



UNIVERSIDADE DE AVEIRO

MODELAÇÃO DE SISTEMAS COMPLEXOS (40780)

*Two-Dimensional Random Walks
First Passage Time and Survival
Probability
Lévy Flights*

Gonçalo Freitas
NºMEC: 98012

Mestrado Em Engenharia Computacional

Docente:
Rui Américo Ferreira da Costa

4 de maio de 2023

Introdução

Two-dimensional Random Walks referem-se a um tipo de modelo matemático utilizado para descrever o movimento aleatório de uma partícula num espaço bidimensional. O caminho percorrido pela partícula é determinado pela sequência de movimentos tomados em cada passo, sendo estes independentes uns dos outros. Em cada passo a partícula pode mover-se numa das 4 direções possíveis (cima, baixo, direita, esquerda) e cada uma destas tem igual probabilidade de acontecer. [1]

First Passage Time (FPT) é um conceito que surge no estudo de *random walks* e refere-se à distribuição de probabilidade do 'tempo' necessário para que uma partícula, sujeita a uma *random walk*, alcance um ponto específico pela primeira vez.

A *Survival Probability* é outro conceito importante no estudo de *random walks* e refere-se à probabilidade de uma partícula, sujeita a um passeio aleatório, ainda não tenha alcançado um ponto específico num determinado momento. Esta probabilidade está intimamente relacionada com a distribuição de probabilidade do FPT e é frequentemente utilizada para estudar o comportamento de *random walks* em longas escalas de tempo. [2]

Os *Lévy flights* são um tipo de *random walks* em que os passos dados pela partícula têm uma distribuição de probabilidade que segue uma lei de potência. Ao contrário do que acontece nas *random walks* convencionais, onde a distância percorrida pela partícula em cada passo é fixa, num *Lévy flight*, os tamanhos dos passos podem variar aleatoriamente. Este tipo de *random walk* é geralmente utilizada para modelar o comportamento de animais na busca por comida ou o movimento de partículas em fluidos turbulentos. [3]

As tarefas propostas para este projeto foram todas resolvidas em `Matlab` e os respetivos códigos foram devidamente anexados. De forma a facilitar, o código está dividido em secções e, para se obter os gráficos aqui apresentados, basta correr as secções indexadas como `Plots` que fazem uso dos ficheiros `.mat`, também anexados, sem assim ter que correr o código para obter os mesmos.

Parte 1 - *Two-Dimensional Random Walks*

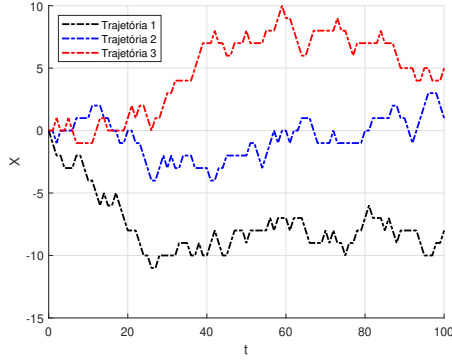
Nesta tarefa serão simuladas *random walks* simétricas numa rede quadrada regular. A cada passo, a partícula pode saltar para a esquerda, para a direita, para cima ou para baixo e cada um destes saltos tem probabilidade igual a $1/4$. Considerou-se que, em cada passo, o tamanho do salto é fixo e toma o valor de 1 unidade.

Task 1.1

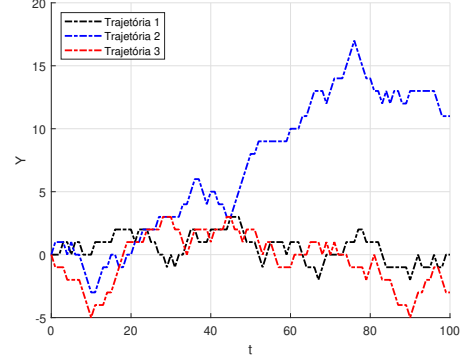
Nesta tarefa o objetivo é realizar uma representação gráfica de 3 *random walks* ao longo de 100 passos, assegurando a independência entre cada passo e entre cada partícula.

Começou-se por definir a posição inicial de cada partícula que, neste caso, foi igual para todas, $(x_0, y_0) = (0, 0)$. De seguida, escolheu-se um número aleatório entre 1 e 4, sendo este valor aquele que define a direção do movimento da partícula no presente passo. Neste caso optou-se que, caso o valor fosse 1 a partícula moveria-se para cima, 2 moveria-se para a direita, 3 para a baixo e, por fim, caso o valor fosse 4 a partícula moveria-se para a esquerda. Repetiu-se então este processo 100 vezes, de forma a simular os 100 passos pedidos no enunciado, para cada uma das 3 partículas.

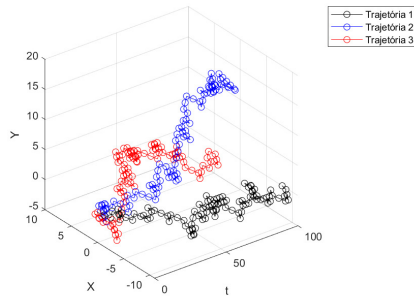
Fixou-se também a *seed* do *Matlab*, [4], com o valor 0, `rng(0)`, de forma a se obter sempre o mesmo resultado. Posto isto, foi possível obter os gráficos apresentados na Figura 1 onde estão traçados os trajetos de cada partícula, em diferentes planos.



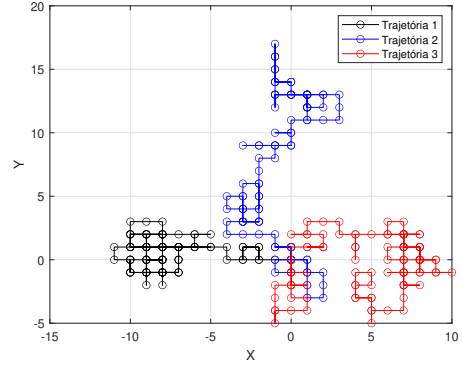
(a) Plano $Y = 0$



(b) Plano $X = 0$



(c) Plano X-Y-t



(d) Plano X-Y

Figura 1: Representações gráficas das 3 *random walks* obtidas

Task 1.2

O objetivo desta tarefa é calcular e representar graficamente a probabilidade experimental, $\hat{P}(x, y; t)$, de encontrar uma partícula num determinado ponto (x, y) após um certo número de passos (*jumps*) t , comparando os resultados obtidos com a distribuição de probabilidade teórica dada pela equação:

$$P(x, y; t) = \frac{1}{\pi t} e^{-\frac{x^2 + y^2}{t}} \quad (1)$$

Além disso, é pedido para se verificar a normalização das probabilidades, as médias ($\langle x \rangle = \langle y \rangle = 0$) e a variância da distância ao ponto inicial da partícula ($\langle x^2 + y^2 \rangle = t$).

Tal como o enunciado refere, para calcular $\hat{P}(x, y; t)$, é preciso calcular esta probabilidade para um número de *jumps* pares, $P_e(x, y; t)$, e ímpares, $P_o(x, y; t)$, e calcular a média entre estes valores, ou seja:

$$\hat{P}(x, y; t) = \frac{1}{2}(P_e(x, y; t) + P_o(x, y; t)) \quad (2)$$

Com cada uma destas probabilidades calculadas como

$$P_i(x, y; t) = \frac{N_i(x, y; t)}{N}; i = e, o \quad (3)$$

Onde $N_i(x, y; t)$ corresponde ao número de vezes que o *random walk* terminou no ponto (x, y) após t *jumps*. Já N corresponde ao número de vezes que *random walk* é simulado.

Ao contrário do que foi feito na **Task 1.1** aqui apenas simulamos uma partícula ao longo de N passos. Quanto aos valores do número de *jumps* pares e ímpares, considerou-se $t_e = 50000$ e $t_o = 50001$, respetivamente.

Pela Lei dos Grandes Números, [5], sabe-se que, quanto maior for o valor de N , mais perto da solução real/teórica serão os resultados. Variando então este valor como $[10, 10^2, 10^3, 10^4, 10^5, 10^6]$, é possível obter a Tabela 1, onde vemos, tal como esperado, que à medida que o valor de N cresce, mais próximo dos valores esperados são os resultados. No entanto, quanto maior for este valor de N maior será a complexidade computacional da simulação, como se pode ver pela coluna "Tempo (s)". Também é possível reparar que os valores médios de x de y são sempre iguais, para qualquer valor de N , como seria de esperar.

N	Norma			$\langle x \rangle$			$\langle y \rangle$			$\langle x^2 + y^2 \rangle$			Tempo (s)
	Obtido	Esperado	Desvio	Obtido	Esperado	Desvio	Obtido	Esperado	Desvio	Obtido	Esperado	Desvio	
10	1	1	0	-53.95	0	53.95	-53.95	0	53.95	56246.7	50000.5	6246.2	0.372350
100	1	1	0	-8.23	0	8.23	-8.23	0	8.23	50802.12	50000.5	801.62	0.224519
10^3	1	1	0	3.911	0	3.911	3.911	0	3.911	50518.849	50000.5	518.349	1.325468
10^4	1	1	0	1.0864	0	1.0864	1.0864	0	1.0864	50371.8098	50000.5	371.3098	13.378046
10^5	1	1	0	-0.13861	0	0.13861	-0.13861	0	0.13861	50119.2898	50000.5	118.7898	139.014623
10^6	1	1	0	-0.025646	0	0.025646	-0.025646	0	0.025646	49969.9796	50000.5	30.5204	1240.636604

Tabela 1: Resultados obtidos, valor esperado e respetivo desvio para cada propriedade pedida.

De forma a obter os valores experimentais apresentados na Tabela 1, usou-se as seguintes equações:

$$\text{Normalização} = \sum_{x,y} \hat{P}(x, y; t) \quad (4)$$

$$\langle x \rangle = \sum_{x,y} x \hat{P}(x, y; t) \quad (5)$$

$$\langle y \rangle = \sum_{x,y} y \hat{P}(x, y; t) \quad (6)$$

$$\langle x^2 + y^2 \rangle = \sum_{x,y} (x^2 + y^2) \hat{P}(x, y; t) \quad (7)$$

Usando os resultados para quando $N = 10^6$, obtemos os seguintes gráficos onde podemos reparar que, qualitativamente, os resultados experimentais vão de acordo com os valores teóricos, apresentando um comportamento semelhante, com maior probabilidade nos pontos mais próximos do ponto inicial e a diminuição desta à medida que se vai

afastando deste ponto. No entanto, se analisarmos os resultados quantitativamente vemos que os valores experimentais são, aproximadamente, 10 vezes maiores que os valores teóricos. Isto deve-se ao facto de que os valores teóricos percorrem todo o espaço de fase, ao contrário dos valores experimentais, o que faz que, como a soma das probabilidades tem que ser igual a 1, segundo a normalização, haja menos valores na zona do pico.

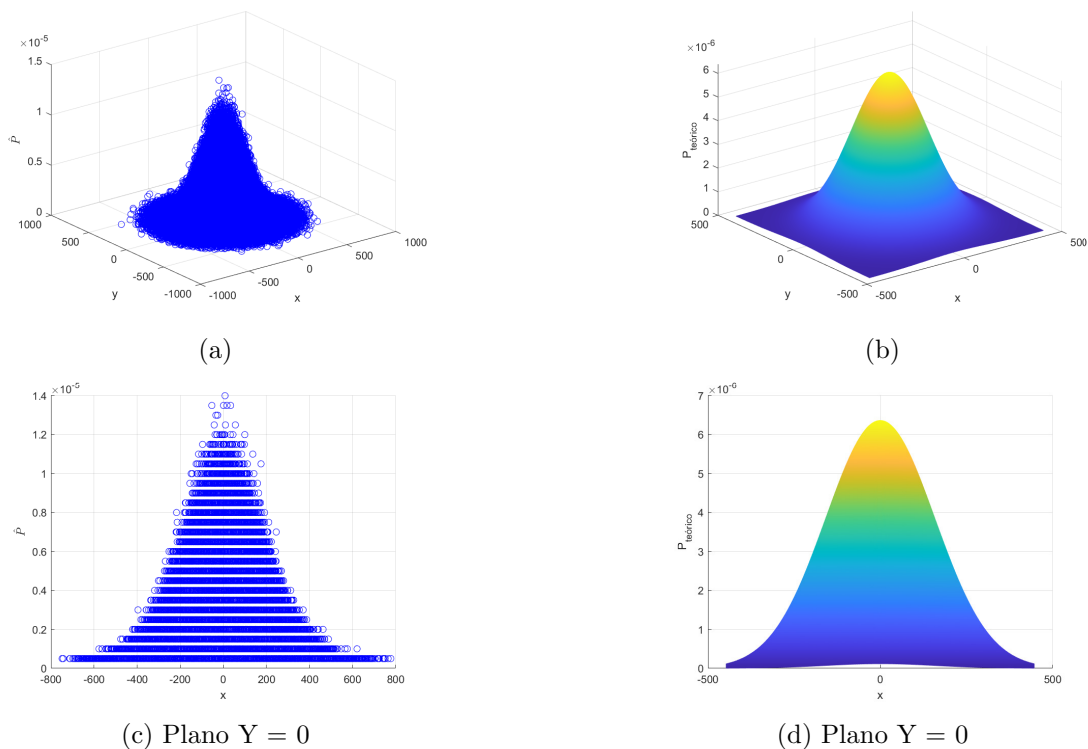
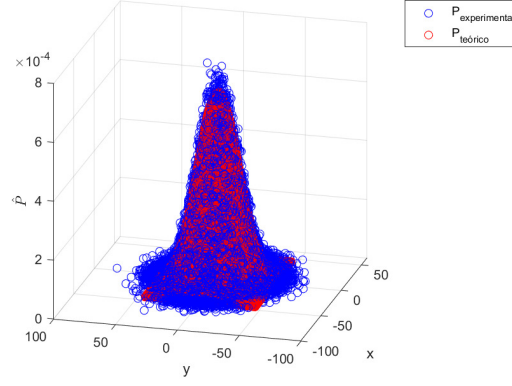


Figura 2: Representação da probabilidade experimental ((a) & (c)) e da probabilidade teórica ((b) & (d)), em diferentes planos

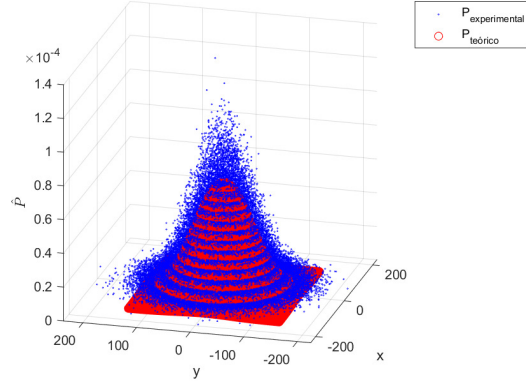
De referir que, para a representação gráfica dos valores teóricos, considerou-se apenas o espaço $(X, Y) \in [-2\sqrt{t}; 2\sqrt{t}]$ pois, fora deste intervalo existem poucos valores que geralmente se consideram como ruído.

Outro estudo que se pode fazer é em relação aos parâmetros t_{even} e t_{odd} , variando-os e vendo que diferenças esta alteração provoca nas probabilidades teóricas e experimentais. Desta forma, variou-se estes valores, para além dos valores acima referidos, como $(t_{even}, t_{odd}) = [(500, 501), (5000, 5001)]$, obtendo-se os seguintes gráficos, representados na Figura 3. A partir desta figura, primeiramente, podemos reparar que os valores experimentais da probabilidade apresentam sempre uma forma semelhante aos valores teóricos mas, à medida que se aumenta os valores de t_{even} e t_{odd} , os valores experimentais vão se afastando, quantitativamente, dos valores teóricos. Por fim, como seria de esperar, é possível concluir que, quanto maior forem estes valores de t , maior é o espaço de fase percorrido, o que faz com que o pico seja menor provocando a diferença entre os valores

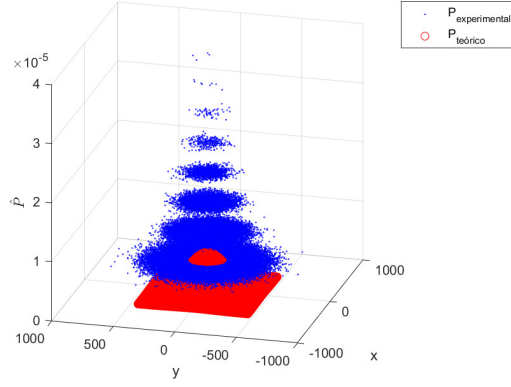
experimentais e teóricos referida anteriormente.



(a) $(t_{even}, t_{odd}) = (500, 501)$



(b) $(t_{even}, t_{odd}) = (5000, 5001)$



(c) $(t_{even}, t_{odd}) = (50000, 50001)$

Figura 3: Representação da probabilidade experimental e da probabilidade teórica, para diferentes valores de (t_{even}, t_{odd}) , assumindo $N = 10^6$

Parte 2 - *First Passage Time e Survival Probability*

Nesta tarefa partimos do código utilizado na tarefa anterior, mantendo a posição inicial da partícula em $(x_0, y_0) = (0, 0)$ mas, com a adição de uma barreira em $x_b = -30$. Caso a partícula atinja esta barreira, a *random walk* termina, caso contrário, esta continua até se realizarem os t jumps, tendo-se, neste caso, definido $t = 50000$. Tal como referido anteriormente, de forma a obter resultados satisfatórios é preciso repetir esta *random walk* várias vezes, tendo-se assim repetido um total de $N = 50000$ vezes. Nas fórmulas apresentadas nesta tarefa, D representa o coeficiente de difusão e toma, em toda a tarefa, o valor de $1/4$.

Task 2.1

Calculamos então o número de partículas que atingiram a barreira pela primeira vez, $N_{fpt}(t)$, no intervalo $[t, t + \Delta t]$, com $\Delta t = 10$ e posteriormente calculamos a *first passage time distribution*, $F(t)$, usando a seguinte fórmula:

$$F(t) = \frac{N_{fpt}(t)}{N\Delta t} \quad (8)$$

Os resultados obtidos estão representados graficamente na Figura 4, a azul, enquanto que a preto temos representados os valores teóricos (previstos), $F_{teor}(t)$, que se podem obter da seguinte forma:

$$F_{teor}(t) = \frac{|x_b - x_0|}{2\sqrt{\pi Dt^3}} \exp\left(-\frac{(x_b - x_0)^2}{4Dt}\right) \quad (9)$$

Analisando o gráfico com a escala logarítmica, Figura 4b, vemos que os desvios dos resultados experimentais para os valores teóricos parecem menores. Isto deve-se ao facto de que a função tem um comportamento exponencial, dada a proporcionalidade com $1/t^{3/2}$. O coeficiente de correlação obtido entre os resultados experimentais e os valores previstos foi de 0.9892, o que mostra excelente similaridade entre estes.

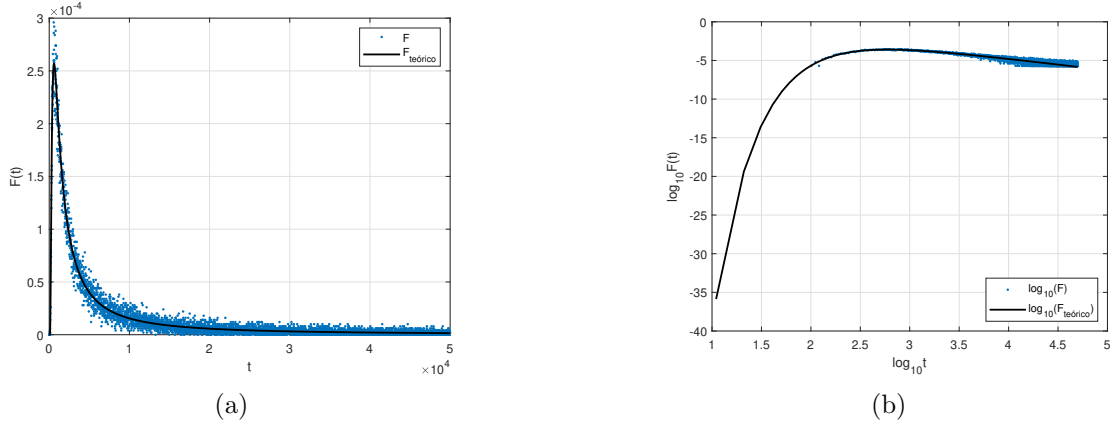


Figura 4: Distribuição do *First Passage Time* em função do número de *jumps* (a) e em escala logarítmica (b)

Quanto à *Survival Probability*, os valores experimentais, $S(t)$, estão representados a tracejado vermelho na Figura 5 e podem ser calculados usando a seguinte fórmula.

$$S(t) = \frac{N_S(t)}{N} \quad (10)$$

Onde $N_S(t)$ representa o número de partículas que 'sobreviveram', isto é, que nunca chegaram à barreira localizada em x_b , após t *jumps*.

Já os valores esperados, $S_{teor}(t)$, calculados usando a Equação 11, encontram-se representados a preto na mesma Figura.

$$S_{teor}(t) = erf\left(\frac{|x_b - x_0|}{2\sqrt{Dt}}\right) \quad (11)$$

Analisando então a Figura 5 reparamos que os valores experimentais e teóricos são bastante próximos uns dos outros, tendo-se obtido um coeficiente de correlação de 0.9999. Esta melhoria no coeficiente de correlação pode estar relacionada com a menor complexidade matemática que a fórmula de $S_{teor}(t)$ apresenta quando comparando com a fórmula de $F_{teor}(t)$.

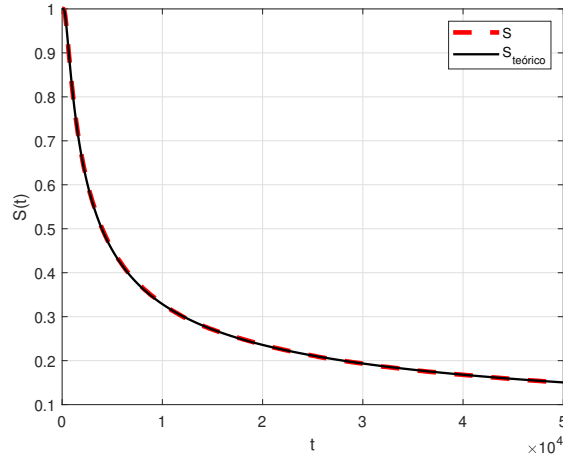


Figura 5: *Survival Probability* em função do número de *jumps*

Proponha uma explicação qualitativa da posição do máximo de $F(t)$

Intuitivamente, podemos pensar que o máximo de $F(t)$ está relacionado com distância entre a posição inicial e a barreira, ou seja, quando a barreira está longe da posição inicial, leva mais tempo (*jumps*) para o processo a atingir. Como resultado, o máximo de $F(t)$ é deslocado para a direita. Por sua vez, quando a barreira está mais perto da posição inicial, será preciso um menor número de *jumps* para que esta seja atingida, deslocando o máximo de $F(t)$ para a esquerda. Esta relação entre o máximo de $F(t)$ com a distância entre o ponto inicial e a barreira pode ser visualizada na Figura 6.

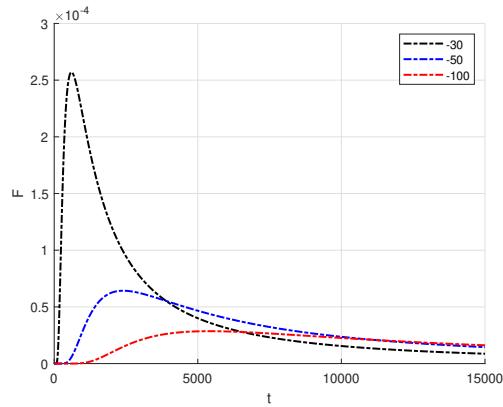


Figura 6: Análise do máximo da distribuição do *First Passage Time* previsto com base no valor de x_b , assumindo $(x_0, y_0) = (0, 0)$ e $D = 1/4$

No entanto, a posição do máximo de $F(t)$ também está relacionada com o coeficiente de difusão D . Este determina a 'rapidez' com que o processo se propaga ao longo do número de *jumps*. Se D for elevado, o processo difundir-se-á rapidamente o que faz com

que o processo demore menos tempo a atingir a barreira, deslocando assim o máximo de $F(t)$ para a esquerda. Por outro lado, se D for pequeno, o processo propagar-se-á mais lentamente, levando mais tempo para que o processo atinja a barreira e o máximo de $F(t)$ é deslocado para a direita, tal como pode ser visualizado na Figura 7.

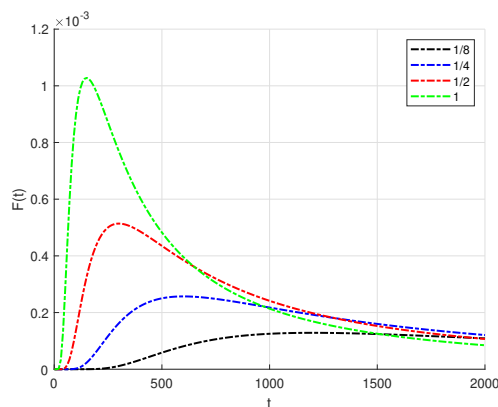


Figura 7: Análise do máximo da distribuição do *First Passage Time* previsto com base no valor de D , assumindo $x_b = -30$ e $(x_0, y_0) = (0, 0)$

Estando no ponto inicial, $(0, 0)$, e medindo o número de partículas que podem chegar a este ponto de diferentes direções. É possível propor um método para descobrir de que lado é o limite?

Podemos calcular, para um número grande de partículas ($N \rightarrow \infty$), a distribuição do *first passage time* e, por sua vez, estimar a *survival probability* ($\partial S(t)/\partial t = -F(t)$), [6].

Por fim, comparando estas probabilidades para diferentes direções podemos descobrir de que lado o limite provavelmente estará. Por exemplo, se a *survival probability* for menor à esquerda do ponto $(0, 0)$, isso sugere que a barreira está provavelmente localizada nessa direção.

Parte 3 - Lévy Flights

Tal como referido na introdução, *Lévy flights* são um tipo de *random walks* em que o tamanho dos passos dados pela partícula pode variar a cada *jump*. A probabilidade, $P(l)$, que um *jump* tenha tamanho l é determinado pela *Lévy flights distribution*:

$$P(l) = Cl^{-\mu} \quad (12)$$

Onde μ representa o índice de estabilidade desta distribuição e, C representa a constante de normalização, para $l_{min} < l < l_{max}$. Se $l_{min} = 1$, o valor de C pode ser calculado por

$$C = \frac{\mu - 1}{1 - l_{max}^{1-\mu}} \quad (13)$$

Após o n -ésimo *jump* a partícula está nas coordenadas $(x_n, y_n) = (x_{n-1}, y_{n-1}) + (l \cos(\varphi), l \sin(\varphi))$. Os saltos são isotrópicos, o que significa que a probabilidade de um salto ter um ângulo φ não depende do ângulo, ou seja, $P(\varphi) = 1/2\pi$

Task 3.1

Nesta tarefa é usado o algoritmo fornecido pela docência, variando o valor de μ como 1.6, 2 e 2.6, assumindo um número total de 1000 *jumps* e considerando, em todos os casos, a posição inicial da partícula no ponto (0,0).

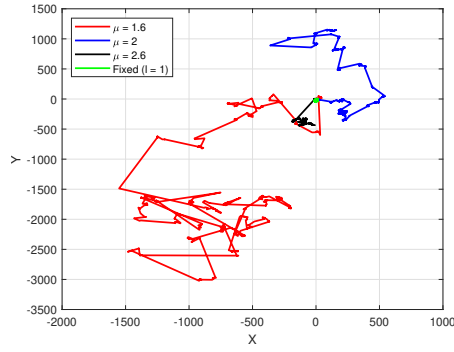
Quando o valor de μ é alterado, a *Lévy flights distribution* também é alterada. Quando μ é menor que 2, a distribuição de probabilidade apresenta uma *heavy tail*, [7], o que significa que há uma maior probabilidade de ocorrerem saltos maiores. Isto explica o porquê de quando $\mu = 1.6$, serem atingidos extremos maiores, tal como se pode ver na Figura 8a.

Por outro lado, quando μ é igual a 2, a distribuição de probabilidade de *Lévy* converge para uma distribuição normal, [8], o que significa que a probabilidade de ocorrerem saltos grandes é menor, como se pode ver na Figura 8a, onde vemos que ainda são atingidos alguns extremos grandes, sendo estes, no entanto, menores do que os observados quando $\mu = 1.6$.

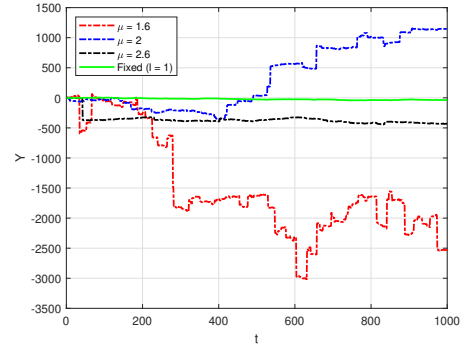
Já quando μ é maior do que 2, a distribuição de probabilidade tem uma *light tail*, [9], o que significa que há uma menor probabilidade de ocorrerem saltos maiores. Este fenómeno explica o porquê de, quando $\mu = 2.6$, haverem poucos extremos atingidos, com os saltos relativamente pequenos e mais regulares, com a trajetória da partícula a se assemelhar a uma *random walk* de salto fixo, como se pode ver nas Figuras 8a e 8b.

Em suma, o valor de μ controla o comportamento dos *Lévy Flights*, influenciando a probabilidade de ocorrerem saltos maiores ou menores, o que afeta diretamente os extremos atingidos e a aparência geral da trajetória da partícula.

No Appendix estão representadas as 4 trajetórias individualmente, no plano X-Y, onde podemos analisar quantitativamente os eixos de forma a confirmar o que foi dito anteriormente.



(a) Plano X-Y



(b) Plano X=0

Figura 8: Representação gráfica dos *Lévy flights* simulados, para diferentes valores de μ e para uma *random walk* de salto fixo

Task 3.2

Nesta tarefa é pedido que se prove, numericamente, que, dado um número x gerado uniformemente no intervalo $[0, 1]$, então os números aleatórios, calculados pela equação:

$$l(x) = [1 - x(1 - l_{max}^{1-\mu})]^{\frac{1}{1-\mu}} \quad (14)$$

são distribuídos de acordo com a *Lévy flights distribution*, dada pela Equação 12.

Para resolver este problema podemos começar por gerar N valores de l , usando a Equação 14, tendo-se assumido $N = 10^6$. De seguida calcula-se, para cada μ , o histograma de contagens de cada valor de l e sobrepõem-se a curva relativa à *Lévy flights distribution*, calculada usando as Equações 12 e 13. Assim, foi possível obter os gráficos representados na Figura 9, onde podemos concluir que os números aleatórios l , gerados a partir de um número x gerado uniformemente no intervalo $[0, 1]$, são, de facto, distribuídos de acordo com a *Lévy flights distribution*.

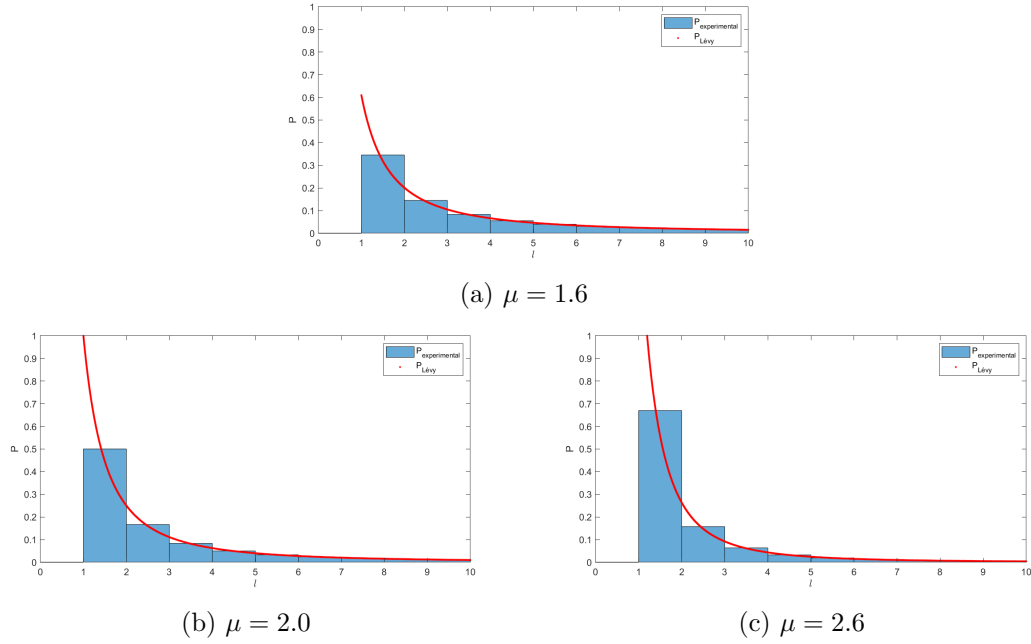


Figura 9: Probabilidade de um *jump* ter comprimento l com base no valor de μ , comparando com a *Lévy flights distribution*

Task 3.3

No caso 1, onde os peixes estão distribuídos homogeneamente no mar, significa que a probabilidade de encontrar peixes é a mesma em todo o mar, pelo que, uma estratégia baseada numa *two-dimensional random walk* (convencional) seria mais eficiente, visto permitir uma exploração sistemática de toda a área do mar.

No entanto, se os peixes estivessem distribuídos em grupos de forma esparsa e imprevisível no mar (caso 2), significa que a densidade dos peixes varia de forma imprevisível e há áreas específicas onde os peixes estão mais concentrados do que outras. Assim, neste caso, uma estratégia baseada nos *Lévy flights* seria mais adequada pois, este tipo de *random walk*, permite uma exploração mais eficiente de uma maior área do mar, o que aumenta a probabilidade de encontrar os peixes que possam estar concentrados em locais específicos.

Referências

- [1] "Introduction to Random Walks and Diffusion". Allan M. Jacobsen e Alexander P. Zijlstra. <https://arxiv.org/abs/0902.4877>
- [2] "First-passage phenomena and their applications". Sidney Redner. <https://arxiv.org/abs/cond-mat/0102171>
- [3] "Lévy Flights and Applications: Introduction". M. F. Shlesinger, G. M. Zaslavsky e U. Frisch. https://link.springer.com/chapter/10.1007/3-540-30551-0_1
- [4] MathWorks. rng - Control random number generator, (2021) <https://www.mathworks.com/help/matlab/ref/rng.html>.
- [5] "Law of large numbers", Encyclopedia of Mathematics, EMS Press, 2001 [1994]
- [6] R. A. da Costa. Slides teóricos, MSC (40780). 2022-23
- [7] "The (Mis)Behavior of Markets: A Fractal View of Risk, Ruin and Reward". Benoit Mandelbrot e Richard L. Hudson (2004).
- [8] "Lévy Flights and Related Topics in Physics". Shlomo Havlin e Armin Bunde (1995).
- [9] "Introduction to Econophysics: Correlations and Complexity in Finance". Rosario N. Mantegna e H. Eugene Stanley (2000).

Appendix

Task 3.1

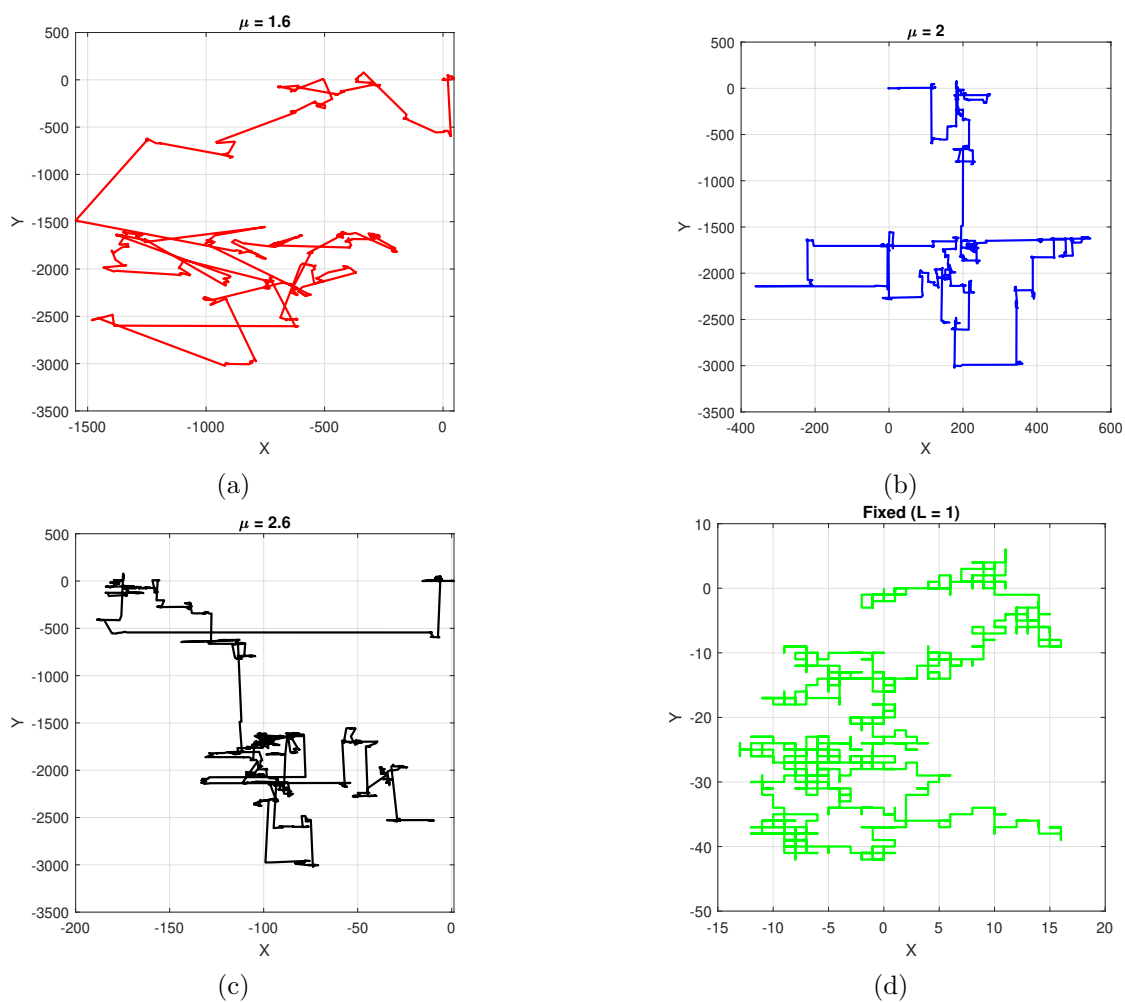


Figura 10: Representações gráficas das trajetórias, no plano X-Y, para diferentes valores de μ e para o caso de *random walks* de salto fixo