CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

MODELAGEM DE VOOS DIFUSIVOS ANÔMALOS ATRAVÉS DA FUNÇÃO DE LEVY E PRESCRIÇÃO PARA AJUSTE DE SEUS PRINCIPAIS PARÂMETROS

LUIZ OTÁVIO RODRIGUES ALVES SERENO

Orientador: Dr. José Luiz Acebal Fernandes
Centro Federal de Educação Tecnológica de Minas Gerais

Coorientadora: Dra. Ana Paula de Paiva Pereira
Universidade Federal de Itajubá

BELO HORIZONTE
FEVEREIRO DE 2023

LUIZ OTÁVIO RODRIGUES ALVES SERENO

MODELAGEM DE VOOS DIFUSIVOS ANÔMALOS ATRAVÉS DA FUNÇÃO DE LEVY E PRESCRIÇÃO PARA AJUSTE DE SEUS PRINCIPAIS PARÂMETROS

Tese apresentada ao Programa de Pós-graduação em Modelagem Matemática e Computacional do Centro Federal de Educação Tecnológica de Minas Gerais, como requisito parcial para a obtenção do título de Doutor em Modelagem Matemática e Computacional.

Área de concentração: Modelagem Matemática e Computacional

Linha de pesquisa: Sistemas Inteligentes

Orientador: Dr. José Luiz Acebal Fernandes
Centro Federal de Educação Tecnológica de Minas Gerais

Coorientadora: Dra. Ana Paula de Paiva Pereira
Universidade Federal de Itajubá

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL
BELO HORIZONTE
FEVEREIRO DE 2023

Sereno, Luiz Otávio Rodrigues Alves
S483m Modelagem de voos difusivos anômalos através da Função de Levy e prescrição para ajuste de seus principais parâmetros / Luiz Otávio Rodrigues Alves Sereno. – 2023.
84 f.

Tese de doutorado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional.
Orientador: José Luiz Acebal Fernandes.
Coorientadora: Ana Paula de Paiva Pereira.
Tese (doutorado) – Centro Federal de Educação Tecnológica de Minas Gerais.

1. Difusão – Modelos matemáticos – Teses. 2. Equações diferenciais – Teses. 3. Cálculo fracionário – Teses. 4. Fatores de transcrição – Teses. I. Fernandes, José Luiz Acebal. II. Pereira, Ana Paula de Paiva. III. Centro Federal de Educação Tecnológica de Minas Gerais. IV. Título.

CDD 519.233

Elaboração da ficha catalográfica pela bibliotecária Jane Marangon Duarte, CRB 6ª 1592 / Cefet/MG

Esta folha deverá ser substituída pela cópia digitalizada da folha de aprovação fornecida pelo Programa de Pós-graduação.

Dedicado a todos os meus familiares e amigos que contribuíram para realização deste trabalho.

Agradecimentos

Agradeço primeiramente ao Criador por me presentear com o dom da vida e por todas as bênçãos que recebi tanto em minha vida pessoal quanto profissional.

Gostaria de agradecer a todas as pessoas que contribuíram, de forma direta ou indireta, para este trabalho, durante estes anos de muito esforço, dedicação e diversas adversidades enfrentadas.

Agradeço de forma especial ao meu orientador Professor Dr. José Luiz Acebal e a minha coorientadora Professora Dra. Ana Paula, que sempre se esforçaram em sanar minhas dúvidas, dar sugestões e mostrar o melhor caminho possível a ser traçado.

Sou imensamente grato a minha esposa Sabrina por me apoiar continuamente e estar ao meu lado, compreendendo as dificuldades enfrentadas durante estes anos. Expresso também gratidão aos meus pais Luiz Fernando e Sueli, que sempre apostaram em meu sucesso acadêmico, incentivando e oferecendo as melhores oportunidades possíveis; e a minha irmã Ana Cláudia, que me apoiou desde o início.

Também agradeço a todos os professores que tive ao longo da minha vida. Em especial, à minha ex-orientadora de Iniciação Científica na Universidade Federal de Viçosa, professora Dra. Marinês Guerrero, e a todos que foram meus professores nos Programas de Mestrado e de Doutorado do CEFET-MG, dos quais destaco o Professor Dr. Rodrigo Cardoso e a Professora Dra. Sandra Mara.

Agradeço a todos os meus amigos, familiares, à Pontifícia Universidade Católica de Minas Gerais e aos meus colegas desta instituição, em especial ao professor Roney, pelo auxílio de sempre. Sou grato ainda pelo apoio e estrutura necessária para a realização deste projeto, oferecidos pelo Centro Federal de Educação Tecnológica de Minas Gerais, instituição que tenho grande carinho, não só por cursar o Doutorado nela, mas ter sido aluno no Ensino Médio, Ensino Técnico e no Mestrado.

“A imaginação é mais importante do que o conhecimento. Conhecimento é limitado, enquanto a imaginação abraça o mundo inteiro.”
(Albert Einstein)

Resumo

A difusão normal é um processo observado em muitos fenômenos na natureza. Para estes processos há simulações e teoria bem desenvolvidas, a saber, o *random walk* e a equação de difusão, ou equação de Fokker-Planck. Porém, em sistemas com interações de longo alcance entre os indivíduos ou entre indivíduos e ambiente, podem haver desvios deste regime e apresentar comportamento anômalo. A essa classe de processos mais gerais, dá-se o nome de difusão anômala. A abordagem teórica desses fenômenos varia de representação analítica por meio de equações diferenciais parciais, equações estocásticas e simulações computacionais que mimetizam o fenômeno. A partir de dados de deslocamento e tempos característicos de um fenômeno em escala microscópica, é possível identificar os parâmetros teóricos dos modelos que descrevem macroscopicamente esses dados. Com essa finalidade, é realizado um estudo teórico em que conjuntos de dados são gerados por simulações com parâmetros microdinâmicos controlados e ajustados a uma função associada ao fenômeno. Neste estudo, são gerados voos via caminhada aleatória de tempo contínuo e os modelos anômalos de difusão são equações diferenciais parciais fracionárias com soluções obtidas por Cálculo Fracionário. É apresentado um algoritmo evolucionário capaz de obter parâmetros que descrevem o fenômeno. Essa obtenção é feita através de uma otimização que ajusta dados simulados para função de Lévy. Além disso, no processo de discretização do modelo teórico associado ao fenômeno surgem novos parâmetros cujos comportamentos são investigados, pois o papel dos mesmos pode influenciar no estudo da difusão anômala. Como aplicação imediata da superdifusão está incluído o estudo de proteínas especializadas, chamadas fatores de transcrição, em células que permitem a transcrição do Ácido desoxirribonucleico (DNA).

Palavras-chave: Difusão Anômala. Análise da Finitude do Modelo. Equações Diferenciais Parciais Fracionárias. Fatores de Transcrição.

Abstract

Normal diffusion is a process observed in many phenomena in nature. For these processes there are well-developed simulations and theory, namely the random walk and the diffusion equation, or Fokker-Planck equation. However, in systems with long-range interactions between individuals or between individuals and the environment, there may be deviations from this regime and present anomalous behaviour. This class of more general processes is called anomalous diffusion. The theoretical approach to these phenomena ranges from analytical representation through partial differential equations, stochastic equations and computational simulations that mimic the phenomenon. From displacement data and characteristic times of a phenomenon on a microscopic scale, it is possible to identify the theoretical parameters of the models that macroscopically describe these data. For this purpose, a theoretical study is carried out in which datasets are generated by simulations with controlled microdynamic parameters and adjusted to a function associated with the phenomenon. In this study, flights are generated via continuous time random walk (CTRW) and the anomalous diffusion models are fractional partial differential equations with solutions obtained by Fractional Calculus. An evolutionary algorithm capable of obtaining parameters describing the phenomenon is presented. This retrieval is done through an optimisation that adjusts simulated data for Lévy's function. Furthermore, in the discretisation process of the theoretical model associated with the phenomenon, new parameters emerge whose behaviours are investigated, as their role can influence the study of anomalous diffusion. As an immediate application of superdiffusion is included the study of specialised proteins, called transcription factors, in cells that allow the transcription of deoxyribonucleic acid (DNA).

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e Siglas

AG	Algoritmo Genético
BFGS	Broyden–Fletcher–Goldfarb–Shanno
CF	Cálculo Fracionário
DNA	Deoxyribonucleic Acid
CF	Cálculo Fracionário
CTRW	Continuous Time Random Walking
EDP	Equação Diferencial Parcial
EDPF	Equação Diferencial Parcial Fracionária
FT	Fatores de Transcrição
PDF	Probability Density Function
PVC	Problema de Valor de Contorno
RL	Riemann-Liouville

Sumário

Capítulo 1

Introdução

No século XIX, o botânico escocês Robert Brown (1773 - 1858), ao observar partículas encontradas em grãos de pólen na água, notou que se moviam de modo errático através desse meio, mas não soube determinar quais mecanismos ocasionavam tal movimento (FORD, 1992). Algumas décadas mais tarde, o cientista alemão Albert Einstein (1879 - 1955) publicou um trabalho detalhando e estabelecendo uma formulação para esse fenômeno, chamado de movimento browniano (MB) em homenagem a Robert Brown (EINSTEIN, 1905). A explicação dada por ele para explicar o fenômeno é de que havia flutuações estatísticas de colisões moleculares (NEWELL, 1923).

O MB é observado em uma ampla variedade de fenômenos físicos e estudado nas últimas décadas por meio de diversas técnicas experimentais. Resumidamente, a causa para o movimento errático é o movimento térmico das moléculas no meio de interesse colidindo com as partículas de pólen de pequena massa que, por terem também pequenas dimensões, não se beneficiam da estabilidade que seria gerada por uma média sobre muitas colisões.

A relação direta entre o percurso quadrático médio, $\langle \xi^2 \rangle$, e o tempo médio entre mudanças, τ, são parâmetros que caracterizam o MB. Esse comportamento apresenta, em escala macroscópica[1], uma distribuição de probabilidade de posições notadamente gaussiana para as partículas sujeitas a esse tipo de fenômeno (SATO; KLAGES, 2019). Isso significa que a distribuição amostral de posição relativa à posição $x = 0$ é descrita pela equação de difusão de Fokker-Planck cuja solução é uma distribuição gaussiana que possui variância evoluindo linearmente com o tempo.

Em muitos sistemas, porém, de diferentes ramos, desde a Biologia e a Física até a Economia, o movimento de uma partícula de interesse não satisfaz as condições da difusão

[1] Conforme dicionário Houaiss, 'macroscópico' "num sistema, caracteriza uma escala de tamanho muito maior que a de seus constituintes". No nosso trabalho, o termo 'macroscópico' refere-se a determinadas situações nas quais um fenômeno é observado em grande escala.

clássica. Correlações entre as partículas observadas e correlações de longo/curto alcance entre partículas e o ambiente produzem, respectivamente, difusão anômala não linear e linear (SOKOLOV; KLAFTER, 2005; SHLESINGER; ZASLAVSKY; FRISCH, 1995; ZABURDAEV; DENISOV; KLAFTER, 2015). Há diversas situações para as quais o comportamento das partículas é anômalo e o deslocamento quadrático médio não é proporcional ao tempo, mas proporcional a alguma potência do tempo, diferente do que ocorre na difusão normal (BENHAMOU, 2018). Além disso, nota-se que a função densidade de probabilidade, em inglês *Probability Density Function* (PDF), observada não é gaussiana (WALPOLE et al., 1993). Houve, ainda, uma generalização do Teorema do Limite Central para o comportamento anômalo das partículas (TSALLIS, 2005).

Em muitos fenômenos são percebidos comportamentos incomuns no movimento em pequena escala e, consequentemente, difusão anômala em grande escala, isto é, observa-se um movimento errático entre partículas quando observadas individualmente, mas um certo padrão quando analisado um conjunto com N partículas. Os voos de certas espécies de aves, por exemplo, apresentam anomalias nas suas trajetórias (KLAFTER; SOKOLOV, 2005). O fenômeno transporte anômalo em plasma turbulento (BALESCU, 1995) também pode ser citado como exemplo. Nota-se, ainda, a difusão anômala em sistemas em meios porosos (KIMMICH, 2012), na difusão em estruturas fractais (STEPHENSON, 1995), na dinâmica populacional (GHAEMI; ZABIHINPOUR; ASGARI, 2009) e em alguns organismos (HILL; HÄDER, 1997). Na engenharia, a difusão anômala de partículas em certos materiais formados por metais ou cerâmica interfere no desempenho e na estabilidade de dispositivos eletrônicos que contêm tais materiais (KUMAR et al., 2018).

Para fenômenos, naturais ou não, observa-se que as causas para difusão anômala são as correlações de longo alcance entre partículas em movimento e/ou com o ambiente em pequena escala (SOKOLOV; KLAFTER, 2005). As partículas podem ter seus deslocamentos influenciados por interações com o ambiente, apresentando saltos maiores ou menores do que no movimento browniano clássico (SHLESINGER; ZASLAVSKY; FRISCH, 1995; ZABURDAEV; DENISOV; KLAFTER, 2015). A difusão anômala de proteínas, por exemplo, ocorre devido à aglomeração das mesmas no citoplasma das células (SCHÜTZ; SCHINDLER; SCHMIDT, 1997) e, recentemente, também foi proposto que entes não-inertes podem alterar fortemente a dinâmica de partículas citoplasmáticas (SABRI et al., 2019). Ainda no ramo da biologia, a difusão anômala afeta os cromossomos humanos, mais especificamente devido à viscoelasticidade do meio circundante a eles (PIERRO et al., 2018). O estado do ambiente pode impor às partículas que sejam limitadas a um local, de tal maneira que as probabilidades de observar longos períodos de descanso entre os deslocamentos sejam maiores do que o esperado no caso clássico. (METZLER; KLAFTER, 2000a). O efeito de gaiola das moléculas em um solvente (EINSTEIN, 1956; KRAMERS, 1940) é um exemplo em que isso ocorre.

Quando as correlações de longo alcance são devidas apenas à interação das partículas com o ambiente, espera-se que os modelos dinâmicos de grande escala sejam lineares. No entanto, se as correlações ocorrerem entre as próprias partículas observadas, espera-se que os modelos dinâmicos associados em larga escala sejam não-lineares (BOLOGNA; TSALLIS; GRIGOLINI, 2000). Um exemplo do último caso é a difusão em meios porosos, em que a posição de muitas partículas vizinhas influencia o tamanho do passo e o tempo de espera de uma partícula individual (TSALLIS; BUKMAN, 1996).

É possível utilizar diversas abordagens a fim de estudar sistemas de difusão anômalos. Existem diversos modelos teóricos e também modelos computacionais propostos. Para compreender o movimento em grande escala, o uso de equações diferenciais parciais (EDP's) mostra-se como uma importante ferramenta na descrição de certos fenômenos. Tais equações podem ser generalizadas através do Cálculo Fracionário (YANG et al., 2017).

A caminhada aleatória no tempo é utilizada na simulação do clássico browniano. O movimento é obtido por passeios aleatórios, um processo markoviano, no qual, em intervalos de tempo, a partícula realiza deslocamentos espaciais aleatórios determinados por uma PDF (METZLER; KLAFTER, 2000a). A caminhada aleatória em tempo contínuo, em inglês *Continuous Time Random Walk* (CTRW)[2], é uma extensão do passeio conhecido como random walk, introduzida por Elliott Waters Montroll e George Herbert Weiss (MONTROLL; WEISS, 1965) como generalização do processo de difusão para descrever a difusão anômala. Ela é caracterizada por um tempo de espera estocástico contínuo entre dois saltos e um valor estocástico contínuo para o comprimento do salto. Se as distribuições de comprimento de salto e tempo de espera forem projetadas de forma adequada, o CTRW consegue simular processos de difusão anômalos controlados de subdifusão a superdifusão (METZLER; KLAFTER, 2000b). O controle permanece em projetar as distribuições de comprimento de salto e tempo de espera, principalmente distribuições de lei de potência para ter momentos adequadamente divergentes. A média finita no tempo de espera e a variância divergente nas distribuições de comprimento do passo geram um processo de superdifusão. Por outro lado, média divergente no tempo de espera e variância finita para os comprimentos de passos geram uma subdifusão. Para ambos finitos, surge uma difusão clássica (PEREIRA et al., 2018).

A Lei de Zipf afirma que as propriedades da universalidade de uma classe de sistemas complexos comumente encontrados em sistemas sociais com um tipo de informação bem caracterizado em suas propriedades interativas e se organizam de tal forma que a informação contida cresce a uma taxa linear de tamanho de classificação crescente (CRISTELLI; BATTY; PIETRONERO, 2012). Tal comportamento gera uma maximização da entropia da informação e uma distribuição da lei de potência para a relação *rank-size* dos

[2]O acrônimo será usado em inglês porque é a forma mais facilmente reconhecida pelos leitores

sistemas (VISSER, 2013). As distribuições de Pareto, lei de potência, Lévy e q-Gaussiana são exemplos de tal classe de distribuições que decaem assintoticamente como distribuições de lei de potência. Conforme o valor do expoente, os momentos estatísticos da distribuição passam a ser divergentes, a distribuição é dita 'de cauda pesada' e, caso a variância não seja bem definida, o sistema é chamado de 'livre de escala'.

O ajuste de distribuições de decaimento da lei de potência tem sido objeto de estudo contínuo devido aos muitos problemas encontrados na recuperação imprecisa do parâmetro de potência, bem como de outros parâmetros de tais distribuições. Em dados empíricos, uma vez que os dados da cauda mostram uma frequência relativamente grande, mas ainda são raros o suficiente para descartar flutuações, os métodos usuais de mínimos quadrados fornecem uma estimativa pobre e imprecisa dos parâmetros e uma grande sensibilidade ao palpite inicial escolhido (CLAUSET; SHALIZI; NEWMAN, 2009). O ajuste linear realizado nas escalas logarítmicas também mostrou ser tendencioso (GOLDSTEIN; MORRIS; YEN, 2004). Além disso, devido às flutuações de frequência na cauda, existem problemas relativos aos bins de distribuição que se tornam cada vez mais esparsos à medida que a distribuição é observada (NEWMAN, 2005). Todo o conjunto de questões torna-se ainda mais crítico quando se considera dados com erros experimentais (KOEN; KONDLO, 2009).

A distribuição de Lévy é uma função especial, com uma cauda de lei de potência que pode ser escrita por uma integral em uma série de Fourier, que possui como característica o decaimento em forma de lei de potência. Métodos de integração numérica incluem um conjunto de parâmetros e uma escolha descuidada deles pode levar a distribuição a exibir uma periodicidade indesejável e ter um conjunto negativo de valores (EDWARDS et al., 2012).

Veremos no trabalho que a função de Lévy é livre de escala em geral e, por isso, apresenta uma dificuldade de ter dados representados por histogramas. A determinação precisa do expoente decrescente também é muito sensível aos valores dos parâmetros. Por estas razões, a definição e busca dos parâmetros devem ser contextualizadas no quadro da análise de Fourier. Existe, ainda, uma necessidade de discretização para implementação computacional. Além disso, seus parâmetros são sensíveis aos valores iniciais dos parâmetros ao aplicar o método do gradiente e a escolha arbitrária de parâmetros no ajuste incorre em custos. Outro fator é a presença de parâmetros inteiros que resiste a métodos gradientes, havendo, portanto, a necessidade de utilizar métodos heurísticos, como algoritmo genético (AG).

A proposta geral da pesquisa consiste no desenvolvimento de um aparato teórico capaz de identificar as características de um processo difusivo anômalo a partir de dados experimentais ou dados de simulações. Para validação deste aparato, a proposta de pesquisa é desenvolver simulações de processos difusivos anômalos lineares e não-lineares. Uma vez

validada, a proposta é a de aplicar esse aparato a dados reais ou a resultados de modelos simulados de processos como, por exemplo, as buscas de sítios no DNA por fatores de transcrição. A primeira etapa é a de abordar Equações Diferenciais Parciais Fracionárias (EDPF) para construção de um modelo teórico de difusão anômala e analisar situações que abrangem o caso linear, simulando voos de partículas, inicialmente, sem correlações de longo alcance e ajustando os dados em uma distribuição de Lévy.

A resolução de um modelo teórico generalizado também é realizada neste trabalho. Trata-se da resolução de um Problema de Valor de Contorno (PVC), que pode ser utilizado como abordagem ao modelar a difusão anômala para o caso geral. Mostraremos que a caminhada aleatória no tempo contínuo com média não-nula leva a uma equação de difusão e advecção fracionária. Em seguida, faremos ajustes para relacionar um modelo teórico com modelos computacionais, visando compreender fenômenos gerais que apresentam difusão anômala. Para isso, serão consideradas equações diferenciais parciais em um PVC com derivadas fracionárias no modelo teórico e simulações computacionais através do método de Monte Carlo pela função inversa para geração dos voos. Utilizamos, ainda, métodos para obtenção dos ajustes, incluindo um AG para casos de superdifusão.

1.1 Justificativa

Determinados fenômenos físicos com dinâmica complexa em pequena escala podem apresentar propriedades ou características cuja descrição fogem da clássica usual, descrita pela difusão normal. Como exemplo, é possível citar um caso especial de difusão anômala no ramo da medicina. Trata-se do movimento de uma partícula radioativa denominada coloide, que pode ser estudado através da dinâmica de Langevin e de um modelo de passeio aleatório, sendo seu entendimento importante para o projeto de tratamentos de câncer radioterapêutico baseados em partículas coloidais radioativas (WILSON, 2019).

Torna-se, dessa forma, relevante avançar na construção de um aparato de estudo para uma melhor compreensão de fenômenos físicos para os quais observa-se anomalia na difusão das partículas para caraterização dos processos difusivos e não lineares. Por exemplo, investigar qual o tipo de fenômeno em pequena escala pode-se esperar ou, dadas certas características em pequena escala, determinar qual modelo em grande escala deve melhor se adequar, procurando, também, qual é o método mais adequado para aferir parâmetros a partir de observações de modelos.

É importante destacar também que situações nas quais as distribuições de probabilidades apresentam cauda pesada estão presentes em diversos ramos da ciência, descritos através de uma lei de potência, sobretudo em se tratando de funções especiais, como é o caso da distribuição de Lévy. Tais situações necessitam de um aparato teórico para

recuperar dados de experimentos.

É possível citar a tentativa de entender o deslocamento praticado por alguns animais na busca de alimentos (BENHAMOU, 2007), o estudo estatístico da quantidade de palavras em uma linguagem (PIANTADOSI, 2014) e no estudo de redes (KRAPIVSKY; REDNER; LEYVRAZ, 2000), por exemplo. No nosso trabalho, mostramos que a solução do modelo teórico para superdifusão é uma função cujo comportamento tende a uma lei de potência, havendo, portanto, uma ligação entre este trabalho e o que é investigado em diversas áreas do conhecimento.

Um dos principais escopos deste estudo é a busca de dados que forneçam informações sobre difusão anômala de forma geral utilizando como ferramenta um modelo computacional, no qual são analisados parâmetros que permitem traduzir uma função real de variáveis reais dadas através de um funcional com integral imprópria em um modelo com somatório e variáveis discretizadas, exigindo, portanto, uma análise mais cuidadosa de fenômenos reais com os dados fornecidos pelo modelo.

No desenvolvimento do trabalho, foram abordados outros estudos referentes a tratamento de dados em decorrência da necessidade de manipular modelos computacionais e na realização de ajustes.

1.2 Organização do trabalho

No Capítulo 2, será feita uma revisão bibliográfica da teoria necessária para resolução analítica do problema teórico, abordando os tópicos de matemática e física necessários para este fim. Também são apresentados estudos teóricos realizados como fundamentação para modelos computacionais utilizados no trabalho. No Capítulo 3, serão abordadas as metodologias necessárias para as simulações computacionais e nos ajustes dos dados obtidos. No Capítulo 4, os principais resultados obtidos ao longo do estudo são analisados e interpretados. No Capítulo 5, será considerada uma situação aplicada de difusão anômala, na qual um estudo sobre a busca de fatores de transcrição para o alvo em cadeias de DNA é realizado. Finalmente, serão apresentadas as considerações finais e conclusões, além de sugestões para possíveis trabalhos futuros.

Capítulo 2

Fundamentação Teórica

Neste capítulo, serão apresentados os fundamentos necessários para o desenvolvimento deste trabalho, reunindo informações necessários para entender a modelagem de voos difusivos anômalos. Apresentamos, inicialmente, uma breve revisão dos conceitos básicos de processos estocásticos, incluindo a definição de caminhada aleatória, processo de Markov e processos de Lévy, fundamentais na modelagem de voos difusivos anômalos. Na sequência, e feita uma descrição detalhada da função de Lévy, seus principais conceitos e características, bem como sua aplicação na modelagem de voos difusivos anômalos. Em outra seção, são apresentadas as principais técnicas do Cálculo Fracionário (CF) para resolução analítica do modelo teórico de difusão anômala. Finalmente, serão abordados os conhecimentos matemáticos necessários para realização de simulações computacionais, como o Método de Monte Carlo pela função inversa e as funções de densidade de probabilidade que serão utilizadas. Também será feita uma introdução aos voos difusivos anômalos e seu comportamento característico em sistemas complexos, presentes em várias áreas, incluindo física, biologia, química e finanças. Finalmente, é feita uma breve discussão sobre as limitações e desafios de modelar voos difusivos anômalos, assim como realizar os ajustes dos dados simulados com os modelos teóricos conhecidos. O capítulo está organizado em nove seções: processos estocásticos, função de Lévy, equações de difusão clássicas, obtendo a equação de difusão generalizada, voos difusivos anômalos e simulação de voos sob ação de difusão anômala e o método de Monte Carlo pela função inversa, ajuste pelo método de análise por dispersão e Limitações e desafios.

2.1 Processos estocásticos

Os processos estocásticos são uma classe de modelos matemáticos usados para descrever o comportamento aleatório de sistemas complexos (SUMIN; DUSHKIN; SMOLENTSEVA, 2018). A caminhada aleatória é um processo estocástico simples em que uma partícula se move em intervalos de tempo discretos em uma direção aleatória.

O processo de Markov é um modelo estatístico que descreve a evolução de um sistema em termos de probabilidades de transição entre estados discretos, sendo que a probabilidade de transição depende apenas do estado atual e não dos estados anteriores. Isso significa que a evolução futura do processo depende apenas do estado atual. Alternativamente, é possível dizer que se trata de um processo em que a probabilidade de um evento futuro depende apenas do estado atual do sistema, não de sua história (PUTERMAN, 1990).

Matematicamente, o processo $\{X_n\}$ é chamado de processo de Markov se para cada n ocorrer:

$$P[X_{n+1} = j | X_0 = i_0, \ldots, X_n = i_n] = P[X_{n+1} = j | X_n = i_n\}, \tag{1}$$

com $j, i_0, \ldots, i_n \in \mathbb{N}$.

Os processos de Lévy são tipos especiais de processos de Markov (METZLER; KLAFTER, 2000b), com propriedades semelhantes às caminhadas aleatórias, mas permitem saltos de grande magnitude. Eles são caracterizados por sua função de densidade de probabilidade de salto, também conhecida como função de Lévy.

Esses processos estocásticos são fundamentais para a modelagem de voos difusivos anômalos, objetos de estudo deste trabalho. A caminhada aleatória pode ser usada, por exemplo, para descrever a trajetória de partículas em um meio poroso (LABOLLE; FOGG; TOMPSON, 1996). Já os processos de Lévy são especialmente importantes na modelagem de voos difusivos anômalos, pois permitem capturar as caudas pesadas da distribuição de probabilidade (PEINKE; BÖTTCHER; BARTH, 2004).

2.2 Função de Lévy

No estudo de fenômenos de difusão anômala, importantes modelos teóricos são considerados. Resolvendo tais modelos, surgem algumas funções especiais atendendo às equações diferenciais parciais e às condições de contorno. Neste contexto, é importante citar a função de Lévy, que será considerada neste estudo, em que trabalharemos sua forma natural e uma maneira adaptada para representá-la computacionalmente.

A função de Lévy é uma função matemática que descreve a distribuição de probabilidade dos saltos em um processo de Lévy (METZLER; KLAFTER, 2000b). Uma de suas características é a presença de caudas pesadas. Essa função tem sido amplamente utilizada na modelagem de voos difusivos anômalos, pois permite capturar a natureza complexa desses voos. Essa função é definida como o seguinte funcional:

$$L_\gamma(x,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-D|\kappa|^\gamma t}\cos(\kappa x)d\kappa. \tag{2}$$

Note que o integrando em (2) é uma função par na variável x. Então podemos reescrever essa equação como segue:

$$L_\gamma(x,t) = \frac{1}{\pi}\int_{0}^{+\infty} e^{-D\kappa^\gamma t}\cos(\kappa x)d\kappa. \tag{3}$$

Importante notar que a função de Lévy possui vários parâmetros, incluindo o índice de estabilidade, que determina a assímetria da distribuição, e o parâmetro de escala, que controla a magnitude dos saltos.

2.2.1 Discretização da Função de Lévy e análise de erros

Para fins de utilização em um modelo computacional, a função de Lévy $L_\gamma(x,t)$ (3), citada na seção anterior, deve ser discretizada. Isso porque a função deve ser implementada e suas variáveis não podem ser representadas no universo dos números reais em um sistema computacional. Por isso, é realizada a mudança $\kappa = \dfrac{n\pi}{\Lambda}$ e é também fixado um valor máximo, Ω, no limite de integração. Além disso, na discretização o diferencial $d\kappa$ torna-se $\Delta\kappa$ e

$$\Delta\kappa = \frac{(n+1)\pi}{\Lambda} - \frac{n\pi}{\Lambda} = \frac{\pi}{\Lambda}. \tag{4}$$

Tem-se então a função discretizada:

$$l_\gamma(x,t) = \frac{1}{\pi}\sum_{n=0}^{\Omega} e^{-D|\frac{n\pi}{\Lambda}|^\gamma t} cos\left(\frac{n\pi}{\Lambda}x\right)\frac{\pi}{\Lambda}. \tag{5}$$

Reescrevendo a Equação (5), segue que:

$$l_\gamma(x,t) = \frac{1}{\Lambda}\sum_{n=0}^{\Omega} e^{-D|\frac{n\pi}{\Lambda}|^\gamma t} cos\left(\frac{n\pi}{\Lambda}x\right). \tag{6}$$

É de interesse neste trabalho que os erros numéricos que ocorrem na discretização e no truncamento do limite de integração sejam determinados e controlados.

É possível mostrar que o período em x de (6) é igual a 2Λ. Parte-se da premissa de que o parâmetro Λ deve ser tal que os dados simulados permaneçam no intervalo $\left[-\frac{\Lambda}{2}, \frac{\Lambda}{2}\right]$, ou seja, $x \in \left[-\frac{\Lambda}{2}, \frac{\Lambda}{2}\right]$. Como $\Delta k = \dfrac{\pi}{\Lambda}$, por analogia, $\Delta x = \dfrac{1}{\Omega}$.

2.2.2 Erro de truncamento

Estimaremos agora o erro ocorrido devido ao truncamento. Para isso, considere (a_n) uma sequência de termos positivos para todo $n \in \mathbb{N}$ e tal que exista N_0 para o qual $f(n) = a_n$ é contínua e decrescente para $n > N_0$. Com estas hipóteses, o Teste da Integral (STEWART, 2013) garante a equivalência:

$$\sum_{n=N_0}^{+\infty} a_n \text{ converge se, e somente se, } \int_{N_0}^{+\infty} f(x)dx \text{ converge.} \tag{7}$$

O resultado acima pode ser demonstrado através das desigualdades:

$$\int_{N_0}^{M+1} f(x)dx \leq \sum_{n=N_0}^{M} a_n \leq a_1 + \int_{N_0}^{M} f(x)dx. \tag{8}$$

Se a série é convergente, as desigualdades (8) se mantêm mesmo quando M tende a $+\infty$.

$$\int_{N_0}^{+\infty} f(x)dx \leq \sum_{n=N_0}^{+\infty} a_n \leq a_1 + \int_{N_0}^{+\infty} f(x)dx. \tag{9}$$

Para estimar o erro de truncamento da série convergente $s = \sum_{n=1}^{+\infty} a_n$, observe que:

$$\sum_{n=1}^{+\infty} a_n = \sum_{n=1}^{M} a_n + \sum_{n=M+1}^{+\infty} a_n, \tag{10}$$

ou seja, é válida a igualdade:

$$s = S_n + R_n, \tag{11}$$

na qual S_n é uma soma parcial e R_n, o resto. O erro ϵ é dado por $\epsilon = |R_n| = |s - S_n|$.

Da segunda desigualdade em (9) e da Equação (10), tem-se:

$$\sum_{n=1}^{M} a_n + \sum_{n=M+1}^{+\infty} a_n \leq a_1 + \int_1^{+\infty} f(x)dx. \tag{12}$$

ou ainda,

$$\sum_{n=M+1}^{+\infty} a_n \leq a_1 + \int_1^{+\infty} f(x)dx - \sum_{n=1}^{M} a_n. \tag{13}$$

Logo,

$$R_n \leq a_1 + \int_1^{+\infty} f(x)dx - \sum_{n=1}^{M} a_n. \tag{14}$$

É possível estimar o erro máximo ocorrido ao truncar a série ao implementar, computacionalmente, a função de Lévy. Isso é feito como segue:

$$\begin{aligned} l_\gamma(x,t) &= \frac{1}{\Lambda} \sum_{n=0}^{+\infty} e^{-D\left|\frac{n\pi}{\Lambda}\right|^\lambda t} \cos\left(\frac{n\pi x}{\Lambda}\right) \\ &= \frac{1}{\Lambda} \sum_{n=0}^{\Omega} e^{-D\left|\frac{n\pi}{\Lambda}\right|^\lambda t} \cos\left(\frac{n\pi x}{\Lambda}\right) + \frac{1}{\Lambda} \sum_{n=\Omega+1}^{+\infty} e^{-D\left|\frac{n\pi}{\Lambda}\right|^\lambda t} \cos\left(\frac{n\pi x}{\Lambda}\right) \end{aligned}$$

Usando a notação R_Ω para a segunda parcela do lado direito da igualdade e supondo que a série é absolutamente convergente, segue que:

$$\begin{aligned} R_\Omega &= \frac{1}{\Lambda} \sum_{n=\Omega+1}^{+\infty} e^{-D\left|\frac{n\pi}{\Lambda}\right|^\lambda t} \cos\left(\frac{n\pi x}{\Lambda}\right) \\ &\leq \frac{1}{\Lambda} \sum_{n=\Omega+1}^{+\infty} \left| e^{-D\left|\frac{n\pi}{\Lambda}\right|^\lambda t} \cos\left(\frac{n\pi x}{\Lambda}\right) \right| \end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{\Lambda}\int_{\Omega}^{+\infty} e^{-D\left|\frac{\epsilon\pi}{\Lambda}\right|^{\lambda}t}\left|\cos\left(\frac{\epsilon\pi x}{\Lambda}\right)\right| d\epsilon \\
&\leq \frac{1}{\Lambda}\int_{\Omega}^{+\infty} e^{-D\left|\frac{\epsilon\pi}{\Lambda}\right|^{\lambda}t} d\epsilon \\
&\leq \frac{1}{\Lambda}\int_{\Omega}^{+\infty} e^{-D\left|\frac{\epsilon\pi}{\Lambda}\right|t} d\epsilon \\
&\leq \frac{1}{\Lambda}\int_{\Omega}^{+\infty} e^{-D\frac{\epsilon\pi}{\Lambda}} d\epsilon \\
&\leq \frac{1}{\Lambda}\lim_{b\to+\infty}\int_{\Omega}^{b} e^{-D\frac{\epsilon\pi}{\Lambda}} d\epsilon \\
&\leq \frac{1}{\Lambda}\lim_{b\to+\infty}\left(\left.\frac{e^{-D\frac{\epsilon\pi}{\Lambda}}}{\frac{-D\pi}{\Lambda}}\right|_{\Omega}^{b}\right) \\
&\leq \frac{1}{\Lambda}\lim_{b\to+\infty}\left(\frac{e^{-D\frac{b\pi}{\Lambda}}}{\frac{-D\pi}{\Lambda}} - \frac{e^{-D\frac{\Omega\pi}{\Lambda}}}{\frac{-D\pi}{\Lambda}}\right) \\
&\leq \frac{1}{\Lambda}\frac{e^{-D\frac{\Omega\pi}{\Lambda}}}{\frac{D\pi}{\Lambda}} \\
&\leq \frac{e^{-D\frac{\Omega\pi}{\Lambda}}}{D\pi}
\end{aligned}
$$

2.3 Modelagem de voos difusivos anômalos

Existem diferentes formas de modelar voos difusivos anômalos. Uma das técnicas mais comuns é a modelagem baseada na função de Lévy, que descreve a distribuição de probabilidade dos saltos em um processo de Lévy. Ela é usada em conjunto com a transformada de Fourier para obter a função de densidade de probabilidade correspondente (MARKSTEINER; ELLINGER; ZOLLER, 1996).

Modelos de voos difusivos anômalos são caracterizados por vários parâmetros, incluindo o índice de estabilidade e o parâmetro de escala da função de Lévy. Esses parâmetros desempenham um papel crítico na descrição do comportamento dos voos difusivos anômalos e devem ser escolhidos com cuidado.

Outra abordagem possível é a modelagem baseada em cálculo fracionário, que envolve a substituição da derivada temporal por uma derivada fracionária. Essa técnica tem sido amplamente utilizada na modelagem de processos estocásticos como voos difusivos anômalos (MEERSCHAERT, 2012). Nesta seção, discutiremos sobre algumas funções especiais e sobre os principais operadores utilizados no cálculo fracionário.

2.3.1 Funções especiais do cálculo fracionário

A resolução de Equações Diferenciais Parciais Fracionárias (EDPF's), utilizadas como modelos da difusão anômala, demanda alguns conhecimentos prévios de um ramo da Matemática denominado Cálculo Fracionário (CF), que, resumidamente, estuda derivadas não inteiras de funções reais e complexas (HERRMANN, 2011). Primeiramente, serão apresentadas algumas funções especiais que são base do CF. Essas são generalizações de algumas funções conhecidas no cálculo clássico. Em seguida, serão apresentadas algumas definições de derivadas fracionárias utilizadas no trabalho.

2.3.2 Função Gama de Euler

A função Gama de Euler pode ser entendida como uma generalização do conceito de fatorial de um número natural, $n!$, de forma que no seu domínio são permitidos valores não-inteiros, incluindo até mesmo complexos (PODLUBNY, 1998).

Seja z um número complexo com a parte real estritamente positiva. A função Gama de Euler, $\Gamma(z)$, é definida através da integral imprópria:

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt, \quad z \neq 0, -1, -2, \ldots \tag{15}$$

Utilizando transformadas de Laplace e a definição da função Gama, obtém-se:

$$B(z,w) = \int_0^1 \tau^{z-1}(1-\tau)^{w-1} d\tau = \frac{\Gamma(z)\Gamma(w)}{\Gamma(z+w)}, \qquad \Re(z) > 0,\ \Re(w) > 0.$$

conhecida como função Beta (DAVIS, 1972).

2.3.3 Funções de Mittag-Leffler

Assim como a função exponencial está presente na resolução de equações diferenciais ordinária com coeficientes constantes, as funções de Mittag-Leffler têm um papel importante no estudo de EDFs lineares com coeficientes constantes (PODLUBNY, 1998; CAMARGO; CAPELAS, 2015). Essas são dadas por:

$$E_{\alpha,\beta}(z) = \sum_{k=0}^{+\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}, \qquad \alpha > 0, \qquad \beta > 0. \tag{16}$$

Considerando $\beta = 1$ em (16), tem-se a função de Mittag-Leffler a um parâmetro

$$E_\alpha(z) = \sum_{k=0}^{+\infty} \frac{z^k}{\Gamma(\alpha k + 1)}. \tag{17}$$

Para $\alpha = 1$, essa função coincide com a exponencial $f(z) = e^z$.

2.3.4 Operadores integro-diferenciais

Serão apresentados na sequência diferentes operadores utilizados no Cálculo Fracionário com intuito de generalizar as derivadas de ordens inteiras conhecidas no Cálculo tradicional. O operador de integral a ser utilizado será denotado por D^{-p} e o de derivadas será D^p.

2.3.4.1 O operador de integral fracionária de Riemann-Liouville

A primeira fórmula de Liouville para a derivada fracionária tem, como ponto de partida, o resultado conhecido para derivadas de ordem inteira (YANG; ZHANG, 2019):

$$\frac{d^k}{dx^k}[e^{ax}] = a^k e^{ax}, \quad k \in \mathbb{Z}_+,$$

a qual Liouville estendeu para derivadas de ordem arbitrárias:

$$D_x^p[e^{ax}] = \frac{d^p}{dx^p}[e^{ax}] = a^p e^{ax}, \quad p \in \mathbb{R}_+.$$

Seja f uma função que pode ser expandida em série, da seguinte forma:

$$f(x) = \sum_{n=0}^{+\infty} c_n e^{a_n x}. \tag{18}$$

Então

$$D_x^p f(x) = \sum_{n=0}^{+\infty} c_n (a_n)^p e^{a_n x}.$$

Essa é conhecida como a primeira fórmula de Liouville para a derivada fracionária. No entanto, a necessidade de convergência da série impõe restrições sobre os valores de p, o que torna essa definição limitada. Diante disso, Liouville formula um segundo conceito como segue:

$$D^{-p}[e^{ax}] = \int_0^{+\infty} u^{a-1}e^{-ux}du, \quad a > 0 \quad x > 0. \tag{19}$$

Fazendo a substituição $t = xu$, a integral em (19) pode ser reescrita da forma:

$$D^{-p}[e^{ax}] = \int_0^{+\infty} \left(\frac{t}{x}\right)^{a-1} e^{-t}\frac{dt}{x} = \frac{1}{x^a}\int_0^{+\infty} t^{a-1}e^{-t}dt = x^{-a}\Gamma(a). \tag{20}$$

De (19) e (20), segue que:

$$\int_0^{+\infty} u^{a-1}e^{-ux}du = x^{-a}\Gamma(a). \tag{21}$$

Aplicando o operador D^p em ambos os lados de (21), segue que:

$$\begin{aligned} D^p\left[\int_0^{+\infty} u^{a-1}e^{-ux}du\right] &= D^p\left[x^{-a}\Gamma(a)\right]. \\ D^p\left[\int_0^{+\infty} u^{a-1}e^{-ux}du\right] &= \Gamma(a)D^p\left[x^{-a}\right]. \\ \Gamma(a)D^p\left[x^{-a}\right] &= (-1)^p\int_0^{+\infty} u^{a-1+p}e^{-ux}du. \end{aligned}$$

Portanto,

$$D^p\left[x^{-a}\right] = \frac{(-1)^p}{\Gamma(a)}\int_0^{+\infty} u^{a-1+p}e^{-ux}du, p > 0, a > 0. \tag{22}$$

Fazendo a mudança $t = xu$ em (22), conclui-se que:

$$\begin{aligned} D^p\left[x^{-a}\right] &= \frac{(-1)^p}{\Gamma(a)}\int_0^{+\infty}\left(\frac{1}{x}\right)^{a-1+p} e^{-t}\frac{dt}{x} \\ &= \frac{(-1)^p}{\Gamma(a)}x^{-a-p}\int_0^{+\infty} t^{a-1+p}e^{-t}dt. \end{aligned}$$

Portanto,

$$D^p\left[x^{-a}\right] = \frac{(-1)^p\Gamma(a+p)}{\Gamma(a)}x^{-a-p}, \qquad p > 0, a > 0. \tag{23}$$

Essa segunda definição aplica-se, no entanto, somente a funções do tipo x^{-a}.

Por sua vez, o matemático alemão Bernhard Riemann (1826-1866), ao procurar por um tipo de generalização da série de Taylor, chega à seguinte expressão:

$$D^{-p}f(x) = \frac{1}{\Gamma(p)} \int_c^x [(x-t)^{(p-1)} f(t) + g(x)]dt, \tag{24}$$

em que $g(x)$ é uma função adicionada para os casos onde o limite inferior c, da lei da soma dos expoentes, ${}_cD_x^{-u}\ {}_cD_x^{-p}f(x) =_c D_x^{-u-p}f(x)$, não serem iguais. Observe que esta teoria é levada à uma contradição no caso em que $x = 0$.

Talvez, essa discussão sobre qual seria o operador fracionário mais adequado tenha sido encerrada por Nikolay Yakovlevich Sonin (1849-1915), ao considerar a integral de Cauchy, escrevendo a derivada n-ésima

$$D^n f(z) = f^{(n)}(z) = \frac{n!}{2\pi i} \int_C \frac{f(t)}{(t-z)^{n+1}} dt. \tag{25}$$

É possível notar que a ordem n da derivada é um número inteiro faz com que o integrando possua um polo de ordem $n+1$, supondo que $f(z)$ seja uma função analítica. Para resolver, pode-se trocar n por $p \in \mathbb{R}$, pois $n! = \Gamma(n+1)$. Feito isso, o denominador do integrando não possui mais um polo quando p for um número real. Em vez disso, ele é agora caracterizado por uma ramificação quando $z = t$, de modo que, para contornar esse problema, deve-se mudar a curva de contorno C para alguma do tipo Bromwich (PODLUBNY, 1998).

Outra maneira de se obter a derivada de ordem p arbitrária é considerar f integrada duas vezes em um intervalo (c, x) (PEREIRA, 2018):

$$\begin{aligned}
D^{-2}f(x) &= \int_c^x dx_1 \int_c^{x_1} f(\xi)d\xi \\
&= \int_c^x f(\xi)d\xi \int_\xi^x dx_1 \\
&= \int_c^x f(\xi)d\xi \cdot x_1|_\xi^x \\
&= \int_c^x (x-\xi)f(\xi)d\xi \\
&= \int_c^x \left[\int_c^{x_1} f(\xi)d\xi\right] dx_1 \\
&= \int_c^x f(\xi)d\xi \int_\xi^x dx_1 x_1.
\end{aligned}$$

Agora, integrada três vezes, tem-se:

$$
\begin{aligned}
D^{-3}f(x) &= \int_c^x \int_c^{x_2} \left[\int_c^{x_1} f(\xi)d\xi\right] dx_1 dx_2 \\
&= \int_c^x \left[\int_c^{x_2} \int_\xi^{x_2} f(\xi)dx_1 d\xi\right] dx_2 \\
&= \int_c^x \int_c^{x_2} f(\xi)d\xi \cdot x_1|_\xi^{x_2}\, dx_2 \\
&= \int_c^x \left[\int_c^{x_2} (x_2-\xi)f(\xi)d\xi\right] dx_2 \\
&= \int_c^x \int_\xi^x (x_2-\xi)f(\xi)dx_2 d\xi \\
&= \int_c^x f(\xi)d\xi\, \frac{(x-\xi)^2}{2}\bigg|_\xi^x = \int_c^x f(\xi)d\xi \frac{(x-\xi)^2}{2}.
\end{aligned}
$$

Portanto,

$$
D^{-3}f(x) = \int_c^x \frac{(x-\xi)^2}{2} f(\xi)d\xi. \tag{26}
$$

Integramos agora quatro vezes

$$
\begin{aligned}
D^{-4}f(x) &= \int_c^x \int_c^{x_3} \int_c^{x_2} \left[\int_c^{x_1} f(\xi)d\xi\right] dx_1 dx_2 dx_3 \\
&= \int_c^x \int_c^{x_3} \int_c^{x_2} f(\xi) \left[\int_\xi^{x_2} dx_1\right] d\xi dx_2 dx_3 \\
&= \int_c^x \int_c^{x_3} \int_c^{x_2} f(\xi)d\xi(x_2-\xi)dx_2 dx_3 \\
&= \int_c^x \int_c^{x_3} \int_c^{x_2} (x_2-\xi)f(\xi)d\xi dx_2 dx_3 \\
&= \int_c^x \int_c^{x_3} f(\xi) \int_\xi^{x_3} (x_2-\xi)dx_2 d\xi dx_3 \\
&= \int_c^x \int_c^{x_3} f(\xi)\, \frac{(x_2-\xi)^2}{2}\bigg|_\xi^{x_3} d\xi dx_3 \\
&= \int_c^x \int_c^{x_3} f(\xi)\frac{(x_3-\xi)^2}{2} d\xi dx_3 \\
&= \int_c^x f(\xi) \int_\xi^x \frac{(x_3-\xi)^2}{2} dx_3 d\xi \\
&= \int_c^x f(\xi)\, \frac{(x_3-\xi)^3}{3.2}\bigg|_\xi^x d\xi.
\end{aligned}
$$

Portanto,

$$D^{-4}f(x) = \int_c^x \frac{(x-\xi)^3}{3!} f(\xi)d\xi. \tag{27}$$

É possível generalizar para ordem n qualquer, como segue:

$$D^{-n}f(x) = \int_c^x \frac{(x-\xi)^{n-1}}{(n-1)!} f(\xi)d\xi, \quad n \in \mathbb{N}. \tag{28}$$

Como $\Gamma(n) = (n-1)!$, segue que:

$$D^{-n}f(x) = \frac{1}{\Gamma(n)} \int_c^x \frac{f(\xi)}{(x-\xi)^{1-n}} d\xi, \quad n \in \mathbb{N}. \tag{29}$$

Substituindo n por um número real positivo p, tem-se:

$$D^{-p}f(x) = \frac{1}{\Gamma(p)} \int_c^x \frac{f(\xi)}{(x-\xi)^{1-p}} d\xi, \quad Re(p) > 0. \tag{30}$$

Essa é a Integral de RL, que será denotada por ${}_c^{RL}D_x^{-p}f(x)$.

Considerando $c = 0$ em (30), retoma-se a definição de Riemann (24), mas sem a função complementar $g(x)$. Quando $c \to \infty$ recupera-se a primeira definição de Liouville para a classe de funções x^{-a} em (23).

Passando para a derivada de ordem fracionária, consideremos (30) para escrever:

$${}_c^{RL}D_x^{-p}f(x) = \frac{1}{\Gamma(p)} \int_c^x (x-\xi)^{p-1} f(\xi)d\xi, \quad Re(p) > 0. \tag{31}$$

Trocando $-p$ por p, tem-se:

$${}_c^{RL}D_x^{p}f(x) = \frac{1}{\Gamma(-p)} \int_c^x (x-\xi)^{-p-1} f(\xi)d\xi, \quad Re(p) > 0. \tag{32}$$

Faz-se a mudança de variáveis $p = k - v$, sendo k o menor inteiro maior que p e, então

$${}_c^{RL}D_x^{p}f(x) =_c D_x^{k-v}f(x) =_c D_x^{k}f(x)_c D_x^{-v}f(x) \tag{33}$$

$$= \frac{d^k}{dx^k}\left[\frac{1}{\Gamma(v)}\int_c^x (x-\xi)^{p-1} f(\xi)d\xi\right] \tag{34}$$

$$= \frac{1}{\Gamma(k-p)}\frac{d^k}{dx^k}\int_c^x \frac{f(\xi)}{(x-\xi)^{p-k+1}}d\xi, \tag{35}$$

onde $k = \lfloor p \rfloor$, ou seja, o menor inteiro maior ou igual a p.

Portanto, temos os operadores integro-diferenciais segundo RL:

$$\text{Integral: } {}_c^{RL}D_x^{-p} f(x) = \frac{1}{\Gamma(p)}\int_c^x (x-\xi)^{p-1} f(\xi)d\xi, \quad \text{Re}(p) > 0. \tag{36}$$

$$\text{Derivada: } {}_c^{RL}D_x^{p} f(x) = \frac{1}{\Gamma(k-p)}\frac{d^k}{dx^k}\int_c^x \frac{f(\xi)}{(x-\xi)^{p-k+1}}d\xi, \quad k = \lceil p \rceil. \tag{37}$$

De (37), obtém-se:

$${}_c^{RL}D_x^{p} f(x) = \frac{d^k}{dx^k}\,{}_c D_x^{-(k-p)} f(x). \tag{38}$$

Assim, a derivada de ordem arbitrária, segundo RL, equivale à derivada de ordem inteira (k) de uma integral de ordem arbitrária.

2.3.5 O operador diferencial fracionário de Caputo

A ordem na qual os operadores integração de ordem não inteira e a derivada de ordem inteira são aplicados se diferem nas derivadas de Caputo e de RL. Isso gera diferenças no resultado da derivada de algumas funções e nas aplicações em determinados fenômenos físicos (PODLUBNY, 1998).

O operador derivada fracionária segundo Caputo, que denotaremos por ${}_c^C D_x^p$, é definido como segue:

$${}_c^C D_x^p f(x) = \frac{1}{\Gamma(k-p)}\int_c^x \frac{f^{(k)}(\xi)}{(x-\xi)^{p+1-k}}d\xi, \tag{39}$$

com $k-1 < p < k$.

Uma vantagem do uso desse operador é que nas equações diferenciais fracionárias (EDFs) as condições impostas podem ser escritas como derivadas de ordem inteira.

2.3.6 O operador de Riesz-Weyl

O uso do operador de Riesz-Weyl simplifica problemas quando se trabalha com o método das transformadas integrais. O mesmo é escrito partindo dos operadores diferenciais segundo RL e Caputo, tomando-se $c \to -\infty$ (SAMKO, 2003).

$$\begin{aligned}
{}^{RL}_{-\infty}D^p_x f(x) &= {}^{C}_{-\infty}D^p_x f(x) \\
&= \frac{1}{\Gamma(k-p)} \int_{-\infty}^{x} \frac{f^{(k)}(\xi)}{(x-\xi)^{p-1-k}} d\xi \\
&= {}^{RL}_{-\infty}D^{-(k-p)}_x f^{(k)}(x) \\
&= {}^{RW}_{-\infty} D^p_x f(x).
\end{aligned}$$

Esse operador tem a vantagem de que, em uma transformada de Fourier, sua forma se mantém simples devido ao fato de $c \to -\infty$.

2.3.7 Transformadas integrais

A metodologia das transformadas integrais permite escrever uma equação diferencial na forma de equação algébrica, normalmente, de mais simples resolução do que a equação original. Resolve-se a equação algébrica e, por meio da respectiva transformada inversa, recupera-se a solução da equação diferencial original (LENZI et al., 2009), (PODLUBNY, 1998). De forma geral, as transformadas integrais são definidas como:

$$\int_I k(s,t) f(t) dt, \tag{40}$$

sendo $k(s,t)$ o núcleo da transformada.

Para a resolução das equações diferenciais fracionárias no tempo e no espaço são, particularmente, importantes a transformada de Laplace e a transformada de Fourier, pois permitem escrever as equações diferenciais na forma de equações algébricas. Ambas serão utilizadas, respectivamente, neste trabalho, para eliminar a dependência temporal ($t > 0$) e para a parte espacial ($-\infty < x < +\infty$). Para a transformada de Laplace tem-se, em (40), $k(s,t) = e^{-st}$ e o intervalo $I = [0, +\infty)$. Na transformada de Fourier, o núcleo é e^{ikt} e o intervalo é $I = (-\infty, +\infty)$.

Neste trabalho, denotaremos a transformada de Laplace de uma função $f(t)$ por $\mathcal{L}\{f(t)\}$ e a transformada de Fourier de $f(t)$ por $\mathcal{F}\{f(t)\}$.

2.3.8 Propriedades necessárias das transformadas de Laplace e de Fourier

Determinadas propriedades são essenciais nas resoluções de problemas de valores iniciais, usados em diversos modelos matemáticos. Citaremos aqui algumas utilizadas na resolução de modelos teóricos do problema estudado neste trabalho.

Para n-ésima derivada de uma função $f(t)$ de classe C^n, tem-se as seguintes fórmulas para transformadas de Laplace e de Fourier, respectivamente:

$$\mathcal{L}\{f^{(n)}(t)\} = s^n\mathcal{L}\{f(t)\} - s^{n-1}f(0) - s^{n-2}f'(0) - \cdots - f^{(n-1)}(0), \tag{41}$$

$$\mathcal{F}\{f^{(n)}(z)\} = i^n k^n \mathcal{F}\{f(z)\}, \tag{42}$$

responsáveis por transformar as equações diferenciais em equações algébricas (BOYCE; DIPRIMA, 1985). Ambas podem ser demonstradas através da integração por partes (OLIVEIRA et al., 2014).

Sejam f e g duas funções no espaço $L^1(\mathbb{R})$, ou seja, funções integráveis em $\mathbb{R}$. A convolução de f e g é definida como:

$$h(x) = \int f(x-\tau)g(\tau)d\tau. \tag{43}$$

A Operação (43), denotada por $f * g$, é comutativa, associativa e distributiva (KATZNELSON, 2004).

Alguns resultados serão necessários na resolução de equações diferencias que modelam o problema a ser estudado. Destacam-se os teoremas da convolução utilizados quando é preciso recuperar o espaço de funções em t. Esses resultados são mostrados a seguir.

Teorema (Convolução): Sejam f e g funções integráveis. Se $F(k) = \mathcal{F}\{f(z)\}$ e $G(k) = \mathcal{F}\{g(z)\}$ denotam as transformadas de Fourier das funções $f(z)$ e $g(z)$, respectivamente, então:

$$\mathcal{F}\{f(z) * g(z)\} = F(k)G(k). \tag{44}$$

Teorema (Convolução): Sejam $F(s) = \mathcal{L}\{f(t)\}$ e $G(s) = \mathcal{L}\{g(t)\}$ as transformadas de Laplace das funções $f(t)$ e $g(t)$, respectivamente, sendo f e g funções contínuas por parte em $[0, \infty)$. Então:

$$\mathcal{L}\{f(t) * g(t)\} = F(s)G(s). \tag{45}$$

O operador de integral a ser utilizado a seguir será denotado por D^{-p} e o de derivadas será D^p. Considere a Integral de Riemann-Liouville (RL), que será denotada por ${}_c^{RL}D_x^{-p}f(x)$.

$$D^{-p}f(x) = \frac{1}{\Gamma(p)} \int_c^x \frac{f(\xi)}{(x-\xi)^{1-p}} d\xi, \qquad Re(p) > 0. \tag{46}$$

É possível verificar que o operador-fracionário de RL pode ser escrito como a convolução das funções $g(t) = t^{p-1}$ e $f(t)$ quando $a = 0$.

$${}_c^{RL}D_t^{-p}f(t) = \frac{1}{\Gamma(p)} \int_0^t (t-\tau)^{p-1} f(\tau) d\tau = \frac{t^{p-1} * f(t)}{\Gamma(p)}. \tag{47}$$

Além disso, a transformada de Laplace da função $f(t) = t^{p-1}, p \in \mathbb{R}_+^*$, é dada por (PODLUBNY, 1998):

$$\mathcal{L}\{t^{p-1}\} = \frac{\Gamma(p)}{s^{-p}}. \tag{48}$$

Utilizando as igualdades em (47) e (48), obtém-se:

$$\mathcal{L}\{{}_c^{RL}D_t^{-p}f(t)\} = \frac{1}{s^p}\mathcal{L}\{f(t)\}. \tag{49}$$

Considere a função $f(t)$ definida em $\mathbb{R}_+$ e sua transformada de Laplace $\mathcal{L}\{f(t)\} = \tilde{f}(s)$ definida para $s \geq c$, $c \in \mathbb{R}$. Caso $p > 0$ com $n = p$, utilizando a transformada da convolução (45) e os operadores íntegro-diferenciais segundo RL, então para $s >$ máx$\{0, c\}$, a transformada de Laplace da derivada fracionária de RL é dada por (PODLUBNY, 1998):

$$\mathcal{L}\{{}_0^{RL}D_t^p f(t)\} = s^p \tilde{f}(s) - \sum_{k=0}^{n-1} s^k [{}_0^{RL}D_t^{p-k-1} f(t)]_{t=0}, \qquad n-1 \leq p < n, \tag{50}$$

desde que existam no ponto $t = 0$ a integral e as derivadas de ordem não inteira.

Com relação à derivada de Caputo, com $a = 0$ e $p > 0$, é possível escrevê-la em termos da convolução entre a derivada de ordem n da função $f(t)$ e a função $g(t) = t^{n-p-1}$ (SALGADO, 2015)

$$ {}_0^C D_t^p f(t) = \frac{t^{n-p-1} * f^{(n)}(t)}{\Gamma(n-p)}. \tag{51} $$

Assim, a transformada de Laplace da derivada de Caputo pode ser obtida através da propriedade da transformada da convolução entre as funções t^{n-p-1} e a função $f^{(n)}(t)$, sabendo que:

$$ \mathcal{L}\{t^{n-p-1}\} = \frac{\Gamma(n-p)}{s^{n-p}}. \tag{52} $$

Tomando uma função $f(t)$ definida em $\mathbb{R}_+$ e diferenciável até ordem $n-1$ em $t = 0^+$ e escrevendo a derivada de Caputo como em (51), ao aplicar a transformada de Laplace, concluímos que a transformada de Laplace da derivada fracionária segundo Caputo é dada por:

$$ \mathcal{L}\{{}_0^C D_t^p f(t)\} = s^p \tilde{f}(s) - \sum_{k=0}^{n-1} s^{p-k-1} f^{(k)}(0), \qquad n-1 < p \leq n. \tag{53} $$

Uma vantagem da utilização desse último operador integro-fracionário é o fato de haver constantes de integração que aparecem no processo de solução de equações diferenciais, que podem ser determinadas por condições iniciais e/ou condições de contorno do problema, que pode ser expressa como derivadas de ordem inteira (PEREIRA, 2018).

A transformada de Fourier da n-ésima derivada da função $h(x)$ com $x \in \mathbb{R}$ é dada por:

$$ \mathcal{F}\{h^{(n)}(x); k\} = (-ik)^n \hat{h}(k). \tag{54} $$

Considere a integral de RL com $a = -\infty$ e $0 < p < 1$. Assim,

$$ {}_{-\infty}^{RL} D_x^{-p} g(x) = \frac{1}{\Gamma(p)} \int_{-\infty}^{x} (x-\tau)^{p-1} g(\tau) d\tau, \qquad 0 < p < 1. \tag{55} $$

Se $h(t) = \dfrac{t^{p-1}}{\Gamma(p)}$, então a transformada de Laplace de $h(t)$ é dada por:

$$ \mathcal{L}\{h(t)\} = \frac{1}{\Gamma(p)} \int_0^{+\infty} t^{p-1} e^{-st} dt = s^{-p}, \qquad 0 < p < 1. \tag{56} $$

Essa integral é convergente se $0 < p < 1$, no caso $s = -ik$ e $k \in \mathbb{R}$.

Para função $h_+(x) = \begin{cases} \dfrac{x^{p-1}}{\Gamma(p)}, \text{ se } x > 0 \\ 0, \text{ se } x \leq 0 \end{cases}$, a transformada de Fourier é dada por:

$$\mathcal{F}\{h_+(t)\} = (-iw)^{-p}, \qquad 0 < p < 1. \tag{57}$$

É possível então escrever a integral fracionária de RL como a convolução das funções $h_+(x)$ e $g(x)$ como segue:

$$ {}^{RL}_{-\infty}D_x^{-p} g(x) = h_+(x) * g(x). \tag{58}$$

A transformada de Fourier é dada por

$$\mathcal{F}\{{}^{RL}_{-\infty}D_x^{-p} g(x)\} = (ik)^{-p}\hat{g}(k). \tag{59}$$

Por outro lado, para o operador diferencial segundo RL tem-se que $p > 0$ e $p = k - v$ com $k = \lceil p \rceil$, que denota a função menor inteiro maior ou igual do que p. Dessa forma, é possível escrever:

$$\mathcal{F}\{{}^{RL}_{-\infty}D_x^{-p} g(x); k\} = \mathcal{F}\{{}^{RW}_{-\infty}D_x^{-(k-p)} g^{(k)}(x); k\} = \mathcal{F}\left\{\frac{1}{\Gamma(k-p)}\int_{-\infty}^{x}(x-\tau)^{(\tau-p)-1} g(\tau)d\tau; k\right\}. \tag{60}$$

Sabendo que a integral em (60) é convergente se $k - 1 < p < 1$, então a derivada fracionária ${}^{RL}_{-\infty}D_x^{-p} g(x)$ pode ser expressa na forma de uma convolução

$$ {}^{RL}_{-\infty}D_x^{-(k-p)} g(x) = \frac{x^{(x-p)-1}}{\Gamma(k-p)} * g^{(k)}(x), \tag{61}$$

na qual o primeiro termo é não-nulo apenas quando $x > 0$.

Finalmente, a transformada de Fourier do operador de Riesz-Weyl é dada por:

$$\mathcal{F}\{{}^{RW}_{-\infty}D_x^{\gamma} g(x); k\} = \mathcal{F}\left\{\frac{\partial^{\gamma} g(x)}{\partial |x|^{\gamma}}; k\right\} = -|k|^{\gamma}\hat{g}(k). \tag{62}$$

com $1 < \gamma < 2$.

2.3.9 Série de Fourier

Existem modelos matemáticos teóricos utilizados no estudo de difusão anômala, cuja solução revela o comportamento de voos subdifusivos ou superdifusivos. Trata-se de equações diferenciais parciais fracionárias cuja solução, com valores de contorno, podem ser obtidas através do Cálculo, importante ramo da matemática. A série de Fourier é uma trigonométrica utilizada na representação de funções periódicas usando senos e cossenos. A forma geral da série é:

$$T(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \operatorname{sen}\left(\frac{n\pi t}{L}\right) \right]. \tag{63}$$

Se $f(t)$ for uma função par, então $b_n = 0$ para todo $n \geq 1$. Além disso, segue que:

$$a_n = \frac{2}{L} \int_0^L f(t) \cos\left(\frac{n\pi t}{L}\right) dt, \text{ para } n \geq 0, \tag{64}$$

sendo $2L$ o período fundamental da função $f(x)$.

2.3.10 Transformada de Fourier de função par

A transformada de Fourier de uma função $f(x)$ é uma função na variável k, denotada por $\hat{f}(k)$ ou por $\mathcal{F}\{f(x)\}$, definida como segue:

$$\hat{f}(\kappa) \equiv \mathcal{F}\{f(x)\} \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(x)\, e^{-i\kappa x} dx. \tag{65}$$

A transformada inversa de $\hat{f}(\kappa)$ é:

$$f(x) \equiv \mathcal{F}^{-1}\{F(\kappa)\} \stackrel{\text{def}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\kappa)\, e^{i\kappa x} d\kappa. \tag{66}$$

Quando as funções $f(t)$ e $\hat{f}(\omega)$ são pares, existem algumas propriedades em relação à simetria e à paridade.

Se $f(t)$ for uma função par, então $\hat{f}(\omega)$ também é. Se $f(t)$ for par, o intervalo de integração pode ser alterado para $[0,\ \infty]$ em lugar de $[-\infty,\ \infty]$, dobrando o valor calculado da integral, como segue:

$$f(x) \equiv \mathcal{F}^{-1}\{F(\kappa)\} = \frac{1}{\pi} \int_{0}^{\infty} F(\kappa)\, e^{i\kappa x} d\kappa. \tag{67}$$

Usando a forma trigonométrica da função $e^{i\kappa x}$ em (67), tem-se também:

$$f(x) \equiv \mathcal{F}^{-1}\{F(\kappa)\} = \frac{1}{\pi}\int_0^\infty F(\kappa)\cos(\kappa x)d\kappa. \tag{68}$$

2.4 Equações de difusão clássicas

Nesta seção, mostraremos como obter a equação de difusão geral a partir da lei da conservação universal (NUSSENZVEIG, 2013). Para isso, será utilizada a Lei de Fick, que estabelece que a taxa de difusão de uma substância é proporcional ao produto da área de superfície através da qual a difusão ocorre e ao gradiente de concentração da substância no meio (PARADISI et al., 2001).

Sejam $u(\vec{r},t)$ um campo conservativo em $\mathbb{R}^3$, $\vec{j}(\vec{r},t)$ o fluxo do estado conservado, $\hat{n}$ um vetor normal unitário apontando para fora na fronteira do domínio físico ω e $S(\vec{r},t)$ uma taxa que representa algum tipo de fonte da quantidade u. A lei de conservação é dada por:

$$\frac{\partial}{\partial t}\int_\omega u(\vec{r},t)dV + \int_{\partial\omega} \overrightarrow{j}(\vec{r},t)\cdot\overrightarrow{n}\,da = \int_\omega S(\vec{r},t)dV. \tag{69}$$

Da Equação (69) e pelo Teorema da Divergência, é possível obter:

$$\frac{\partial}{\partial t}\int_\omega u(\vec{r},t)dV + \int_\omega \overrightarrow{\nabla}\cdot\overrightarrow{j}(\vec{r},t)dV = \int_\omega S(\vec{r},t)dV. \tag{70}$$

Utilizando (70), obtém-se a equação diferencial parcial:

$$\frac{\partial u(\vec{r},t)}{\partial t} + \overrightarrow{\nabla}\cdot\vec{j}(\vec{r},t) = S(\vec{r},t). \tag{71}$$

No caso clássico, existem duas fontes de fluxo nesta situação. Primeiro, o fluxo difusivo surge devido à difusão. Isso é tipicamente aproximado pela primeira lei de Fick:

$$\vec{j}_{\text{difusão}} = -D\,\nabla u, \tag{72}$$

isto é, o fluxo do material de difusão (relativo ao movimento do *bulk*) em qualquer parte do sistema é proporcional ao gradiente de concentração local, onde D é o coeficiente de difusão. Segundo, quando há convecção ou fluxo geral, existe um fluxo associado chamado fluxo advectivo:

$$\vec{j}_{\text{advecção}} = \vec{v}\, u. \tag{73}$$

O fluxo total é dado pela soma de (72) e (73). Nos problemas físicos de transporte é, em geral, possível identificar dois tipos de fluxos, também chamados de correntes, a saber:

$$\vec{j} = \vec{j}_{\text{difusão}} + \vec{j}_{\text{advecção}}. \tag{74}$$

Substituindo as correntes (72) e (73) na Equação (71) e considerando que o valor S é nulo, tem-se:

$$\frac{\partial u}{\partial t} + \nabla \cdot (-D\nabla u + \vec{v}\, u) = 0,$$

ou ainda,

$$\frac{\partial u}{\partial t} = D\nabla^2 u - \nabla \vec{v}\, u. \tag{75}$$

Essa é a equação de difusão normal.

Neste trabalho, propomos uma nova interpretação da Equação (74) utilizando o Cálculo Fracionário, sendo $\vec{j}_{\text{advecção}}$ a corrente advectiva definida por:

$$\vec{j}_{\text{advecção}} = \vec{v}(\vec{r})u(\vec{r}, t). \tag{76}$$

Consideramos, também, a corrente $\vec{j}_{\text{difusão}}$ como:

$$\vec{j}_{\text{difusão}} = -D(\vec{r})\vec{\nabla}^{\epsilon} u(\vec{r}, t). \tag{77}$$

Inserindo as correntes (76) e (77) na Equação (71), obtém-se:

$$\frac{\partial u(\vec{r}, t)}{\partial t} + \vec{\nabla} \cdot \left(\vec{j}_{\text{advecção}}(\vec{r}, t) + \vec{j}_{\text{difusão}}(\vec{r}, t)\right) = S(\vec{r}, t), \tag{78}$$

ou,

$$\frac{\partial u(\vec{r}, t)}{\partial t} + \vec{\nabla} \cdot \left(\vec{v}(\vec{r})\, u(\vec{r}, t) - D(\vec{r})\vec{\nabla} u(\vec{r}, t)\right) = S(\vec{r}, t). \tag{79}$$

Aplicando o operador em (79), é possível obter a equação:

$$\frac{\partial u(\vec{r},t)}{\partial t} - (\vec{\nabla} D(\vec{r})) \cdot (\vec{\nabla} u(\vec{r},t)) - D(\vec{r}) \nabla^2 u(\vec{r},t) + (\vec{\nabla} \cdot \vec{v}(\vec{r}))\, u(\vec{r},t) + (\vec{v}(\vec{r}) \cdot \vec{\nabla})\, u(\vec{r},t) = S(\vec{r},t). \tag{80}$$

Restringindo ao caso em que $\vec{v}$ e D são constantes, tem-se:

$$\frac{\partial u(\vec{r},t)}{\partial t} - D\nabla^2 u(\vec{r},t) + \vec{v} \cdot \vec{\nabla}\, u(\vec{r},t) = S(\vec{r},t). \tag{81}$$

Assumindo não haver fontes, isto é, $S(\vec{r},t) = 0$, segue que:

$$\frac{\partial u(\vec{r},t)}{\partial t} = D\vec{\nabla}^2 u(\vec{r},t) - \vec{v} \cdot \vec{\nabla}\, u(\vec{r},t). \tag{82}$$

Em uma dimensão, tem-se a equação de difusão:

$$\frac{\partial u(x,t)}{\partial t} = D\frac{\partial^2 u(x,t)}{\partial x^2} - v\frac{u(x,t)}{\partial x}. \tag{83}$$

2.5 Obtendo a equação de difusão generalizada

Na Seção (2.4) foi mostrado como obter a equação de difusão clássica a partir da Lei da Conservação Universal. O objetivo, agora, é obter uma equação de difusão que generaliza o caso anterior. Isso pode ser feito a partir de modelos teóricos de difusão anômala. Uma abordagem utilizada é conhecida como caminhada aleatória no tempo contínuo, em inglês, *continuous time random walk* (CTRW). Trata-se de um método de simulação utilizado no estudo da difusão anômala linear (METZLER; KLAFTER, 2000a). Nele, são mimetizados processos de interação entre as partículas em movimento com o ambiente em escala micro[1], permitindo a compreensão de características que definem o movimento difusivo em escala macro. Nela, as partículas estão em movimentos erráticos, tanto o comprimento de um salto como o tempo de espera entre um salto e outro não são constantes como no movimento aleatório (VLAHOS et al., 2008).

A Figura (1) mostra uma partícula em movimento realizando saltos em diferentes posições e com tempos de espera entre tais saltos. A PDF $\Psi(x,t)$ indica qual é a probabilidade de uma partícula realizar um salto com comprimento x e ter um tempo de espera t. Assim, os saltos e os tempos de espera são variáveis aleatórias obtidas a partir dessa PDF.

[1] Usaremos a nomenclatura escala micro ou escala microscópica para nos referir à escala menor do fenômeno em que os entes são individualizados e interagem por meio de forças de natureza física e, com isso, alteram seus microestados.

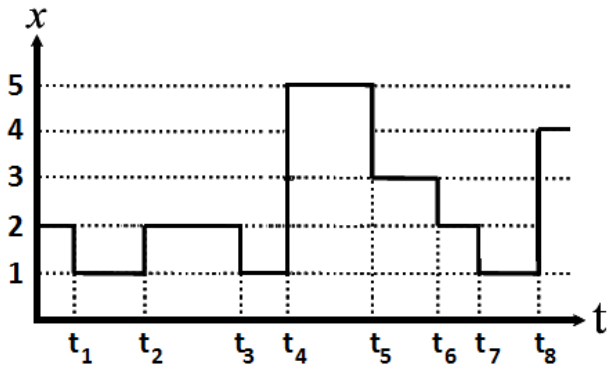


Figura 1 – Esquema para caminhada aleatória no tempo contínuo (CTRW). No eixo das abscissas são mostrados valores diferentes de instantes de tempos e, no eixo das ordenadas, as posições de uma partícula que está executando o CTRW.

Considerando que a função $\eta(x,t)$ denota a probabilidade por unidade de deslocamento e tempo de um caminhante aleatório que saiu de x no tempo t, para a posição x' no tempo t', tem-se:

$$\eta(x,t) = \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \eta(x',t')\Psi(x-x',t-t')dt'dx' + \delta(x)\delta(t), \tag{84}$$

na qual $\delta(x)\delta(t)$ é a condição inicial do *random walk* (METZLER; KLAFTER, 2000a). Portanto, $\Psi(x,t)$ se comporta como um *kernel* que promove a evolução de $\eta(x,t)$.

Por outro lado, parte-se da premissa de que existem duas PDFs que governam o deslocamento de partículas que sofrem a difusão anômala em uma dimensão: uma, $\lambda(x)$, corresponde às frequências de comprimentos de saltos realizados aleatoriamente e a outra, $\psi(t)$, corresponde às frequências dos valores de tempo de espera entre os saltos (METZLER; KLAFTER, 2000a). Se $\lambda(x)$ e $\psi(t)$ representam, respectivamente, a distribuição de probabilidades marginais para as variáveis aleatórias referentes ao comprimento dos saltos e ao tempo de espera entre os saltos, tem-se então as funções:

$$\lambda(x) = \int_{-\infty}^{+\infty} \Psi(x,t)dt, \tag{85}$$

$$\psi(t) = \int_{0}^{+\infty} \Psi(x,t)dx. \tag{86}$$

Sejam δx uma variável aleatória para o comprimento de saltos e δt uma variável aleatória correspondente ao tempo de espera, ambas obtidas de (85) e (86), respectivamente. É possível haver uma dependência entre elas ou não. Neste trabalho, consideraremos a

última possibilidade, permitindo que a PDF $\Psi(x,t)$ seja escrita de uma forma desacoplada como segue:

$$\Psi(x,t) = \lambda(x)\psi(t). \tag{87}$$

Seja $\phi(t)$ a probabilidade da partícula permanecer na mesma posição. Como a integral $\int_0^t \psi(t')dt'$ representa a probabilidade de ela realizar um passo em um intervalo entre 0 e t e não há outra probabilidade para um possível desenvolvimento temporal, segue que a probabilidade da partícula permanecer em uma mesma posição por um tempo t é dada por:

$$\phi(t) = 1 - \int_0^t \psi(t')dt'. \tag{88}$$

A PDF $u(x,t)$ da partícula se encontrar em x no tempo t é obtida através das probabilidades das PDF de chegada, $\eta(x,t)$, e de permanência, $\phi(t)$, multiplicando-as e integrando no tempo total, como segue:

$$u(x,t) = \int_0^t \eta(x, t-t'')\phi(t'')dt''. \tag{89}$$

Substituindo (84) em (89), tem-se:

$$u(x,t) = \int_0^t \left(\int_{-\infty}^{+\infty} \int_0^{+\infty} \eta(x',t')\Psi(x-x', t-t'-t'')dt'dx' + \delta(x)\delta(t-t'') \right) \phi(t'')dt''. \tag{90}$$

Invertendo a ordem de integração em (90) e fazendo $T = t' + t''$, conclui-se que:

$$u(x,t) = \int_{-\infty}^{+\infty} \int_0^{+\infty} \left(\left[\int_0^t \eta(x', T-t'')\Phi(t'')dt'' \right] \Psi(x-x', t-T) + \phi(t)\delta(x) \right) dTdx', \tag{91}$$

ou seja,

$$u(x,t) = \int_{-\infty}^{+\infty} \int_0^{+\infty} u(x',T)\Psi(x-x', t-T) + \phi(t)\delta(x)dTdx'. \tag{92}$$

Aplicando então a transformada de Laplace na variável temporal t, segue que:

$$\tilde{u}(x,s) = \int_0^{+\infty} u(x,t)e^{-st}dt \tag{93}$$

$$\tilde{u}(x,s) = \int_0^{+\infty} \left(\int_{-\infty}^{+\infty} dx' \int_0^{+\infty} u(x',T)\Psi(x-x',t-T)dT + \phi(t)\delta(x) \right) e^{-st}dt. \tag{94}$$

Assim, pelo Teorema da Convolução para transformada de Laplace (45), obtém-se:

$$\tilde{u}(x,s) = \int_{-\infty}^{+\infty} u(x',s)\Psi(x-x',s)dx' + \Phi(s)\delta(x). \tag{95}$$

Aplicando então a transformada de Fourier na variável espacial, tem-se:

$$\hat{\tilde{u}}(k,s) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{+\infty} u(x',s)\Psi(x-x',s)dx' + \Phi(s)\delta(x) \right) e^{ikx}dx. \tag{96}$$

Aplicando o Teorema da Convolução para transformada de Fourier (44), obtém-se:

$$\hat{\tilde{u}}(k,s) = \hat{\tilde{u}}(k,s)\Psi(k,s) + \tilde{\Phi}(s)\hat{u}(k,0), \tag{97}$$

onde $\mathcal{F}\{\delta(x);k\} = \hat{u}(k,0)$.

Agora, aplicando a transformada de Laplace em (88), segue que

$$\tilde{\Phi}(s) = \frac{1-\tilde{\Psi}(s)}{s}. \tag{98}$$

Utilizando (98) em (97), tem-se a relação algébrica:

$$\hat{\tilde{u}}(k,s) = \frac{1-\Psi(s)}{s} \frac{\hat{\tilde{u}}(k,0)}{1-\hat{\tilde{\Psi}}(k,s)}, \tag{99}$$

denominada fórmula de Montroll-Weiss (MONTROLL; WEISS, 1965).

Como já mencionado, consideramos neste trabalho o caso no qual a distribuição $\Psi(x,t)$ está desacoplada, ou seja, a distribuição para o tamanho dos passos e a distribuição do tempo de espera são independentes. Supondo que o valor médio no tempo $\langle t \rangle$ e o segundo momento da posição $\langle x^2 \rangle$ são ambos finitos, pode-se escrever a transformada no tempo e na posição de $\Psi(x,t)$ como

$$\Psi(k,s) = \Psi(s)\lambda(k), \tag{100}$$

onde:

$$\Psi(s) = \int_0^{+\infty} \Psi(t)e^{-st}dt \tag{101}$$

$$\lambda(k) = \int_{-\infty}^{+\infty} \lambda(x)e^{-ikx}dx \tag{102}$$

É possível reescrever as distribuições (101) e (102) usando a representação em série da função exponencial da seguinte forma:

$$\begin{aligned}
\Psi(s) &= \int_0^{+\infty} w(t)\left(\sum_{n=0}^{+\infty}\frac{(-st)^n}{n!}\right)dt \\
&= \sum_{n=0}^{+\infty}(-1)^n\frac{s^n}{n!}\int_0^{+\infty} w(t)t^n dt \\
&= \sum_{n=0}^{+\infty}(-1)^n\frac{s^n}{n!}\langle t^n\rangle
\end{aligned} \tag{103}$$

$$\begin{aligned}
\lambda(k) &= \int_{-\infty}^{+\infty} \lambda(x)\left(\sum_{n=0}^{+\infty}\frac{(-ikx)^n}{n!}\right)dx \\
&= \sum_{n=0}^{+\infty}\frac{(-ik)^n}{n!}\int_{-\infty}^{+\infty} \lambda(x)x^n dx \\
&= \sum_{n=0}^{+\infty}(-i)^n\frac{k^n}{n!}\langle x^n\rangle
\end{aligned} \tag{104}$$

A partir de (103) e (104), é possível obter uma aproximação para as distribuições $\Psi(s)$ e $\lambda(k)$ considerando apenas os termos de menor potência nas séries quando k e s tendem a 0, como segue:

$$\Psi(s) \sim 1 - s\langle t\rangle + O(s^2), \tag{105}$$

$$\lambda(k) \sim 1 - \frac{k^2}{2}\langle x^2\rangle. \tag{106}$$

Substituindo (105) e (106) na relação de Montroll-Weiss (99):

$$\hat{\tilde{u}}(k,s) = \frac{1-\tilde{\Psi}(s)}{s}\frac{\hat{u}(k,0)}{1-(1-ikx_0-\frac{\sigma^\gamma}{2}|k|^\gamma)(1-(\tau s)^\alpha)}$$
$$\hat{\tilde{u}}(k,s) = \frac{1-(1-(\tau s)^\alpha)}{s}\frac{\hat{u}(k,0)}{1-[1-(\tau s)^\alpha-ikx_0-\frac{\sigma^\gamma}{2}|k|^\gamma]}$$
$$\hat{\tilde{u}}(k,s) = \frac{(\tau s)^\alpha}{s}\frac{\hat{u}(k,0)}{(\tau s)^\alpha+ikx_0+\frac{\sigma^\gamma}{2}|k|^\gamma}$$
$$\hat{\tilde{u}}(k,s) = s^{\alpha-1}\frac{\hat{u}(k,0)}{s^\alpha+\frac{ikx_0}{\tau^\alpha}+\frac{\sigma^\gamma}{2\tau^\alpha}|k|^\gamma}$$

Considerando então que seja válida a relação

$$\frac{x_0}{\tau^\alpha} = \frac{x_0}{\tau\tau^{\alpha-1}} = \frac{v}{\tau^{\alpha-1}} = v\tau^{1-\alpha} = \nu_\alpha,$$

onde $v = \dfrac{x_0}{\tau}$, é possível concluir que:

$$\hat{\tilde{u}}(k,s)\left(s^\alpha + ik\nu_\alpha + \frac{\sigma^\gamma}{2\tau^\alpha}|k|^\gamma\right) = s^{\alpha-1}\hat{u}(k,0)$$

$$s^\alpha\hat{\tilde{u}}(k,s) + ik\nu_\alpha\hat{\tilde{u}}(k,s) - s^{\alpha-1}\hat{u}(k,0) + \frac{\sigma^\gamma}{2\tau^\alpha}|k|^\gamma\hat{\tilde{u}}(k,s) = 0$$

$$\mathcal{L}^{-1}\{ik\nu_\alpha\hat{\tilde{u}}(k,s)\} + \mathcal{L}^{-1}\{s^\alpha\hat{\tilde{u}}(k,s) - s^{\alpha-1}\hat{u}(k,0)\} + \mathcal{L}^{-1}\left\{\frac{\sigma^\gamma}{2\tau^\alpha}|k|^\gamma\hat{\tilde{u}}(k,s)\right\} = 0$$

$$ik\nu_\alpha\hat{u}(k,t) + \frac{\partial^\alpha\hat{u}}{\partial t^\alpha}(k,t) + \frac{\sigma^\gamma}{2\tau^\alpha}|k|^\gamma\hat{u}(k,t) = 0$$

$$\mathcal{F}^{-1}\left\{ik\nu_\alpha\hat{u}(k,t) + \frac{\partial^\alpha\hat{u}}{\partial t^\alpha}(k,t) + \frac{\sigma^\gamma}{2\tau^\alpha}|k|^\gamma\hat{u}(k,t)\right\} = 0$$

$$\nu_\alpha\frac{\partial u}{\partial x}(x,t) + \frac{\partial^\alpha u}{\partial t^\alpha}(x,t) + \frac{\sigma^\gamma}{2\tau^\alpha}\frac{\partial^\gamma u}{\partial t^\gamma}(x,t) = 0$$

Portanto, tem-se a equação de difusão generalizada:

$$\frac{\partial^\alpha u}{\partial t^\alpha}(x,t) - D\frac{\partial^\gamma u}{\partial t^\gamma}(x,t) + \nu_\alpha\frac{\partial u}{\partial x}(x,t) = 0, \tag{107}$$

onde $D = -\dfrac{\sigma^{\gamma}}{2\tau^2}$.

2.6 Voos difusivos anômalos

Os voos difusivos anômalos são caracterizados por uma distribuição de probabilidade não gaussiana e caudas pesadas, que diferem dos voos normais (ou difusão normal) encontrados em sistemas onde as partículas se movem aleatoriamente e de forma uniforme. A trajetória de um voo difusivo anômalo consiste em um caminho errático que pode apresentar períodos prolongados de tempo com pequenas variações, seguidos de mudanças abruptas de direção. Essas mudanças repentinas de direção resultam em uma trajetória altamente ramificada e fractal.

Os voos difusivos anômalos são encontrados em vários sistemas complexos, como, por exemplo, na migração de animais, na dispersão de partículas em meio poroso, na difusão de moléculas em células biológicas e na dinâmica de preços de ativos financeiros. A caracterização e modelagem de voos difusivos anômalos são importantes para entender o comportamento desses sistemas complexos e prever seus resultados.

2.7 Simulação de voos sob ação de difusão anômala e o método de Monte Carlo pela função inversa

Nas seções anteriores, vimos como obter um modelo teórico para difusão e difusão anômala. Nesta seção, analisamos as ferramentas teóricas necessárias para simulação de voos que estão sob ação de difusão anômala.

Para simular o movimento de partículas com dinâmica em escala micro para fenômenos com difusão anômala, é possível utilizar diversas abordagens. Dentre elas, destaca-se o CTRW. Nele, cada partícula pode apresentar voos nos quais ocorrem tempos de espera e saltos variados (COMPTE, 1997). Com intuito de realizar sorteios para simular tais movimentos, faz-se o uso do Método de Monte Carlo pela Função Inversa. Trata-se de uma técnica estatística utilizada para gerar números aleatórios que seguem uma distribuição desejada. Ele funciona gerando números aleatórios uniformes e transformando-os através da inversa da função de distribuição cumulativa (FDC) da distribuição desejada (ROBERT; CASELLA; CASELLA, 1999).

Nesta seção, o Método de Monte Carlo pela Função Inversa é aplicado sobre as funções que serão utilizadas como leis de potências nas simulações computacionais. Tratam-se das distribuições para os comprimentos dos saltos e para os tempos de espera das partículas.

Inicialmente, serão determinados os fatores de normalização, κ, para, a seguir, encontrar as funções inversas para área, utilizadas para obter números aleatórios entre 0 e 1.

Sejam κ e ϵ duas constantes reais positivas e considere a função:

$$p(x - x_0) = \frac{\kappa}{(|x - x_0| + \epsilon)^r}, \qquad r > 1. \tag{108}$$

Note que, pela forma como está definida, essa função é uma distribuição de probabilidade que não possui pontos de singularidade, motivo de sua escolha.

É possível encontrar κ para que a função seja uma PDF:

$$\int_{-\infty}^{+\infty} p(x - x_0)dx = \int_{-\infty}^{x_0} p(x - x_0)dx + \int_{x_0}^{+\infty} p(x - x_0)dx = 1.$$

$$\int_{-\infty}^{x_0} \frac{\kappa}{(x - (x_0 + \epsilon))^r}dx + \int_{x_0}^{+\infty} \frac{\kappa}{(x - (x_0 - \epsilon))^r}dx = 1.$$

$$\kappa \lim_{a \to -\infty} \int_a^{x_0} \frac{1}{(x_0 + \epsilon - x)^r}dx + \kappa \lim_{b \to +\infty} \int_{x_0}^{b} \frac{1}{(x - (x_0 - \epsilon))^r}dx = 1.$$

$$\kappa \lim_{a \to -\infty} \left(\frac{\epsilon^{1-r}}{r-1} - \frac{(x_0 + \epsilon - a)^{1-r}}{r-1} \right)_a^{x_0} + \kappa \lim_{b \to +\infty} \left(\frac{(b - (x_0 - \epsilon))^{1-r}}{1-r} - \frac{\epsilon^{1-r}}{1-r} \right)_{x_0}^{b} = 1.$$

$$\kappa \frac{\epsilon^{1-r}}{r-1} + \kappa \frac{\epsilon^{1-r}}{r-1} = 1.$$

$$\kappa = \frac{r-1}{2\epsilon^{1-r}}, \qquad r > 1,\ \epsilon > 0. \tag{109}$$

Portanto, com a expressão calculada em (109), tem-se:

$$p(x - x_0) = \begin{cases} \dfrac{r-1}{2\epsilon^{1-r}} \dfrac{1}{(x - (x_0 - \epsilon))^r}, & \text{se } x > x_0, \\ \dfrac{r-1}{2\epsilon^{1-r}} \dfrac{1}{|x - (x_0 + \epsilon)|^r}, & \text{se } x \leq x_0. \end{cases} \tag{110}$$

$$p(x - x_0) = \frac{r-1}{2\epsilon^{1-r}} \frac{1}{(|x - x_0| + \epsilon)^r}. \tag{111}$$

A função $A(x)$ a seguir fornece a área entre a curva de $p(x - x_0)$ e o eixo x, de $-\infty$ até x.

$$A(x) = \begin{cases} \displaystyle\int_{-\infty}^{x} p(t - x_0)dt, \text{ se } x \leq x_0, \\ \displaystyle\frac{1}{2} + \int_{x_0}^{x} p(t - x_0)dt, \text{ se } x > x_0. \end{cases} \tag{112}$$

Portanto,

$$A(x) = \begin{cases} \dfrac{(x_0 + \epsilon - x)^{1-r}}{2\epsilon^{1-r}}, \text{ se } x \leq x_0, \\ \dfrac{1}{2} + \dfrac{\epsilon^{1-r} - (x - (x_0 - \epsilon))^{1-r}}{2\epsilon^{1-r}}, \text{ se } x > x_0. \end{cases} \tag{113}$$

Para utilizar o Método de Monte Carlo pela função inversa, deve-se encontrar a função $x(A)$, inversa de $A(x)$. Ela deve ser construída em duas sentenças devido à lei de formação da função $A(x)$.

Para $0 \leq A \leq \dfrac{1}{2}$, tem-se:

$$A = \frac{(x_0 + \epsilon - x)^{1-r}}{2\epsilon^{1-r}}.$$
$$2A\epsilon^{1-r} = (x_0 + \epsilon - x)^{1-r}.$$
$$\sqrt[1-r]{2A\epsilon^{1-r}} = x_0 + \epsilon - x.$$
$$\epsilon \sqrt[1-r]{2A} - x_0 - \epsilon = -x.$$
$$x = x_0 + \epsilon - \epsilon \sqrt[1-r]{2A}.$$

Para $\dfrac{1}{2} < A \leq 1$, segue que:

$$A = \frac{1}{2} + \frac{\epsilon^{1-r} - (x - (x_0 - \epsilon))^{1-r}}{2\epsilon^{1-r}}.$$
$$2\epsilon^{1-r}\left(A - \frac{1}{2}\right) = \epsilon^{1-r} - (x - (x_0 - \epsilon))^{1-r}.$$
$$2\epsilon^{1-r}\left(A - \frac{1}{2}\right) - \epsilon^{1-r} = -(x - (x_0 - \epsilon))^{1-r}.$$
$$(x - (x_0 - \epsilon))^{1-r} = \epsilon^{1-r} - \epsilon^{1-r}(2A - 1).$$
$$(x - (x_0 - \epsilon))^{1-r} = \epsilon^{1-r}[1 - (2A - 1)].$$
$$x - (x_0 - \epsilon) = \epsilon \sqrt[1-r]{2 - 2A}.$$
$$x = x_0 - \epsilon + \epsilon \sqrt[1-r]{2 - 2A}.$$

Portanto,

$$x(A) = \begin{cases} x_0 + \epsilon - \epsilon \sqrt[1-r]{2A}, \text{ se } 0 \leq A \leq \dfrac{1}{2}, \\ x_0 - \epsilon + \epsilon \sqrt[1-r]{2 - 2A}, \text{ se } \dfrac{1}{2} < A \leq 1. \end{cases} \tag{114}$$

Utilizando uma função semelhante, vamos considerar agora os tempos de espera com uma nova distribuição.

Consideremos agora a função distribuição de probabilidade para os tempos de espera $p(t)$:

$$p(t - t_0) = \frac{\kappa}{(|t - t_0| + \epsilon)^s}, \tag{115}$$

ou seja,

$$p(t) = \frac{\kappa}{(t + \epsilon)^s}. \tag{116}$$

Para obter κ de forma que a função seja uma PDF, deve-se ter a integral de $p(t)$ sobre seu domínio igual a 1, ou seja, é preciso que:

$$\int_0^{+\infty} p(t)dt = 1.$$

Substituindo (116) na integral, segue que:

$$\int_0^{+\infty} \frac{\kappa}{(t + \epsilon)^s} dt = 1.$$

Efetuando os cálculos de forma semelhante ao que foi feito anteriormente sobre $p(x)$, obtém-se

$$p(t) = \frac{s - 1}{\epsilon^{1-s}} \frac{1}{(t + \epsilon)^s}. \tag{117}$$

A função $A(t)$ a seguir fornece a área entre o eixo x e a curva do gráfico de $p(t)$ de 0 até t.

$$A(t) = \int_0^t p(w)dw.$$

$$A(t) = \int_0^t \frac{s-1}{\epsilon^{1-s}} \frac{1}{(w+\epsilon)^s} dw.$$

$$A(t) = \frac{\epsilon^{-s+1} - (t+\epsilon)^{-s+1}}{\epsilon^{1-s}}. \tag{118}$$

Para aplicar o Método de Monte Carlo pela função inversa, é necessário obter a inversa $t(A)$ da função em (118). Isso pode ser feito como segue:

$$A\epsilon^{1-s} = \epsilon^{-s+1} - (t+\epsilon)^{-s+1}.$$

Efetuando os cálculos, obtém-se uma expressão para t, o que possibilita escrever a lei de formação da função:

$$t = \sqrt[-s+1]{\epsilon^{-s+1} - A\epsilon^{1-s}} - \epsilon, \qquad 0 \leq A \leq 1 \tag{119}$$

2.8 Ajuste pelo método de análise por dispersão

Há métodos para estabelecer uma relação entre os modelos teóricos e computacionais, produzindo voos bem controlados com intuito de obter os parâmetros em grande escala μ e D_μ, que significam o expoente e o coeficiente de difusão, respectivamente. Dentre eles podemos citar a calibração via método de análise da medida da dispersão e via método de otimização dos modelos teóricos (PEREIRA, 2018).

Em situações nas quais ocorre o fenômeno físico de difusão, o deslocamento quadrático médio $\langle X^2(t) \rangle$ das partículas possui uma dependência com o tempo t, em geral, na forma de uma lei de potências t^μ, ou seja, podemos afirmar que

$$\langle X^2(t) \rangle = 2Dt^\mu, \tag{120}$$

onde D é o coeficiente de difusão e μ é o expoente (KLAFTER; SHLESINGER; ZUMOFEN, 1996).

Foram geradas por simulação N séries para obter os parâmetros citados do movimento em escala macro por meio de regressão, sendo possível analisar se os voos gerados são

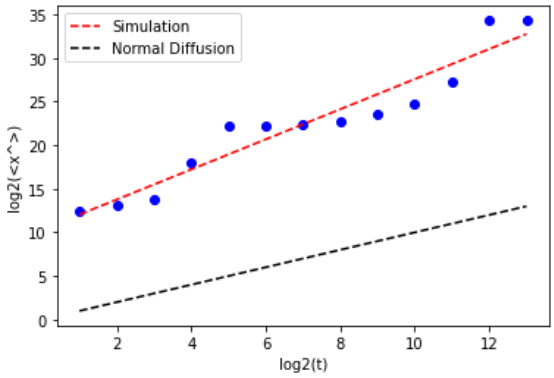


Figura 2 – Gráfico em escala $\log \times log$ com a variância da posição das partículas em função de diferentes tempos de corte considerados nas simulações de um voo superdifusivo. Devido às diferenças de variações das variãncias é observado o fenômeno conhecido como escadas do diabo.

subdifusivos, caso μ seja menor que 1, superdifusivo, quando $\mu > 1$, ou se ocorre a difusão normal ($\mu = 1$) (PEREIRA et al., 2017).

Além disso, aplicando o logaritmo nos dois lados da Equação (120), segue que:

$$\log\langle X^2(t)\rangle = \mu \log\, t + \log\, 2D, \tag{121}$$

cujo gráfico é uma reta no plano $\log t \;\times\; \log\langle X^2(t)\rangle$. Na Equação (121) fica, portanto, estabelecida uma relação entre as escalas de tempo e da dispersão da variância com os parâmetros em escala macro, ou seja, o expoente de difusão μ e o coeficiente de difusão D. De posse de dados de experimento ou simulações, essa relação permite determinar de forma simples, via regressão linear, os parâmetros μ e D do processo difusivo. Contudo, deve-se notar que não há, *a priori*, uma distinção sobre a natureza do processo difusivo anômalo estudado, se linear ou não-linear. Além disso, limitações computacionais podem dificultar a obtenção dos parâmetros através deste método, fazendo com que apareçam no gráfico em escala log log um fenômeno conhecido como escada do diabo (BAK; BRUINSMA, 1982), como observado na Figura (2).

2.9 Limitações e desafios

A modelagem de voos difusivos anômalos apresenta várias limitações e desafios na prática. Um dos principais desafios é a falta de dados experimentais para diversos fenômenos que poderiam ser usados para validar os modelos. Os voos difusivos anômalos ocorrem em sistemas complexos e muitas vezes é difícil medir a trajetória precisa das partículas, tornando a validação dificultada (BOUCHAUD; GEORGES, 1990).

Além disso, a complexidade dos sistemas em que os voos difusivos anômalos ocorrem

apresenta outro desafio. Modelar esses sistemas de forma precisa e realista pode ser difícil, e muitas vezes é necessário fazer simplificações que podem levar a erros na modelagem (SOKOLOV, 2012).

Outra limitação é a dependência dos modelos em relação aos parâmetros escolhidos. Os modelos de voos difusivos anômalos são caracterizados por vários parâmetros, e a escolha desses parâmetros pode afetar significativamente os resultados.

Outro fato importante é que os modelos de voos difusivos anômalos podem ser computacionalmente intensivos, especialmente quando se utiliza modelagem baseada em cálculo fracionário. Isso pode tornar a modelagem de sistemas complexos com grande número de partículas impraticável em termos computacionais. Tal questão será discutida no Capítulo 4.

Em relação aos ajustes dos dados simulados com um modelo teórico, diversas questões devem ser consideradas. A escolha de um algoritmo adequado para o ajuste e consequente obtenção dos parâmetros torna-se um importante desafio. Na literatura, algumas comparações podem ser realizadas com ajustes disponíveis (PEREIRA et al., 2018). Neste trabalho apresentamos algumas alternativas e analisar suas vantagens e limitações.

Para que o ajuste seja realizado de fato, há outros fatores que devem ser considerados e analisados neste estudo, como a montagem do histograma para o conjunto de dados simulados, por exemplo, que serão tratados no Capítulo 4.

Capítulo 3

Metodologia

O presente capítulo descreverá a metodologia utilizada neste trabalho. Como proposta inicial, buscou-se desenvolver e validar um aparato teórico para identificação de características de um processo difusivo anômalo a partir de dados experimentais ou simulações. A proposta geral da pesquisa é desenvolver uma abordagem teórica a partir de equações diferenciais parciais fracionárias (EDPF). Pretende-se analisar e caracterizar o comportamento linear e não-linear de modelos feitos por simulações e dados experimentais.

A primeira etapa consiste em partir de um modelo teórico de difusão anômala na forma de uma EDPF para descrever o fenômeno e determinar seus parâmetros. Isso pode ser feito simulando controladamente voos de partículas sem correlações de longo alcance e ajustando os dados a uma distribuição de Lévy. Uma vez validado com dados sintetizados, a ideia é aplicar esse aparato a dados reais ou a resultados de modelos simulados de processos como as buscas de sítios no DNA por fatores de transcrição.

3.1 Metodologia para a geração de voos anômalos

Para geração de voos superdifusivos ou subdifisivos controlados foi utilizado o CTRW em um conjunto de partículas utilizando o método de Monte Carlo pela função inversa, discutido no Capítulo 2. Para isso, foram consideradas leis de potências que fornecem os passos e tempos de esperas obtidos como discutido na seção 2.7. A fundamentação teórica para construção do modelo computacional é apresentada no Capítulo 2. Foi escolhida a linguagem de programação *Python* com intuito de simular a dinâmica de processos em difusão anômala. A geração de números aleatórios neste contexto é relevante e, por isso, o assunto é tratado a seguir, nesta seção. As simulações citadas permitem a obtenção da distribuição de frequência das partículas e o controle do tipo de voo é feito pela escolha dos expoentes r e s presentes na Tabela 1, que determinam as leis de potências apresentadas no Capítulo 2 utilizadas nas simulações.

Tabela 1 – Relação entre parâmetros e tipos de voos.

s	r	Tipo de voo
$1 < s < 2$	$r > 3$	subdifusivo
$s > 2$	$r > 3$	difusão normal
$s > 2$	$2 < r < 3$	superdifusivo

Nas simulações, foi utilizada a função *random()*. A sua eficácia se deve ao fato do Python, versão 2.9, utilizar o gerador Mersenne Twister. Ele é capaz de produzir pontos flutuantes com precisão de 53 bits e possui um período de $2^{19937} - 1$. A implementação subjacente em C é rápida e segura para *threads*. O Mersenne Twister é um dos geradores de números aleatórios mais amplamente testados existentes. No entanto, sendo determinístico, é importante destacar que não é adequado para todos os fins. Em criptografia, por exemplo, torna-se inadequado.

3.2 Condicionamento da função de ajuste

A função de Lévy, apresentada no Capítulo 2, é uma distribuição estatística que descreve processos difusivos anômalos. É de interesse neste estudo determinar os parâmetros que ajustam a curva aos dados obtidos nas simulações, conforme Seção 3.1. Para isso, existe a necessidade de implementação computacional desta função que, como discutido na Seção 2.2.1, gera novos parâmetros que são inteiros. Por isso, é realizada uma discretização. Isso significa que a integral imprópria é reescrita na forma de somatório finito, no qual o diferencial torna-se um subintervalo da reta. Este fato contrasta com o comportamento assintótico da função que decai por lei de potências que, frequentemente, descreve fenômenos livres de escala. Como consequência, a informação sobre as distribuições estão espalhadas por grandes intervalos na reta. A discretização, porém, leva à necessidade de utilizar um AG no ajuste, pois o cálculo do gradiente fica possibilitado por haver parâmetros inteiros. Por isso, trabalhamos com a função discretizada e os novos parâmetros também são ajustados no nosso trabalho. Tal ajuste será discutido na Seção 3.3.

3.3 Metodologia para ajuste

Após obter os dados dos voos superdifusos, foram utilizados métodos de minimização para que fossem realizados ajustes para obtenção de parâmetros da função $l(x,t)$ (132). Dentre eles o método de e Broyden–Fletcher–Goldfarb–Shanno (BFGS), tipo de algoritmo para otimização multidimensional (LIU; NOCEDAL, 1989) e um AG, tipo de algoritmo evolucionário que é probabilístico e baseado no princípio do darwinismo, no qual se observa a evolução de indivíduos, havendo a sobrevivência dos mais aptos e cruzamento dos mais aptos, forma em que a busca é realizada (PACHECO et al., 1999).

No AG utilizado, inicialmente, é realizada uma captura dos elementos da população com melhor desempenho. Em seguida, é feita uma seleção dos indivíduos que sofrem um cruzamento e uma mutação entre eles. Há então uma avaliação parcial, dos resultados, um processo de elitismo para finalmente ocorrer uma última avaliação. Em cada iteração, os dados são convertidos do tipo *float* para *string*, representados por cadeias de caracteres de zeros e uns.

Para assegurar a acurácia do método, foram considerados parâmetros fixos obtidos em um ajuste, como modelos teóricos. A eles foram adicionados ruídos a uma taxa de 5% de erro para se ter dados sintetizados. Em seguida, foi aplicado o método de AG, considerado nesta parte do trabalho para recuperação dos parâmetros presentes no modelo teórico. Aplicando diversos testes, percebeu-se que fixando o parâmetro A e fazendo a busca para os demais parâmetros atinge-se uma precisão adequada. Os resultados dos ajustes serão apresentados no Capítulo 4.

3.4 Metodologia para estudo de busca de FT em cadeias de DNA

Uma situação na qual pode-se observar o fenômeno de difusão anômala é na busca de fatores de transcrição (FT), proteínas especializadas, em um alvo nas cadeias de DNA. Para isso, foram desenvolvidos sistemas computacionais que simulam o movimento das proteínas dentro das células, supondo que os deslocamentos são realizados por meio de difusão facilitada, processo no qual os FTs fazem um deslocamento em 3D, mas ao se encontrarem com o filamento de DNA, a busca é feita em 1D, podendo se soltar dele ou continuar até que o alvo seja atingido. Foram simuladas, por meio de algoritmos implementados em Python, as buscas de diferentes FTs e observada a eficiência da busca e a posição das partículas após um certo tempo. Nosso objetivo foi mostrar que este tipo de busca pode ser caracterizado como exemplo de fenômeno superdifusivo.

Capítulo 4

Resultados e Análises

4.1 Modelo teórico

Se em um determinado processo ocorre a difusão anômala linear, então este pode ser modelado através de uma equação de difusão com derivadas fracionárias. Nesse modelo, o tempo de dispersão e a forma da distribuição são caracterizados pelos seus parâmetros. Em geral, o processo difusivo modelado neste capitulo será anômalo.

Para proporcionar uma abordagem conceitual adequada, é possível tratar a implementação computacional da distribuição para o problema de otimização a partir de um PVC com EDPF com derivadas fracionárias na derivada espacial, com condições de contorno reflexivas, ou de isolamento. Considere o seguinte problema de valor de contorno (PVC):

$$\begin{cases} \dfrac{\partial u}{\partial t} = D\dfrac{\partial^{\gamma} u}{\partial x^{\gamma}} \\ u_x\left(-\Lambda, t\right) = 0 \\ u_x\left(\Lambda, t\right) = 0 \\ u\left(x, 0\right) = N\delta(x) \end{cases} \tag{122}$$

A equação do problema em (122) é uma EDPF com derivada espacial fracionária e pode ser utilizada como modelo para o caso da superdifusão. A ordem fracionária da derivada γ é tal que se $\gamma = 2$, recupera-se a difusão clássica e se $1 < \gamma < 2$, ocorre uma difusão anômala. A condição inicial $u(x, 0) = N\delta(x)$ estabelece que no instante de tempo $t = 0$ as N partículas estão em $x = 0$. As condições de contorno $u_x\left(\pm\Lambda, t\right) = 0$, devido às derivadas parciais, especificam haver isolamento reflexivo nas fronteiras $x = \pm\Lambda$ e que, portanto, não há fluxo de partículas na escala 2Λ. Essa condição é necessária para conservação da probabilidade, mas não é suficiente para interpretação probabilística porque deve-se garantir também que a função seja positivo-definida e integrável.

Um ponto-chave para ser considerado é que a implementação computacional é finita, gerando escala para o fenômeno que, em princípio, pode ser livre de escala. Por outro lado, a fixação arbitrária de parâmetros tão grandes como desejado gera elevados custos computacionais. A questão se resume a quais são os valores suficientemente grandes para os parâmetros. Além disso, deve-se saber quais são os parâmetros relevantes no problema. A resposta será dada através do tratamento adequado do PVC (122) por meio do uso da análise de Fourier, explorando a relação entre Série de Fourier e Transformadas de Fourier.

No caso $\Lambda \to \infty$, usando a identidade de Riez-Weyl para transformada de Fourier da derivada fracionária $\mathcal{F}\left\{u(x,t)\right\} = |k|^{\gamma}\,\mathcal{F}\left\{u(x,t)\right\}$ na EDPF, mostra-se que o problema (122) possui a distribuição de Lévy como solução (SAMKO et al., 1993).

$$L_{\gamma}(x,t) = \frac{N}{2\pi}\int_{-\infty}^{+\infty} e^{-D|k|^{\gamma}t}\cos kx\; dk, \qquad x \in (-\infty,\infty), \tag{123}$$

Trata-se de uma função especial com decaimento em lei de potência (DOSS; MAURO, 2021), sendo possível mostrar que $L_{\gamma}(x,t) \sim 1/x^{\gamma+1}$ quando $x \to \infty$ (LUTZ, 2003). Tal função possui como característica uma cauda pesada, em geral, com decaimento por leis de potência, que dependendo do expoente pode ser uma distribuição livre de escala.

Para implementação computacional, a função (123) deve ser discretizada. Considerando que $\cos kx$ é função par e inserindo na integral $\sum_{n=0}^{\infty}\delta(k-k_n)\Delta k$, com $k_n = \frac{n\pi}{\Lambda}$, $n = 0,1,2,\cdots$ portanto, $\Delta k = k_{n+1} - k_n = \frac{\pi}{\Lambda}$. Isso é feito através da mudança $\kappa = \dfrac{n\pi}{\Lambda}$, $n = 0,1,2,\cdots$. Além disso, na discretização o diferencial $d\kappa$ se torna $\Delta\kappa$ e $\Delta\kappa = \dfrac{(n+1)\pi}{\Lambda} - \dfrac{n\pi}{\Lambda} = \dfrac{\pi}{\Lambda}$. Assim, tem-se a função discretizada

$$l_{\gamma}(x,t) = \frac{N}{\pi}\int_{0}^{\infty}\sum_{n=0}^{\infty}\delta(k-k_n)\Delta k\; e^{-D|k|^{\gamma}t}\cos(kx)dk = \frac{N}{\pi}\sum_{n=0}^{+\infty} e^{-D|\frac{n\pi}{\Lambda}|^{\gamma}t}\frac{\pi}{\Lambda}\cos\frac{n\pi x}{\Lambda}\,.$$

Então, segue que

$$l_{\gamma}(x,t) = \frac{N}{\Lambda}\sum_{n=0}^{+\infty} e^{-D|\frac{n\pi}{\Lambda}|^{\gamma}t}\cos\frac{n\pi x}{\Lambda}\,. \tag{124}$$

Por outro lado, usando a forma complexa da série de Fourier para gerar o *ansatz*, a série $\displaystyle\sum_{n=-\infty}^{+\infty} c_n(t)e^{i\frac{n\pi}{\Lambda}x}$ com Λ finito e os coeficientes $c_n(t)$ dependentes do tempo, na equação diferencial fracionária do problema (122), segue que

$$\sum_{n=-\infty}^{+\infty} c_n(t)\frac{\partial^\alpha}{\partial x^\alpha}e^{-i\left(\frac{n\pi}{\Lambda}\right)x} = \frac{1}{D}\sum_{n=-\infty}^{+\infty}\frac{\partial}{\partial t}c_n(t)e^{-i\left(\frac{n\pi}{\Lambda}\right)x} \tag{125}$$

Se considerarmos $\frac{\partial^\gamma}{\partial x^\gamma}e^{-i\left(\frac{n\pi}{\Lambda}\right)x} = -\left|\frac{n\pi}{\Lambda}\right|^\gamma c_n(t)$, teremos

$$\frac{\partial}{\partial t}c_n(t) = -D\left|\frac{n\pi}{\Lambda}\right|^\gamma c_n(t)\,, \tag{126}$$

cuja solução é $c_n(t) = E_n e^{-D\left|\frac{n\pi}{\Lambda}\right|^\gamma t}$, assim

$$u(x,t) = \sum_{n=-\infty}^{+\infty} E_n e^{-D\left|\frac{n\pi}{\Lambda}\right|^\gamma t} e^{i\frac{n\pi}{\Lambda}x} \tag{127}$$

Usando a condição inicial, segue que

$$u(x,0) = \sum_{n=-\infty}^{+\infty} E_n e^{i\left|\frac{n\pi}{\Lambda}\right|^\gamma x} = N\delta(x)\,, \tag{128}$$

que é, formalmente, a Série de Fourier da função δ. Aplicando a fórmula dos coeficientes, obtém-se

$$E_n = \frac{N}{2\Lambda}\int_{-\Lambda}^{\Lambda}\delta(x)e^{i\left|\frac{n\pi}{\Lambda}\right|^\gamma x}dx = \frac{N}{2\Lambda} \tag{129}$$

Substituindo os valores E_n, obtidos em (129), na Equação (127), segue que

$$u(x,t) = \frac{N}{2\Lambda}\sum_{n=-\infty}^{+\infty} e^{-D\left|\frac{n\pi}{\Lambda}\right|^\gamma t} e^{i\frac{n\pi}{\Lambda}x}. \tag{130}$$

As condições de contorno impõem que a derivada parcial $u_x(x,t)$em relação a x se anule em $x = \pm\Lambda$. Tal condição elimina os senos e a parte imaginária. Além disso, os cossenos são pares e a solução pode ser escrita como

$$u(x,t) = \frac{N}{\Lambda}\sum_{n=0}^{\infty} e^{-D\left|\frac{n\pi}{\Lambda}\right|^\gamma t}\cos\frac{n\pi x}{\Lambda} \tag{131}$$

Este resultado é precisamente o obtido na discretização da $L_\gamma(x,t)$ na Equação (124), completando a análise e dando sentido ao parâmetro de escala no contexto da implementa-

ção computacional. O último parâmetro a ser alterado na implementação computacional é a finitude no número de termos, $n_{max} = \Omega$, assim,

$$l_\gamma(x,t) \approx \frac{N}{\Lambda} \sum_{n=0}^{\Omega} e^{-D|\frac{n\pi}{\Lambda}|^\gamma t} \cos \frac{n\pi x}{\Lambda}\,, \tag{132}$$

O PVC (122) com as condições de contorno de isolamento, estas sobre as derivadas, u_x são as que levam à solução de Lévy $L_\gamma(x,t)$, função que é uma ferramenta que pode ser utilizada para avaliar dados e determinar os atributos estatísticos de difusão anômala. Tal função possui como característica decaimento como lei de potência (DOSS; MAURO, 2021). É possível mostrar que $L_\gamma(x,t) \to 1/x^\gamma$ quando $x \to \infty$ ou seja, a potência de decaimento é o próprio expoente γ (LUTZ, 2003).

Por exemplo, pode ser empregado para combinar dados experimentais e identificar os parâmetros de difusão anômala examinando o deslocamento quadrático médio das partículas ao longo do tempo e comparando-o com a função de Lévy. Se um ajuste adequado for obtido, os parâmetros da função de Lévy podem ser usados para inferir o processo de difusão. Além disso, pode ser aplicado para analisar dados de sistemas financeiros e extrair informações sobre o comportamento dos preços dos ativos, modelar o movimento de partículas em meios porosos e entender o impacto da porosidade na difusão anômala.

Tabela 2 – Tabela de parâmetros presentes na função de Lévy discretizada e suas descrições.

Parâmetro	Símbolo	Domínio	Descrição
Coeficiente de Difusão	D	$(0,+\infty)$	Controla a taxa de espalhamento das partículas
Expoente anômalo	γ	$(1,2)$	Determina o grau de desvio da difusão normal
Amplitude	A	$(0,+\infty)$	Controla a magnitude geral da função de densidade de probabilidade
Parâmetro de escala	Λ	$(0,+\infty)$	Determina a escala na qual as condições de contorno são impostas

Como os dados são finitos, a taxa de eventos que formam a cauda da distribuição fica prejudicada em relação ao centro. Como o problema é finito, ele necessariamente possui escala em x. Esta escala é manifesta na forma das condições de contorno como feito acima conforme o Problema (122). Uma consequência é que a solução discreta $l_\gamma(x,t) = l_\gamma(x+2\Lambda,t)$ é periódica de período 2Λ.

Uma forma de obter o ajuste entre dados e a função discretizada de Lévy seria o método de gradiente através do algoritmo de BFGS (FLETCHER, 2013). Porém, dependendo do conjunto de dados, ele é incapaz de produzir bons ajustes pela impossibilidade de calcular

gradiente de função com variáveis inteiras, pois é projetado para ajustar funções contínuas e deriváveis. Ele usa a derivada da função para calcular a direção de máximo crescimento e, em seguida, atualiza os valores das variáveis de acordo com essa direção (CHONG; ZAK, 2013). Como as funções com variáveis inteiras não são contínuas e nem deriváveis, o método do gradiente não é apropriado para esse tipo de ajuste. Outro fator dificultador é a sensibilidade do chute inicial, necessário na busca, tornando o processo de otimização ainda mais complexo.

Uma alternativa para a realização de ajustes é o uso de algoritmos genéticos. Trata-se de um tipo de algoritmo evolucionário utilizado em otimização (KRAMER, 2017). A função objetivo a ser minimizada é dada por (133).

$$S = \sum_{j=0}^{n-1} |p_j - L(x_j, A, \gamma, D, \Lambda)|, \tag{133}$$

onde n é o comprimento do vetor de dados x, p_j é o j-ésimo elemento do vetor de probabilidades, x_j é o j-ésimo elemento do vetor x.

Os parâmetros a serem otimizados são $A, \gamma, D, \Lambda, \Omega$, dados na Tabela 2. Todos são do tipo *float*, com exceção de Ω, o qual é do tipo *integer*. O AG não necessita de um chute inicial para cada parâmetro como no BFGS, mas é preciso estabelecer uma faixa de valores para cada um, que devem ser conforme a natureza do problema, tomando α, por exemplo, entre 1 e 2. Para o parâmetro A, foi calculada a área aproximada do histograma e estabelecido um intervalo em torno deste valor. O coeficiente de difusão D foi tomado entre 1 e 100. Para as variáveis Ω e Λ, a princípio foram estabelecidos valores quaisquer. Entretanto, como será discutido, após diversos testes, consideramos valores mais elevados para esses parâmetros.

No AG utilizado no nosso trabalho consideramos as etapas de *captura* dos elementos da população, *seleção* entre os melhores, *cruzamento* entre os selecionados, *mutação* entre eles e *avaliação* final na função objetivo. Como critério de parada, é fixado um número máximo de gerações que, ao ser atingido, retorna parâmetros otimizados para função discretizada de Lévy (132). Um fluxograma é mostrado na Figura 3. Nesse modelo, os parâmetros α e γ são as ordens das derivadas fracionárias e podem variar em $0 < \alpha < 1$ e $1 < \gamma < 2$. Note que $\alpha = 1$ e $\gamma = 1$ recuperam a difusão normal como na Equação (83). Além disso, o termo $v\dfrac{\partial u}{\partial x}$ contém um parâmetro v, que indica a velocidade das partículas. e, quando v é não-nulo, ocorre um fenômeno chamado "*drift*".

Para resolver o PVC (122), primeiramente, aplica-se a transformada de Laplace na variável t nos dois lados da equação diferencial, obtendo a igualdade em (134).

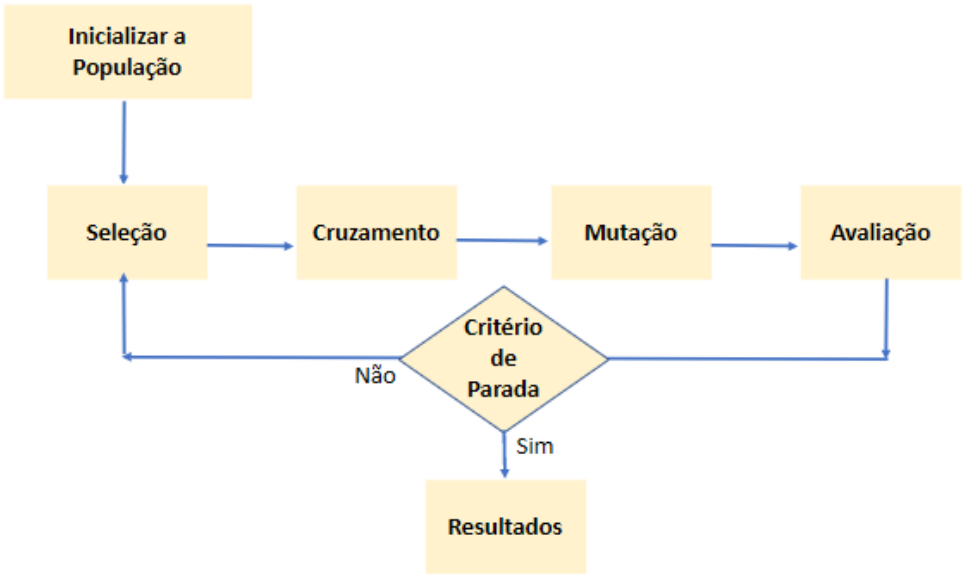


Figura 3 – Fluxograma do algoritmo genético utilizado neste trabalho para otimização que determina valores para os parâmetros da função de Levy discretizada, modelo para fenômenos superdifusivos. O algoritmo é uma heurística que possui como etapas principais em uma estrutura de repetição a seleção dos indivíduos, o cruzamento entre os selecionados, uma mutação e uma avaliação, parando quando é atingido o critério escolhido.

$$s^{\alpha}\tilde{u}(x,s) - s^{\alpha-1}\tilde{u}(x,0) = D\frac{\partial^{\gamma}\tilde{u}(x,s)}{\partial|x|^{\gamma}} - v\frac{\partial\tilde{u}(x,s)}{\partial x}. \tag{134}$$

Aplicando a transformada de Fourier em (134), tem-se

$$s^{\alpha}\hat{\tilde{u}}(k,s) - s^{\alpha-1}\hat{\tilde{u}}(k,0) = -D|k|^{\gamma}\hat{\tilde{u}}(k,s) - v(-ik)\hat{\tilde{u}}(k,s). \tag{135}$$

Colocando $\hat{\tilde{u}}(k,s)$ em evidência na Equação (135), pode-se obter

$$\hat{\tilde{u}}(k,s) = \frac{s^{\alpha-1}\hat{\tilde{u}}(k,0)}{s^{\alpha} - (ivk - D|k|^{\gamma})}. \tag{136}$$

Considere a função de Mittag-Leffler de ordem α, $E_{\alpha}(z)$, como na Subseção 2.3.3. Conforme mostrado em (MAINARDI, 2010), segue que

$$\mathcal{L}\{E_{\alpha}(-\lambda t^{\alpha})\} = \frac{s^{\alpha-1}}{s^{\alpha}+\lambda}, \quad Re(s) > |c|^{\frac{1}{\alpha}}, \quad c \in \mathbb{C}, \quad 0 < \alpha \leq 1. \tag{137}$$

Aplicando a transformada inversa de Laplace em (136) e usando o resultado (137), obtém-se

$$\hat{u}(k,t) = E_\alpha([ivk - D|k|^\gamma]t^\alpha)\hat{u}(k,0). \tag{138}$$

Aplicando a transformada de Fourier inversa em (138) e utilizando o Teorema da Convolução (44), é possível escrever a solução do problema da seguinte forma

$$u(x,t) = G_{\alpha,\gamma} * u(x,0) = G_{\alpha,\gamma} * \delta(x) = G_{\alpha,\gamma}(x,t). \tag{139}$$

Considerando a condição inicial $g(x') = \delta(x')$ em (139), a solução se resume a

$$G_{\alpha,\gamma}(x,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ikx}E_\alpha([ivk - D|k|^\gamma]t^\alpha)dk, \tag{140}$$

onde $G_{\alpha,\gamma}$, conhecida como função de Green (PODLUBNY, 1998), é definida por

$$G_{\alpha,\gamma} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ikx}E_\alpha((-ik)^\gamma Dt^\alpha)dk. \tag{141}$$

Para o caso particular da equação (122) no qual a velocidade é nula, $v = 0$, em (140), tem-se

$$G_{\alpha,\gamma}(x,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ikx}E_\alpha([-D|k|^\gamma]t^\alpha)dk, \tag{142}$$

Se o modelo teórico (122) descreve o caso subdifusivo, então $0 < \alpha \leq 1$ e $\gamma = 2$ (PEREIRA, 2018).

$$G_{\alpha,2}(x,t) = \frac{t^{-\alpha/2}}{2\sqrt{D}}\sum_{n=0}^{+\infty}\frac{(-z)^n}{n!\Gamma\left(\frac{-\alpha n}{2} + 1 - \frac{\alpha}{2}\right)}, \qquad z = \frac{|x|}{\sqrt{D}t^{\alpha/2}}. \tag{143}$$

Se o modelo teórico (122) descreve o caso superdifusivo, então $\alpha = 1$ e $1 < \gamma \leq 2$ (PEREIRA, 2018).

$$G_{1,\gamma}(x,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-D|k|^\gamma t}\cos(kx)dk. \tag{144}$$

A função em (144) também é conhecida como função de Lévy (PODLUBNY, 1998) e é denotada como $L_\gamma(x,t)$.

4.1.1 Modelando voos subdifusivos e superdifusivos

Como casos especiais do modelo teórico estudado nesta seção, as equações vistas a seguir modelam os casos superdifusivo e subdifusivo. Para obter um modelo em grande escala para a difusão anômala com voos superdifusivos, é possível utilizar a equação diferencial parcial fracionária espacial (EDFE) (PEREIRA, 2018)

$$\frac{\partial u}{\partial t} = D_\gamma \frac{\partial^\gamma u(x,t)}{\partial |x|^\gamma}. \tag{145}$$

Trata-se de um caso particular da equação (122), cuja solução é dada por (144).

Também é possível utilizar um modelo teórico referente ao caso de voos subdifusivos através da equação diferencial parcial fracionária temporal (EDFT) (PEREIRA, 2018)

$$\frac{\partial^\alpha u}{\partial t^\alpha} = D_\alpha \frac{\partial^2 u(x,t)}{\partial x^2}. \tag{146}$$

cuja solução é dada por (143).

Na próxima seção analisaremos os voos que ocorrem sob o fenômeno de superdifusão e os ajustes feitos através da solução da Equação (145).

4.2 Resultados das simulações computacionais

Como forma de estudar o fenômeno de difusão anômala através de métodos estocásticos, neste trabalho buscou-se uma função de lei de potência de forma a evitar singularidades e que pudesse ser usada como uma PDF. A partir do método de Monte Carlo pela função inversa, discutido no Capítulo 2, e baseado nas funções oriundas do Cálculo Fracionário (CF), também discutidas neste capítulo, buscamos fazer um ajuste com o modelo teórico e as simulações.

Com o Método de Monte Carlo são gerados dois vetores, δ_k e δ_t, ambos com N números aleatórios. Eles representam os tamanhos dos saltos e os tempos de espera, respectivamente. Utilizando os vetores, são construídos dois outros vetores, $\sum \delta_x$ e em $\sum \delta_t$, que representam as somas parciais e, com esses últimos, tem-se um voo. Repete-se o processo descrito para N voos. Considera-se então um ponto t_0 e observa-se quais os valores alcançados nos tamanhos dos saltos até a posição correspondente ao índice para o qual o vetor acumulado do tempo de espera atingiu t_0. Um esquema teórico é representado na Tabela 3. Do lado esquerdo, são mostrados os valores sorteados a partir das duas leis de potência consideradas no Capítulo 2. Os valores em δ_x representam os comprimentos de saltos e, em δ_t estão os tempos de espera. Como os voos são considerados em uma

única dimensão, valores positivos em δ_x significam saltos em um sentido positivo. No lado direito da tabela, são mostrados os vetores $\sum \delta_x$ e em $\sum \delta_t$. Para um voo específico, são indicados em cada coluna a posição e o instante no qual a partícula está nesta localização. A partir de valores como na parte (b) da Tabela 3, é possível gerar gráficos para visualizar o comportamento dos voos.

Considerando no esquema hipotético um tempo de corte $t_0 = 0,7$, o algoritmo localiza a posição no qual uma partícula se encontra imediatamente após esse instante. Dessa forma, um vetor de posições armazena esses valores, que corresponde à terceira linha da Tabela 4. Finalmente, são obtidos parâmetros de grande escala da difusão anômala através da análise da evolução temporal da dispersão/variância dos deslocamentos.

Tabela 3 – Entendendo o algoritmo: esquema com voos hipotéticos.

(a) Comprimentos dos saltos e valores de tempos de espera

Voo 1					
δ_x	1	4	-2	3	1
δ_t	0,1	0,3	0,2	0,1	0,5
Voo 2					
δ_x	2	-4	3	4	2
δ_t	0,1	0,6	0,7	0,4	0,2
Voo 3					
δ_x	5	-2	1	3	4
δ_t	0,5	0,1	0,1	0,4	0,3
Voo 4					
δ_x	3	5	-4	-7	2
δ_t	0,2	0,2	0,3	0,6	0,2
Voo 5					
δ_x	-5	3	4	2	1
δ_t	0,1	0,3	0,5	0,1	0,1

(b) Posições das partículas em um determinado instante de tempo

Voo 1					
$\sum \delta_x$	1	5	3	**6**	7
$\sum \delta_t$	0,1	0,4	0,6	0,7	1,2
Voo 2					
$\sum \delta_x$	2	**-2**	1	5	7
$\sum \delta_t$	0,1	0,7	1,4	1,8	2
Voo 3					
$\sum \delta_x$	5	3	**4**	7	11
$\sum \delta_t$	0,5	0,6	0,7	1,1	1,4
Voo 4					
$\sum \delta_x$	3	8	**4**	-3	-1
$\sum \delta_t$	0,2	0,4	0,7	1,3	1,5
Voo 5					
$\sum \delta_x$	-5	-2	**2**	4	5
$\sum \delta_t$	0,1	0,4	0,9	1	1,1

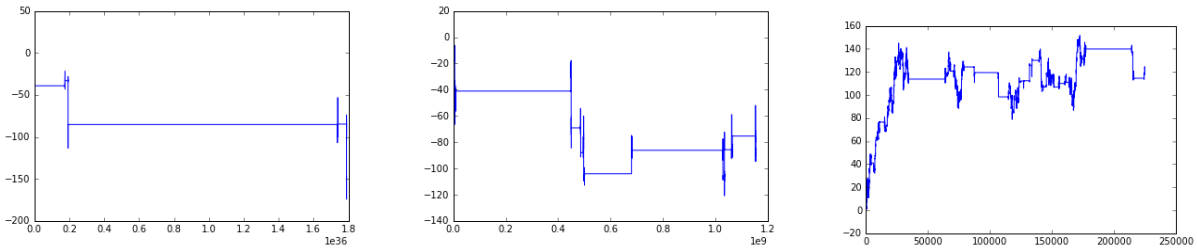
Tabela 4 – Considerando os tempos de corte em cada voo.

voo	1	2	3	4	5
índice	4	2	3	3	2
posição	**6**	**-2**	**4**	**4**	**2**

Os parâmetros s e r, presentes nas leis de potência utilizadas nas simulações, determinam os tipos de voos. A relação entre os parâmetros e tipos de voos estão sumarizados na Tabela 1 (PEREIRA, 2018),

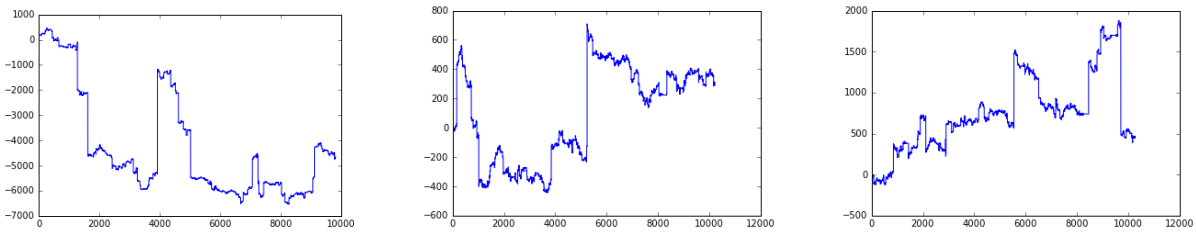
Nas Figuras 4, 5 e 6 são mostrados os comportamentos de alguns voos superdifusivos, subdifusos e com difusão normal, individualmente. Note que nas Figuras 4 e 5, em voos subdifusivos os tempos de espera são maiores e para voos superdifusivos ocorrem saltos maiores em intervalos de tempo muito curtos. Na Figura 6, são gerados alguns voos com

os parâmetros $s = 3$ e $r = 4$, que correspondem à difusão normal, conforme a Tabela 1. Nas Figuras 7, 8 e 9 um conjunto maior de voos é observado para valores fixados dos parâmetros r e s.



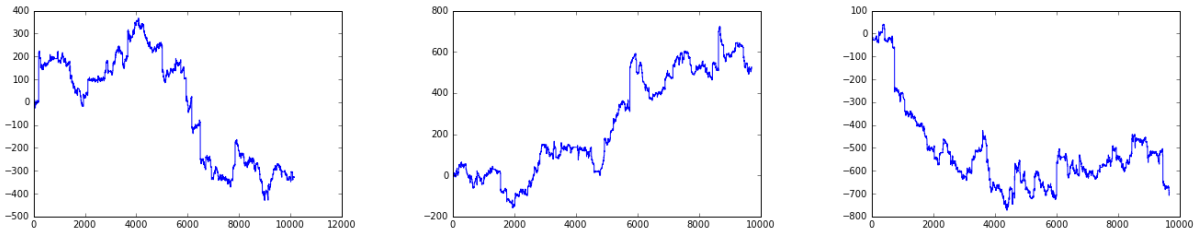
(a) $r = 4$ e $s = 1,1$ (b) $r = 4$ e $s = 1,5$ (c) $r = 4$ e $s = 1,9$

Figura 4 – Gráficos exibindo o comportamento de voos subdifusivos para o expoente $r = 4$, referente a PDF de comprimento de saltos, e diferentes valores de s, expoente da PDF de tempos de espera. O eixo das abscissas mostra a medida do tempo e o eixo das ordenadas, as posições. Nos três casos, são observados longos tempos de espera entre diferentes saltos.



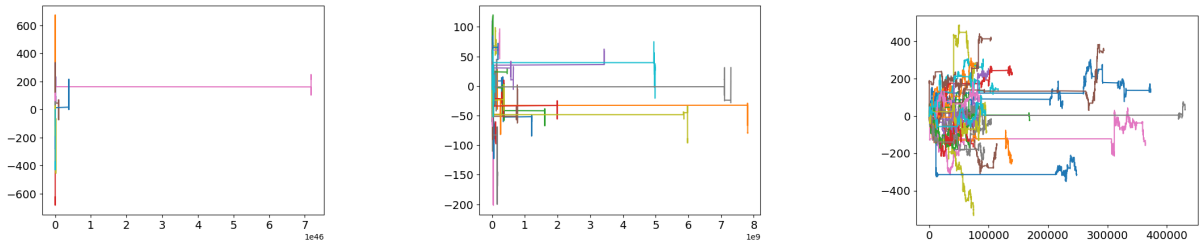
(a) $r = 2,3$ e $s = 3$ (b) $r = 2,5$ e $s = 3$ (c) $r = 2,7$ e $s = 3$

Figura 5 – Gráficos exibindo o comportamento de voos subdifusivos para diferentes valores do expoente r, referente a PDF de comprimento de saltos, e expoente $s = 3$, da PDF de tempos de espera. O eixo das abscissas mostra a medida do tempo e o eixo das ordenadas, as posições. Nos três casos, observa-se grandes saltos.



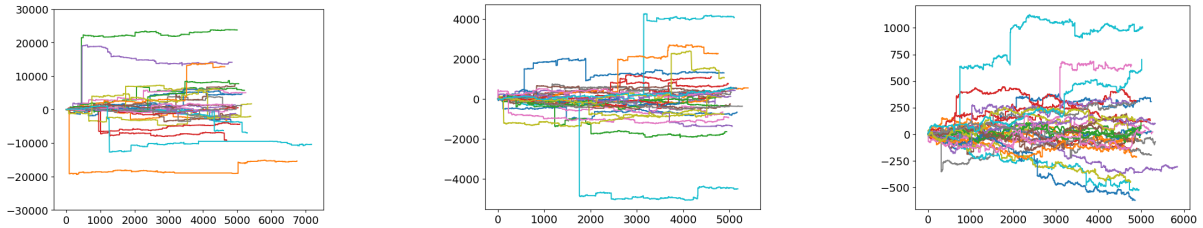
(a) $r = 4$ e $s = 3$ (b) $r = 4$ e $s = 3$ (c) $r = 4$ e $s = 3$

Figura 6 – Gráficos exibindo o comportamento de voos difusivos normais para os expoentes $r = 4$ e $s = 3$, referente, respectivamente, a PDF de comprimento de saltos e ao tempos de espera. O eixo das abscissas mostra a medida do tempo e o eixo das ordenadas, as posições. Nos casos exibidos, não ocorre longos tempos de espera e nem grandes saltos.



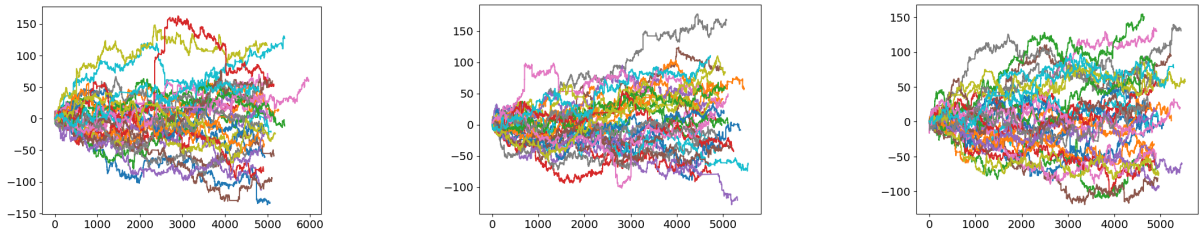
(a) $r = 4$ e $s = 1,1$ (b) $r = 4$ e $s = 1,5$ (c) $r = 4$ e $s = 1,9$

Figura 7 – Gráficos com conjuntos de voos subdifusivos para o expoente $r = 4$ da PDF de comprimento de saltos e diferentes valores do expoente s para o PDF dos tempos de espera. Cada voo é mostrado em uma cor diferente, nos quais o eixo das abscissas mostra a medida do tempo e o eixo das ordenadas, as posições.



(a) $r = 2,1$ e $s = 3$ (b) $r = 2,5$ e $s = 3$ (c) $r = 2,9$ e $s = 3$

Figura 8 – Gráficos com conjuntos de voos superdifusivos para diferentes valores do expoente r da PDF de comprimento de saltos e expoente $s = 3$ para o PDF dos tempos de espera. Cada voo é mostrado em uma cor diferente, nos quais o eixo das abscissas mostra a medida do tempo e o eixo das ordenadas, as posições.



(a) $r = 4$ e $s = 3$ (b) $r = 4$ e $s = 3$ (c) $r = 4$ e $s = 3$

Figura 9 – Gráficos com conjuntos de voos superdifusivos para o valor do expoente $r = 4$ da PDF de comprimento de saltos e expoente $s = 3$ para o PDF dos tempos de espera. Cada voo é mostrado em uma cor diferente, nos quais o eixo das abscissas mostra a medida do tempo e o eixo das ordenadas, as posições.

Com intuito de investigar estruturas e dinâmicas de sistemas que contêm muitas partículas que se movimentam de forma irregular em pequena escala, mas regular em grande escala, foram calculados, através de simulações, os parâmetros para os movimentos em grande escala. A obtenção desses dados foi dada por meio de uma calibração na qual foi gerado um número N suficientemente grande de voos. O objetivo é utilizar tais parâmetros para confrontar com os obtidos nas simulações com os movimentos em pequena escala usando uma formulação através do CRTW.

Para os casos subdifusivos, o modelo teórico fornece a solução (143). Nesse caso, o deslocamento quadrático médio das partículas é dado por (PEREIRA, 2018)

$$
\begin{aligned}
\langle x^2(t)\rangle &= \int_{-\infty}^{+\infty} x^2 G_\alpha(x,t)dx. \\
{}_0^{Ca}D_t^{\alpha} \int_{-\infty}^{+\infty} G(x,t)x^2 dx &= 2D_\alpha. \\
{}_0^{Ca}D_t^{-1}\left[{}_0^{Ca}D_t^{\alpha}(\langle x^2(t)\rangle)\right] &= {}_0^{Ca}D_t^{-1}\left[2D_\alpha\right]. \\
{}_0^{Ca}D_t^{\alpha-1}(\langle x^2(t)\rangle) &= 2D_\alpha t. \\
{}_0^{Ca}D_t^{-(\alpha-1)}({}_0^{Ca}D_t^{\alpha-1}(\langle x^2(t)\rangle)) &= 2D_\alpha t.
\end{aligned}
$$

Segue que

$$
{}_0^{Ca}D_t^{1-\alpha} t = \frac{\Gamma(2)}{\Gamma(1+\alpha)}, \quad 0<\alpha<1. \tag{147}
$$

Consequentemente,

$$
\langle x^2(t)\rangle = \frac{2D_\alpha}{\Gamma(\alpha+1)} t^\alpha, \quad 0<\alpha<1. \tag{148}
$$

Aplicando logaritmo nos dois lados da Equação (148), conclui-se que

$$
\log\langle x^2(t)\rangle = \alpha \log t + \log\left(\frac{2D_\alpha}{\Gamma(\alpha+1)}\right). \tag{149}
$$

Comparando a Equação (149) com (121), é possível obter as relações

$$
\alpha_d = \mu \qquad \text{e} \qquad D_\alpha = D_\mu \Gamma(\alpha_d+1). \tag{150}
$$

Em relação aos casos superdifusivos, o deslocamento quadrático médio é dado por (PEREIRA, 2018)

$$
\langle x^2(t)\rangle = 2D_\gamma t^{2/\gamma}, \quad 1<\gamma<2. \tag{151}
$$

Aplicando logaritmo nos dois lados de (151), tem-se

$$
\log\langle x^2(t)\rangle = \frac{2}{\gamma}\log t + \log(2D_\gamma). \tag{152}
$$

Ao comparar (152) com (121), obtém-se as relações

$$\gamma_d = \frac{2}{\mu} \qquad \text{e} \qquad D_{\gamma_d} = D_\mu. \tag{153}$$

Nas simulações, nota-se que o modelo é sensível à escolha das potências r e s das leis de potência para o comprimento de salto e para os tempos de espera, respectivamente.

Para a realização de ajustes, existem algumas alternativas disponíveis na literatura. No caso subdifusivo, por exemplo, é possível utilizar o método de Nelder-Mead. Ajustes foram realizados nos histogramas obtidos da geração dos voos simulados. As funções (143) e (144) foram usadas para alguns casos subdifusivos e superdifusivos, respectivamente, sendo a última a função de Lévy discutida no Capítulo 2. Os parâmetros obtidos podem ser verificados nas Tabelas 5 e 6, considerando diferentes valores dos expoentes s e r, referentes a PDFs de tempos de espera e de comprimentos de saltos, respectivamente. Os resultados dos ajustes são mostrados nas Figuras 10 e 11. Cabe ressaltar que no caso superdifusivo, nesta parte do trabalho, foi utilizado o algoritmo de BFGS para superdifusão, prejudicando a qualidade do ajuste. Um método alternativa será discutido na sequência do texto. Por outro lado, até aqui, fica estabelecida uma conexão entre um caso particular do modelo teórico estudado e o modelo computacional adotado. Almeja-se estender a conexão para o caso teórico mais geral com uma adaptação do modelo computacional utilizado até então.

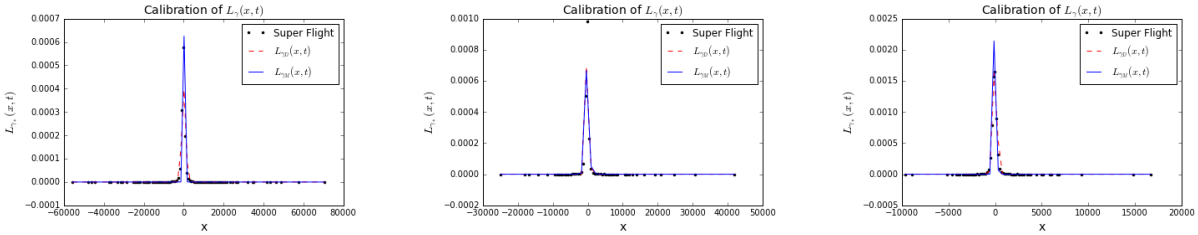
Tabela 5 – Parâmetros obtidos nos ajustes para o caso superdifusivo com expoente $s = 3$.

r	γ	D_γ
$2,3$	1,3348	0,9993
$2,7$	1,6437	0,9259
$2,9$	1,5988	0,9224

Tabela 6 – Parâmetros obtidos nos ajustes para o caso subdifusivo com expoente $r = 4$.

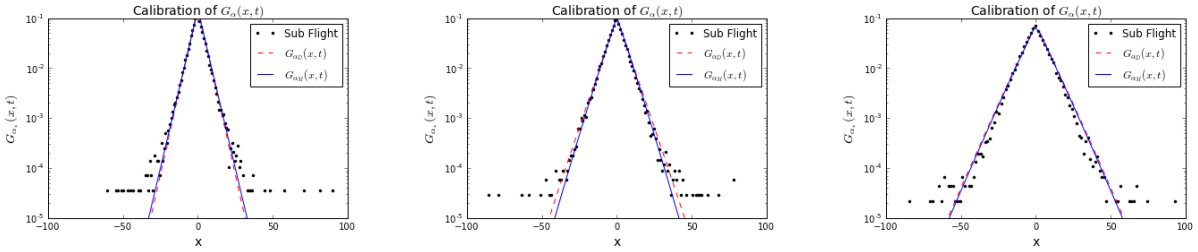
s	α	D_α
$1,9$	0,6695	0,7548
$1,7$	0,6942	0,2060
$1,5$	0,2345	2,3632
$1,3$	0,0010	11,9297

Como mencionado anteriormente, tais ajustes, especialmente no caso superdifusivo tornam-se dificultados pela forma dos histogramas. Como existe uma grande dispersão de dados, é característico dos casos superdifusivos possuir uma pequena quantidade de pontos com frequências muito elevadas e uma abundância de pontos com frequência mais baixa, longe da média, constituindo uma grande cauda, ou seja, como os dados são finitos, a taxa de eventos que formam a cauda da distribuição fica prejudicada em relação ao

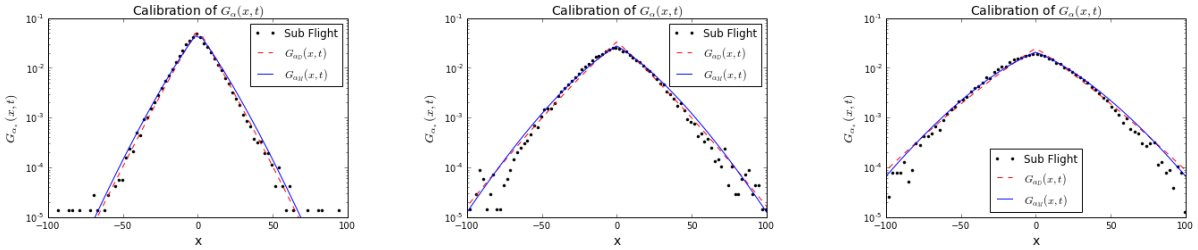


(a) $r = 2,3$ (b) $r = 2,7$ (c) $r = 2,9$

Figura 10 – Ajustes realizados com o algoritmo de BFGS para histogramas de dados obtidos para simulações com $N = 30000$ (número de voos), expoente da PDF de tempos de espera $s = 3$ e diferentes valores do expoente r da PDF dos comprimentos de saltos, contemplando casos superdifusivos.



(a) Caso $s = 1,3$ (b) Caso $s = 1,4$ (c) Caso $s = 1,5$



(d) Caso $s = 1,7$ (e) Caso $s = 1,8$ (f) Caso $s = 1,9$

Figura 11 – Ajustes realizados com o algoritmo de Nelder-Mead para histogramas de dados obtidos para simulações com $N = 30000$ (número de voos), expoente da PDF de comprimentos de salto $r = 4$ e diferentes valores do expoente s da PDF de tempos de espera, contemplando casos subdifusivos.

centro, criando também uma escala para o problema em x. É importante recordar que, computacionalmente, a função de Lévy $L_\gamma(x,t)$ não pode ser representada com variáveis reais, já que máquinas são incapazes de representar todos os números, havendo uma finitude no problema. Por isso, é necessária uma discretização, discutida no Capítulo 2. Assim, o número de parâmetros a serem obtidos é ainda maior que o necessário para descrever a função, havendo uma variável Λ, que está relacionada com a periodicidade da função, e uma variável Ω, que surge devido ao truncamento feito na série utilizada para representar a função de Lévy discretizada, pois a função original é uma integral imprópria. Esse último parâmetro indica a quantidade de termos a serem considerados no somatório.

Dessa forma, o método de BFGS torna-se incapaz de produzir bons ajustes. Além disso, um chute inicial para cada parâmetro deve ser feito, tornando o processo de otimização ainda mais complexo. Uma alternativa encontrada no trabalho foi a utilização de um algoritmo evolucionário, conhecido como AG, adaptado ao problema, visando obter os parâmetros que descrevem a superdifusão, incluindo obtenção e análise dos parâmetros. Utilizando os princípios do darwinismo, este método probabilístico produz soluções em um espaço factível que sofre evoluções até atingir resultados próximos do ideal, com uma precisão adequada comprovada por testes realizados ao longo do trabalho. Para este tipo de algoritmo, não há a necessidade de um chute inicial para cada um dos parâmetros envolvidos no problema, apenas uma delimitação do espaço no qual a busca deverá ser realizada.

4.2.1 Escolha do histograma para o ajuste

A impossibilidade de utilizar a escala logarítmica como uma ferramenta adequada para obtenção de parâmetros que descrevem a superdifusão tem como implicação a necessidade de utilizar outras estratégias como o ajuste da própria distribuição. Entretanto, já é sabido que, para caudas decaindo em leis de potência, a produção de histogramas é problemática.

Os dados no processo de superdifusão são obtidos fazendo-se uma seção dos voos em um tempo arbitrariamente grande para obter os dados das distribuições de Lévy. Para construção dos histogramas, é necessário escolher uma quantidade adequada de arestas no histograma para otimizar a sua largura. O uso de tamanhos pode influenciar na interpretação dos resultados e levar à análise incorreta do problema (WHITE; ENQUIST; GREEN, 2008). Por isso, uma escolha a ser observada é a largura das barras e qual a quantidade a ser utilizada, pois para um tamanho de compartimento muito pequeno a altura da barra em cada compartimento sofrerá uma grande flutuação estatística devido à falta de amostras em cada compartimento. Além disso, se for escolhido um tamanho de compartimento muito grande, o histograma não conseguirá representar corretamente a forma da distribuição pela resolução não ser boa o suficiente neste caso (SHIMAZAKI; SHINOMOTO, 2007). Existe, portanto, trade-off na escolha da largura e da quantidade de compartimentos (WHITE; ENQUIST; GREEN, 2008).

Nos histogramas foi utilizado o método de otimização de largura binária através do comando *histogram edge bin* do *Numpy* com *Python (3.9 Version)*. Um exemplo que demonstra a escolha ótima em busca de um perfil próximo ao suave é mostrado na Figura 12.

Consideramos o topo de cada aresta do histograma e selecionamos o ponto médio. Isso gera um gráfico com pontos discretos no plano bidimensional no qual os ajustes com a Função de Lévy podem ser realizados.

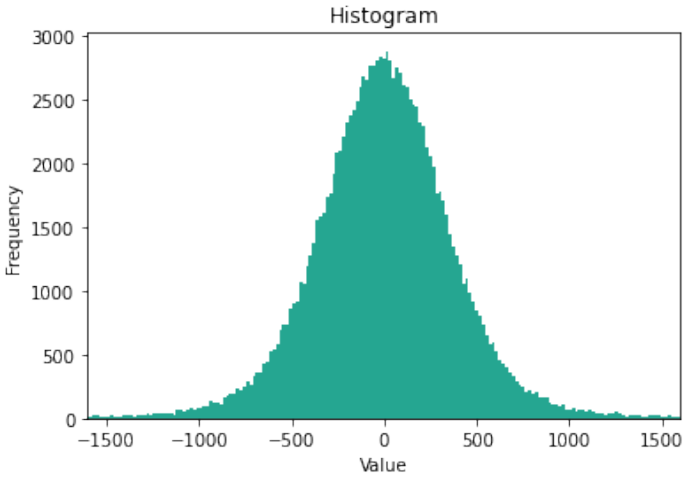


Figura 12 – Exemplo de histograma exibido com otimização na escolha das larguras. Os dados foram obtidos a partir de simulações realizadas com expoentes $r = 2,9$ e $s = 3$, das PDFs de comprimentos de saltos e tempos de espera, respectivamente, exibindo as frequências relativas nos eixos das ordenadas e diferentes valores para posições das partículas no eixo das abscissas.

Como o número de voos N é finito e computacionalmente limitado pelo custo, o número de barras/pontos na distribuição torna-se limitado. Além disso, o fenômeno é essencialmente livre de escala e pode ter caudas pesadas e longas, importantes para caracterização dos parâmetros do processo. Consequentemente, sobram, comparado à cauda, poucos pontos na região central da distribuição (ver Fig. 13). Isto torna o problema de ajuste difícil.

4.2.2 Calibração do Algoritmo Genético

Como mencionado, a utilização de um AG para realização dos ajustes foi utilizada devido às circunstâncias que envolveram a realização do trabalho. Para atestar a capacidade de obtenção de bons ajustes, foram considerados modelo teóricos para um conjunto de parâmetros e adicionado ruído a cada um deles, ou seja, criados pontos sintetizados como explicado na Seção 3.3. Isso foi feito adicionando variações de até 5% nas coordenadas y dos pontos do histograma gerado a partir de modelos teóricos. Os resultados são apresentados na Tabela 7 para diferentes valores de r, expoente da PDF de comprimento de salto, considerando a primeira linha como o modelo teórico e a segunda linha os parâmetros recuperados a partir do ajuste.

Atestando a qualidade do ajuste, também foi calculado o coeficiente de correlação de Pearson (SEDGWICK, 2012), uma medida estatística que mostra o grau de correlação entre duas variáveis da escala métrica, expressando a relação linear entre duas variáveis. Ele varia entre -1 e 1, onde -1 indica uma relação negativa perfeita, 1 indica uma relação positiva perfeita e 0 indica ausência de relação linear.

O coeficiente de correlação de Pearson é dado pela Equação (154).

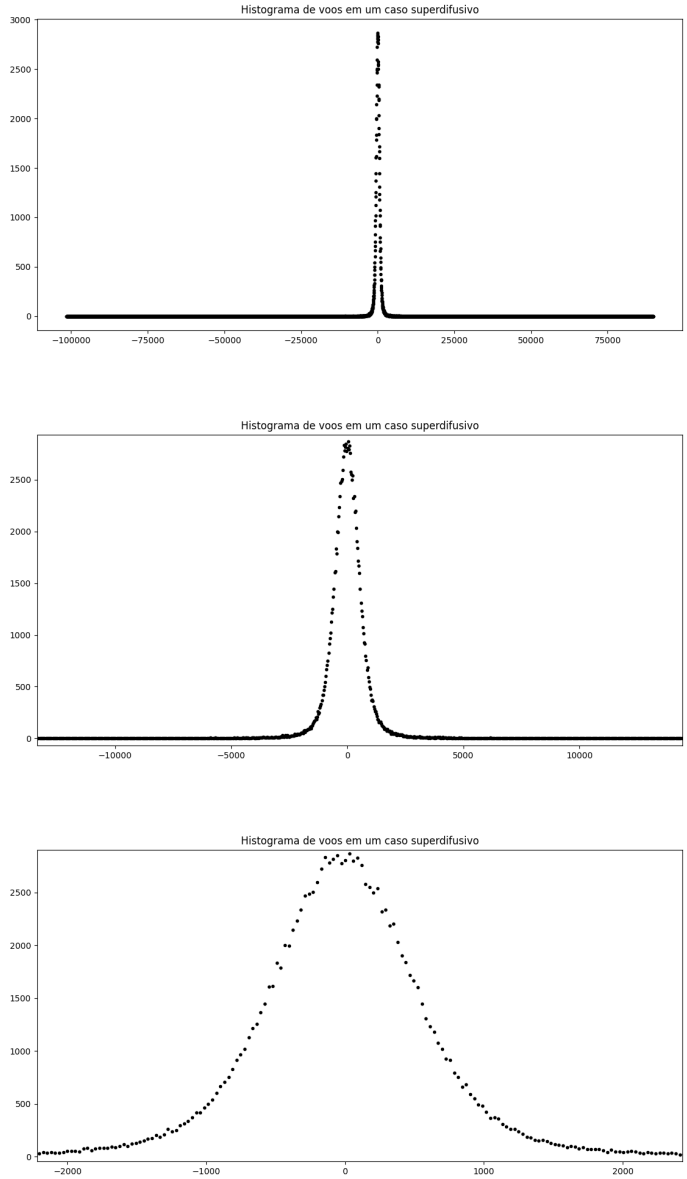


Figura 13 – Exemplo de dispersão de dados na distribuição de um voo superdifusivo. As figuras diferem apenas em escala. Nesta figura as barras do histograma são substituídas por pontos.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{154}$$

onde x_i e y_i são as observações das duas variáveis, $\bar{x}$ e $\bar{y}$ são as médias das observações das respectivas variáveis.

Para calibrar o modelo, foram realizados novos ajustes com aumento no número de pontos no histograma para 100 e no número de indivíduos utilizados no AG para 400, em casos onde a diferença do modelo teórico e do ajuste foi mais acentuada. Os resultados

Tabela 7 – Tabela com resultados dos ajustes realizados a partir de dados sintetizados, criados com adição de ruídos para atestar a qualidade do algoritmo genético utilizado no trabalho. Para diferentes casos superdifusivos, foram considerados modelos teóricos para um conjunto de parâmetros e adicionado ruído a eles. Foram adicionadas variações de até 5% nas coordenadas y dos pontos do histograma gerado a partir de modelos teóricos.de parâmetros presentes na função de Lévy discretizada.

Caso $r : 2,1$	γ	D	Λ	C.C.Pearson
	1,516	47,898	65500,372	-
	1,523	51,035	64872,753	0,999561
Caso $r : 2,2$	γ	D	Λ	C.C.Pearson
	1,765	124,927	40868,172	-
	1,750	110,883	40674,648	0,999723
Caso $r : 2,3$	γ	D	Λ	C.C.Pearson
	1,766	52,786	29529,541	-
	1,767	53,147	29560,214	0,999782
Caso $r : 2,4$	γ	D	Λ	C.C.Pearson
	1,171	0,782	8416,227	-
	1,188	0,881	8389,014	0,99978
Caso $r : 2,5$	γ	D	Λ	C.C.Pearson
	1,942	42,229	14085,542	-
	1,929	38,803	13938,066	0,999785
Caso $r : 2,6$	γ	D	Λ	C.C.Pearson
	1,691	5,279	9784,682	-
	1,698	5,500	9843,991	0,99979
Caso $r : 2,7$	γ	D	Λ	C.C.Pearson
	1,827	7,234	7438,759	-
	1,845	8,033	7483,495	0,999932
Caso $r : 2,8$	γ	D	Λ	C.C.Pearson
	1,774	3,715	4506,354	-
	1,782	3,871	4537,667	0,999797
Caso $r : 2,9$	γ	D	Λ	C.C.Pearson
	1,564	0,782	5288,328	-
	1,587	0,893	5320,279	0,999817

são mostrados na Tabela 8.

4.2.3 Análise do local de ajuste

Pela dificuldade em ajustar todo o conjunto de dados mencionada na Seção 4.2.1, inicialmente, no nosso trabalho foi proposta a realização de ajuste nos pontos suficientemente próximos da média, que possuem elevada frequência, como observado nas figuras.

Consideramos um corte dos pontos, que ficam restritos ao intervalo $[-\zeta, \zeta]$, onde ζ é um ponto de inflexão obtido ao interpolar os dados como uma função polinomial de terceiro grau. Para isso, foram considerados algoritmos de derivação discreta para o cálculo da

Tabela 8 – Tabela com resultados dos ajustes realizados.

		γ	D	Λ	C.C.Pearson
Caso $r : 2,2$	Modelo teórico	1,765	124,927	40868,172	-
	Ajuste	1,757	117,477	40475,75	0,999723
	Variação	-0,45 %	-6,12 %	-0,99 %	-
		γ	D	Λ	C.C.Pearson
Caso $r : 2,7$	Modelo teórico	1,827	7,234	7438,759	-
	Ajuste	1,82	7,0	7347,121	0,999789
	Variação	-0,36%	-3,26%	-1,22%	
		γ	D	Λ	C.C.Pearson
Caso $r : 2,9$	Modelo teórico	1,564	0,782	5288,328	-
	Ajuste	1,558	0,757	5305,334	0,99978
	Variação	-0,44 %	-3,13 %	0,23 %	

segunda derivada e de sua raiz. Através do AG, os parâmetros A, γ, D, Λ foram obtidos, representando respectivamente a área abaixo da curva, o expoente de difusão, o coeficiente de difusão e o semi período da função de ajuste para diferentes casos superdifusivos. Os resultados estão mostrados na Tabela 9.

Tabela 9 – Parâmetros da função de Lévy obtidos através de um AG para diferentes casos de superdifusão variando o valor do expoente r e fixando $s = 3$. A qualidade de cada ajuste é atestada pelo coeficiente de correlaçaõ de Pearson.

r	A	γ	D	Λ
2,1	34221213,4	1,61192571	92,5708700	39921,8182
2,2	22912614,9	1,68132942	63,1476051	28741,8182
2,3	15543415,8	1,82404692	78,7878788	20298,1818
2,4	10709694,0	1,82111437	36,2658847	15255,4545
2,5	8253186,29	1,84946237	25,4154448	9198,2014
2,6	5327404,46	1,96187683	24,6793744	10408,1818
2,7	5505129,89	1,97360704	20,5278592	3333,39198
2,8	3312310,15	1,92179863	7,5268817	5560,90909
2,9	2553489,46	1,83382209	2,9325513	4857,27273

Fixando o parâmetro A obtido nos ajustes anteriores e utilizando como base teórica a acurácia comprovada para testes realizados com dados sintetizados, foram obtidos os parâmetros γ, D, Λ com uma precisão superior à observada nos ajustes anteriores. Os dados são mostrados na Tabela 10 e ajustes referentes à essa tabela são mostrados nas Figuras 14, 15, 16, 17 e 18. Em cada caso ajustado, os parâmetros A, γ, D e Λ foram inicialmente obtidos, cuja descrição de cada um na função discretizada de Lévy (132) é mostrada na Tabela 2.

Na Tabela 10 são revelados os comportamentos de voos superdifusivos. Para casos com menores valores do expoente r, onde os efeitos da superdifusão são mais acentuados, são observados maiores valores do coeficiente de difusão D e do parâmetro Λ. Por outro

Tabela 10 – Parâmetros da função de Lévy discretizada obtidos através de um AG para diferentes casos de superdifusão variando o valor do expoente r, referente a PDF de comprimento de saltos, e fixando o expoente $s = 3$, referente a PDF de tempos de espera. Existe uma tendência de aumento de γ e diminuição dos parâmetros D e de Λ ao aumentar r. Valores do coeficiente de correlação de Pearson próximos a 1 atestam a qualidade do ajuste em todos os casos apresentados.

r	γ	D	Λ	C. C. de Pearson
2,1	1,619	101,249	43789,1224	0,960254
2,2	1,669	58,401	30924,8478	0,988045
2,3	1,819	75,823	20296,1976	0,992932
2,4	1,813	34,511	15444,8429	0,995161
2,5	1,834	22,874	9357,3491	0,995060
2,6	1,954	22,581	11925,5019	0,995137
2,7	1,952	18,475	3499,8987	0,992074
2,8	1,907	6,9610	5627,7704	0,996702
2,9	1,842	3,0475	4771,3327	0,995722

lado, existe uma tendência de aumento para o expoente γ quando o expoente r é também aumentado.

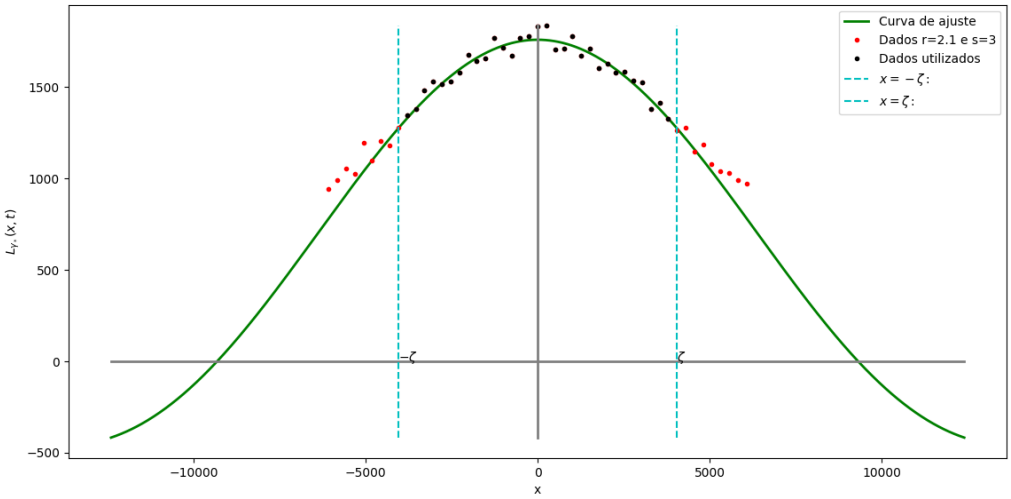


Figura 14 – Resultado do ajuste realizado com um algoritmo genético sobre o histograma de dados para o caso em expoentes $s = 3$ e $r = 2,1$, referentes a PDFs de tempos de espera e comprimentos de saltos, respectivamente, para determinar os parâmetros na função de Lévy. O ajuste foi feito apenas sobre os pontos localizados entre os pontos de corte indicados por retas verticais, em azul, escolhidos sobre os pontos de inflexão do polinômio de 3º grau que ajusta os pontos.

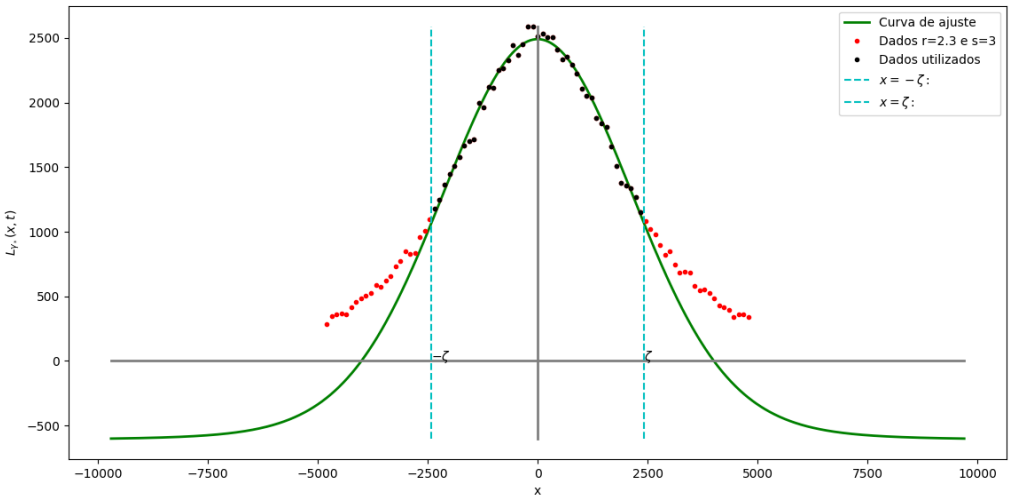


Figura 15 – Resultado do ajuste realizado com um algoritmo genético sobre o histograma de dados para o caso em expoentes $s = 3$ e $r = 2,3$, referentes a PDFs de tempos de espera e comprimentos de saltos, respectivamente, para determinar os parâmetros na função de Lévy. O ajuste foi feito apenas sobre os pontos localizados entre os pontos de corte indicados por retas verticais, em azul, escolhidos sobre os pontos de inflexão do polinômio de 3º grau que ajusta os pontos.

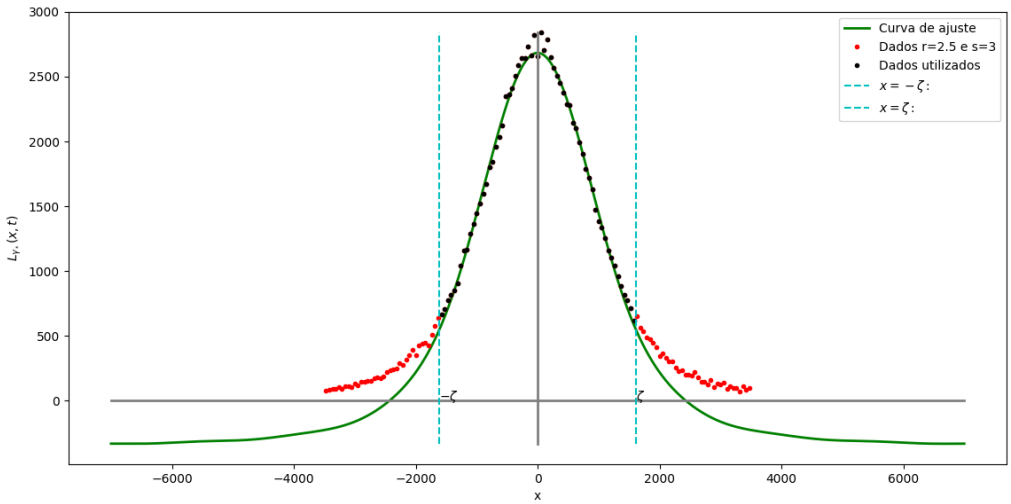


Figura 16 – Resultado do ajuste realizado com um algoritmo genético sobre o histograma de dados para o caso em expoentes $s = 3$ e $r = 2,5$, referentes a PDFs de tempos de espera e comprimentos de saltos, respectivamente, para determinar os parâmetros na função de Lévy. O ajuste foi feito apenas sobre os pontos localizados entre os pontos de corte indicados por retas verticais, em azul, escolhidos sobre os pontos de inflexão do polinômio de 3º grau que ajusta os pontos.

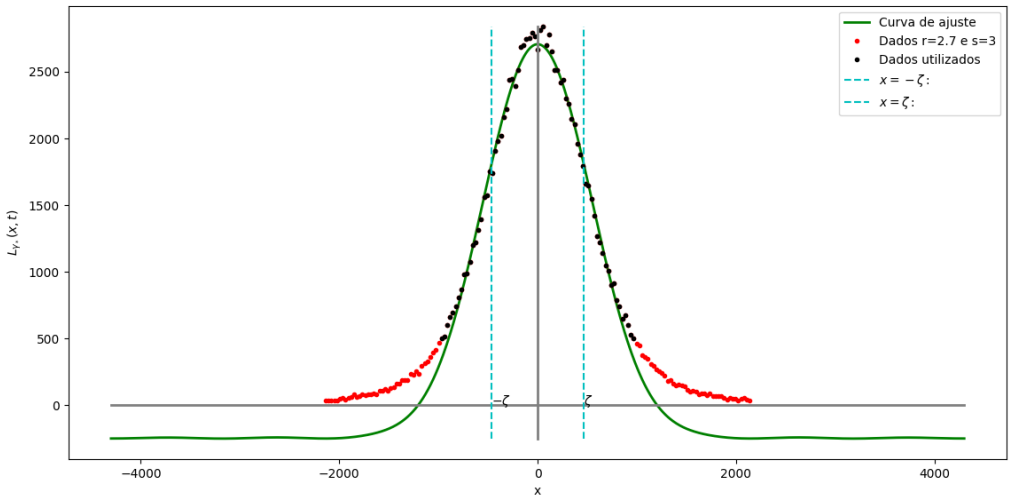


Figura 17 – Resultado do ajuste realizado com um algoritmo genético sobre o histograma de dados para o caso em expoentes $s = 3$ e $r = 2,7$, referentes a PDFs de tempos de espera e comprimentos de saltos, respectivamente, para determinar os parâmetros na função de Lévy. O ajuste foi feito apenas sobre os pontos localizados entre os pontos de corte indicados por retas verticais, em azul, escolhidos sobre os pontos de inflexão do polinômio de 3º grau que ajusta os pontos.

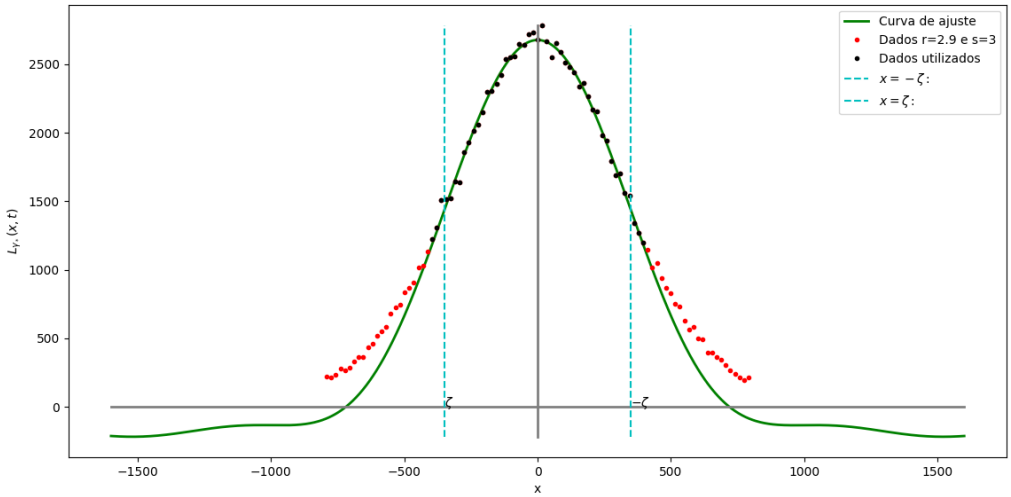


Figura 18 – Resultado do ajuste realizado com um algoritmo genético sobre o histograma de dados para o caso em expoentes $s = 3$ e $r = 2,9$, referentes a PDFs de tempos de espera e comprimentos de saltos, respectivamente, para determinar os parâmetros na função de Lévy. O ajuste foi feito apenas sobre os pontos localizados entre os pontos de corte indicados por retas verticais, em azul, escolhidos sobre os pontos de inflexão do polinômio de 3º grau que ajusta os pontos.

4.2.4 Finitude do número de voos N

Fenômenos reais, na prática, são realizados em função de um tempo contínuo. Entretanto, devido às limitações computacionais existentes em qualquer sistema que utilize algoritmos computacionais, os dados são, na prática, discretos, mesmo em supercomputadores. Mostramos nesta seção que a utilização de valores maiores do número de voos na geração de dados possuem um efeito que suaviza as indesejadas escadas do diabo.

Quando um fenômeno de superdifusão é modelado através de uma equação diferencial fracionária e uma solução é obtida analiticamente, o fenômeno é "sem escala". Porém, quando um N finito é escolhido na simulação e a integral é resolvida discretamente no ajuste feito por otimização, em ambos os casos, uma escala é introduzida ao problema no momento da observação no tempo de corte. Em outras palavras, se uma solução é periódica, seu período é, necessariamente, uma escala. Em relação à escada do diabo, a conformação se deve ao fato de que o problema é uma tendência à descamação. Isso significa que uma ou algumas partículas extrapolam qualquer escala de saltos e tornam a variância descontínua. Se fosse possível simular o movimento de infinitas partículas, esse comportamento de "extrapolação de escala" seria mais frequente e contínuo. Então, dado um tempo inicial de observação, a distribuição para um número N finito de partículas é tal que, em uma certa escala à direita e à esquerda, o número de partículas é zero. Isso pode estar relacionado a condições de contorno homogêneas nessas escalas. Em nosso estudo é fixado um tempo de corte e investigamos a existência de uma sucessão de N's que estabilizam os parâmetros dos valores, fazendo com que as escadas do diabo desapareçam ou tenham um efeito menos intensificado.

A Figura 19 mostra diferentes resultados obtidos variando o número de voos para um caso de superdifusão, com expoentes $r = 2,5$ e $s = 3$. Quanto maior o valor de N, mais alinhados estão os pontos, permitindo obter valores mais confiáveis para os parâmetros por meio da regressão linear.

4.3 Ajustando todo conjunto de dados para um caso superdifusivo

Conforme discutido na Subseção 4.2.3, uma maneira de ajustar os dados da amostra é restringir um intervalo no histograma, removendo a cauda e ajustando apenas entre os pontos de inflexão. Outra tentativa foi ajustar apenas na cauda, mas isso prejudica o ajuste anterior, como pode ser visto na Figura 20.

Por outro lado, analisando o modelo teórico, quando o valor de Λ é alto, a curva oscila e se torna negativa, mas se aproxima do eixo das abcissas com uma amplitude variável

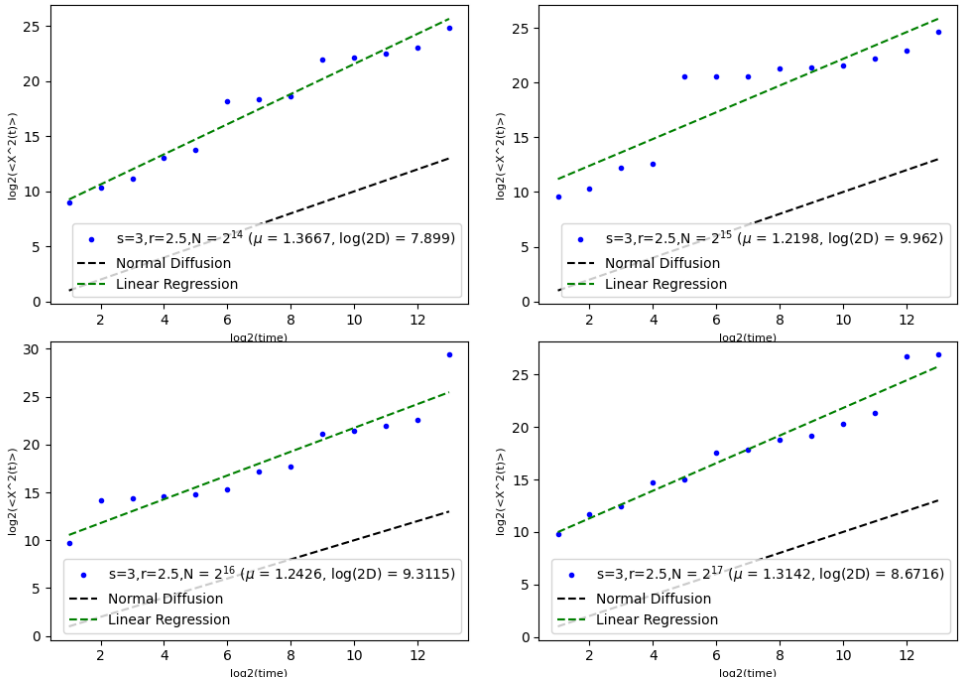


Figura 19 – Ajustes por dispersão realizados através de dados obtidos com diferentes números de voos N utilizados em simulações no mesmo caso superdifusivo $s = 3$ e $r = 2,5$. Em cada ajuste, é exibido o comportamento da variância das posições em função de cada tempo de corte. O aumento do valor de N suaviza o efeito das escadas do diabo e diferentes resultados foram encontrados para os parâmetros D e γ na função Lévy.

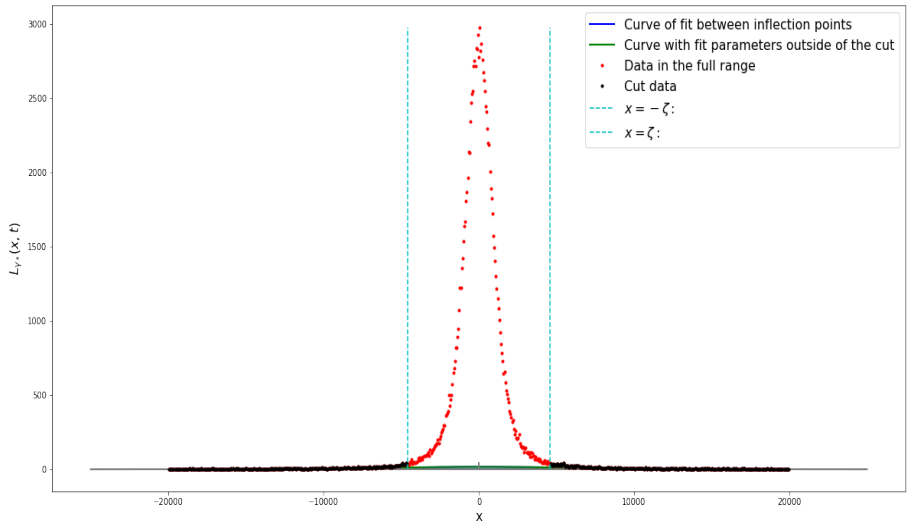


Figura 20 – Tentativa de ajuste, considerando apenas os dados da cauda para o caso $N = 2^{17}$ (número de voos) e expoentes $r = 2,5$ e $s = 3$, referentes ao comprimento de saltos e tempo de espera, respectivamente. O resultado mostra que a curva não se ajusta devidamente ao dados do histograma, como esperado.

de $1/x$, conhecida como “envelope”. Essa característica é desejável e é o objetivo para ajustar dados com altos valores de Ω na função $l_\gamma(x,t)$. Podemos notar que, quando os parâmetros Λ e Ω tendem ao infinito na função discretizada (132), esta tende para a função

Lévy original (3), ou seja, $l_\gamma(x,y) \to L_\gamma(x,t)$ quando $\Lambda, \Omega \to \infty$. Porém, realizar avaliações da função nessas condições tende a aumentar muito o custo computacional, tornando-se um obstáculo para um bom ajuste.

Apesar da dificuldade mencionada, considerando valores maiores desses parâmetros, obteve-se um ajuste que abrange todo o conjunto de dados, como mostra a Figura 21.

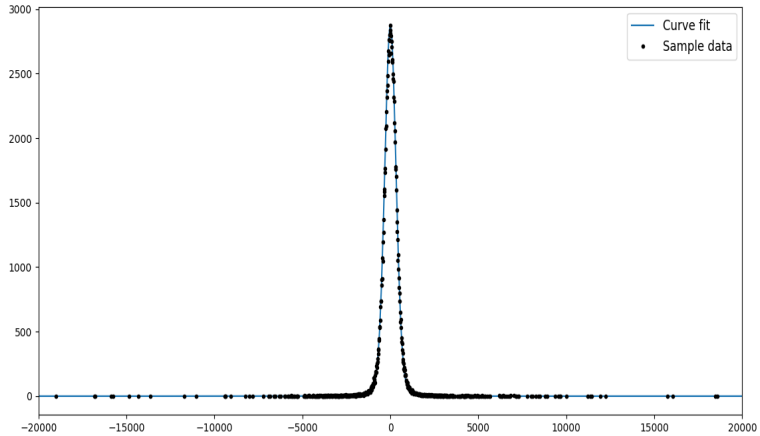


Figura 21 – Ajuste considerando todos os dados para o caso superdifusivo com $N = 2^{17}$ (número de voos) e valores de expoentes $r = 2,9$ e $s = 3$, referentes a PDFs de comprimento de salto e tempo de espera. Nesta otimização foram utilizados valores maiores para os parâmetros Ω e Λ em comparação aos ajustes realizados anteriormente.

Utilizando o coeficiente de correlação de Pearson como medida estatística entre a função com os parâmetros obtidos e os dados, obteve-se o valor $0,999595$, mostrando a qualidade do ajuste realizado.

4.3.1 Prescrição para realização do ajuste

No nosso trabalho, trabalhamos com determinados modelos e funções para difusão anômala cujos ajustes apresentavam uma série de dificuldades. Desde limitações computacionais até a interpretação dos parâmetros e como limitá-los. No caso específico da função discretizada (132), houve uma tentativa inicial de utilizar um algoritmo baseado no Método de Gradiente, o que resultou em ajustes que não são adequados aos dados do problema. Isso porque neste caso éramos forçados a manter os valores de Ω fixado com valores arbitrários. Uma consequência disso é que os valores de Λ diminuíram quando o valor do expoente r foi aumentado, sendo que A também foi aumentado. Esta metologia mostrou haver uma impossibilidade do conjunto formado por todos os dados se ajustarem corretamente à curva da função. Em outras palavras, os parâmetros Λ e A mostraram colineridade ou não se mostraram suficientemente independentes.

Sugerimos que a melhor forma de ajustar é manter o intervalo de ajuste em cada caso, utilizando o AG e, através dele, obter os valores adequados para todos os parâmetros da função de Lévy discretizada, incluindo Ω. Isso, porém, eleva o custo computacional do sistema pelo fato de a função objetivo do AG ter que avaliar a diferença de todos os dados considerados com a curva e de calcular um número maior de termos na avaliação da função $l_\gamma(x,t)$.

Capítulo 5

Busca de Fatores de Transcrição em Cadeias de DNA

Dentre diversos fenômenos nos quais se observa difusão anômala, pode-se citar o movimento de estruturas dentro das células de seres vivos. Nesta parte do trabalho, será abordada a busca de fatores de transcrição em cadeias de ácido desoxibonucleico ou, em inglês, *deoxyribonucleic acid* (DNA). Trata-se de proteínas importantes para reprodução de seres vivos e que devem encontrar um sítio ao longo do DNA para que a transcrição de fato ocorra. É de interesse descobrir como essa busca ocorre. Neste trabalho foram desenvolvidas simulações através do Método de Monte Carlo para estudar o mecanismo de busca desses fatores de transcrição, apresentando resultados a respeito do movimento (SAXTON, 2017).

Determinados estudos sugerem que os fatores de transcrição se movam dentro do citoplasma das células através de uma subdifusão, graças à presença de diversas outras estruturas no ambiente (ZAID; LOMHOLT; METZLER, 2009). Por outro lado, dependendo do modelo utilizado, conclui-se que a proteína reguladora do gene se move de forma superdifusiva ao longo do DNA, permitindo que ela encontre seu alvo rapidamente. Isso devido às longas caudas de potência dos saltos disponíveis para a proteína em um filamento de DNA, proporcionando, portanto, que a eficiência da busca seja aumentada, graças à superdifusão. Experimentos *in vivo* realizados na proteína lac repressora mostraram que ela pode atingir o alvo muito mais rapidamente que através da difusão normal, sendo o movimento superdifusivo capaz de acelerar tal busca (ROSA et al., 2010).

Presente na célula de todos os seres vivos, o DNA é uma molécula heteropolímera que, dentre outras funções, é responsável pela regulação genética. Apesar da grande variedade de seres vivos e da quantidade de informações genéticas contidas no DNA, é possível afirmar que todos têm uma estrutura fisicamente idêntica conhecida como B-DNA (ORDÓÑEZ, 2016). Essa consiste em duas estruturas suporte helicoidalmente torcidas

compostas de fosfato e açúcar com pares de bases de dois tipos, AT e GC, sendo A, T, C e G nucleotídeos presentes no DNA. A superfície da dupla hélice não é cilíndrica. Ela possui sulcos maiores e menores, importantes para o funcionamento do DNA, pois, na célula, numerosas proteínas devem se ligar a um local específico na molécula de DNA. Elas realizam essas tarefas nesses sulcos (FRANK-KAMENETSKII, 1997).

Em células eucarióticas, fatores de transcrição (FTs) são exemplos de tais proteínas. Elas se ligam ao DNA para permitir uma ligação entre o acido ribonucleico ARN-polimerase e o DNA, o que possibilita a transcrição e a futura tradução. Mais especificamente, para que a ARN-polimerase consiga encontrar o ponto de início da transcrição, ela necessita dos FT's, responsáveis por se ligar a sequências específicas de DNA (promotores) que caracterizem o local de início da transcrição, e recrutar a ARN-polimerase a esses sítios (SERENO; ACEBAL, 2018). Cada FT pode ter de um a dezenas de tais sítios específicos sobre o DNA. No local de ligação, o FT forma um complexo de DNA-proteína estável que pode ativar ou reprimir a transcrição dos genes próximos, dependendo do mecanismo de controle (SLUTSKY; MIRNY, 2004). Dessa forma, alguns fatores de transcrição também são capazes de inibir este processo, como o lac repressor, sobre o qual o mecanismo de ligação e reconhecimento de FT no DNA é estudado em (BAUER; METZLER, 2012).

Um exemplo no qual a questão discutida se mostra importante ocorre quando um vírus α-phage injeta seu DNA em uma bactéria do tipo *Escherichia coli*. Para as células infectadas, isso se torna uma corrida contra o tempo: para sobreviver, elas dependem da capacidade de enzima de restrição adequada para encontrar e reconhecer o local específico no DNA viral e depois cortá-la, tornando-o assim inoperável e inofensivo à bactéria. Se a enzima de restrição levar muito tempo para localizar seu alvo, a célula é morta. Esse sistema de restrição depende, portanto, da capacidade de determinadas proteínas de localizar seus respectivos locais-alvo específicos no DNA e fazer isso o mais rapidamente possível a fim de que a célula sobreviva ao ataque do vírus (HU; GROSBERG; SHKLOVSKII, 2006).

Há mais de quatro décadas, observou-se *in vitro* que a taxa de associação de fatores de transcrição do tipo lac repressor era de $7\times10^9 M^{-1}s^{-1}$, um resultado a princípio contraditório com a fórmula de difusão controlada de Smoluchowski, que produz um resultado de aproximadamente $10^8 M^{-1}s^{-1}$. Acreditava-se até então que essa diferença poderia ser justificada devido a efeitos da atração eletrostática entre o local carregado positivamente no FT e os grupos de fosfatos carregados positivamente no DNA (BAUER; METZLER, 2012). Mas em 1978, o trabalho (BERG; BLOMBERG, 1978) justificou as diferenças entre as taxas observadas experimentalmente e o resultado da fórmula de Smoluchowski. Nesse trabalho, afirmou-se que a princípio, os FTs se ligariam ao DNA através de uma busca em 3D dentro da célula, mas sem se ligar necessariamente ao seu local específico na molécula. Na sequência, os FTs executam uma busca unidimensional, sendo capazes também de se

mover pelo DNA ou de se soltar dele fazendo pequenos saltos para retornar a ele ou voltar à massa para executar uma nova busca tridimensional. Esse modelo teórico, denominado difusão facilitada, é aceito pela maioria dos pesquisadores e explica uma relação não monotônica da dependência da concentração de sal (BAUER; METZLER, 2012).

Recentemente, o estudo do processo de busca de fatores de transcrição nas cadeias de DNA tem ganhado motivação devido a melhorias na análise entre DNA e proteínas de ligação em uma única molécula, tornando-se capazes de ocorrer inclusive em células vivas, como feito em (MATHELIER et al., 2016) e (BAUER; METZLER, 2013), por exemplo.

Os resultados de estudos nessa linha podem afetar a compreensão dos princípios moleculares, biofísicas de regulação da transcrição, e melhorar significativamente a nossa capacidade de prever como as variações nas sequências de DNA, ou seja, mutações ou polimorfismos e concentrações de proteína, influenciam programas de expressão gênica em células vivas (AFEK et al., 2014). Apesar de ser largamente aceita, a forma geral da difusão facilitada a princípio apresentaria um problema conhecido como paradoxo velocidade-estabilidade. Por um lado, as proteínas percorrem o DNA unidimensionalmente, com constante de difusão proporcional a $\exp(-a^2)$, devido à rugosidade no potencial, onde a é a raiz quadrada média do potencial. Os FTs podem apenas se movimentar em direção ao alvo se $a < K_B T$. Por outro lado, a proteína deve permanecer ligada à sequência alvo tempo suficiente para que a expressão gênica real tome lugar, o que requer $a > 5K_B T$ (BAUER; METZLER, 2012). Aqui K_B representa a constante de Boltzmann e T a temperatura. Esses limites foram obtidos por (MIRNY et al., 2009). Para resolver o paradoxo exposto, podemos usar a ideia de que os fatores de transcrição podem assumir dois estados no processo: o de busca e o de reconhecimento (BAUER; METZLER, 2012). Dessa forma, tais proteínas podem assumir velocidades maiores quando estão no estado busca e, quando presentes no contorno da molécula de DNA, podem fazer atingir o alvo específico, com velocidade reduzida, permanecendo a partir desse momento no estado de reconhecimento.

Além disso, conforme (GOMEZ; KLUMPP, 2016), há diferentes tipos de obstáculos observados nas células, que estão presentes nas mesmas devido a outras proteínas. Cada um dos obstáculos pode implicar em algum tipo de comportamento dos resultados estudados, devido às diferentes dinâmicas apresentadas por eles. Por haver uma interação entre as partículas dentro das células, fazendo com o que o movimento de uma seja afetado pelo movimento das demais, espera-se que a difusão seja anômala com o caso não-linear, como visto nos Capítulos 2 e 4.

Um modelo computacional é proposto na Seção 5.2 com intuito de permitir um melhor estudo da busca de fatores de transcrição em células eucariotas, levando-se em consideração também a presença de obstáculos e suas influências nos resultados, como as taxas de associação das ligações e o tempo médio de localização dos alvos no DNA.

5.1 Entendendo o problema

Uma melhor compreensão do ácido desoxiribonucléico (DNA) é preciso entendendo melhor o papel desempenhado pelos fatores de transcrição, presentes nas células, cuja função é se ligar ao DNA e formar um complexo proteína-DNA para permitir ou impedir a transcrição do ácido. Para atingir tal objetivo, a investigação do tempo médio com que elas ocorrem se torna essencial. Além disso, também é necessário analisar os dados de como a busca é realizada, comparando dados do momento em que os FT se deslocam no citoplasma e de valores que se deslocam na cadeia de DNA. Esses são objetos centrais neste estudo, que avaliará tais dados para células eucariotas. Essa tarefa é realizada ao longo das últimas décadas abordando modelos teóricos baseados em difusão facilitada, modelo que se tornou uma chave para a compreensão da regulação genética, inclusive células em procariotas (BAUER; METZLER, 2012). Neste sentido, o problema estudado caracteriza-se como a obtenção de dados que descrevam o comportamento de determinadas proteínas das células no seu processo de busca alvo em moléculas de DNA, considerando que ocorre difusão facilitada, tipo de busca mais aceito pelos pesquisadores na atualidade (GOMEZ; KLUMPP, 2016).

A finalidade principal neste fragmento do estudo é entender melhor o funcionamento do DNA, analisando como os fatores de transcrição se ligam aos seus sítios específicos. Para isto, foram simuladas buscas considerando o processo de difusão facilitada considerando-o como uma difusão anômala, no qual, em cada intervalo de tempo, as proteínas podem dar passos de tamanhos variados, com diferentes tempos de espera.

5.2 Modelo

Com intuito de analisar como os fatores de transcrição buscam seus alvos nas moléculas de DNA, podemos utilizar modelos matemáticos, baseados em equações diferenciais parciais (EDP), que descrevem os movimentos destas proteínas em busca de seus sítios específicos, como feito por Bauer e Metzler (2012) e, em seguida, através de métodos analíticos, resolver estas equações. Uma forma de realizar esta tarefa é utilizar transformadas de Laplace para reescrever as EDP's como equações diferenciais ordinárias (EDO's), ou seja, equações nas quais funções incógnitas possuem apenas uma variável.

Alternativamente, é possível modelar as células eucariotas na presença do DNA através de modelos em redes. Em (GOMEZ; KLUMPP, 2016) utilizou-se um modelo deste tipo, no qual cada unidade poderia ser preenchida com a menor parte do DNA. Essa molécula foi colocada de forma linear com um tamanho específico, sendo adicionados diferentes tipos de obstáculos que existem nas células, como outras proteínas, por exemplo, com objetivo de simular a busca dos fatores de transcrição no seu alvo.

Como observado experimentalmente, o movimento das proteínas nas células analisadas *in vivo* com obstáculos sofre um processo de difusão anômala. Diferentemente de (BAUER; METZLER, 2012), que considerou a difusão normal, com movimento Browniano, foi adotado neste trabalho a difusão anômala.

Com auxílio deste modelo, foi simulado o processo de difusão facilitada, no qual os FT's alternam entre uma busca tridimensional na célula e busca unidimensional sobre o DNA. Quando a busca é realizada no citoplasma, a cada passo de simulação o FT pode se mover com igual probabilidade para qualquer uma de 6 direções possíveis. Na busca unidimensional, quando tais proteínas estão ligadas à molécula de DNA, o FT tem a possibilidade de executar três ações possíveis, podendo se mover em dois sentidos, permanecendo ligado à cadeia de DNA, ou voltar ao citoplasma com uma probabilidade p_{off}.

5.3 Resultados

Implementamos um código que simula a difusão facilitada de fatores de transcrição procurando por um alvo em uma cadeia de DNA em um ambiente tridimensional. Ele usa uma abordagem de caminhada aleatória em que cada proteína começa em uma posição aleatória em uma faixa definida e se move no espaço 3D com uma direção a princípio aleatória, mas com uma força de atração que existe entre ela e a cadeia de DNA. Uma vez que um fator de transcrição esteja suficientemente próximo da cadeia de DNA (em uma distância inferior a 1 unidade do eixo x), ele se move ao longo da fita em uma dimensão, mais rapidamente, porém com uma pequena probabilidade p_{off} de dissociar e se mover no espaço 3D novamente. O código também verifica se uma proteína alcançou o alvo na cadeia de DNA e mantém o registro do número de iterações e proteínas que alcançaram o alvo e no final calcula a distância média de todas as proteínas ao alvo. Nas Figuras 22 são mostrados diferentes momentos da simulação para um caso onde foram consideradas 1000 proteínas na busca. Conforme aumenta o número de iterações, uma quantidade maior de FTs atinge o alvo ba cadeia de DNA.

Foram utilizadas na implementação as bibliotecas Numpy, para manipular arrays multidimensionais, e Matplotlib para criar gráficos 3D.

Para mostrar que a busca é superdifusiva, foi calculada, em cada iteração, a distância média entre as proteínas e o ponto de busca no DNA ao longo do tempo. A busca é considerada superdifusiva quando a distância média diminui mais rapidamente do que o esperado pela difusão normal. O resultado é apresentado na Figura 23.

Concluímos notando que a busca de fatores de transcrição ao alvo realizada pelo processo de difusão facilitada, no qual FT's fazem uma busca em 3D na massa até atingir a

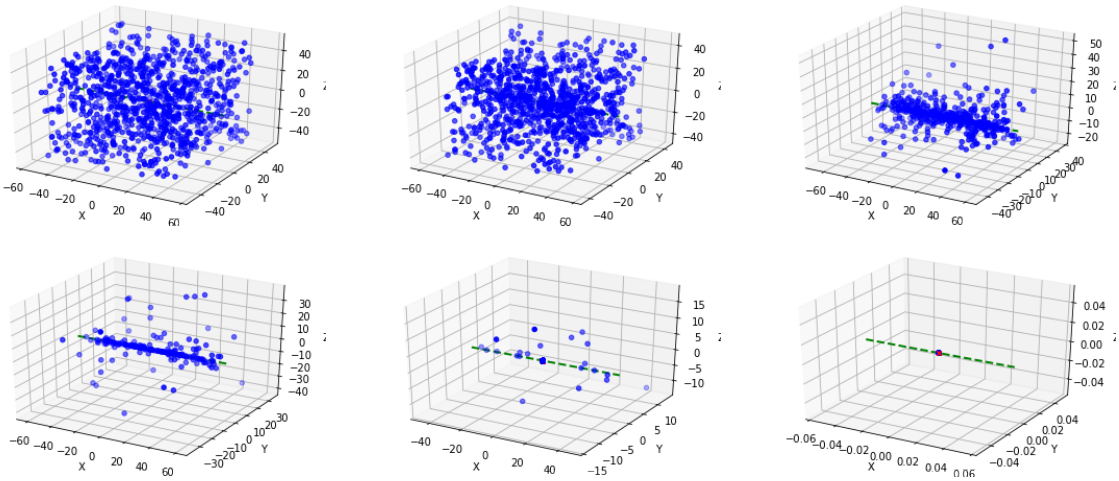


Figura 22 – Gráficos que ilustram o comportamento de fatores de transcrição na busca de um alvo em uma cadeia de DNA em diferentes etapas da simulação em um caso com 1000 partículas. A fita de DNA é mostrada em uma linha tracejada e está sobre o eixo x.

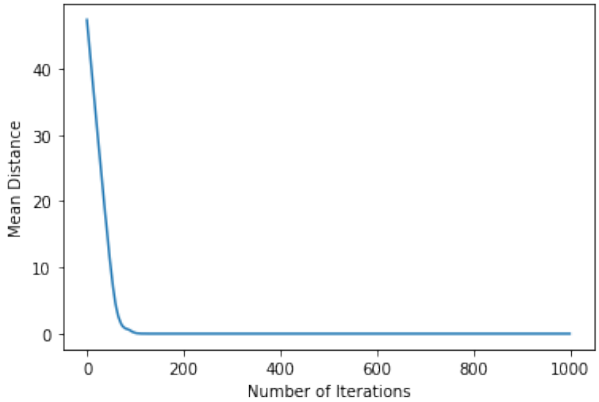


Figura 23 – Gráfico que mostra a distância média das partículas ao alvo, no final de cada iteração para um caso em que foram consideradas 1000 partículas.

cadeia de DNA, é, empiricamente, mais eficiente que a difusão normal. Quando isto ocorre, a busca é feita de modo unidimensional com uma constante de difusão diferente. Preso à cadeia, o FT pode retornar à massa ou se movimentar em busca do seu alvo no DNA.

Capítulo 6

Conclusão

Neste trabalho, foi desenvolvido um estudo da difusão anômala através de abordagens distintas, com um modelo teórico e modelos computacionais, estabelecendo uma conexão entre aquelas, com intuito de entender fenômenos gerais nos quais não podem ser observadas as leis usuais de difusão normal.

Sob um ponto de vista diverso ao já adotado na bibliografia, foram levantados resultados que ajudam a descrever este importante fenômeno físico presente em inúmeras situações. Notamos que existe uma grande sensibilidade nos parâmetros do modelo nos quais pequenas alterações produzem diferentes resultados no sistema.

Foram considerados diferentes métodos de ajuste, incluindo o método de L-BFGS, modificação do clássico método do gradiente, mas devido à restrição na solução do PVC houve uma opção pela utilização de um algoritmo genético, um método de otimização, importante classe dos algoritmos evolucionários.

Finalmente, foi realizado um estudo da busca de fatores de transcrição por alvos no DNA e feita uma conexão com a teoria da superdifusão, pois mostramos por simulações computacionais que o método de busca mais aceito na bibliografia, chamado de difusão facilitada, é compatível com o fenômeno de superdifusão, estudado anteriormente no trabalho.

Acreditamos, ainda, que o trabalho trouxe uma contribuição para o estudo que envolvem tratamento de dados, por descrever possibilidades de ajustes que podem ser aplicadas em diferentes áreas do conhecimento e não apenas nos problemas apresentados neste texto.

6.1 Trabalhos Futuros

Como sugestão de continuidade deste trabalho, ainda com a temática da difusão anômala, destacamos o estudo de condições nas quais a interação entre os entes do

fenômeno é considerada. Para esse fim, poderão ser utilizadas simulações computacionais e um modelo teórico composto por equações diferenciais fracionárias não-lineares.

Um outro próximo passo para este estudo é a análise de casos mais gerais das EDPs, incluindo outros termos como os que podem provocar o efeito de “drift”.

Além disso, os recursos computacionais desenvolvidos no nosso trabalho podem ser utilizados para estudos em áreas diversas das utilizadas por nós, como sistemas financeiros, por exemplo.

Vale ressaltar que estamos trabalhando na otimização do algoritmo genético utilizado com intuito de obter maior precisão nos ajustes.

Referências

AFEK, A. et al. Protein-DNA binding in the absence of specific base-pair recognition. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 111, n. 48, p. 17140–17145, 2014. Citado na página 72.

BAK, P.; BRUINSMA, R. One-dimensional ising model and the complete devil's staircase. **Physical Review Letters**, APS, v. 49, n. 4, p. 249, 1982. Citado na página 39.

BALESCU, R. Anomalous transport in turbulent plasmas and continuous time random walks. **Physical Review E**, APS, v. 51, n. 5, p. 4807, 1995. Citado na página 2.

BAUER, M.; METZLER, R. Generalized facilitated diffusion model for DNA-binding proteins with search and recognition states. **Biophysical journal**, Elsevier, v. 102, n. 10, p. 2321–2330, 2012. Citado 4 vezes nas páginas 71, 72, 73 e 74.

BAUER, M.; METZLER, R. In vivo facilitated diffusion model. **PLoS One**, Public Library of Science, v. 8, n. 1, p. e53956, 2013. Citado na página 72.

BENHAMOU, M. Lecture on the anomalous diffusion in condensed matter physics. **Materials and Devices**, v. 3, n. 2, 2018. Citado na página 2.

BENHAMOU, S. How many animals really do the Lévy walk? **Ecology**, Wiley Online Library, v. 88, n. 8, p. 1962–1969, 2007. Citado na página 6.

BERG, O. G.; BLOMBERG, C. Association kinetics with coupled diffusion: Iii. ionic-strength dependence of the lac repressor-operator association. **Biophysical chemistry**, Elsevier, v. 8, n. 4, p. 271–280, 1978. Citado na página 71.

BOLOGNA, M.; TSALLIS, C.; GRIGOLINI, P. Anomalous diffusion associated with nonlinear fractional derivative fokker-planck-like equation: exact time-dependent solutions. **Physical Review E**, v. 62, p. 2213, 2000. Citado na página 3.

BOUCHAUD, J.-P.; GEORGES, A. Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications. **Physics reports**, Elsevier, v. 195, n. 4-5, p. 127–293, 1990. Citado na página 39.

BOYCE, W. E.; DIPRIMA, R. C. **Equações diferenciais elementares e problemas de valores de contorno**. [S.l.]: Guanabara Dois, 1985. Citado na página 21.

CAMARGO, R. F.; CAPELAS, E. O. **Cálculo Fracionário**. 1a. ed.. ed. [S.l.]: Editora livraria da física, 2015. Citado na página 13.

CHONG, E. K.; ZAK, S. H. **An introduction to optimization**. [S.l.]: John Wiley & Sons, 2013. v. 75. Citado na página 48.

CLAUSET, A.; SHALIZI, C. R.; NEWMAN, M. E. J. Power-law distributions in empirical data. **SIAM review**, SIAM, v. 51, n. 4, p. 661–703, 2009. Citado na página 4.

COMPTE, A. Continuous time random walks on moving fluids. **Physical Review E**, APS, v. 55, n. 6, p. 6821, 1997. Citado na página 34.

CRISTELLI, M.; BATTY, M.; PIETRONERO, L. There is more than a power law in zipf. **Scientific reports**, Springer, v. 2, n. 1, p. 1–7, 2012. Citado na página 3.

DAVIS, P. J. Gamma function and related functions. **Handbook of mathematical functions with formulas, graphs, and mathematical tables**, Dover New York, p. 253–293, 1972. Citado na página 13.

DOSS, K.; MAURO, J. C. Theory of structural relaxation in glass from the thermodynamics of irreversible processes. **Physical Review E**, APS, v. 103, n. 6, p. 062606, 2021. Citado 2 vezes nas páginas 45 e 47.

EDWARDS, A. M. et al. Incorrect likelihood methods were used to infer scaling laws of marine predator search behaviour. **PLoS ONE**, Public Library of Science San Francisco, USA, 2012. Citado na página 4.

EINSTEIN, A. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. **Annalen der physik**, v. 4, 1905. Citado na página 1.

EINSTEIN, A. **Investigations on the Theory of the Brownian Movement**. [S.l.]: Courier Corporation, 1956. Citado na página 2.

FLETCHER, R. **Practical methods of optimization**. [S.l.]: John Wiley & Sons, 2013. Citado na página 47.

FORD, B. J. Brownian movement in clarkia pollen: a reprise of the first observations. **MICROSCOPE-LONDON THEN CHICAGO-**, MCCRONE RESEARCH INSTITUTE, v. 40, p. 235–235, 1992. Citado na página 1.

FRANK-KAMENETSKII, M. D. Biophysics of the DNA molecule. **Physics Reports**, Elsevier, v. 288, n. 1-6, p. 13–60, 1997. Citado na página 71.

GHAEMI, M.; ZABIHINPOUR, Z.; ASGARI, Y. Computer simulation study of the lévy flight process. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 388, n. 8, p. 1509–1514, 2009. Citado na página 2.

GOLDSTEIN, M. L.; MORRIS, S. A.; YEN, G. G. Problems with fitting to the power-law distribution. **The European Physical Journal B-Condensed Matter and Complex Systems**, Springer, v. 41, p. 255–258, 2004. Citado na página 4.

GOMEZ, D.; KLUMPP, S. Facilitated diffusion in the presence of obstacles on the DNA. **Physical Chemistry Chemical Physics**, Royal Society of Chemistry, v. 18, n. 16, p. 11184–11192, 2016. Citado 2 vezes nas páginas 72 e 73.

HERRMANN, R. **Fractional calculus: an introduction for physicists**. [S.l.]: World Scientific, 2011. Citado na página 13.

HILL, N. A.; HÄDER, D.-P. A biased random walk model for the trajectories of swimming micro-organisms. **Journal of Theoretical Biology**, Elsevier, v. 186, n. 4, p. 503–526, 1997. Citado na página 2.

HU, T.; GROSBERG, A. Y.; SHKLOVSKII, B. How proteins search for their specific sites on DNA: the role of DNA conformation. **Biophysical journal**, Elsevier, v. 90, n. 8, p. 2731–2744, 2006. Citado na página 71.

KATZNELSON, Y. **An introduction to harmonic analysis**. [S.l.]: Cambridge University Press, 2004. Citado na página 21.

KIMMICH, R. **NMR: Tomography, diffusometry, relaxometry**. [S.l.]: Springer Science & Business Media, 2012. Citado na página 2.

KLAFTER, J.; SHLESINGER, M. F.; ZUMOFEN, G. Beyond brownian motion. **Physics today**, [New York, American Institute of Physics], v. 49, n. 2, p. 33–39, 1996. Citado na página 38.

KLAFTER, J.; SOKOLOV, I. M. Anomalous diffusion spreads its wings. **Physics world**, IOP Publishing, v. 18, n. 8, p. 29, 2005. Citado na página 2.

KOEN, C.; KONDLO, L. Fitting power-law distributions to data with measurement errors. **Monthly Notices of the Royal Astronomical Society**, The Royal Astronomical Society, v. 397, n. 1, p. 495–505, 2009. Citado na página 4.

KRAMER, O. Genetic algorithms. In: **Genetic algorithm essentials**. [S.l.]: Springer, 2017. p. 11–19. Citado na página 48.

KRAMERS, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. **Physica**, Elsevier, v. 7, n. 4, p. 284–304, 1940. Citado na página 2.

KRAPIVSKY, P. L.; REDNER, S.; LEYVRAZ, F. Connectivity of growing random networks. **Physical review letters**, APS, v. 85, n. 21, p. 4629, 2000. Citado na página 6.

KUMAR, A. et al. Anomalous diffusion along metal/ceramic interfaces. **Nature communications**, Nature Publishing Group, v. 9, n. 1, p. 5251, 2018. Citado na página 2.

LABOLLE, E. M.; FOGG, G. E.; TOMPSON, A. F. Random-walk simulation of transport in heterogeneous porous media: Local mass-conservation problem and implementation methods. **Water Resources Research**, Wiley Online Library, v. 32, n. 3, p. 583–593, 1996. Citado na página 8.

LENZI, E. K. et al. Solutions for diffusion equation with a nonlocal term/solucoes para a equacao de difusao com um termo nao-local. **Acta Scientiarum. Technology**, Universidade Estadual de Maringa, v. 31, n. 1, p. 81–87, 2009. Citado na página 20.

LIU, D. C.; NOCEDAL, J. On the limited memory bfgs method for large scale optimization. **Mathematical programming**, Springer, v. 45, n. 1-3, p. 503–528, 1989. Citado na página 42.

LUTZ, E. Anomalous diffusion and tsallis statistics in an optical lattice. **Physical Review A**, APS, v. 67, n. 5, p. 051402, 2003. Citado 2 vezes nas páginas 45 e 47.

MAINARDI, F. **Fractional Calculus and Waves in Linear Viscoelasticity: An Introduction to Mathematical Models**. 2. ed. [S.l.]: Imperial College Press, 2010. Citado na página 49.

MARKSTEINER, S.; ELLINGER, K.; ZOLLER, P. Anomalous diffusion and lévy walks in optical lattices. **Physical Review A**, APS, v. 53, n. 5, p. 3409, 1996. Citado na página 12.

MATHELIER, A. et al. DNA shape features improve transcription factor binding site predictions in vivo. **Cell systems**, Elsevier, v. 3, n. 3, p. 278–286, 2016. Citado na página 72.

MEERSCHAERT, M. M. Fractional calculus, anomalous diffusion, and probability. **Fractional dynamics: recent advances**, World Scientific, p. 265–284, 2012. Citado na página 12.

METZLER, R.; KLAFTER, J. The random walk's guide to anomalous diffusion: A fractional dynamics approach. **Physics Reports**, v. 339, p. 1–77, 2000. Citado 4 vezes nas páginas 2, 3, 28 e 29.

METZLER, R.; KLAFTER, J. The random walk's guide to anomalous diffusion: a fractional dynamics approach. **Physics reports**, Elsevier, v. 339, n. 1, p. 1–77, 2000. Citado 2 vezes nas páginas 3 e 8.

MIRNY, L. et al. How a protein searches for its site on dna: the mechanism of facilitated diffusion. **Journal of Physics A: Mathematical and Theoretical**, IOP Publishing, v. 42, n. 43, p. 434013, 2009. Citado na página 72.

MONTROLL, E. W.; WEISS, G. H. Random walks on lattices. ii. **Journal of Mathematical Physics**, AIP, v. 6, n. 2, p. 167–181, 1965. Citado 2 vezes nas páginas 3 e 31.

NEWELL, L. C. Robert brown and the discovery of the brownian movement. **Industrial & Engineering Chemistry**, ACS Publications, v. 15, n. 12, p. 1279–1281, 1923. Citado na página 1.

NEWMAN, M. E. Power laws, pareto distributions and zipf's law. **Contemporary physics**, Taylor & Francis, v. 46, n. 5, p. 323–351, 2005. Citado na página 4.

NUSSENZVEIG, H. M. **Curso de física básica: Mecânica (vol. 1)**. [S.l.]: Editora Blucher, 2013. v. 394. Citado na página 26.

OLIVEIRA, D. d. S. d. et al. Derivada fracionária e as funções de mittag-leffler. [sn], 2014. Citado na página 21.

ORDÓÑEZ, H. J. L. **Interacción de ligandos orgánicos y complejos metálicos con ácidos nucleicos**. Tese (Doutorado) — Universidad de Burgos, 2016. Citado na página 70.

PACHECO, M. A. C. et al. Algoritmos genéticos: princípios e aplicações. **ICA: Laboratório de Inteligência Computacional Aplicada. Departamento de Engenharia Elétrica. Pontifícia Universidade Católica do Rio de Janeiro. Fonte desconhecida**, v. 28, 1999. Citado na página 42.

PARADISI, P. et al. The fractional fick's law for non-local transport processes. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 293, n. 1-2, p. 130–142, 2001. Citado na página 26.

PEINKE, J.; BÖTTCHER, F.; BARTH, S. Anomalous statistics in turbulence, financial markets and other complex systems. **Annalen der Physik**, Wiley Online Library, v. 13, n. 7-8, p. 450–460, 2004. Citado na página 8.

PEREIRA, A. P. et al. Parameter calibration between models and simulations: Connecting linear and non-linear descriptions of anomalous diffusion. **Physica A: Statistical Mechanics and its Applications**, Elsevier, 2018. Citado 2 vezes nas páginas 3 e 40.

PEREIRA, A. P. P. **Calibração de Parâmetros entre as Escalas de Voos Difusivos Anômalos: Prescrição para Corresponder Simulação e Modelos**. Tese (Doutorado) — Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, 2018. Citado 7 vezes nas páginas 16, 23, 38, 50, 51, 52 e 55.

PEREIRA, A. P. P. et al. Simulation and calibration between parameters of continuous time random walks and subdifusive model. **TEMA**, Scielo, v. 18, p. 305 – 315, 2017. Citado na página 39.

PIANTADOSI, S. T. Zipf's word frequency law in natural language: A critical review and future directions. **Psychonomic bulletin & review**, Springer, v. 21, p. 1112–1130, 2014. Citado na página 6.

PIERRO, M. D. et al. Anomalous diffusion, spatial coherence, and viscoelasticity from the energy landscape of human chromosomes. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 115, n. 30, p. 7753–7758, 2018. Citado na página 2.

PODLUBNY, I. **Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications**. [S.l.]: Elsevier, 1998. v. 198. Citado 6 vezes nas páginas 13, 16, 19, 20, 22 e 50.

PUTERMAN, M. L. Markov decision processes. **Handbooks in operations research and management science**, Elsevier, v. 2, p. 331–434, 1990. Citado na página 8.

ROBERT, C. P.; CASELLA, G.; CASELLA, G. **Monte Carlo statistical methods**. [S.l.]: Springer, 1999. v. 2. Citado na página 34.

ROSA, M. A. D. de la et al. Dynamic strategies for target-site search by dna-binding proteins. **Biophysical journal**, Elsevier, v. 98, n. 12, p. 2943–2953, 2010. Citado na página 70.

SABRI, A. et al. Elucidating the origin of heterogeneous anomalous diffusion in the cytoplasm of mammalian cells. **arXiv preprint arXiv:1910.00102**, 2019. Citado na página 2.

SALGADO, G. H. O. **Métodos numéricos para solução de equações diferenciais segundo a derivada de Caputo**. Tese (Doutorado) — Universidade Federal de Minas Gerais, Departamento de Engenharia Elétrica, Curso de Pós-Graduação em Engenharia Elétrica, Belo Horizonte, 2015. Citado na página 23.

SAMKO, S. G. Fractional weyl-riesz integrodifferentiation of periodic functions of two variables via the periodization of the riesz kernel. **Applicable Analysis**, Taylor & Francis, v. 82, n. 3, p. 269–299, 2003. Citado na página 20.

SAMKO, S. G. et al. **Fractional integrals and derivatives**. [S.l.]: Gordon and breach science publishers, Yverdon Yverdon-les-Bains, Switzerland, 1993. v. 1. Citado na página 45.

SATO, Y.; KLAGES, R. Anomalous diffusion in random dynamical systems. **Physical review letters**, APS, v. 122, n. 17, p. 174101, 2019. Citado na página 1.

SAXTON, M. J. Diffusion of dna-binding species in thenucleus: A transient anomalous subdiffusion model. **Biophysical Journal**, Elsevier, v. 112, n. 3, p. 149a, 2017. Citado na página 70.

SCHÜTZ, G. J.; SCHINDLER, H.; SCHMIDT, T. Single-molecule microscopy on model membranes reveals anomalous diffusion. **Biophysical journal**, Elsevier, v. 73, n. 2, p. 1073–1080, 1997. Citado na página 2.

SEDGWICK, P. Pearson's correlation coefficient. **Bmj**, British Medical Journal Publishing Group, v. 345, 2012. Citado na página 59.

SERENO, L. O. R.; ACEBAL, J. L. Busca alvo de fatores de transcrição em cadeias de adn. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, v. 6, n. 1, 2018. Citado na página 71.

SHIMAZAKI, H.; SHINOMOTO, S. A method for selecting the bin size of a time histogram. **Neural computation**, MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , v. 19, n. 6, p. 1503–1527, 2007. Citado na página 58.

SHLESINGER, M. F.; ZASLAVSKY, G. M.; FRISCH, U. Lévy flights and related topics in physics. **Lecture notes in physics**, Springer, v. 450, p. 52, 1995. Citado na página 2.

SLUTSKY, M.; MIRNY, L. A. Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. **Biophysical journal**, Elsevier, v. 87, n. 6, p. 4021–4035, 2004. Citado na página 71.

SOKOLOV, I. M. Models of anomalous diffusion in crowded environments. **Soft Matter**, Royal Society of Chemistry, v. 8, n. 35, p. 9043–9052, 2012. Citado na página 40.

SOKOLOV, I. M.; KLAFTER, J. From diffusion to anomalous diffusion: a century after einstein's brownian motion. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, American Institute of Physics, v. 15, n. 2, p. 026103, 2005. Citado na página 2.

STEPHENSON, J. Some non-linear diffusion equations and fractal diffusion. **Physica A**, v. 222, p. 234–247, 1995. Citado na página 2.

STEWART, J. **Cálculo volume 2**. [S.l.]: 2013, 2013. Citado na página 10.

SUMIN, V.; DUSHKIN, A.; SMOLENTSEVA, T. Mathematical models to determine stable behavior of complex systems. In: IOP PUBLISHING. **Journal of Physics: Conference Series**. [S.l.], 2018. v. 1015, n. 3, p. 032136. Citado na página 7.

TSALLIS, C. Nonextensive statistical mechanics, anomalous diffusion and central limit theorems. **Milan Journal of Mathematics**, Springer, v. 73, n. 1, p. 145–176, 2005. Citado na página 2.

TSALLIS, C.; BUKMAN, D. J. Anomalous diffusion in the presence of external forces: Exact time-dependent solutions and their thermostatistical basis. **Phys. Rev. E**, v. 54, p. R2197, 1996. Citado na página 3.

VISSER, M. Zipf's law, power laws and maximum entropy. **New Journal of Physics**, IOP Publishing, v. 15, n. 4, p. 043021, 2013. Citado na página 4.

VLAHOS, L. et al. Normal and anomalous diffusion: A tutorial. **arXiv preprint arXiv:0805.0419**, 2008. Citado na página 28.

WALPOLE, R. E. et al. **Probability and statistics for engineers and scientists**. [S.l.]: Macmillan New York, 1993. v. 5. Citado na página 2.

WHITE, E. P.; ENQUIST, B. J.; GREEN, J. L. On estimating the exponent of power-law frequency distributions. **Ecology**, Wiley Online Library, v. 89, n. 4, p. 905–912, 2008. Citado na página 58.

WILSON, G. S. **Anomalous diffusion and self-propulsion of radioactive colloidal particles**. Tese (Doutorado) — Kansas State University, 2019. Citado na página 5.

YANG, X.-J. et al. General fractional-order anomalous diffusion with non-singular power-law kernel. Thermal Science, 2017. Citado na página 3.

YANG, Y.; ZHANG, H. H. Fractional calculus with its applications in engineering and technology. **Synthesis Lectures on Mechanical Engineering**, Morgan & Claypool Publishers, v. 3, n. 1, p. 1–107, 2019. Citado na página 14.

ZABURDAEV, V.; DENISOV, S.; KLAFTER, J. Lévy walks. **Reviews of Modern Physics**, APS, v. 87, n. 2, p. 483, 2015. Citado na página 2.

ZAID, I. M.; LOMHOLT, M. A.; METZLER, R. How subdiffusion changes the kinetics of binding to a surface. **Biophysical journal**, Elsevier, v. 97, n. 3, p. 710–721, 2009. Citado na página 70.