of software development, several repositories have been created to support the implementation of various methodologies and algorithms developed in this work. These include:

- GPA (2D and 3D): https://github.com/rsautter/GPA
- GPA 1D: https://github.com/rsautter/GPA1D
- χ-Space: https://github.com/rsautter/Chi-Space
- Milli-Millennium query: https://github.com/rsautter/MillenniumQuery
- SCGLE: https://github.com/rsautter/Noisy-Complex-Ginzburg-Landau

These contributions collectively enhance the scientific understanding of GPA and provide accessible tools.

APPENDIX C - SURROGATE

Surrogate techniques are valuable methods used to transform a given time series into a new series while altering specific properties while preserving others. Various techniques have been developed and studied extensively in the literature. For a comprehensive overview of popular surrogate techniques, Schreiber and Schmitz (2000) provide a review that highlights their characteristics and applications.

One commonly used surrogate technique is shuffling, which involves randomizing the order of the original time series while preserving its statistical distribution. This operation breaks the temporal structure of the data, making it useful for validating certain analyses. The shuffling approach has been applied to validate the GPA1D.

A more refined surrogate technique is the Amplitude Adjusted Fourier Transform (AAFT), originally proposed by Theiler et al. (1992). The AAFT technique aims to preserve the Power Spectrum Density (PSD) of the original time series while randomizing the phases. To achieve this, the technique involves performing a Fourier transform on the time series, randomizing the phases of the Fourier coefficients, and then applying an inverse Fourier transform to obtain the surrogate time series. By maintaining the PSD, the AAFT method ensures that the surrogate data exhibit similar frequency content as the original series.

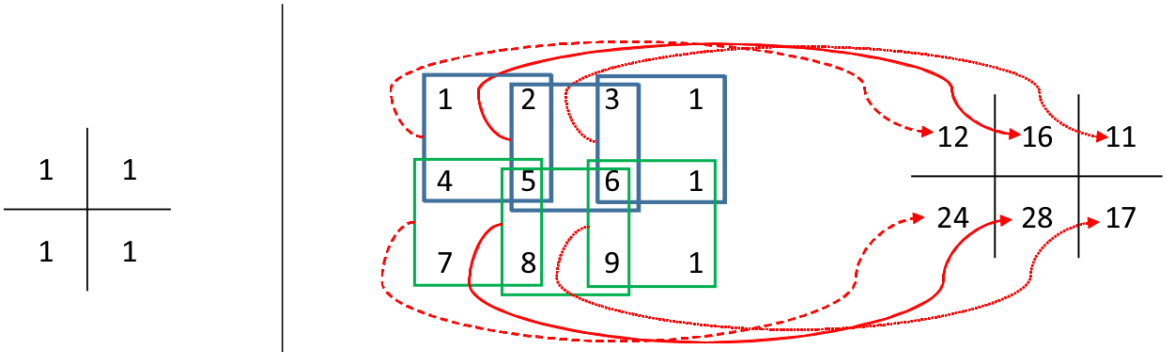
The advantage of the AAFT technique over simple shuffling lies in its ability to preserve the power spectrum characteristics of the original data. This refinement enables a more accurate assessment of the significance of observed patterns, as it accounts for the frequency distribution present in the original time series. Consequently, the AAFT method is often preferred when analyzing time-dependent data where both temporal structure and frequency content play crucial roles.

APPENDIX D - DEEP LEARNING PERSPECTIVES

Deep Neural Networks (DNN) have emerged as the new paradigm for machine learning analysis. These techniques have shown significant improvements, particularly when incorporating handcrafted features (LI et al., 2019; NANNI et al., 2020; MECHETER et al., 2022). Therefore, it is crucial to extend the GPA technique to the Deep Learning into the context of data science (ROSA, 2020).

Typically, the incorporation of handcrafted features in DNN models involves fusion algorithms (LI et al., 2019; NANNI et al., 2020) or data augmentation techniques (MECHETER et al., 2022). However, the GPA technique can also be integrated directly into the architecture of a DNN. Similar to a convolution operation, the GPA technique can be generalized as a neuronal operation that leverages spatial properties. In a convolution operation, a moving kernel with weights is applied to different segments of the matrices, as illustrated in Figure D.1.

Figure D.1 - Illustration of the convolution operation



(a) A 2×2 kernel

(b) The convolution input and output

SOURCE: Wu (2017).

A possible G_2 neuron approximation would be given by Equation D.1. In this equation, $v_{i,j}$ represents the spatial gradient, $w_{i,j}$ are the weights, and $m_{i,j}$ is the asymmetry binary matrix.

$$y_{i,j} = \frac{\sum m_{i,j} w_{i,j} |v_{i,j}|}{|\sum m_{i,j} w_{i,j} v_{i,j}|} \tag{D.1}$$

The asymmetry binary matrix captures the information about the concentric asymmetry, where asymmetrical vectors are represented as ones, and the remaining vectors are represented as zeros. It is important to note that the term $|\sum m_{i,j} w_{i,j} v_{i,j}|$ should not be zero to prevent issues such as overflow, underflow, or divergence. To handle such cases, we can define the values $y_{i,j} = \pm 1$ in the event of overflow or underflow, respectively, and $y_{i,j} = 0$ when $\sum |v_{i,j}| = 0$.

To optimize the weights of the neurons, the Gradient Descent Method can be applied. This numerical method requires the derivatives of the operation with respect to the weights. By differentiating Equation D.1 in relation to the weights, we obtain Equation D.2. This Equation also comes with a limitation in the denominator term. However, in this scenario we can set the derivative as zero to avoid numerical problems, without changing the weights of the neurons.

$$\frac{\partial y_{i,j}}{\partial w_{i,j}} = \sum m_{i,j} |v_{i,j}| \left(\frac{|\sum m_{i,j} w_{i,j} v_{i,j}| - w_{i,j} |\sum m_{i,j} v_{i,j}|}{|\sum m_{i,j} w_{i,j} v_{i,j}|^2} \right) \tag{D.2}$$

As the convolutional neural networks, the proposed method can be stacked in layers, which may contain an activation functions and other types of neurons.

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.