

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
DEPARTAMENTO DE ESTATÍSTICA

**INFERÊNCIA DA CONECTIVIDADE NEURONAL  
E O MODELO DE GALVES E LÖCHERBACH**

**Vitor Gomes Schiavone**

**Trabalho de Conclusão de Curso**



UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
DEPARTAMENTO DE ESTATÍSTICA

INFERÊNCIA DA CONECTIVIDADE NEURONAL E O  
MODELO DE GALVES E LÖCHERBACH

**Vitor Gomes Schiavone**

**Orientador: Ricardo Felipe Ferreira**

**Coorientador: Rodrigo Felipe de Oliveira Pena**

Trabalho de Conclusão de Curso apresentado ao Departamento de Estatística da Universidade Federal de São Carlos - DEs-UFSCar, como parte dos requisitos para obtenção do título de Bacharel em Estatística.

São Carlos  
Outubro de 2022



FEDERAL UNIVERSITY OF SÃO CARLOS  
EXACT AND TECHNOLOGY SCIENCES CENTER  
DEPARTMENT OF STATISTICS

NEURONAL CONNECTIVITY INFERENCE AND THE  
GALVES-LÖCHERBACH MODEL

**Vitor Gomes Schiavone**

**Advisor: Ricardo Felipe Ferreira**

**Co-advisor: Rodrigo Felipe de Oliveira Pena**

Bachelors dissertation submitted to the Department of Statistics, Federal University of São Carlos - DEs-UFSCar, in partial fulfillment of the requirements for the degree of Bachelor in Statistics.

São Carlos  
September 2022



Vitor Gomes Schiavone

INFERÊNCIA DA CONECTIVIDADE NEURONAL E O  
MODELO DE GALVES E LÖCHERBACH

Este exemplar corresponde à redação final do trabalho de conclusão de curso devidamente corrigido e defendido por Vitor Gomes Schiavone e aprovado pela banca examinadora.

Aprovado em dia 16 de Setembro de 2022

Banca Examinadora:

- Ricardo Felipe Ferreira (Orientador)
- Alessandro Giacomo Grimbart Gallo
- Márcio Luis Lanfredi Viola





*Dedico este trabalho aos meus familiares, amigos e professores, os quais sempre me incentivaram a acreditar em meu potencial.*



# Agradecimentos

Na vida passamos por diversos desafios, os quais são mais facilmente superados com o auxílio de familiares, amigos e fé e, por isso, trago aqui os agradecimentos a todos que fizeram parte dessa etapa de minha vida.

Inicialmente, agradeço a Deus, por me conceder a calma, a sabedoria e a capacidade necessária para concluir não só este trabalho, como também o curso como um todo. Posteriormente, agradeço a minha família, a minha mãe Adriana, ao meu pai Luiz e a minha irmã Ana Luiza, os quais sempre me incentivaram, apoiaram e, acima de tudo, foram as inspirações para que eu seguisse os meus objetivos.

Agradeço também, ao meu orientador Ricardo Ferreira que acreditou em meu potencial durante todo o trabalho, compartilhando todo o conhecimento necessário e tendo completa paciência. Em seguida, agradeço ao meu coorientador Rodrigo Pena que se dedicou para adicionar conhecimentos esclarecedores sobre as questões de neurociência em nosso trabalho, refinando-o ainda mais. Sem dúvidas o trabalho não seria o mesmo sem as participações de Ricardo e Rodrigo.

Além disso, agradeço aos grandes amigos que fiz durante a graduação, os quais viveram esses desafios simultaneamente a mim e me apoiaram até o fim. Em especial, agradeço Luben Cabezas, André Oliveira, Luiz Piccin, Reinaldo Anacleto, Vitor Ramos e a todo o grupo do Uniforme.

Por fim, agradeço a todos do Departamento de Estatística da Universidade Federal de São Carlos e reconheço todas as pessoas que passaram pela minha vida e impactaram na pessoa que me tornei hoje.



*“São os detalhes”*

(Professor José Carlos Martinez)



# Resumo

Uma das questões mais importantes em neurociência é entender a dinâmica do comportamento animal, isto é, a ação conjunta de um grande número de neurônios, partes do corpo, e o ambiente. A relação entre esse comportamento e a maneira com a qual os neurônios interagem uns com os outros é uma questão desafiadora, uma vez que experimentos reais têm acesso somente a uma pequena parte do sistema neural. Alternativamente, modelos matemáticos genéricos são considerados na modelagem desse fenômeno. Neste trabalho, a atividade de cada neurônio é representada por um processo estocástico a tempo discreto, cujas variáveis aleatórias indicam se houve ou não um disparo em um dado instante de tempo. Para cada neurônio, a probabilidade de se observar um disparo em um dado instante de tempo depende da evolução dos neurônios pré-sinápticos a partir do seu último disparo. Vamos considerar a modelagem dessa probabilidade a partir do modelo de Galves e Löcherbach. Quando um neurônio dispara, seu potencial de membrana é resetado para o seu potencial de repouso e sinais elétricos são gerados, modificando o potencial de membrana de todos os neurônios pós-sinápticos. A relação entre um neurônio e seus neurônios pré-sinápticos e pós-sinápticos definem um grafo orientado com peso. O objetivo deste trabalho é estudar, via simulação, a performance do processo de estimação desse grafo a partir da observação da atividade neuronal de um conjunto finito de neurônios ao longo de intervalo de tempo limitado, considerando como subjacente o modelo de Galves e Löcherbach.

**Palavras-chave:** *Cadeias estocásticas com memória de alcance variável, Modelo de Galves e Löcherbach, Inferência da conectividade neuronal.*





# Abstract

One of the most important issues in neuroscience is to understand the dynamics of animal behavior, that is, the joint action of large numbers of neurons, body parts, and the environment. The relationship between this behavior and the way that neurons interact with each other is a challenging question, since real experiments only have access to a small part of the neural system. Alternatively, generic mathematical models are considered to describe this phenomenon. In this work, the activity of each neuron is represented by a discrete-time stochastic process, whose random variables indicate whether or not there was a neuronal spike at a given instant of time. For each neuron, the probability of observing a neuronal spike at a given instant of time depends on the evolution of presynaptic neurons since their last spiking time. We use the Galves-Löcherbach model to model this spike probability. When a neuron spikes, its membrane potential is reset to its resting potential and electrical signals are generated, modifying the membrane potential of all postsynaptic neurons. The relationship between a neuron and its presynaptic and postsynaptic neurons defines a weighted oriented graph. The objective of this work is studying, from a computation point of view, the performance of the estimation process of the connectivity graph from the observation of the neuronal activity of a finite set of neurons over a limited time interval, considering the Galves-Löcherbach model as underlying model.

**Keywords:** *Stochastic chains with variable-range memory, Galves-Löcherbach model, Neuronal connectivity inference.*



# Lista de Figuras

2.1	Divisões do cérebro . . . . .	26
2.2	Divisões morfológicas . . . . .	27
2.3	Comportamento da voltagem da membrana . . . . .	29
3.1	Representação pictorial da árvore $\tau_1$ definida no Exemplo 3.8. Adaptado de Girardi (2021). . . . .	41
3.2	Representação pictorial da árvore $\tau_2$ definida no Exemplo 3.8. Adaptado de Girardi (2021). . . . .	41
3.3	Representação pictorial da árvore de contexto probabilística. . . . .	42
4.1	Erro quadrático médio empírico calculado a partir com estimativas do peso sináptico $W_{3 \rightarrow 5}$ de uma rede com 5 neurônios, via método da máxima verossimilhança com LASSO, e considerando cinco diferentes tamanhos amostrais: $T=120$ , $T=1200$ , $T=12000$ , $T=24000$ e $T=48000$ . . . . .	60
4.2	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão ou erro quadrático médio empírico menor que 0.05, para uma rede de 5 neurônios e cinco diferentes tamanhos de amostra: $T = 120$ , $T = 1200$ , $T = 12000$ , $T = 24000$ e $T = 48000$ . Código da cor: Verde - 100% na proporção de acertos da conexão ou $ \hat{W}_{j \rightarrow i}^{(T)}  < 0.05$ , Vermelho - caso contrário. . . . .	63
4.3	Ilustração do grafo de conectividade por meio da proporção de acertos dos tipos de conexão entre os neurônios, considerando apenas os realmente conectados, para uma rede de 5 neurônios e cinco diferentes tamanhos de amostra: $T = 120$ , $T = 1200$ , $T = 12000$ , $T = 24000$ e $T = 48000$ . Código da cor: Verde - 100% dos tipos de conexão estimadas corretamente, Vermelho - menos de 100% dos tipo de conexão estimadas corretamente. . . . .	64



# Sumário

<b>1</b>	<b>Introdução</b>	<b>21</b>
<b>2</b>	<b>Neurobiologia</b>	<b>25</b>
2.1	Conceitos básicos em neurociência . . . . .	25
2.1.1	O cérebro . . . . .	25
2.1.2	O neurônio . . . . .	26
2.1.3	Transmissão e processamento de informação no cérebro . . . . .	28
2.2	Modelos neurais . . . . .	29
2.2.1	Modelos de um único neurônio . . . . .	29
2.2.2	Modelos de massa neural . . . . .	31
2.3	Estocasticidade no cérebro . . . . .	32
<b>3</b>	<b>Inferência da conectividade neural</b>	<b>35</b>
3.1	Notações, definições e conceitos preliminares . . . . .	35
3.2	Cadeia estocástica com memória de alcance variável . . . . .	37
3.3	Modelo de Galves e Löcherbach . . . . .	43
3.4	Grafo de conectividade . . . . .	47
3.5	Estimação do grafo de conectividade . . . . .	50
<b>4</b>	<b>Aplicação Computacional</b>	<b>53</b>
4.1	Hipóteses . . . . .	53
4.1.1	Geração da rede neuronal . . . . .	54
4.1.2	Estimação do Grafo de Conectividade . . . . .	57
4.2	Resultados . . . . .	58
4.2.1	Estimação da matriz de pesos sinápticos . . . . .	58
4.2.2	Detecção de conexões sinápticas . . . . .	61

<b>5</b>	<b>Considerações Finais</b>	<b>67</b>
	<b>Referências Bibliográficas</b>	<b>69</b>

# Capítulo 1

## Introdução

Neurociência é o estudo científico do sistema nervoso. O cérebro é o principal e mais complexo órgão desse sistema. Ele é composto por bilhões de neurônios ([Herculano-Houzel, 2009](#)), que são células eletricamente excitáveis especializadas na troca de informação, i.e., elas são capazes de se comunicar com outros neurônios e outros tipos de células por meio de junções especializadas chamadas sinapses, nas quais sinais elétricos podem ser transmitidos de uma célula para outra. As sinapses podem ser elétricas ou químicas. Nas sinapses elétricas, ocorre o fluxo direto de íons entre as células, enquanto que, nas sinapses químicas, um neurotransmissor é liberado após o disparo de um neurônio (pré-sináptico) e se liga a uma proteína em uma célula receptora (pós-sináptica) fazendo com que uma despolarização no seu potencial de membrana ocorra ([Koch, 2004](#)).

O potencial de membrana é a diferença de potencial elétrico entre o meio intracelular e o meio extracelular. O potencial de repouso ocorre quando essa diferença se mantém aproximadamente constante devido a uma estabilização do fluxo iônico. Por outro lado, o potencial de ação pode ocorrer quando um estímulo (externo ou sináptico) provoca uma variação brusca nessa diferença de potencial elétrico levando a uma descarga elétrica. Potenciais de ação em neurônios são também conhecidos como impulsos nervosos ou disparos, e a sequência temporal dos potenciais de ação gerados por um neurônio é chamada de trem de disparo neuronal. A atividade neuronal é, então, manifestada pela emissão de sequências de disparos neuronais de milhões de neurônios e o processamento cerebral é decorrente da interação dessas sequências ([Gerstner \*et al.\*, 2014](#)).

Matemática é fundamental para entender e acompanhar o progresso da neurociência ([Ermentrout e Terman, 2010](#)). Nesse sentido, um modelo de neurônio é uma equação ou um conjunto de equações que descrevem a evolução do potencial de membrana do

neurônio e, possivelmente de variáveis adicionais, de forma que a modelagem é biologicamente plausível, matematicamente tratável e numericamente eficiente (Cessac, 2010). Registros da atividade neuronal têm revelado que os disparos elétricos podem ocorrer de forma espontânea e irregular (Crochet *et al.*, 2011; Naud e Gerstner, 2012) e variam mesmo quando o neurônio é exposto aos mesmos tipos de estímulos (Bair e Koch, 1996). Além disso, durante um dado período, diferentes sequências de disparos ocorrem mesmo sem alterações no estímulo e o fato de uma única sequência em particular ocorrer em meio a todas as sequências possíveis sugere uma estrutura probabilística para descrição matemática e o tratamento de fenômenos neuronais (London *et al.*, 2002). Portanto, utilizamos um processo estocástico para modelar a atividade de um neurônio ao longo do tempo.

Na literatura, há trabalhos que mostram ser inadequada a utilização de uma descrição Markoviana para os processos estocásticos utilizados na modelagem da atividade neuronal (Friston, 2010; Truccolo *et al.*, 2010; Cessac, 2011). Na verdade, a probabilidade de disparo de um neurônio depende da atividade acumulada do sistema após o seu último disparo. Dessa forma, a evolução temporal de cada neurônio pode ser modelada através de uma cadeia estocástica de memória com alcance variável (Galves e Löcherbach, 2013, 2016).

Neste trabalho, vamos considerar um conjunto finito de neurônios cuja atividade é modelada pelo modelo a tempo discreto introduzido por Galves e Löcherbach (2016), para o qual a probabilidade de um neurônio disparar em um determinado tempo, dada a configuração passada da rede, depende de alguns poucos neurônios e de poucos instantes de tempo no passado.

A atividade neuronal pode ser observada diretamente, enquanto interações entre as estruturas neuronais podem somente ser inferidas dos dados. Portanto, estamos interessados em estimar o grafo de interação de um grupo de neurônios a partir do modelo de Galves e Löcherbach. Para isso, vamos resolver um conjunto de equações não-lineares que satisfazem o critério de máxima verossimilhança com penalização sobre a norma  $\ell_1$  dos coeficientes. Pretendemos avaliar a consistência e a consistência dos estimadores por meio de simulações computacionais. Até onde vai nosso conhecimento, não há na literatura trabalho algum que trata da consistência do método da máxima verossimilhança com regularização  $\ell_1$  dos coeficientes para o modelo de Galves e Löcherbach. Portanto, as evidências empíricas obtidas neste trabalho complementam, em um certo sentido, os estudos sobre a inferência da conectividade funcional entre neurônios.



Este trabalho está organizado da seguinte maneira. No próximo capítulo, apresentamos definições e noções básicas sobre neurociência. No Capítulo 3, revisamos alguns conceitos básicos da teoria de processos estocásticos, introduzimos a classe das cadeias com memória de alcance variável, apresentamos o modelo de Galves e Löcherbach e tratamos da estimação do grafo de interação através do método da máxima verossimilhança com penalização sobre a norma  $\ell_1$  dos coeficientes. No Capítulo 4, apresentamos um estudo de simulação para avaliar a consistência do método de estimação proposto. O Capítulo 5 encerra essa monografia com algumas considerações finais e os próximos passos deste trabalho.



# Capítulo 2

## Neurobiologia

O cérebro, órgão encarregado de controlar ações importantes do corpo humano, além de ser composto por subáreas, é constituído por células nervosas denominadas neurônios, células responsáveis por receber e transmitir sinais entre si e para outras células do organismo a partir de disparos elétricos. Para que ocorra um disparo é preciso que, dentro do núcleo da célula, em sua membrana, ocorra um estímulo suficientemente grande para gerar um potencial de ação e, conseqüentemente, o disparo. Nesse sentido, para compreender o comportamento neuronal, foram propostos modelos neurais que observam tanto um único neurônio quanto um conjunto, sendo que os modelos que consideram um conjunto de neurônios são denominados modelos de massa neural. A seguir, falaremos sobre a anatomia do cérebro, a estocasticidade do cérebro, sobre os neurônios e os modelos neurais.

### 2.1 Conceitos básicos em neurociência

O sistema nervoso é responsável por captar, processar e transmitir estímulos ao longo do corpo humano. Esse sistema é composto por diferentes componentes, mas principalmente pelo cérebro. Neste momento, vamos entender o funcionamento cerebral e como ocorrem as atividades neuronais a partir das características anatômicas do cérebro e dos neurônios.

#### 2.1.1 O cérebro

Responsável por todo o comportamento humano e principal componente do sistema nervoso central, o cérebro é a estrutura viva mais complexa que conhecemos. Ele controla

tudo que fazemos, coordenando os pensamentos, emoções e ações.

Anatomicamente, o cérebro é bilateral e essencialmente simétrico, ele tem em média  $1.500\text{ cm}^3$  sendo um dos maiores dentre os mamíferos terrestres e é composto pelos seguintes principais componentes: tronco cerebral, cerebelo, tálamo, hipotálamo, corpo caloso e córtex cerebral (Moya, 2011). As principais funções cognitivas são de responsabilidade do córtex, uma massa cinzenta que cobre os hemisférios cerebrais e que é dividida em quatro partes: lobo frontal, parietal, temporal e occipital, sendo que esses lobos tem funções especializadas (Kandel *et al.*, 2000).

Além disso, o cérebro contém dentre 15 a 33 bilhões de neurônios espalhados por todo tecido nervoso, sendo células especializadas capazes de processar e transmitir informações para outras células nervosas, músculos ou glândulas.

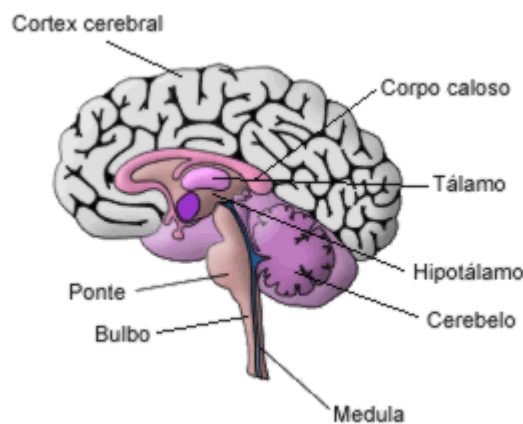


Figura 2.1: **Divisões do cérebro.** De cima para baixo: Córtex cerebral, Corpo caloso, Tálamo, Hipotálamo, Cerebelo e Tronco cerebral representado pela Ponte, Bulbo e Medula. Adaptado de: <https://www.sobiologia.com.br/conteudos/FisiologiaAnimal/nervoso8.php>.

### 2.1.2 O neurônio

O cérebro é constituído de bilhões de neurônios, e apesar de podermos classificá-los em diversos diferentes tipos, eles contém uma mesma arquitetura básica em comum, sendo compostos por um corpo celular, axônio, dendritos e sinapses. Cada um desses componentes tem um papel diferente na geração e transmissão dos sinais entre as células nervosas (Moya, 2011).

O corpo celular é o centro metabólico da célula responsável por armazenar o núcleo e, conseqüentemente, os genes e o retículo endoplasmático em que as proteínas das células

são sintetizadas. Ademais, do corpo celular originam-se outras duas partes do neurônio, os dendritos, que se ramificam e são responsáveis por receber sinais de outras células, e o axônio, que ao contrário dos dendritos, se estende ao longo da célula e é encarregado de transmitir sinais elétricos, chamados de potencial de ação, para outros neurônios. Curiosamente, um axônio pode transmitir o potencial de ação ao longo de distâncias que variam de 0,1 milímetros a 3 metros, e esses sinais são impulsos elétricos de até 100 milivolts (Kandel *et al.*, 2000).

Por fim, a sinapse é o ponto em que dois neurônios se unem ou se comunicam, existindo as sinapses químicas e as elétricas. As sinapses químicas comunicam-se a partir de neurotransmissores que são liberados após a membrana atingir um potencial de ação, sendo que os neurotransmissores se ligam à célula que recebe a informação e a torna mais propensa a disparar. Já as elétricas conectam diretamente um neurônio a outro e a comunicação ocorre a partir de um fluxo de íons entre as células. É importante saber que a célula que transmite o sinal é chamada de pré-sináptica e a que recebe o sinal é chamada de pós-sináptica.

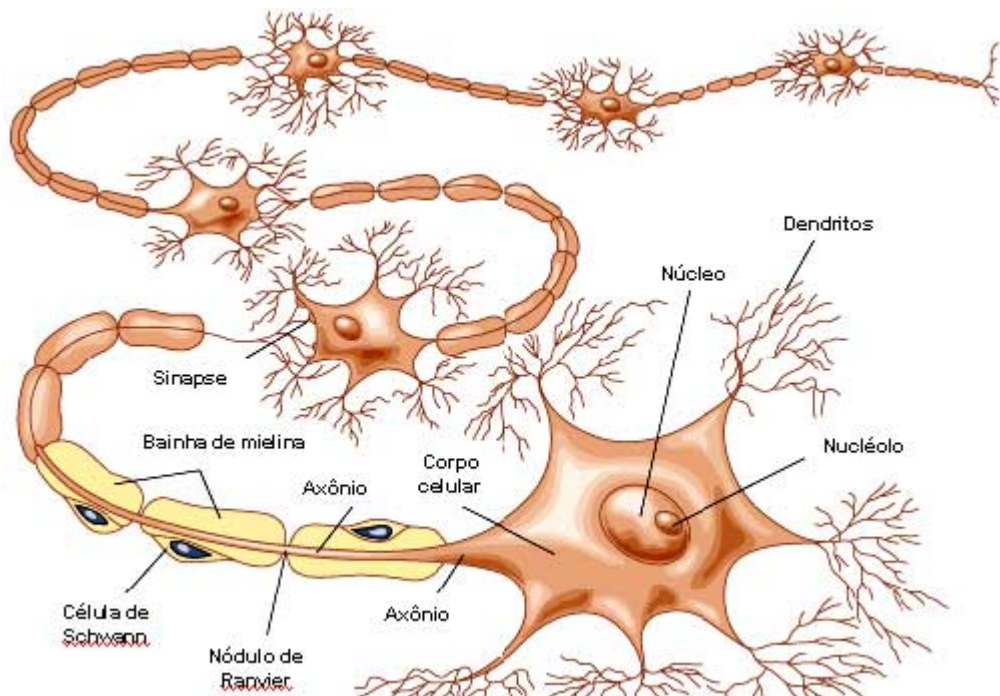


Figura 2.2: **Divisões morfológicas do neurônio.** Da esquerda para a direita: Sinapse, Axônio e seus agregados, Corpo Celular composto por Núcleo e Dendritos. Adaptado de: <https://www.sobiologia.com.br/conteudos/Histologia/epitelio27.php>.

### 2.1.3 Transmissão e processamento de informação no cérebro

Sabendo a anatomia do neurônio e quais as funções de seus componentes, podemos falar sobre como ocorre a transmissão e o processamento de informação entre eles. De forma geral, para que ocorra um disparo, é necessário que a membrana, elemento presente no núcleo e responsável por realizar os disparos, acumule um potencial suficientemente grande a partir de um estímulo elétrico que gera uma despolarização na célula.

Mais profundamente, quando a membrana de um neurônio está polarizada, dizemos que ela está em repouso, o que significa que há uma diferença de cargas elétricas do meio interno para o meio externo da membrana, em média, uma célula em repouso apresenta um potencial de -65 mV. Assim polarizada, em cargas negativas no meio interno e positivas no meio externo, as cargas elétricas se organizam em um estado estacionário e estabelecem uma energia elétrica potencial através da membrana: o potencial de ação ou potencial de repouso.

Dessa forma, no momento em que chega um estímulo ao neurônio, cuja membrana está em repouso, se ele for suficientemente grande, a célula se torna mais permeável e pode ser que ocorra uma grande entrada de cargas positivas na célula e uma pequena vazão de cargas negativas, resultado em uma inversão das cargas ao redor da membrana, o que a deixa despolarizada e gera o potencial de ação. Dessa maneira, a despolarização se propaga pelo neurônio e, conseqüentemente, gera um disparo. Por fim, imediatamente após a passagem do impulso, a membrana sofre a repolarização, recuperando o estado de repouso (Moya, 2011).

Resumidamente, quando a membrana de um neurônio pré-sináptico está eletricamente estimulada, sua sinapse se abre e envia os neurotransmissores para uma célula pós-sináptica, em seguida, os receptores que se localizam nos dendritos da célula pós-sináptica coletam os sinais e o enviam para o núcleo. O núcleo por sua vez, considerando que o estímulo foi suficientemente grande, gera um disparo de saída e envia empulsos elétricos para os axônios, que transmitem para outros neurônio. E nesse sentido, os impulsos são processados e retransmitidos.

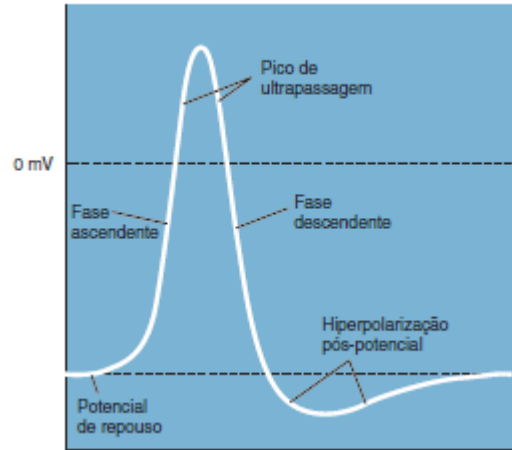


Figura 2.3: **Comportamento da voltagem da membrana.** Da esquerda para a direita em branco o potencial de repouso, que ascende até ultrapassar o limiar de 0 mV tracejado, gerando o potencial de ação. Após isso a voltagem descende e retorna aos níveis do potencial de repouso. Adaptado de [Bear \*et al.\* \(2017\)](#).

## 2.2 Modelos neurais

Existem diversas formas de estudar o cérebro. Podemos subdividi-lo em áreas, de acordo com suas funções, ou podemos ir além e estudar o comportamento de cada neurônio individualmente. Nesse sentido, a partir dessas duas ideias, modelos neurais foram propostos para entender o comportamento dos neurônios, tanto como um conjunto quanto de forma individual. Assim, a seguir falaremos sobre o modelo de Huxley e Hodgkin e o modelo integra e dispara, que modelam o comportamento de apenas um neurônio, e abordaremos modelos de massa neural, que abrangem um conjunto de neurônios e buscam estimar o comportamento da rede.

### 2.2.1 Modelos de um único neurônio

A utilidade de um modelo de um único neurônio é a de determinar a dinâmica de um neurônio de uma forma detalhada, focando em diferentes parâmetros a uma escala microscópica. Existem diversos modelos de um único neurônio que consideram diferentes tipos de equações e parâmetros, mas todos têm como função prever, qual será a saída, dado que houve uma entrada induzida pelo disparo de outro neurônio ou por um estímulo externo. Apesar de haver diversos modelos para descrever o comportamento neuronal, neste sentido, todos eles consideram como variável mais importante a voltagem da membrana no neurônio. A seguir, serão apresentados dois modelos de um único neurônio:

modelo de Huxley e Hodkin, e o modelo Integra e Dispara.

### Modelo de Huxley e Hodkin

Como visto anteriormente, impulsos elétricos transmitidos entre os neurônios a partir dos canais iônicos da membrana podem excitar um potencial de ação, que faz com que os neurônios pós-sinápticos (aqueles que recebem os sinais) disparem. Pensando nisso, para modelar o comportamento dos impulsos neuronais, a variável considerada como sendo a mais importante é a voltagem na membrana do neurônio. Desse modo, em 1952, Hodgkin e Huxley realizaram experimentos com axônios de lulas, com o intuito de mensurar e entender o comportamento da corrente elétrica que leva a membrana a gerar o potencial de ação.

A dinâmica do modelo é descrita a partir de equações diferenciais, que descrevem o comportamento do potencial de membrana em função de diversos canais iônicos, diferentes sinapses e da geometria espacial específica. As equações do modelo são

$$I(t) = I_c(t) + I_K + I_{Na} + I_L$$

e

$$C \frac{dv(t)}{dt} = I(t) - g_k n^4 (u - E_k) - g_k m^3 h (u - E_{Na}) - g_{RL} (u - E_L),$$

em que  $v(t)$  é o potencial de membrana,  $n^4$  representa a variável de abertura do canal de potássio,  $m^3$  abertura do canal de sódio,  $h$  a variável de fechamento do canal de sódio;  $g_k$ ,  $g_{Na}$  e  $g_L$  representam as condutâncias de cada um dos respectivos íons e são tratadas como constantes;  $E_K$ ,  $E_{Na}$  e  $E_L$  são os potenciais de reversão desses íons; e  $I_c(t)$ ,  $I_K(t)$ ,  $I_{Na}(t)$  e  $I_L$  são as intensidades das correntes que fluem pelos respectivos canais iônicos (Girardi, 2021).

### Modelo integra e dispara

O modelo integra e dispara foi introduzido por Lapicque em 1907 e, assim como o modelo anterior, é um modelo de um único neurônio e é estruturado em função do potencial de membrana. Ele explica de forma simplificada o modelo de um neurônio considerando a membrana como um circuito resistor capacitor. Além disso, esse modelo permite a inclusão de ruídos no potencial de membrana, o que pode ser útil para experimentos com



ruídos endógenos. As equações do modelo são:

$$\begin{cases} \frac{dv(t)}{dt} = \frac{v(t)}{\tau_m} + I_{ext} + I_{sin}(t) + D\xi(t), & 0 < v(t) < V_{th} \\ v(t_0^+) = 0 \\ v(t_0^-) = V_{th}, \end{cases} \quad (2.1)$$

$$I_{sin}(t) = g \sum_{\text{disparos}} f(t - t_{\text{disparo}}) \quad (2.2)$$

e

$$f(t) = \exp\left(\frac{-t}{\tau_1}\right) - \exp\left(\frac{-t}{\tau_2}\right), \quad (2.3)$$

em que  $v(t)$  é o potencial de membrana no instante de tempo  $t$ ,  $V_{th}$  é o limiar para a geração de um disparo,  $I_{ext}$  é a corrente de um estímulo externo,  $I_{sin}$  é a soma das correntes sinápticas,  $D$  é a intensidade do ruído,  $\xi(t)$  representa o ruído branco normal com  $E(\xi) = 0$  e  $\text{Var}(\xi) = 2\delta(t - t')$ ,  $\tau_1$  e  $\tau_2$  são constantes de tempo que caracterizam as correntes sinápticas,  $\tau_m$  é a constante de tempo da membrana e  $g$  é a força da conexão entre dois neurônios.

Desse modo, um disparo ocorre quando  $v(t)$  atinge um potencial suficientemente grande  $V_{th}$  e, em seguida, a célula retorna ao potencial de repouso. O disparo causado por esse neurônio é transmitido para um célula pós-sináptica, e essa entrada é definida como  $I_{sin}$ .

No sistema de equações (2.1), conforme o tempo passa e o neurônio recebe disparos de outros neurônios, a voltagem do potencial de membrana vai se alterando, subindo até que atinge o limiar  $V_{th}$ . Apenas lembrando que, apesar de em uma rede todos os neurônios estarem conectados e todos os disparos serem transmitidos para toda a rede neural, este modelo descreve o comportamento dinâmico de apenas um neurônio (Moya, 2011).

## 2.2.2 Modelos de massa neural

Dada a complexidade de se trabalhar com o cérebro e ao alto número de neurônios que ele possui, surgiu o interesse em se modelar a atividade neural de forma mais abrangente, ou seja, de forma que pudessemos entender o comportamento de uma rede de neurônios e não de apenas um, se aproximando mais da realidade. Esses modelos são chamados de

modelos de campo médio e são ferramentas importantes nos estudos das oscilações neurais e na descrição da evolução espaço-temporal de variáveis como a taxa média de disparo.

A ideia desses modelos é selecionar um conjunto de neurônios e estimar o comportamento da rede a partir das propriedades médias, interações e comportamentos dos neurônios pertencentes a ela. A vantagem é que esses modelos são mais simples e fáceis de analisar quando comparados aos modelos de um neurônio.

Neste trabalho, vamos considerar a abordagem de modelos de massa neural para descrever o comportamento neuronal, selecionando um conjunto finito de neurônios e observando alguns instantes de tempo no passado a partir do modelo proposto por [Galves e Löcherbach \(2013\)](#), que é um modelo funcional de massa neural que tem como principal componente uma função do potencial de membrana dos neurônios.

## 2.3 Estocasticidade no cérebro

Quando a atividade de um neurônio é modelada matematicamente, podemos simular o seu comportamento ao longo do tempo por meio de algoritmos computacionais, de forma a se obter como resultado uma sequência regular de disparos. No entanto, registros da atividade neuronal revelam uma atividade espontânea irregular dos neurônios ([Crochet et al., 2011](#); [Naud e Gerstner, 2012](#)) e uma variabilidade da resposta dos neurônios a certos estímulos ([Bair e Koch, 1996](#)), ou seja, existem evidências que corroboram para a veracidade da hipótese de que haja ruídos no cérebro, os quais podem ser visto como flutuações aleatórias que não fazem parte do sinal de disparo.

Em um experimento, [Holt et al. \(1996\)](#) registraram a atividade neuronal do córtex visual de um gato *in vivo* e *in vitro*. A partir disso, foi identificado que, no animal *in vitro*, a maioria dos neurônios disparavam regularmente após receber uma injeção de corrente constante. Porém, por outro lado, no animal *in vivo*, os neurônios responderam a corrente de forma irregular. Apesar dos trens de disparos variarem bastante de um experimento para o outro, esses resultados fortalecem a hipótese da presença de ruídos no cérebro ([Moya, 2011](#)).

Além disso, não se sabe exatamente de onde se originariam esses ruídos, pois eles podem surgir de diferentes fontes, como dos canais iônicos ou até mesmo dos milhares de disparos produzidos por outros neurônios e a atividade de outras células.

Desse modo, esses dados experimentais sugerem que os neurônios, as sinapses e o

sistema neuronal são inerentemente estocásticos (Buesing *et al.*, 2011). Portanto, não é simples modelar esse comportamento, uma vez que, em um dado período de tempo, diferentes sequências de disparos ocorrem mesmo sem alterações no estímulo e o fato de uma única sequência em particular ocorrer em meio a todas as sequências possíveis sugere uma estrutura probabilística para a descrição do comportamento dos neurônios e, nesse sentido, utiliza-se um processo estocástico para modelar a atividade neuronal ao longo do tempo.



# Capítulo 3

## Inferência da conectividade neural

Dadas as características neurobiológicas do cérebro, temos recursos para apresentar as metodologias que serão utilizadas para realizar a inferência da conectividade neural. Assim, neste capítulo, apresentaremos inicialmente notações e conceitos iniciais que serão utilizados em seguida para definir, nesse mesmo capítulo, uma cadeia estocástica com memória de alcance variável, ou seja, uma cadeia estocástica que se assemelha ao comportamento que queremos descrever (conectividade neural), um modelo apropriado para descrever o comportamento dessa espécie de cadeia e, por fim, a técnica para realizar a inferência do grafo de conectividade.

### 3.1 Notações, definições e conceitos preliminares

Suponha que estamos monitorando a atividade de um neurônio durante uma janela de tempo finita  $[0, T]$ , em que o tempo é medido em milisegundos (ms) e  $T$  é um número real positivo. Para fins de registro, é útil considerar uma partição  $\{0, \delta, 2\delta, 3\delta, \dots, T\}$  do intervalo  $[0, T]$ , em que  $\delta$ ,  $\delta < T$ , é um número inteiro positivo denominado *tamanho do bin*. O tamanho do bin é da ordem de poucos milisegundos e o registro de um trem de disparo neuronal pode ser representado por uma sequência  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ , em que

$$x_t = \begin{cases} 1, & \text{se observamos um disparo do neurônio no } t\text{-ésimo bin,} \\ 0, & \text{caso contrário.} \end{cases}$$

Registros da atividade neuronal têm revelado que os disparos elétricos podem ocorrer de forma espontânea e irregular (Crochet *et al.*, 2011; Naud e Gerstner, 2012) e variam

mesmo quando o neurônio é exposto aos mesmos tipos de estímulos (Bair e Koch, 1996). Além disso, durante um dado período, diferentes sequências de disparos ocorrem mesmo sem alterações no estímulo e o fato de uma única sequência em particular ocorrer em meio a todas as sequências possíveis sugere uma estrutura probabilística para descrição matemática e o tratamento de fenômenos neuronais (London *et al.*, 2002). Portanto, utilizamos um processo estocástico para modelar a atividade de um neurônio ao longo do tempo.

**Definição 3.1 (Processo Estocásticos)** *Um processo estocástico é uma sequência de variáveis aleatórias  $\mathbf{X} := \{X_t : t \in \mathbb{T}\}$  definidas em um mesmo espaço de probabilidades  $(\Omega, \mathcal{A}, P)$  com valores em um alfabeto  $A \subset \mathbb{R}$ .*

Nesta monografia, vamos assumir que o processo estocástico  $\mathbf{X}$ , que modela a atividade de um neurônio ao longo do tempo, é um processo a tempo discreto. Portanto, assumimos, sem perda de generalidade,  $\mathbb{T} = \mathbb{Z}$ , em que  $\mathbb{Z}$  é o conjunto dos números inteiros. Nesse contexto,  $\mathbf{X} := \{X_t : t \in \mathbb{Z}\}$  é uma cadeia estocástica a tempo discreto tal que

$$X_t = \begin{cases} 1, & \text{se o neurônio disparar no instante de tempo } t, \\ 0, & \text{caso contrário.} \end{cases}$$

Portanto, neste trabalho, o alfabeto  $A$ , no qual o processo estocástico  $\mathbf{X}$  assume valores, é tal que  $A = \{0, 1\}$ .

Além disso, como estamos assumindo uma modelagem estocástica para a atividade neuronal, é interessante, em um contexto inferencial, considerar um conjunto de probabilidades de transição que regem a dinâmica do sistema, invariantes ao longo do tempo, e analisar os resultados estocásticos do experimento sem depender explicitamente da escolha do tempo inicial da cadeia (Girardi, 2021). Desse modo, consideraremos neste estudo cadeias estocásticas a tempo discreto que são homogêneas e estacionárias.

**Definição 3.2 (Homogeneidade)** *Um processo estocástico a tempo discreto  $\mathbf{X} := \{X_t : t \in \mathbb{Z}\}$ , definido em um espaço de probabilidade apropriado  $(\Omega, \mathcal{A}, P)$  com valores em um alfabeto finito  $A$ , é dito ser homogêneo quando as probabilidades de transição são invariantes ao longo do tempo, ou seja, os conjuntos de probabilidades condicionais que regem a dinâmica do sistema são independentes do valor de  $t \in \mathbb{Z}$ .*

**Definição 3.3 (Estacionariedade)** *Um processo estocástico a tempo discreto  $\mathbf{X} :=$*

$\{X_t : t \in \mathbb{Z}\}$ , definido em um espaço de probabilidade apropriado  $(\Omega, \mathcal{A}, P)$  com valores em um alfabeto finito  $A$ , é dito ser estacionário quando as distribuições conjuntas não dependem do tempo da escolha inicial, ou seja,

$$P\left(\bigcap_{t=m}^n \{X_t = a_t\}\right) = P\left(\bigcap_{t=m}^n \{X_{t+1} = a_t\}\right),$$

quaisquer que sejam  $m, n \in \mathbb{Z}$  e  $a_i \in A$  sendo  $i = m, m+1, \dots, n$ .

Neste momento, para prosseguir, vamos fixar algumas notações que serão utilizadas ao longo do texto. Considerando  $m$  e  $n$  dois números inteiros tais que  $m \leq n$ , denotamos por  $a_m^n$  a sequência finita  $(a_m, \dots, a_n)$ , por  $a_{-\infty}^n$  a sequência infinita à esquerda  $(\dots, a_{n-1}, a_n)$  e por  $a_m^{+\infty}$  a sequência infinita à direita  $(a_m, a_{m+1}, \dots)$ , de forma que  $a_i \in A$  para todo  $i \in \mathbb{Z}$ .

Para qualquer  $m \leq n$ , a quantidade de símbolos pertencentes à sequência  $a_m^n$  é dada por  $|a_m^n| = n - m + 1$ . Além disso, para todo  $n \in \mathbb{Z}$ , vamos considerar que  $a_{n+1}^n = \emptyset$ , visto que  $|a_{n+1}^n| = n - (n+1) + 1 = 0$ . Considerando duas sequências  $v$  e  $v'$ , denotamos por  $vv'$  a sequência de tamanho  $|v| + |v'|$  obtida a partir da concatenação de duas sequências. A junção de duas sequências é também estendida ao caso em que  $v$  representa uma sequência semi-infinita, isto é,  $v = v_{-\infty}^{-1}$ . Se  $n$  é um inteiro positivo e  $v$  uma sequência finita de símbolos em  $A$ , denotamos por  $v^n = vv\dots v$  a junção de  $n$  vezes a sequência  $v$ .

Nesse contexto, denotaremos o conjunto de todas as sequências finitas e o conjunto de todas as sequências infinitas à esquerda, respectivamente, por

$$\mathbb{A}^* = \bigcup_{j=0}^{+\infty} \mathbb{A}^{\{-j, \dots, -1\}} \quad \text{e} \quad \mathbb{A}^{-\mathbb{Z}} = \mathbb{A}^{\{\dots, -2, -1\}}.$$

No caso em que  $j = 0$  temos a sequência vazia denotada por  $\emptyset$  (Gallo, 2011).

## 3.2 Cadeia estocástica com memória de alcance variável

De acordo com as características apresentadas sobre o cérebro, neurônios e como as sequências de disparos funcionam e podem ser modeladas probabilisticamente, nesta monografia, trabalharemos com cadeias estocásticas de alcance variável para descrever o comportamento da atividade neuronal. A seguir, por meio do Exemplo 3.4, entenderemos o motivo.

**Exemplo 3.4** *Considere uma cadeia estocástica a tempo discreto  $V = \{V_t : t \geq 0\}$ , em que  $V_t$  denota o potencial de membrana no tempo  $t \geq 0$  de um dado neurônio. Vamos assumir que  $V_0 = v_R$  e, para  $t \geq 1$ , a cadeia se comporta da seguinte forma:*

$$V_t = \begin{cases} v_R, & \text{se o neurônio disparar no tempo } t, \\ \rho V_{t-1} + \mu, & \text{caso contrário,} \end{cases}$$

em que  $0 \leq \rho \leq 1$  modela o efeito da corrente de vazamento,  $v_R \in \mathbb{R}$  representa o potencial de repouso da membrana e  $\mu \in \mathbb{R}$  é uma corrente constante de entrada. Quando o neurônio dispara no tempo  $t$ ,  $V_t$  retorna ao potencial de repouso, caso contrário, multiplica-se o potencial da membrana no estado anterior pelo efeito da corrente  $\rho$  e soma-se com uma corrente de entrada  $\mu$ . Desse modo, o espaço de estados da cadeia estocástica  $V$  é

$$\mathcal{V} := \{v_R, \rho v_R + \mu, \rho^2 v_R + \mu(1 + \rho), \rho^3 v_R + \mu(1 + \rho + \rho^2), \dots\},$$

e a probabilidade de um neurônio disparar num instante de tempo  $t \in \mathbb{Z}$  condicionada ao potencial de membrana no instante  $t - 1$  é dada por

$$P(X_t = 1 \mid V_{t-1} = v_{t-1}) = \varphi(v_{t-1}),$$

em que  $v_{t-1} \in \mathcal{V}$ ,  $\varphi : \mathbb{R} \rightarrow [0, 1]$  é uma função crescente denominada taxa de disparo e  $X := \{X_t : t \in \mathbb{Z}\}$  é um cadeia estocástica a tempo discreto tal que

$$X_t = \begin{cases} 1, & \text{se } V_t = v_R, \\ 0, & \text{caso contrário.} \end{cases}$$

Uma pergunta natural é a respeito da Markovianidade da cadeia  $X$ . Para responder a pergunta, temos que entender como simular  $X_t$  a partir dos estados anteriores  $X_{t-1}, \dots, X_0$ . Então, para  $t \in \mathbb{Z}$ , tome  $U_t$  como uma variável aleatória com distribuição uniforme  $[0, 1]$  e considere que  $X_{t-1} = 1$  e então  $V_{t-1} = v_R$  e

$$X_t = \begin{cases} 1, & \text{caso } U_t \leq \varphi(v_R), \\ 0, & \text{caso contrário.} \end{cases}$$

Se  $X_{t-1} = 0$ , então  $V_{t-1} \neq v_R$  e seria necessário observar o valor de mais um estado



no passado  $X_{t-2}$ . No caso de  $X_{t-2} = 1$ , teríamos  $V_{t-2} = v_R$ ,  $V_{t-1} = \rho v_R + \mu$  e

$$X_t = \begin{cases} 1, & \text{caso } U_t \leq \varphi(\rho v_R + \mu), \\ 0, & \text{caso contrário.} \end{cases}$$

Contudo, se  $X_{t-2} = 0$ , então  $V_{t-2} \neq v_R$  e precisaríamos olhar  $X_{t-3}$  e assim sucessivamente. Logo, pode-se notar que para descobrir qual foi o último disparo ocorrido existe uma dependência variável em relação a quantos estados no passado observar. Desse modo, se denotarmos o último disparo realizado pelo neurônio antes do instante  $t$  por

$$L_t = \sup \{0 \leq s \leq t : X_s = 1\},$$

segue que  $X_t$  é uma função de  $(X_{t-1}, X_{t-2}, \dots, X_{L_{t-1}}, U_t)$ . Portanto, podemos notar a partir deste algoritmo que  $X$  não é uma cadeia de Markov, pois a dependência do passado não tem um tamanho fixo e sim é uma função do passado que varia a cada estado e por isso a cadeia que será utilizada é de alcance variável, cuja definição forneceremos a posteriori. (*Ost e Reynaud-Bouret, 2020*).

Para definir formalmente a dinâmica de uma cadeia estocástica com memória de alcance variável precisamos apresentar alguns conceitos tais como o de sufixo, de árvore, de árvore completa e de uma árvore de contexto probabilística.

**Definição 3.5 (Sufixo)** Dados  $m$  e  $n$  números inteiros positivos. Uma sequência de estados  $s_{-m}^{-1}$  é sufixo de uma outra sequência  $v_{-n}^{-1}$  se  $m < n$  e  $v_{-m}^{-1} = s_{-m}^{-1}$ .

**Definição 3.6 (Propriedade do Sufixo)** Um subconjunto  $\tau$  de  $A^* \cup A^{-\mathbb{N}}$  é uma árvore se nenhuma sequência de estados  $s \in \tau$  é um sufixo de outra sequência  $v \in \tau$ . Essa propriedade é denominada propriedade do sufixo.

Somado a isso, podemos denotar a altura da árvore  $\tau$  por

$$h(\tau) = \sup \{|a| : a \in \tau\}.$$

Informalmente falando, a altura da árvore  $\tau$  é a maior distância entre a raiz da árvore (desenhada no topo) e uma de suas folhas (desenhada no fim). Caso  $h(\tau) < +\infty$ , dizemos que  $\tau$  é limitada, isto é, todas as sequências da árvore  $\tau$  são finitas e com tamanho menor

ou igual a  $h(\tau)$ . Por outro lado, se  $h(\tau) = +\infty$ , dizemos que  $\tau$  é ilimitada, ou seja, existe alguma sequência infinita em  $\tau$ .

**Definição 3.7 (Árvore Completa e Árvore de Contexto)** *Uma árvore  $\tau$  é completa se qualquer elemento  $a_{-\infty}^{-1}$  de  $A^{-\mathbb{N}}$  tem um sufixo pertencente a  $\tau$ . Pela propriedade do sufixo, temos que esse sufixo é único. Chamamos esse sufixo de contexto da sequência  $a_{-\infty}^{-1}$  e é denotado por  $c_\tau(a_{-\infty}^{-1})$ . Uma árvore completa é denominada uma árvore de contexto.*

Cada sequência  $a_{-n}^{-1} \in \tau$ , com  $n$  sendo um número finito, pode ser vista como um ramo que liga a raiz da árvore até uma folha. Denomina-se a folha da árvore como sendo uma sequência  $c_{-j}^{-1}$  de  $j$  arestas começando no topo e rotuladas, de cima para baixo, por  $c_{-1}, c_{-2}, \dots, c_{-j}$ . No caso de sequências infinitas, temos sequências infinitas a esquerda, ou seja, folhas infinitas. As sequências finitas ou infinitas  $l_{-n}^{-1}$  tais que  $l_{-|n|}^{-1}$  é sufixo de  $a_{-|n|}^{-1}$  para qualquer  $a_{-|n|}^{-1} \in \tau$ , com  $|a| < |l|$ , são denominadas de nós, e quando eles são finitos, são denominados de nós internos. Os filhos do nó  $l_{-|n|}^{-1}$  são sequências  $bl_{-|n|}^{-1}$ ,  $b \in A$ . A raiz é identificada como sequência vazia (Girardi, 2021).

**Exemplo 3.8** *Considere  $A = \{0, 1\}$  como alfabeto e considere as duas árvores definidas sobre  $A$ .*

1. *A primeira árvore é dada por  $\tau_1 = \{000, 10, 01, 11\}$ . Logo,  $\tau_1$  é limitada, com  $h(\tau_1) = 3$ . Entretanto,  $\tau_1$  não é uma árvore completa, pois a sequência 000 pode ser substituída por 00 de forma que a propriedade do sufixo ainda seja satisfeita. Logo,  $\tau_1$  não é árvore de contexto.*
2. *A segunda árvore é dada por  $\tau_2 = 0^{\mathbb{Z}^-} \cup \bigcup_{i \geq 1} 10_{-i}^{-1}$ .  $\tau_2$  é ilimitada com  $h(\tau_2) = +\infty$ . Além disso,  $\tau_2$  é árvore de contexto, pois é uma árvore completa e cada uma de suas sequências satisfazem a propriedade do sufixo.*

Dessa forma, são apresentadas as Figuras 3.1 e 3.2, que mostram as representações pictoriais das árvores  $\tau_1$  e  $\tau_2$  do Exemplo 3.8.

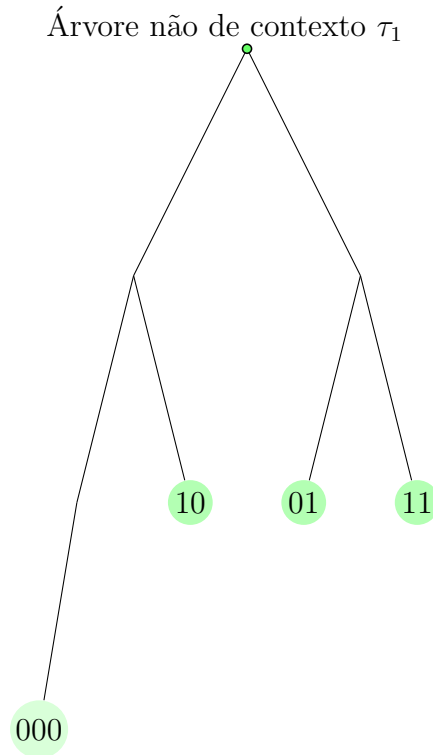


Figura 3.1: Representação pictorial da árvore  $\tau_1$  definida no Exemplo 3.8. Adaptado de Girardi (2021).

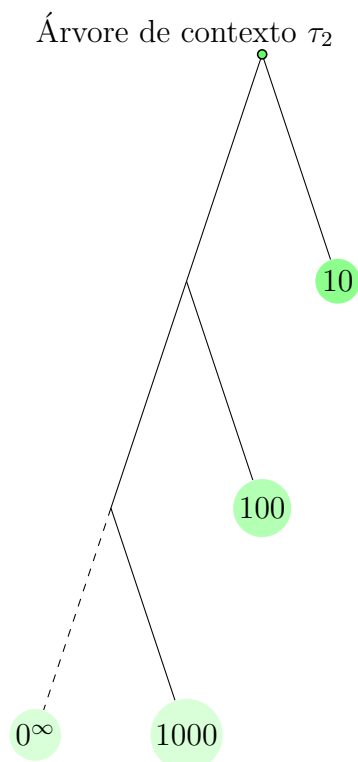


Figura 3.2: Representação pictorial da árvore  $\tau_2$  definida no Exemplo 3.8. Adaptado de Girardi (2021).

**Definição 3.9 (Árvore de Contexto Probabilística)** *Uma árvore de contexto probabilística com símbolos em  $A$  é um par ordenado  $(\tau, Q)$  tal que*

1.  $\tau$  é uma árvore de contexto;
2.  $Q := \{Q(b | a) : b \in A \text{ e } a \in \tau\}$  é uma família de probabilidades de transição sobre  $A$ .

A árvore de contexto probabilística é vista como uma representação de uma família de probabilidades de transição tais que, para cada passado, precisamos considerar apenas um contexto desse passado para definir a transição do sistema para o estado seguinte. A seguir, no Exemplo 3.3, apresentamos uma árvore de contexto probabilística.

**Exemplo 3.10** *Seja  $\tau = \{00, 10, 1\}$  uma árvore de contexto probabilística dada pelo par ordenado  $(\tau, Q)$ , tal que  $Q := \{Q(b | a) : b \in A \text{ e } a \in \tau\}$  é uma família de probabilidades de transição sobre  $A$ . Podemos escrever  $Q$  da seguinte forma*

$$Q := \begin{array}{c} \begin{array}{cc} & 0 & 1 \\ 00 & \begin{pmatrix} 1 - \alpha & \alpha \end{pmatrix} \\ 01 & \begin{pmatrix} \beta & 1 - \beta \end{pmatrix} \\ 1 & \begin{pmatrix} 1 - \gamma & \gamma \end{pmatrix} \end{array} \end{array}$$

em que  $\alpha, \beta, \gamma \in (0, 1)$ . Podemos representar a árvore de  $(\tau, Q)$  pictorialmente conforme a Figura 3.3

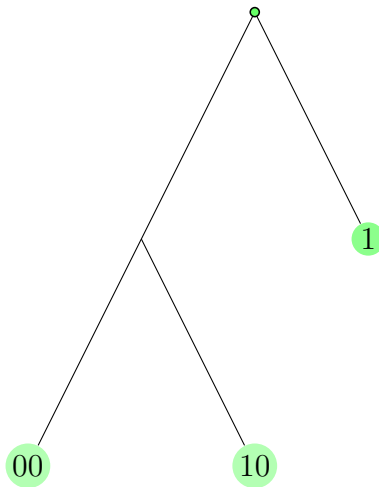


Figura 3.3: Representação pictorial da árvore de contexto probabilística.

Agora, tendo as definições necessárias, pode ser definida formalmente a cadeia com memória de alcance variável.

**Definição 3.11 (Cadeia com Memória de Alcance Variável)** *Dizemos que uma cadeia estocástica  $X := \{X_t : t \in \mathbb{Z}\}$  em um espaço de probabilidade apropriado  $(\Omega, \mathcal{A}, P)$  é compatível com uma árvore de contexto  $(\tau, Q)$  se para  $P$ -quase toda  $a_{-\infty}^{-1} \in A^{-\mathbb{N}}$  e qualquer  $a \in A$  tivermos*

$$P(X_0 = a \mid X_{-\infty}^{-1} = a_{-\infty}^{-1}) = Q(a \mid c_\tau(a_{-\infty}^{-1})).$$

*Essas cadeias são chamadas de cadeias de memória com alcance variável.*

Em inferência estatística, cadeias de Markov de ordem finita exigem muitos parâmetros para serem inferidos. Já, se uma cadeia  $X$  é de memória com alcance variável compatível com uma árvore de contexto  $\tau$  de forma que  $h(\tau) = j < \infty$ , então uma coleção de  $(|A| - 1)|\tau|$  probabilidades de transição basta para descrever a cadeia. Assim, é vantajoso, inferencialmente, considerar cadeias compatíveis com árvores de contexto probabilística, visto que em um processo de estimação na abordagem Markoviana a quantidade de parâmetros a serem estimados é superior ao da abordagem via cadeias de memória com alcance variável ([Girardi, 2021](#)).

**Observação 3.12** *A classe das cadeias estocásticas com memória de alcance variável foi introduzida por [Rissanen \(1983\)](#). Essa classe ficou popular na comunidade estatística e probabilística com o trabalho de [Bühlmann et al. \(1999\)](#) que cunhou o termo cadeias de Markov de tamanho variável para se referir às cadeias compatíveis com árvores cujos contextos eram todos finitos. Mais tarde, surge o termo cadeias estocásticas com memória de alcance variável proposto por [Galves e Löcherbach \(2008\)](#) para referir-se às cadeias que são compatíveis com árvores que possuem contextos infinitos.*

### 3.3 Modelo de Galves e Löcherbach

Sejam  $I$  um conjunto finito de neurônios e  $\mathbf{X} := \{X_t^{(i)} : t \in \mathbb{Z} \text{ e } i \in I\}$  uma cadeia estocástica homogênea que assume valores em  $\{0, 1\}^{I \times \mathbb{Z}}$ , definida num espaço de probabilidade adequado  $(\Omega, \mathcal{A}, P)$ . Para cada neurônio  $i \in I$ , a cada instante de tempo  $t \in$

$\mathbb{Z}$ ,

$$X_t^{(i)} = \begin{cases} 1, & \text{se o neurônio } i \text{ dispara no instante de tempo } t, \\ 0, & \text{caso contrário.} \end{cases}$$

Nesse sentido, a configuração global dos neurônios no instante de tempo  $t \in \mathbb{Z}$  é denotada por  $\mathbf{X}_t := \{X_t^{(i)} : i \in I\}$  e o caminho do neurônio  $i \in I$  é denotado por  $\mathbf{X}^{(i)} := \{X_t^{(i)} : t \in \mathbb{Z}\}$ .

Define-se, para qualquer instante de tempo  $t \in \mathbb{Z}$ , a filtragem

$$\mathcal{F}_t = \sigma(\mathbf{X}_s, s \in \mathbb{Z}, s \leq t),$$

em outras palavras,  $\mathcal{F}_t$  é a sigma-álgebra gerada pelos eventos até o instante de tempo  $t$ . Dessa forma, pode-se introduzir a dinâmica do processo. Em cada instante de tempo  $t \in \mathbb{Z}$ , condicionando em todo o passado, os neurônios se atualizam de forma independente, isto é, o disparar ou não dos neurônios está condicionado ao passado da cadeia e não às ações que ocorrem naquele instante de tempo  $t$ . Matematicamente, isso significa que para qualquer subconjunto de neurônios  $J \subset I$ ,  $a_i \in \{0, 1\}$ ,  $i \in J$ , temos

$$P\left(\bigcap_{i \in J} \{X_t^{(i)} = a_i\} \middle| \mathcal{F}_{t-1}\right) = \prod_{i \in J} P(X_t^{(i)} = a_i \mid \mathcal{F}_{t-1}).$$

Além disso, para cada neurônio  $i \in I$  em cada instante de tempo  $t \in \mathbb{Z}$ , condicionalmente a todo passado, a probabilidade do neurônio  $i$  disparar é uma função do seu potencial de membrana acumulado até o instante de tempo  $t - 1$ . Em outras palavras, para cada neurônio  $i \in I$  e cada instante de tempo  $t \in \mathbb{Z}$ , temos

$$P\left(X_t^{(i)} = 1 \middle| \mathcal{F}_{t-1}\right) = \phi_i\left(V_{t-1}^{(i)}\right), \quad (3.13)$$

em que  $\mathbf{V} := \{V_t^{(i)} : t \in \mathbb{Z} \text{ e } i \in I\}$  é um processo estocástico também definido em  $(\Omega, \mathcal{A}, P)$  que toma valores em  $\mathbb{R}$  e tal que  $V_t^{(i)}$  é o potencial de membrana do neurônio  $i$  no instante de tempo  $t$ , e  $\phi_i : \mathbb{R} \rightarrow [0, 1]$  é uma função crescente chamada função taxa de disparo do neurônio  $i$ .

O potencial de membrana de um neurônio  $i \in I$  é afetado pelas ações de todos os outros neurônios da rede que estão interagindo com ele. Nesse sentido, a probabilidade do neurônio  $i \in I$  disparar, condicionalmente a todo passado da rede, cresce monotonicamente.

mente com o seu potencial de membrana. Quando um neurônio  $i \in I$  dispara, seu potencial de membrana reseta para o potencial de repouso. Quando um neurônio pré-sináptico  $j \in I$ ,  $j \neq i$ , dispara, o potencial de membrana do neurônio  $i$  cresce proporcionalmente ao seu peso sináptico  $W_{j \rightarrow i} \in \mathbb{R}$ . Em outras palavras,  $W_{j \rightarrow i}$  descreve o efeito que o disparo do neurônio pré-sináptico  $j$  tem sobre o neurônio  $i$ . Se  $W_{j \rightarrow i} > 0$ , então o efeito que o neurônio  $j$  tem sobre o neurônio  $i$  é excitatório, e se  $W_{j \rightarrow i} < 0$  esse efeito é inibitório. Além disso, assumimos que  $W_{j \rightarrow j} = 0$  para todo neurônio  $j \in I$ .

Assumindo essa descrição, para cada neurônio  $i \in I$  em cada instante de tempo  $t \in \mathbb{Z}$ , o modelo proposto por Galves e Löcherbach (2016) para modelar a conectividade neuronal é tal que

$$V_t^{(i)} := \sum_{j \in I} W_{j \rightarrow i} \sum_{s=L_t^{(i)}+1}^t g_i(t-s) X_s^{(j)}, \quad (3.14)$$

em que  $g_j : \mathbb{N} \rightarrow \mathbb{R}_+$  é uma função que descreve o efeito de vazamento no canal de vazamento da membrana do neurônio  $j$ , que é um canal iônico sem mecanismos de bloqueio que permite um fluxo de íons (entrada e saída) na membrana sem impedâncias, e

$$L_t^{(i)} := \sup \{s \leq t : X_s^{(i)} = 1\}$$

é o último instante de tempo em que se observou um disparo do neurônio  $i$  antes do instante de tempo  $t$ . Aqui, adotamos a convenção que  $L_t^{(i)} = -\infty$  sempre quando  $X_s^{(i)} = 0$  para todo  $s < t$ . Portanto,  $X_t^{(i)} = 1$  implica em  $V_t^{(i)} = 0$ . Nesse sentido, podemos reescrever o Modelo (3.13) da seguinte forma

$$P(X_t^{(i)} = 1 \mid \mathcal{F}_{t-1}) = \phi_i \left( \sum_{j \in I} W_{j \rightarrow i} \sum_{s=L_{t-1}^{(i)}+1}^{t-1} g_i(t-s) X_s^{(j)} \right), \quad (3.15)$$

qualquer que seja o neurônio  $i \in I$  e o instante de tempo  $t \in \mathbb{Z}$ . Observe que a contribuição do neurônio  $j \in I$  no potencial de membrana do neurônio  $i \in I$  é excitatório ou inibitório, dependendo do sinal do peso sináptico  $W_{j \rightarrow i}$ . Essa é uma situação observada em redes neurais biológicas em que os neurônios podem ser excitados ou inibidos pela ação dos outros.

**Exemplo 3.16** *Para compreender o sentido neurobiológico por trás do modelo de Galves e Löcherbach (2016), considere uma amostra  $\mathbf{x} := \{x_t^{(i)} : i \in \{1, 2, 3\} \text{ e } t \in \{1, 2, 3, 4, 5\}\}$*

da atividade de uma rede com três neurônios observados ao longo de cinco instantes de tempo, a qual é dada pela seguinte matriz

$$\mathbf{x} := \begin{array}{c} \begin{array}{ccccc} & 1 & 2 & 3 & 4 & 5 \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \end{array}$$

em que cada entrada indica se o neurônio da linha  $i$  disparou no instante de tempo  $t$ , sendo  $i$  e  $t$  números naturais tais que  $1 \leq i \leq 3$  e  $1 \leq t \leq 5$ .

Assim, tendo os instantes de tempo observados, podemos inicialmente descrever, por exemplo, o potencial de membrana do neurônio  $i = 1$  no instante  $t = 6$ , com base no passado da cadeia. Sabemos que o potencial de membrana do neurônio 1 no instante de tempo 6, com base na Equação (3.14) é dada por

$$V_6^{(1)} = \sum_{j \in I} W_{j \rightarrow 1} \sum_{s=L_i^{(1)}+1}^t g_1(t-s) X_s^{(j)}.$$

Utilizando a amostra observada  $\mathbf{x}$ , segue que o potencial de membrana observado é

$$\begin{aligned} V_6^{(1)} &= W_{2 \rightarrow 1} \cdot \{ [g_1(6-3) \cdot x_3^2] + [g_1(6-4) \cdot x_4^2] + [g_1(6-5) \cdot x_5^2] + \\ &\quad + [g_1(6-6) \cdot x_6^2] \} + W_{3 \rightarrow 1} \cdot \{ ([g_1(6-3) \cdot x_3^3] + [g_1(6-4) \cdot x_4^3] + \\ &\quad + [g_1(6-5) \cdot x_5^3] + [g_1(6-6) \cdot x_6^3]) \} \\ &= W_{2 \rightarrow 1} \cdot (\{g_1(3) \cdot 1\} + \{g_1(2) \cdot 1\} + \{g_1(1) \cdot 0\} + \{g_1(0) \cdot 0\}) + \\ &\quad + W_{3 \rightarrow 1} \cdot (\{g_1(3) \cdot 0\} + \{g_1(2) \cdot 0\} + \{g_1(1) \cdot 1\} + \{g_1(0) \cdot 0\}) \\ &= W_{2 \rightarrow 1} \cdot \{g_1(3) + g_1(2)\} + W_{3 \rightarrow 1} \cdot g_1(1) \end{aligned} \tag{3.17}$$

Desse modo, a Expressão (3.17) revela que o potencial de membrana de um neurônio  $i$  em um dado instante de tempo  $t$ , na realidade, pode ser descrito como a soma dos vazamentos de potencial de membrana nos instantes passados a esse tempo em que os neurônios vizinhos dispararam, multiplicados pelos respectivos pesos sinápticos da vizinhança.

Se assumirmos, para o modelo de Galves e Löcherbach, que  $\phi_i$  é uma função logística, para todo  $i \in I$ , e que certas suposições sobre os parâmetros e a função de vazamento sejam satisfeitas (ver Galves e Löcherbach (2016)), então a cadeia  $\mathbf{X}$  que modela a atividade



da rede de neurônios será estacionária, o que implica que a dinâmica da cadeia  $\mathbf{X}$  é totalmente caracterizada pelas probabilidades de transição especificadas em (3.15).

### 3.4 Grafo de conectividade

Sejam  $I$  um conjunto finito de neurônios e  $F \subset I$  uma região amostrável finita. As interações que ocorrem entre os neurônios de  $F$  podem ser representadas por um grafo, que é um conjunto de vértices interconectados por arestas.

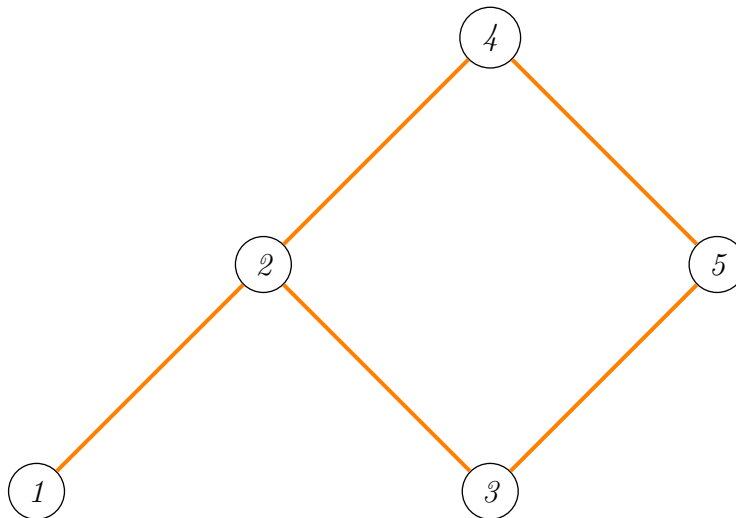
**Definição 3.18 (Grafo)** *Um grafo  $G$  é um par  $(V, E)$  em que  $V$  é um conjunto de vértices (objetos em estudo) e  $E$  é um conjunto de arestas (conexões entre os objetos em estudo), isto é,*

$$E := \{(i, j) : i, j \in V\}.$$

**Exemplo 3.19** *Considere uma região amostrável  $F = \{1, 2, 3, 4, 5\}$  com 5 neurônios, então  $F$  é o conjunto de vértices de um grafo  $G$ . Defina o conjunto  $E$  de relações entre esses neurônios da seguinte maneira*

$$E = \{(1, 2), (2, 3), (2, 4), (3, 5), (4, 5)\}.$$

Logo,  $E$  é o conjunto de arestas do grafo  $G$ . Nesse caso, o grafo  $G = (V, E)$  associado ao conjunto de neurônios  $F$  é não-orientado, isto é, não existe uma relação ordinária entre os vértices e, portanto, sua representação pictórica é dada da seguinte maneira:



Entretanto, para representar adequadamente a situação que estamos estudando, devemos utilizar um grafo orientado e com pesos, pois como vimos, as relações entre os

neurônios, que são caracterizadas pelos pesos sinápticos, podem ser excitatórias ou inibitórias, isto é, um neurônio pode contribuir para que outro dispare ou pode inibir o seu disparo, assim, os neurônios não possuem relações de mesmo peso com toda a rede.

**Definição 3.20 (Grafo orientado e com pesos)** *Um grafo  $G$  é um par  $(V, E)$  em que  $V$  é um conjunto de vértices e  $E$  é um conjunto de arestas sobre  $V$ . O grafo  $G$  é orientado quando o conjunto  $E$  de arestas define uma relação ordinária sobre o conjunto  $V$  de vértices. Quando a relação ordinária é impactada por características do problema de forma a fortalecer ou enfraquecer a conexão entre os vértices, pode-se atribuir pesos às arestas. Em outras palavras, podemos considerar uma matriz real  $W$  de ordem  $|V| \times |V|$  tal que o elemento  $W_{i \rightarrow j}$  que está na linha  $i$  e coluna  $j$  da matriz  $W$  representa o impacto que o vértice  $i$  possui sobre o vértice  $j$ , em que  $i, j \in V$ . Nesse caso,  $G$  é dito ser um grafo orientado com matriz de pesos  $W$ .*

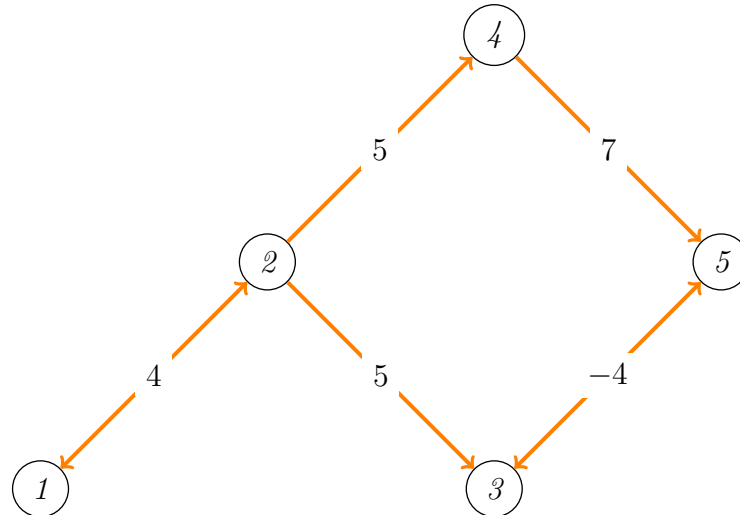
**Exemplo 3.21** *Considere uma região amostrável  $F$  com 5 neurônios, então  $F = \{1, 2, 3, 4, 5\}$  é o conjunto de vértices de um grafo  $G$ . Defina o conjunto  $E$  de relações entre os neurônios tal que*

$$E = \{(1, 2), (2, 3), (2, 4), (3, 5), (3, 4)\}.$$

*Logo,  $E$  é o conjunto de arestas do grafo  $G$ . Ao grafo  $G = (V, E)$ , associamos uma matriz  $W$  de pesos sinápticos dada por*

$$W := \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 0 & 4 & 0 & 0 & 0 \\ 4 & 0 & 5 & 5 & 0 \\ 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 7 \\ 0 & 0 & -4 & 0 & 0 \end{pmatrix} \end{matrix}$$

*em que  $W_{i \rightarrow j}$  representa o efeito dos neurônio pré-sináptico  $i$  sobre o neurônios pós-sináptico  $j$ , em que  $i, j \in F$ . Nesse caso,  $G = (F, E)$  é um grafo orientado com pesos e sua representação pictórica é dada da seguinte maneira:*



No Exemplo 3.21, podemos observar que os neurônios pré-sinápticos excitatórios são os neurônios 1, 2 e 4, sendo o neurônio 4 aquele com maior peso sináptico associado. Por outro lado, os neurônios pré-sinápticos inibitórios da rede são os neurônios 3 e 5 que inibem um ao outro com peso sináptico igual a  $-4$ .

Dada uma região amostrável finita  $F$  de neurônios com grafo de conectividade  $G = (F, E)$ , podemos definir para cada neurônio  $i \in F$  a sua vizinhança de interação, ou seja, o conjunto de todos os neurônios  $j \in F$ ,  $j \neq i$  conectados com o neurônio  $i$ .

**Definição 3.22** Dada uma região amostrável finita  $F$  de neurônios com grafo de conectividade  $G = (F, E)$  e matriz de pesos sinápticos  $W$ , a vizinhança de interação de um neurônio  $i \in F$  é o conjunto  $\mathcal{V}_i$  tal que

$$\mathcal{V}_i := \{j \in F \setminus \{i\} : W_{j \rightarrow i} \neq 0\}.$$

**Exemplo 3.23** Tome o grafo de conectividade do Exemplo 3.21. Considerando o conjunto de neurônios  $F = \{1, 2, 3, 4, 5\}$ , podemos determinar a vizinhança de interação de cada neurônio  $i \in F$  por meio do grafo de conectividade. De acordo com a Definição 3.22, a vizinhança de interação de um neurônio  $i \in F$  é formada por neurônios  $j \in F$  que excitam ou inibem o neurônio  $i$ , sendo  $i \neq j$ . Portanto, podemos determinar a vizinhança de interação dos neurônios pertencentes ao conjunto  $F$ , a partir dos pesos dos neurônios pré-sinápticos. Nesse sentido,

$$\mathcal{V}_1 = \{2\}; \mathcal{V}_2 = \{1\}; \mathcal{V}_3 = \{2, 5\}; \mathcal{V}_4 = \{2\}; \mathcal{V}_5 = \{3, 4\}.$$

### 3.5 Estimação do grafo de conectividade

Para inteiros positivos  $N$  e  $T$ , consideramos um conjunto finito  $I$  de neurônios e uma região amostrável finita  $F \subset I$  com  $N$  neurônios, observada ao longo de um intervalo de tempo, cuja discretização possui  $T$  janelas temporais. Assim, dada a configuração  $\mathbf{x}_1^T := \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\} \in \{0, 1\}^{N \times T}$  obtida a partir da cadeia estocástica  $\mathbf{X}$ , estamos interessados em estimar, para qualquer neurônio  $i \in F$ , a vizinhança de interação  $\mathcal{V}_i$ .

Sabemos que a vizinhança de interação é representada por um grafo de conectividade direcionado com pesos, e que ela é composta pelos neurônios  $j$  cujos pesos sinápticos  $W_{j \rightarrow i}$  são não nulos, sendo  $j \neq i$ . Então, estimar a vizinhança de interação é análogo a estimar o grafo de conectividade, que por sua vez é determinado pelos pesos sinápticos, ou seja, se estimarmos o efeito que os neurônios pré-sinápticos tem sobre os neurônios pós-sinápticos, podemos dizer se há conectividade entre eles ou não e, conseqüentemente, estaremos estimando a vizinhança de interação. Portanto, inferir a matriz de pesos sinápticos é análogo a estimar a vizinhança de interação (Ost e Reynaud-Bouret, 2020).

Nesse contexto, estamos interessados em estimar, a partir da amostra observada  $\mathbf{x}_1^T$ , a matriz de pesos sinápticos

$$W := \begin{pmatrix} W_{1 \rightarrow 1} & W_{1 \rightarrow 2} & \dots & W_{1 \rightarrow N} \\ W_{2 \rightarrow 1} & W_{2 \rightarrow 2} & \dots & W_{2 \rightarrow N} \\ \vdots & \vdots & & \vdots \\ W_{N \rightarrow 1} & W_{N \rightarrow 2} & \dots & W_{N \rightarrow N} \end{pmatrix}$$

Além disso, para a função taxa de disparo  $\phi_i$ , vamos considerar a função logística, visto que, é comumente utilizada em neurôciência computacional para fins de modelagem (Jolivet *et al.*, 2006) e, por ter o conjunto imagem no intervalo aberto  $]0, 1[$ , é plausível para descrever o comportamento de probabilidades. Assim, a função taxa de disparo é uma função  $\phi_i : \mathbb{R} \rightarrow ]0, 1[$  tal que

$$\phi_i(x) = \frac{\exp(x)}{1 + \exp(x)}, \quad (3.24)$$

para todo neurônio  $i \in I$ .

Desse modo, podemos reescrever a Equação (3.13) com base na função logística defi-

nida por (3.24), obtendo

$$P\left(X_t^{(i)} = 1 \mid \mathcal{F}_{t-1}\right) = \frac{\exp(V_{t-1}^{(i)})}{1 - \exp(V_{t-1}^{(i)})}. \quad (3.25)$$

Dessa forma, dada a configuração  $\mathbf{x}_1^T$ , podemos escrever, para cada neurônio  $i \in I$ , a função de log-verossimilhança associada ao vetor de pesos sinápticos

$$W^{(i)} = (W_{1 \rightarrow i}, W_{2 \rightarrow i}, \dots, W_{N \rightarrow i})^\top$$

da seguinte maneira

$$\begin{aligned} \ell(W^{(i)}, \mathbf{x}_1^T) &:= \sum_{t=1}^T \left[ V_{t-1}^{(i)} - \log\left(1 - \exp(V_{t-1}^{(i)})\right) \right] \\ &:= \sum_{t=1}^T \left[ \left( \sum_{j=1}^N W_{j \rightarrow i} \sum_{s=L_{t-1}^{(i)}+1}^{t-1} g_i(t-s)x_s^{(j)} \right) \right. \\ &\quad \left. - \log\left(1 - \exp\left( \sum_{j=1}^N W_{j \rightarrow i} \sum_{s=L_{t-1}^{(i)}+1}^{t-1} g_i(t-s)x_s^{(j)} \right)\right) \right]. \end{aligned}$$

Definida a função de log-verossimilhança do modelo, é possível realizar a estimação dos parâmetros. Para inferir sobre a conectividade da cadeia, como visto no Exemplo 3.16, são utilizadas como covariáveis as informações de disparo ou não disparo dos neurônios amostrados em cada instante de tempo e, portanto, as covariáveis desse modelo são os estados de cada neurônio em cada instante, ou seja, cada instante  $t$  da cadeia  $\mathbf{X}$  representa uma covariável. Desse modo, o número de covariáveis pode ser muito grande e, inclusive, maior que a quantidade de observações (neurônios) e, por essa razão, é interessante utilizar uma regularização sobre os coeficientes do modelo no processo de estimação a fim de obter um modelo esparsos.

Neste trabalho, vamos utilizar uma regularização  $\ell_1$  sobre os coeficientes do modelo, com o objetivo de selecionar as conexões neuronais que irão constituir o grafo de conectividade. Este método foi proposto por Tibshirani (1996) em um contexto de regressão linear e denominado *Least Absolute Shrinkage and Selection Operator* (LASSO). A técnica consiste em adotar os coeficientes do modelo por meio da resolução de um problema de otimização com uma restrição sobre os coeficientes dada pela norma  $\ell_1$ .

Nesse contexto, dada a amostra observada  $\mathbf{x}_1^T$ , o vetor de pesos sinápticos estimado  $\hat{W}^{(i)}$  através do método LASSO é definida como

$$\hat{W}^{(i)} := \arg \min_{W^{(i)}} \left( -\ell(W^{(i)}, \mathbf{x}_1^T) + \lambda \sum_{j=1}^N |W_{j \rightarrow i}| \right), \quad (3.26)$$

em que  $\ell(W^{(i)}, \mathbf{x}_1^T)$  é a função de log-verossimilhança calculada anteriormente,  $W_{j \rightarrow i}$  são os pesos sinápticos e  $\lambda$  é um parâmetro de penalização escolhido por validação-cruzada, buscando minimizar o risco de predição, ou seja, escolhemos um  $\lambda$  que gere equilíbrio entre viés e variância. Nesse contexto, o que estamos fazendo com a expressão (3.26) para determinar os estimadores é minimizar a log-verossimilhança negativada sob a norma  $\ell_1$  dos pesos sinápticos  $W_{j \rightarrow i}$  multiplicados pela constante  $\lambda$ . Especificamente, a constante  $\lambda$  tem a funcionalidade de ponderar a penalidade, quando é igual a zero o estimador será dado simplesmente pela minimização da log-verossimilhança negativa, ou seja, é o próprio estimador de máxima verossimilhança. Contudo, quando  $\lambda$  tende ao infinito, a expressão é dominada pela soma dos pesos sinápticos e, logo, para minimizar a expressão (3.26) os pesos sinápticos tendem a ser muito pequenos, próximos ou até iguais a zero. Portanto, há uma menor variância dos estimadores quando  $\lambda$  é grande, porém, um viés maior, visto que os pesos tendem a ser todos baixos, por outro lado, quando  $\lambda$  é pequeno, os estimadores podem ser não-viesados mas com variância significativa.

Assim, há uma incerteza sobre a consistência e resultados em relação ao estimador definido sobre a norma  $\ell_1$ , devido tanto a questão da definição da constante  $\lambda$  adequada, quanto por não haver trabalhos que avaliam o uso do LASSO para a estimação dos parâmetros do modelo de Galves e Löcherbach. Então, o objetivo do trabalho é, além de inferir o grafo de conectividade neuronal, avaliar, via simulações computacionais, a consistência do estimador LASSO da matriz de pesos sinápticos  $W$ .

# Capítulo 4

## Aplicação Computacional

Registros da atividade neuronal revelam que os disparos elétricos dos neurônios podem ocorrer de forma espontânea e irregular e sofrem variações mesmo quando o neurônio é exposto aos mesmos estímulos (Crochet *et al.*, 2011; Naud e Gerstner, 2012; Bair e Koch, 1996; Buesing *et al.*, 2011), consequência da existência de ruídos nos sinais elétricos. Essas observações empíricas sugerem uma estrutura estocásticas para a descrição matemática de fenômenos neurais. Nesse sentido, propomos, neste trabalho, a utilização do modelo de Galves e Löcherbach (2016) para descrever a atividade de um neurônio ao longo do tempo. Por meio desse modelo de neurônio, geramos computacionalmente uma rede de neurônios estocásticos a partir de suposições neurobiologicamente plausíveis, e propomos realizar a estimação do grafo de conectividade a partir de um procedimento do tipo LASSO.

Neste capítulo, avaliamos a performance do método de estimação proposto por meio de réplicas de simulações de Monte Carlo. Usamos tais simulações para comparar nossas estimativas com os valores dos parâmetros que foram fixados. Nesse contexto, obtivemos evidências a favor da consistência do estimador proposto e também do método de seleção de variáveis.

### 4.1 Hipóteses

Devido à existência de diversas hipóteses relacionadas à maneira como se dá o armazenamento de informação pelos neurônios, sobre a forma que eles a utilizam, e o impacto que estímulos externos possuem na habilidade de um neurônio trocar informação, vamos fixar, nesta seção, as restrições com base na neurobiologia que levamos em consideração para realizar o estudo computacional.

As simulações foram feitas em R, versão 8.16, e o código foi disponibilizado publicamente no github do autor <sup>1</sup> juntamente com a documentação para uso futuro da comunidade científica. Neste contexto, as amostras de disparos neuronais consideradas neste trabalho foram obtidas a partir da simulação do modelo de Galves e Löcherbach. Dada uma amostra, a estimação do grafo de conectividade neuronal foi realizada a partir do estimador de máxima verossimilhança com regularização  $\ell_1$  dos coeficientes. Contudo, considerar apenas uma amostra não nos fornece informação suficiente a respeito do desempenho do método de estimação proposto nesta monografia. Assim, utilizamos réplicas de simulação Monte Carlo para comparar nossas estimativas com os verdadeiros valores dos parâmetros e, assim, estudar o comportamento assintótico desses estimadores. A seguir, descrevemos em mais detalhes como cada etapa deste estudo de simulação foi realizada.

### 4.1.1 Geração da rede neuronal

O estudo da conectividade neuronal via simulação é uma viabilização prática para nos ajudar no entendimento de como se dão as conexões sinápticas, sem precisar de coletas amostrais de um organismo biológico. Desta forma, o primeiro passo é gerar uma amostra dos disparos neuronais de uma rede de neurônios.

Considerando a complexidade do algoritmo implementado, o qual realiza muitas operações matriciais, optamos por simular uma rede composta por apenas cinco neurônios. Simulamos a atividade dessa rede em cinco janelas temporais:  $T = 120$ ,  $T = 1.200$ ,  $T = 12.000$ ,  $T = 24.000$  e  $T = 48.000$ .

Para realizar a simulação dessa rede, foi utilizada, como função de ativação, a taxa de disparo logística  $\phi_i$ , que foi definida na Equação (3.24). Nesse sentido, independente da janela temporal, dado um neurônio  $i \in F$ ,  $F := \{1, 2, 3, 4, 5\}$ , geramos a  $t$ -ésima observação  $x_t^{(i)}$  do neurônio  $i$ ,  $t = 1, 2, \dots, T$ , da seguinte maneira:

$$x_t^{(i)} = \begin{cases} 1, & \text{se } u_t \leq \phi(v_{t-1}^{(i)}) , \\ 0, & \text{caso contrário,} \end{cases} \quad (4.1)$$

em que  $u_t$  é um número pseudo-aleatório gerado a partir de uma variável aleatória com

---

<sup>1</sup><https://github.com/Schiavone662/TG---Estima-o-do-grafo-de-conectividade-neuronal-.git>



distribuição uniforme no intervalo  $[0, 1]$  e  $v_{t-1}^{(i)}$  é o potencial de membrana do neurônio  $i$  no instante de tempo  $t - 1$ .

Nesse trabalho,  $v_t^{(i)}$  é dado de acordo com o modelo de [Galves e Löcherbach \(2016\)](#), isto é,

$$v_t^{(i)} := \sum_{j \in I} W_{j \rightarrow i} \sum_{s=L_t^{(i)}+1}^t g_i(t-s)x_s^{(j)}, \quad (4.2)$$

em que  $W_{i \rightarrow j}$  é o peso sináptico que o neurônio pré-sináptico  $i \in F$  exerce sobre o neurônio pós-sináptico  $j \in F$  e  $g_i : \mathbb{N} \rightarrow \mathbb{R}_+$  é a função que descreve o efeito de vazamento no canal iônico da membrana do neurônio  $i \in F$ . Para as simulações desta monografia, definimos

$$g_i(t-s) := \frac{1}{2^{t-L_t^{(i)}}}. \quad (4.3)$$

Do ponto de vista neurobiológico, optar por uma função de vazamento com essa expressão analítica, significa afirmar que o efeito do vazamento do canal iônico do neurônio  $i$  é maior quanto mais próximo o tempo do último disparo do neurônio  $i$  estiver do instante de tempo  $t$  ([Gerstner et al., 2014](#)).

Uma vez definida a função  $g_i$  a ser utilizada nas simulações, basta fixarmos a matriz de pesos sinápticos  $W$ . Para definir essa matriz, consideramos alguns fatores neurobiológicos conhecidos. Por exemplo, definimos que a rede a ser gerada será composta por 80% dos neurônios como sendo excitatórios e 20% como sendo inibitórios. Além disso, fixamos o peso sináptico que os neurônios inibitórios exercem sobre os outros neurônios como sendo quatro vezes mais forte que os pesos sinápticos dos neurônios excitatórios. Esse tipo de matriz de pesos sinápticos é comum no córtex cerebral ([Kandel et al., 2000](#)). Desse modo, definimos a matriz de pesos sinápticos  $W$  da seguinte maneira

$$W = \begin{pmatrix} 0 & 0 & 5 & 0 & 5 \\ 0 & 0 & 5 & 5 & 0 \\ 5 & 5 & 0 & 0 & 5 \\ 0 & 5 & 0 & 0 & 5 \\ -20 & 0 & -20 & -20 & 0 \end{pmatrix}.$$

O tempo médio que um neurônio leva para disparar após o seu último disparo depende do neurônio que estamos estudando. Sem perda de generalidade, vamos assumir que estamos trabalhando com neurônios corticais, os quais, em média, levam de 1 a 3

milissegundos (ms) para disparar (ver Softky Koch (1993)). Na prática, em geral, a discretização temporal utilizada no processo de amostragem das respostas temporais de um neurônio (disparar ou ficar em repouso) leva em consideração intervalos da ordem de 0,30ms. Esta é a ordem que utilizamos para a discretização temporal deste projeto. Assim, para iniciar o algoritmo de geração da cadeia, fixamos os quatro primeiros instantes de tempo de todos os neurônios da rede de forma que observamos pelo menos um disparo para cada neurônio. Note que é razoável considerarmos 4 instantes iniciais pois, em média, precisamos observar a atividade do neurônio em 4 intervalos de tempo de tamanho 0,30ms. Portanto, neste trabalho, a amostra da atividade de um dado neurônio é uma sequência binária  $x_{-3}^T$  tal que  $x_t = 1$  se, e somente se, observa-se um disparo no  $t$ -ésimo intervalo de 0,30ms da discretização, em que  $t = -3, -2, \dots, T$ .

**Algoritmo.** Pseudo algoritmo para a geração das cadeias neuronais.

- **Entrada:**  $x_{-3}^0$  (passado) e matriz  $W$  de pesos sinápticos.
- **Saída:** a amostra  $\mathbf{x}_i$  da atividade neuronal de todos os neurônios  $i$  da rede.

- Fixar o número  $N$  de neurônios.
- Fixar o tamanho  $T$  da amostra a ser gerada.
- Fixar uma sequência inicializadora  $x_{-3}^0$
- Fixar uma matriz de pesos sinápticos  $W$ .

**para**  $t \leftarrow 1$  até  $T$  **faça:**

**para**  $i \leftarrow 1$  até  $N$  **faça:**

Potencial  $\leftarrow \phi(V_t^{(i)})$

$U \leftarrow$  Número Pseudo-aleatório da Uniforme(0,1)

**Se**  $U \leq$  Potencial **então**

$x_t^{(i)} = 1$

**senão**

$x_t^{(i)} = 0$

**fim se**

**fim para**

**fim para**

### 4.1.2 Estimação do Grafo de Conectividade

Gerada a atividade da rede neuronal por meio do modelo de [Galves e Löcherbach \(2016\)](#), podemos realizar a estimação do grafo de conectividade neuronal, por meio da estimação da matriz de pesos sinápticos  $W$ . O método de estimação utilizado foi apresentado em detalhes na seção anterior, o qual estima a partir da função de verossimilhança com regularização  $\ell_1$  dos coeficientes do modelo. A inferência da conectividade neuronal é importante, pois nos ajuda a entender como o cérebro opera.

Para realizar a estimação do grafo de conectividade utilizamos a função `glmnet` do *R*. Dessa forma, para cada neurônio  $i \in F$  e cada tamanho amostral  $T$  fixado, foi necessário considerar o vetor de respostas, a matriz de covariáveis e o vetor de parâmetros dados, respectivamente, por

$$\mathbf{y}_T^{(i)} = \begin{bmatrix} x_1^{(i)} \\ x_2^{(i)} \\ \vdots \\ x_T^{(i-1)} \end{bmatrix}, \quad \mathbf{A}_T^{(i)} = \begin{bmatrix} a_1^{(1)} & a_1^{(2)} & \cdots & a_1^{(N)} \\ a_2^{(1)} & a_2^{(2)} & \cdots & a_2^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ a_T^{(1)} & a_T^{(2)} & \cdots & a_T^{(N)} \end{bmatrix} \quad \text{e} \quad \mathbf{W}_{\cdot \rightarrow i} = \begin{bmatrix} W_{1 \rightarrow i} \\ W_{2 \rightarrow i} \\ \vdots \\ W_{N \rightarrow i} \end{bmatrix},$$

em que

$$a_t^{(j)} = \sum_{s=L_t^{(i)}+1}^{t-1} \frac{1}{2^{t-L_t^{(i)}}} x_s^{(j)}$$

para  $j = 1, 2, \dots, N$  e  $t = 1, 2, \dots, T$ .

Note que para cada neurônio  $i \in F$  há uma matriz de covariáveis  $A_T^{(i)}$ , uma vez que  $a_t^{(j)}$  é diferente de acordo com o neurônio  $i$  que estamos considerando. Consequentemente, no caso da nossa rede com 5 neurônios, existem cinco diferentes matrizes de dados que serão utilizadas em cinco diferentes modelos independentes. Tais modelos são independentes, pois essa é uma das suposições do modelo de Galves e Löcherbach. Para realizar a estimação das cinco colunas  $\mathbf{W}_{\cdot \rightarrow i}$  da matriz  $\hat{W}$ , utilizamos, respectivamente, a matriz  $A_T^{(i)}$ . Dessa forma, a partir da amostra gerado, obtemos a matriz de pesos sinápticos estimada

$$\hat{W} = \begin{pmatrix} \hat{W}_{1 \rightarrow 1} & \hat{W}_{1 \rightarrow 2} & \hat{W}_{1 \rightarrow 3} & \hat{W}_{1 \rightarrow 4} & \hat{W}_{1 \rightarrow 5} \\ \hat{W}_{2 \rightarrow 1} & \hat{W}_{2 \rightarrow 2} & \hat{W}_{2 \rightarrow 3} & \hat{W}_{2 \rightarrow 4} & \hat{W}_{2 \rightarrow 5} \\ \hat{W}_{3 \rightarrow 1} & \hat{W}_{3 \rightarrow 2} & \hat{W}_{3 \rightarrow 3} & \hat{W}_{3 \rightarrow 4} & \hat{W}_{3 \rightarrow 5} \\ \hat{W}_{4 \rightarrow 1} & \hat{W}_{4 \rightarrow 2} & \hat{W}_{4 \rightarrow 3} & \hat{W}_{4 \rightarrow 4} & \hat{W}_{4 \rightarrow 5} \\ \hat{W}_{5 \rightarrow 1} & \hat{W}_{5 \rightarrow 2} & \hat{W}_{5 \rightarrow 3} & \hat{W}_{5 \rightarrow 4} & \hat{W}_{5 \rightarrow 5} \end{pmatrix}.$$

Outra consideração comum a todos tamanhos amostrais  $T$  fixados é a respeito da exclusão de 20% das amostras geradas. Realizamos esta ação, para dissipar a influência que os valores iniciais das cadeias possuem sobre a estacionariedade.

A seguir descrevemos os resultados obtidos com o estudo de simulação proposto.

## 4.2 Resultados

Nesta seção, são explorados os resultados do método de estimação proposto visando estudar o seu desempenho. Foram geradas 100 réplicas de cada tamanho amostral especificado na seção anterior e foram realizadas as estimações dos pesos sinápticos para cada réplica. Em seguida, para cada tamanho amostral, calculamos o erro quadrático médio empírico e as proporções de conexões sinápticas corretamente identificadas bem como a proporção de conexões excitatórias e inibitórias estimadas corretamente.

### 4.2.1 Estimação da matriz de pesos sinápticos

Parâmetros	Simulado	T=120	T=1200	T=12000	T=24000	T=48000
$W_{2 \rightarrow 1}$	0	9.9996	0.3670	0.0253	0.0177	0.0076
$W_{3 \rightarrow 1}$	5	27.0727	2.5154	0.2585	0.1727	0.1279
$W_{4 \rightarrow 1}$	0	18.8718	0.7610	0.0477	0.0295	0.0091
$W_{5 \rightarrow 1}$	-20	49.9191	4.9356	0.5718	0.3455	0.2711
$W_{1 \rightarrow 2}$	0	5.1192	0.1354	0.0141	0.0049	0.0027
$W_{3 \rightarrow 2}$	5	39.1353	0.6969	0.0698	0.0427	0.0372
$W_{4 \rightarrow 2}$	5	72.5528	0.8887	0.0729	0.0553	0.0465
$W_{5 \rightarrow 2}$	0	0.9658	0.0717	0.0058	0.0029	0.0014
$W_{1 \rightarrow 3}$	5	30.1170	1.7695	0.1948	0.1015	0.0687
$W_{2 \rightarrow 3}$	5	14.0979	1.1467	0.1296	0.0879	0.0620
$W_{4 \rightarrow 3}$	0	19.5080	0.9478	0.0581	0.0222	0.0121
$W_{5 \rightarrow 3}$	-20	87.5167	3.5163	0.4797	0.2679	0.2230
$W_{1 \rightarrow 4}$	0	8.4005	1.1818	0.0947	0.0376	0.0128
$W_{2 \rightarrow 4}$	5	17.2148	1.5866	0.1102	0.0804	0.0787
$W_{3 \rightarrow 4}$	0	10.9257	0.8219	0.0713	0.0381	0.0093
$W_{5 \rightarrow 4}$	-20	55.4463	4.1943	0.3586	0.2818	0.2495
$W_{1 \rightarrow 5}$	5	76.9120	1.1409	0.1136	0.0818	0.0468
$W_{2 \rightarrow 5}$	0	1.2327	0.1184	0.0110	0.0046	0.0022
$W_{3 \rightarrow 5}$	5	124.4655	0.8803	0.0986	0.0729	0.0422
$W_{4 \rightarrow 5}$	5	71.8773	0.7411	0.1099	0.0669	0.0416

Tabela 4.1: Erro quadrático médio empírico calculado a partir de cem estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, via método da máxima verossimilhança com LASSO, e considerando cinco diferentes tamanhos amostrais:  $T=120$ ,  $T=1200$ ,  $T=12000$ ,  $T=24000$  e  $T=48000$ .

Por meio da Tabela 4.1, podemos observar os erros quadráticos médios empíricos calculados a partir das 100 réplicas, considerando cada parâmetro da matriz de pesos sinápticos e os cinco tamanhos amostrais. Nota-se que, na medida em que os tamanhos amostrais crescem, o erro quadrático médio empírico das estimativas diminuem e tendem a se aproximar de zero, não havendo sequer um caso em que o erro não diminuiu conforme o tamanho amostral aumentou, o que é um indicativo de que o método de estimação proposto é consistente para o modelo de Galves e Löcherbach (2016).

Além disso, podemos observar por meio da Figura 4.1 que há um indicativo de que a convergência do erro quadrático médio empírico para zero tende a decair exponencialmente. Para ilustrar esse fato, utilizamos o erro quadrático médio do peso sináptico  $W_{3 \rightarrow 5}$ , pois foi o que apresentou o maior erro quadrático médio empírico quando  $T = 120$ .

Com essas análises, temos um indicativo de que, considerando este cenário utilizando nas simulações, o método de estimação proposto produz estimativas consistentes dos parâmetros que representam os pesos sinápticos no modelo de Galves e Löcherbah. Em outras palavras, temos evidências de que a seguinte afirmação seja verdadeira:

**Afirmção 4.4** *Seja  $I$  uma rede finita de neurônios e para cada neurônio  $i \in I$  em cada instante de tempo  $t \in \mathbb{Z}$ , condicionalmente a todo passado, considere que a probabilidade do neurônio  $i \in I$  disparar é dada pela dinâmica do modelo de Galves e Löcherbach, isto é,*

$$P(X_t^{(i)} = 1 \mid \mathcal{F}_{t-1}) = \phi_i \left( \sum_{j \in I} W_{j \rightarrow i} \sum_{s=L_{t-1}^{(i)}+1}^{t-1} g_i(t-s) X_s^{(j)} \right). \quad (4.5)$$

*Dada uma região amostrável finita  $F \subset I$ , com  $N$  neurônios, observada ao longo de um intervalo de tempo, cuja discretização possui  $T$  janelas temporais, de forma que  $\mathbf{x}_1^T := \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\} \in \{0, 1\}^{N \times T}$  é a configuração observada da cadeia  $\mathbf{X}$ , segue que para todo neurônio  $i, j \in F$ , se, para cada instante de tempo  $T$ , considerarmos  $\hat{W}_{j \rightarrow i}^{(T)}$  a estimativa de máxima verossimilhança com LASSO do peso sináptico  $W_{j \rightarrow i}$ , então*

$$\lim_{T \rightarrow \infty} E \left[ \left( \hat{W}_{j \rightarrow i}^{(T)} - W_{j \rightarrow i} \right)^2 \right] = 0.$$

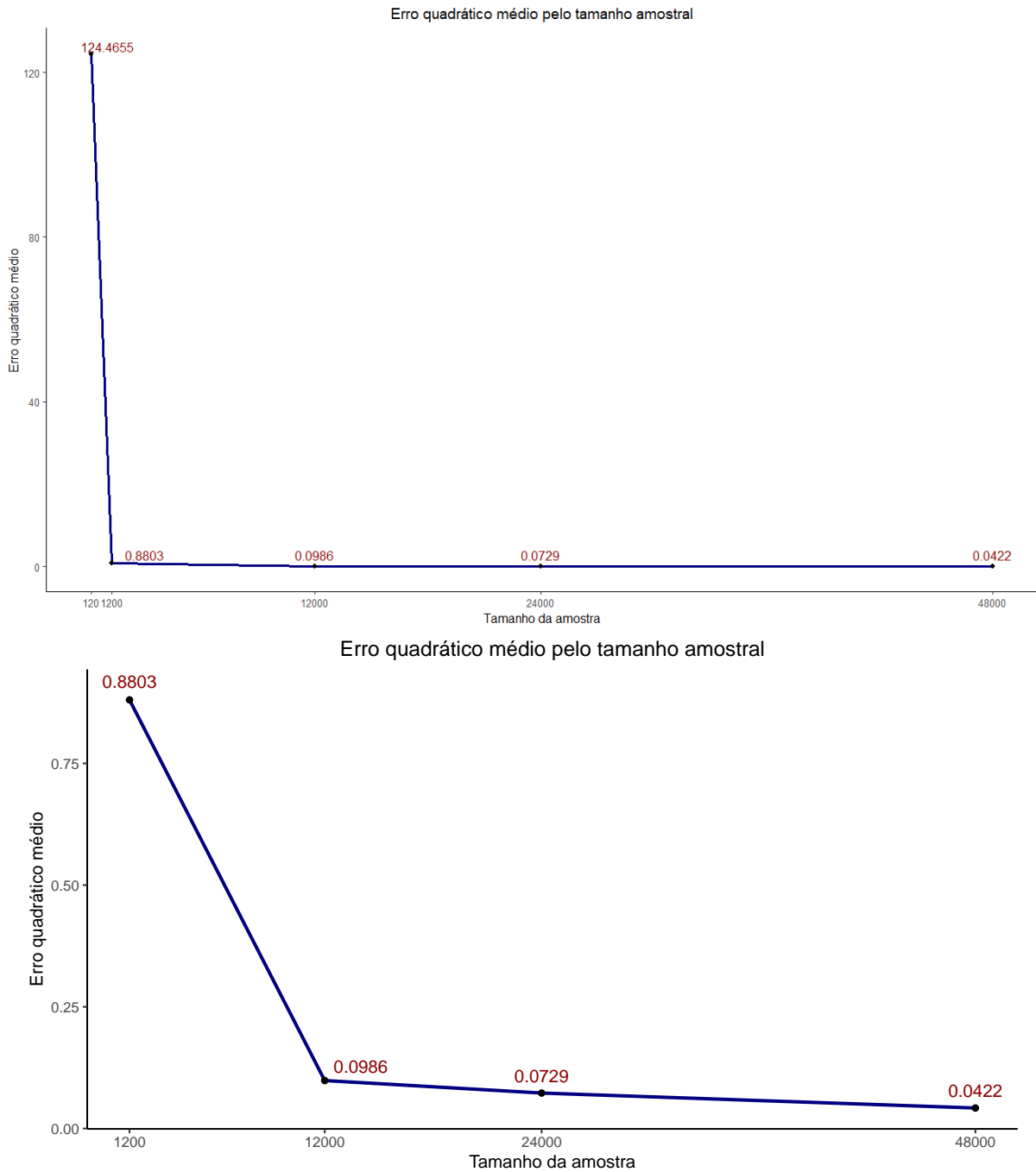


Figura 4.1: Erro quadrático médio empírico calculado a partir com estimativas do peso sináptico  $W_{3 \rightarrow 5}$  de uma rede com 5 neurônios, via método da máxima verossimilhança com LASSO, e considerando cinco diferentes tamanhos amostrais:  $T=120$ ,  $T=1200$ ,  $T=12000$ ,  $T=24000$  e  $T=48000$ .

## 4.2.2 Detecção de conexões sinápticas

Parâmetros	Simulado	T=120	T=1200	T=12000	T=24000	T=48000
$W_{2 \rightarrow 1}$	Desconectados	0.33	0.26	0.25	0.28	0.32
$W_{3 \rightarrow 1}$	Conectados	0.54	1.00	1.00	1.00	1.00
$W_{4 \rightarrow 1}$	Desconectados	0.51	0.28	0.27	0.32	0.52
$W_{5 \rightarrow 1}$	Conectados	1.00	1.00	1.00	1.00	1.00
$W_{1 \rightarrow 2}$	Desconectados	0.56	0.37	0.31	0.43	0.51
$W_{3 \rightarrow 2}$	Conectados	0.98	1.00	1.00	1.00	1.00
$W_{4 \rightarrow 2}$	Conectados	0.96	1.00	1.00	1.00	1.00
$W_{5 \rightarrow 2}$	Desconectados	0.45	0.39	0.38	0.37	0.54
$W_{1 \rightarrow 3}$	Conectados	0.70	1.00	1.00	1.00	1.00
$W_{2 \rightarrow 3}$	Conectados	0.80	1.00	1.00	1.00	1.00
$W_{4 \rightarrow 3}$	Desconectados	0.46	0.25	0.25	0.31	0.42
$W_{5 \rightarrow 3}$	Conectados	1.00	1.00	1.00	1.00	1.00
$W_{1 \rightarrow 4}$	Desconectados	0.48	0.35	0.25	0.36	0.48
$W_{2 \rightarrow 4}$	Conectados	0.66	1.00	1.00	1.00	1.00
$W_{3 \rightarrow 4}$	Desconectados	0.53	0.29	0.24	0.29	0.55
$W_{5 \rightarrow 4}$	Conectados	1.00	1.00	1.00	1.00	1.00
$W_{1 \rightarrow 5}$	Conectados	0.95	1.00	1.00	1.00	1.00
$W_{2 \rightarrow 5}$	Desconectados	0.30	0.31	0.22	0.26	0.32
$W_{3 \rightarrow 5}$	Conectados	0.96	1.00	1.00	1.00	1.00
$W_{4 \rightarrow 5}$	Conectados	0.98	1.00	1.00	1.00	1.00

Tabela 4.2: Proporção de conexões identificadas corretamente em cem estimativas de uma rede com 5 neurônios, via método da máxima verossimilhança com LASSO, e considerando cinco diferentes tamanhos amostrais:  $T=120$ ,  $T=1200$ ,  $T=12000$ ,  $T=24000$  e  $T=48000$ .

A Tabela 4.2 exibe a proporção de vezes em que a conexão sináptica foi identificada corretamente pelo método de estimação proposto. Nota-se que quando os neurônios estão conectados, conforme o tamanho da amostra cresce, as proporções de acertos também crescem. Entretanto, observa-se que quando os neurônios são desconectados, em alguns casos, essa proporção não seguiu esse comportamento, e isso ocorre porque para que os neurônios sejam considerados desconectados os parâmetros estimados devem ser exatamente iguais zero, ou seja, a estimativa do parâmetro deve ser exatamente igual ao valor do parâmetro gerado. Todavia, o que de fato ocorre são valores suficientemente próximos de zero. Apesar disso, a partir da amostra de tamanho 12.000 as proporções de conexões acertadas apresentaram um aumento e, devido ao fato de o erro quadrático médio se aproximar de zero, isto é, das estimativas se aproximarem do verdadeiro parâmetro, espera-se que essas proporções aumentem ainda mais na medida em que a amostra aumenta.

Com essas análises, temos um indicativo de que a seleção de variáveis do método de estimação proposto é consistente. Em outras palavras, temos evidências de que a seguinte

afirmação seja verdadeira:

**Afirmção 4.6** *Seja  $I$  uma rede finita de neurônios e para cada neurônio  $i \in I$  em cada instante de tempo  $t \in \mathbb{Z}$ , condicionalmente a todo passado, considere que a probabilidade do neurônio  $i \in I$  disparar é dada pela dinâmica do modelo de Galves e Löcherbach, isto é,*

$$P(X_t^{(i)} = 1 \mid \mathcal{F}_{t-1}) = \phi_i \left( \sum_{j \in I} W_{j \rightarrow i} \sum_{s=L_{t-1}^{(i)}+1}^{t-1} g_i(t-s) X_s^{(j)} \right). \quad (4.7)$$

*Dada uma região amostrável finita  $F \subset I$ , com  $N$  neurônios, observada ao longo de um intervalo de tempo, cuja discretização possui  $T$  janelas temporais, de forma que  $\mathbf{x}_1^T := \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\} \in \{0, 1\}^{N \times T}$  é a configuração observada da cadeia  $\mathbf{X}$ . Nesse contexto, para todo neurônio  $i, j \in F$ , se, para cada instante de tempo  $T$ , considerarmos  $\hat{W}_{j \rightarrow i}^{(T)}$  a estimativa de máxima verossimilhança com LASSO do peso sináptico  $W_{j \rightarrow i}$  e*

$$\hat{V}_i^{(T)} := \left\{ j \in F : \hat{W}_{j \rightarrow i}^{(T)} \neq 0 \right\}$$

*é a estimativa da vizinhança  $V_i$  do neurônio  $i \in F$ , então*

(a) *para todo neurônio  $j \in F - V_i$ , nós temos*

$$P\left(j \in \hat{V}_i^{(T)}\right) = \lim_{T \rightarrow \infty} P\left(\hat{W}_{j \rightarrow i}^{(T)} \neq 0\right) = 0.$$

(b) *para todo neurônio  $j \in V_i$ , nós temos*

$$P\left(j \notin \hat{V}_i^{(T)}\right) = \lim_{T \rightarrow \infty} P\left(\hat{W}_{j \rightarrow i}^{(T)} = 0\right) = 0.$$

Na Figura 4.2, podemos observar como se deu a estimação do grafo de conectividade ao longo dos diferentes tamanhos de amostras considerados neste trabalho. Para superar a dificuldade mencionada com relação à detecção correta de neurônios desconectados, definimos que se a estimativa  $\hat{W}_{j \rightarrow i}^{(T)}$  do peso sináptico entre dois neurônios desconectados  $i$  e  $j$  para o tamanho amostral  $T$  for suficientemente próxima de zero, então há evidências para afirmar que tais neurônios não possuem uma conexão significativa. Neste trabalho, definimos o nível de precisão da estimação como sendo 5%. Assim, se  $\left| \hat{W}_{j \rightarrow i}^{(T)} \right| < 0.05$ , então consideramos que os neurônios  $i$  e  $j$  estão desconectados.



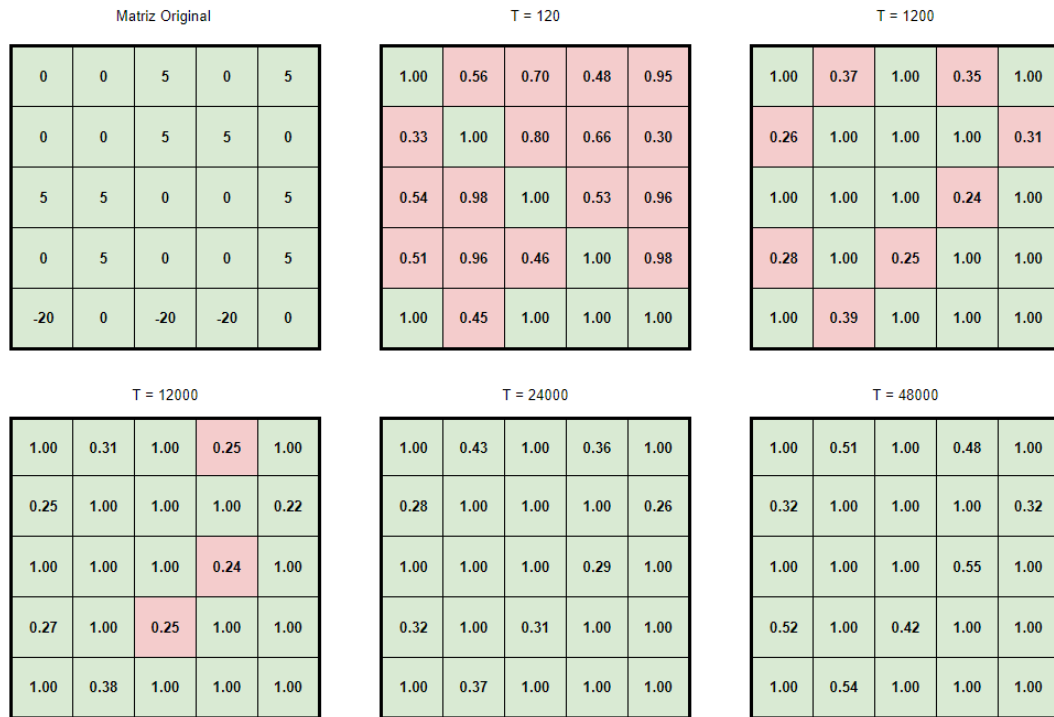


Figura 4.2: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão ou erro quadrático médio empírico menor que 0.05, para uma rede de 5 neurônios e cinco diferentes tamanhos de amostra:  $T = 120$ ,  $T = 1200$ ,  $T = 12000$ ,  $T = 24000$  e  $T = 48000$ . Código da cor: Verde - 100% na proporção de acertos da conexão ou  $|\hat{W}_{j \rightarrow i}^{(T)}| < 0.05$ , Vermelho - caso contrário.

Parâmetros	Simulado	T=120	T=1200	T=12000	T=24000	T=48000
$W_{3 \rightarrow 1}$	5	0.50	1.00	1.00	1.00	1.00
$W_{5 \rightarrow 1}$	-20	1.00	1.00	1.00	1.00	1.00
$W_{3 \rightarrow 2}$	5	0.98	1.00	1.00	1.00	1.00
$W_{4 \rightarrow 2}$	5	0.96	1.00	1.00	1.00	1.00
$W_{1 \rightarrow 3}$	5	0.65	1.00	1.00	1.00	1.00
$W_{2 \rightarrow 3}$	5	0.77	1.00	1.00	1.00	1.00
$W_{5 \rightarrow 3}$	-20	1.00	1.00	1.00	1.00	1.00
$W_{2 \rightarrow 4}$	5	0.65	1.00	1.00	1.00	1.00
$W_{5 \rightarrow 4}$	-20	1.00	1.00	1.00	1.00	1.00
$W_{1 \rightarrow 5}$	5	0.94	1.00	1.00	1.00	1.00
$W_{3 \rightarrow 5}$	5	0.96	1.00	1.00	1.00	1.00
$W_{4 \rightarrow 5}$	5	0.97	1.00	1.00	1.00	1.00

Tabela 4.3: Proporção de neurônios excitatórios ou inibitórios identificados corretamente em cem estimativas de uma rede com 5 neurônios, via método da máxima verossimilhança com LASSO, e considerando cinco diferentes tamanhos amostrais:  $T=120$ ,  $T=1200$ ,  $T=12000$ ,  $T=24000$  e  $T=48000$ .

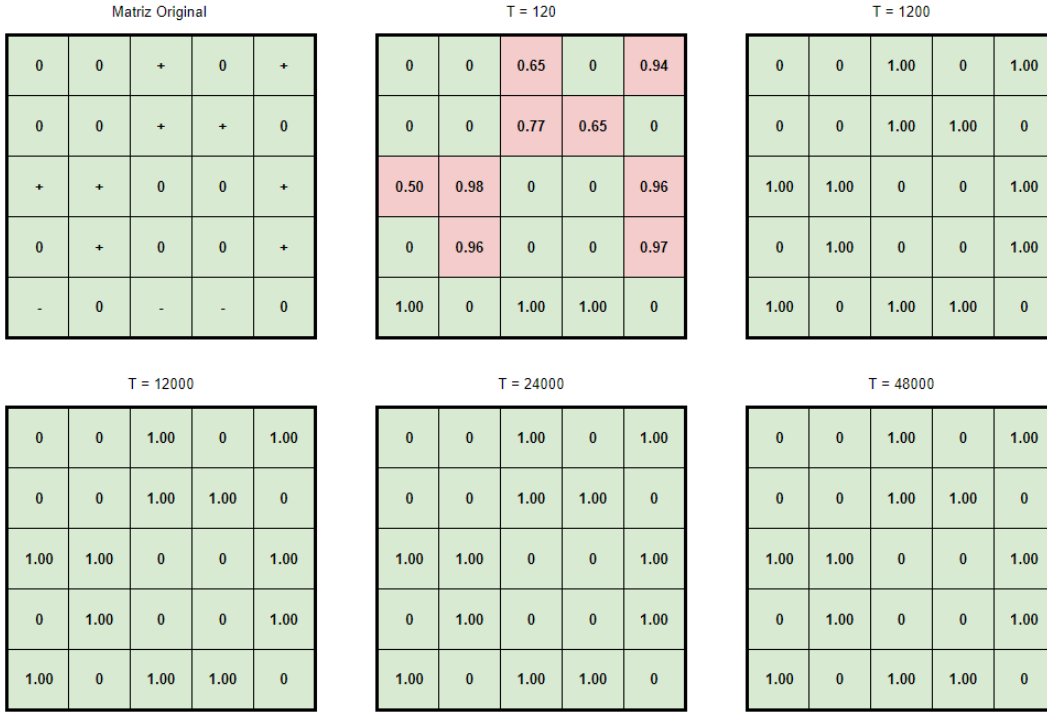


Figura 4.3: Ilustração do grafo de conectividade por meio da proporção de acertos dos tipos de conexão entre os neurônios, considerando apenas os realmente conectados, para uma rede de 5 neurônios e cinco diferentes tamanhos de amostra:  $T = 120$ ,  $T = 1200$ ,  $T = 12000$ ,  $T = 24000$  e  $T = 48000$ . Código da cor: Verde - 100% dos tipos de conexão estimadas corretamente, Vermelho - menos de 100% dos tipos de conexão estimadas corretamente.

Por fim, a Tabela 4.3 e a Figura 4.3 ilustram a proporção de vezes em que a estimativa acertou o tipo da conexão entre os neurônios, se eram excitatórias ou inibitórias. Nota-se que as proporções tendem a crescer conforme a amostra cresce, não havendo sequer um caso em que a proporção não aumentou conforme o tamanho amostral aumentou, o que é um indicativo de que o método de estimação proposto consegue identificar de forma consistente os tipos de conexões. Em outras palavras, temos indicativos de que a seguinte afirmação seja verdadeira:

**Afirmção 4.8** *Seja  $I$  uma rede finita de neurônios e para cada neurônio  $i \in I$  em cada instante de tempo  $t \in \mathbb{Z}$ , condicionalmente a todo passado, considere que a probabilidade do neurônio  $i \in I$  disparar é dada pela dinâmica do modelo de Galves e Löcherbach, isto é,*

$$P(X_t^{(i)} = 1 \mid \mathcal{F}_{t-1}) = \phi_i \left( \sum_{j \in I} W_{j \rightarrow i} \sum_{s=L_{t-1}^{(i)}+1}^{t-1} g_i(t-s) X_s^{(j)} \right). \quad (4.9)$$

Dada uma região amostrável finita  $F \subset I$ , com  $N$  neurônios, observada ao longo de um

intervalo de tempo, cuja discretização possui  $T$  janelas temporais, de forma que  $\mathbf{x}_1^T := \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\} \in \{0, 1\}^{N \times T}$  é a configuração observada da cadeia  $\mathbf{X}$ , segue que para todo neurônio  $i, j \in F$ , se, para cada instante de tempo  $T$ , considerarmos  $\hat{W}_{j \rightarrow i}^{(T)}$  a estimativa de máxima verossimilhança com LASSO do peso sináptico  $W_{j \rightarrow i}$ , então

$$\lim_{T \rightarrow \infty} P \left( \text{sgn}(\hat{W}_{j \rightarrow i}^{(T)}) \neq \text{sgn}(W_{j \rightarrow i}) \right) = 0,$$

em que  $\text{sgn}$  representa função sinal.

Até onde vai nosso conhecimento, não há na literatura trabalho algum que trata da consistência do método da máxima verossimilhança com LASSO para o modelo de Galves e Löcherbach. Portanto, tais evidência empíricas complementam, em um certo sentido, os estudos sobre a inferência da conectividade funcional entre neurônios.



# Capítulo 5

## Considerações Finais

Neste trabalho, estudamos o funcionamento do cérebro e, em particular, como as células que o compõe, os neurônios, realizam a transmissão e recepção de informações. Apesar de na maioria dos casos entendermos os mecanismos dos neurônios e os processos químicos e biológicos que ocorrem para que eles disparem individualmente, é de interesse dos neurocientistas compreender como neurônios se conectam e interagem entre si.

Nesse sentido, nesta monografia, utilizamos o modelo a tempo discreto introduzido por Galves e Löcherbach (2016) para modelar a atividade de um conjunto amostrável de neurônios. Propomos inferir o grafo de interação neuronal a partir da solução de um conjunto de equações não-lineares que satisfazem o critério de máxima verossimilhança com penalização sobre a norma  $\ell_1$  dos coeficientes; e, avaliar, via simulações computacionais, a consistência desse processo de estimação.

Assim, foram simuladas cem cadeias neuronais para diferentes tamanhos amostrais, considerando a forma como o modelo de Galves e Löcherbach (2016) descreve o comportamento neuronal e as características neurobiológicas dos neurônios e, em seguida, foram realizadas as estimativas dessas cadeias por meio do método de máxima verossimilhança com LASSO e analisados os resultados. Nesse contexto, obtivemos como resultados, evidências a favor da consistência do estimador proposto, isto é, o erro quadrático médio se aproxima de zero na medida em que o tamanho amostral cresce. Além disso, as simulações trazem também evidências de que o método de seleção de variável proposto é consistente, uma vez que a proporção de conexões identificadas corretamente aumenta na medida em que o tamanho amostral cresce.

Portanto, os próximos passos naturais deste trabalho são realizar a demonstração matemática das Afirmações 4.4, 4.6 e 4.8 e conduzir uma aplicação em dados reais obtidos

por meio de sinais eletrofisiológicos. Diante do que expomos nesta monografia, acreditamos que os trabalhos futuros mencionados são altamente exequíveis. Portanto, esperamos que, no futuro, tenhamos a oportunidade de conduzir tais estudos ou vê-los publicados.

# Referências Bibliográficas

- Bair, W. e Koch, C. (1996). Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Computation*, **8**(6), 1185–1202.
- Bear, M. F., Connors, B. W. e Paradiso, M. A. (2017). *Neurociência: desvendando o sistema nervoso*. Artmed Editora, fourth edition.
- Buesing, L., Bill, J., Nessler, B. e Maass, W. (2011). Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS computational biology*, **7**(11), e1002211.
- Bühlmann, P., Wyner, A. J. et al. (1999). Variable length Markov chains. *The Annals of Statistics*, **27**(2), 480–513.
- Cessac, B. (2010). A view of neural networks as dynamical systems. *International Journal of Bifurcation and Chaos*, **20**(06), 1585–1629.
- Cessac, B. (2011). A discrete time neural network model with spiking neurons: Ii: Dynamics with noise. *Journal of Mathematical Biology*, **62**(6), 863–900.
- Crochet, S., Poulet, J. F., Kremer, Y. e Petersen, C. C. (2011). Synaptic mechanisms underlying sparse coding of active touch. *Neuron*, **69**(6), 1160–1175.
- Ermentrout, B. e Terman, D. H. (2010). *Mathematical foundations of neuroscience*, volume 35. Springer.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, **11**(2), 127–138.
- Gallo, S. (2011). Chains with unbounded variable length memory: perfect simulation and a visible regeneration scheme. *Advances in Applied Probability*, **43**(3), 735–759.

- Galves, A. e Löcherbach, E. (2008). Stochastic chains with memory of variable length. Festschrift in honour of the 75th birthday of Jorma Rissanen.
- Galves, A. e Löcherbach, E. (2013). Infinite systems of interacting chains with memory of variable length—a stochastic model for biological neural nets. *Journal of Statistical Physics*, **151**(5), 896–921.
- Galves, A. e Löcherbach, E. (2016). Modeling networks of spiking neurons as interacting processes with memory of variable length. *Journal de la Société Française de Statistique*, **157**(1), 17–32.
- Gerstner, W., Kistler, W. M., Naud, R. e Paninski, L. (2014). *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press.
- Girardi, V. (2021). *Inferência da conectividade neuronal via estimação de medidas da teoria da informação*. Dissertação de mestrado, Universidade Federal de São Carlos, The address of the publisher. An optional note.
- Herculano-Houzel, S. (2009). The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in Human Neuroscience*, **3**, 31.
- Holt, G. R., Softky, W. R., Koch, C. e Douglas, R. J. (1996). Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *Journal of neurophysiology*, **75**(5), 1806–1814.
- Jolivet, R., Rauch, A., Lüscher, H.-R. e Gerstner, W. (2006). Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *Journal of Computational Neuroscience*, **21**(1), 35–49.
- Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S., Hudspeth, A. J. e Mack, S. (2000). *Principles of neural science*, volume 4. McGraw-hill New York.
- Koch, C. (2004). *Biophysics of computation: information processing in single neurons*. Oxford University Press.
- London, M., Shreibman, A. e Segev, I. (2002). Estimating information theoretic quantities of spike-trains using the context tree weighting algorithm. *Nature Neuroscience*, **5**.



- Moya, J. G. (2011). *Integration of the information in complex neural networks with noise*. Tese de doutorado, Universitat Politècnica de Catalunya. Escola Tècnica Superior d'Enginyeries . . . .
- Naud, R. e Gerstner, W. (2012). The performance (and limits) of simple neuron models: generalizations of the leaky integrate-and-fire model. Em *Computational Systems Neurobiology*, páginas 163–192. Springer.
- Ost, G. e Reynaud-Bouret, P. (2020). Sparse space–time models: Concentration inequalities and lasso. Em *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 56, páginas 2377–2405. Institut Henri Poincaré.
- Rissanen, J. (1983). A universal data compression system. *IEEE Transactions on Information Theory*, **29**(5), 656–664.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, **58**(1), 267–288.
- Truccolo, W., Hochberg, L. R. e Donoghue, J. P. (2010). Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes. *Nature Neuroscience*, **13**(1), 105–111.