UNIVERSIDADE FEDERAL DO PARANÁ

FELIPE EDUARDO LOPES DA CRUZ

ESTUDO SOBRE O COMPRIMENTO MÍNIMO DAS DIAGONAIS EM GRÁFICOS DE RECORRÊNCIA

CURITIBA - PR

2021

FELIPE EDUARDO LOPES DA CRUZ

ESTUDO SOBRE O COMPRIMENTO MÍNIMO DAS DIAGONAIS EM GRÁFICOS DE RECORRÊNCIA

Dissertação apresentada ao curso de Pós-Graduação em Física, Setor de Ciências Exatas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Física.

Orientador: Prof. Dr. Thiago de Lima Prado.

CURITIBA - PR
2021

Catalogação na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

C957e Cruz , Felipe Eduardo Lopes da
Estudo sobre o comprimento mínimo das diagonais em gráficos de recorrência [recurso eletrônico] / Felipe Eduardo Lopes da Cruz– Curitiba, 2021.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas. Programa de Pós-graduação em Física.

Orientador: Prof. Dr. Thiago de Lima Prado

1. Análise de séries temporais. I. Universidade Federal do Paraná. II. Prado, Thiago de Lima. III. Título.

CDD: 515.39

Bibliotecária: Roseny Rivelini Morciani CRB-9/1585

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em FÍSICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **FELIPE EDUARDO LOPES DA CRUZ** intitulada: **"Estudo Sobre o Comprimento Mínimo das Diagonais em Gráficos de Recorrência"** , sob orientação do Prof. Dr. THIAGO DE LIMA PRADO, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 01 de Outubro de 2021.

Assinatura Eletrônica
01/10/2021 17:46:12.0
THIAGO DE LIMA PRADO
Presidente da Banca Examinadora

Assinatura Eletrônica
01/10/2021 18:27:33.0
SERGIO ROBERTO LOPES
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica
01/10/2021 18:10:20.0
ELBERT EINSTEIN NEHRER MACAU
Avaliador Externo (UNIVERSIDADE FEDERAL DE SÃO PAULO)

Agradecimentos

Os agradecimentos mais necessários serão direcionados ao meu orientador, Professor Dr. Thiago de Lima Prado, pois este trabalho apenas foi finalizado devido ao apoio e paciência do mesmo desde o início do projeto, onde sempre conversávamos acerca de possíveis implementações e diferentes maneiras de se conduzir a produção desta dissertação.

À Universidade Federal do Paraná, por ceder a infraestrutura e recursos necessários para a conclusão deste trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, que através do Programa de Excelência Acadêmica - PROEX, forneceu apoio financeiro para a realização desta dissertação, sem o qual ela não seria finalizada.

Aos membros da Coordenação e da Secretaria do Programa de Pós-Graduação em Física da Universidade Federal do Paraná, que sempre estavam dispostos a ajudar quando necessário.

RESUMO

Neste trabalho será abordado o conceito e algumas propriedades de séries temporais, assim como serão citadas formas de analisá-las. Em particular, será discutido o método que utiliza gráficos de recorrências. Dentro deste contexto, serão mostrados algumas formas de se analisar um gráfico de recorrências a partir de seus quantificadores utilizados no âmbito da análise de quantificação de recorrências. Dentre estes quantificadores, há um em particular chamado *determinismo*, que possui em sua definição um parâmetro livre denotado por l_{min}, que atualmente é definido por critérios pouco objetivos. O objetivo deste trabalho é discutir a forma atual de escolha deste parâmetro, assim como propor um critério de escolha para o l_{min} com base em duas restrições e no princípio de maximização de entropia. O critério proposto foi aplicado para séries com diferentes comportamentos dinâmicos no cálculo do determinismo, e foi visto que de fato é possível encontrar o parâmetro em questão de forma que as restrições sejam satisfeitas. Os resultados mostram também que é possível gerar valores para o determinismo em maior concordância com seu comportamento.

Palavras-chave: séries temporais, análise de quantificação de recorrências, determinismo, parâmetro livre.

ABSTRACT

In this work the concept and some properties of time series will be addressed, as well as ways to analyze them will be mentioned. In particular, the method of recurrence plots will be discussed. Within this context, some ways to analyze a recurrence graph from its quantifiers used in the scope of recurrence quantification analysis will be shown. Among these quantifiers there is one in particular called *determinism*, which has in its definition a free parameter denoted by l_{min}, which is currently defined by non-objective criteria. The objective of this work is to discuss the current way of choosing this parameter, as well as to propose a choice criterion for the l_{min} based on two restrictions and the principle of entropy maximization. The proposed criterion was applied to series with different dynamic behaviors in the determinism calculation and it was seen that it is in fact possible to find the parameter in question so that the restrictions are satisfied. The results also show that it is possible to generate values for determinism in greater agreement with its behavior.

Keywords: time series, recurrence quantification analysis, determinism, free parameter.

Lista de Figuras

Figura 1 – Exemplo de séries temporais encontradas em diversas áreas do conhecimento. Na figura (a) é ilustrada a série temporal do volume de chuvas na região de Campina Grande - PB no período de 1961-2018, retirada da referência [31]. Na figura (b) é vista a série temporal do número diário de novas infecções por COVID-19 até 02 de junho de 2020, retirada da referência [32]. Em (c) vemos a série temporal de sucessivas medições da intensidade de um LASER NMR, retirada da referência [7]. 17

Figura 2 – Um ruído branco uniformemente distribuído (acima) dividido em diferentes intervalos e as esperanças $\mathbb{E}(x)$ da série dentro de seus respectivos intervalos (abaixo). Como a esperança é constante no tempo, esta série é estacionária. 19

Figura 3 – Um seno contaminado com 25% de ruído branco (acima) dividido em diferentes intervalos e a esperança $\mathbb{E}(x)$ da série dentro de seus respectivos intervalos (abaixo). Há uma dependência da esperança com o tempo, fazendo com que esta não seja uma série estacionária. 20

Figura 4 – Ilustração de um volume dependente de ε em um espaço de estados mostrando um segmento de órbita que será considerado recorrente. Figura retirada da referência [25]. 26

Figura 5 – Ilustração de tipologia regular. Em (a) vemos uma série temporal periódica senoidal com 5 períodos, em (b) vemos sua trajetória reconstruída em um espaço de estados através de um *embedding* e, em (c), o gráfico de recorrências associado à órbita. 27

Figura 6 – Ilustração de tipologia caótica. Em (a) vemos uma série temporal caótica correspondente à componente x do sistema de Rössler, em (b) vemos sua trajetória reconstruída em um espaço de estados através de um *embedding* e, em (c), o gráfico de recorrências associado à órbita. 28

Figura 7 – Ilustração de tipologia homogênea. Em (a) vemos uma série temporal estocástica sem autocorrelação (ruído branco), em (b) vemos sua trajetória reconstruída em um espaço de estados através de um *embedding* e, em (c), o gráfico de recorrências associado à órbita. . 29

Figura 8 – Ilustração de tipologia com perturbação. Em (a) vemos uma série temporal estocástica com autocorrelação (ruído browniano), em (b) vemos sua trajetória reconstruída em um espaço de estados através de um *embedding* e, em (c), o gráfico de recorrências associado à órbita. 30

Figura 9 – Acima: o diagrama de órbitas do mapa logístico para o parâmetro nos intervalos 3 até 4. Abaixo: o determinismo do mapa logístico para cada um dos parâmetros. O quantificador foi calculado a partir de um gráfico de recorrências gerado com um ε igual a 0,05. Vemos que há máximos em transições ordem $\rightarrow$ caos e vice-versa. 32

Figura 10 – Acima: diagrama de órbitas do mapa logístico para $r \in [3, 4]$. Abaixo: LAM para o mesmo intervalo do parâmetro. O quantificador foi calculado a partir de um gráfico de recorrências gerado com um ε igual a 0,05. Os picos deste quantificador ocorrem nas transições caos $\rightarrow$ caos. 34

Figura 11 – Ilustração de uma matriz de recorrência evidenciando um microestado e seu respectivo histograma. Em (a), vemos o gráfico de recorrências de uma série de 6 pontos, evidenciando uma matriz de microestados. Imagem adaptada da referência [73]. Em (b), vemos o histograma $P(\cdot)$ dos microestados existentes na figura 11a. 37

Figura 12 – Ilustração de como obter o ε para um RP através da maximização da entropia de microestados. Em (a), vemos o seno utilizado para ilustrar a maximização da entropia de microestados. Em (b), a entropia de microestados da série da figura 12a em função do erro ε. Em (c), vemos o gráfico de recorrências da série da figura 12a gerado com o ε obtido na figura 12b. 38
Figura 13 – Comparação entre os determinismos de um seno com 5 períodos e da componente x do oscilador de Chua obtidos no tempo, em unidades de iteração, onde ambas as séries possuem o mesmo passo temporal. 40
Figura 14 – Representações gráficas de funções seno com a mesma quantidade de pontos ($N = 500$) e diferentes números de oscilações em um mesmo intervalo. Em (a), $\rho = 5$. Em (b), $\rho = 270$. 42
Figura 15 – Senos com o mesmo número de oscilações ($\rho = 10$) e diferentes quantidades de pontos em um mesmo intervalo. Uma caixa está selecionando uma região para ser vista posteriormente. Em (a), $N = 25$. Em (b), $N = 1100$. 42
Figura 16 – Ilustração da importância da resolução em uma série temporal por meio da ênfase da curva subjacente γ, evidenciando a quantidade n de pontos nelae a separação entre estes pontos. Em (a), $n = 2$. Em (b), $n = 55$. Figura ilustrativa com o objetivo de evidenciar a distância entre dois pontos na série, explicitando o fato de que a importância de dois pontos consecutivos diminui a medida em que o número de pontos total da órbita aumenta. Isto ocorre pois o segmento da curva representada por dois pontos consecutivos fica cada vez menos relevante diante do formato completo da curva. 43
Figura 17 – Ilustração dos determinismos de um seno (ponto) e de um ruído branco (estrela) e a suas distância até seus valores desejados. O parâmetro de assimetria é a diferença nos comprimentos de ambas as barras. Neste exemplo, o parâmetro A seria |1 - 0,459 - 0,215| = 0,326. 44
Figura 18 – Ilustração de como se minimiza o parâmetro de assimetria A em função de l_{min}, evidenciando qual o valor de l_{min} escolher. 45
Figura 19 – Valores médios para os l_{min} em função da resolução da série (acima) e, determinismos de um seno com uma oscilação e de um ruído branco (abaixo). Percebe-se que na figura abaixo os determinismos do seno convergem para 1 ao passo em que os determinismos do ruído branco convergem para 0 de forma simétrica com relação às restrições para resoluções maiores. As curvas em preto são os valores médios das quantidades em questão e os sombreados cinzas representam os desvios-padrões. 46
Figura 20 – Valores médios para os l_{min} em função da resolução da série (acima) e, determinismos de um seno com quatro oscilação e de um ruído branco (abaixo). Percebe-se que na figura abaixo os determinismos do seno convergem para 1 ao passo em que os determinismos convergem para 0 de forma simétrica com relação às condições de contorno para resoluções maiores. As curvas em preto são os valores médios das quantidades em questão e os sombreados cinzas representam os desvios-padrões. 47
Figura 21 – Valores médios para os l_{min} em função da resolução da série para $\rho = \{1, 3, 4\}$ e seus respectivos N_c. O gráfico mais acima corresponde a uma oscilação, o do meio à 3 oscilações e o mais abaixo a 4 oscilações. 48

Figura 22 – Derivada de segunda ordem da curva $\langle l_{min} \rangle \times N$ para $\rho = 4$ oscilações. Confirma-se que realmente há um mínimo na segunda derivada no N_c. Figura obtida por meio de diferenciação numérica, o que resulta em perda de pontos no processo, neste caso, pontos finais e iniciais. Diferenciação realizada levando em conta 75 vizinhos para os cálculos dos coeficientes angulares das retas locais. 49
Figura 23 – Extrapolação do modelo logarítmico da curva da figura 20, evidenciando a concordância do modelo com os testes feitos. A curva extrapolada (tracejada) foi modelada apenas com os dados para $N \geq N_c$. Vemos que ela concorda com os testes realizados por meio da minimização do parâmetro de assimetria (pontos). 50
Figura 24 – N_c para algumas quantidades de períodos, mostrando uma relação do tipo lei de potência. Também pode ser possível utilizar um modelo linear, obteríamos, com isso, cerca de 125 pontos por períodos/pseudo-períodos. 50
Figura 25 – Ilustração do seno que será analisado nesta seção. Será composto por 5 períodos e 1000 pontos. 51
Figura 26 – Seno com 5 períodos e seus primeiros 20 pontos destacados no início da série. A cada iteração da série um ponto novo é adicionado, com os quais um novo determinismo é obtido, começando com os primeiros 20 pontos. 52
Figura 27 – Curvas de determinismo calculados ao longo do tempo para o seno com 5 períodos e ruído branco. A curva cinza no fundo serve como guia e ilustra até que ponto da série foi utilizado para os determinismos correspondentes. As curvas em azul são os determinismos de um seno e de um ruído branco obtidos com $l_{min} = 2$. As curvas em preto são os determinismos das mesmas séries, porém, com o método proposto. As duas curvas com valores mais altos correspondem aos senos e as duas com valores mais baixos aos ruídos. 53
Figura 28 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados para o cálculo dos determinismos do seno com 5 períodos no tempo. A curva cinza no fundo serve como guia para ilustrar qual seção da série foi utilizada para o cálculo de l_{min} e ρ. 53
Figura 29 – Seno com 5 períodos e uma janela, evidenciando quais pontos serão necessários para analisar a janela em questão. As janelas serão compostas por 100 pontos e a próxima sempre começará 25 pontos após o ponto inicial da anterior, em outras palavras, a janela se desloca 25 pontos por vez. 54
Figura 30 – Determinismos calculados para cada janela para um seno com 5 períodos e um ruído branco. A curva cinza no fundo serve como guia. Os determinismos estão situados no centro da janela de 100 pontos na qual eles foram calculados. Os pontos em azul são os determinismos de um seno e de um ruído branco obtidos com $l_{min} = 2$. Os pontos em preto são os determinismos das mesmas séries, porém, com o método proposto. As duas curvas com valores mais altos correspondem aos senos e as duas com valores mais baixo aos ruídos. 55
Figura 31 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados para o cálculo dos determinismos do seno com 5 períodos para cada janela. A curva cinza no fundo serve como guia para ilustrar quais as posições das janelas na série foram utilizadas para os cálculos de l_{min} e ρ. Cada ponto está no centro de sua respectiva janela. 55

Figura 32 – Seno com 5 períodos com uma resolução de 20 pontos, evidenciando quais pontos serão necessários para análise e qual a curva subjacente da série em questão. 56
Figura 33 – Determinismos para um seno com 5 períodos e um ruído branco em função de sua resolução. As curvas em azul correspondem aos determinismos obtidos com $l_{min} = 2$. As curvas em preto correspondem aos determinismos obtidos com o critório proposto. As duas curvas com os valores mais altos foram obtidas para os senos. As duas curvas com os valores mais baixos representam os determinismos para os ruídos. 57
Figura 34 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada resolução para o cálculo dos determinismos do seno com 5 períodos. 57
Figura 35 – Componente x do sistema de Lorenz com condições iniciais $(0, 10, 0)$ em sua região caótica mais conhecida. Parâmetros são $(\sigma, \varrho, \beta) = (10, 20, 8/3)$. 58
Figura 36 – Determinismos calculados ao longo do tempo para a componente x do sistema de Lorenz e ruído branco. A curva cinza serve como guia e ilustra até que ponto da série foi utilizado para os determinismos correspondentes. Em azul, os determinismos da componetne x do sistema de Lorenz obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco. 59
Figura 37 – Valores de $\langle l_{min} \rangle$ obtidos via o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada ponto da série para o cálculo dos determinismos da componente x do sistema de Lorenz no tempo. A curva cinza serve como guia para ilustrar qual seção da série foi utilizada para o cálculo de l_{min} e ρ. 59
Figura 38 – Determinismos calculados para cada janela para a componente x do sistema de Lorenz e um ruído branco. A curva cinza ao fundo serve como guia. Os determinismos estão situados no centro da janela de 100 pontos na qual eles foram calculados. Em azul, os determinismos da componente x do sistema de Lorenz obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco, obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco. 60
Figura 39 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada janela para o cálculo dos determinismos da componente x do sistema de Lorenz para cada janela. A curva cinza ao fundo serve como guia para ilustrar quais as posições das janelas na série foram utilizadas para os cálculos de l_{min} e ρ. Cada ponto está no centro de sua respectiva janela. 61
Figura 40 – Determinismos para a x do sistema de Lorenz e um ruído branco obtidos em função de sua resolução. A curva em azul corresponde aos determinismos da componente x do sistema de Lorenz obtidos com $l_{min} = 2$. As curvas em preto correspondem aos determinismos obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco. 62
Figura 41 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de pseudo-períodos (azul) utilizados para cada resolução no cálculo dos determinismos da componente x do sistema de Lorenz. 62
Figura 42 – Componente x do oscilador de Chua com condições iniciais $(7/10, 0, 0)$ em uma região caótica. A série possui 2000 pontos no intervalo temporal de 0 a 75 unidades e possui parâmetros $(\alpha = 156/10, \beta = 2558/100, m_0 = -8/7, m_1 = -5/7)$. 63

Figura 43 – Determinismos calculados ao longo do tempo para a componente x do oscilador de Chua e ruído branco. A curva cinza serve como guia e ilustra até que ponto da série foi utilizado para os determinismos correspondentes. Em azul, os determinismos da componetne x do oscilador de Chua obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco. 64
Figura 44 – Valores de $\langle l_{min} \rangle$ obtidos via o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada ponto da série para o cálculo dos determinismos da componente x do oscilador de Chua no tempo. A curva cinza serve como guia para ilustrar qual seção da série foi utilizada para o cálculo de l_{min} e ρ. 64
Figura 45 – Determinismos calculados para cada janela para a componente x do oscilador de Chua e um ruído branco. A curva cinza ao fundo serve como guia. Os determinismos estão situados no centro da janela de 100 pontos na qual eles foram calculados. Em azul, os determinismos da componente x do oscilador de Chua obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco, obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco. 65
Figura 46 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada janela para o cálculo dos determinismos da componente x do oscilador de Chua para cada janela. A curva cinza ao fundo serve como guia para ilustrar quais as posições das janelas na série foram utilizadas para os cálculos de l_{min} e ρ. Cada ponto está no centro de sua respectiva janela. 66
Figura 47 – Determinismos para a componente x do oscilador de Chua e um ruído branco obtidos em função de sua resolução. A curva em azul corresponde aos determinismos da componente x do oscilador de Chua obtidos com $l_{min} = 2$. As curvas em preto correspondem aos determinismos obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco. 66
Figura 48 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de pseudo-períodos (azul) utilizados para cada resolução no cálculo dos determinismos da componente x do oscilador de Chua. 67
Figura 49 – Processo de Wiener unidimensional com 1000 pontos. 68
Figura 50 – Determinismos calculados ao longo do tempo para o movimento browniano e ruído branco. A curva cinza serve como guia e ilustra até que ponto da série foi utilizado para os determinismos correspondentes. Em azul, os determinismos do movimento browniano obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco. 69
Figura 51 – Valores de $\langle l_{min} \rangle$ obtidos via o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada ponto da série para o cálculo dos determinismos do movimento browniano no tempo. A curva cinza serve como guia para ilustrar qual seção da série foi utilizada para o cálculo de l_{min} e ρ. 69

Figura 52 – Determinismos calculados para cada janela para o movimento browniano e um ruído branco. A curva cinza ao fundo serve como guia. Os determinismos estão situados no centro da janela de 100 pontos na qual eles foram calculados. Em azul, os determinismos do movimento browniano obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco, obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco. 70
Figura 53 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada janela para o cálculo dos determinismos do movimento browniano para cada janela. A curva cinza ao fundo serve como guia para ilustrar quais as posições das janelas na série foram utilizadas para os cálculos de l_{min} e ρ. Cada ponto está no centro de sua respectiva janela. 71
Figura 54 – Determinismos para o movimento broaniano e um ruído branco obtidos em função de sua resolução. A curva em azul corresponde aos determinismos do movimento browniano obtidos com $l_{min} = 2$. As curvas em preto correspondem aos determinismos obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco. . . 71
Figura 55 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de pseudo-períodos (azul) utilizados para cada resolução no cálculo dos determinismos do movimento browniano. 72

Conteúdo

Lista de Figuras **6**

Conteúdo **12**

1 INTRODUÇÃO **14**

2 CONCEITOS DE SÉRIES TEMPORAIS **16**
2.1 Séries Temporais Estacionárias **17**
2.2 Séries Temporais Não Estacionárias **18**
2.3 Componentes de Séries Temporais **20**
2.4 Métodos de Análise de Séries Temporais **21**

3 INTRODUÇÃO AOS SISTEMAS DINÂMICOS E GRÁFICOS DE RECORRÊNCIAS . **23**
3.1 Conceitos de Sistemas Dinâmicos **23**
3.2 Gráficos de Recorrências **25**
3.2.1 Aspectos Macroscópicos dos Gráficos de Recorrências 26
3.2.2 Variações dos Gráficos de Recorrências 29
3.3 Análise de Quantificação de Recorrências **30**
3.3.1 Entropia de Microestados 36
3.4 Critério Atual de Escolha do Parâmetro l_{min} **39**
3.5 Problemas com o Critério Atual **39**

4 MÉTODOS DE DEFINIÇÃO DA DIAGONAL MÍNIMA DE RECORRÊNCIAS **41**
4.1 Sistemas Contínuos **41**
4.2 Sistemas Discretos **45**

5 RESULTADOS E DISCUSSÃO **46**
5.1 Seno com 5 Períodos **51**
5.1.1 Análise no tempo 52
5.1.2 Análise com defasagem 54
5.1.3 Análise na resolução 55
5.2 Componente x do Sistema de Lorenz **57**
5.2.1 Análise no tempo 58
5.2.2 Análise com defasagem 60
5.2.3 Análise na resolução 61
5.3 Componente x do Oscilador de Chua **62**
5.3.1 Análise no tempo 63
5.3.2 Análise com defasagem 65
5.3.3 Análise na resolução 65
5.4 Movimento Browniano **67**
5.4.1 Análise no tempo 68

5.4.2 Análise com defasagem 70
5.4.3 Análise na resolução 71

6 CONCLUSÃO 73

BIBLIOGRAFIA 75

A EXEMPLOS DE SISTEMAS DINÂMICOS 80
A.1 Exemplos de Mapas Determinísticos 80
A.2 Exemplos de Fluxos Determinísticos 80

B PROCESSOS ESTOCÁSTICOS 82

1 Introdução

Observar a natureza é, e sempre foi, um aspecto essencial no processo de geração de conhecimento. Isto ocorre desde a antiguidade, onde as observações da atração e repulsão entre objetos atritados com âmbar na Grécia deram origem ao Eletromagnetismo [1]. Como um exemplo mais recente, podemos citar a observações do dinamarquês Tycho Brahe, que serviram como base para os trabalhos de Johannes Kepler acerca do movimento dos corpos celestes [2]. A obtenção de dados sempre esteve presente no processo de produção científica e tecnológica.

Existem várias maneiras de se obter dados a partir de observações de fenômenos naturais. Pode ser via experimentação em laboratório [3], onde os resultados obtidos são valores numéricos coletados a cada experimento, ou até mesmo através de observações de uma determinada grandeza ao longo do tempo [4]. Quando falamos em observações e experimentações, todavia, não estamos exclusivamente falando de fenômenos naturais, mas também de sociais [5] ou financeiros [6], por exemplo. A coleta de dados está presente em todas as áreas do conhecimento.

Para que seja possível atingir os objetivos de uma pesquisa, entretanto, apenas a obtenção dos dados necessários não é suficiente. Devemos também ser capazes de analisá-los [7] de forma a extrair a maior quantidade de informações possíveis. Novamente, pode-se citar o caso das observações de Tycho Brahe que, mesmo possuindo rigor e precisão incomparáveis para a época, apenas puderam expor peculiares comportamentos elípticos das órbitas celestes após serem analisadas adequadamente por Kepler. Isso nos mostra claramente que é necessário investir também no desenvolvimento das ferramentas de análise dos dados, sejam elas analíticas ou tecnológicas.

Sob este contexto, a Estatística [8, 9] proporciona uma miríade de ferramentas úteis e muito utilizadas na análise de dados, assim como conceitos importantes. Estatística, no entanto, não é a única área do conhecimento que possui interesse na análise dos dados encontrados. Matemática e Física, em particular, também possuem conceitos e teorias próprias, como Teoria de Equações Diferenciais [10, 11], Topologia [12], Sistemas Dinâmicos e Teoria Ergódica [13]. Estes conceitos e teorias também são úteis para extrair informações relevantes de observações e experimentos. Com isso, passamos a olhar para dados de uma forma diferente da Estatística. A ênfase não mais está puramente nos momentos estatísticos e distribuições de probabilidade [14], mas sim, nas propriedades dinâmicas e topológicas do sistema analisado. Métodos baseados na Teoria de Sistemas Dinâmicos estão sendo difundidos cada vez mais entre várias áreas, fazendo com que novas formas de análise de dados sejam continuamente desenvolvidas. Vemos isto nas mais diversas situações, como sistemas financeiros [6], fenômenos físicos [15] ou até mesmo medicina [16]. Em várias situações, os dados obtidos apresentam dependência temporal (*séries temporais*) [17], que são objetos que possuem grande importância acadêmica, financeira, empresarial e industrial [18].

Quando trabalhamos com estas séries, estamos intrinsecamente lidando com sistemas que variam no tempo, também chamados de *sistemas dinâmicos* [19]. Pode-se dizer que o estudo sobre este tipo de sistema, onde sua configuração varia com o tempo, começou no final do século XIX, com Henri Poincaré [20], em seu trabalho sobre *recorrência*. Após esta publicação, este campo de pesquisa se expandiu em larga escala, atraindo para si não apenas Matemáticos, mas também Físicos [7] e Engenheiros [21]. Com isso, várias descobertas foram feitas, tanto sobre grandezas invariantes [22] que podem ser encontradas nestes sistemas, até a maneiras de controlá-los [23]. Em particular, o assunto que abriu este campo de pesquisa é a base de um dos métodos de análise de séries temporais focado em uma abordagem visual [24], o que facilita a

identificação do tipo de sistema estudado. Este método, chamado *gráfico de recorrências* [25], assim como muitos outros, também é utilizado por profissionais de diversas áreas. Como toda ferramenta, entretanto, seja ela tecnológica ou analítica, sempre há espaço para melhorias [26], fazendo com que pesquisas com o objetivo de aprimorar os métodos já existentes estejam sempre sendo desenvolvidas [27].

A estrutura deste trabalho é formada de modo que os capítulos anteriores serão sempre necessários para os próximos. No capítulo 2, intitulado *Conceitos de Séries Temporais*, será abordado o que são estas séries, onde e como são utilizadas. No capítulo 3, denominado *Gráficos de Recorrências*, haverá a apresentação das técnicas e ferramentas fundamentais de gráficos de recorrências. No capítulo 4, chamado de *Métodos*, usaremos os conceitos abordados nos capítulos anteriores para construir a base teórica da proposta deste trabalho. Nele serão mostrados quais os pontos mais importantes e possíveis aplicações da proposta, assim como seus limites de aplicações. No capítulo 5, intitulado de *Resutados*, o método construído no capítulo 4 será utilizado para analisar alguns diferentes sistemas, onde a compreensão sobre o que foi visto nos capítulos anteriores é necessária para a correta interpretação dos resultados. Seguindo, temos o capítulo 6, *Conclusão*, no qual os principais aspectos acerca dos resultados serão abordados. Ao fim, temos dois apêndices, um mostando alguns exemplos de sistemas dinâmicos, e outro nos dando uma breve e simples definição sobre o que são processos estocásticos, ambos conceitos fundamentais que permeiam grande parte deste trabalho.

2 Conceitos de Séries Temporais

Uma série temporal é um conjunto de elementos oriundos de observações ou medições feitas de forma sequencial no tempo, ou seja, uma série temporal é um conjunto munido de um ordenamento [9]. Outra característica que, mesmo não sendo presente em todos os casos, também é importante, é a uniformidade no espaçamento temporal dos elementos. Isso que dizer que as observações são, de forma geral, feitas em intervalos uniformes de tempo, porém, como já mencionado, isso não ocorre para todas as séries [28, 29].

Quando uma série temporal é dependente de um parâmetro contínuo, ou seja, quando as observações são feitas de maneira contínua no tempo ela é chamada de *série temporal contínua* [8]. Séries deste tipo podem ser denotadas por

$$s = \{x(t);\ t \in \mathbb{R}\},$$

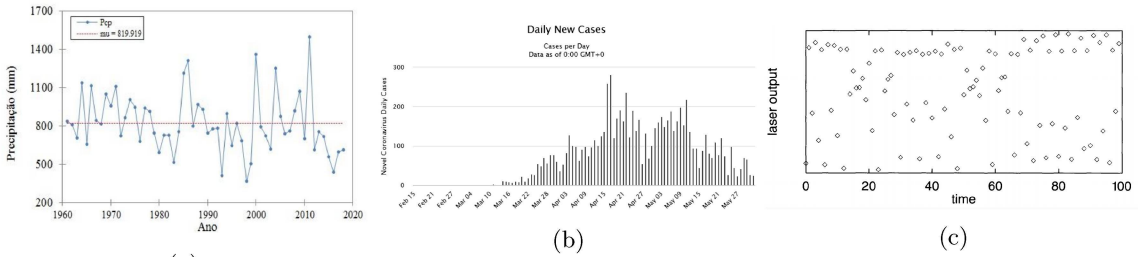
onde o intervalo de tempo em que a série se encontra é subconjunto dos corpo dos reais. Este intervalo é dado por $[t_i, t_f]$. No caso de observações separadas por um intervalo de tempo bem definido, dizemos que as séries resultantes destas observações são *séries temporais discretas* [8]. Séries assim podem ser denotadas por

$$s = \{x_n;\ n \in \mathbb{N}\},$$

onde o índice n é discreto, diferentemente do parâmetro contínuo t mostrado no caso anterior.

Encontrar modelos que sejam capazes de descrever a série estudada é um dos principais objetivos de quem realiza a análise. Há, ainda, objetivos adicionais que podem variar desde modelagem de uma série para *prever* valores futuros [17], encontrar relações entre diferentes variáveis ou apenas analisar diferentes características da série como periodicidade ou estocasticidade [30], por exemplo. Conjuntos de pontos com estas características podem ser obtidos de várias maneiras, por exemplo, via medições do volume pluviométrico em um determinado local durante um determinado período de tempo [31]. Neste caso, medições mensais desta grandeza em alguma cidade geraria uma série temporal *discreta*, pela natureza das observações (mensais). Um outro exemplo seria o conjunto de valores obtidos via observação do número de casos confirmados de COVID-19 em algum país [32], também gerando uma série discreta devido ao fato de que o número de infectados sempre é subconjunto do corpo dos naturais. Como último exemplo, pode-se mencionar a coleção de observações da intensidade de um *laser* em laboratório [7], neste caso, a série pode ser tanto contínua como discreta, dependendo da natureza do parâmetro. Por exemplo, se a intensidade é medida variando-se continuamente sua energia ou ângulo de incidência em alguma superfície, ou se ela é obtida variando outro parâmetro de forma discreta.

A partir de séries temporais pode-se obter quantidades importantes que descrevem o fenômeno subjacente à série, como a média de seus valores e sua variância [14], por exemplo. Com estas quantidades é possível determinar vários aspectos da série em questão [17]. Um exemplo seria a estimativa de crescimento ou decrescimento das variáveis analisadas, neste caso, a média seria, respectivamente, crescente ou decrescente com o tempo. Além das características que definem séries temporais, há aquelas que as *caracterizam*, ou seja, que as separam em grupos [9]. São eles o das séries temporais *estacionárias* e o das *não estacionárias*. Estes dois grupos são muitos distintos e, mesmo que possuam uma nomenclatura similar, são tratados de formas muito diferentes [7]. Eles também são encontrados com frequências muito diferentes, onde as séries



(a) (b) (c)

Figura 1 – Exemplo de séries temporais encontradas em diversas áreas do conhecimento. Na figura (a) é ilustrada a série temporal do volume de chuvas na região de Campina Grande - PB no período de 1961-2018, retirada da referência [31]. Na figura (b) é vista a série temporal do número diário de novas infecções por COVID-19 até 02 de junho de 2020, retirada da referência [32]. Em (c) vemos a série temporal de sucessivas medições da intensidade de um LASER NMR, retirada da referência [7].

não estacionárias são as mais comuns por uma grande margem. Vejamos quais as características de cada um destes grupos.

2.1 Séries Temporais Estacionárias

Seja uma série temporal $s = \{x_1, x_2, ..., x_N\} = \{x_n\}_{n=1}^{N}$ para $n, N \in \mathbb{N}$ cujos elementos seguem uma distribuição de probabilidades $P(\{x_n^*\}_{n=1}^{N^*})$, onde $\{x_n^*\}_{n=1}^{N^*}$ é o conjunto de todos os elementos únicos de s e N^* é a cardinalidade deste conjunto. Para que a série s seja considerada estacionária, primeiro precisamos definir um subconjunto de s com tamanho $h+1$, chamado também de "janela". Em outras palavras, devemos criar subséries de s denotadas por s_i, de forma que cada uma seja dada por

$$s_i = \{x_i, x_{i+1}, ..., x_{i+h}\},$$

onde i é um índice inteiro que vai de 1 até $N - h$. Para que s seja então considerada estacionária, deve ocorrer o seguinte:

Definição 1 (Estacionariedade estrita [9]) *Uma série é* ***estritamente estacionária*** *se o comportamento probabilístico de uma janela* $\{x_{t_1}, ..., x_{t_k}\}$ *for idêntico ao da janela defasada* $\{x_{t_1+h}, ..., x_{t_k+h}\}$*. Isso implica em*

$$Probabilidade(x_{t_1} \leq c_1, ..., x_{t_k} \leq c_k) = Probabilidade(x_{t_1+h} \leq c_1, ..., x_{t_k+h} \leq c_k) \tag{2.1}$$

para todo $k = 1, 2, ...$*, todos os tempos* $t_1, t_2, ...$*, todo número* $c_1, c_2, ...$ *e todos os intervalos* $h = 0, \pm 1, \pm 2, ...$

Em outras palavras, a distribuição de probabilidades $P(s_i)$ dos elementos em um intervalo começando em x_i deve ser igual a distribuição de probabilidades de um intervalo começando em x_j. Como consequência, $P(s_i) = P(s_j),\ \forall i, j \in [1, N-h]$. Temos então que as propriedades estatísticas da série não podem depender do intervalo observado [8, 9].

Definamos agora a função *esperança* $\mathbb{E}(x)$ [9, 14] de uma variável aleatória discreta da seguinte maneira:

$$\mathbb{E}(x) = \sum_{n=1}^{N} x_n P(x_n).$$

A esperança é, então, a soma de cada elemento da série multiplicado pela probabilidade de ser encontrado. Porém, no caso contínuo, isso fica um pouco mais complicado. Para definir uma função média de uma série temporal, nos casos em que é possível, primeiro devemos definir uma função *densidade de probabilidade* [33], sendo ela dada por

$$p(x) = \frac{dP(x)}{dx}.$$

Com isso, caso a função $p(x)$ exista, podemos definir a esperança de uma variável aleatória contínua de forma que [9, 14]

$$\mathbb{E}(x) = \int_{t_i}^{t_f} xp(x)dx.$$

Uma outra função de suma importância na análise de séries temporais é a sua *variância*, denotada por $\mathrm{Var}(x)$ [9, 14], cuja relação com a esperança é dada por

$$\mathrm{Var}(x) = \mathbb{E}(x^2) - \mathbb{E}(x)^2.$$

Para que a série seja estacionária vimos que suas propriedades estatísticas não devem depender dos intervalos analisados. Este é um caso de *estacionariedade estrita*, porém, nem sempre é factível ou simples de se verificar estacionariedade desta forma. Para contornar este problema alguns critérios menos rigorosos foram definidos, são eles [9]

- A média (valor esperado/esperança) da série não pode depender do tempo;
- A variância da série não pode depender do tempo.

Quando uma série apenas apresenta as propriedades mencionadas a pouco, diz-se que ela possui uma *estacionariedade fraca*. Com isso, vemos que as séries dos exemplos da figura 1 não são estacionárias, pelo menos em uma análise mais superficial. Isso ocorre pois na figura 1a os valores da série tendem a se afastar cada vez mais de sua média ao passar do tempo, implicando em uma variância dependente do tempo. Na figura 1b, vemos que para uma janela no início da série a média será menor do que se esta janela estivesse mais no centro da abscissa. No terceiro caso, mostrado na figura 1c vemos que existem certos intervalos com uma densidade maior de pontos, fazendo com que a média seja dependente do intervalo escolhido.

Para exemplificar o que foi dito pode-se utilizar uma série estocástica sem correlação [8, 9, 33] e calcular sua média em diferentes intervalos. Desta forma é possível verificar que sua média de fato não depende do intervalo. Vemos isso na figura 2, onde a série foi dividida em dez intervalos, nos quais foram calculadas as médias dos elementos de dentro de cada intervalo. Com isso, foi constatado que estas médias estão todas próximas da média global da série, que deveria ser 0. Isto é indício de que ela é estacionária.

2.2 Séries Temporais Não Estacionárias

Este grupo representa a maior parcela das séries encontradas de forma experimental ou observacional [7]. Neste caso as séries não possuem uma constância em suas médias e variâncias, dificultando suas análises, tal como pode ser observado em gráficos de flutuações de preço de uma determinada ação na bolsa de valores [34].

Uma equação diferencial [10, 11] simples pode gerar uma série temporal, assim como uma componente de um sistema de equações também pode ser considerada uma série temporal. Como exemplo, a equação

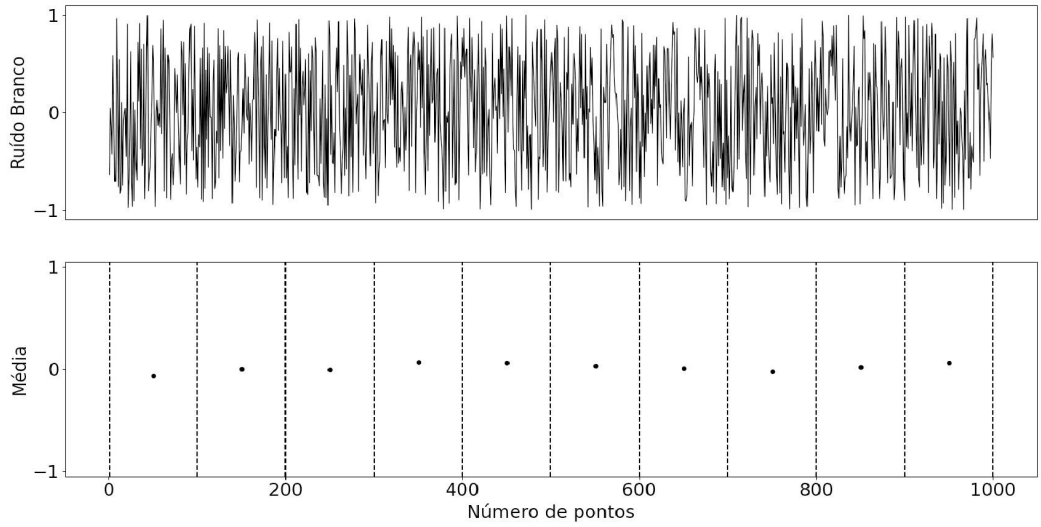


Figura 2 – Um ruído branco uniformemente distribuído (acima) dividido em diferentes intervalos e as esperanças $\mathbb{E}(x)$ da série dentro de seus respectivos intervalos (abaixo). Como a esperança é constante no tempo, esta série é estacionária.

$$\ddot{x} + x = 0 \tag{2.2}$$

resulta em uma função senoidal, que é uma função periódica. Ao analisar a série gerada através da equação 2.2 contaminada com uma componente estocástica [33], vemos que a série resultante não é estacionária. Quando funções periódicas são submetidas à análise de estacionariedade como apresentada por [9], vemos de início que podemos ter um caso não estacionário. Isso ocorre pois os elementos contidos em janelas centradas nos valores mais altos possuirão uma média mais alta do que os elementos comportados nos intervalos que contém os *vales*.

Percebe-se claramente que as médias da figura 3 variam ao transladar o intervalo no qual elas foram obtidas. Séries temporais não estacionárias não possuem as mesmas restrições das estacionárias, e podem apresentar eventos recorrentes denominado por *sazonalidae*. Elas podem possuir uma *tendência*, que é a propensão em crescer ou decrescer seus valores conforme o tempo passa e podem ainda possuir alguma espécie de componente aleatória. Esses conceitos podem ser vistos ao decompor a série como

$$s_n = T_n + C_n + \epsilon_n,$$

chamada de decomposição *aditiva* [9]. Podemos ainda decompor a série de forma *multiplicativa*, como

$$s_n = T_n C_n \epsilon_n,$$

onde o lado direito apresenta as componentes de *tendência*, *cíclica* (ou sazonal) e o *ruído*, respectivamente. Estas componentes são fatores importantes na hora de caracterizar uma série temporal, pois estas propriedades afetam os possíveis modelos ou técnicas que possam ser empregados em sua análise.

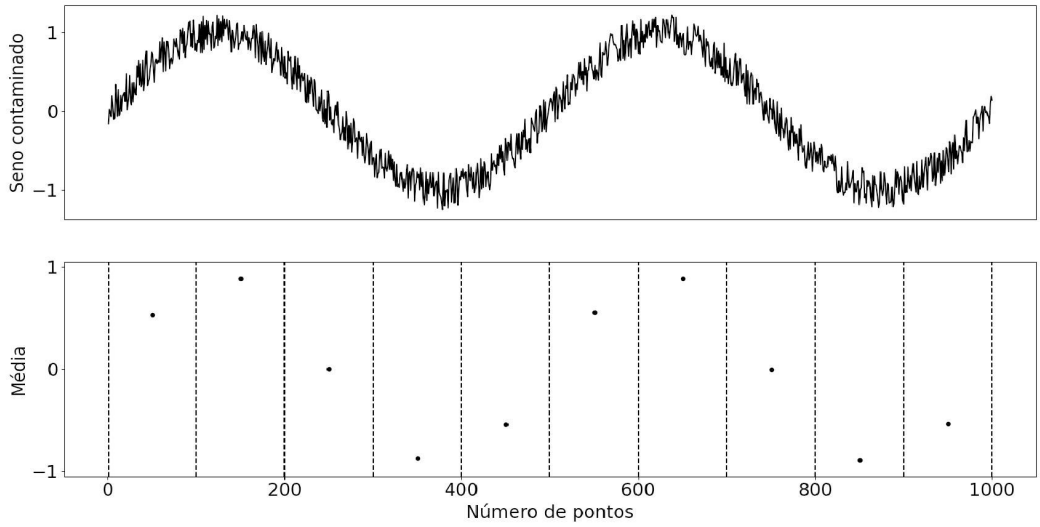


Figura 3 – Um seno contaminado com 25% de ruído branco (acima) dividido em diferentes intervalos e a esperança $\mathbb{E}(x)$ da série dentro de seus respectivos intervalos (abaixo). Há uma dependência da esperança com o tempo, fazendo com que esta não seja uma série estacionária.

2.3 Componentes de Séries Temporais

- **Tendência [8, 9, 17, 18]**: Conhecida também por *trend*, é a componente causadora de crescimento ou decrescimento em uma série temporal. Em outras palavras, esta componente descreve se a série analisada está crescendo ou decrescendo ao longo do tempo. A tendência pode ou não ser linear, porém, sempre que uma série apresenta alguma tendência, a série não será estacionária. Isso acontece pois a média de subconjuntos da série seguirão aumentando ou diminuindo conforme o tipo de tendência, o que viola até mesmo a condição de estacionariedade fraca. É possível extrair esta componente quando se sabe seu formato através da subtração ou divisão do modelo da tendência pela série (depende da decomposição feita).

- **Sazonalidade [8, 9, 17, 18]**: É o comportamento oscilatório da série em curto prazo, ou seja, sua periodicidade. Séries sazonais, por definição não podem ser estacionárias pois suas médias dependem de qual subintervalo está sendo considerado para seu cálculo (depende da defasagem). Entretanto, existem maneiras de se extrair a componente sazonal de uma série, de modo similar ao que é feito com a tendência. Neste caso, o processo leva a perda de uma quantidade significativa de pontos da série, correspondendo aos pontos constituintes do primeiro período.

- **Ciclicidade [8, 9, 17, 18]**: Similar à sazonalidade, a ciclicidade também contém informação acerca das oscilações encontradas nas séries temporais. Entretanto, diferentemente da anterior, a ciclicidade apresenta oscilações em escalas maiores e não precisa ser regular. Isto significa que uma série pode apresentar ciclos de diferentes durações. Um ciclo pode ainda conter várias oscilações menores oriundas da sazonalidade.

- **Ruído [8, 9, 17, 18]**: Também chamado de *componente aleatória*, é o que resta da série quando se remove as outras componentes. Um dos ruídos mais comuns é chamado de *ruído branco uniformemente*

distribuído e é descrito por uma distribuição de probabilidade de média zero e variância finita, onde seus valores máximos e mínimos são finitos. Cada um dos valores possíveis possuem a mesma probabilidade de ocorrência. Um ruído branco uniformemente distribuído é uma série estacionária onde cada um de seus elementos é independente dos demais. Neste trabalho, por simplicidade, será chamado apenas de *ruído branco.*

2.4 Métodos de Análise de Séries Temporais

Existem inúmeras maneiras de se analisar séries temporais. Vários destes métodos são vastamente conhecidos e ensinados logo no início da formação dos pesquisadores. Regressão linear é um exemplo de método ensinado em disciplinas de Estatística para uma vasta gama de cursos. Alguns métodos lineares muito utilizados são:

- Regressão Linear [35];
- Análise espectral [36, 37];
- Filtros lineares [38];
- Modelos preditivos (ARMA, ARIMA, SARIMA, etc.) [17].

Estes métodos, entretanto, podem não ser robustos o suficiente para analisar todos os dados encontrados, principalmente no caso de o sistema observado ser não linear. Isto dificulta a análise e pode introduzir dinâmicas complicadas como caos [7, 19, 21] ou até mesmo hiper-caos [39]. Nos casos em que se faz necessário a implementação de métodos não lineares, normalmente se inicia com a reconstrução de espaço de estados por *coordenadas de atraso* (*delay coordinates*) [21, 40]. Esta etapa é chamada de *embedding* [21, 41, 42] do sistema em um espaço m-dimensional. Isso é feito pois estes métodos são em sua maioria baseados na Teoria de Sistemas Dinâmicos, impondo que o comportamento do sistema a ser estudado no espaço de estados carrega informação acerca de sua dinâmica. As coordenadas de atraso são obtidas como

$$\vec{x}_i = \sum_{j=1}^{m} s_{i+(j-1)\tau} \vec{e}_j,$$

onde $\vec{x}_i$ é o i-ésimo ponto da órbita no espaço de estados reconstruído, m é a dimensão de *embedding*, $s_{i+(j-1)}$ são os valores da série temporal, τ é o *delay* utilizado na amostragem dos valores da série e $\vec{e}_j$ são versores do espaço reconstruído.

Existe mais de uma forma de se obter um valor adequado para o *delay*, porém, um dos mais utilizados é o valor correspondente ao primeiro mínimo da informação mútua [43] dá série. A informação mútua $I(X, Y)$ de duas variáveis X e Y é dada por

$$I(X, Y) = H(X) + H(Y) - H(X, Y),$$

na qual $H(X)$ e $H(Y)$ são as entropias de informação da variável X e $H(X, Y)$ é a entropia conjunta das duas variáveis. A entropia de informação é definida como

$$H(X) = -\sum_{x \in X} p(x) \log(p(x)), \tag{2.3}$$

com $p(x)$ sendo a probabilidade de se encontrar a variável nos estados X. Para a entropia conjunta, a expressão é a mesma, porém, $p(x)$ é substituída por $p(x, y)$, que é a probabilidade conjunta destas duas

variáveis. A informação mútua, então, nos diz a quantidade de informação que é obtida para X ao observar Y, pode-se dizer que é uma medida de dependência entre X e Y. Ao minimizarmos esta quantidade em função de um atraso temporal, e para valores de uma mesma série, obtemos o atraso temporal (chamado de *time delay*) que nos possibilita gerar vetores de base do espaço de estados reconstruído que sejam o mais independentes quanto possível.

A reconstrução de um espaço de estados é necessária para várias formas de análise de dados, como por exemplo:

- Obtenção de expoentes de Lyapunov [44];
- Determinação da existência de atratores e suas propriedades (bacias, fronteiras, dimensões, etc.);
- Seções de Poincaré [21];
- Cálculo de entropia de Kolmogorov-Sinai [22];
- Gráficos de recorrências [25, 24, 45, 46].

Para o cálculo de dimensão do espaço onde se quer realizar o *embedding* do sistema existem teoremas [42, 41] que nos permitem encontrar uma maneira que garanta um *difeomorfismo*[1] entre o espaço reconstruído e o espaço original. Para isso, se utiliza o conceito de *dimensão de correlação* [7, 21]. A dimensionalidade m do espaço reconstruído pode ser obtida de modo que

$$m \geq 2D_2 + 1,$$

onde D_2 representa a dimensão de correlação [47, 48, 49] da órbita do sistema. A dimensão de correlação, muitas vezes chamada de "dimensão fractal", é um dos possíveis cálculos para a dimensionalidade de um conjunto de pontos. No caso de um sistema dinâmico, o conjunto em questão é a união dos pontos da órbita do sistema. Esta dimensão é obtida através da expressão

$$D_2 = \lim_{r \to 0} \frac{\log C(r)}{\log(r)}. \tag{2.4}$$

Na expressão 2.4 o termo $C(r)$ é chamado de *integral de correlação* e é definido por [50]

$$C(r) \equiv \lim_{r \to 0} \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \Theta(r - \left\| \vec{x}_i - \vec{x}_j \right\|), \tag{2.5}$$

na qual $\Theta(\cdot)$ é a *função de Heaviside* (função degrau), N é o número de pontos da órbita e r é um parâmetro que descreve se os pontos em questão estão dentro da mesma região ou não.

Dentre os possíveis métodos de análise de séries com base em Teoria de Sistemas Dinâmicos, neste trabalho será abordado com mais detalhes a construção e utilização de gráficos de recorrências.

[1] Uma bijeção entre varidades suaves em que tanto ela própria quando sua inversa são também suaves.

3 Introdução aos Sistemas Dinâmicos e Gráficos de Recorrências

3.1 Conceitos de Sistemas Dinâmicos

Sistemas dinâmicos [19] são famílias de transformações f que levam um vetor $x \in \mathcal{M}$ em sua imagem $f(x) \in \mathcal{M}$, e assim sucessivamente. Uma maneira mais simples de explicar o que são sistemas dinâmicos seria dizer que eles são conjuntos de valores que evoluem no tempo, gerando uma trajetória em um espaço de estados. Estes sistemas podem aparecer em diversas åreas do conhecimento como Física [21], Engenharia [19], Economia [51] e Medicina [52], para citar alguns exemplos. Podemos mencionar quatro categorias de sistemas dinâmicos, sendo elas:

- Determinísticos com tempo discreto;
- Determinísticos com tempo contínuo;
- Probabilísticos com tempo discreto;
- Probabilísticos com tempo contínuo.

De forma geral, podemos dizer que um sistema dinamico é um objeto que evolui no tempo e possui como lei de composição [13]

$$(f^{\tau_1} \circ f^{\tau_2})(x) = f^{\tau_1}(f^{\tau_2}(x)) = f^{\tau_1+\tau_2}(x), \tag{3.1}$$

na qual $\tau_1, \tau_2 \in \mathcal{T}$ e $f^0(x) = x$. Com isso, essencialmente temos dois casos:

1. $\mathcal{T} = \mathbb{N}$;
2. $\mathcal{T} = \mathbb{R}$,

sendo que se o primeiro caso é verdadeiro, o sistema é chamado de *mapa*, e que se o segundo caso é verdadeiro é chamado de *fluxo*. Em todos os casos, o conjunto de pontos

$$\mathcal{O}(x) = \bigcup_{\tau \in \mathcal{T}} f^{\tau}(x) \tag{3.2}$$

denota a *órbita* do sistema em seu espaço de estados, a qual precisa existir e ser única para cada condição inicial $f^0(x) = x_0$, *para sistemas dinâmicos determinísticos*. Isso significa que para esta classe de sistemas não podem haver intersecções da órbita com ela mesma em nenhum ponto do espaço de estados [21]. Para sistemas probabilísticos, entretanto, isso não é uma exigência.

No espaço de estados, que é o espaço gerado pelas componentes de x (que não é necessariamente unidimensional), podem ser encontrados em alguns casos, pontos, ou conjuntos de pontos, que atraem ou repelem as trajetórias do sistema. Estes pontos podem ser [7]

- **Pontos fixos [19, 21]**: Conjunto de apenas um ponto x^* que atrai ou repele as trajetórias quando a condição inicial x_0 é diferente de x^*, que mapeia a si mesmo quando $x_0 = x^*$. Desta forma, $f^\tau(x^*) = x^*$, $\forall\tau \in \mathcal{T}$. Quando o ponto fixo *atrai* as órbitas ele é chamado de *estável* e quando *repele* é chamado de *instável*. Eles representam estados de equilíbrio do sistema. Há a possibilidade de que um mesmo ponto fixo seja estável por uma direção e instável por outra, sendo chamado de ponto fixo *parcialmente estável*. Pontos fixos dominam a dinâmica de sistemas unidimensionais.

- **Ciclos limite [19, 21]**: São conjuntos de pontos que formam uma curva fechada no espaço de estados que atraem ou repelem assintoticamente as órbitas do sistema. A órbita $\mathcal{O}(x)$ se aproxima ou se afasta destes pontos de forma a circundar este ciclo, sendo um conjunto de estados oscilatórios do sistema em equilíbrio. São encontrados em sistemas dinâmicos com duas ou mais dimensões.

- **Atratores e atratores estranhos**[1] **[19, 21]**: São conjuntos de pontos que atraem o sistema e podem ser encontrados em sistemas a partir de qualquer dimensão, porém, quando o sistema possui três dimensões ou mais podemos encontrar atratores dito *estranhos*, o que significa que são um conjunto de pontos em uma determinada região do espaço de estados que possui geometria fractal, como dimensionalidade fracionária e autossimilaridade. São encontrados em sistemas caóticos e uma vez que a órbita do sistema entra em um atrator ela dificilmente sairá e será confinada nele.

Um mapa, ou sistema dinâmico com tempo discreto, pode ser representado pela seguinte expressão [21]:

$$x_{n+1} = \mathbf{M}[x_n], \tag{3.3}$$

onde $\mathbf{M}$ é uma matriz que representa a relação entre as componentes de x nos tempos n e $(n+1)$. É possível que a matriz $\mathbf{M}$ seja linear, ou seja, que suas linhas e colunas apenas contenham termos lineares nas variáveis e sem produtos entre elas, ou pode ser que ela seja não linear. Dinâmicas lineares são mais fáceis de se analisar. No caso de a matriz $\mathbf{M}$ ser invertível, o sistema pode ser analisado de forma simples, porém, quando $\mathbf{M}$ não é invertível, é possível encontrar *caos* mesmo em mapas unidimensionais.

No caso de fluxos, ou sistemas dinâmicos com tempo contínuo, podemos representá-los por uma família de curvas em um espaço vetorial, usualmente gerado por uma equação diferencial ordinária da forma [21]:

$$\frac{dx}{dt} = F(x), \tag{3.4}$$

onde x não é necessariamente unidimensional e F denota o sistema de equações diferenciais de primeira ordem. Para esta classe de sistemas, quando temos um fluxo não linear com dimensão maior ou igual a três, é possível encontrar comportamento caótico.

Para ambas as classes de sistemas, tanto para mapas quanto para fluxos, existem pontos que fazem com que a órbita volte a visitá-los a partir de uma quantidade de tempo, estes são chamados de pontos *periódicos*. As órbitas associadas a eles são chamadas de *órbitas periódicas*. Se aplicarmos sucessivamente a transformação f em um ponto x_p e após um período de tempo igual a p obtermos o próprio ponto x_p, dizemos que este ponto/órbita possui período p [19, 21]. É possível mostrar que se um sistema possui uma órbita periódica de período 3, então ele possui uma órbita periódica para todos os períodos possíveis, ou seja, este sistema será caótico.

[1] Atratores *estranhos* são conjuntos de pontos no espaço de estados que formam um objeto com dimensão não inteira, denominados de *fractais*. Este conjunto de pontos atrai a órbita do sistema, sendo que a ela fica então confinada dentro deste conunto.

3.2 Gráficos de Recorrências

Proposto por Eckmann *et al.* em 1987 [24], este método pode ser utilizado para analisar tanto séries temporais estacionárias como não estacionárias. Ele se baseia no Teorema da Recorrência de Poincaré [20], enunciado em 1890, que diz:

Teorema 1 (Teorema de Recorrências de Poincaré) *Se um fluxo preserva seu volume*[2] *e possui apenas órbitas limitadas, então, para cada conjunto aberto no espaço onde a dinâmica ocorre, qualquer órbita intercepta este conjunto infinitas vezes.*

Isto significa que, para cada fluxo que preserva medida, cada região do espaço onde o fluxo ocorre será visitado infinitas vezes, desde que esperemos um tempo suficientemente longo.

Com isso, conceitos da Teoria de Sistemas Dinâmicos também podem, e são, utilizados no lugar de apenas Estatística no âmbito de análise de dados [7, 23]. Este método é utilizado por profissionais de diversas áreas, como por exemplo, economia, na análise do mercado financeiro [6], em telecomunicações, para analisar sinais de áudio [53], ou até mesmo em medicina, analisando taxas de batimentos cardíacos a fim de detectar anomalias [16]. Ele consiste em quantificar as recorrências de uma órbita em um espaço de estados [25, 46], que pode ser reconstruído através de coordenadas de atraso ou não e representá-las na forma de uma imagem. Estes gráficos são representações visuais de uma matriz R preenchida por valores $\{0, 1\}$. Esta matriz recebe o nome de *matriz de recorrência* [24] cujas entradas têm a forma

$$R_{ij} := \Theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|), \tag{3.5}$$

na qual $\Theta(\cdot)$ é a função degrau, ou *função de Heaviside*, $\vec{x}_i$ é o i-ésimo vetor que descreve a órbita do sistema no espaço de estados. Em outras palavras, $\vec{x}_i$ é o i-ésimo ponto da trajetória no espaço de estados (lembrando que em um sistema dinâmico o ordenamento dos elementos é de suma importância). Na expressão 3.5, ε é a maior distância possível que dois pontos podem ter entre si para que sejam considerados "recorrentes".

O conceito de recorrência aparece quando consideramos um volume m-dimensional de tamanho fixo (para um dado ε) centrado em um ponto $\vec{x}_i$ [25]. Dizemos que este ponto *recorre* com todos os outros pontos $\vec{x}_j$ que estiverem a uma distância de $\vec{x}_i$ menor do que ε. Em outras palavras, todos os pontos dentro deste volume são pontos *recorrentes* entre si. A expressão 3.5 nos diz então que o valor "1" é colocado na matriz R em todas as posições (i, j) correspondentes aos pontos contidos no mesmo ε-volume [45]. Uma consequência disto é que a diagonal principal sempre conterá os valores "1", pois $\|\vec{x}_i - \vec{x}_i\| = 0 \leq \varepsilon$, $\forall i \in \mathbb{N}$. O restante dos elementos da matriz R são nulas. O gráfico de recorrências (RP), como imagem, é então formado através do mapeamento os valores "1" por *pixels* pretos (por convenção) e os valores "0" por *pixels* brancos. O gráfico de recorrências, por conseguinte, nada mais é do que a representação visual dos valores de uma matriz.

A partir destes gráficos é possível perceber que, para cada tipo de dinâmica do sistema, existem estruturas bem definidas e, apenas analisando visualmente estas estruturas, podemos dizer se o sistema é periódico, caótico ou estocástico [25, 45, 46]. É possível também verificar se ele possui transições na dinâmica, tendência ou até mesmo o valor dos períodos e pseudo-períodos[3] pode ser encontrado. Estas estruturas que aparecem nos gráficos de recorrências podem ser divididas em duas categorias, são elas as *tipologias* (aspecto geral) e *textutas* (estruturas internas) [25]. Falaremos primeiro das diferentes tipologias encontradas nestes gráficos.

[2] Também chamado de *medida*, usualmente é medida de *Lebesgue* μ, porém, esta não é a única.

[3] Pseudo-períodos correspondem ao tempo em que uma série leva para repetir de forma aproximada seu comportamento, similar ao conceito de período. Entretanto, este termo é válido para séries aperiódicas e pode ser estimado através de métodos como periodogramas ou funções de autocorrelação, por exemplo.

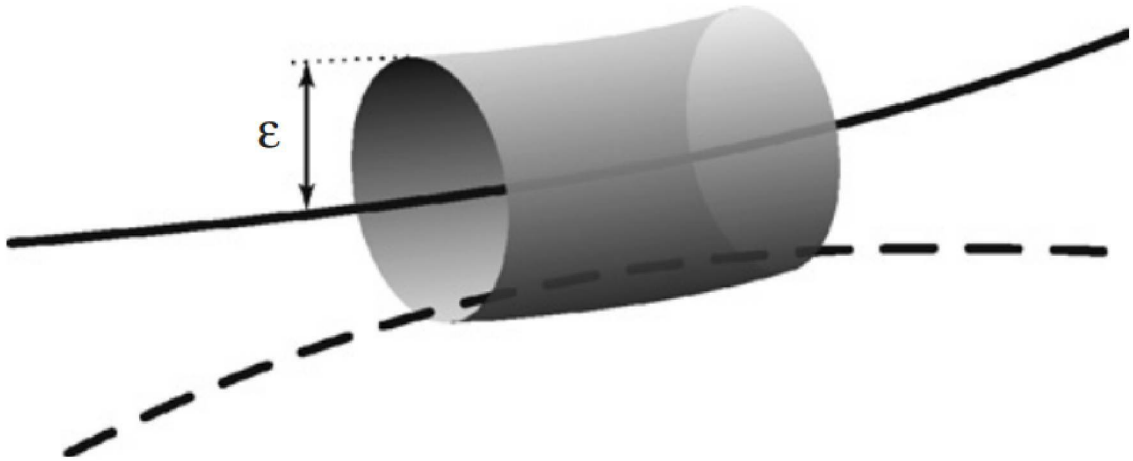


Figura 4 – Ilustração de um volume dependente de ε em um espaço de estados mostrando um segmento de órbita que será considerado recorrente. Figura retirada da referência [25].

A tipologia de um gráfico de recorrências nos dá informação sobre a natureza dinâmica do sistema, ou seja, se ele é periódico, caótico ou estocástico, refletindo esta natureza no aspecto macroscópico destes gráficos.

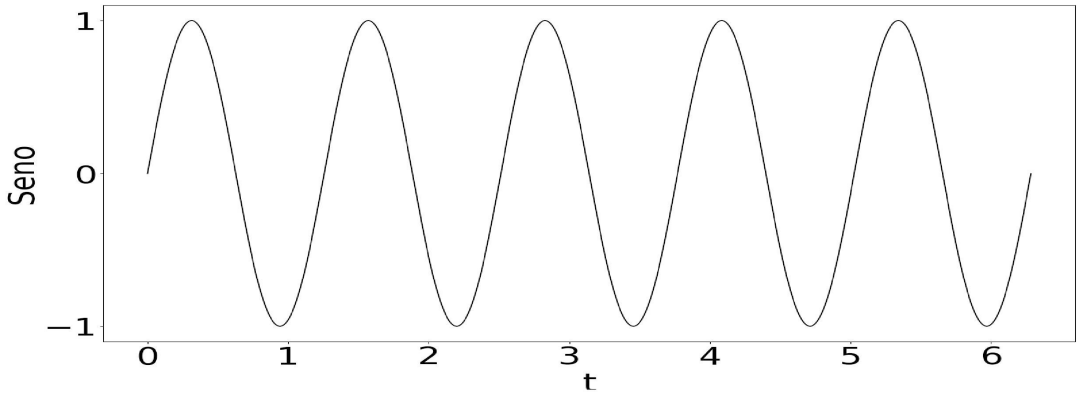
3.2.1 Aspectos Macroscópicos dos Gráficos de Recorrências

Regular [25, 54]: Gráficos gerados a partir de séries regulares, ou seja, que possuem periodicidade, tendem a apresentar uma maior quantidade de linhas diagonais, onde estas linhas, em geral, são mais longas. Vejamos agora o motivo disso ocorrer. Quando uma órbita possui dois pontos, digamos, $\vec{x}_i$ e $\vec{x}_j$, dentro de um mesmo ε-volume[4], um *pixel* no gráfico de recorrências é colorido de preto na posição (i, j), portanto, se os próximos pontos $\vec{x}_{i+1}$ e $\vec{x}_{j+1}$ também recorrerem entre si, um *pixel* na posição (i+1, j+1) também será colorido de preto. Isso continuará ocorrendo para todos os índices (i+k, j+k) cujos pontos correspondentes recorrem, gerando assim, *pixels* pretos em diagonal. Isto significa que um segmento inteiro da trajetória recorreu, ou seja, que existe uma subsequência da órbita se repetindo pelo menos uma vez. Por este motivo estas estruturas são mais facilmente encontradas em gráficos de recorrências de sistemas periódicos, onde há por definição repetições na trajetória. A figura 5 mostra uma série temporal senoidal, sua órbita gerada a partir de um *embedding* e o gráfico de recorrências correspondente, evidenciando as longas linhas diagonais.

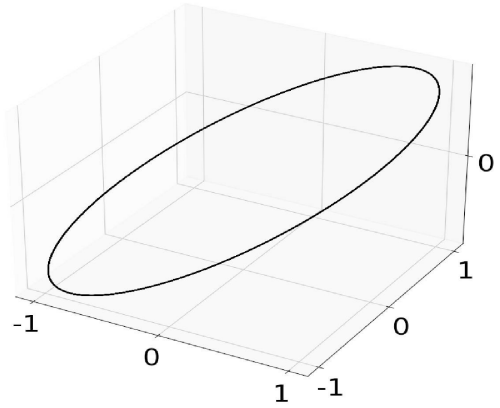
É possível determinar o valor dos períodos, em unidades de iteração, da série estudada a partir do espaçamento entre duas linhas diagonais. No exemplo utilizado (Figura 5) ela corresponde a cerca de 220 pontos. A série em questão possui 5 períodos no intervalo $[0, 2\pi]$ e 1000 pontos, portanto, percebemos que a distância obtida equivale à um intervalo de duração $1.32 \approx \frac{2\pi}{5}$ unidades de tempo, que é o valor correto para o período da série.

Caótico [25, 54]: Quando o sistema estudado possui uma dinâmica caótica, o gráfico de recorrências correspondente irá refletir este comportamento de forma a deixá-la evidente. Gráficos desta natureza apresentam estruturas muito similares àqueles gerados a partir de sistemas periódicos, pois também contém uma

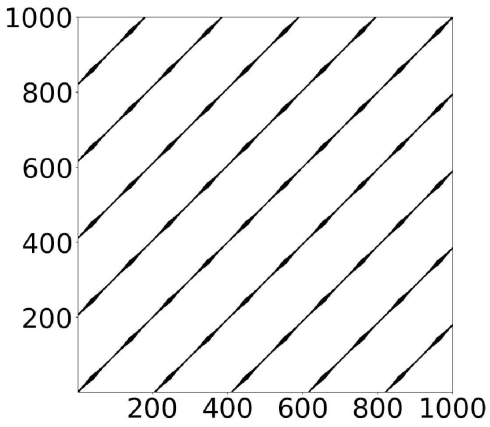
[4] Uma bola m-dimensional de raio ε, usualmente denotada por $B_\varepsilon(\vec{x})$, na qual m é tomado como a dimensão do espaço em que ela está inserida.



(a)



(b)



(c)

Figura 5 – Ilustração de tipologia regular. Em (a) vemos uma série temporal periódica senoidal com 5 períodos, em (b) vemos sua trajetória reconstruída em um espaço de estados através de um *embedding* e, em (c), o gráfico de recorrências associado à órbita.

quantidade significativa de linhas diagonais. Entretanto, a maior diferença se encontra na distribuição dos comprimentos destas linhas. Em contraste com gráficos de séries periódicas, que possuem as linhas diagonais com tamanhos menos diversificados, os gráficos de recorrências de séries caóticas consistem em várias linhas diagonais de diferentes tamanhos, cada uma associada a alguma frequência de oscilação. Como séries caóticas possuem comportamentos menos previsíveis e menos repetitivos do que séries periódicas, elas oscilam, de modo geral, com um espectro maior de frequências. Isso resulta em sequências mais distintas de pontos recorrentes, culminando assim em uma distribuição com mais linhas diagonais possíveis. De forma simplificada, cada linha diagonal está relacionada a alguma frequência de oscilação, como séries caóticas oscilam em várias frequências, seus gráficos de recorrências apresentam linhas diagonais mais diversificadas. Como exemplo vemos (figura 6) a componente x do sistema de Rössler [55], dado pelas equações

$$\begin{aligned} \dot{x} &= -y - z, \\ \dot{y} &= x + ay, \\ \dot{z} &= b + z(x - c), \end{aligned} \tag{3.6}$$

em uma região caótica, sua órbita reconstruída e seu respectivo gráfico de recorrências. Assim como no caso

periódico, é possível estimar visualmente os pseudo-períodos de uma série caótica.

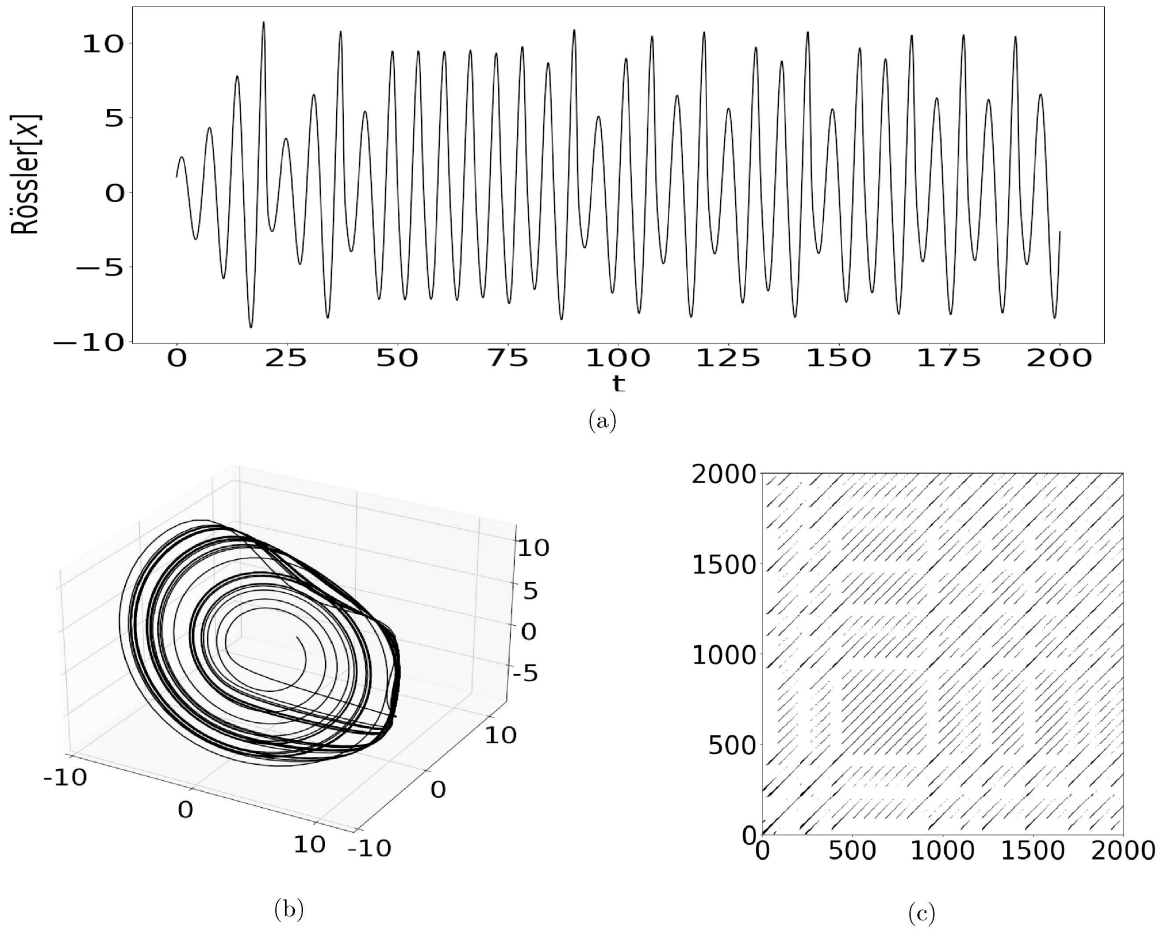


Figura 6 – Ilustração de tipologia caótica. Em (a) vemos uma série temporal caótica correspondente à componente x do sistema de Rössler, em (b) vemos sua trajetória reconstruída em um espaço de estados através de um *embedding* e, em (c), o gráfico de recorrências associado à órbita.

Homogêneo [25, 54]: Ao aplicar este método na análise de séries estocásticas obtemos outras duas tipologias, uma relacionada aos sistemas sem autocorrelação (homogênea) e outra correspondente a sistemas com algum grau de correlação, embora ainda estocásticos. Falaremos primeiro sobre sistemas não autocorrelacionados. Quando gera-se uma série estocástica sem autocorrelação como um ruído branco, sabe-se que pontos sequencialmente recorrentes podem ocorrer com probabilidade não nula, porém, isto é extremamente improvável. A probabilidade de ocorrência segmentos recorrentes é muito pequena, isso se reflete no fato que valores pequenos de ε geram gráficos compostos majoritariamente por pontos isolados. Como consequência, as linhas diagonais em um gráfico de recorrências gerado a partir de uma série destas serão mais curtas e aparecerão em menor número. Séries como esta são usualmente estacionárias, o que faz com que haja recorrências, mesmo que esporádicas, por toda a extensão do gráfico. Para exemplificar, vemos na figura 7 uma série de números pseudo-aleatórios[5] gerados por computador, sua órbita reconstruída e seu

[5] Uma sequência numérica gerada artificialmente, não sendo *de facto* aleatória.

respectivo gráfico de recorrências.

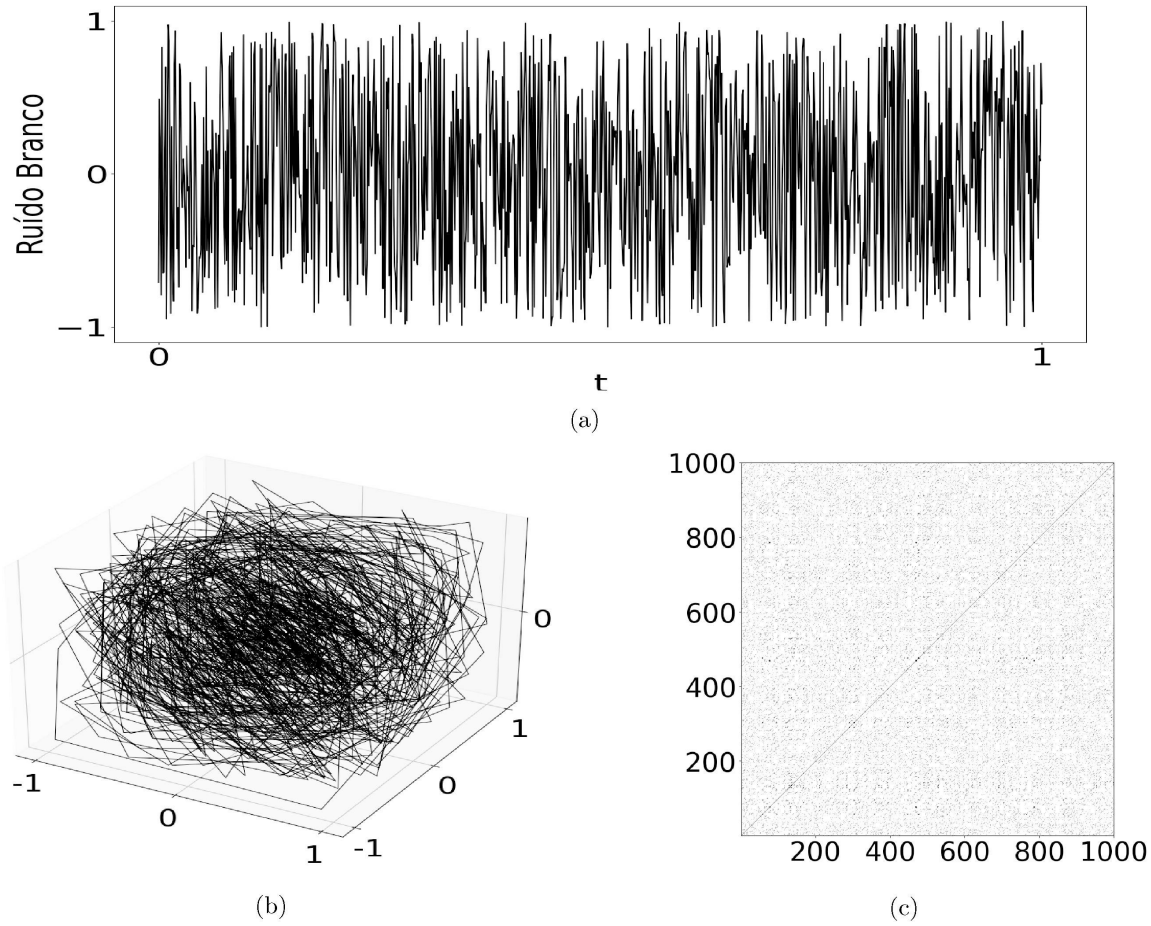
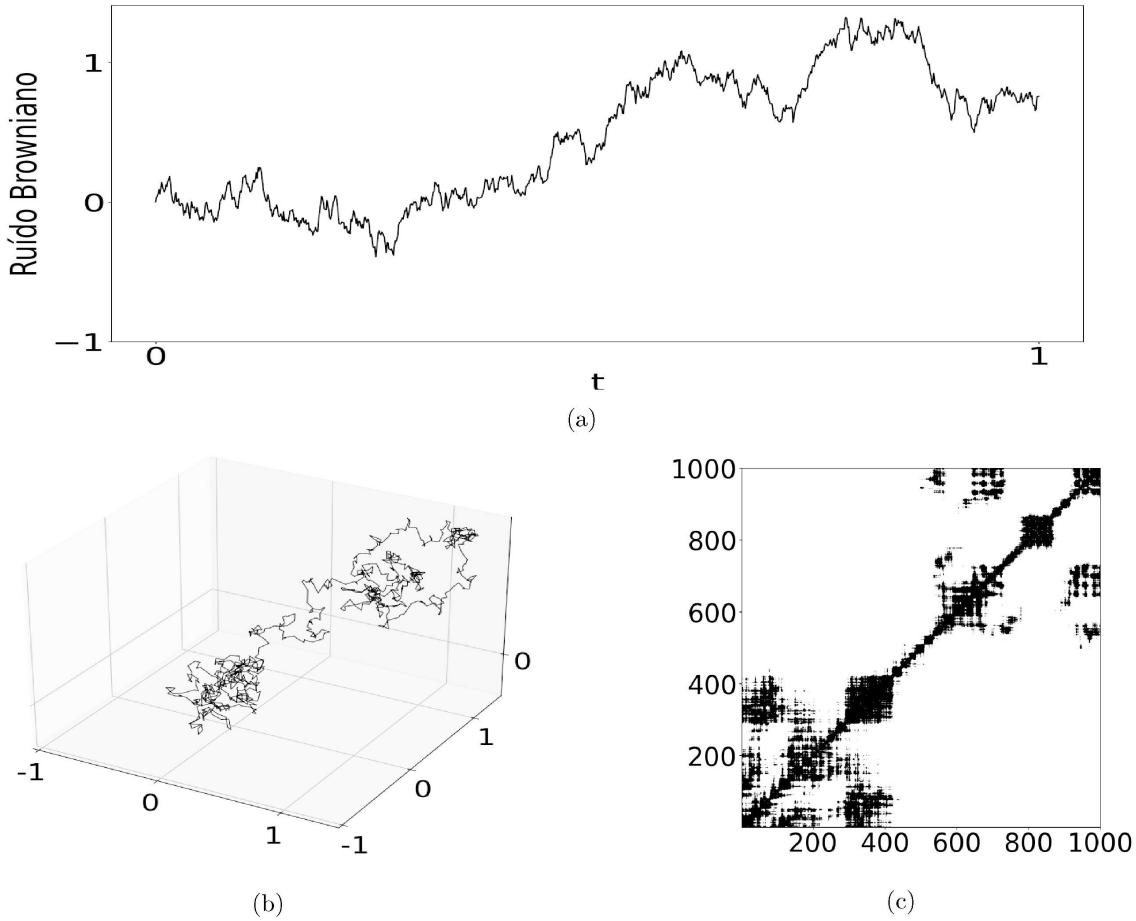


Figura 7 – Ilustração de tipologia homogênea. Em (a) vemos uma série temporal estocástica sem autocorrelação (ruído branco), em (b) vemos sua trajetória reconstruída em um espaço de estados através de um *embedding* e, em (c), o gráfico de recorrências associado à órbita.

Com Perturbação [25, 54] (*Disrupted*): No caso de séries estocásticas com algum grau de correlação, pode-se citar um *ruído browniano*, também chamado de *ruído vermelho* ou *ruído marrom*. Este é um tipo de ruído que pode apresentar uma determinada tendência, e por isso, mesmo que no infinito ele seja estacionário, isso não é visto na prática para séries a tempos finitos. Sinais desta natureza aparecem frequentemente em economia, na forma de flutuações de preços no mercado financeiro e produzem gráficos de recorrências cuja densidade de pontos diminui a medida que nos afastamos da diagonal principal, devido ao fato de que os pontos iniciais já não recorrem mais com os do final da série. Gráficos de recorrências oriundos de séries desta natureza apresentam desdobramentos diferentes dos outros vistos até então, formando estruturas complexas, sem muitas diagonais longas ou pontos isolados, o que os tornam únicos e fáceis de identificar (Figura 8).

3.2.2 Variações dos Gráficos de Recorrências

Existem outros tipos de gráficos de recorrências, utilizados tanto na análise de uma única série de cada vez, como é o caso dos RPs descritos nesta seção, como para analisar mais de uma série ao mesmo tempo,



(a)

(b)

(c)

Figura 8 – Ilustração de tipologia com perturbação. Em (a) vemos uma série temporal estocástica com autocorrelação (ruído browniano), em (b) vemos sua trajetória reconstruída em um espaço de estados através de um *embedding* e, em (c), o gráfico de recorrências associado à órbita.

como os *Cross Recurrence Plots* (CRP) [56, 57] e os *Joint Recurrence Plots* (JRP) [58, 59], assim como muitas outras variações, cada uma com suas vantagens e aplicações. Além da tipologia, também podemos analisar as *texturas* de um gráfico de recorrências, que são estruturas "microscópicas" do RP, em contraste com as tipologias, que são estruturas macroscópicas. As texturas são as estruturas menores como pontos isolados, linhas verticais e diagonais que formam as tipologias, por isso são chamadas de "microscópicas". Para estudá-las de forma mais detalhada, surgiu uma extensão do método descrito nesta seção chamado de *Análise de Quantificação de Recorrências*, conhecida por RQA (*Recurrence Quantification Analysis*). Para uma descrição mais detalhada da teoria e da matemática por trás destas ferramentas vide a referência [25].

3.3 Análise de Quantificação de Recorrências

As texturas que podem surgir em um gráfico de recorrências carregam muita informação acerca do sistema que está sendo analisado. Pensando nisso, Joseph P. Zbilut e Charles L. Webber Jr. elaboraram em 1992 [60] uma maneira de extrair estas informações diretamente das estruturas que formam as texturas dos gráficos de recorrências, baseada em cálculos estatísticos envolvendo as distribuições de pontos isolados, assim como de

linhas diagonais e verticais. Este método passou a ser conhecido por Análise de Quantificação de Recorrências (RQA). Extensões e melhorias no método foram propostas por diversos autores desde então, o que ampliou ainda mais a eficiência da técnica. As estatísticas das texturas dos gráficos de recorrências então são combinadas a fim de formar alguns valores, que são utilizados para estudar o comportamento do sistema, esses valores são chamados de "quantificadores de recorrência". Os quantificadores utilizados mais frequentemente são:

Taxa de Recorrência [60, 61, 62] (*Recurrence Rate - RR*): É a fração de pontos que recorrem no RP, dada por:

$$RR = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} R_{ij},$$

na qual N é o número de pontos da trajetória no espaço de estados. Ela nos dá a probabilidade de um ponto da órbita recorrer a uma certa região dos espaço de estados, dado um determinado ε, a medida em que o sistema evolui. Para casos realistas, onde temos uma série finita, a RR representa a densidade de pontos no RP e pode ser relacionada à *soma de correlação* [63]. No limite de $\varepsilon \to 0$ ela é igual a integral de correlação (eq. 2.5). Este quantificador pode ser utilizado no estabelecimento de um valor para o erro ε, pois vários pesquisadores escolhem um determinado ε que mantém a RR em um certo valor pré-definido [64].

Determinismo [60, 61, 62] (*Determinism - DET*): Este quantificador representa, em teoria, a predictabilidade do sistema, no sentido em que ele quantifica sequências de pontos que recorrem simultaneamente. No lugar de quantificar apenas pontos recorrentes isolados, o DET nos dá uma estatística para subconjuntos da órbita que são recorrentes, não apenas estados isolados. O determinismo é dado pela expressão

$$DET = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}, \tag{3.7}$$

na qual $P(l)$ é o número de diagonais de tamanho l e $P(\cdot)$ é o histograma das diagonais e pode ser obtido por meio da expressão [25]

$$P(l) = \sum_{i=1}^{N} \sum_{j=1}^{N} (1 - R_{i-1,j-1})(1 - R_{i+l,j+l}) \prod_{k=0}^{l-1} R_{i+k,j+k}. \tag{3.8}$$

Como o índice do somatório no denominador começa em 1, vemos que o denominador da expressão 3.7 é igual ao quantificador RR multiplicado por N^2. No numerador, o parâmetro l_{min} é o comprimento mínimo que uma sequência recorrente de pontos precisa ter para ser considerada uma diagonal, portanto, $l_{min} \geq 2$. De modo geral, a escolha desta quantidade não é embasada em argumentos ou conceitos bem estabelecidos na literatura. Este quantificador é frequentemente utilizado para encontrar mudanças no comportamento dinâmico do sistema, como transições de regimes da forma ordem $\to$ caos [65] ou até mesmo caos $\to$ caos [66].

Para encontrar transições de regime na dinâmica de sistemas, podemos, por exemplo, obter o determinismo do sistema para diversos valores de algum parâmetro. A figura 9 reflete o último caso, na qual ela mostra o diagrama de órbitas para o mapa logístico, dado pela equação

$$x_{n+1} = r x_n (1 - x_n), \tag{3.9}$$

e seu determinismo em função de seu parâmetro, expondo a correlação entre o quantificador e as janelas periódicas (determinismo mais alto) e as regiões caóticas (determinismo mais baixo).

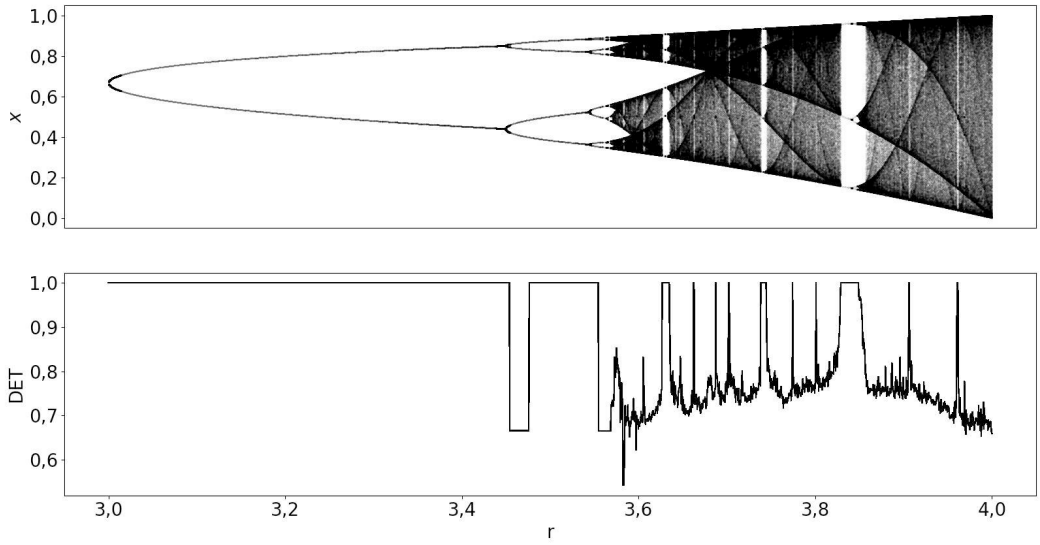


Figura 9 – Acima: o diagrama de órbitas do mapa logístico para o parâmetro nos intervalos 3 até 4. Abaixo: o determinismo do mapa logístico para cada um dos parâmetros. O quantificador foi calculado a partir de um gráfico de recorrências gerado com um ε igual a 0,05. Vemos que há máximos em transições ordem $\rightarrow$ caos e vice-versa.

Ratio [60, 61, 62] (*RATIO*): É a razão do determinismo pela taxa de recorrência, dado pela expressão:

$$\begin{aligned} RATIO &= \frac{DET}{RR} \\ &= \frac{\frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}}{\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} R_{ij}} \\ &= \frac{\frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}}{\frac{1}{N^2}\sum_{l=1}^{N} lP(l)}. \end{aligned}$$

Como já vimos que o somatório no denominador do determinismo, $\sum_{l=1}^{N} lP(l)$ pode ser escrito como $\sum_{i=1}^{N}\sum_{j=1}^{N} R_{ij}$, que é igual a $N^2(RR)$, a $RATIO$ pode ser escrita como:

$$RATIO = N^2 \frac{\sum_{l=l_{min}}^{N} lP(l)}{\left(\sum_{l=1}^{N} lP(l)\right)^2}.$$

Assim como outros quantificadores, a $RATIO$ pode ser utilizada como ferramenta na análise de séries temporais.

Diagonal Média [60, 61, 62] (*Mean Diagonal - L*): Este quantificador é o valor médio do comprimento das diagonais no RP. Ele nos dá informação sobre o tamanho médio das seções recorrentes na trajetória em

questão, ou, em outras palavras, o tempo médio de predição. Este quantificador é útil para complementar os resultados obtidos com o determinismo e é dado pela expressão

$$L = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=l_{min}}^{N} P(l)}.$$

Diagonal Máxima [60, 61, 62] (*Longest Diagonal Line - L_{max}*): É o maior comprimento das diagonais em um gráfico de recorrências, sem considerar a diagonal principal (também chamada de *Line of Identity* - LOI) e é dada por

$$L_{max} = \max\{l_k\}_{k=1}^{N_l},$$

onde N_l é número de linhas diagonais no RP, sem a LOI. Este quantificador nos mostra qual o tamanho da maior subsequência recorrente em uma trajetória ou série temporal.

Divergência [25, 46] (*Divergence - DIV*): É o inverso da diagonal máxima L_{max}, sendo definido pela expressão

$$DIV = \frac{1}{L_{max}}.$$

Existe uma relação [67] entre a divergência DIV de uma série com seu maior expoente de Lyapunov, na qual as duas quantidades possuem tendências similares.

Entropia de Shannon das Diagonais [60, 61, 62] (*Shannon's Entropy of Diagonal Lines - ENTR*): É a entropia de informação [43] da distribuição das linhas diagonais, dada pela equação

$$ENTR = -\sum_{l=l_{min}}^{N} p(l)\log\left(p(l)\right),$$

na qual $p(l)$ é a probabilidade de se encontrar uma linha diagonal de comprimento l, e é definida como

$$p(l) = \frac{P(l)}{\sum_{l=l_{min}}^{N} P(l)}.$$

Este quantificador fornece uma maneira de quantificar a complexidade da série, onde uma série estocástica que apresenta uma quantidade baixa de linhas diagonais teria uma entropia mais baixa, mostrando assim, que não há complexidade, ou seja, não há uma regra por trás dela, ao passo que uma série com mais recorrências implicaria em mais diagonais, resultando em uma entropia/complexidade maior.

Assim como os quantificadores foram definidos anteriormente com base nas estruturas diagonais, outros foram criados de forma análoga, porém, com base nas estruturas formadas por linhas verticais (ou horizontais, já que RPs são simétricos com relação à LOI). Vejamos de forma rápida quais são os quantificadores mais utilizados:

Laminaridade [66] (*Laminarity - LAM*): É o análogo baseado em linhas verticais do determinismo, nos mostra a fração de pontos do RP que formam estruturas verticais/horizontais, e é dado pela expressão

$$LAM = \frac{\sum_{v_{min}}^{N} vP(v)}{\sum_{v=1}^{N} vP(v)}.$$

Este quantificador é um dos mais utilizados na busca por transições caos-caos, pois ele nos fornece informações acerca dos estados intermitentes do sistema analisado, assim como estados laminares. Isso ocorre pois as estruturas verticais possuem a forma $\{(i,j),(i+1,j),\ldots,(i+k,j)\}$ onde k é o comprimento da linha vertical, isso significa que um mesmo estado $\vec{x}(j)$ recorre com outros k estados. Isso mostra que o sistema não mudou de estado, ou está mudando muito lentamente, o que reflete no nome "laminaridade". Nesta expressão, $P(v)$ é o número de linhas verticais de tamanho v e $P(\cdot)$ é o histograma de todas estas linhas. Também temos que v_{min} é o tempo mínimo, em unidades de iteração, que um estado precisa se manter estacionário para ser considerado uma linha vertical e contribuir para a LAM. Pode-se calcular o quantificador laminaridade a medida em que o sistema evolui.

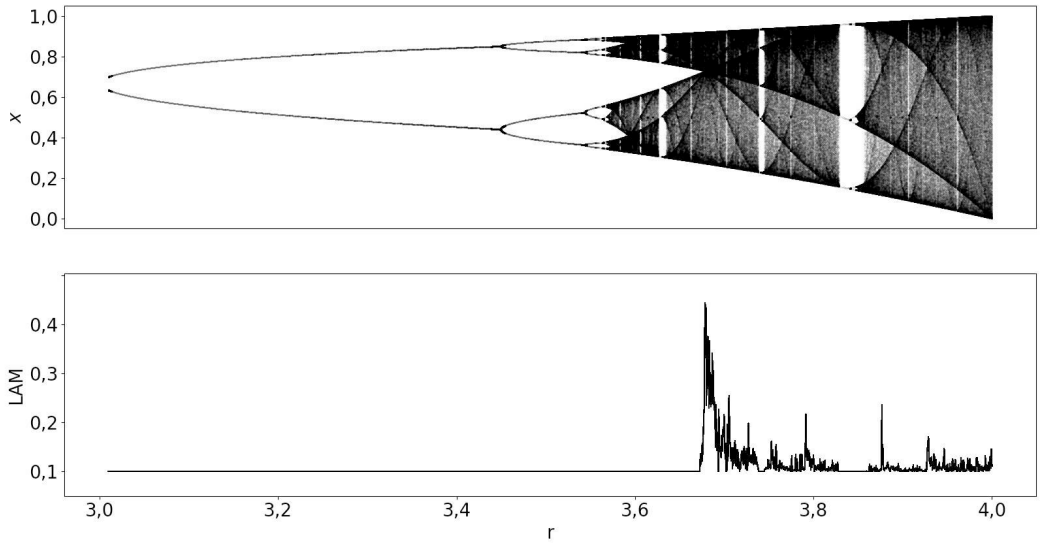


Figura 10 – Acima: diagrama de órbitas do mapa logístico para $r \in [3,4]$. Abaixo: LAM para o mesmo intervalo do parâmetro. O quantificador foi calculado a partir de um gráfico de recorrências gerado com um ε igual a 0,05. Os picos deste quantificador ocorrem nas transições caos $\rightarrow$ caos.

Tempo de Aprisionamento [66] (*Trapping Time - TT*): Este quantificador é o análogo vertical da diagonal média L, ou seja, o tempo TT é o comprimento médio das linhas verticais em um RP. Esta estatística corresponde ao tempo médio em que o sistema se mantém em algum estado laminar, por este motivo recebeu o nome "tempo de *aprisionamento*". O TT é dado pela expressão

$$TT = \frac{\sum_{v=v_{min}}^{N} vP(v)}{\sum_{v=v_{min}}^{N} P(v)}.$$

Vertical Máxima [66] (*Longest Vertical Line - V_{max}*): Como o nome já nos mostra, é o comprimento da maior linha vertical encontrada no RP. Este quantificador é útil para determinar o maior período de tempo que um sistema se mantém estacionário ou laminar. É obtido por

$$V_{max} = \max\{v_k\}_{k=1}^{N_v},$$

na qual N_v é o número de linhas verticais no gráfico de recorrências. Existe um análogo com linhas verticais para quantificador que usa linhas diagonais, porém, foram citados apenas os mais utilizados.

Também é possível estimar estes quantificadores a partir de um *lag*, ou seja, a partir de uma determinada distância da diagonal principal.

Tendência [60, 61, 62] (*TREND*): É obtida observando como a densidade de pontos se comporta ao nos afastarmos da diagonal principal e é dada por

$$TREND = \frac{\sum_{\tau=1}^{\tilde{N}}(\tau - \frac{\tilde{N}}{2})(RR_\tau - \langle RR_\tau \rangle)}{\sum_{\tau=1}^{\tilde{N}}(\tau - \frac{\tilde{N}}{2})^2},$$

na qual RR_τ é a taxa de recorrência dos pontos em cada linha diagonal separadamente, com τ sendo a distância da linha à LOI. Por exemplo, se quisermos saber a taxa de recorrência dos pontos na diagonal imediatamente abaixo da diagonal principal temos que $\tau = -1$, se quisermos saber a RR para a 5ª diagonal acima da LOI, usamos $\tau = +5$, e assim por diante. Esta taxa de recorrência por linha é dada por

$$RR_\tau = \frac{1}{N - \tau} \sum_{i,i+\tau} R_{i,i+\tau}.$$

Todas os outros quantificadores que dependem de estruturas diagonais também podem ser definidos de maneira similar.

Um novo quantificador chamado de *lacunaridade* (*lacunarity*) foi proposto em 2021 [68] e não depende das estruturas previamente mencionadas, como linhas verticais, diagonais ou *delays*, mas sim em sub regiões do gráfico de recorrências e será definida a seguir.

Lacunaridade [68] - (*Lacunarity* - Λ): A lacunaridade é definida através da determinação de um tamanho w para matrizes internas do RP. Encontra-se a "massa" desta região denotada por $\bar{M}$, como sendo a soma dos valores nela e, posteriormente, encontra-se a distribuição de probabilidades $P_w(\bar{M})$ dessas "massas". Os momentos de ordem até dois são então obtidos como

$$Z^{(q)}(w) = \sum_{\bar{M}} \bar{M}^q P_w(\bar{M}),$$

onde a uma lacunaridade é obtida como

$$\tilde{\Lambda} = \frac{Z^{(2)}(w)}{[Z^{(1)}(w)]^2}.$$

É então obtido o quantificador complementar à laminaridade denotado por $\tilde{\Lambda}^\dagger$ dado pela soma dos pontos dentro das mesmas sub regiões, porém, com os valores trocados ($1 \leftrightarrow 0$), fazendo com que o quantificador utilizado na pratica seja

$$\Lambda = 2 - \left(\frac{1}{\tilde{\Lambda}(w)} + \frac{1}{\tilde{\Lambda}^\dagger(w)} \right). \tag{3.10}$$

O objetivo que este quantificador se propõe a satisfazer é a detecção das lacunas (ou *gaps*) nos gráficos de recorrências, com isso oferecendo mais uma maneira de quantificar a complexidade do sistema estudado.

3.3.1 Entropia de Microestados

Vimos na seção sobre gráficos de recorrências que, ao gerá-los, é necessário encontrar um parâmetro denotado por ε, e corresponde à distância máxima em que dois pontos no espaço de estados devem ter entre si para serem considerados recorrentes. Entretanto, com a proposta do quantificador *entropia de microestados*, denotada por S_d, este ε deixa de ser um parâmetro livre, o que remove uma possível arbitrariedade do método. Normalmente, o método de escolha do parâmetro ε, mesmo com sua importância, não é bem definido, sendo ainda objeto de estudo [69, 70]. Com a proposta da entropia de microestados, isso deixa de ser um problema. Isso ocorre pois cada valor da distância ε irá resultar em um valor diferente para S_d, ou seja, $S_d = S_d(\varepsilon)$, e o erro escolhido é, então, aquele que maximiza esta entropia. A maximização de funções de entropia é algo comum e muito utilizado em várias áreas, inclusive na Física [71, 72], se mostrando até os dias de hoje, um conceito de extrema importância. Em sistemas dinâmicos isso não seria diferente. Estes *microestados*, são sub-regiões da matriz R e assumem a forma de matrizes menores que constituem a matriz de recorrência R (Figura 11a).

Proposto em 2018 [73], este quantificador olha para as estruturas dos RPs de uma maneira distinta dos outros vistos até então. No lugar de analisar as estatísticas de pontos isolados e estruturas diagonais e verticais, S_d analisa subregiões da matriz R, obtendo assim os chamados "microestados". Seja uma matriz de recorrência R de dimensões $N \times N$ e uma matriz *quadrada* de microestados M de dimensão $d \times d$, com $N \geq d \geq 2$. O número total de microestados *distintos* uns dos outros é, então, $N_d = 2^{d^2}$, por exemplo, para microestados 2×2 existem 16 possibilidades, para 3×3 existem $N_3 = 2^{3^2} = 2^9 = 512$ possibilidades, e assim por diante, mostrando assim, que o conjunto $\{M_n\}_{n=1}^{N_d}$ de matrizes distintas cresce muito rapidamente. Ao analisar um RP a partir deste método, uma fração das matrizes de recorrência M é então coletada de forma aleatória, após isto, um histograma é feito e a entropia é então calculada a partir dele. A partir deste histograma, obtêm-se a entropia da forma

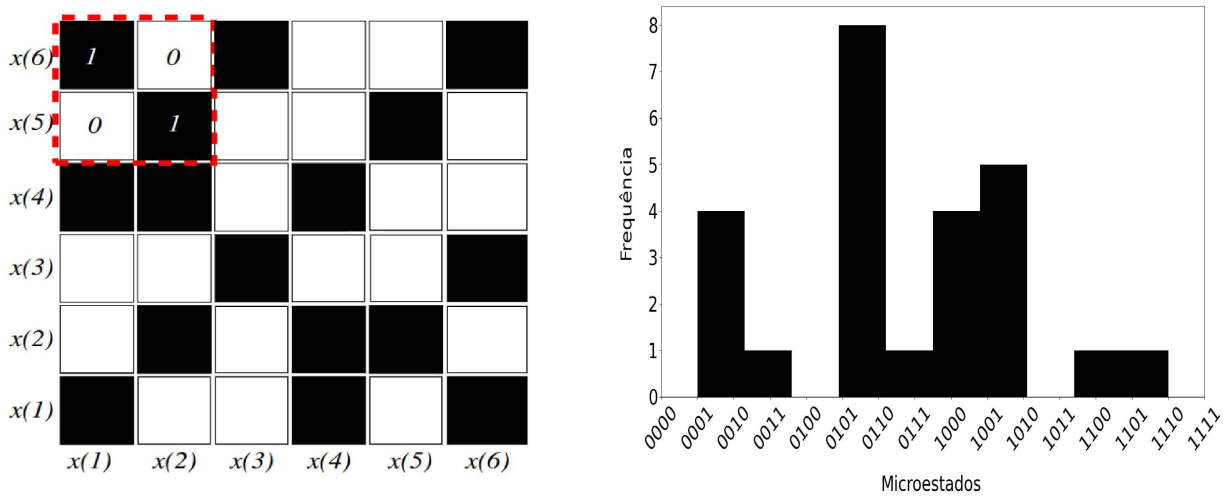
$$S_d = -\sum_{n=1}^{N_d} p(M_n) \log\left(p(M_n)\right),$$

com $p(M_n)$ sendo a probabilidade de se encontrar o microestado M_n no histograma gerado, sendo dada pela equação

$$p(M_n) = \frac{P(M_n)}{\sum_{n=1}^{N_d} P(M_n)}, \tag{3.11}$$

na qual $P(M_m)$ é o número de vezes que o microestado M_n aparece no histograma.

Para referenciar os microestados, pode-se utilizar uma representação binária, como ilustrado na figura 11a. Nela, o microestado destacado (canto superior esquerdo da figura) é denotado por "1001", sendo que existem outros que podem ser denotados por "0001" ou "1110", por exemplo. Esta forma de rotular os microestados deixa mais claro o motivo da expressão $N_d = 2^{d^2}$. Estes microestados levam em consideração não apenas as recorrências pontuais ou sequenciais, mas também relações entre diferentes configurações dos elementos de uma série. Possibilitando assim, obter informações relevantes acerca da dinâmica do sistema de forma muito mais eficiente do que os quantificadores clássicos. Até agora neste trabalho, o conceito de *embedding* em um espaço de estados m-dimensional e de um *time delay* τ foi muito importante para o cálculo dos quantificadores vistos, porém, a necessidade de um *embedding* é questionável [74, 75]. O motivo disto ocorrer é que toda a informação extraída desta forma também se faz presente na própria série, ou seja, não há um



(a) (b)

Figura 11 – Ilustração de uma matriz de recorrência evidenciando um microestado e seu respectivo histograma. Em (a), vemos o gráfico de recorrências de uma série de 6 pontos, evidenciando uma matriz de microestados. Imagem adaptada da referência [73]. Em (b), vemos o histograma $P(\cdot)$ dos microestados existentes na figura 11a.

ganho efetivo de informação ao se realizar um *embedding*. A entropia de microestados S_d proposta consegue resultados excepcionais mesmo para gráficos de recorrências gerados a partir das próprias séries, ou seja, sem usar órbitas reconstruídas. Isso equivale à um *embedding* em um espaço unidimensional, onde apenas recorrências escalares[6] são consideradas. Para exemplificar, a figura 12b mostra como um ε é escolhido para o gráfico de recorrências de um seno, sem *embedding*, por meio da maximização de sua entropia de microestados e o RP correspondente.

Para uma série estocástica sem autocorrelação nós podemos esperar encontrar todos os tipos de microestados com probabilidades iguais, da forma $p(M_n) = p = N_d^{-1}$, $\forall n \in [1, N_d]$, fazendo com que a expressão 3.3.1 se torne

$$\begin{aligned} S_d &= -\sum_{n=1}^{N_d} N_d^{-1} \log\left(N_d^{-1}\right) \\ &= N_d^{-1} N_d \log\left(N_d\right) \\ &= \log\left(2^{d^2}\right) \\ &= d^2 \log(2), \end{aligned} \tag{3.12}$$

na qual as matrizes de microestados possuem dimensão $d \times d$. Outro resultado é o menor valor possível para a entropia máxima, que ocorre quando todas as matrizes de microestados são iguais, ou seja, há apenas probabilidade de encontrar M_n, para algum $n = n^*$, e esta probabilidade é $p(M_{n^*}) = 1$, com isso

$$S_d = -\sum_{n=1}^{N_d} 1 \log(1) = 0, \tag{3.13}$$

[6] Recorrências dos valores da própria série, sem envolver a norma de vetores m-dimensionais.

que é o mínimo teórico. O último caso de interesse ocorre quando há apenas dois microestados, digamos M_n e M_m, que resulta em $p(M_n) = p(M_m) = \frac{1}{2}$, correspondendo ao seguinte valor para a entropia:

$$S_d = -\sum_{n=1}^{N_d} \frac{1}{2} \log\left(\frac{1}{2}\right) = \log(2). \tag{3.14}$$

A maximização da entropia leva à geração de gráficos de recorrências que nos permitem extrair o máximo de informação possível sobre a série com base nas inter-relações entre seus elementos. Estas relações se refletem em sua distribuição de microestados, que levam em consideração mais do que apenas recorrências simples, mas também as relações entre elementos que não necessariamente recorrem entre si. Isso é observado pois as matrizes de microestados contém tanto pontos recorrentes como não recorrentes, o que torna este método mais geral, além de ser uma alternativa viável para a determinação da distância mínima ε.

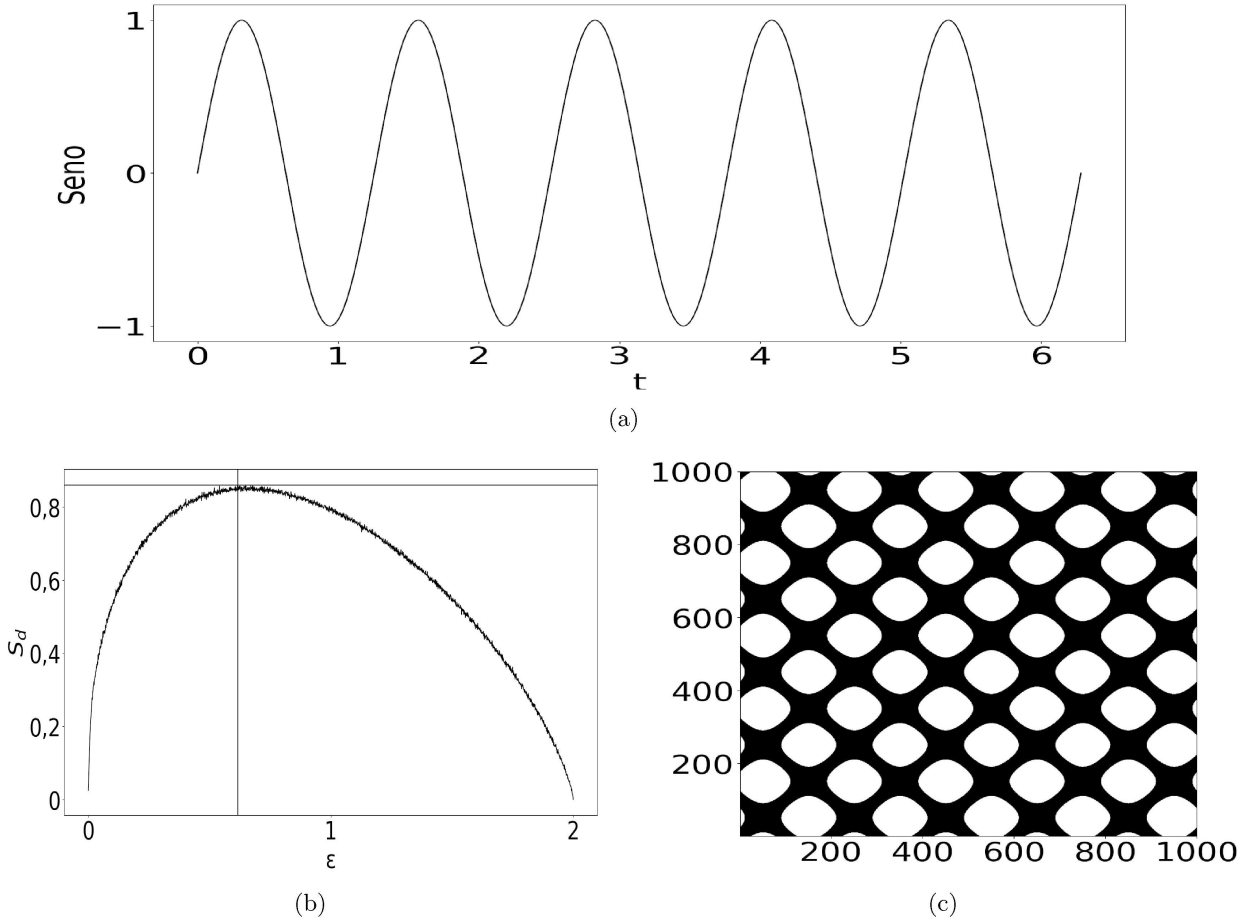


Figura 12 – Ilustração de como obter o ε para um RP através da maximização da entropia de microestados. Em (a), vemos o seno utilizado para ilustrar a maximização da entropia de microestados. Em (b), a entropia de microestados da série da figura 12a em função do erro ε. Em (c), vemos o gráfico de recorrências da série da figura 12a gerado com o ε obtido na figura 12b.

3.4 Critério Atual de Escolha do Parâmetro l_{min}

Para poder utilizar o quantificador *determinismo* são necessários três parâmetros, sendo dois deles explícitos na expressão para DET e um implícito. As equações 3.7 e 3.8 nos dizem que

$$DET = \frac{\overbrace{\sum_{l=l_{min}}^{N} l \sum_{i=1}^{N}\sum_{j=1}^{N}(1-R_{i-1,j-1}(\varepsilon))(1-R_{i+l,j+l}(\varepsilon))\prod_{k=0}^{l-1}R_{i+k,j+k}(\varepsilon)}^{P(l)}}{\underbrace{\sum_{l=1}^{N} l \sum_{i=1}^{N}\sum_{j=1}^{N}(1-R_{i-1,j-1}(\varepsilon))(1-R_{i+l,j+l}(\varepsilon))\prod_{k=0}^{l-1}R_{i+k,j+k}(\varepsilon)}_{P(l)}}. \tag{3.15}$$

Vemos com isso que N e l_{min} aparecem de forma explícita na expressão 3.15 nos limites dos somatórios, porém, ε aparece apenas implicitamente como um parâmetro da matriz de recorrência $R(\varepsilon)$. O parâmetro N representa a quantidade de pontos presentes na série a ser analisada enquanto ε representa o erro, ou *threshold*, utilizado na geração da matriz de recorrência. O parâmetro l_{min} é o comprimento mínimo necessário (ou quantidade de pontos) que as linhas diagonais devem ter para que o trecho correspondente da série seja considerada uma subsequência relevante de pontos recorrentes [25, 45, 46]. Em qualquer análise é possível saber o valor de N, o que o remove da lista de parâmetros livres. No caso de ε, ele pode ser obtido de várias maneiras. Considerando como método padrão para a sua obtenção a maximização da entropia de microestados, também podemos removê-lo da lista de parâmetros livres. Contudo, l_{min} ainda não possui uma maneira automatizada de ser obtido, ou seja, há um parâmetro livre que introduz arbitrariedade nas análises feitas com os quantificadores que o utilizam. Neste trabalho nos focaremos no determinismo.

Atualmente não há uma grande variedade de trabalhos que tentam encontrar o l_{min} de modo a melhorar a qualidade das análises [76], o que leva pesquisadores a considerarem o valor $l_{min} = 2$ como sendo algo natural e mais adequado para a maioria dos casos. O trabalho mais recente data de 2014 e nos mostra uma análise do l_{min} baseado em qual valor para este parâmetro maximiza a variância do determinismo ao obtê-lo para diferentes configurações de um sistema dinâmico. Com isso, a premissa é que o quantificador em questão seria mais eficiente em nos mostrar as transições na dinâmica do sistema.

3.5 Problemas com o Critério Atual

O valor atualmente empregado com mais frequência, entretanto, não é capaz de gerar resultados satisfatórios no que tange à distinção de séries temporais com diferentes dinâmicas. Percebe-se isto pois este critério gera determinismos praticamente iguais para séries caóticas, estocásticas e até mesmo periódicas. Isso ocorre pois ao maximizar a entropia de microestados obtemos valores muito altos para ε, que geram determinismos de aproximadamente 0,75 para ruídos brancos, o que claramente dificulta as análises realizadas. A dificuldade ocorre pois o quantificador em questão perde sua capacidade de dizer se uma série possui algum comportamento dinâmico completamente aleatório ou não.

Utilizando um seno com 5 oscilações e a primeira componente do oscilador de Chua[7], ambos com 1000 pontos, podemos deixar claro o que foi dito. Vemos que os determinismos das duas séries, mesmo que possuam dinâmicas diferentes (uma é periódica e a outra é caótica), são praticamente os mesmos quando calculados no

[7] Visto com mais detalhes na seção sobre resultados.

tempo[8]. A figura 13 ilustra esta afirmação, mostrando que os determinismos de uma série caótica possuem o valor 1, impossibilitando a distinção de diferentes regimes dinâmicos para a série.

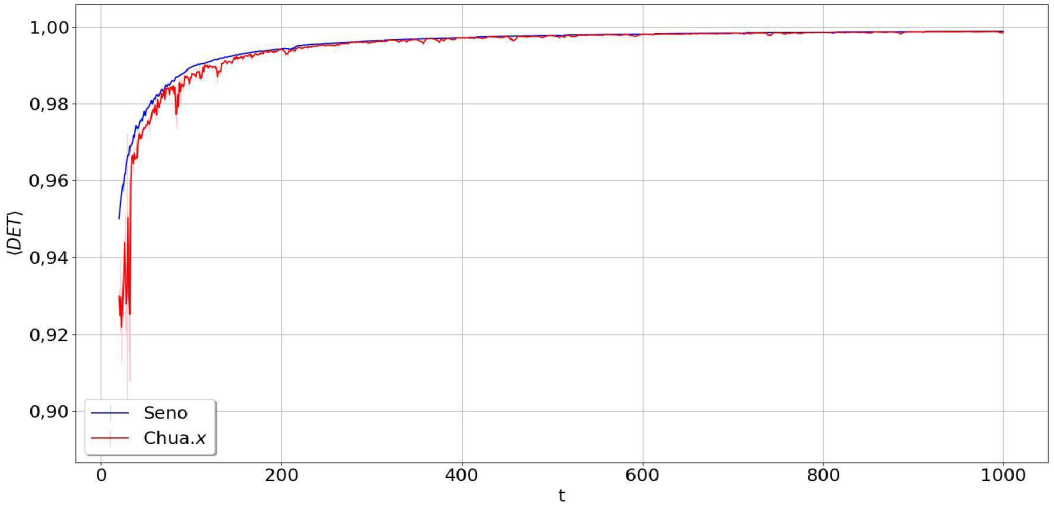


Figura 13 – Comparação entre os determinismos de um seno com 5 períodos e da componente x do oscilador de Chua obtidos no tempo, em unidades de iteração, onde ambas as séries possuem o mesmo passo temporal.

Apenas com este caso já é possível ver que o critério atual de escolha do l_{min} não é sempre adequado. Com isso em mente, alguns pesquisadores escolhem outros valores para o l_{min} ao realizarem seus trabalhos. Porém, não há uma maneira de se obter este parâmetro de forma autônoma e que gere valores adequados, esta é a maior motivação para este trabalho.

[8] Esta maneira de se obter o determinismo também será vista na seção sobre resultados.

4 Métodos de Definição da Diagonal Mínima de Recorrências

Tendo em vista a clara discrepância entre a proposta do quantificador DET e seu comportamento real, propomos uma alternativa no processo de escolha do l_{min}. Esta alternativa consiste em fazer com que o determinismo obedeça duas restrições, para defini-las, utilizaremos o fato de que uma série periódica precise, necessariamente, ter um alto nível de determinismo. Uma outra condição de contorno que será utilizada corresponde ao outro extremo, ou seja, uma série aleatória sem autocorrelação precisa, necessariamente, ter um determinismo mais baixo quanto possível.

Quando falamos em escolher o tamanho mínimo de uma diagonal em um gráfico de recorrências é preciso ter cuidado, pois este parâmetro possui várias sutilezas. Ao utilizar um valor inadequado, podemos obter resultados desde verdadeiros, passando por ambíguos e até falsos. Isto dificulta as análise feitas, o que pode ser visto pelo fato de séries caóticas serem tão determinísticas quanto séries perfeitamente periódicas, o que é errado. Sabemos que mesmo séries caóticas possuem uma lei de formação, como os exemplos já mencionados (mapa logístico, sistema de Lorenz e sistema de Rössler, por exemplo), mas também sabemos que estas leis de formação sempre são menos determinísticas do que séries periódicas. Isso faz com que em algumas análises nós tenhamos uma ideia equivocada ao tentar entender os motivos das oscilações em uma curva de determinismo, pois nem sempre um mínimo representa uma região caótica e nem sempre um máximo representa uma região periódica.

Consideremos uma série $s(t) \in \mathbb{S} \subset \mathbb{R}$ com N pontos, onde sua variável dependente é $t \in \mathbb{I} \subset \mathbb{R}$, onde $s(t)$ possui um número ρ de oscilações para $t \in \mathbb{I}$. Estes valores são importantes para as balizas que foram estipuladas, pois mesmo um seno (ou qualquer outra série) pode ser interpretado como uma série totalmente diferente quando $\rho \to \infty$ ou $N \to 0$. Veremos agora como isso pode ser facilmente verificado. Considere que o espaço onde a série está contida seja $\mathbb{U} = \mathbb{I} \times \mathbb{S}$. Por exemplo, se uma série for formada a partir da função

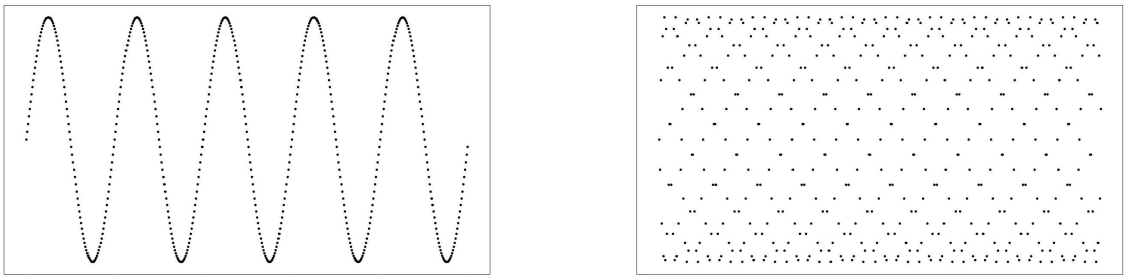
$$s(t) = \frac{1}{2}\left(\sin(2\pi \cdot \rho \cdot t) + 1\right), \tag{4.1}$$

vemos que $\mathbb{S} = [0, 1]$, que é o conjunto dos possíveis valores para a série e, se a analisarmos para $t \in [0, 1] = \mathbb{I}$, ela apresentaria exatamente ρ oscilações, e o espaço $\mathbb{U}$ seria o quadrado unitário $[0, 1] \times [0, 1]$. Para este caso, quando $\rho \gg 1$ e N fixo, temos o comportamento mostrado na figura 14.

Com isso vemos que ρ é importante para a análise, portanto, deve influenciar no cálculo de l_{min}. Agora vamos ver que N também tem importância (Figura 15).
Vemos que existe uma grande diferença entre as figuras 15b e 15a, mesmo que a série seja essencialmente a mesma. Com isso fica claro que N também é um parâmetro de grande importância na análise, o que é óbvio, quanto maior a amostra mais precisas são as estatísticas que se pode obter.

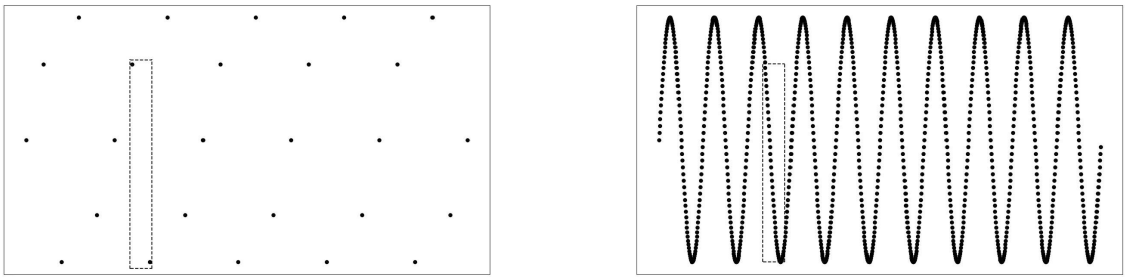
4.1 Sistemas Contínuos

Podemos inicialmente pensar de uma maneira qualitativa sobre isso. Vemos então que quando N aumenta (quando a *resolução* da série aumenta), os pontos da série ficam cada vez mais próximos uns dos outros, fazendo com que uma recorrência de apenas dois pontos seja muito mais importante para baixas resoluções do que para N grande. Veremos agora o motivo disso acontecer. Ao pegarmos um trecho γ da série S,



(a) (b)

Figura 14 – Representações gráficas de funções seno com a mesma quantidade de pontos ($N = 500$) e diferentes números de oscilações em um mesmo intervalo. Em (a), $\rho = 5$. Em (b), $\rho = 270$.



(a) (b)

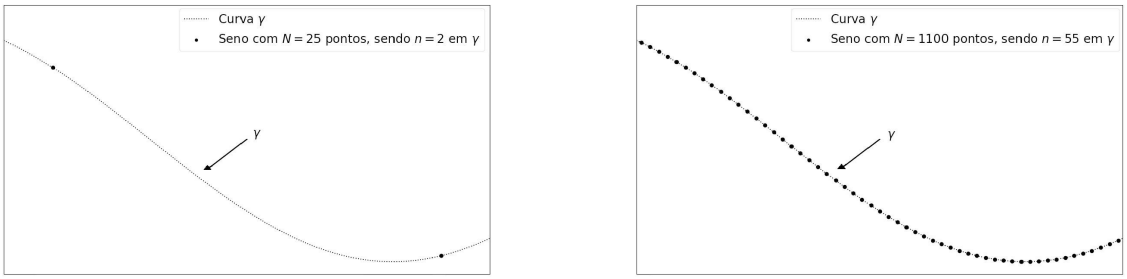
Figura 15 – Senos com o mesmo número de oscilações ($\rho = 10$) e diferentes quantidades de pontos em um mesmo intervalo. Uma caixa está selecionando uma região para ser vista posteriormente. Em (a), $N = 25$. Em (b), $N = 1100$.

ele conterá uma determinada quantidade $n \leq N$ de pontos. À medida que N aumenta[1], se mantivermos o comprimento da curva γ constante, veremos uma aglomeração cada vez maior de pontos em γ. Isso significa que esta curva comportará um número de elementos $n' \geq n$, culminando em uma diminuição da distância entre os pontos contidos neste trecho.

Podemos simplificar com um exemplo. Suponha que γ seja a curva que interpola os pontos da série em alguma região de $\mathbb{U}$, ilustrada na figura 15 para $t \in [0,235,\ 0,285]$. Como consequência, γ é o formato subjacente da série para esta região de $\mathbb{U}$. A figura 16 ilustra isso, onde vemos claramente o formato de γ dentro da caixa destacada nas figuras 15a e 15b e o número de pontos dentro de γ. Fica claro que o espaçamento entre os pontos diminui ao passo em que N aumenta, fazendo com que sejam então necessários mais pontos para representar a mesma região. No caso das figuras 16, precisamos de apenas 2 pontos para representar o retângulo ressaltado para $N = 25$, porém precisamos de 55 pontos para $N = 1100$.

Ao tentar obter um l_{min} adequado, precisamos destes conceitos, mesmo que sejam simples, eles são necessários. Podemos afirmar que como a distância entre dois pontos em γ sempre diminui à medida que N aumenta, vemos que o "volume" comportado por estes dois pontos fica cada vez menor. Isso faz com que eles estejam cada vez mais próximos de serem um único ponto do que uma curva com alguma relevância para a

[1] É importante frisar que o conceito de o número de pontos da série "aumentar" não significa evolução temporal, mas sim um aumento em sua resolução, pois mantemos $\mathbb{I}$ fixo, lembrando que este é o intervalo em que se encontra a variável independente.



(a) (b)

Figura 16 – Ilustração da importância da resolução em uma série temporal por meio da ênfase da curva subjacente γ, evidenciando a quantidade n de pontos nelae a separação entre estes pontos. Em (a), $n = 2$. Em (b), $n = 55$. Figura ilustrativa com o objetivo de evidenciar a distância entre dois pontos na série, explicitando o fato de que a importância de dois pontos consecutivos diminui a medida em que o número de pontos total da órbita aumenta. Isto ocorre pois o segmento da curva representada por dois pontos consecutivos fica cada vez menos relevante diante do formato completo da curva.

dinâmica do sistema. Como o l_{min} representa o *menor* trecho da órbita (ou da série, no caso de não utilizar *embedding*) que pode ser considerado uma trajetória relevante, vemos claramente que se N aumenta, l_{min} também deve aumentar. Isso mostra que $l_{min} = l_{min}(N)$. Para o caso do número de oscilações ρ, podemos fazer uma análise qualitativa completamente análoga à anterior e, com o auxílio da figura 14, concluir que $l_{min} = l_{min}(N, \rho)$. Vemos isto pois ao manter N fixo e aumentar ρ estamos na verdade aumentando o comprimento total da curva γ, diminuindo assim a densidade de pontos dentro dela, resultando em imagens como as 16. Estas implicações validam nossa conclusão sobre a dependência de l_{min} com ρ.

A hipótese de que há uma curva subjacente γ não é verdadeira para todos os casos, pois em séries temporais oriundas de observações de fenômenos reais há sempre um componente aleatório. Isso faz com que não haja diferenciabilidade nestas séries, em contraste com componentes de sistemas dinâmicos artificiais, que normalmente não possuem a componente erro ϵ. Entretanto, pode-se descartar a necessidade de que a série seja diferenciável ou suave e apenas requerer que ela seja contínua, pois o teorema de Bolzano-Cauchy[2] [77]. Ele afirma que

Teorema 2 (Teorema do Valor Intermediário) *Para qualquer função contínua f no domínio $[a, b]$, existe algum $c \in [a, b]$ tal que $f(a) \leq f(c) \leq f(b)$.*

Ele é condição suficiente para que o método funcione. Com isso em mente, podemos começar a definir como será o processo de escolha do l_{min}. Para isso, chamaremos um seno com N pontos e ρ períodos de $s(t; N, \rho)$ e um ruído branco com N pontos de $w(t; N)$. Precisamos exigir que o determinismo do seno seja tão próximo de 1[3] quanto o determinismo de um ruído branco esteja de 0. Com isso, temos o sistema

$$\begin{aligned} DET[s(t; N, \rho)] &\approx 1, \\ DET[w(t; N)] &\approx 0. \end{aligned} \tag{4.2}$$

A metodologia utilizada para isto foi determinar um *parâmetro de assimetria* A, definido como

[2] Também conhecido por *Teorema do Valor Intermediário.*

[3] A escolha do seno foi feita com base no conceito de que um seno é uma sequência de oscilações simples, onde todas as outras formas de oscilação podem ser escritas como combinações lineares de senos.

$$A = |1 - DET[s(t; N, \rho)] - DET[w(t; N)]|. \tag{4.3}$$

Este parâmetro foi definido desta forma utilizando as equações 4.2, pois queremos que seja satisfeito o sistema

$$\begin{aligned} 1 - DET_{\text{seno}} &\approx 0, \\ DET_{\text{ruído branco}} - 0 &\approx 0. \end{aligned} \tag{4.4}$$

Sempre obteremos um resultado positivo isoladamente para cada uma das diferenças, porém, queremos a distância $d(1 - DET_{\text{seno}},\ DET_{\text{ruído branco}})$ entre elas, por isso A é definido com a função módulo. Podemos entender melhor qual o significado de A através da figura 17.

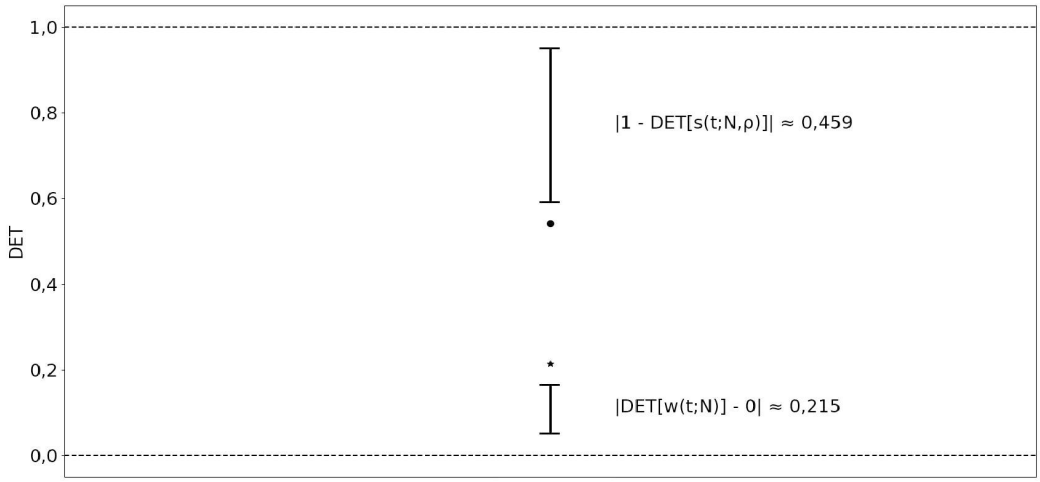


Figura 17 – Ilustração dos determinismos de um seno (ponto) e de um ruído branco (estrela) e a suas distância até seus valores desejados. O parâmetro de assimetria é a diferença nos comprimentos de ambas as barras. Neste exemplo, o parâmetro A seria |1 - 0,459 - 0,215| = 0,326.

O parâmetro A depende implicitamente do l_{min} por meio dos determinismos e nos mostra o quão assimétrico os determinismos são com relação às condições de contorno. É importante ressaltar que como $w(t)$ não possui nenhum número válido para ρ, nós apenas precisamos que $s(t)$ e $w(t)$ tenham o mesmo N. O objetivo é então encontrar um l_{min} que faça $A \to 0$. Portanto, o cálculo de A é feito para vários l_{min} e, em seguida, é selecionado o parâmetro adequado por meio da minimização de A, como ilustrado na figura 18. Podemos simplificar dizendo que ao minimizar A, estamos escolhendo um l_{min} que gera determinismos com as barras na figura 18 que tenham o mesmo comprimento uma da outra.

Desta maneira, podemos escolher o l_{min} que corresponde à um determinado N e ρ e que, quando utilizado, resulta nos determinismos que mais se aproximam simultaneamente das condições 4.2, dentro das limitações da série. Com isso, podemos obter um l_{min} que resulta no determinismo mais próximo do desejado quanto possível. Obviamente, há situações onde a densidade de pontos não é suficiente para uma análise adequada. Entretanto, mesmo nestes casos, o resultado é satisfatório no sentido de ser mais realista do que os obtidos atualmente (como o determinismo de um ruído branco ser $0,75$). Mesmo que as séries estocásticas utilizadas no método como condições de contorno não sejam *de facto* aleatórias, nós ainda podemos utilizá-las, pois estas séries pseudo-aleatórias são muito semelhantes à séries intrinsecamente/verdadeiramente aleatórias.

Isso implica em determinismos também similares para elas, sendo ainda um resultado satisfatório. O papel do determinismo como identificador de mudanças no comportamento de séries temporais ainda pode ser realizado com este l_{min}, porém, com a vantagem de nos dar o valor mais próximo das condições de contorno.

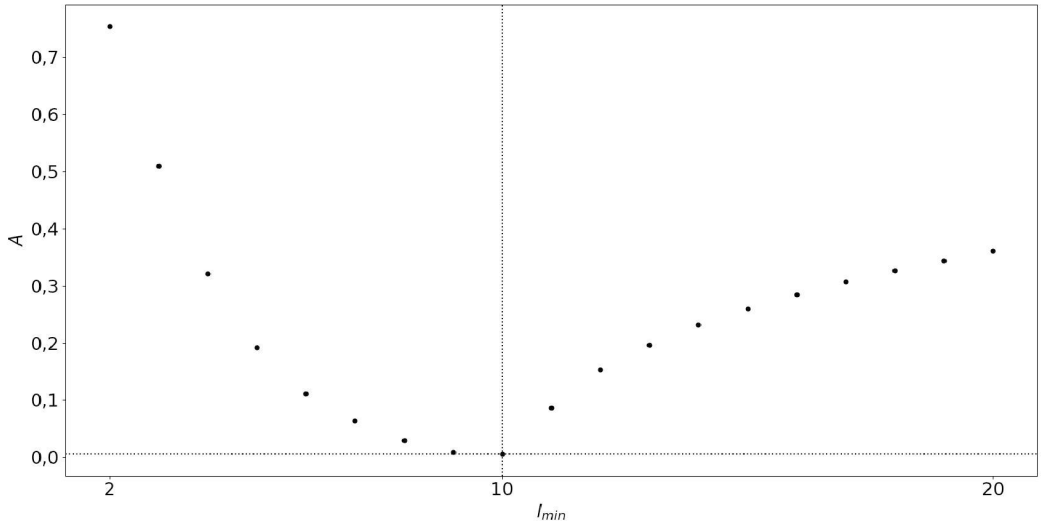


Figura 18 – Ilustração de como se minimiza o parâmetro de assimetria A em função de l_{min}, evidenciando qual o valor de l_{min} escolher.

4.2 Sistemas Discretos

Para o caso de mapas, o conceito de uma curva subjacente que interpola os pontos da série ainda vale, porém, a densidade de pontos nesta curva já não é mais independente do intervalo temporal. O intervalo passa a ser $\mathbb{I} \subset \mathbb{N}$, e o tempo $t \in \mathbb{I}$ passa a ser $t = n$, com n inteiro. Nesse caso, o passo temporal é constante e igual a $1 = (n + 1) - n$, fazendo com que a resolução da série seja uniforme para todo o tempo. Isso mostra que não podemos alterar sua resolução sem evoluir o sistema.

No caso de sistemas contínuos supõe-se uma curva γ e o fato de que o número de pontos nesta curva aumenta conforme N aumenta resulta em uma dependência de l_{min} com N. Para mapas, isso é diferente. Como cada oscilação de um sistema discreto sempre irá possuir o mesmo número de pontos, é perdido um grau de liberdade, pois não podemos mais alterar de forma independente o intervalo temporal e o número de pontos da série. Com isso em mente, ocorre uma dificuldade intrínseca no cálculo de determinismos para mapas com o critério proposto. Uma solução seria a definição de condições de contorno adequadas à sistemas discretos. Este trabalho foca-se em sistemas contínuos.

5 Resultados e Discussão

Veremos agora como este critério de escolha do l_{min} afeta o comportamento dos determinismos. Faremos isto inicialmente para as séries utilizadas para calibrar o método (senos e ruído branco) e, posteriormente, iremos utilizar este método para analisar outros sistemas.

Observa-se que a minimização do parâmetro de assimetria impacta no determinismo de uma maneira peculiar, mostrando que de fato há uma convergência para os valores desejados ao aumentar o número de pontos da série (Figura 19).

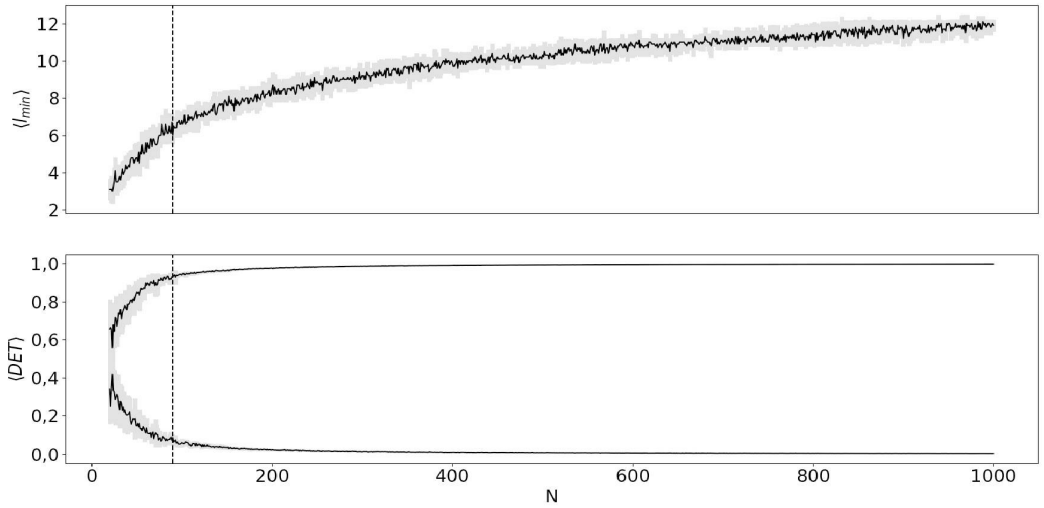


Figura 19 – Valores médios para os l_{min} em função da resolução da série (acima) e, determinismos de um seno com uma oscilação e de um ruído branco (abaixo). Percebe-se que na figura abaixo os determinismos do seno convergem para 1 ao passo em que os determinismos do ruído branco convergem para 0 de forma simétrica com relação às restrições para resoluções maiores. As curvas em preto são os valores médios das quantidades em questão e os sombreados cinzas representam os desvios-padrões.

Para gerar a figura 19, o parâmetro de assimetria A foi calculado e minimizado várias vezes para cada resolução das séries, ou seja, para senos e ruídos brancos ambos com 20 pontos, 21 pontos, ..., 1000 pontos. Isso resulta em vários l_{min} para o mesmo N, gerando determinismos para cada um destes valores. Ocorreram pequenas variações devido à maximização da entropia de microestados, que gera o ε para o gráfico de recorrências. Isso é observado pois os microestados são amostrados aleatoriamente, resultando em um ε levemente diferente a cada cálculo e, consequentemente, em determinismos diferentes. Estes determinismos, então, geram parâmetros de assimetria também ligeiramente diferentes, mas isso apenas introduz pequenas flutuações que não alteram a dependência geral do l_{min} com N. Um fato curioso é que os l_{min} obtidos desta forma possuem um comportamento aparentemente logarítmico com o número de pontos da série, tornando possível sua determinação mesmo sem precisar calcular o parâmetro A. Podemos modelar a curva empiricamente da utilizando a expressão

$$\langle l_{min} \rangle = a + b \cdot \log(N), \tag{5.1}$$

o que mesmo sendo pouco rigoroso, serviria para o propósito do método. Apenas os l_{min} para os N adequados podem ser utilizados para calcular os determinismos de séries com N pontos. Por exemplo, para calcular o determinismo de um seno com uma oscilação e com 400 pontos, deveríamos utilizar um l_{min} sendo aproximadamente 10, pois, pelo gráfico, $l_{min}(N, \rho) = l_{min}(400, 1) \approx 10$. Algumas hipóteses foram leventadas acerta do formato da equação 5.1, uma seria o fato de que a entropia de microestados tem alguma influência no modelo logarítmico. Outra possível causa para o comportamento observado seria uma relação do número de pontos em uma linha diagonal com uma resolução ao número de pontos da mesma linha ao aumentar a resolução, caso o aspecto do gráfico seja mantido à medida que N cresce.

Ao utilizar um seno com outra quantidade de oscilações para calibrar o l_{min} obtêm-se uma lista de valores distintos dos mostrados na figura 19. Isso acontece pois o método, por construção, nos dá o l_{min} mais adequado[1] para um certo N e um certo ρ, ou seja, para cada número de períodos o critério proposto gera uma lista diferente de l_{min}.

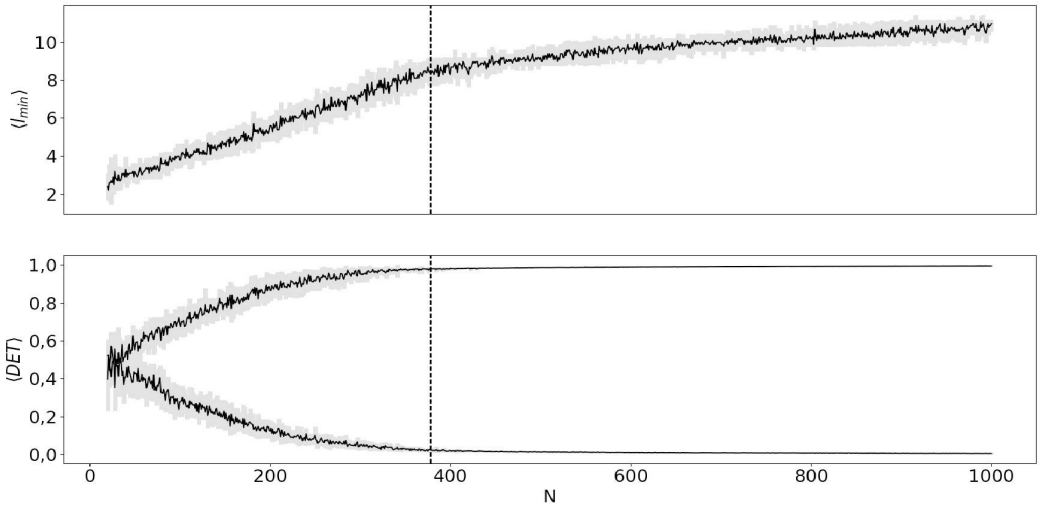


Figura 20 – Valores médios para os l_{min} em função da resolução da série (acima) e, determinismos de um seno com quatro oscilação e de um ruído branco (abaixo). Percebe-se que na figura abaixo os determinismos do seno convergem para 1 ao passo em que os determinismos convergem para 0 de forma simétrica com relação às condições de contorno para resoluções maiores. As curvas em preto são os valores médios das quantidades em questão e os sombreados cinzas representam os desvios-padrões.

Vemos que nas imagens 19 e 20 foram traçadas algumas linhas verticais, onde para a figura 19 seu significado talvez não seja muito evidente no gráfico de $\langle l_{min} \rangle \times N$, mas é possível ver claramente nos determinismos que ela nos mostra a resolução necessária para que o método forneça resultados satisfatórios. A expressão 5.1 é válida para valores de resolução após esta linha vertical. Para os gráficos das médias dos l_{min}, por outro lado, percebemos na figura 20 que essa convergência do método ocorre em um ponto onde a curva dos l_{min} sofre uma alteração em sua taxa de crescimento. Esta resolução mínima/necessária para a convergência de

[1] Aqui a palavra "adequado" se refere ao l_{min} que satisfaz da melhor forma possível as condições de contorno impostas, ou seja, o minimizador do parâmetro de assimetria.

método será chamada de N *crítico* e denotada por N_c. Vemos que quanto mais a série oscila, mais pontos ela precisa ter para que seja descrita dentro do calibre proposto pelas equações 4.2. Isso pode ser visto ao perceber que o N *crítico* é uma função que cresce de uma forma aparentemente monotônica com o número de oscilações da série, fazendo com que o critério proposto resulte em curvas diferentes de $\langle l_{min} \rangle \times N$ para cada número de períodos analisados.

Este comportamento ocorre não só para os casos mostrados, onde $\rho = \{1, 4\}$, mas para qualquer número de oscilações (figura 21), mesmo para ρ grande, como 10, 20 ou 30, por exemplo. Estes valores podem ser encontrados facilmente através da diferenciação em segunda ordem da curva $\langle l_{min} \rangle \times N$, onde a mudança do comportamento da inclinação da curva gera um mínimo na segunda derivada. Isso pode ser feito pois antes do N_c a curva $\langle l_{min} \rangle \times N$ se comporta aproximadamente de forma linear, e após, como a equação 5.1, desta forma

$$\frac{d^2}{dN^2} \langle l_{min} \rangle \approx \begin{cases} 0 & , \text{se } N < N_c \\ -\frac{1}{N^2} & , \text{se } N > N_c. \end{cases} \tag{5.2}$$

Como na prática a derivada de primeira ordem não é estritamente uma função degrau, a queda de 0 para $-1/N_c^2$ é algo similar à uma sigmoide (uma "versão contínua" da função degrau). A partir disso, podemos dizer que há um mínimo na derivada de segunda ordem pois ela vai continuamente[2] de 0 até $-1/N_c^2$ e, de volta para 0, à medida que N aumenta. O mínimo, então, está em $N = N_c$ (figura 22). Assim é possível encontrar a partir de qual valor de resolução podemos obter resultados dentro das restrições requeridas, tornando o cálculo do N_c acessível.

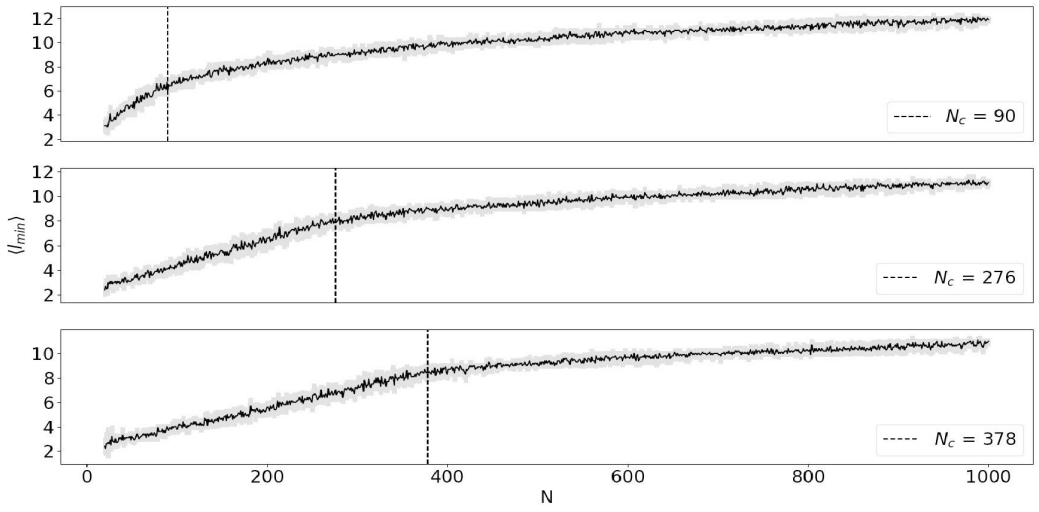


Figura 21 – Valores médios para os l_{min} em função da resolução da série para $\rho = \{1, 3, 4\}$ e seus respectivos N_c. O gráfico mais acima corresponde a uma oscilação, o do meio à 3 oscilações e o mais abaixo a 4 oscilações.

Mesmo que *a priori* isso não seja evidente, é possível mostrar que as curvas após o N_c são logarítmicas. Tal fato pode ser observado para quatro oscilações extrapolando a curva teórica 5.1 para até 10^4 pontos

[2] De forma aproximada, pois o domínio é discreto ($N \in \mathbb{N}$). Entretanto, como $\langle l_{min} \rangle \in \mathbb{R}$ e como há muitos valores de N podemos aproximar para algo contínuo com certa precisão.

com os parâmetros a e b obtidos a partir dos dados com $N \geq N_c$. O resultado obtido ao encontrar o l_{min} para $\{2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000\}$ pontos pode ser visto na figura 23. Tendo isso em mente, é possível modelar as curvas de l_{min} em função de N e ρ e obter mais rapidamente os valores, pois isso pode diminuir o tempo computacional oriundo da minimização do parâmetro de assimetria.

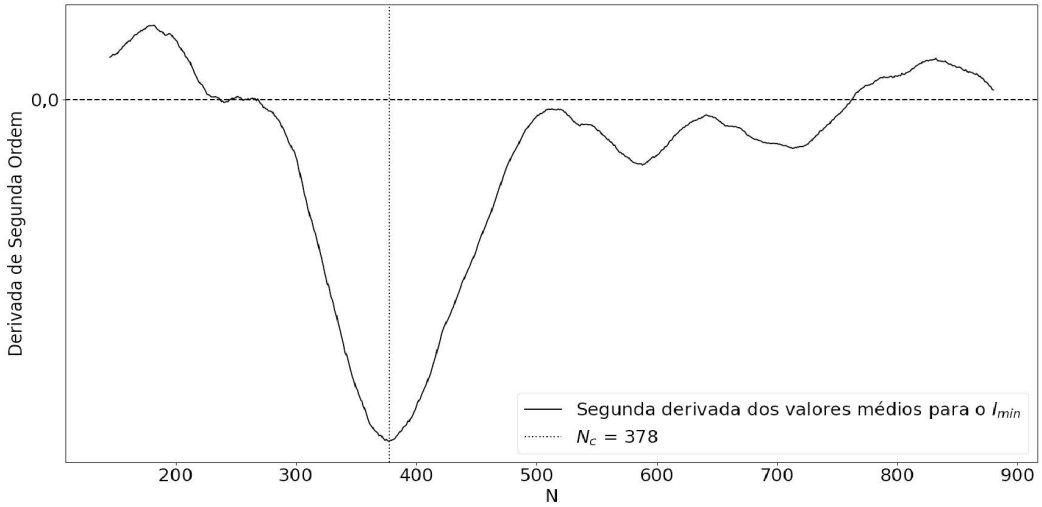


Figura 22 – Derivada de segunda ordem da curva $\langle l_{min} \rangle \times N$ para $\rho = 4$ oscilações. Confirma-se que realmente há um mínimo na segunda derivada no N_c. Figura obtida por meio de diferenciação numérica, o que resulta em perda de pontos no processo, neste caso, pontos finais e iniciais. Diferenciação realizada levando em conta 75 vizinhos para os cálculos dos coeficientes angulares das retas locais.

Vemos que de fato os determinismos das figuras 19 e 20 para valores menores do que o N_c possuem variâncias muito maiores do que para $N \geq N_c$. Isso, aliado ao fato de que os determinismos para $N \geq N_c$ possuem valores próximos dos que foram estipulados no sistema 4.2, corrobora com a hipótese de que N_c é de fato o número de pontos necessário para a análise. Para justificar a necessidade de ter um certo número de pontos para o critério de escolha proposto funcione, podemos usar o fato de que ele depende de duas séries *simultaneamente*. Como consequência, mesmo que possamos ver a estrutura periódica de um seno mesmo com poucos pontos disponíveis, nós ainda precisaríamos de amostras maiores para poder *distinguir* este seno de um ruído branco. Nós precisamos de mais pontos pois estamos definindo padrões com os quais podemos comparar outras séries, o que não seria possível com baixas resoluções. Este é um problema de *distinção* de séries. A figura 24 foi gerada ao obter os valores para N_c para cada número de períodos de um seno, por meio da minimização da segunda derivada de $\langle l_{min} \rangle \times N$, para vários números de oscilações. Vemos que o N_c se é crescente com uma lei de potência muito próximo ao linear.

Esta constatação concorda com o argumento do crescimento monotônico de $N_c(\rho)$, pelo menos neste intervalo de ρ. Para que seja possível tirar proveito deste conceito podemos dizer que 100 pontos por período/pseudo-período são o suficiente para uma análise adequada para séries que possuam de 1 a 20 períodos. Isso já seria o suficiente para fazer com que o sistema 4.2 seja satisfeito, mesmo que de forma aproximada.

Os determinismos para senos e ruídos brancos convergem para seus valores desejados por construção, ou seja, o método foi feito para isso. Contudo, podemos utilizar este critério de escolha do l_{min} para analisar outras séries. Notadamente serão analisadas um seno com 5 períodos, a componente x do sistema de

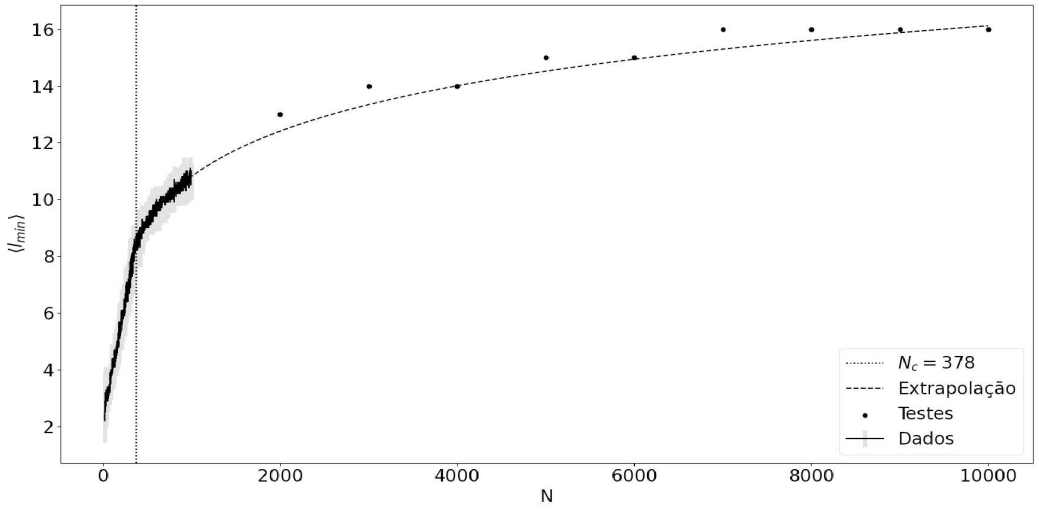


Figura 23 – Extrapolação do modelo logarítmico da curva da figura 20, evidenciando a concordância do modelo com os testes feitos. A curva extrapolada (tracejada) foi modelada apenas com os dados para $N \geq N_c$. Vemos que ela concorda com os testes realizados por meio da minimização do parâmetro de assimetria (pontos).

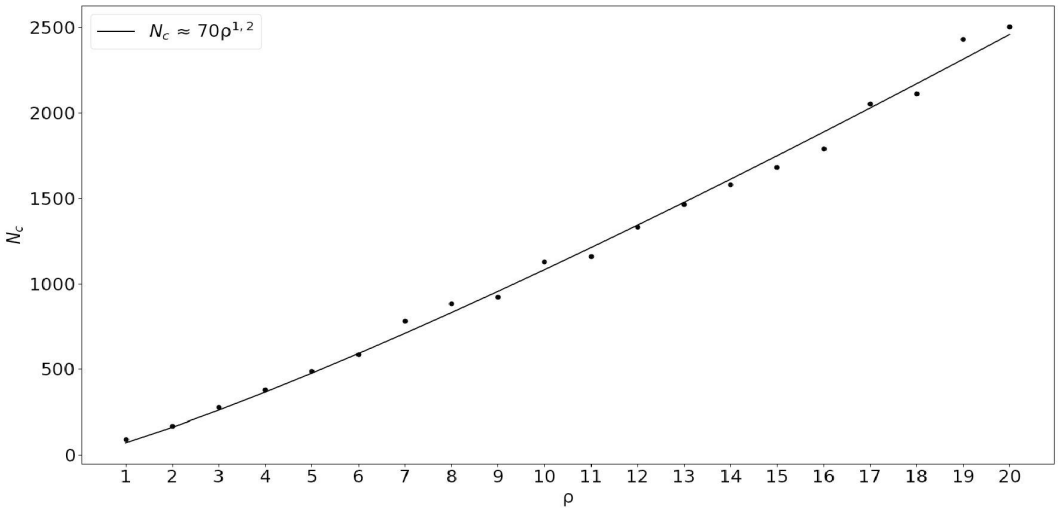


Figura 24 – N_c para algumas quantidades de períodos, mostrando uma relação do tipo lei de potência. Também pode ser possível utilizar um modelo linear, obteríamos, com isso, cerca de 125 pontos por períodos/pseudo-períodos.

Lorenz, componente x do oscilador de Chua e um ruído browniano. As análises subsequentes serão feitas de várias maneiras, sendo elas no tempo, com defasagem e na resolução, onde cada uma destas maneiras será brevemente descrita a seguir:

- **No tempo**: como o nome sugere, o determinismo da série é calculado para cada novo ponto adicionado

à série a partir de 20 pontos[3]. Temos assim um determinismo para os primeiros 20 pontos, outro determinismo para os primeiros 21 pontos, e assim por diante.

- **Com defasagem**: comumente chamada de "janela", a defasagem é algum subintervalo da série. Para este tipo de análise, a série será dividida em um certo número de janelas e um valor para o determinismo será encontrado para cada janela. Este método gera um número menor de pontos de determinismo.
- **Na resolução**: assim como o nome sugere, este tipo de análise é feita mantendo o intervalo temporal constante (diferindo da análise no tempo) e variando o número de pontos da série espalhados uniformemente sobre todo o intervalo temporal. Alterando assim a distância média entre os pontos na série.

5.1 Seno com 5 Períodos

Vimos que no critério proposto para a escolha do l_{min} são utilizados senos e ruídos brancos para calibrar, de certa forma, o determinismo da série a ser analisada. Isso nos diz que ao usá-lo para analisar um seno com ρ períodos, na verdade estamos apenas calibrando o l_{min} para ρ oscilações. O método vai convergir por construção. Entretanto, podemos analisar um seno como uma forma de fixar as maneiras de utilização do método e ver como ele se comporta. Nesta seção, o seno utilizado terá 1000 pontos e 5 períodos (figura 25). É importante ressaltar que o objetivo destas análises é apenas comparar os resultados obtidos pelos dois critérios de escolha do l_{min}, o atual e o proposto. Mesmo que este seja um caso onde não há componente aleatória, é ainda importante que façamos testes com sistemas dinâmicos artificiais como este para que seja possível verificar como o critério atual se comporta para apenas depois testá-lo em cenários mais realistas.

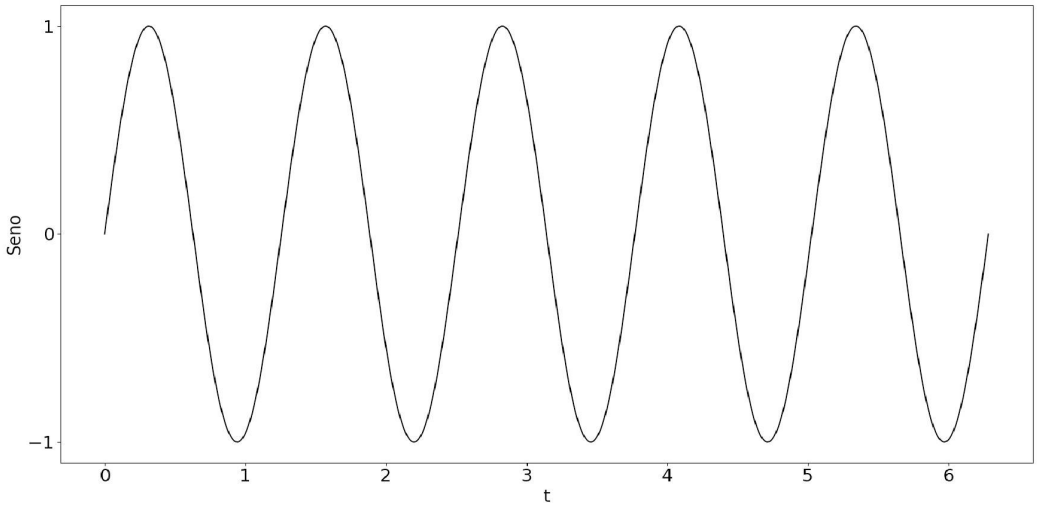


Figura 25 – Ilustração do seno que será analisado nesta seção. Será composto por 5 períodos e 1000 pontos.

[3] Uma quantidade menor de pontos torna o método inefetivo.

5.1.1 Análise no tempo

Ao pegar apenas os 20 primeiros pontos da série mostrada na figura 25 obtemos a figura 26, onde nela estão explícitos os pontos da série que serão utilizados para obter o primeiro valor do determinismo.

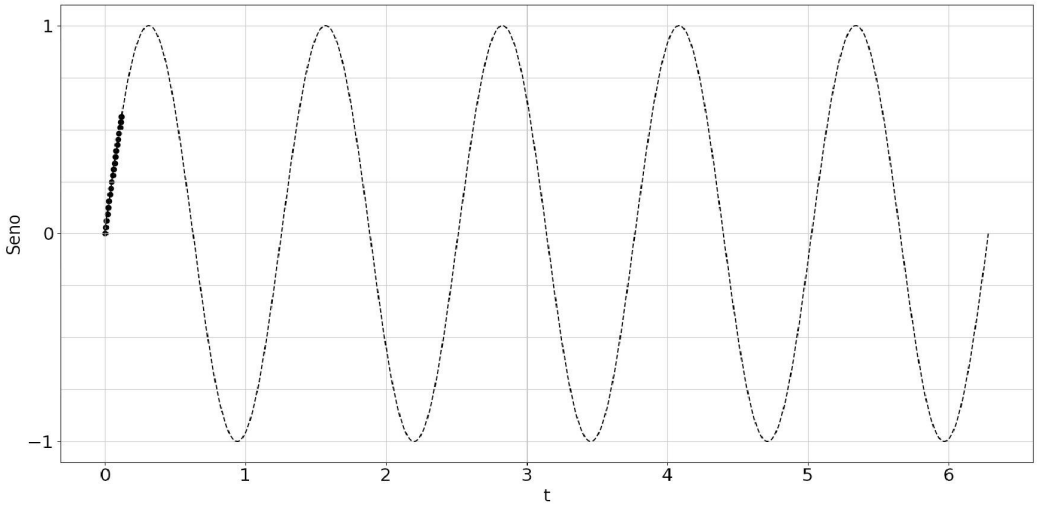


Figura 26 – Seno com 5 períodos e seus primeiros 20 pontos destacados no início da série. A cada iteração da série um ponto novo é adicionado, com os quais um novo determinismo é obtido, começando com os primeiros 20 pontos.

Calculando quantos pseudo-períodos a série composta por estes 20 pontos possui podemos gerar um seno com este mesmo número de pseudo-períodos e também com 20 pontos. Com isso encontra-se uma nova entropia de microestados máxima para este segmento e minimiza-se o parâmetro de assimetria. Desta forma, encontra-se um l_{min} adequado para esta configuração de (N, ρ). Feito isso, pode-se obter o determinismo apenas desta seção da série. Fazendo isto para 20 pontos até a extensão total da série obtemos a figura 27.

Vemos nela que os determinismos dos senos possuem valores muito próximos para ambos os critérios de escolha do l_{min}. Mesmo havendo certa variação dos valores do determinismo obtido com o critério proposto, esta variação é muito pequena para ser relevante. A configuração (N, ρ) é utilizada como base para o cálculo dos determinismos, como esta configuração varia ponto a ponto, o l_{min} varia, fazendo assim com que o determinismo também sofra pequenas alterações. Entretanto, vemos o benefício do método proposto nos determinismos de ruídos brancos, dados pelas duas curvas de determinismo com os valores menores. Como os determinismos para os senos não possuem quase nenhuma diferença, podemos concluir que o método proposto é capaz de encontrar os parâmetros l_{min} mais adequados, levando em conta suas restrições.

Podemos ver de forma mais clara como o processo de escolha do l_{min} ocorreu (figura 28), pois para cada trecho da série há um valor correspondente para N e para ρ. O número de pseudo-períodos foi estimado através de uma variação de periodograma chamada periodograma de Lomb-Scargle[4] [28, 78]. Este método de determinação do número de pseudo-períodos foi idealizado para encontrar o valor dos períodos de séries com amostragem não uniforme, ou seja, com os valores da série possuindo espaçamentos temporais diferentes. Entretanto ele é robusto o suficiente para ser utilizado nestes casos. A lista de $\langle l_{min} \rangle$ na figura 28 não é

[4] Este foi a maneira escolhida para determinar a quantidade ρ de pseudo-períodos neste caso, porém, isso pode ser feito da forma que o pesquisador achar melhor.

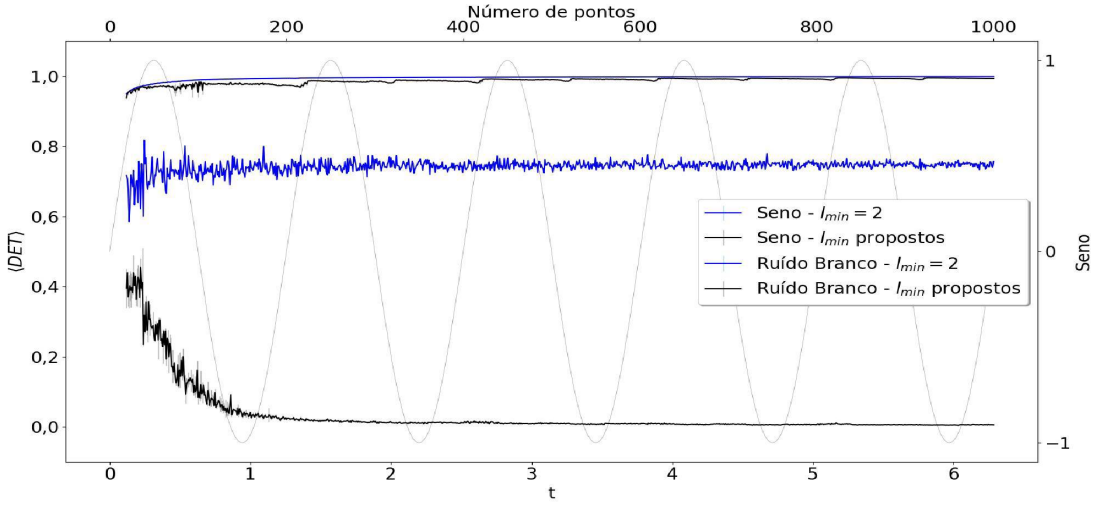


Figura 27 – Curvas de determinismo calculados ao longo do tempo para o seno com 5 períodos e ruído branco. A curva cinza no fundo serve como guia e ilustra até que ponto da série foi utilizado para os determinismos correspondentes. As curvas em azul são os determinismos de um seno e de um ruído branco obtidos com $l_{min} = 2$. As curvas em preto são os determinismos das mesmas séries, porém, com o método proposto. As duas curvas com valores mais altos correspondem aos senos e as duas com valores mais baixos aos ruídos.

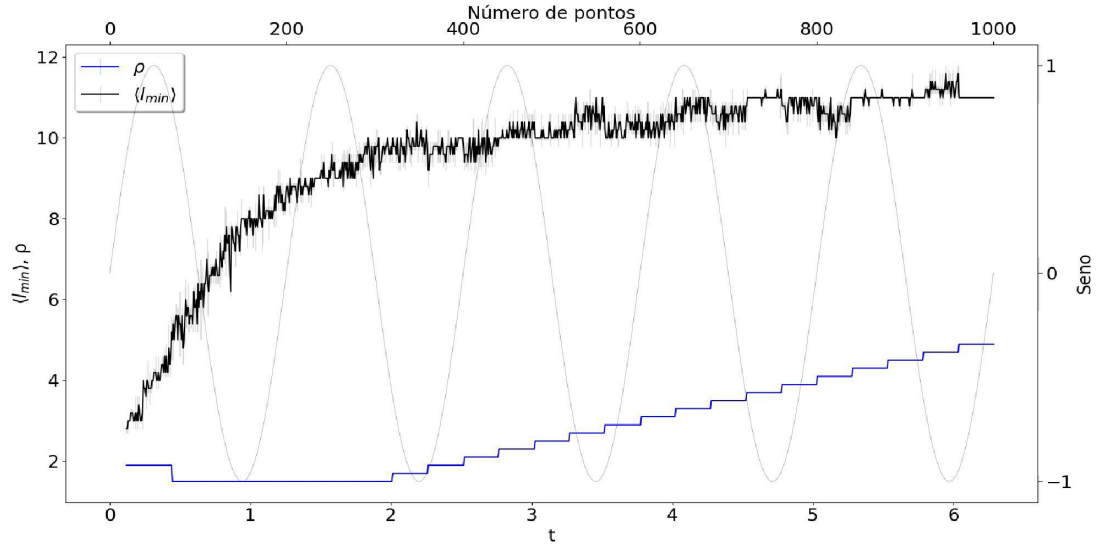


Figura 28 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados para o cálculo dos determinismos do seno com 5 períodos no tempo. A curva cinza no fundo serve como guia para ilustrar qual seção da série foi utilizada para o cálculo de l_{min} e ρ.

exatamente igual as listas mostradas nas figuras 19, 20 e 21 pois neste caso não há um valor fixo para ρ. Contudo, ela pode ser considerada uma combinação das curvas de $\langle l_{min} \rangle$ para diferentes ρ.

5.1.2 Análise com defasagem

Nesta subseção o mesmo seno com 5 períodos mostrado na figura 25 será analisado, porém, de forma "janelada". Em vez de estudarmos como o determinismo se comporta ao evoluir a série no tempo, estudaremos como o l_{min} escolhido de forma dinâmica resultará em determinismos para certas janelas (ou defasagens) da série. De forma análoga à figura 26, que foi utilizada para ilustrar como os determinismos foram obtidos na subseção anterior, a figura 29 ilustra como e quais pontos serão utilizados na análise subsequente. Todas as janelas serão compostas por 100 pontos e a próxima sempre começará 25 pontos após o ponto inicial da anterior, em outras palavras, a janela se desloca 25 pontos por vez.

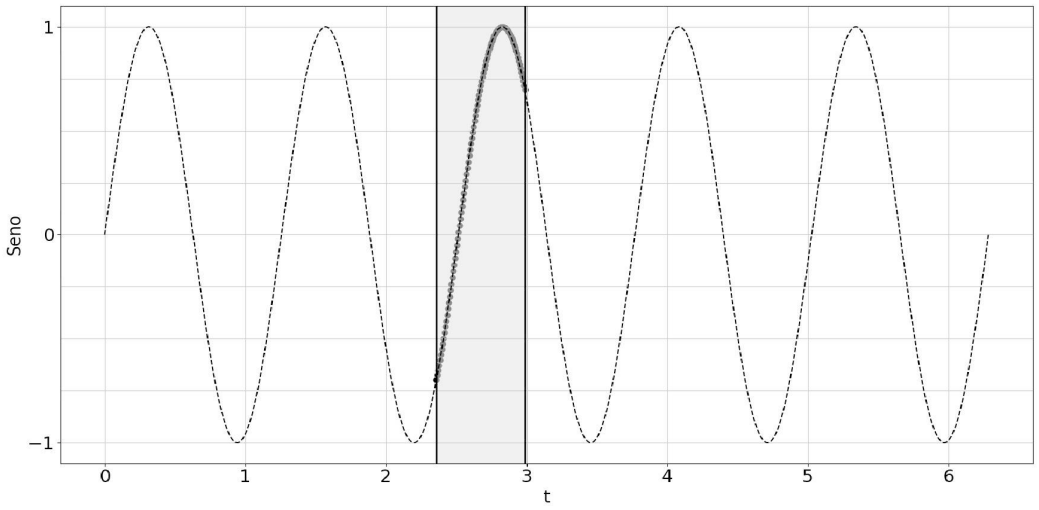


Figura 29 – Seno com 5 períodos e uma janela, evidenciando quais pontos serão necessários para analisar a janela em questão. As janelas serão compostas por 100 pontos e a próxima sempre começará 25 pontos após o ponto inicial da anterior, em outras palavras, a janela se desloca 25 pontos por vez.

Com isso, podemos estimar o número de pseudo-períodos[5] ρ para cada janela, começando com o conjunto de pontos $[1, 100]$ e obter um ε maximizando a entropia de microestados para cada janela. Com isso pode-se minimizar o parâmetro de assimetria, obtendo assim um l_{min}, consequentemente calculando um determinismo para cada uma das janelas.

A lista de $\langle l_{min} \rangle$ gerada e as estimativas de ρ para cada janela podem ser vistas na figura 31.
Vemos que as estimativas para o número de pseudo-períodos da série em cada janela possui uma certa sazonalidade, onde ora $\rho = 2$ e ora $\rho = 1,5$. Ao levar o número de pontos de cada janela em consideração, os l_{min} ficaram entre 5 e 7, lembrando que os pontos na imagem 31 representam as médias de l_{min}. Com isso, conclui-se novamente que, para séries senoidais, o critério proposto de escolha do l_{min} não causa desvios significativos nos determinismos e, adicionalmente, consegue gerar o resultado desejado. Simultaneamente consegue-se que um ruído branco com o mesmo número de pontos que a série tenha um determinismo baixo, satisfazendo assim, as restrições impostas inicialmente.

[5] Neste caso não será utilizado o termo "período" pois as janelas não conseguem comportar um período inteiro da série.

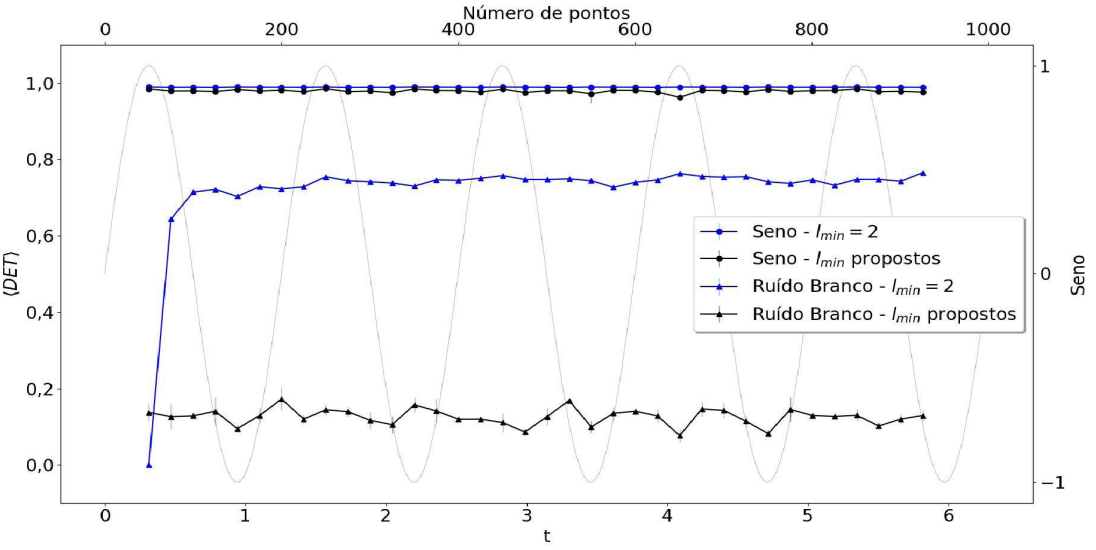


Figura 30 – Determinismos calculados para cada janela para um seno com 5 períodos e um ruído branco. A curva cinza no fundo serve como guia. Os determinismos estão situados no centro da janela de 100 pontos na qual eles foram calculados. Os pontos em azul são os determinismos de um seno e de um ruído branco obtidos com $l_{min} = 2$. Os pontos em preto são os determinismos das mesmas séries, porém, com o método proposto. As duas curvas com valores mais altos correspondem aos senos e as duas com valores mais baixo aos ruídos.

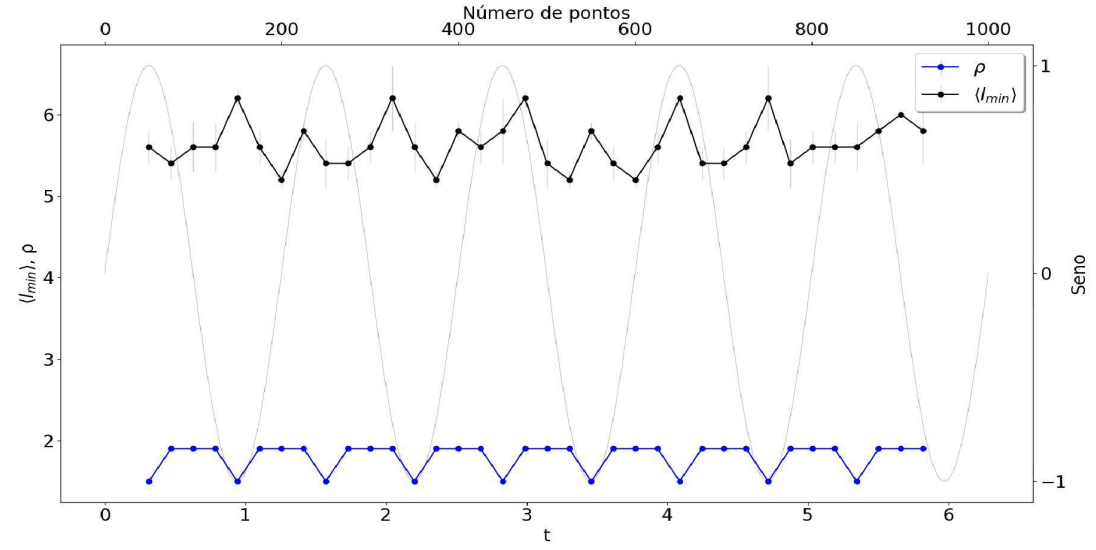


Figura 31 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados para o cálculo dos determinismos do seno com 5 períodos para cada janela. A curva cinza no fundo serve como guia para ilustrar quais as posições das janelas na série foram utilizadas para os cálculos de l_{min} e ρ. Cada ponto está no centro de sua respectiva janela.

5.1.3 Análise na resolução

Para este tipo de análise, o intervalo temporal é mantido constante durante todo o processo, mas o número de pontos será gradualmente aumentado ao longo de todo o intervalo temporal, fazendo com que a distância

média entre os pontos diminua. Os pontos sempre preenchem toda a extensão da série, não apenas um determinado segmento, como no caso da análise no tempo. A figura 32 mostra quais são os primeiros 20 pontos do seno que serão utilizados na análise nesta subseção, ilustrando a diferença entre este método e a análise no tempo. Esta não é uma maneira usual de análise de séries temporais, porém, esta análise tem como objetivo verificar as consequências da alteração da resolução no processo de escolha do novo l_{min}. Em outras palavras, as diferenças entre os resultados obtidos com o critério atual e o proposto são os pontos mais importantes e estas diferenças que são os objetos de interesse.

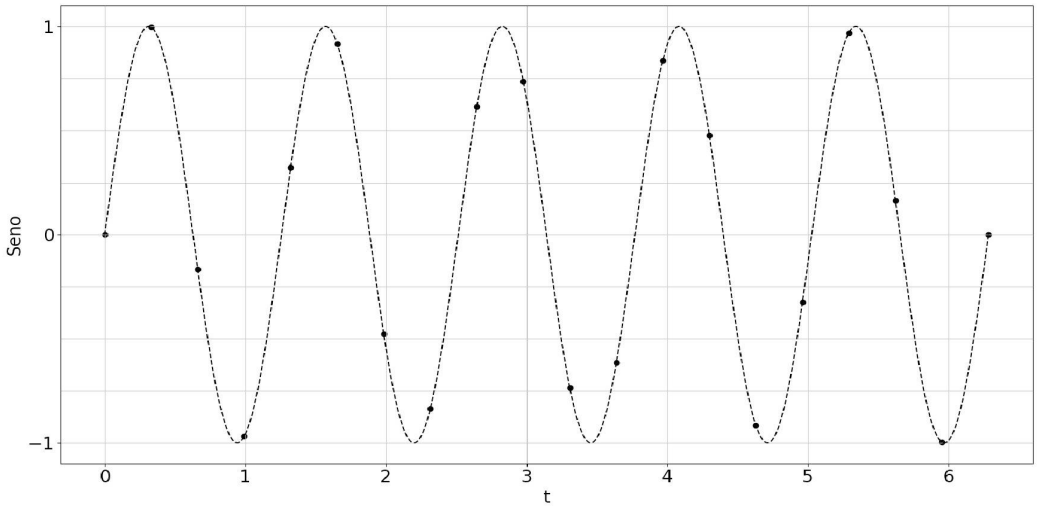


Figura 32 – Seno com 5 períodos com uma resolução de 20 pontos, evidenciando quais pontos serão necessários para análise e qual a curva subjacente da série em questão.

A figura 33 nos mostra como o determinismo se comporta ao manter o intervalo temporal constante enquanto a resolução da série é alterada. Vemos que para este tipo de análise precisamos de mais pontos para poder descrever a série em questão de forma correta.
Isso ocorre pois os pontos estão dispostos por toda a extensão da curva, fazendo com que a estimativa de ρ seja quase sempre constante para este caso, implicando em um único N_c para a série como um todo. A figura 34 mostra quais os $\langle l_{min} \rangle$ e as estimativas para ρ foram utilizados para cada resolução, confirmando o argumento sobre a constância de ρ.
A lista de $\langle l_{min} \rangle$ na figura 34 tenderá ao mesmo comportamento das listas vistas em 21 para um número maior de cálculos de l_{min} para cada valor de resolução. Entretanto, nesta imagem, foram amostrados poucos l_{min} por N, fazendo com que a curva não ficasse tão suave. Vemos que neste caso houve o mesmo fenômeno observado nas outras duas subseções anteriores, os determinismos obtidos com os l_{min} propostos satisfazem as restrições impostas, na medida do possível, a partir de uma certa resolução. Os determinismos obtidos com $l_{min} = 2$ não as satisfazem sequer aproximadamente, mantendo-se altos para ambas as séries. Com isso, conclui-se que, para séries senoidais, o método proposto é capaz de distinguir muito mais as séries através de seus determinismos sem comprometer resultados já conhecidos. Podemos extrapolar estes resultados para séries periódicas com baixos níveis de ruído, ou seja, *SNR* (*signal to noise ratio*) alto.

Nas próximas análises os determinismos dos ruídos brancos obtidos com l_{min} igual à dois serão omitidos para diminuir a poluição visual nos gráficos subsequentes e pelo motivo de que estas curvas poderão ser vistas

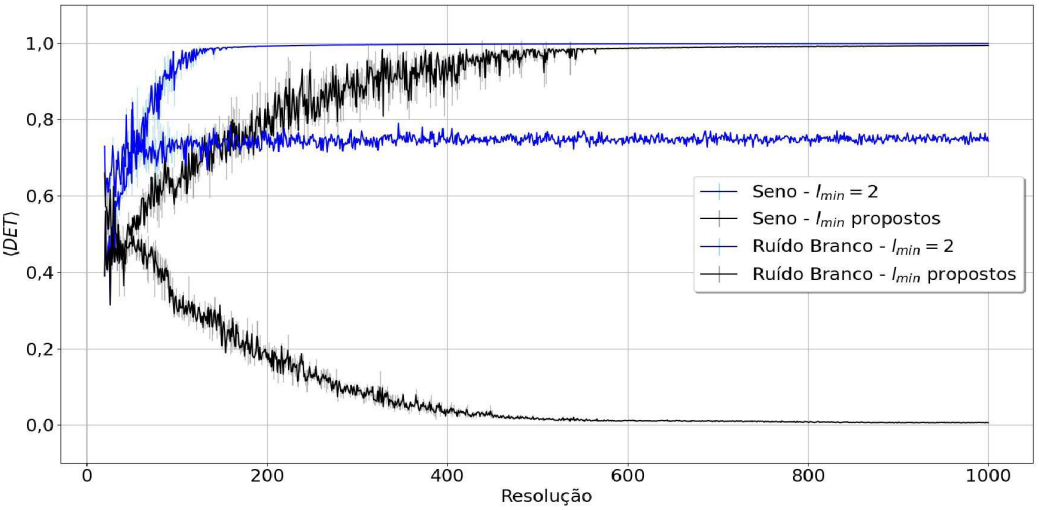


Figura 33 – Determinismos para um seno com 5 períodos e um ruído branco em função de sua resolução. As curvas em azul correspondem aos determinismos obtidos com $l_{min} = 2$. As curvas em preto correspondem aos determinismos obtidos com o critório proposto. As duas curvas com os valores mais altos foram obtidas para os senos. As duas curvas com os valores mais baixos representam os determinismos para os ruídos.

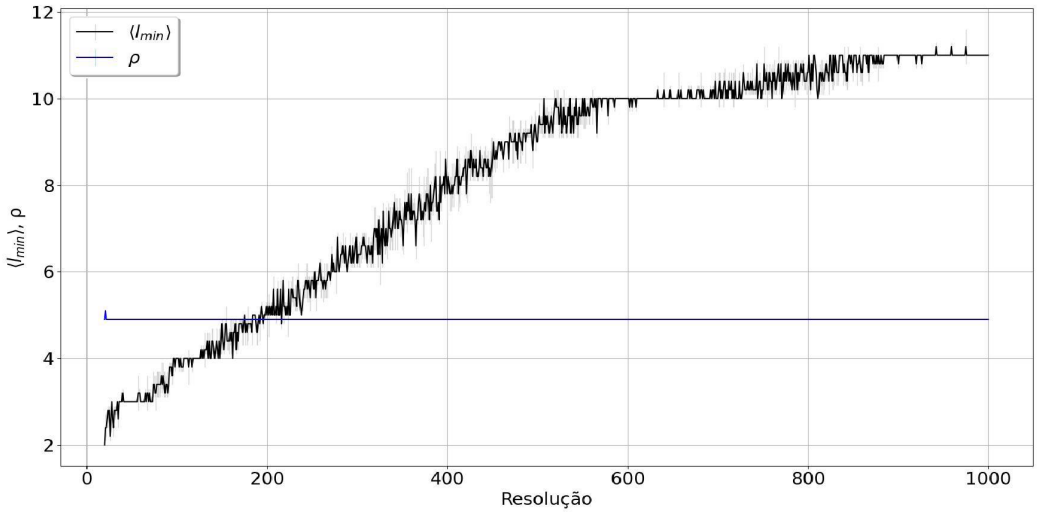


Figura 34 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada resolução para o cálculo dos determinismos do seno com 5 períodos.

nas imagens 27, 30 e 33.

5.2 Componente x do Sistema de Lorenz

O sistema de Lorenz consiste em 3 equações diferenciais ordinárias não lineares proposto em 1963 por Edward Lorenz [79] com o objetivo de servir como um simples modelo para dinâmica atmosférica dado pelas equações

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\varrho - z) - y, \\ \dot{z} &= xy - \beta z. \end{aligned} \tag{5.3}$$

A série analisada nesta subseção será a componente x do sistema de Lorenz, ilustrada na figura 35. A variável x é proporcional ao movimento convectivo da atmosfera, a variável y é proporcional à diferença de temperatura entre a fração de fluido descendendo e a fração ascendendo por convecção, e a variável z é proporcional à distorção do perfil linear da temperatura na vertical do sistema.

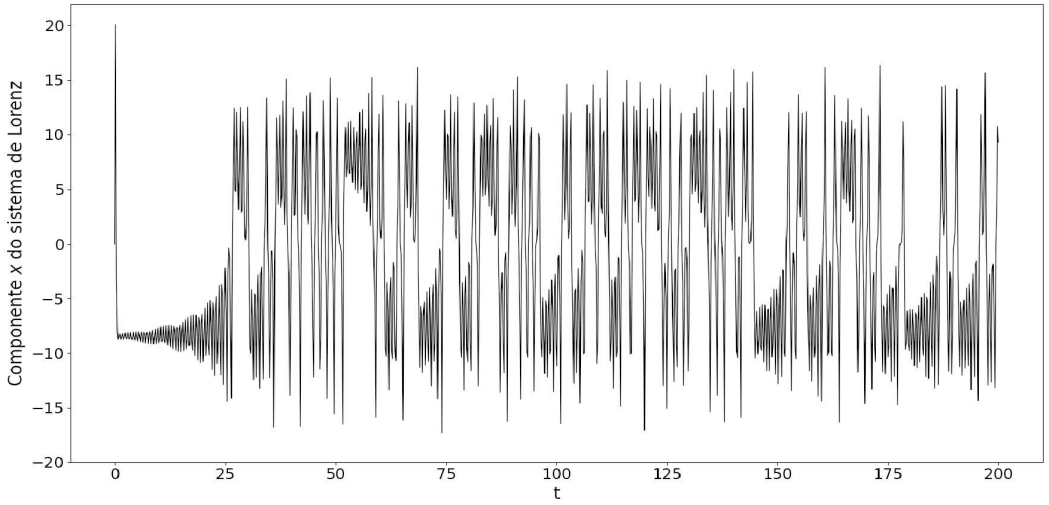


Figura 35 – Componente x do sistema de Lorenz com condições iniciais $(0, 10, 0)$ em sua região caótica mais conhecida. Parâmetros são $(\sigma, \varrho, \beta) = (10, 20, 8/3)$.

Este sistema de equações possui três pontos fixos $\vec{x}_i^*$ (com i sendo $\{1, 2, 3\}$). Todos eles são iguais a zero para $\varrho \leq 1$ (neste caso ϱ faz referência ao parâmetro do sistema 5.3, não ao número de pseudo-períodos da série) e, para $\varrho \geq 1$, dados em função dos parâmetros como mostrado nas equações

$$\begin{aligned} \vec{x}_1^* &= (0, 0, 0), \\ \vec{x}_2^* &= \left(\sqrt{\beta(\varrho - 1)}, \sqrt{\beta(\varrho - 1)}, \varrho - 1\right), \\ \vec{x}_3^* &= \left(-\sqrt{\beta(\varrho - 1)}, -\sqrt{\beta(\varrho - 1)}, \varrho - 1\right). \end{aligned} \tag{5.4}$$

As condições iniciais utilizadas nas análises a seguir serão $(x_0, y_0, z_0) = (0, 10, 0)$, resultando em pontos fixos diferentes da condição inicial, garantindo uma dinâmica não trivial. Estes parâmetros também garantem que estaremos analisando a série em uma região caótica. Nesta subseção, a série utilizada mostrada na figura 35 terá 2000 pontos.

5.2.1 Análise no tempo

Esta análise é executada de modo completamente análogo ao que foi feito para o seno.

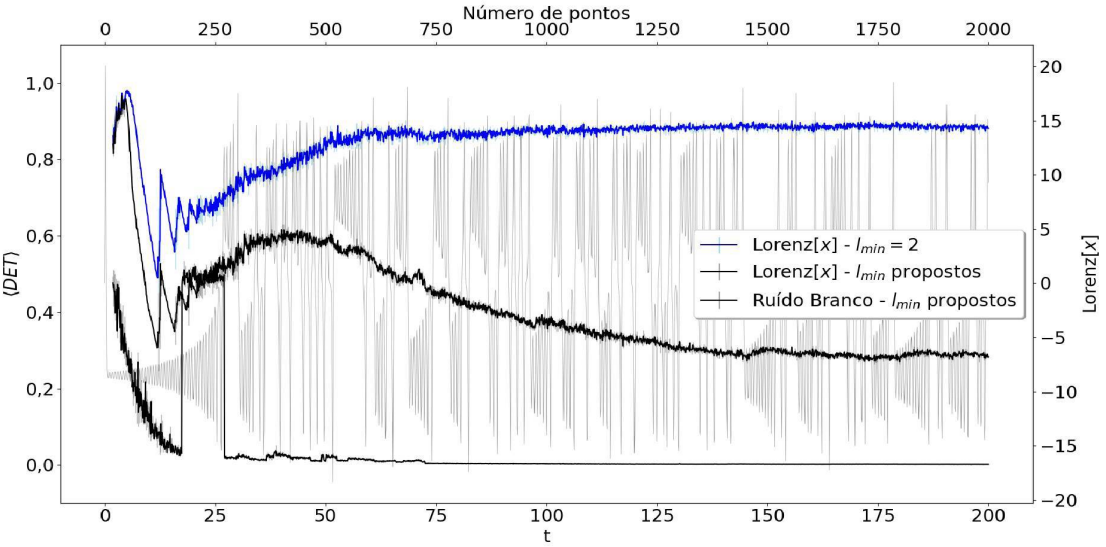


Figura 36 – Determinismos calculados ao longo do tempo para a componente x do sistema de Lorenz e ruído branco. A curva cinza serve como guia e ilustra até que ponto da série foi utilizado para os determinismos correspondentes. Em azul, os determinismos da componetne x do sistema de Lorenz obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco.

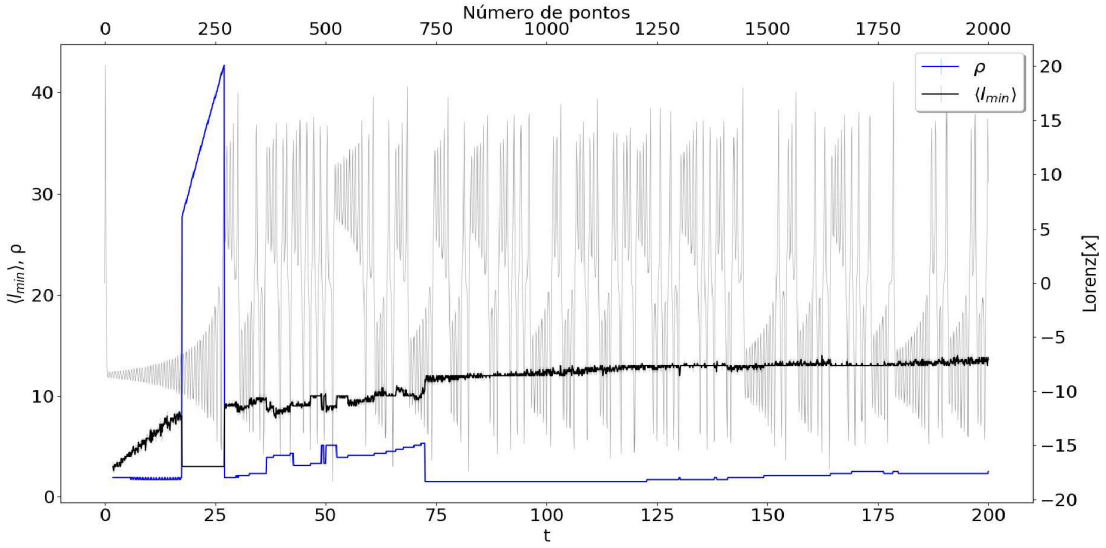


Figura 37 – Valores de $\langle l_{min} \rangle$ obtidos via o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada ponto da série para o cálculo dos determinismos da componente x do sistema de Lorenz no tempo. A curva cinza serve como guia para ilustrar qual seção da série foi utilizada para o cálculo de l_{min} e ρ.

A série em questão é evoluída no tempo a partir dos 20 pontos com t começando em 0 e gerado um novo determinismo correspondente a cada novo ponto.

Vemos uma diferença clara nos valores do determinismo pouco antes dos 250 pontos (figure 36), quando a série deixa de ser algo próximo de um seno com amplitude crescente e passa a se comportar de forma mais

caótica. Nesse momento, os determinismos passam a tender aos valores $0,85$ e $0,3$ para $l_{min} = 2$ e com os l_{min} propostos, respectivamente. Na região dos 175 até aos 260 pontos, aproximadamente, vemos que há um aumento abrupto nos valores do determinismo de um ruído branco com os l_{min} propostos. Isso significa que, para este número de pontos da série, há uma estimativa grande de pseudo-períodos em comparação com o número de pontos.

O que ocorre na faixa dos $175 - 260$ pontos é uma estimativa diferente do número de pseudo-períodos nesta faixa (figura 37), o que confirma que esta é uma etapa importante do processo e deve ser feita com cautela. Nesta subseção também foi observado uma distinção muito clara nos determinismos com ambos os critérios de escolha do l_{min}, mostrando mais uma vez que o proposto apresenta algumas vantagens em relação ao atual.

5.2.2 Análise com defasagem

Para a análise com defasagens foi escolhido uma janela também de 100 pontos com um incremento de 25 pontos entre o início de duas consecutivas, assim como na seção anterior. Lembrando que a escolha do tamanho e da distância entre o inícios de duas janelas consecutivas é extremamente importante e pode alterar os resultados. Ao utilizar as defasagens mencionadas e o obter os determinismos para cada uma delas foi gerada a figura 38.

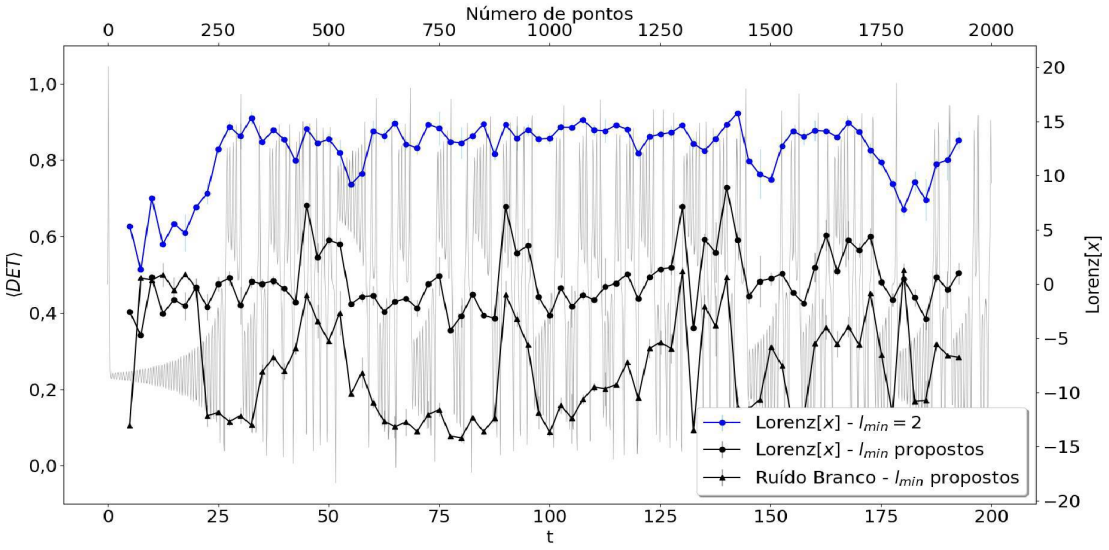


Figura 38 – Determinismos calculados para cada janela para a componente x do sistema de Lorenz e um ruído branco. A curva cinza ao fundo serve como guia. Os determinismos estão situados no centro da janela de 100 pontos na qual eles foram calculados. Em azul, os determinismos da componente x do sistema de Lorenz obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco, obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco.

Podemos observar o comportamento dos l_{min} utilizados e quais as estimativas da quantidade de pseudo-períodos em cada janela (figura 39). Há uma correlação visível entre o determinismo de um ruído branco com os l_{min} propostos e as estimativas de número de pseudo-períodos, o que indica que a estimativa de ρ talvez possa ser utilizada de outras formas nestas análises. Algo interessante que ocorre, entretanto, é o fato de que os l_{min} propostos sempre serem pequenos, não diferindo muito dos valores utilizados atualmente, como 2 ou

5. Isso é indício de que não há prejuízo em utilizar tal método. Vemos na figura 39 que regiões que possuem valores de N e ρ abaixo do N_c possuem determinismos mais altos para o ruído branco. Isso nos mostra que regiões onde $N \geq N_c$ são de fato importantes para uma correta utilização do método proposto.

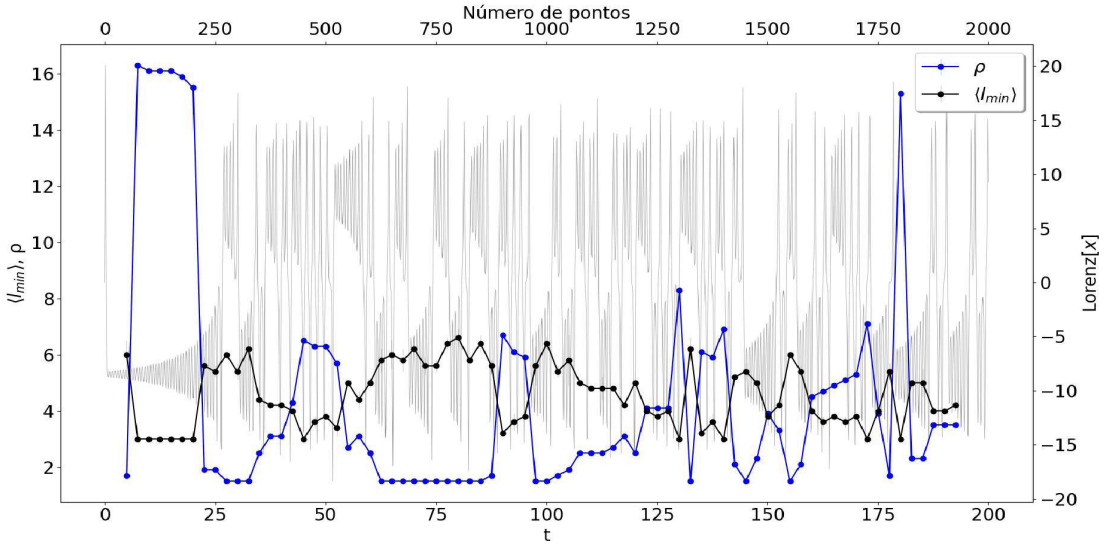


Figura 39 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada janela para o cálculo dos determinismos da componente x do sistema de Lorenz para cada janela. A curva cinza ao fundo serve como guia para ilustrar quais as posições das janelas na série foram utilizadas para os cálculos de l_{min} e ρ. Cada ponto está no centro de sua respectiva janela.

5.2.3 Análise na resolução

Ao utilizar um pacote computacional que permitisse gerar uma solução densa de um sistema de equações diferenciais [80, 81], é possível obter uma curva subjacente contínua como trajetória em um espaço de estados. Isso nos permite selecionar apenas a componente x, obtendo assim, uma série temporal em que é possível variar sua resolução sem alterar seu formato. Isso foi feito variando o número de pontos dentro da curva gerada de 20 até 2000 pontos (figura 40).

Os determinismos do ruído branco e da série analisada obtidos com os l_{min} propostos são praticamente os mesmos para baixas resoluções e se distinguem cada vez mais ao passo que a resolução aumenta. É importante salientar também que, desta maneira, o determinismo da componente x do sistema analisado obtido com $l_{min} = 2$ possui valores muito próximos do determinismo de um ruído branco para resoluções muito maiores. Para 1000 pontos na série eles diferem por menos de 0.1. As estimativas do número de pseudo-períodos e os respectivos $\langle l_{min} \rangle$ podem ser vistos na figura 41.

Percebe-se que até aproximadamente os 250 pontos de resolução a estimativa de pseudo-períodos é muito alta em comparação ao número de pontos, fazendo com que esta região esteja atrás do N_c, ou seja, não há pontos o suficiente para dizer algo relevante sobre a série.

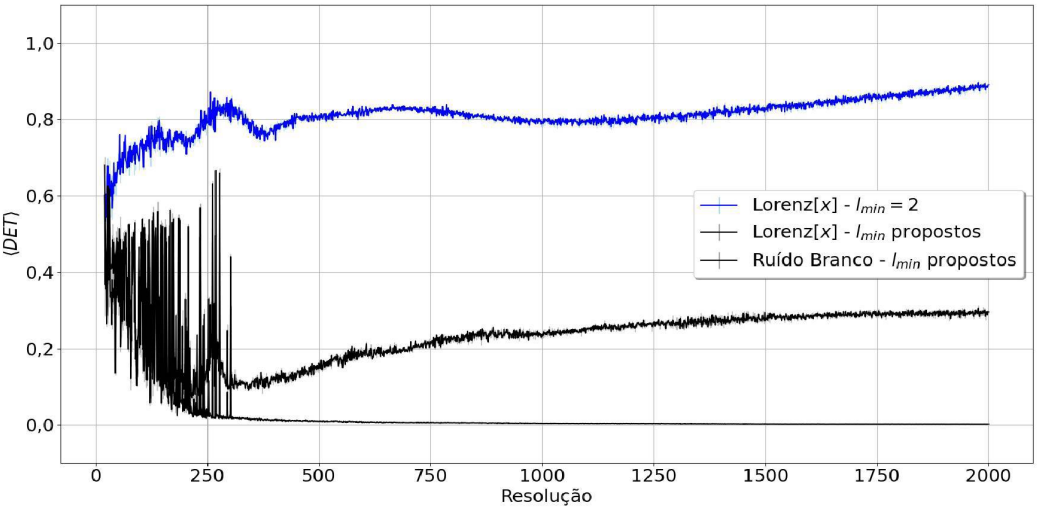


Figura 40 – Determinismos para a x do sistema de Lorenz e um ruído branco obtidos em função de sua resolução. A curva em azul corresponde aos determinismos da componente x do sistema de Lorenz obtidos com $l_{min} = 2$. As curvas em preto correspondem aos determinismos obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco.

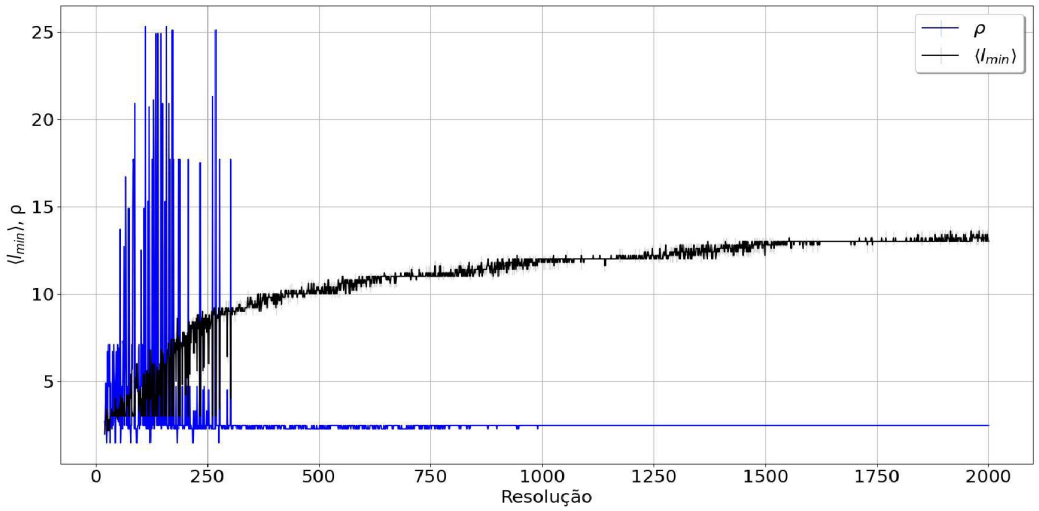


Figura 41 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de pseudo-períodos (azul) utilizados para cada resolução no cálculo dos determinismos da componente x do sistema de Lorenz.

5.3 Componente x do Oscilador de Chua

Este oscilador foi proposto na forma de um circuito eletrônico na década de 80 com o objetivo de mostrar que era possível observar caos em sistemas reais, e não apenas em equações abstratas [15], como pensavam na época. Foi verificado que caos está, de fato, presente neste circuito [82]. Este oscilador é dado pelas equações

$$\begin{aligned} \dot{x} &= \alpha(y - h(x)), \\ \dot{y} &= x - y + z, \\ \dot{z} &= -\beta y, \end{aligned} \tag{5.5}$$

onde $h(x)$ é uma função não linear dada por

$$h(x) \coloneqq m_1 x + \frac{1}{2}(m_0 - m_1)[|x+1| - |x-1|].$$

A função $h(x)$ representa a relação suave por partes entre a corrente em um resistor não linear em função da tensão elétrica no mesmo. A série analisada nesta subseção será a primeira componente deste sistema e pode ser vista na figura 42.

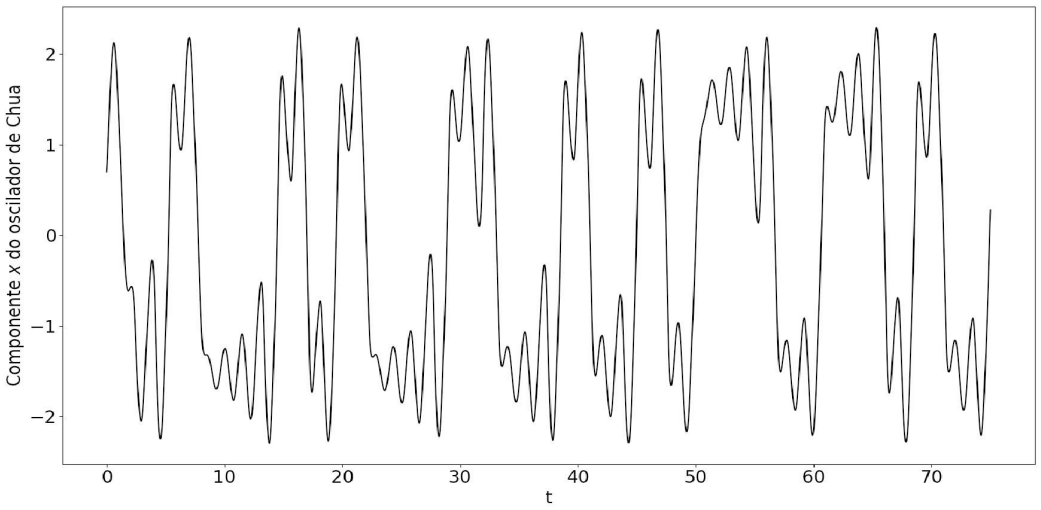


Figura 42 – Componente x do oscilador de Chua com condições iniciais $(7/10, 0, 0)$ em uma região caótica. A série possui 2000 pontos no intervalo temporal de 0 a 75 unidades e possui parâmetros ($\alpha = 156/10, \beta = 2558/100, m_0 = -8/7, m_1 = -5/7$).

Os parâmetros α, β, m_0 e m_1 representam capacitâncias, resistências e combinações delas no circuito gerador do sistema.

5.3.1 Análise no tempo

Para a análise no tempo da componente x do oscilador de Chua usaremos a mesma metodologia, começando com os 20 primeiros pontos até a totalidade da série (figura 43). Vemos nitidamente que os determinismos em azul, correspondendo à componente x do oscilador de Chua obtidos com $l_{min} = 2$, possuem valores muito próximos de 1, de forma muito similar aos determinismos de
um seno (figura 27). Isso nos mostra que, com o critério atual de escolha do l_{min}, é impossível distinguir séries completamente diferentes em algumas situações. Isso, contudo, é possível com o critério proposto. Ele mantém o determinismo da série analisada consistentemente abaixo de 1, sendo aproximadamente 0,8 para resoluções maiores que 250 pontos, ao passo que também mantém o determinismo de um ruído branco em 0. Os $\langle l_{min} \rangle$ e as estimativas de pseudo-períodos ρ podem ser vista na figura 44. Percebemos a mesma ordem

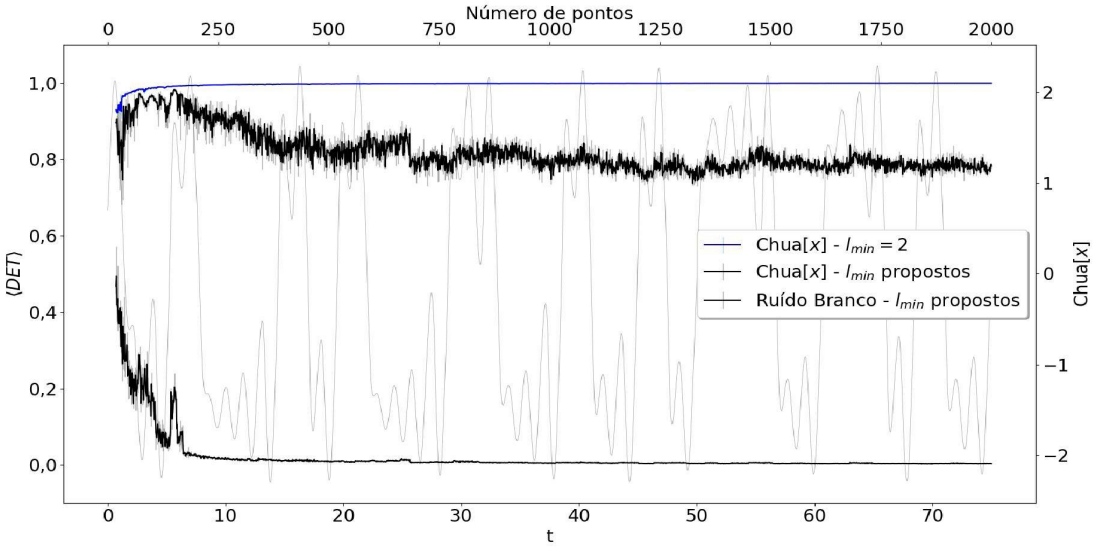


Figura 43 – Determinismos calculados ao longo do tempo para a componente x do oscilador de Chua e ruído branco. A curva cinza serve como guia e ilustra até que ponto da série foi utilizado para os determinismos correspondentes. Em azul, os determinismos da componetne x do oscilador de Chua obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco.

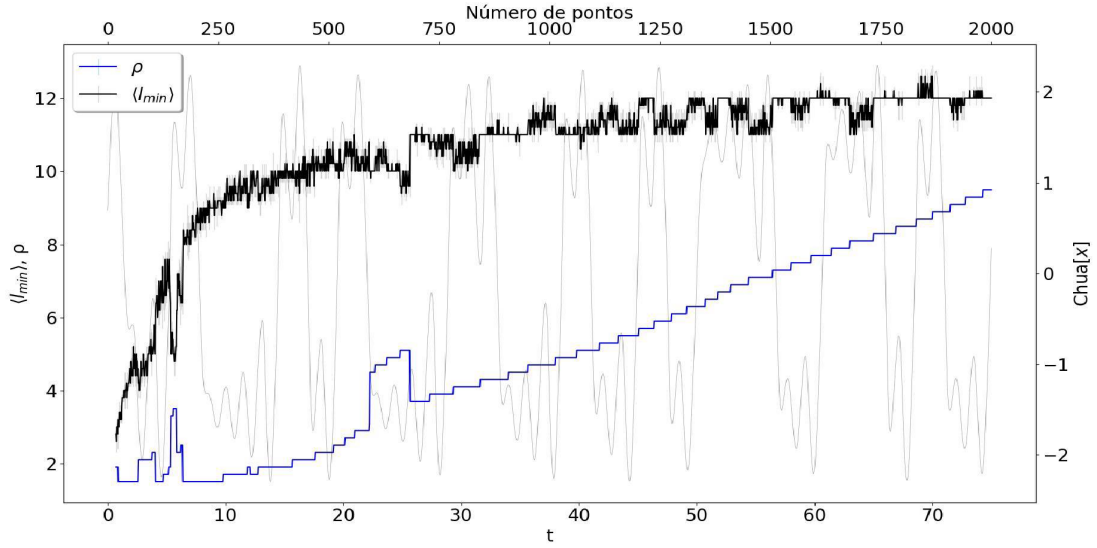


Figura 44 – Valores de $\langle l_{min} \rangle$ obtidos via o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada ponto da série para o cálculo dos determinismos da componente x do oscilador de Chua no tempo. A curva cinza serve como guia para ilustrar qual seção da série foi utilizada para o cálculo de l_{min} e ρ.

de grandeza nos $\langle l_{min} \rangle$, que não chegam a passar de 13, onde a estimativa de ρ nos mostra que foi estimado que a série tem cerca de 10 pseudo-períodos quando todos os 2000 pontos foram analisados. Conclui-se que o critério proposto tem vantagem ao atualmente utilizado neste tipo de análises, pois possibilita uma melhor distinção entre diferentes séries.

Neste sistema vemos que o número de pseudo-períodos, visto como a curva em azul na figura 44, lembra muito o que foi visto na figura 28. Isso ocorre pelo fato de as oscilações da série da figura 42 serem relativamente regulares, fazendo com que recorrências ocorram com mais frequência. Com isso, o número de linhas diagonais dos gráficos de recorrências gerados a cada ponto aumenta, resultando em um aumento do determinismo.

5.3.2 Análise com defasagem

Para a análise com defasagens são utilizadas janelas com 100 pontos que caminharão 25 pontos a cada iteração, assim como nos casos anteriores (figura 45).

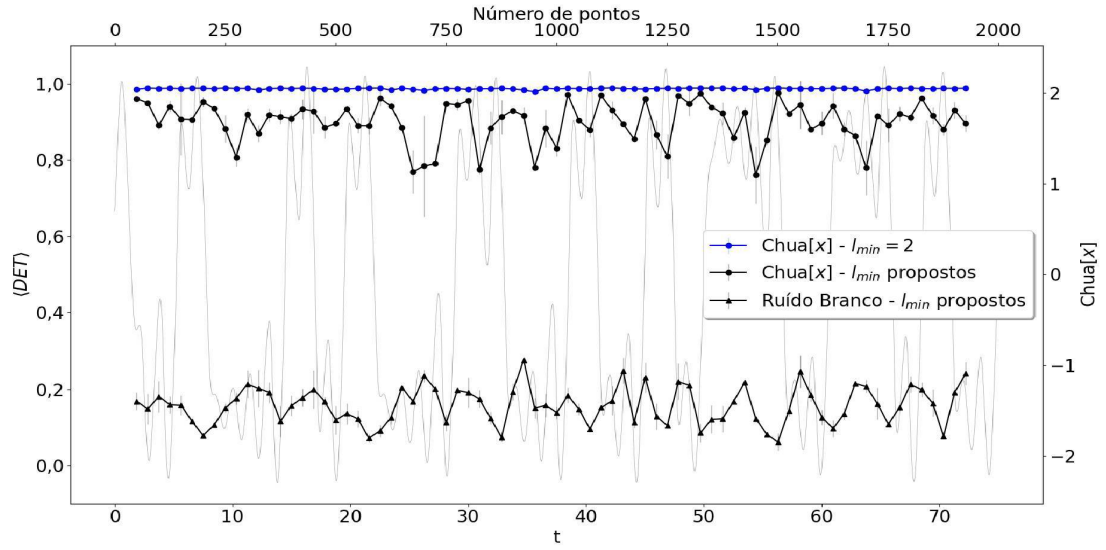


Figura 45 – Determinismos calculados para cada janela para a componente x do oscilador de Chua e um ruído branco. A curva cinza ao fundo serve como guia. Os determinismos estão situados no centro da janela de 100 pontos na qual eles foram calculados. Em azul, os determinismos da componente x do oscilador de Chua obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco, obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco.

Aqui também é possível ver que o critério atual não é capaz de diferenciar esta série e oscilações senoidais simples. Entretanto, sabemos que esta série está em um regime caótico e que a premissa da análise de quantificação de recorrência com o determinismo é ser capaz de distinguir os regimes dinâmicos de um sistema. Isto não acontece com $l_{min} = 2$. Os valores dos determinismos são essencialmente os mesmos de um seno, que é uma série periódica. Mais uma vez, conclui-se que o critério proposto é superior na distinção de séries, mesmo quando utilizamos defasagens nas análises. Podemos ver os $\langle l_{min} \rangle$ utilizados e as estimativas de ρ na figura 46.

5.3.3 Análise na resolução

Ao utilizar o mesmo pacote computacional empregado no sistema de Lorenz foi possível realizar análises variando a resolução desta série (figura 47).

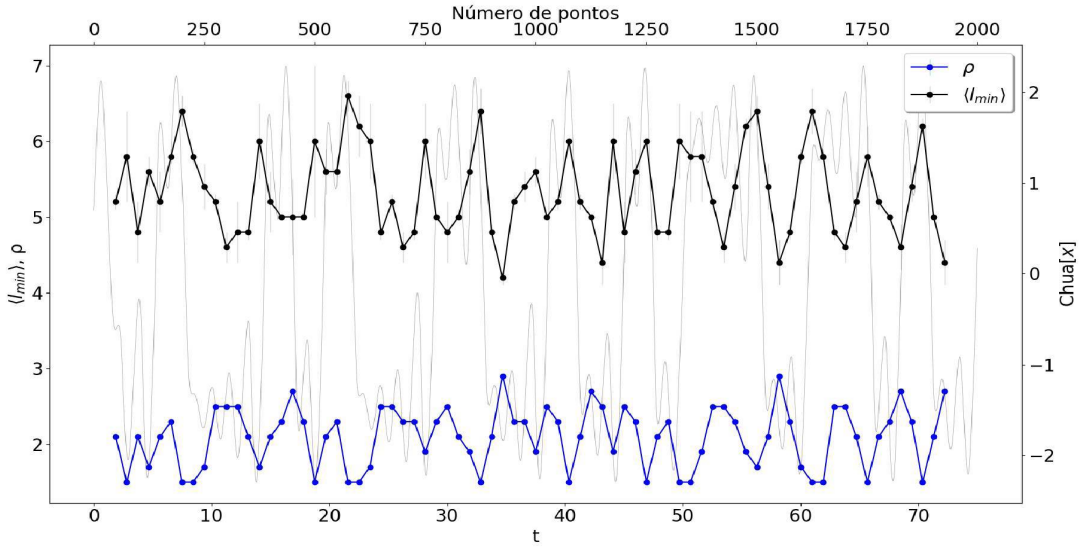


Figura 46 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada janela para o cálculo dos determinismos da componente x do oscilador de Chua para cada janela. A curva cinza ao fundo serve como guia para ilustrar quais as posições das janelas na série foram utilizadas para os cálculos de l_{min} e ρ. Cada ponto está no centro de sua respectiva janela.

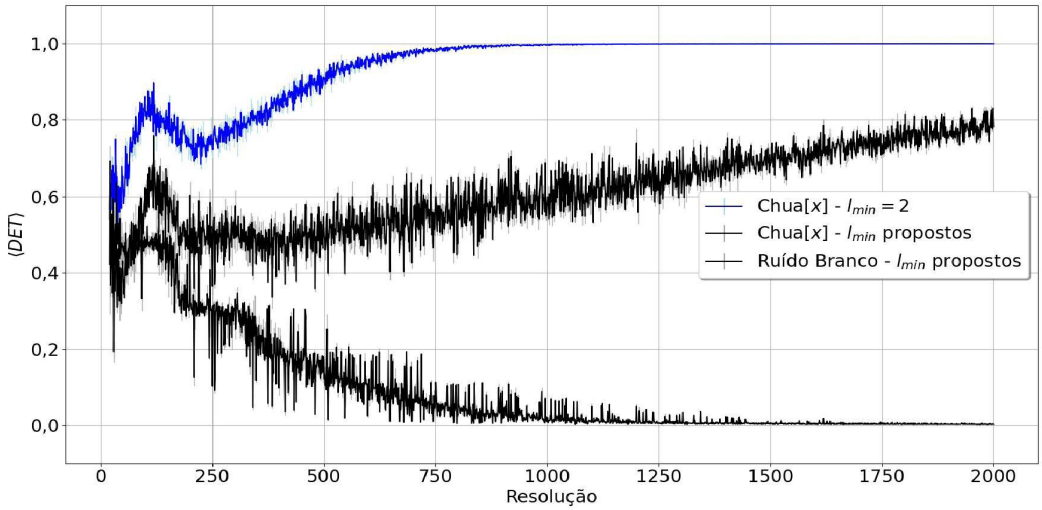


Figura 47 – Determinismos para a componente x do oscilador de Chua e um ruído branco obtidos em função de sua resolução. A curva em azul corresponde aos determinismos da componente x do oscilador de Chua obtidos com $l_{min} = 2$. As curvas em preto correspondem aos determinismos obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco.

Vemos que os determinismos obtidos com $l_{min} = 2$ vão muito rapidamente à 1, porém, neste caso, também percebemos que os determinismos obtidos com os l_{min} propostos também crescem, fazendo com que eles tendam à 1 para maiores resoluções. Entretanto, o adequado seria utilizar valores de N próximos ao N_c, este valor de resolução já é suficiente para que as restrições sejam satisfeitas.

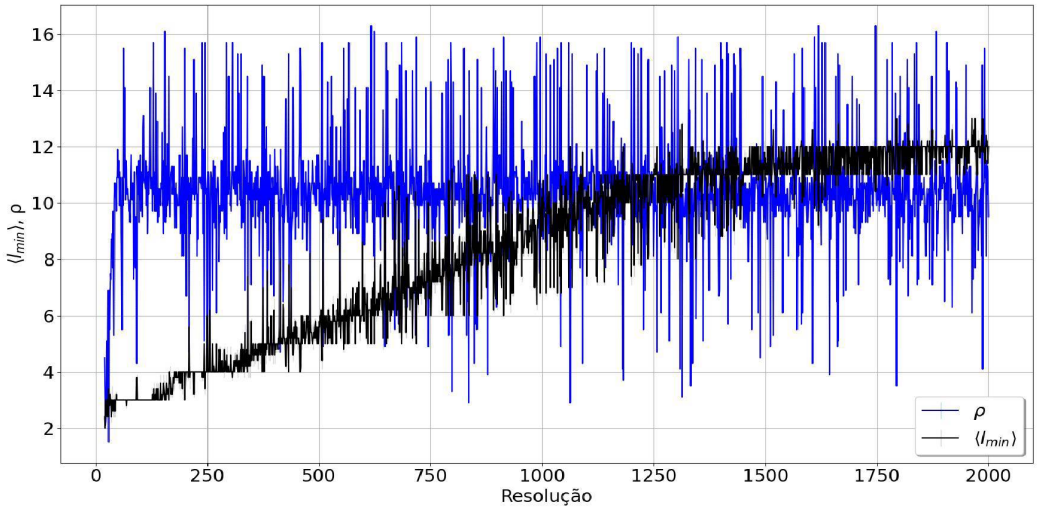


Figura 48 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de pseudo-períodos (azul) utilizados para cada resolução no cálculo dos determinismos da componente x do oscilador de Chua.

Nesta subseção vimos que o método proposto de obtenção autônoma do l_{min} faz com que os determinismos para alguns sistemas tendam ao valor 1, mesmo que estes sistemas sejam caóticos. Contudo, vale ressaltar que este fenômeno também é observado atualmente com o $l_{min} = 2$ e, mesmo nestes casos, o determinismo é substancialmente diferente nas proximidades do N_c, fazendo com que o método proposto de fato seja capaz de distinguir corretamente o comportamento de diferentes séries. A curva de determinismo para a componente x o oscilador de Chua obtido com $l_{min} = 2$ é praticamente indistinguível da curva de determinismo de um seno, mesmo para baixas resoluções, o que não ocorre com o critério proposto.

5.4 Movimento Browniano

Nesta subseção abordaremos como o critério proposto de escolha do parâmetro l_{min} altera os resultados de análises via determinismo para uma série estocástica oriunda de um processo de Wiener [83] unidimensional. Este processo, que podemos denotar por s_t, possui algumas propriedades importantes:

- É *quase sempre* contínuo;
- É um processo de Markov [84, 33], ou seja, apenas s_t influencia no valor de s_{t+dt};
- É um Martingale [85], ou seja, a distribuição de probabilidades de s_{d+dt} é centrada em s_t;
- É um processo de Lévy, ou seja, a série gerada a partir das diferenças entre elementos consecutivos $s_1 = s_{t_1}$ e $s_2 = s_{t_2}$, para $0 \leq t_1 \leq t_2$, possui uma distribuição gaussiana da forma

$$p(s_{t_2 - t_1}) = \frac{1}{\sqrt{2\pi(t_2 - t_1)}} e^{-\frac{(s_2 - s_1)^2}{2(t_2 - t_1)}}; \tag{5.6}$$

- A variável s_t não é monótona em nenhum intervalo temporal, porém, o teorema de Bolzano-Cauchy é válido para ela;

- A variável s_t possui uma densidade de probabilidade gaussiana da forma

$$p(s_t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}. \tag{5.7}$$

Esta série é chamada de movimento browniano padrão (figura 49). O fato de ser contínuo significa que podemos aplicar o nosso método e o fato de ser um proecesso markoviano nos diz que o único valor que influencia a próxima iteração é o atual, o que faz com que seu comportamento seja difícil de prever. Este tipo de processo possui uma memória curta, fazendo com que possamos supor um determinismo no mínimo diferente do obtido para um seno, que é uma série previsível.

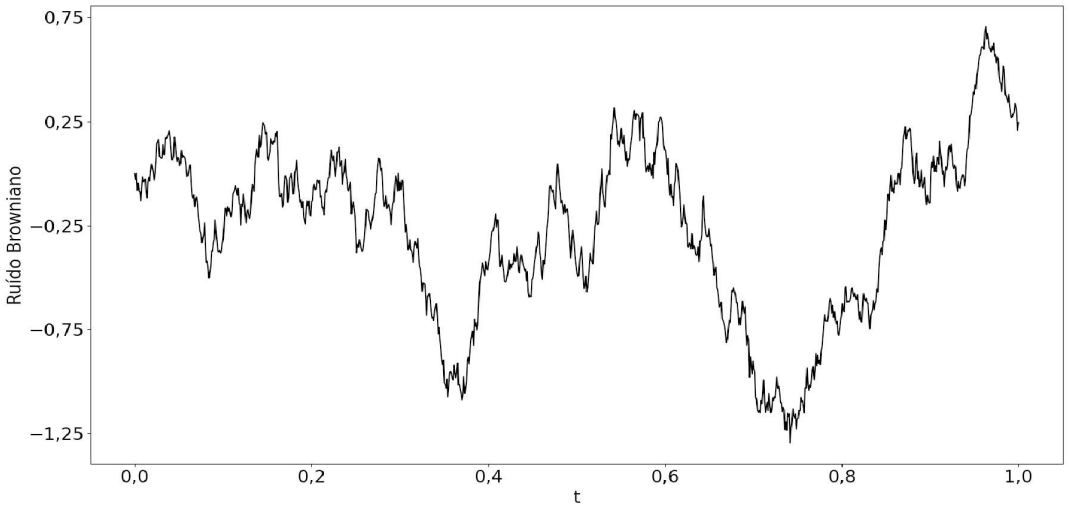


Figura 49 – Processo de Wiener unidimensional com 1000 pontos.

Uma maneira de compreender onde podemos nos deparar um movimento browniano seria ao tentar descrever a dinâmica de uma partícula submetida à uma força de caráter estocástico $F(t)$ e uma força de arrasto proporcional à sua velocidade, isto resulta na equação

$$m\frac{dv}{dt} = -\alpha v + F(t), \tag{5.8}$$

que é chamada de equação de Langevin [86]. Nela, temos que

$$\langle F(t) \rangle = 0, \tag{5.9}$$

e que

$$\langle F(t)F(t') \rangle \propto \delta(t - t'). \tag{5.10}$$

5.4.1 Análise no tempo

A análise nesta subseção será análoga aos outros casos, também será iniciada aos 20 pontos e terminará

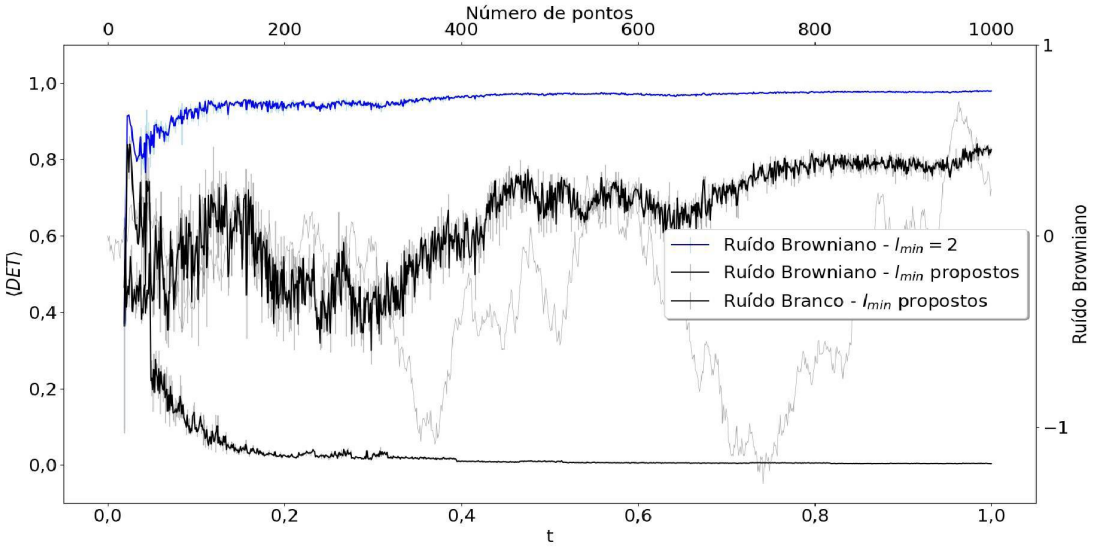


Figura 50 – Determinismos calculados ao longo do tempo para o movimento browniano e ruído branco. A curva cinza serve como guia e ilustra até que ponto da série foi utilizado para os determinismos correspondentes. Em azul, os determinismos do movimento browniano obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco.

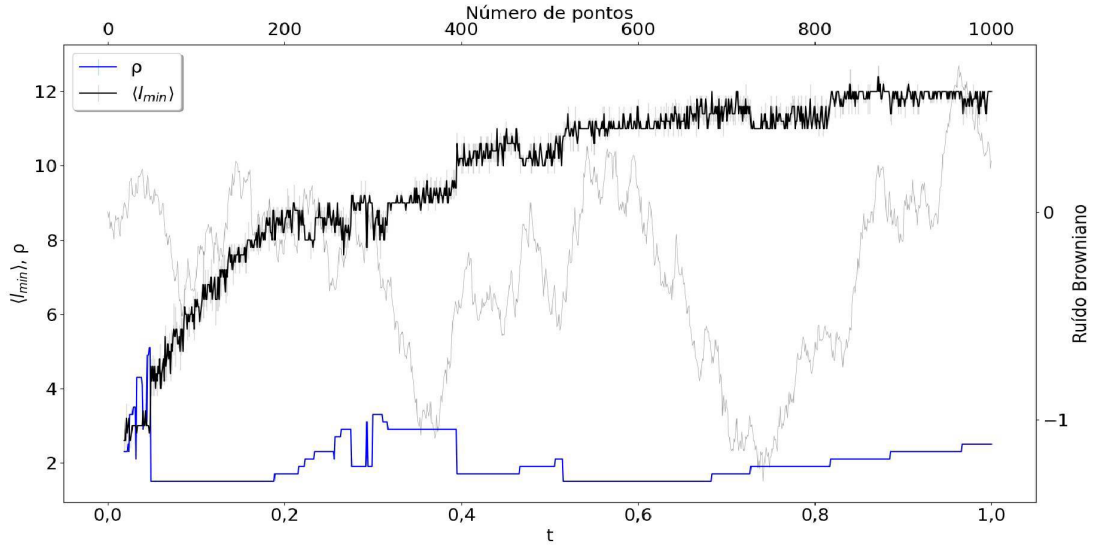


Figura 51 – Valores de $\langle l_{min} \rangle$ obtidos via o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada ponto da série para o cálculo dos determinismos do movimento browniano no tempo. A curva cinza serve como guia para ilustrar qual seção da série foi utilizada para o cálculo de l_{min} e ρ.

nos 1000 pontos, onde a série estará completa. Os determinismos podem ser vistos na figura 50. Vemos que, para este caso, mesmo sendo uma série estocástica, os determinismos obtidos (figura 50) com $l_{min} = 2$ são praticamente os mesmos dos obtidos para um seno (série periódica) e para o oscilador de Chua (série caótica).

Isso confirma que o critério atual de escolha do l_{min} não é capaz de diferenciar a dinâmica de algumas séries, o que é possível com o critério proposto. Vemos isto na curva preta entre $0,4$ e $0,8$ (figura 50).

A figura 51 mostra que as estimativas do número de pseudo-períodos são baixas e estão em torno de 2, fazendo com que precisemos de poucos pontos para alcançar a região onde $N \approx N_c$.

5.4.2 Análise com defasagem

As defasagens nesta subseção também serão de tamanho 100 e passo 25 pontos, e os determinismos obtidos para cada uma destas janelas podem ser vistos na figura 52.

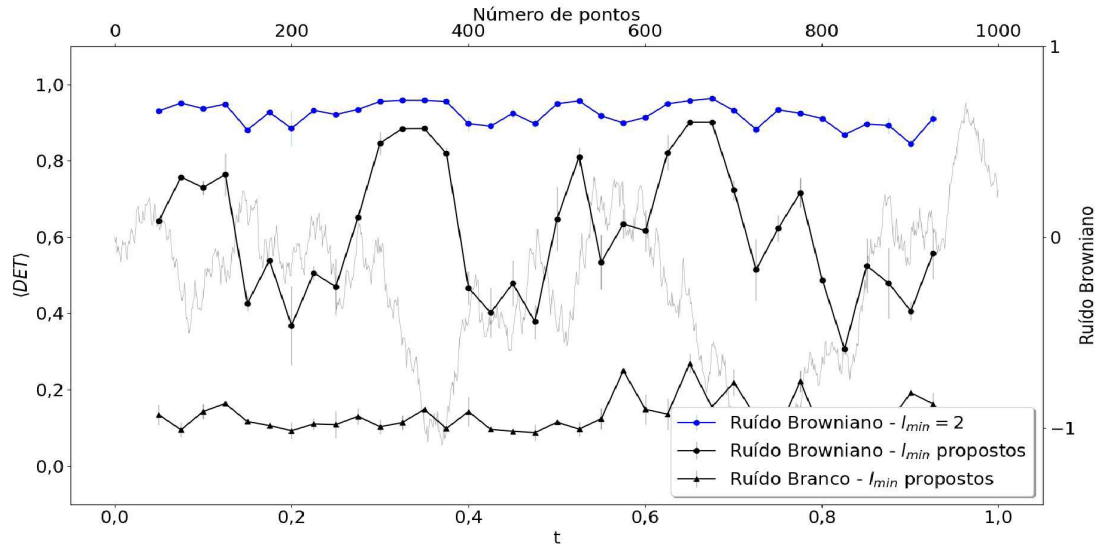


Figura 52 – Determinismos calculados para cada janela para o movimento browniano e um ruído branco. A curva cinza ao fundo serve como guia. Os determinismos estão situados no centro da janela de 100 pontos na qual eles foram calculados. Em azul, os determinismos do movimento browniano obtidos com $l_{min} = 2$. Em preto, os determinismos da série em questão e do ruído branco, obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco.

Vemos que, mesmo analisando esta série utilizando janelas, o determinismo se mantém muito próximo ao de um seno quando utilizamos o valor atual para o l_{min}. Contudo, ao utilizar os valores escolhidos de forma autônoma por meio da minimização do parâmetro de assimetria, obtemos valores completamente diferentes. Isto evidencia as mudanças no comportamento da série, pois vemos que onde há quedas nos valores da série há máximos locais nos determinismos. Com isso, confirma-se mais uma vez que é possível extrair informação acerca do comportamento da série com o método proposto de forma mais eficiente do que é feito atualmente. A lista de $\langle l_{min} \rangle$ e das estimativas de ρ podem ser vistas na figura 55.

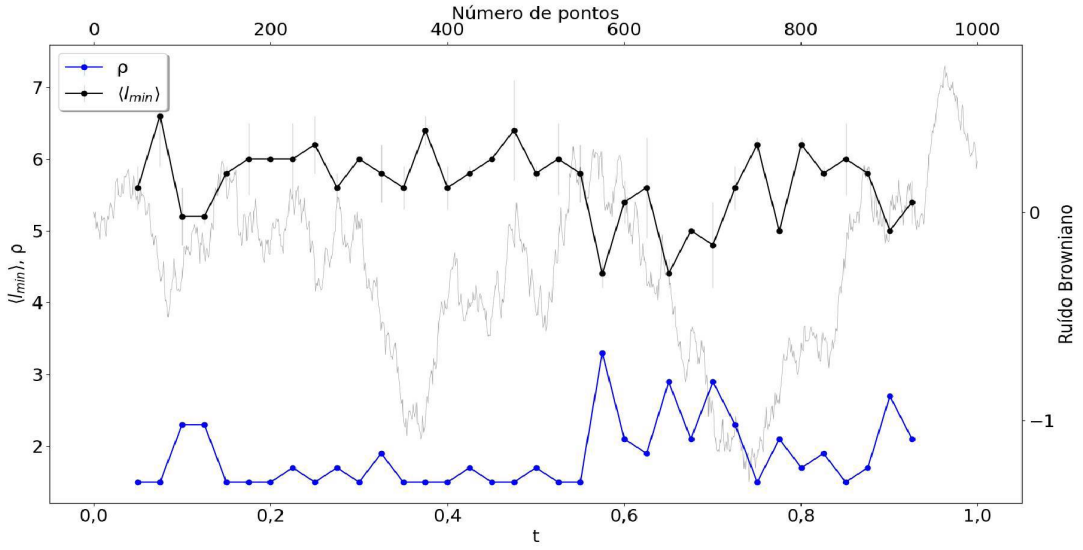


Figura 53 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de períodos (azul) utilizados em cada janela para o cálculo dos determinismos do movimento browniano para cada janela. A curva cinza ao fundo serve como guia para ilustrar quais as posições das janelas na série foram utilizadas para os cálculos de l_{min} e ρ. Cada ponto está no centro de sua respectiva janela.

5.4.3 Análise na resolução

A análise se dá também de forma análoga, começando com 20 pontos de resolução até todos os 1000 pontos (figura 54).

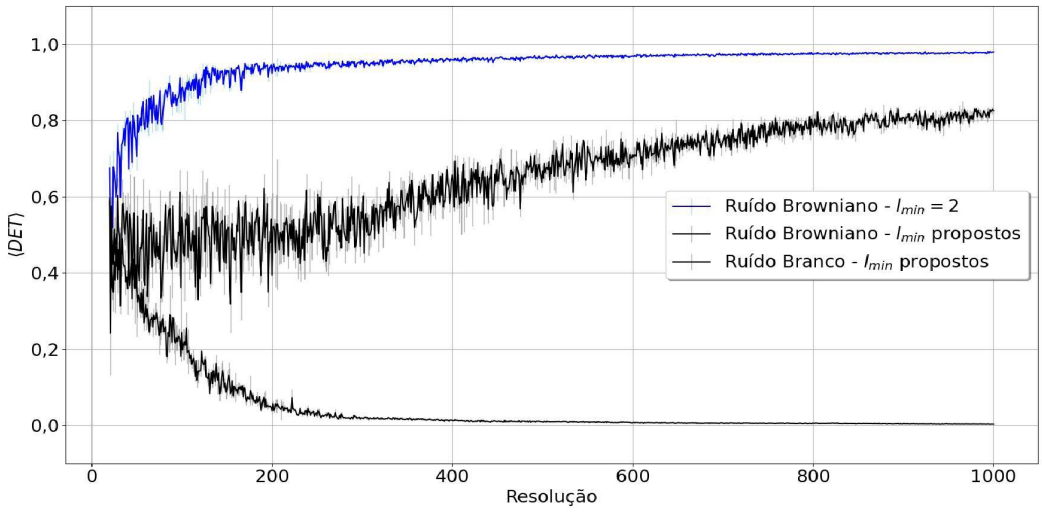


Figura 54 – Determinismos para o movimento broaniano e um ruído branco obtidos em função de sua resolução. A curva em azul corresponde aos determinismos do movimento browniano obtidos com $l_{min} = 2$. As curvas em preto correspondem aos determinismos obtidos com o critério proposto. A curva com os valores mais baixos corresponde ao ruído branco.

Como esta série é oriunda de um processo contínuo, foi possível gerar uma solução densa em pontos, de forma análoga ao que foi feito para os sistemas dinâmicos anteriores.

Uma outra propriedade muito importante de um movimento browniano que não foi mencionada é que ele não é diferenciável *quase nunca*. Isso significaria que, a *priori*, uma análise feita variando sua resolução não faça sentido. Entretanto, como na prática o que fazemos é utilizar intervalos temporais finitos, mesmo que pequenos, o teorema de Bolzano-Cauchy, também conhecido como *teorema do valor intermediário*, nos diz que é possível traçar uma curva subjacente aos valores da série, mesmo que haja variância não nula. Assim teríamos uma curva contínua que representa de forma razoável o processo, mesmo que não seja de forma exata. Do ponto de vista de análise de recorrência, isso implica apenas aumenta o nível de detalhes de alguma das estruturas internas do gráfico de recorrências. Como consequência, o número de linhas diagonais aumentaria, gerando determinismos maiores ao passo em que a resolução aumenta, pois cada linha do gráfico tenderia a ser maior do que apenas dois pontos consecutivos. Vemos este aumento do número de linhas diagonais implicando em um aumento do determinismo ao alterar a resolução em todos os sistemas até agora.

Vemos que na resolução necessária para que o determinismo de um ruído branco seja próximo de zero o determinismo do ruído browniano é aproximadamente 0,6 com o critério proposto e, com o valor atualmente aceito para o l_{min}, ele está próximo de 1 (figura 55).

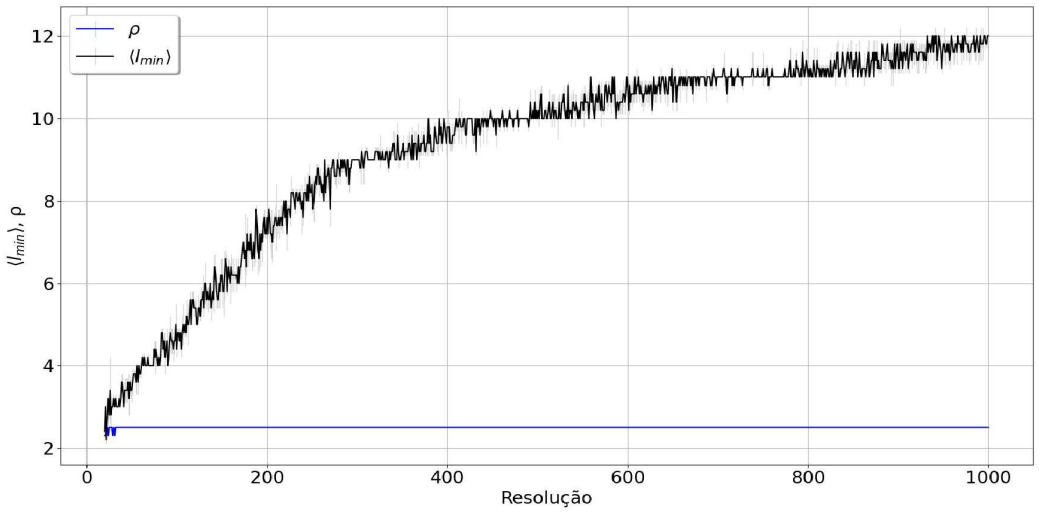


Figura 55 – Valores de $\langle l_{min} \rangle$ obtidos com o método proposto (preto) e estimativa do número de pseudo-períodos (azul) utilizados para cada resolução no cálculo dos determinismos do movimento browniano.

Percebe-se que o número de pseudo-períodos utilizados para a análise na resolução é majoritariamente 2,5 para toda a resolução (com excessão das resoluções próximas aos 20 pontos). Isso faz com que a combinação de N e ρ forneça valores confiáveis para resoluções próximas do N_c, mesmo para baixas resoluções. "Confiabilidade", neste caso, implica em determinismos para $N \geq N_c$.

6 Conclusão

Nesta dissertação vimos que o estado atual da pesquisa feita no âmbito da análise de quantificação de recorrência possui pontos pouco explorados pela literatura. Este trabalho busca apresentar propostas de aprimoramento, particularmente no que tange à escolha dos parâmetros associados aos quantificadores de recorrência, mais especificamente, o determinismo.

Ao postular duas condições de contorno para os determinismos de senos e ruídos brancos foi possível obter um critério de escolha do parâmetro l_{min} de maneira automática, apenas dependendo do número de pontos e da quantidade de períodos, ou pseudo-períodos apresentados pela série analisada. Desta forma, a arbitrariedade indesejada é removida por completo. Um passo extra que é introduzido, no entanto, é a obtenção desta quantidade de períodos, ou pseudo-períodos, porém, na vasta maioria dos casos, é algo simples e rápido de se obter. Para isso, há atualmente vários métodos teóricos como periodogramas e extremização da função de autocorrelação, assim como pacotes computacionais que nos permitem aplicar estes algorítimos de forma fácil e rápida.

A obtenção do determinismo, obtido com o método proposto, pode ser resumido como uma comparação entre os ciclos da série analisada com a mesma quantidade de oscilações simples com o mesmo número de pontos. Além disso, durante o processo de investigação do método, foi encontrada uma maneira de verificar se o resultado obtido pode ser considerado satisfatório ou não, por meio do conceito de *N crítico*. Esta quantidade nos mostra se o determinismo calculado sofreu o reescalonamento correto ou não. Este conceito é muito importante e pode ser utilizado na análise com variação de parâmetros de sistemas dinâmicos. Mesmo que em algumas regiões o determinismo não seja elucidativo, o conceito de *N crítico* é capaz de mostrar se o resultado obtido possui algum significado ou não, onde para valores baixos de pontos por período, ou pseudo-período, os determinismos obtidos podem não ser confiáveis.

Em comparação com os determinismos obtidos com o critério atual de escolha do l_{min}, vimos uma grande vantagem do critério proposto na distinção de séries temporais com dinâmicas diferentes. O critério atual resulta em várias séries possuindo os mesmos valores para o quantificador, potencialmente induzindo à erros e ambiguidades nas análises. Ao considerarmos o conceito de *N crítico* e o critério proposto, no entanto, este problema pode ser facilmente contornado. Isso foi notadamente visto na análise do oscilador de Chua, onde as curvas de determinismo para este sistema em um estado caótico eram as mesmas curvas obtidas para um seno. Atualmente não seria possível distinguir a dinâmica de duas séries completamente diferentes, o que foi corrigido com uma simples aplicação de condições de contorno.

Em suma, a simples introdução de um parâmetro automaticamente obtido a partir dos determinismos de um seno e de um ruído branco possibilita a obtenção de resultados mais satisfatórios no que diz respeito à distinção de séries do que o critério atual. A arbitrariedade na hora de utilizar o quantificador determinismo é reduzida ao máximo, pois, com o conceito de microestados, os gráficos de recorrências não necessitam mais do parâmetro *threshold*. Somado à isso, apenas sabendo o número de pontos da série e a quantidade de períodos ou pseudo-períodos presentes nela, é possível obter de forma autônoma também o l_{min}. O resultado é uma diminuição de um possível viés na hora de calcular o quantificador determinismo.

Possibilidades de extensão do método incluem a definição de condições de contorno para mapas, análogas às que foram estipuladas para sistemas contínuos e, extrapolar o procedimento para o quantificador laminaridade. No último caso, vemos claramente o motivo, pois a laminaridade é equivalente ao determinismo, em termos de formulação matemática. Outra linha de pensamento que podem ser seguida é a determinação da

relação entre as quantidades associadas ao l_{min} e entropia de informação. Os resutados apresentados refletem uma maneira de se trabalhar com análise de recorrência, quando se maximiza a entropia dos microestados da série.

Bibliografia

[1] J. D. Jackson, *Classical Electrodynamics*. John Wiley & Sons, 3 ed., 1998.

[2] B. W. Carroll, D. A. Ostlie, *An Introduction to Modern Astrophysics*. Cambridge University Press, 2 ed., 2017.

[3] M. Frazier, B. Taddese, B. Xiao, T. Antonsen, E. Ott and S. M. Anlage, "Nonlinear time reversal of classical waves: Experiment and model," *Physical Review E*, vol. 88, 2021.

[4] S. V. George, S. Kachhara, R. Misra and G. Ambika, "Early warning signals indicate a critical transition in Betelgeuse," *Astronomy & Astrophysics*, vol. 640, 2020.

[5] R. Fusaroli, I. Konvalinka and S. Wallot, "Analyzing Social Interactions: The Promises and Challenges of Using Cross Recurrence Quantification Analysis," in *Translational Recurrences: From Mathematical Theory to Real-World Applications*, 2014.

[6] P. M. Addo, M. Billio and D. Guégan, "Nonlinear dynamics and recurrence plots for detecting financial crisis," *The North American Journal of Economics and Finance*, vol. 26, 2013.

[7] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge University Press, 2 ed., 2014.

[8] P. J. Brockwell1 and R. A. Davis, *Time Series: Theory and Methods*. Springer, 2 ed., 1991.

[9] R. H. Shumway and D. S. Sotffer, *Time Series Analysis and Its Applications: With R Examples*. Springer, 4 ed., 2017.

[10] J. K. Hale, "Ordinary Differential Equations," 2009.

[11] V. Arnold, *Équations Différentielles Ordinaires*. Mir, 1 ed., 1974.

[12] G. Carlsson, "Topology and Data," *Bulletin of the American Mathematical Society*, vol. 46, 2009.

[13] K. Oliveira and M. Viana, *Fundamentos da Teoria Ergódica*. SBM, 2 ed., 2019.

[14] P. A. Morettin and W. de O. Bussab, *Estatística Básica*. Saraiva Uni, 9 ed., 2017.

[15] L. O. Chua, "The Genesis of Chua's Circuit," *International Journal of Electronics and Communications*, vol. 46, 1992.

[16] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim and J. S. Suri, "Heart rate variability: a review," *Medical and Biological Engineering and Computing*, vol. 11, 2006.

[17] G. E. P. Box, G. M. Jenkins, G. C. Reinsel and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, 5 ed., 2015.

[18] P. H. Franses, D. van Dijk and A. Opschoor, *Time Series Models for Business and Economic Forecasting*. Cambridge University Press, 2 ed., 2014.

[19] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Westview Press, 1 ed., 2001.

[20] H. Poincaré, *Sur le problème des trois corps et les équations de la dynamique*, vol. 13. Acta Math, 1890.

[21] E. Ott, *Chaos in Dynamical Systems.* Cambridge University Press, 1 ed., 1993.

[22] P. Grassberger and I. Procaccia, "Estimation of the Kolmogorov entropy from a chaotic signal," *Physical Review A*, vol. 9, 1983.

[23] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control.* Cambridge University Press, 1 ed., 2019.

[24] J. P. Eckmann, S. O. Kamphorst and D. Ruelle, "Recurrence Plot for Dynamical Systems," *Europhysics Letters*, vol. 4, 1987.

[25] N. Marwan, J. Kurths, M. Thiel and M. C. Romano, "Recurrence plots for the analysis of complex systems," *Physics Reports*, vol. 438, 2007.

[26] N. Marwan, "How to avoid potential pitfalls in recurrence plot based data analysis," *International Journal of Bifurcation and Chaos*, vol. 21, 2010.

[27] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan and J. Kurths, "Recurrence Networks: a novel paradigm for nonlinear time series analysis," *New Journal of Physics*, vol. 12, 2010.

[28] N. R. Lomb, "Least-Squares Frequency Analysis of Unequally Sampled Data," *Astrophysics and Space Sciences*, vol. 39, 1976.

[29] I. Ozken, D. Eroglu, S. F. M. Breitenbach, N. Marwan, L. Tan, U. Tirnakli and J. Kurths, "Recurrence plot analysis of irregularly sampled data," *Physical Review E*, vol. 98, 2018.

[30] G. Orlando and G. Zimatore, "Recurrence quantification analysis of business cycles," *Chaos, Solitons & Fractals*, vol. 110, 2018.

[31] K. S. Santos, A. C. de Souza, M. T. Silva and P. V. de Azevedo, "Análise de séries temporais de variáveis metereológicas na cidade de Campina Grande-PB," in *Anais do XI Workshop Brasileiro de Micrometeorologia*, 2019.

[32] D. Kadaa, A. Kouidere, O. Balatif and L. el houssine, "Mathematical modeling of the spread of COVID-19 among different age groups in Morocco: Optimal control approach for intervention strategies," *Chaos, Solitons & Fractals*, vol. 141, 2020.

[33] T. Tomé and M. J. de Oliveira, *Dinâmica Estocástica e Irreversibilidade.* Edusp, 2 ed., 2014.

[34] W. Bao, J. Yue and Y. Rao, "A deep learning framework for financial time series using stacked autoencoders and longshort term memory," *PLOS ONE*, vol. 12, 2017.

[35] N. R. Draper and H. Smith, *Applied Regression Analysis.* John Wiley & Sons, 3 ed., 1998.

[36] E. Kreyszig, *Introductory Functional Analysis With Applications.* John Wiley & Sons, 1 ed., 1978.

[37] L. Koopmans, *The Spectral Analysis of Time Series.* Springer, 1 ed., 1995.

[38] J. Benesty, M. G. Christensen and J. R. Jensen, *Signal Enhancement with Variable Span Linear Filters.* Springer, 1 ed., 2016.

[39] O. E. Rössler, "An equation for hyperchaos," *Physics Letters A*, vol. 71, 1979.

[40] N. H. Packard, J. P. Crutchfield, J. D. Farmer and R. S. Shaw, "Geometry from a time series," *Physical Review Letters*, vol. 45, 1980.

[41] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence: Lecture Notes in Mathematics*, vol. 898, Springer, 1981.

[42] H. Whitney, "Differentiable Manifolds," *Annals of Mathematics*, vol. 37, 1936.

[43] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, 1948.

[44] A. Pikovsky and A. Politi, *Lyapunov Exponents.* Cambridge University Press, 1 ed., 2016.

[45] C. L. Webber Jr. and N. Marwan, *Recurrence Quantification Analysis: Theory and Best Practices.* Springer, 1 ed., 2015.

[46] B. Goswami, "A Brief Introduction to Nonlinear Time Series Analysis and Recurrence Plots," *Vibration*, vol. 2, 2019.

[47] F. Takens, "On the numerical determination of the dimension of an attractor," in *Dynamical Systems and Bifurcations: Lecture Notes in Mathematics*, vol. 1125, Springer, 1985.

[48] J. Theiler, "Lacunarity in a best estimator of fractal dimension," *Physics Letters A*, vol. 133, 1988.

[49] S. Borovkova, R. Burton and H. Dehling, "Consistency of the Takens estimator for the correlation dimension," *Annals of Applied Probability*, vol. 9, 1999.

[50] P. Grassberger and I. Procaccia, "Characterization of Strange Attractors," *Physical Review Letters*, vol. 50, 1983.

[51] M. Lines, *Nonlinear Dynamical Systems in Economics.* Springer, 1 ed., 2005.

[52] T. Jackson and A. Radunskaya, *Applications of Dynamical Systems in Biology and Medicine.* Springer, 1 ed., 2015.

[53] M. Small, J. Zhang and X. Xu, "Transforming Time Series into Complex Networks," *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering: Complex Sciences*, vol. 5, 2009.

[54] N. Marwan, "A historical review of recurrence plots," *The European Physical Journal Special Topics*, vol. 164, 2008.

[55] O. E. Rössler, "An equation for continuous chaos," *Physics Letters A*, vol. 57, 1976.

[56] J. P. Zbilut, A. Giuliani and C. L. Webber Jr., "Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification," *Physics Letters A*, vol. 246, 1998.

[57] N. Marwan and J. Kurths, "Nonlinear analysis of bivariate data with cross recurrence plots," *Physics Letters A*, vol. 302, 2002.

[58] A. Porta, G. Baselli, N. Montano, T. G. Ruscone, F. Lombardi, A. Malliani and S. Cerutti, "Classification of coupling patterns among spontaneous rhythms and ventilation in the sympathetic discharge of decerebrate cats," *Biological Cybernetics*, vol. 75, 1996.

[59] M. C. Romano, M. Thiel, J. Kurths and W. von Bloh, "Multivariate recurrence plots," *Physics Letters A*, vol. 330, 2004.

[60] J. P. Zbilut and C. L. Webber Jr., "Embeddings and delays as derived from quantification of recurrence plots," *Physics Letters A*, 1992.

[61] C. L. Webber Jr. and J. P. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plots strategies," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 76, 1994.

[62] J. P. Zbilut and C. L. Webber Jr., "Recurrence quantification analysis: Introduction and historical context," *International Journal of Bifurcation and Chaos*, vol. 17, 2007.

[63] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica D: Nonlinear Phenomena*, vol. 9, 1983.

[64] G. B. Mindlin and R. Gilmore, "Topological analysis and synthesis of chaotic time series," *Physica D: Nonlinear Phenomena*, vol. 58, 1992.

[65] L. L. Trulla, A. Giuliani, J. P. Ziblut and C. L. Webber Jr., "Recurrence quantification analysis of the logistic equation with transients," *Physics Letters A*, vol. 223, 1996.

[66] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan and J. Kurths, "Recurrence plot based measures of complexity and their application to rate variability data," *Physical Review E*, vol. 66, 2002.

[67] J. Gao and H. Cai, "On the structures and quantifiaction of recurrence plots," *Physics Letters A*, vol. 270, 2000.

[68] T. Braunn, V. R. Unni, R. I. Sujith, J. Kurths and N. Marwan, "Detection of dynamical regime transitions with lacunarity as a multiscale recurrence quantification measure," *Nonlinear Dynamics, An International Journal of Nonlinear Dynamics and Chaos in Engineering Systems*, vol. 104, 2021.

[69] D. Yang, W. X. Ren, Y. D. Hu and D. Li, "Selection of optimal threshold to construct recurrence plot for structural operational vibration measurements," *Journal of Sound and Vibration*, vol. 349, 2015.

[70] K. K. Hauke, R. V. Donner, J. Heitzig and N. Marwan, "Recurrence threshold selection for obtaining robust recurrence characteristics in different embedding dimensions," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, 2018.

[71] E. T. Jaynes, "Information Theory and Statistical Mechanics," *Physical Review*, vol. 106, 1957.

[72] E. T. Jaynes, "Information Theory and Statistical Mechanics II," *Physical Review*, vol. 108, 1957.

[73] G. Corso, T. de L. Prado, G. Z. dos S. Lima, J. Kurths, S. R. Lopes, "Quantifying entropy using recurrence matrix microstates," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, 2018.

[74] J. Iwanski and E. Bradley, "Recurrence plots of experimental Data: To embed or not to embed?," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 8, 1988.

[75] T. K. March, S. C. Chapman and R. O. Dendy, "Recurrence plot statistics and the effect of embedding," *Physica D*, vol. 200, 2005.

[76] B. Babaei, R. Zarghami, S. Hossein, R. S. Gharebagh and N. Mostoufi, "Selection of minimal length of line in recurrence quantification analysis," *Physica A*, 2014.

[77] H. Guidorizzi, *Um Curso de Cálculo: Vol 1.* LTC, 5 ed., 2001.

[78] J. D. Scargle, "Studies in Astronomical Time Series Analysis II. Statistical Aspects of Spectral Analysis of Unevenly Spaced Data," *The Astrophysical Journal*, 263.

[79] E. N. Lorenz, "Deterministic Nonperiodic Flow," *Journal of the Atmospheric Sciences*, vol. 20, 1963.

[80] C. Rackauckas and Q. Nie, "DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia," *Journal of Open Research Software*, 2017.

[81] J. Bezanson, S. Karpinski, V. B. Shah and A. Edelman, "Julia: A Fast Dynamic Language for Technical Computing," Tech. Rep. arXiv:1209.5145 [cs.PL], ArXiV, September 2012.

[82] L. O. Chua, M. Komuro and T. Matsumoto, "The Double Scroll Family," *IEEE Transactions on Circuits and Systems*, vol. 33, 1986.

[83] N. Wiener, *Norbert Wiener: Collected Works Vol. 1.* MIT Press, 1 ed., 1976.

[84] R. A. Howard, *Dynamic Probabilistic Systems, Volume 1: Markov Chains.* Dover Publications, 1 ed., 2007.

[85] D. Williams, *Probability with Martingales.* Cambridge University Press, 1 ed., 1991.

[86] P. Langevin, "Sur la théorie du mouvement brownien," *Comptes Rendus de l'Académie des Sciences*, vol. 146, 1908.

A Exemplos de Sistemas Dinâmicos

A.1 Exemplos de Mapas Determinísticos

Veremos alguns exemplos de sistemas dinâmicos determinísticos com tempo discreto.

Mapa Logístico: Definido pela equação

$$x_{n+1} = rx_n(1 - x_n), \tag{A.1}$$

é um mapa unidimensional, não linear e não invertível, portanto, caótico, que pode ser utilizado para modelar sistemas biológicos como flutuações populacionais.

Mapa de Hénon: Definido por

$$\begin{aligned} x_{n+1} &= 1 - ax_n^2 + y_n, \\ y_{n+1} &= bx_n, \end{aligned} \tag{A.2}$$

é um mapa bidimensional, não linear e não invertível, que apresenta várias características de sistemas caóticos, inclusive um atrator estranho.

Mapa Padrão: Dado pelas equações

$$\begin{aligned} \theta_{n+1} &= \theta_n + p_{n+1}, \\ p_{n+1} &= p_n + k\sin(\theta_n), \end{aligned} \tag{A.3}$$

é um mapa conservativo bidimensional, caótico que pode ser obtido ao realizar uma seção de Poincarè[1] em um sistema de um rotor pulsado.

A.2 Exemplos de Fluxos Determinísticos

Veremos alguns exemplos de sistemas dinâmicos com tempo contínuo.

Sistema de Lorenz: Definido pelas equações

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z, \end{aligned} \tag{A.4}$$

é um sistema não linear e caótico simplificado utilizado para modelar o movimento de convecção atmosférica, mostrando que pode-se encontrar caos em sistemas desta natureza.

[1] É um método de se obter um mapa $(m-1)$-dimensional a partir de um sistema contínuo m-dimensional por meio da delimitação de um plano $(m-1)$-dimensional no espaço de fase do sistema contínuo. Neste plano, coletam-se os pontos que o cruzam, sendo que o conjunto de pontos coletados é então um mapa com dimensão menor e que preserva todas as propriedades do sistema original e é chamado de "Seção de Poincarè".

Oscilador de Chua: Gerado a partir das equações

$$\begin{aligned} \dot{x} &= \alpha(y - h(x)), \\ \dot{y} &= x - y + z, \\ \dot{z} &= -\beta z, \end{aligned} \tag{A.5}$$

onde o termo $h(x)$ é dado por

$$h(x) \coloneqq m_1 x + \frac{1}{2}(m_0 - m_1)(|x + 1| - |x - 1|), \tag{A.6}$$

é um sistema não linear e caótico com três dimensões obtido ao modelar um circuito eletrônico com um componente resistivo com resposta não linear. Isso permitiu com que caos pudesse ser verificado experimentalmente, em contraste com os modelos puramente teóricos até então.

Sistema de Rössler: Dado pelas equações

$$\begin{aligned} \dot{x} &= -y - z, \\ \dot{y} &= x + ay, \\ \dot{z} &= b + z(x - c), \end{aligned} \tag{A.7}$$

é um sistema não linear caótico que foi proposto para ser o sistema mais simples possível que é capaz de apresentar caos, tendo apenas um termo não linear.

B Processos Estocásticos

Uma variável aleatória X_τ que depende do tempo τ é chamada de processo estocástico [33] e pode ser completamente descrita para valores de τ dentro de um intervalo $[a, b] \in \mathcal{T}$ através de uma distribuição de probabilidades

$$\mathcal{P}(\{n_\tau\}), \forall \tau \in [a, b]. \tag{B.1}$$

A função $\mathcal{P}$ representa a distribuição de probabilidades de que a variável X_τ valha n_τ no instante τ, onde τ está dentro de um intervalo temporal definido por $a \leq b$ e $[a, b] \subset \mathcal{T}$, com

1. $\mathcal{T} = \mathbb{N}$;
2. $\mathcal{T} = \mathbb{R}$.

No apêndice sobre Sistemas Dinâmicos foram abordados apenas sistemas chamados de *determinísticos*, nesse contexto, processos estocásticos podem ser considerados sistemas dinâmicos *probabilísticos*. Eles também representam a evolução temporal de alguma quantidade. Veremos a seguir exemplos breves de processos estocásticos.

Processos de Markov: Esta categoria de processos diz respeito a todas as sequências $\{X_\tau\}_{\tau \in \mathcal{T}}$ em que a próxima iteração apenas depende do valor atual da variável, gerando assim uma "falta de memória" no sistema. A previsão de valores futuros se dá através de uma matriz de probabilidades chamada de *matriz de transição*, denotada por T. Esta matriz representa a probabilidade de transição de um estado i para um estado j, dado que o sistema estava previamente no estado i. A evolução de um sistema desta maneira se dá por sucessivas aplicações da matriz de transição T em um vetor de estado m-dimensional X, onde

- Se o processo for a tempo discreto a aplicação ocorre com $X_{\tau+n} = T^n \cdot X_n$;
- Se o processo for a tempo contínuo a aplicação ocorre com $X_{\tau+\delta\tau} = T(\delta\tau) \cdot X_\tau$,

Para o caso discreto, o índice n representa um expoente e, no caso contínuo, $\delta\tau$ representa um argumento, com T passando a ser um *operador*. No caso contínuo, o operador T pode ser entendido como uma integral do estado X_τ com o *kernel* de evolução, que é uma função que nos dá a evolução contínua para o processo e a integração ocorre entre $[\tau, \tau + \delta\tau]$.

Processo Gaussiano: São sequências de variáveis aleatórias com domínio contínuo que possuem densidade de probabilidades normal do tipo

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{B.2}$$

na qual μ é a média e σ^2 é a variância da distribuição. Quaisquer aplicações de funcionais lineares também resultará em uma distribuição de probabilidades normal.

Processo de Lévy: São processos estocásticos que possuem a diferença entre elementos consecutivos na sequência $\{X_\tau\}_{\tau\in\mathcal{T}}$ independentes e que seguem uma distribuição de probabilidades normal e estacionários. Estas diferenças satisfazem a propriedade

$$\lim_{\delta\tau\to 0} P(|X_{\tau+\delta\tau} - X_\tau| - \varepsilon) = 0, \tag{B.3}$$

ou seja, a distribuição de probabilidades da variável X_τ é contínua quando consideramos incrementos temporais infinitesimalmente pequenos.

Martingale: É um caso particular de um processo de Markov, onde o valor da variável aleatória $\{X_\tau\}_\tau$ no tempo τ é igual ao valor esperado do próximo elemento da sequência. Desta maneira,

$$\mathbb{E}(X_{\tau+\delta\tau}) = X_\tau, \tag{B.4}$$

na qual $\delta\tau$ representa o incremento temporal da sequência.

Processo de Wiener: É ao mesmo tempo um processo de Markov, um processo de Lévy e um Martingale, sendo denotado por W_t, que possui as seguintes propriedades:

- $W_0 = 0$, quase sempre;
- Segue distribuição de probabilidades normal;
- Possui incrementos normalmente distribuídos, assim como seus elementos;
- É contínuo quase sempre.

É usualmente chamado de movimento browniano e é obtido quando alguma quantidade está sujeita à "forças" de caráter aleatório, sendo modelado através da equação de Langevin:

$$\frac{dV}{dt} = -\alpha V + F(t), \tag{B.5}$$

na qual $V = dX/dt$ é a "velocidade" associada à variável aleatória X. Mesmo que palavras como "força" e "velocidade" tenham sido utilizadas, estas grandezas não são necessariamente oriundas da Física, pois processos como este podem ser observados comumente também em sistemas financeiros, por exemplo.