

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

**ESTIMAÇÃO, SELEÇÃO E PREDIÇÃO DE
REDES DE NEURÔNIOS ESTOCÁSTICOS COM
MEMÓRIA DE ALCANCE VARIÁVEL**

Matheus Elias Pacola

Trabalho de Conclusão de Curso

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

ESTIMAÇÃO, SELEÇÃO E PREDIÇÃO DE REDES DE
NEURÔNIOS ESTOCÁSTICOS COM MEMÓRIA DE
ALCANCE VARIÁVEL

Matheus Elias Pacola

Orientador: Prof. Dr. Ricardo Felipe Ferreira

Trabalho de Conclusão de Curso apresentado
como parte dos requisitos para obtenção do
título de Bacharel em Estatística.

São Carlos
Fevereiro de 2024

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

ESTIMATION, SELECTION AND PREDICTION OF
NETWORKS WITH VARIABLE LENGHT MEMORY
STOCHASTIC NEURONS

Matheus Elias Pacola

Advisor: Prof. Dr. Ricardo Felipe Ferreira

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
February 2024

Matheus Elias Pacola

ESTIMAÇÃO, SELEÇÃO E PREDIÇÃO DE REDES DE
NEURÔNIOS ESTOCÁSTICOS COM MEMÓRIA DE
ALCANCE VARIÁVEL

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

Aprovado em 02 de fevereiro de 2024.

Banca Examinadora:

- Prof. Dr. Ricardo Felipe Ferreira
- Profa. Dra. Andressa Cerqueira
- Prof. Dr. Renato Jacob Gava

Aos meus pais, fonte de inspiração e apoio incondicional.

Agradecimentos

Agradeço à minha família, aos amigos que estiveram comigo em cada desafio e celebração, e em particular ao meu orientador, o Prof. Ricardo, cuja experiência e orientação foram fundamentais para este trabalho. Não posso deixar de agradecer à minha companheira Julia pelo seu apoio e compreensão ao longo desta jornada. Além disso, um carinho especial ao Paco, meu gato, que passou tantos dias ao meu lado, tornando os momentos de trabalho mais acolhedores com sua presença.

“A educação é a arma mais poderosa que você pode usar para mudar o mundo.”
(Paulo Freire)

Resumo

O sistema nervoso é submetido a diversas fontes de ruído. Registros da atividade neuronal revelam que parte desses ruídos devem-se aos disparos elétricos ocorrerem de forma espontânea e irregular e sofrerem variações mesmo quando o neurônio é exposto aos mesmos estímulos. Essas observações empíricas sugerem uma estrutura probabilística para a descrição matemática e o tratamento de fenômenos neurais. 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. Essa atividade é afetada pelas ações de todos os outros neurônios que estão interagindo com ele. Neste trabalho, a probabilidade de cada neurônio disparar condicionada a atividade passada da rede é tão maior quanto mais longe estiver o último disparo desse neurônio no passado em questão. Dessa forma, os neurônios das redes que pretendemos estudar possuem uma atividade estocástica com memória de alcance variável. Uma das questões mais importantes em neurociência é identificar as conexões que definem o circuito neural e o funcionamento do cérebro. Nesta monografia, vamos estimar a conectividade neuronal a partir da estimação dos parâmetros do modelo subjacente pelo método de máxima verossimilhança com e sem regularização dos parâmetros e, ainda, pelo método de máxima verossimilhança em conjunto com uma nova metodologia proposta, baseada na distância euclidiana entre modelos. A partir dos modelos ajustados, conduzimos um estudo de simulação para avaliar suas performances na predição futura da rede, na estimação da matriz de pesos sinápticos e no grafo de conectividade. Além disso, realizamos uma análise comparativa para avaliar como os diferentes métodos de seleção de variáveis abordados afetam o desempenho do modelo de rede neural de alcance variável na previsão do comportamento futuro de uma rede real, a qual foi obtida por meio de registros eletrofisiológicos.

Palavras-chave: *Cadeias estocásticas de memória de alcance variável, estimação do grafo de conectividade neuronal, predição da atividade neuronal, sistema estocástico de neurônios interagentes, estimação da matriz de pesos sinápticos, grafo de conectividade.*

Abstract

The nervous system is subjected to various sources of noise. Neuronal activity recordings reveal that part of this noise is due to electrical firings occurring spontaneously and irregularly, with variations even when the neuron is exposed to the same stimuli. These empirical observations suggest a probabilistic structure for the mathematical description and treatment of neural phenomena. In this work, the activity of each neuron is represented by a discrete-time stochastic process, whose random variables indicate whether or not a firing occurred at a given time instant. This activity is affected by the actions of all other neurons interacting with it. In this work, the probability of each neuron firing conditioned on the past activity of the network is higher the farther away the last firing of that neuron was in the past under consideration. Thus, the neurons in the networks we intend to study exhibit stochastic activity with variable-range memory. One of the most important questions in neuroscience is to identify the connections that define the neural circuit and brain function. In this monograph, we will estimate neuronal connectivity from the estimation of the parameters of the underlying model by maximum likelihood method with and without regularization of the parameters, and also by maximum likelihood method together with a new proposed methodology based on the Euclidean distance between models. From the fitted models, we conducted a simulation study to evaluate their performances in the future prediction of the network, in the estimation of the synaptic weights matrix, and in the connectivity graph. Additionally, we performed a comparative analysis to assess how the different variable selection methods addressed affect the performance of the variable-range neural network model in predicting the future behavior of a real network, which was obtained through electrophysiological recordings.

Keywords: *Stochastic variable-range memory chains, estimation of neuronal connectivity graph, prediction of neuronal activity, stochastic system of interacting neurons, estimation of synaptic weights matrix, connectivity graph.*

Lista de Figuras

2.1	Representação das estruturas das árvores definidas no exemplo.	51
2.2	Representação da estrutura da árvore τ_4 truncada no nível 3.	53
3.1	Representação pictórica do grafo $G = (F, E)$, em que $F = \{1, 2, 3, 4\}$ e $E = \{(1, 2), (1, 4), (2, 4), (3, 4)\}$	63
3.2	Representação pictórica do grafo orientado com pesos $G = (F, E)$, em que $F = \{1, 2, 3, 4\}$ e $E = \{(1, 2), (1, 4), (2, 4), (3, 4)\}$ e matriz de pesos W	65
4.1	Matriz de pesos sinápticos W para o cenário 1.	75
4.2	Matriz de pesos sinápticos W para o cenário 2.	75
4.3	Matriz de pesos sinápticos W para o cenário 3.	76
4.4	Matriz de pesos sinápticos W para o cenário 4.	76
4.5	Matriz de pesos sinápticos W para o cenário 5.	77
4.6	Matriz de pesos sinápticos W para o cenário 6.	77
A.1	Gráfico dos valores do erro quadrático médio empírico no cenário 1 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.	103
A.2	Gráfico dos valores do erro quadrático médio empírico no cenário 1 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.	103
A.3	Gráfico dos valores do erro quadrático médio empírico no cenário 1 em relação ao tamanho da amostra para o método de estimação proposto. . . .	104
A.4	Gráfico dos valores do erro quadrático médio empírico no cenário 2 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.	106

A.5	Gráfico dos valores do erro quadrático médio empírico no cenário 2 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.	106
A.6	Gráfico dos valores do erro quadrático médio empírico no cenário 2 em relação ao tamanho da amostra para o método de estimação proposto. . . .	107
A.7	Gráfico dos valores do erro quadrático médio empírico no cenário 3 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.	109
A.8	Gráfico dos valores do erro quadrático médio empírico no cenário 3 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.	109
A.9	Gráfico dos valores do erro quadrático médio empírico no cenário 3 em relação ao tamanho da amostra para o método de estimação proposto. . . .	110
A.10	Gráfico dos valores do erro quadrático médio empírico no cenário 4 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.	112
A.11	Gráfico dos valores do erro quadrático médio empírico no cenário 4 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.	112
A.12	Gráfico dos valores do erro quadrático médio empírico no cenário 4 em relação ao tamanho da amostra para o método de estimação proposto. . . .	113
A.13	Gráfico dos valores do erro quadrático médio empírico no cenário 5 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.	115
A.14	Gráfico dos valores do erro quadrático médio empírico no cenário 5 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.	115
A.15	Gráfico dos valores do erro quadrático médio empírico no cenário 5 em relação ao tamanho da amostra para o método de estimação proposto. . . .	116
A.16	Gráfico dos valores do erro quadrático médio empírico no cenário 6 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.	118

A.17	Gráfico dos valores do erro quadrático médio empírico no cenário 6 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.	118
A.18	Gráfico dos valores do erro quadrático médio empírico no cenário 6 em relação ao tamanho da amostra para o método de estimação proposto. . . .	119
B.1	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 4 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$	121
B.2	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 1 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$	122
B.3	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 1 para o método de do grafo de conectividade inspirado proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$	122
B.4	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 2 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$	123
B.5	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 2 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$	123

B.6	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 2 para o método de do grafo de conectividade inspirado proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.	124
B.7	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 3 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.	124
B.8	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 3 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.	125
B.9	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 3 para o método de do grafo de conectividade proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.	125
B.10	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 4 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.	126
B.11	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 4 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.	126
B.12	Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 4 para o método de do grafo de conectividade inspirado proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.	127

B.13 Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 5 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$	127
B.14 Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 5 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$	128
B.15 Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 5 para o método de do grafo de conectividade inspirado proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$. . .	128
B.16 Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 6 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$	129
B.17 Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 6 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$	129
B.18 Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 6 para o método de do grafo de conectividade inspirado proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$. . .	130

Lista de Tabelas

5.1	Proporção dos tipos de conexões sinápticas detectadas por cada método. . .	91
5.2	Métricas da matriz de confusão obtida na previsão do comportamento da rede no banco de dados reais com 34 neurônios monitorados por EEG. . . .	92
A.1	Erro quadrático médio empírico calculado para o cenário 1 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	101
A.2	Erro quadrático médio empírico calculado para o cenário 1 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	102
A.3	Erro quadrático médio empírico calculado para o cenário 1 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	102
A.4	Erro quadrático médio empírico calculado para o cenário 2 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	104

A.5	Erro quadrático médio empírico calculado para o cenário 2 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: T=120, T=1200, T=12000, T=24000.	105
A.6	Erro quadrático médio empírico calculado para o cenário 2 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: T=120, T=1200, T=12000, T=24000.	105
A.7	Erro quadrático médio empírico calculado para o cenário 3 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: T=120, T=1200, T=12000, T=24000.	107
A.8	Erro quadrático médio empírico calculado para o cenário 3 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: T=120, T=1200, T=12000, T=24000.	108
A.9	Erro quadrático médio empírico calculado para o cenário 3 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: T=120, T=1200, T=12000, T=24000.	108
A.10	Erro quadrático médio empírico calculado para o cenário 4 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: T=120, T=1200, T=12000, T=24000.	110

A.11	Erro quadrático médio empírico calculado para o cenário 4 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	111
A.12	Erro quadrático médio empírico calculado para o cenário 4 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	111
A.13	Erro quadrático médio empírico calculado para o cenário 5 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	113
A.14	Erro quadrático médio empírico calculado para o cenário 5 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	114
A.15	Erro quadrático médio empírico calculado para o cenário 5 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	114
A.16	Erro quadrático médio empírico calculado para o cenário 6 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	116

A.17	Erro quadrático médio empírico calculado para o cenário 6 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	117
A.18	Erro quadrático médio empírico calculado para o cenário 6 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$	117
A.19	Distância euclidiana média entre a matriz estimada e a matriz de pesos sinápticos original tamanho de T nos métodos abordados para o cenário 7.	119
B.1	Média de acertos de conexões sinápticas nas 100 réplicas simuladas no cenário 7 por todos os métodos com medida de precisão $\epsilon = 0, 2$	130
C.1	Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 1 para todos os métodos empregados no estudo com variações da proporção de treino e teste.	131
C.2	Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 2 para todos os métodos empregados no estudo com variações da proporção de treino e teste.	132
C.3	Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 3 para todos os métodos empregados no estudo com variações da proporção de treino e teste.	132
C.4	Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 4 para todos os métodos empregados no estudo com variações da proporção de treino e teste.	133
C.5	Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 5 para todos os métodos empregados no estudo com variações da proporção de treino e teste.	133
C.6	Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 6 para todos os métodos empregados no estudo com variações da proporção de treino e teste.	134

C.7	Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 7 para todos os métodos empregados no estudo com variações da proporção de treino e teste.	134
-----	--	-----

Sumário

1	Introdução	33
2	Modelagem da atividade neuronal	37
2.1	Notações, definições e conceitos preliminares.	37
2.2	Cadeias de Markov	40
2.2.1	Cadeias de Markov de ordem 1	40
2.2.2	Cadeias de Markov de ordem k	45
2.3	Cadeias com memória de alcance variável	49
3	Sistemas de neurônios estocásticos interagentes	57
3.1	Rede de neurônios	57
3.2	Modelo para neurônios estocásticos interagentes	59
3.3	Grafo de conectividade	62
3.4	Estimação dos pesos sinápticos e do grafo de conectividade	66
3.4.1	Método da máxima verossimilhança com e sem regularização dos parâmetros	67
3.4.2	Método proposto	70
3.5	Predição da atividade neuronal	71
4	Aplicação Computacional	73
4.1	Hipóteses e cenários	74
4.2	Simulação	78
4.2.1	Geração da atividade neuronal	78
4.2.2	Estimação dos pesos sinápticos e do grafo de conectividade	80
4.2.3	Predição da atividade neuronal	82
4.3	Resultados	83

5	Aplicação em dados eletrofisiológicos	89
5.1	Pré-processamento dos dados	89
5.2	Estimação dos pesos sinápticos	90
5.3	Estimação do grafo de conectividade	91
5.4	Previsão da atividade da rede	91
6	Considerações finais	95
	Referências Bibliográficas	97
A	Tabelas e gráficos com os resultados da estimação da matriz de pesos sinápticos	101
A.1	Cenário 1	101
A.2	Cenário 2	104
A.3	Cenário 3	107
A.4	Cenário 4	110
A.5	Cenário 5	113
A.6	Cenário 6	116
A.7	Cenário 7	119
B	Gráficos com os resultados da estimação do grafo de conectividade	121
B.1	Cenário 1	121
B.2	Cenário 2	123
B.3	Cenário 3	124
B.4	Cenário 4	126
B.5	Cenário 5	127
B.6	Cenário 6	129
B.7	Cenário 7	130
C	Tabelas com os resultados da predição da atividade da rede neuronal	131
C.1	Cenário 1	131
C.2	Cenário 2	132
C.3	Cenário 3	132
C.4	Cenário 4	133
C.5	Cenário 5	133

C.6 Cenário 6	134
C.7 Cenário 7	134

Capítulo 1

Introdução

Neurociência é um campo de pesquisa relativamente novo, que teve origem quando [y Cajal \(1888\)](#) propôs o papel dos neurônios como componente fundamental do sistema nervoso. O sistema nervoso representa uma rede de comunicação do organismo. O cérebro é o principal e mais complexo órgão desse sistema, o qual é composto por bilhões de neurônios, que são células eletricamente excitáveis especializadas na troca de informação. A comunicação entre os neurônios ocorre por meio das sinapses, as quais podem ser elétricas ou químicas. Nas sinapses elétricas, ocorre um fluxo direto de íons entre as células, enquanto que nas sinapses químicas, após o disparar de um neurônio (pré-sináptico) um neurotransmissor é liberado e se liga a uma proteína em uma célula receptora (pós-sináptica) fazendo com que uma despolarização localizada no seu potencial de membrana ocorra. O potencial de membrana é a diferença de potencial elétrico entre o meio intracelular e o meio extracelular, onde se destacam dois estados: (i) o potencial de repouso, no qual essa diferença se mantém aproximadamente constante devido a uma estabilização do fluxo iônico e (ii) o potencial de ação, no qual um estímulo (externo ou sináptico) provoca uma variação brusca nessa diferença de potencial elétrico, levando a uma descarga elétrica. Em neurociência, os potenciais de ação são conhecidos como impulsos nervosos ou disparos neuronais e a sequência de potenciais de ação geradas por um neurônio ao longo do tempo é chamada de trem de disparos neuronal. A atividade neural 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](#)).

A primeira descrição matemática do potencial de membrana neuronal deve-se ao trabalho de [Hodgkin e Huxley \(1952\)](#) que além de descreverem matematicamente o meca-

nismo envolvido na geração de potenciais de ação em neurônios, trouxeram uma simulação computacional dessas equações matemáticas sendo a primeira reprodução *in silico* do potencial de membrana. Desde então muitos modelos matemáticos foram desenvolvidos para descrever a dinâmica neuronal (Gerstner e Brette, 2009; Galves e Löcherbach, 2013).

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). Esses dados experimentais sugerem que os neurônios, as sinapses e o sistema neural são inerentemente estocásticos (Buesing *et al.*, 2011). Portanto, a descrição matemática dos fenômenos neurais pode ser tratada em termos probabilísticos. Nesse sentido, a evolução da atividade neuronal ao longo do tempo pode ser modelada por um processo estocástico.

Um neurônio dispara quando seu potencial de membrana ultrapassa um certo limiar. Em seguida, ele retorna ao seu potencial de repouso e se mantém nesse estado até que outro desequilíbrio eletrolítico ocorra. Dessa forma, após seu último disparo, o neurônio acumula potencial de membrana até que este seja suficientemente grande para permitir um disparo. Assim, é razoável considerar que a probabilidade de um neurônio disparar condicionada ao conhecimento da sua resposta temporal passada é tão maior quanto mais longe estiver o último disparo desse neurônio no passado em questão. Isso implica que o processo estocástico que modela a atividade desse neurônio não possui uma memória Markoviana de ordem finita, como mostra alguns trabalhos da literatura (Friston, 2010; Truccolo *et al.*, 2010; Cessac, 2010). A atividade de um neurônio é, portanto, modelada por um processo estocástico cuja dependência com o passado é de alcance variável.

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) e o termo cadeias estocásticas com memória de alcance variável deve-se a Galves e Löcherbach (2008). As probabilidades de transição desse tipo de processo estocástico não dependem de uma quantidade fixa de estados passados, levando em conta, por outro lado, a estrutura de dependência presente nos dados. Assim, podemos representar o conjunto das sequências de estados passados como uma árvore, denominada árvore de contexto. Considerando uma memória de tamanho variável temos modelos mais informativos, flexíveis e parcimoniosos do que as cadeias de Markov.

Entender como o sistema neural integra, codifica, transmite e processa informação é

fundamental para entender o funcionamento do cérebro. Nesse sentido, identificar as conexões que definem o circuito neural é essencial para explicar como tais representações são produzidas, prever como a rede se comportará em novas situações e extrair algoritmos baseados no comportamento do cérebro para aplicações técnicas (Sporns *et al.*, 2005). Para prever o comportamento da rede neuronal a partir de uma amostra de registros históricos, uma abordagem comum é primeiro utilizar dados de treinamento para ajustar um modelo, e então utilizar as probabilidades condicionais do modelo para prever a atividade neuronal.

Neste trabalho, consideramos um conjunto finito de neurônios cuja atividade é modelada por um modelo estatístico a tempo discreto no qual a probabilidade de um neurônio disparar em um determinado instante de tempo dada a configuração passada da rede possui memória de alcance variável. Para isso, realizamos a estimação do grafo de conectividade neuronal pelo método de máxima verossimilhança com e sem regularização ℓ_1 dos parâmetros e pelo método de máxima verossimilhança junto à uma nova metodologia de seleção de variáveis que iremos propor baseada na distância euclidiana entre modelos. Em um primeiro momento, simulamos amostras para o ajuste dos modelos e, a partir disso, avaliamos consistência da performance dos ajustes na predição do comportamento futuro da rede e na estimação da matriz de pesos sinápticos e grafo de conectividade. Esse estudo de simulação foi conduzido considerando diferentes tipos de rede com diferentes tamanhos de amostra e proporção entre treino e teste. Além disso, o poder preditivo dos modelos gerados por cada um também foram comparados, bem como a qualidade das estimativas.

A avaliação da performance do método de estimação proposto foi realizada por meio de réplicas de simulações de Monte Carlo. Utilizamos essas simulações para comparar nossas estimativas com os valores dos parâmetros fixados por meio do erro quadrático médio empírico. Nesse contexto, evidenciamos a consistência do estimador proposto, assim como a eficácia dos métodos de seleção de variáveis. Por outro lado, avaliamos a performance dos preditores obtidos a partir do ajuste do modelo proposto sob os diferentes contextos de seleção de variáveis por meio de medidas obtidas a partir da matriz de confusão tais como o coeficiente de correlação de Matthews (MCC), a G-Média, a sensibilidade, a especificidade, o valor preditivo positivo (*VPP*) e o valor preditivo negativo (*VPN*). Os resultados obtidos indicam que, em cenários de microcircuitos, o método proposto se sobressai em termos de precisão na previsão de valores, visto que inclui seleção de

variáveis, porém sem introduzir viés nas estimativas. Por outro lado, em macrocircuitos, sua performance é comparável à abordagem de máxima verossimilhança com regularização ℓ_1 dos parâmetros.

Por fim, também aplicamos as três metodologias discutidas a banco de dados reais que contém informações obtidas a partir de um monitoramento EEG em diversas partes do cérebro. Ao final, comparamos os resultados na predição do comportamento da rede e discutimos a robustez e eficácia do método proposto nesse contexto de experimentos não simulados.

Este trabalho está estruturado em quatro capítulos. No Capítulo 2, apresentaremos uma abordagem detalhada dos conceitos relacionados à modelagem da atividade neuronal a partir de processos estocásticos. Esse capítulo se desdobra em três seções distintas: a primeira estabelece as notações, definições e conceitos preliminares essenciais; a segunda promove uma breve discussão sobre as propriedades das cadeias de Markov; e a terceira realiza uma minuciosa definição das características das cadeias estocásticas com memória de alcance variável. Além disso, No Capítulo 3, direcionaremos o foco à definição e ao funcionamento de uma rede de neurônios na primeira seção. Em sequência, procederemos com a formulação do modelo para neurônios estocásticos interagentes, o qual será utilizado nesta pesquisa. Adicionalmente, a Seção 3.3 será dedicada à exploração dos grafos de conectividade neuronal, seguida pela articulação das definições referentes aos métodos empregados na estimação do grafo, abordados em 3.4. Posteriormente, no Capítulo 4, temos a discussão a cerca da aplicação computacional e dos resultados obtidos. Por último, apresentamos os detalhes da aplicação dos métodos de estimação abordados em dados reais no Capítulo 5.

Capítulo 2

Modelagem da atividade neuronal

A modelagem matemática da atividade neuronal tem como objetivo descrever a natureza e o comportamento dos disparos neuronais, permitindo uma compreensão mais profunda das sequências de potenciais de ação geradas por um neurônio. Além disso, a modelagem matemática nos permite fazer previsões sobre o funcionamento das células nervosas com base nas informações adquiridas. Neste capítulo, discutimos as características iniciais da modelagem da atividade neuronal, incluindo conceitos preliminares, definições e propriedades. O objetivo é apresentar como a atividade neuronal pode ser modelada a partir de processos estocásticos a fim de criar uma base teórica para o entendimento do modelo de neurônios estocásticos interagentes apresentado no próximo capítulo. Nesse sentido, organizamos este capítulo da seguinte maneira: na Seção 2.1 fixamos as notações, definições e conceitos preliminares, que estabelecem a terminologia e as convenções usadas ao longo do texto; em seguida, na Seção 2.2 apresentamos uma breve discussão a respeito das cadeias de Markov e, por fim, na Seção 2.3 definimos as cadeias estocásticas com memória de alcance variável. Uma discussão comparando o uso dessas cadeias em um contexto de modelagem neuronal também é apresentada.

2.1 Notações, definições e conceitos preliminares.

O sistema nervoso é caracterizado pela presença de células sensíveis aos estímulos externos e internos do corpo, os neurônios. Essas células produzem os chamados trens de disparos elétricos, que podem ser observados em exames convencionais. O registro dessas atividades deve ser feito de forma efetiva para o uso da modelagem matemática. Sendo assim, assumamos que estejamos monitorando a atividade de um axônio durante um período

de tempo de T milissegundos, medida convencional para esse contexto, e que dividimos o intervalo $[0, T]$ em frações iguais e sequenciais $\{t_0, t_1, t_2, \dots, t_n\}$ de tempo, em que n é um número inteiro e positivo e cada partição é também medida em milissegundos (ms). Nesse caso, podemos tomar os tempos t_i , $i = 1, 2, \dots, n$, comportando-se como uma progressão aritmética, em que o valor da razão de progressão é dado por $h := \frac{T}{n}$ e o valor inicial da progressão é t_0 . O valor h é um número natural suficientemente pequeno e não negativo, chamado de **tamanho do bin**. A partir de agora, ao menos que seja mencionado, ao falarmos de **bin** chamaremos de tempo t .

As observações do comportamento das sequências geradas pelos disparos sugerem uma estrutura probabilística para descrição matemática e o tratamento de fenômenos neuronais, justificada pelo conhecimento de que existem diferentes sequências de disparos que podem ocorrer mesmo sem alteração do estímulo dentro de um intervalo de tempo $[0, T]$, entretanto, em meio a todas as sequências possíveis, uma única em particular acontece (London *et al.*, 2002).

Dado o contexto de estocasticidade e a proposta de modelar o comportamento de um neurônio ao longo de um período de tempo de T ms, podemos utilizar um processo estocástico a tempo discreto para descrever a atividade neuronal. Sendo assim, gerando uma base com fundamentos probabilísticos para o comportamento dos disparos de neurônios.

Definição 2.1 (Processo estocástico) *Um processo estocástico a tempo discreto é uma sequência de variáveis aleatórias $\mathbf{X} := \{X_t : t \in \mathbb{T}\}$, definidas em um mesmo espaço de probabilidades $(\Omega, \mathcal{F}, P)$ com valores em um alfabeto A , em que $\mathbb{T}$ é um conjunto de índices enumeráveis.*

Neste trabalho, vamos assumir, sem perda de generalidade, que $\mathbb{T} = \mathbb{N}$, em que $\mathbb{N}$ denota o conjunto dos números naturais. Nesse sentido, para uma solução que melhor se encaixa ao problema de modelar a atividade neuronal, definimos uma cadeia estocástica a tempo discreto $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$ tal que

$$X_t := \begin{cases} 1, & \text{se for observado um disparo do neurônio no } t\text{-ésimo bin,} \\ 0, & \text{caso contrário,} \end{cases}$$

para todo $t = 1, 2, \dots, n$.

Deste modo, ao avaliarmos a estrutura do processo estocástico $\mathbf{X}$ definido para modelagem da atividade neuronal, notamos que o alfabeto A , no qual $\mathbf{X}$ assume valores, é

tal que $A = \{0, 1\}$.

Ademais, em um contexto inferencial, ao definirmos a modelagem estocástica para a atividade neuronal é de grande interesse para o estudo que haja diminuição da complexidade do problema. Para isso, buscamos a redução do número de parâmetros a serem estimados, portanto, assumimos que o conjunto de probabilidades de transição atrelado ao sistema é invariável ao longo do tempo, e que as distribuições conjuntas não dependem da escolha do tempo inicial. Logo, estamos considerando um processo estocástico **estacionário e homogêneo**.

Definição 2.2 (Homogeneidade) *Um processo estocástico a tempo discreto $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$, definido em um espaço de probabilidade $(\Omega, \mathcal{F}, P)$ com valores em um alfabeto finito A , é homogêneo quando as probabilidades de transição não variam ao longo do tempo, ou seja, não há dependência entre o tempo $t \in \mathbb{N}$ e as probabilidades condicionais do sistema. Ou seja,*

$$P(X_t = b \mid \bigcap_{k=-\infty}^{t-1} \{X_k = a_k\}) = P(X_{t+n} = b \mid \bigcap_{k=-\infty}^{t-1} \{X_{(k+n)} = a_k\}),$$

para qualquer $n \in \mathbb{N}$ e $a_k, b \in A$ sendo $t \in \mathbb{N}$.

Definição 2.3 (Estacionariedade) *Um processo estocástico a tempo discreto $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$ definido em um espaço de probabilidade $(\Omega, \mathcal{F}, P)$ com valores em um alfabeto finito A , é estacionário quando as distribuições conjuntas não dependem da escolha do tempo 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 $m, n \in \mathbb{N}$ e $a_t \in A$ sendo $t = m, m+1, \dots, n$.

Durante o desenvolvimento do trabalho é essencial que tenhamos notações para alguns elementos importantes do estudo que serão utilizados com frequência. Podemos notar que o alicerce do registro neuronal é construído com base em sequências, logo, ao considerarmos m e n dois números inteiros tais que $m \leq n$, denotamos por a_m^n a sequência finita $(a_m, \dots, a_n)$. Sendo assim, a partir da mesma lógica de notação em que m define a indexação inicial e n a final, denotamos por $a_m^{+\infty}$ a sequência infinita à direita $(a_m, a_{m+1}, \dots)$

e por $a_{-\infty}^n$ a sequência infinita à esquerda $(\dots, a_{n-1}, a_n)$, em que $a_i \in A$ para todo $i \in \mathbb{N}$.

Além disso, para qualquer $m \leq n$, podemos denotar a quantidade de elementos da sequência a_m^n por $|a_m^n| = n - m + 1$. Ainda, consideraremos que a_{n+1}^n é uma sequência vazia para todo $n \in \mathbb{N}$, dado que $|a_{n+1}^n| = 1 - 1 = 0$, por isso $a_{n+1}^n = \emptyset$.

Ao tomarmos o caso de duas sequências v e v' , podemos concatená-las e denotar o resultado como vv' , estendendo-o também para situações em que v representa uma sequência semi-infinita, ou seja, $v = v_{-\infty}^{-1}$. O tamanho da concatenação de duas sequências finitas é dado pela soma dos tamanhos individuais de cada uma, isto é, $|v| + |v'|$. Ainda, denotamos por $v^n = vv\dots v$ a junção de n vezes a sequência v se n é um número inteiro positivo e v uma sequência finita de elementos em um alfabeto A .

Ainda, podemos denotar o conjunto de todas as sequências finitas por

$$A^* = \bigcup_{j=0}^{+\infty} A^j,$$

em que A^j é o conjunto de todas as sequências de tamanho j ; e o conjunto de todas as sequências infinitas à esquerda por $A^{-\infty}$. Temos também a sequência vazia, denotada por $\emptyset$, em casos que $j = 0$. Por fim, denotamos por X_m^n o vetor aleatório $(X_m, X_{m+1}, \dots, X_n)$

2.2 Cadeias de Markov

As cadeias de Markov são processos estocásticos em que a probabilidade de transição para um novo estado depende da análise de um número fixo e finito de estados no passado. Essa quantidade é o que denominamos ordem da cadeia. Por exemplo, para cadeias de Markov de ordem 1 as probabilidades de transição dependem da análise de apenas 1 estado anterior em relação ao atual, enquanto para cadeias de Markov de ordem k , essa dependência se estende para k passos no passado, em que k é um número inteiro positivo. Nesse sentido, vamos iniciar definindo matematicamente a situação mais simples, ou seja, de ordem 1.

2.2.1 Cadeias de Markov de ordem 1

Definição 2.4 *Um processo estocástico a tempo discreto $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$, definido em um dado espaço de probabilidade $(\Omega, \mathcal{F}, P)$ com valores em um alfabeto finito A , é dito*

ser uma cadeia de Markov quando $\mathbf{X}$ possui a propriedade de Markov, i.e.,

$$P(X_t = a_t | X_{-\infty}^{t-1} = a_{-\infty}^{t-1}) = P(X_t = a_t | X_{t-1} = a_{t-1}),$$

quaisquer que sejam $t \in \mathbb{N}$ e $a_{-\infty}^{t-1} \in A^\infty$.

Portanto, como dito anteriormente, nas cadeias de Markov de ordem 1, a probabilidade do estado atual do sistema depende apenas da condição imediatamente anterior a ele. Sendo assim, as probabilidades condicionais especificam as transições da cadeia a partir de 1 passo no passado.

Anteriormente, discutimos a importância na diminuição do número de parâmetros para inferência da atividade neuronal e como a propriedade de homogeneidade nos ajuda nesse quesito. Desse modo, vamos considerar um sistema homogêneo para as cadeias de Markov, o que faz com que, nesse caso, as probabilidades de transição entre os estados não dependam do valor de instante tempo $t \in \mathbb{N}$. Logo, como temos um espaço de estados finito, podemos descrever a dinâmica do sistema a partir de uma matriz, a qual denominamos matriz de transição.

Definição 2.5 Uma cadeia de Markov homogênea $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$ de ordem 1 definida no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ com valores em alfabeto finito A , é dita possuir uma matriz de transição Q quando

$$Q = \begin{pmatrix} Q(a_1|a_1) & Q(a_2|a_1) & \dots & Q(a_n|a_1) \\ Q(a_1|a_2) & Q(a_2|a_2) & \dots & Q(a_n|a_2) \\ \vdots & \vdots & \ddots & \vdots \\ Q(a_1|a_n) & Q(a_2|a_n) & \dots & Q(a_n|a_n) \end{pmatrix}$$

em que, sem perda de generalidade, podemos assumir $A := \{a_1, a_2, \dots, a_n\}$ e as entradas da matriz Q são tais que

$$Q(b|a) := P(X_t = b | X_{t-1} = a),$$

quaisquer que sejam $a, b \in A$ e $t \in \mathbb{N}$.

Em situações práticas, podemos dizer que uma matriz de transição rege a evolução da cadeia, em que as linhas indicam a condição passada do sistema, enquanto o estado atual é dado pelas colunas. Logo, cada valor da matriz é uma probabilidade de transição entre esses estados. Desse modo, Q é uma matriz estocástica, ou seja,

$$0 \leq Q(\cdot|a) \leq 1, \quad \forall a \in A,$$

$$\sum_{b \in A} Q(b|a) = 1.$$

No caso de um sistema Markoviano de ordem 1, o valor que a variável aleatória X_0 assume é denominado por estado inicial da cadeia. Deste modo, atribuímos uma probabilidade à ocorrência de cada um dos possíveis estados iniciais e, portanto, o sistema é iniciado de maneira estocástica.

Definição 2.6 *Uma cadeia de Markov homogênea $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$ de ordem 1 definida no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ com valores em alfabeto finito A , possui distribuição inicial μ_0 sobre A quando*

$$\mu_0(a) := P(X_0 = a),$$

qualquer que seja $a \in A$.

A partir disso, constatamos que μ_0 é a distribuição de probabilidade de X_0 . Ainda, sem perda de generalidade, podemos assumir $A = \{a_1, a_2, \dots, a_n\}$ e representar a distribuição inicial pelo vetor coluna

$$\mu_0 = \begin{pmatrix} \mu_0(a_1) \\ \mu_0(a_2) \\ \vdots \\ \mu_0(a_n) \end{pmatrix}$$

Observação 2.7 *Em um abuso de notação, vamos utilizar μ_0 tanto para denotar o vetor quanto para a medida de probabilidade.*

Sendo assim, no caso da cadeia de Markov de ordem 1, concluímos que a sua condição

inicial é, basicamente, a distribuição inicial dada pelo vetor coluna μ_0 , e que a dinâmica do sistema é fixada a partir de uma matriz de transição Q . Desse modo, se $\mathbf{X}$ é uma cadeia de Markov de ordem 1 com alfabeto finito A segue, pela regra do produto de Bayes que podemos encontrar em (Mood *et al.*, 1974, p. 37), no Teorema 31 que

$$P(X_0 = a_0, \dots, X_t = a_t) = P(X_0 = a_0) \prod_{i=1}^t P(X_i = a_i | X_{i-1} = a_{i-1}),$$

quaisquer que sejam $a_t \in A$ e $t \in \mathbb{N}$. Assim, uma vez que definimos a matriz de transição Q e a distribuição inicial μ_0 da cadeia $\mathbf{X}$, podemos escrever a distribuição conjunta da cadeia $\mathbf{X}$ como sendo

$$P(X_0 = a_0, \dots, X_t = a_t) = \mu_0(a_0)Q(a_1|a_0)\dots Q(a_t|a_{t-1}), \quad (2.8)$$

quaisquer que sejam $a_t \in A$ e $t \in \mathbb{N}$. Portanto, como vimos na equação (2.8), podemos concluir que a distribuição da cadeia de Markov $\mathbf{X}$ é, de fato, determinada pela sua condição inicial μ_0 e pela matriz de transição Q .

Assim como definimos a distribuição inicial μ_0 , podemos estender o resultado para cada instante de tempo $t \in \mathbb{N}$ de uma cadeia de Markov homogênea $\mathbf{X}$ de ordem 1 com alfabeto finito A , definindo a distribuição de X_t por meio de uma medida de probabilidade μ_t sobre o alfabeto A tal que

$$\mu_t(a) := P(X_t = a),$$

para qualquer $a \in A$. Desse modo, μ_t é a distribuição da cadeia de Markov de ordem 1 no t -ésimo instante de tempo. Podemos representar a distribuição μ_t de forma matricial da mesma maneira como representamos μ_0 . Assim, podemos nos basear no teorema da probabilidade total para escrever

$$\mu_t^\top = \mu_{t-1}^\top Q,$$

em que $\top$ denota a transposta de um vetor ou de uma matriz. Recursivamente, podemos utilizar Q^t como notação para t -vezes a matriz Q e escrever

$$\mu_t^\top = \mu_0^\top Q^t.$$

Podemos notar que $\mu_t(a)$ depende do instante de tempo $t \in \mathbb{N}$ para todo $a \in A$. Entretanto, como discutimos na Seção 2.1, no contexto desse trabalho é interessante considerarmos uma cadeia com distribuição estacionária para que não haja uma dependência explícita da escolha do tempo inicial do sistema, com o objetivo de diminuir a quantidade de parâmetros a serem estimados. Logo, vamos assumir $\mu_t(a) = \mu_{t-1}(a)$ para todo $t \in \mathbb{N}$.

Definição 2.9 *Uma distribuição de probabilidade π sobre um alfabeto finito A é dita ser a distribuição estacionária da matriz de transição Q , quando*

$$\pi^\top = \pi^\top Q. \quad (2.10)$$

Em particular,

$$\pi(a) = \sum_{b \in A} \pi(b)Q(a|b), \quad \forall a \in A,$$

a qual é denominada equação de balanço do estado a .

Definição 2.11 *Seja $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$ uma cadeia de Markov homogênea de ordem 1 definida no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ com valores em alfabeto finito A , matriz de transformação Q e μ_0 como distribuição inicial. Dizemos que $\mathbf{X}$ é uma cadeia estacionária se $\mu_0 = \pi$, em que π é uma distribuição estacionária de Q .*

Sendo assim, se $\mathbf{X}$ é uma cadeia estacionária, então, para todo $t \geq 0$, podemos iterar t vezes a equação 2.10 e obter

$$\pi^\top = \pi^\top Q^t,$$

em que Q é matriz de transição de $\mathbf{X}$ e π é uma distribuição estacionária de Q . Logo, temos um sistema que não depende explicitamente da escolha do tempo inicial $t \in \mathbb{N}$.

Uma cadeia de Markov de ordem 1 homogênea com matriz de transição Q assumindo valores em um alfabeto finito A tem a propriedade de possuir pelo menos uma distribuição estacionária π sobre A . Além disso, mesmo que a sequência $\{X_t : t \in \mathbb{N}\}$ não convirja para um estado fixo, a sequência de distribuições associadas $\{\mu_t : t \in \mathbb{N}\}$ converge para a distribuição estacionária π , sendo assim, podemos interpretar π como um comportamento assintótico de μ_t .

Em termos da neurociência, utilizar uma cadeia de Markov de ordem 1 para modelar a atividade neuronal pode ser consideravelmente restritivo, uma vez que o sistema evolui apenas com base no seu estado imediatamente anterior. Entretanto, existem evidências empíricas de que a dependência da cadeia, neste caso, é de ordem maior que 1 (Karnup ; Zhadin 1983; González-Montoro et al. 2011; Foik et al. 2014) dada a autocorrelação temporal dos disparos neuronais. Logo, podemos considerar um modelo um pouco menos restritivo e adotar, por exemplo, uma cadeia de Markov de ordem k , sendo $k > 1$.

2.2.2 Cadeias de Markov de ordem k

Como dito no início da Seção 2.2, a quantidade de passos anteriores que devem ser analisados para obtermos um resultado probabilístico em relação ao estado atual define a ordem dos modelos de Markov. Desse modo, a dinâmica das cadeias Markov de ordem k , assim como nas de ordem 1, também é regida por regras probabilísticas que dependem de uma quantidade fixa de estados no passado. Portanto, em uma cadeia de Markov de ordem k , a probabilidade do sistema assumir um determinado estado de um alfabeto finito A no instante atual de tempo t depende de k passos no passado, sendo $k > 1$ um número inteiro positivo.

Definição 2.12 *Um processo estocástico a tempo discreto $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$ definido no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ com valores em alfabeto finito A , é dito ser uma cadeia de Markov de ordem k , sendo k um inteiro positivo, quando*

$$P(X_t = a_t | X_0^{t-1} = a_0^{t-1}) = P(X_t = a_t | X_{t-k}^{t-1} = a_{t-k}^{t-1}),$$

quaisquer que sejam $t \in \mathbb{N}$ com $t \geq k$, $a_t \in A$ e $a_0^{t-1} \in A^t$.

Ou seja, a probabilidade de transição do sistema para um determinado estado depende somente dos últimos k estados. Além disso, uma cadeia de Markov de ordem k , assim como no caso de ordem 1, pode ter sua dinâmica descrita a partir de uma matriz de transição, quando o alfabeto A é finito e assumimos homogeneidade, uma vez que as probabilidades de transição não dependem do valor do instante de tempo $t \in \mathbb{N}$ com $t \geq k$.

Definição 2.13 *Uma cadeia de Markov homogênea $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$ de ordem finita*

k definida no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ com valores em alfabeto finito A , possui matriz de transição Q quando as entradas da matriz são tais que para cada $a_0^{k-1} \in A^k$ temos

$$Q(b|a_0^{k-1}) := P(X_t = b | X_{t-k}^{t-1} = a_0^{k-1}),$$

para qualquer $b \in A$ e $t \in \mathbb{N}$.

Assim como no caso da cadeia de Markov de ordem 1, a matriz de transição Q para a ordem k também representa a dinâmica da evolução do sistema. Entretanto, agora estamos olhando para a passagem de k estados passados até o presente, por exemplo, se fixarmos $k = 6$, então cada entrada da matriz Q é uma probabilidade de transição da condição do sistema em 6 passos anteriores até o atual. Portanto, Q é uma matriz estocástica, ou seja,

$$\begin{aligned} 0 \leq Q(\cdot|a_0^k) \leq 1, \quad \forall a_0^k \in A^k, \\ \sum_{b \in A} Q(b|a_0^k) = 1. \end{aligned}$$

Como vimos anteriormente, cadeias de Markov de ordem 1 têm a condição inicial representada por um único escalar $a \in A$. Entretanto, ao aumentar a ordem da cadeia para k , $k > 1$, a condição inicial é um vetor de dimensão k . Além disso, em cada um dos possíveis vetores de estados iniciais é atribuída uma probabilidade. Logo, os modelos de ordem k também podem ser iniciados de maneira estocástica.

Definição 2.14 *Uma cadeia de Markov homogênea $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$ de ordem finita k definida no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ com valores em um alfabeto finito A , possui distribuição inicial μ_0 sobre A^k quando*

$$\mu_0(a_0^{k-1}) := P(X_0^{k-1} = a_0^{k-1}),$$

qualquer que seja $a_0^{k-1} \in A^k$.

Desse modo, segue que μ_0 é a distribuição de probabilidade de X_0^{k-1} . Além disso, podemos representar a distribuição inicial por um vetor coluna, em que cada entrada é

cada uma das combinações de ordem k dos elementos de A .

De forma semelhante ao que fizemos na Subseção 2.2.1, podemos encontrar a distribuição do sistema, uma vez que, para cadeias de Markov de ordem k , a condição inicial é dada pela distribuição inicial e a dinâmica é fixada a partir da matriz de transição Q .

Além disso, em geral, podemos notar uma grande semelhança entre as cadeias de Markov de ordem 1 e de ordem k , isso porque, na verdade, todo sistema de ordem k no alfabeto A pode ser reescrito como uma cadeia de Markov de ordem 1 no alfabeto A^k . O exemplo a seguir ilustra como podemos interpretar cadeias de Markov de ordem k reescrevendo-as em sistemas de ordem 1 em um contexto neurobiológico.

Exemplo 2.15 *Suponha que estejamos interessados em monitorar a atividade de um neurônio ao longo do tempo e que modelamos a evolução desse sistema por meio de uma cadeia de Markov homogênea $\mathbf{X} := \{X_t : t \in \mathbb{N}\}$ de ordem 2 definida em um espaço de probabilidade $(\Omega, \mathcal{F}, P)$ adequado, com valores no alfabeto finito $A = \{0, 1\}$ e com uma matriz de transição Q dada por*

$$Q := \begin{matrix} & \begin{matrix} 0 & 1 \end{matrix} \\ \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix} & \begin{pmatrix} 1-p & p \\ q & 1-q \\ 1-s & s \\ r & 1-r \end{pmatrix} \end{matrix}$$

em que $p, q, s, r \in (0, 1)$.

Para encontrarmos a cadeia estacionária π associada à matriz de transição Q , devemos encontrar π tal que as equações de balanço sejam satisfeitas, ou seja,

$$\pi(00) = (1-p)\pi(00) + (1-s)\pi(01),$$

$$\pi(01) = q\pi(10) + r\pi(11),$$

$$\pi(10) = p\pi(00) + s\pi(01),$$

$$\pi(11) = (1-q)\pi(10) + (1-r)\pi(11).$$

Desse modo,

$$\begin{aligned}
\pi(00) &= \frac{r(s-1)}{r(s-1) + p(-2r+q-1)}, \\
\pi(01) &= -\frac{pr}{r(s-1) + p(-2r+q-1)}, \\
\pi(10) &= -\frac{pr}{r(s-1) + p(-2r+q-1)}, \\
\pi(11) &= \frac{p(q-1)}{r(s-1) + p(-2r+q-1)}.
\end{aligned}$$

Sendo assim, se $\mu_0 = \pi := (\pi(00), \pi(01), \pi(10), \pi(11))^\top$, segue que a cadeia de Markov $\mathbf{X}$ de ordem 2 com matriz de transição Q é estacionária, já que não depende explicitamente da escolha do tempo inicial.

Agora, vamos assumir que $Y_t = (X_{t-1}, X_t)$, para todo $t \in \mathbb{N}$. Assim, podemos escrever as probabilidades de transição do processo $\mathbf{Y} := \{Y_t : t \in \mathbb{N}\}$ da seguinte maneira

$$P(X_t = a_t | X_{t-1} = a_{t-1}, X_{t-2} = a_{t-2}) = P(Y_t = u_t | T_{t-1} = u_{t-1}),$$

em que $u_t = (a_{t-1}, a_t) \in A^2$, $t \in \mathbb{N}$. Logo, vamos tratar dois estados da cadeia compilados em uma única variável u_t , porém no alfabeto A^2 , o que nos possibilita escrever a matriz de transição Q^* de $\mathbf{Y}$ como

$$Q^* := \begin{matrix} & \begin{matrix} 00 & 01 & 10 & 11 \end{matrix} \\ \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix} & \begin{pmatrix} 1-p & p & 0 & 0 \\ 0 & 0 & q & 1-q \\ 1-s & s & 0 & 0 \\ 0 & 0 & r & 1-r \end{pmatrix} \end{matrix}$$

em que $p, q, s, r \in (0, 1)$.

Logo, $\mathbf{Y}$ é uma cadeia de Markov homogênea de ordem 1 sobre A^2 .

Para encontrarmos a cadeia estacionária π^* associada à matriz de transição Q^* , devemos encontrar π^* tal que as equações de balanço sejam satisfeitas, ou seja,

$$\begin{aligned}
\pi^*(00) &= (1-p)\pi^*(00) + (1-s)\pi^*(10), \\
\pi^*(01) &= p\pi^*(00) + s\pi^*(10), \\
\pi^*(10) &= q\pi^*(01) + r\pi^*(11), \\
\pi^*(11) &= (1-q)\pi^*(01) + (1-r)\pi^*(11).
\end{aligned}$$

Desse modo,

$$\begin{aligned}
\pi^*(00) &= \frac{r(s-1)}{r(s-1) + p(-2r+q-1)}, \\
\pi^*(01) &= -\frac{pr}{r(s-1) + p(-2r+q-1)}, \\
\pi^*(10) &= -\frac{pr}{r(s-1) + p(-2r+q-1)}, \\
\pi^*(11) &= \frac{p(q-1)}{r(s-1) + p(-2r+q-1)}.
\end{aligned}$$

Sendo assim, se $\mu_0^ = \pi^* := (\pi^*(00), \pi^*(01), \pi^*(10), \pi^*(11))^\top$, segue que a cadeia de Markov $\mathbf{Y}$ de ordem 1 com matriz de transição Q^* é estacionária, já que não depende explicitamente da escolha do tempo inicial.*

Finalmente, ao compararmos π e π^ , podemos notar que $\pi = \pi^*$. Logo, mostramos que a existência de cadeias estacionárias de Markov de ordem k pode ser analisada a partir do estudo da existência de cadeias estacionárias de Markov de ordem 1 ao adaptá-las no contexto do problema, uma vez que é possível encontrar a distribuição considerando tanto uma cadeia de ordem k quanto de ordem 1.*

2.3 Cadeias com memória de alcance variável

Cadeias estocásticas com memória de tamanho variável, também conhecidas como modelos de árvore de contexto, são ferramentas importantes em diversos campos do conhecimento. Isso inclui não apenas a neurociência, mas também áreas como bioinformática (Bejerano e Yona, 2001) e linguística (Galves *et al.*, 2004), entre outras. Essas cadeias possuem uma dinâmica relativamente semelhante ao que mostramos nas Subseções 2.2.1 e 2.2.2, entretanto, a ordem ou tamanho da sua memória é variável, ou seja, depende de

um número não-fixo de estados no passado.

Os modelos de árvore de contexto são denominados dessa forma, pois é possível representar de forma pictória, por meio de uma árvore ilustrativa, o conjunto das possíveis sequências do passado que são relevantes para definir a transição da cadeia para um novo estado. Além disso, essas cadeias são consideravelmente mais flexíveis e parcimoniosas quando comparadas às de ordem fixa, uma vez que elas incorporam memória ao sistema apenas onde é necessário e podem gerar modelos mais eficientes para a predição de dados.

As definições a seguir servirão como alicerce para a definição formal de uma cadeia estocástica com memória de alcance variável que serão utilizadas neste trabalho para modelar a atividade de uma rede de neurônios interagentes.

Definição 2.16 (Sufixo) *Dados m e n números inteiros positivos. Uma sequência de estados s_{-m}^{-1} é sufixo de uma outra sequência de estados v_{-n}^{-1} se $m < n$ e $v_{-m}^{-1} = s_{-m}^{-1}$. A notação para esse caso é $s_{-m}^{-1} \preceq v_{-n}^{-1}$. Especificamente, se $|s_{-m}^{-1}| < |v_{-n}^{-1}|$, então dizemos que s é um sufixo adequado de v e denotamos por $s_{-m}^{-1} \prec v_{-n}^{-1}$.*

A principal motivação dessa definição deriva do pensamento de que os valores que uma cadeia estocástica $\mathbf{X}$ assume no alfabeto finito A podem ser interpretados como palavras e não estritamente como sequências numéricas. Por exemplo, se assumirmos duas sequências $w = \{0, 0, 1\}$ e $v = \{0, 1\}$, ao encararmos a combinação de 0 e 1 de forma linguística, então a palavra 01 é sufixo da palavra 001.

Definição 2.17 (Propriedade do Sufixo) *Um subconjunto τ de $A^* \cup A^{-\infty}$ é uma árvore se satisfaz a propriedade do sufixo, ou seja, se nenhuma sequência de estados $s \in \tau$ é um sufixo de outra sequência $v \in \tau$. Se, além disso, uma árvore τ satisfaz a propriedade de irredutibilidade, que afirma que nenhuma sequência pertencente a τ pode ser substituída por um sufixo adequado sem violar a propriedade de sufixo, então ela é chamada de árvore irredutível.*

Uma árvore τ pode ser graficamente representada desde que atenda a propriedade do sufixo. A representação de uma árvore é uma forma pictória de visualizar as diferentes sequências de um processo em um determinado alfabeto finito A . Na representação, a raiz da árvore é o nó inicial, denotado por λ , a partir do qual se ramificam as arestas que representam as possíveis transições para os símbolos do alfabeto A . Logo, cada sequência é representada como um caminho da raiz λ , até uma folha, ou seja, um nó que não possui

filhos. Nós que possuem filhos são denominados nós internos. Além disso, as sequências são escritas da direita para a esquerda, diferentemente da leitura convencional.

Exemplo 2.18 Vamos considerar um alfabeto finito $A = \{0, 1\}$ e

$$\tau_1 = \{001, 01, 1, 00\},$$

$$\tau_2 = \{100, 10, 1\},$$

$$\tau_3 = \{00, 10, 1\}.$$

Sendo assim, podemos notar que τ_3 é uma árvore irredutível em A , uma vez que ela satisfaz a propriedade do sufixo e nenhuma sequência pertencente a ela pode ser substituída por um sufixo adequado sem violar a propriedade. Em contrapartida, τ_1 não pode ser considerada uma árvore em A , já que 01 é sufixo de 001 , logo, violando a propriedade do sufixo. Por fim, τ_2 é uma árvore em A , já que não possui nenhuma sequência $v \in \tau_2$ que seja sufixo de outra sequência $w \in \tau_2$, entretanto, a árvore não é irredutível, pois a sequência 100 pode ser reduzida para 00 .

Por satisfazerem a propriedade do sufixo, τ_2 e τ_3 podem ser representadas de forma pictórica por

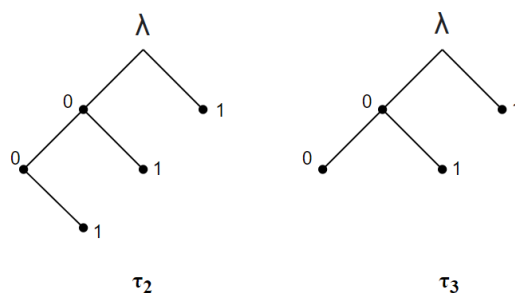


Figura 2.1: Representação das estruturas das árvores definidas no exemplo.

Definição 2.19 (Altura da árvore) A altura de uma árvore τ é determinada pelo maior

caminho entre a raiz da árvore e a maior sequência que pertence a τ , ou seja,

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

A altura da árvore determina também uma propriedade importante de τ . Sempre que $h(\tau) < +\infty$, dizemos que τ é uma árvore limitada e todas as suas sequências são finitas. Por outro lado, se $h(\tau) = +\infty$, isso significa que τ é ilimitada, o que implica que pelo menos uma sequência em τ é infinita.

Exemplo 2.20 Vamos considerar as mesmas árvores definidas no Exemplo 2.18. Desse modo, temos $h(\tau_2) = 3$ e $h(\tau_3) = 2$.

Além disso, a partir da definição, podemos afirmar que tanto τ_2 quanto τ_3 são árvores limitadas, uma vez que $h(\tau_2) < +\infty$ e o mesmo vale para $h(\tau_3)$. Agora, consideremos $\tau_4 = \{10^k : k = 0, 1, \dots\} \cup 0^\infty$, em que $\{10^k : k = 0, 1, \dots\}$ representa as sequências obtidas ao concatenar 1 com 0 em k iterações, ou seja, $\{1, 10, 100, 1000, \dots\}$. Podemos concluir que τ_4 é uma árvore em $A = \{0, 1\}$, pois atende à propriedade do sufixo. A particularidade de τ_4 em relação às árvores τ_2 e τ_3 se dá pelo fato de que ela possui folhas infinitas, ou seja, $h(\tau_4) = +\infty$. Logo, a árvore τ_4 é ilimitada.

A seguir, vamos estudar a definição de árvore truncada.

Definição 2.21 (Árvore truncada) Dada uma árvore $\tau \in A^* \cup A^{-\infty}$ e $K \in \mathbb{N}$, vamos denotar por $\tau|_K$ a árvore truncada no nível K , ou seja,

$$\tau|_K = \{w \in \tau : h(w) \leq K\} \cup \{w \in A^K : w \prec u \text{ para algum } u \in \tau\}$$

Exemplo 2.22 Vamos considerar a árvore $\tau_4 = \{10^k : k = 0, 1, \dots\} \cup 0^\infty$ do exemplo 2.20 e $K = 3$. Sendo assim, a árvore truncada é dada por

$$\tau_4|_3 = \{000, 100, 10, 1\}.$$

Sendo assim, a representação pictórica de τ_4 é dada por

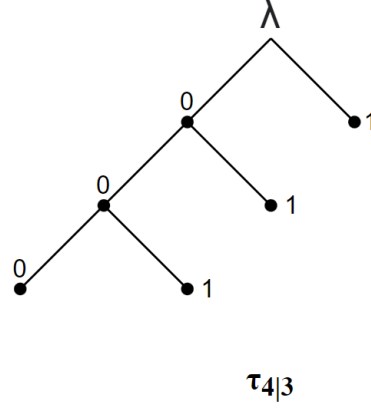


Figura 2.2: Representação da estrutura da árvore τ_4 truncada no nível 3.

Definição 2.23 (Árvore Completa e Árvore de contexto) *Uma árvore τ é completa se qualquer elemento $a_{-\infty}^{-1}$ de $A^{-\infty}$ tem um sufixo pertencente a τ . Pela propriedade do sufixo, temos que ele é único, uma vez que nenhuma sequência de estados $s \in \tau$ é um sufixo de outra sequência $v \in \tau$. Uma árvore completa também é denominada como árvore de contexto. Chamamos qualquer sufixo $v \in \tau$ de contexto da sequência $a_{-\infty}^{-1}$, desde que τ seja uma árvore completa, e o denotamos por $c_{\tau}(a_{-\infty}^{-1})$.*

Exemplo 2.24 *Vamos considerar $\tau_1 = \{10, 00, 1\}$. Sendo assim, podemos notar que τ_1 satisfaz a propriedade do sufixo e que, além disso, qualquer elemento $a_{-\infty}^{-1}$ de $A^{-\infty}$ possui sufixo pertencente à árvore, logo, τ_1 é uma árvore completa. Além disso, se analisarmos, por exemplo, a sequência $a_{-\infty}^{-1} = \dots 001$, então o contexto de $a_{-\infty}^{-1}$ é dado por $c_{\tau_1}(a_{-\infty}^{-1}) = 1$.*

Por outro lado, vamos levar em consideração $\tau_2 = \{100, 10, 1\}$. Logo, τ_2 satisfaz a propriedade do sufixo e pode ser considerada uma árvore. Entretanto, nem toda sequência $a_{-\infty}^{-1}$ de $A^{-\infty}$ possui sufixo pertencente à árvore, como, por exemplo, $a_{-\infty}^{-1} = \dots 00$. Portanto, τ_2 não é uma árvore completa.

Definição 2.25 (Árvore de Contexto Probabilística) *Uma árvore de contexto probabilística com símbolos em um alfabeto finito A é um par (τ, Q) tal que*

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

Ou seja, a árvore de contexto probabilística pode ser vista como uma representação de uma família de probabilidades de transição. Sendo assim, devemos considerar cada

estado passado e seus respectivos contextos para definirmos a transição do sistema para o estado seguinte.

Definição 2.26 (Cadeia com Memória de Alcance Variável) *Dizemos que uma cadeia estocástica $\mathbf{X} := \{X_t : t \in \mathbb{Z}\}$ definida no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ e com valores em um alfabeto finito A é compatível com uma árvore de contexto probabilística (τ, Q) quando para P -quase toda sequência $a_{-\infty}^{-1} \in A^{-\infty}$ e qualquer $a \in A$ tivermos*

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

Chamamos essa cadeia de cadeia estocástica com memória de alcance variável.

Em relação à estimação, a quantidade de parâmetros em uma cadeia de Markov de ordem finita que devem ser inferidos é maior comparado ao caso de cadeias com memória de alcance variável, quando compatíveis a uma árvore de contexto limitada τ , uma vez que, nesse caso, precisamos apenas de uma coleção de $(|A| - 1)|\tau|$ probabilidades de transição para descrever o sistema. Sendo assim, é inferencialmente vantajoso considerar cadeias com memória de alcance variável.

Nesse sentido, podemos utilizar uma cadeia estocástica com memória de alcance variável para modelar a atividade neuronal, levando em consideração que cada possível disparo de um neurônio depende do histórico da atividade até o último sinal elétrico emitido e, conseqüentemente, do potencial de membrana acumulado até o momento presente. A seguir, veremos um exemplo detalhado sobre a modelagem com esse tipo de cadeia aplicada ao contexto neuronal.

Exemplo 2.27 *Vamos considerar uma cadeia estocástica a tempo discreto $\mathbf{V} = \{V_t : t \in \mathbb{N}\}$, sendo que o potencial de membrana de um dado neurônio no tempo $t \geq 0$ é denotado por V_t . Além disso, assumimos que V_t no tempo 0 é dado por $V_0 = v_R$, ou seja, v_R é o potencial de repouso da membrana. Desse modo, para $t \geq 1$, o comportamento da cadeia é da seguinte forma:*

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

em que $0 \leq \rho \leq 1$ modela o efeito da corrente de vazamento e $\mu \in \mathbb{R}$ é uma corrente constante de entrada. Como o potencial de membrana sofre uma queda abrupta com o disparo do neurônio, escrevemos $V_t = v_R$ caso um disparo do neurônio seja observado, por outro lado, se não houver disparo multiplica-se o potencial da membrana no estado v_{t-1} no tempo anterior $t - 1$ pelo efeito da corrente ρ e soma-se com uma corrente de entrada μ . Desse modo, o espaço de estados da cadeia estocástica $\mathbf{V}$ é

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

Nesse sentido, podemos definir a probabilidade condicional com relação ao potencial de membrana no instante $t - 1$ de um neurônio efetuar um disparo em $t \in \mathbb{Z}$ da seguinte maneira

$$P(X_t = 1 | V_{t-1} = v_{t-1}) = \phi(v_{t-1}),$$

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

Podemos compreender a dinâmica do alcance da memória da cadeia por meio da simulação dos valores X_t a partir dos estados $X_{t-1}, \dots, X_0$. Para isso, tomamos U_t , para $t \in \mathbb{Z}$ como uma variável aleatória com distribuição uniforme no intervalo $[0, 1]$. Se considerarmos $X_{t-1} = 1$ e então $V_{t-1} = v_R$ e

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

A simulação com valores da distribuição uniforme $[0, 1]$ é válida uma vez que

$$P(X_t = 1) = P(U_t \leq \phi(v_{t-1})) = \phi(v_{t-1}).$$

A partir disso, vamos supor agora que $X_{t-1} = 0$, desse modo, temos $V_{t-1} \neq v_R$ e seria necessário observar o valor de mais um estado no passado X_{t-2} , uma vez que o ponto de referência para a construção da cadeia é, neste contexto, sempre uma função do potencial de repouso da membrana. No caso de $X_{t-2} = 1$, teríamos $V_{t-2} = v_R$, $V_{t-1} = \rho v_R + \mu$ e

$$X_t = \begin{cases} 1, & \text{caso } U_t \leq \phi(\rho v_R + \mu), \\ 0, & \text{caso contrário.} \end{cases} \quad (2.28)$$

Contudo, se $X_{t-2} = 0$, então $V_{t-2} \neq v_R$ e deveríamos avaliar X_{t-3} e assim sucessivamente. Logo, cria-se uma dependência do valor X_t com um número variável de estados no passado. Desse modo, se denotarmos o último disparo realizado pelo neurônio antes do instante t por

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

então X_t é uma função de $\{X_{t-1}, X_{t-2}, \dots, X_{L_{t-1}}, U_t\}$. Logo, podemos notar, por meio desse algoritmo, que a dependência do valor X_t no passado não é fixa, mas, sim, uma função do passado que varia a cada estado, uma vez que buscamos olhar para o último disparo do neurônio antes do tempo t .

Como dito anteriormente, utilizar uma Cadeia de Markov de ordem 1 para modelar a atividade neuronal pode ser consideravelmente restritivo. Entretanto, modelar esses fenômenos a partir de uma Cadeia de Markov de ordem k pode, também, apresentar restrições severas, uma vez que os disparos dos neurônios estão sujeitos a mudanças ao longo do tempo e podem ser influenciados por uma ampla gama de fatores internos e externos, logo, a modelagem dessas atividades não depende de um número fixo de estados no passado, como visto no Exemplo 2.27. Por outro lado, as cadeias estocásticas com memória de alcance variável levam em consideração a variabilidade inerente a esses eventos, permitindo um ajuste menos restritivo.

Em resumo, os processos estocásticos com memória variável fornecem um arcabouço teórico robusto e flexível, que permite uma representação mais precisa das atividades neuronais quando comparados às cadeias Markovianas, pois geram estruturas adequadas à natureza desses fenômenos.

Capítulo 3

Sistemas de neurônios estocásticos interagentes

Como dissemos anteriormente, temos evidências empíricas de que a atividade dos disparos neuronais é inerentemente estocástica e, sabendo disso, o nosso principal objetivo é conseguir fazer a predição de redes de neurônio estocásticos com memória de alcance variável partindo dos conceitos estudados no Capítulo 2. Para isso, primeiramente, vamos discutir a definição de uma rede de neurônios e explicar seu funcionamento. Em sequência, na Seção 3.2, vamos definir e detalhar o modelo não Markoviano que será utilizado para realizar a predição. Além disso, na Seção 3.3, vamos introduzir os grafos de conectividade neuronal, suas propriedades e particularidades. Por fim, discutiremos os métodos de estimação dos grafos de conectividade de uma rede neuronal que serão considerados para esta pesquisa. Ainda, vale ressaltar que o modelo que definiremos neste capítulo é inspirado no modelo proposto por Galves e Löcherbach (2016), o qual remonta características da modelagem neurobiológicas também consideradas por autores como De Masi *et al.* (2015), Duarte e Ost (2014), Fournier e Löcherbach (2016), Robert e Touboul (2016), Duarte *et al.* (2015), Hodara e Löcherbach (2017), Yaginuma (2016), Baccelli e Taillefumier (2019) e Baccelli *et al.* (2022).

3.1 Rede de neurônios

O cérebro é o principal e mais complexo órgão do sistema nervoso, o qual é composto por bilhões de neurônios, que são células eletricamente excitáveis especializadas na troca de informação. A comunicação entre os neurônios é realizada através de sinapses,

que podem ser de natureza elétrica ou química. Nas sinapses químicas, o neurônio pré-sináptico libera um neurotransmissor, uma substância química endógena que transmite informações, que se liga a uma proteína em um neurônio pós-sináptico ocasionando uma despolarização localizada em seu potencial de membrana. Já nas sinapses elétricas, ocorre um fluxo direto de íons entre as células.

O potencial de membrana refere-se à diferença de carga elétrica entre o interior e o exterior da célula em dois estados principais: potencial de repouso e potencial de ação. Chamamos de potencial de repouso quando a diferença de carga elétrica intra e extracelular é mantida em um valor aproximadamente constante devido à regulação do fluxo de íons. Em contrapartida, o potencial de ação, também conhecido como impulso nervoso ou disparo neuronal, é um estado em que ocorre uma alteração rápida na diferença de potencial elétrico entre o interior e o exterior da célula, desencadeada por um estímulo, seja ele externo ou sináptico, resultando em um disparo elétrico. Em um contexto neurobiológico, a sequência de disparos neuronais gerada por um neurônio ao longo do tempo é chamada de trem de disparo. Desse modo, podemos dizer que a atividade neuronal se manifesta por meio da ocorrência de trem de disparos em diversos neurônios do nosso cérebro. Além disso, o processamento cerebral resulta da interação dessas sequências geradas ([Gerstner et al., 2014](#)).

Quando o potencial de membrana de um neurônio excede um limiar específico, ocorre a despolarização da célula, desencadeando um disparo neuronal. Em seguida, o neurônio retorna ao seu potencial de repouso e permanece nesse estado até que haja outro desequilíbrio eletrolítico. Dessa forma, após seu último disparo, o neurônio acumula potencial de membrana até que este seja suficientemente grande para permitir um novo disparo. Portanto, podemos inferir que a probabilidade de um neurônio disparar, levando em consideração seu histórico temporal passado, é maior quanto mais distante estiver o último disparo desse neurônio no passado em questão, ou seja, existe uma relação de crescimento monótono entre o aumento do potencial de membrana e a probabilidade do neurônio disparar. Isso implica que o processo estocástico que modela a atividade desse neurônio não possui uma memória Markoviana de ordem finita, como mostra alguns trabalhos da literatura ([Friston, 2010](#); [Truccolo et al., 2010](#); [Cessac, 2010](#)). A atividade de um neurônio é, portanto, modelada por um processo estocástico cuja dependência com o passado é de alcance variável. Como vimos na Seção 2.3, as probabilidades de transição desse tipo de processo estocástico não dependem de uma quantidade fixa de estados passados, levando

em conta, por outro lado, a estrutura de dependência presente nos dados.

Além disso, o potencial de membrana de um neurônio é afetado pela atividade de todos os outros neurônios que interagem com ele na rede. Sendo assim, para modelarmos a atividade neuronal não podemos isolar o comportamento de um único neurônio para prever o próximo disparo, mas considerar todas as interações pré-sinápticas que contribuem para o aumento ou diminuição do potencial de membrana.

3.2 Modelo para neurônios estocásticos interagentes

Galves e Löcherbach (2016) exploraram sistemas interagentes de processos estocásticos com memória de alcance variável, os quais podem ser utilizados na modelagem da atividade neuronal. Neste trabalho, inspirados pelo modelo proposto por esses autores, vamos modelar uma rede de neurônios estocásticos interagentes com memória de alcance variável a partir de um modelo que leva em consideração a função de vazamento como sendo a função identidade.

Seja 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 $\{0, 1\}$, definida em um espaço de probabilidade adequado $(\Omega, \mathcal{F}, P)$. Para cada neurônio $i \in I$, a cada instante de tempo $t \in \mathbb{Z}$, definimos

$$X_t^{(i)} = \begin{cases} 1, & \text{se houver disparo do neurônio } i \text{ no tempo } t, \\ 0, & \text{caso contrário.} \end{cases}$$

Sendo assim, vamos fixar a notação $\mathbf{X}_t := \{X_t^{(i)} : i \in I\}$ para a configuração global dos neurônios no instante de tempo $t \in \mathbb{Z}$ e a notação $\mathbf{X}^{(i)} := \{X_t^{(i)} : t \in \mathbb{Z}\}$ para a trajetória do neurônio $i \in I$.

Assim, podemos descrever a dinâmica do processo da seguinte maneira. Em cada instante de tempo $t \in \mathbb{Z}$, levando em consideração todo o passado, os neurônios são atualizados de forma independente. Isso implica que, para qualquer subconjunto $J \subset I$, onde $a_i \in \{0, 1\}$ para $i \in J$, temos a seguinte relação:

$$P\left(\bigcap_{i \in J} \{X_t^{(i)} = a_i\} \mid \mathbf{X}_{-\infty}^{t-1} = \mathbf{x}_{-\infty}^{t-1}\right) = \prod_{i \in J} P(X_t^{(i)} = a_i \mid \mathbf{X}_{-\infty}^{t-1} = \mathbf{x}_{-\infty}^{t-1}).$$

Além disso, como vimos anteriormente, para cada neurônio $i \in I$ em cada instante de tempo $t \in \mathbb{Z}$, considerando todo o passado, a probabilidade do neurônio i disparar é determinada pelo 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 a seguinte relação:

$$P\left(X_t^{(i)} = 1 \mid \mathbf{X}_{-\infty}^{t-1} = \mathbf{x}_{-\infty}^{t-1}\right) = \phi\left(V_{t-1}^{(i)}\right), \quad (3.1)$$

em que $\mathbf{V} := \left\{V_t^{(i)} : t \in \mathbb{Z} \text{ e } i \in I\right\}$ é um processo estocástico que toma valores em $\mathbb{R}$ tal que $V_t^{(i)}$ representa o potencial de membrana do neurônio i no instante de tempo t , e $\phi : \mathbb{R} \rightarrow [0, 1]$ é uma função que modela probabilidade de disparo do neurônio i , conhecida como *taxa de disparo neuronal*.

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 com o seu potencial de membrana. Quando um neurônio $i \in I$ dispara, ele retorna ao seu potencial de repouso e se mantém nesse estado até que outro desequilíbrio eletrolítico ocorra. Entretanto, quando um neurônio pré-sináptico $j \in I$, $j \neq i$, dispara, o potencial de membrana do neurônio i cresce proporcionalmente a um peso sináptico $\omega_{j \rightarrow i}$ que modela o efeito que o disparo do neurônio pré-sináptico j tem sobre o neurônio i , em que $\omega_{j \rightarrow i} \in \mathbb{R}$. Quando o neurônio pré-sináptico j tem um efeito excitatório sobre o neurônio pós-sináptico i temos $\omega_{j \rightarrow i} > 0$. Nesse caso, observamos um aumento no potencial de membrana do neurônio i . Por outro lado, quando o efeito do neurônio pré-sináptico j sobre o neurônio pós-sináptico i é inibitório, temos $\omega_{j \rightarrow i} < 0$, caso em que observamos uma queda no potencial de membrana do neurônio i . Nesse contexto, definimos $\omega_{j \rightarrow j} = 0$ e, neste caso, o único efeito sobre o potencial de membrana do neurônio j que é considerado deve-se ao efeito dos neurônios externos à rede em estudo. Além disso, os pesos sinápticos devem satisfazer

$$\sup_{i \in I} \sum_{j \in I} |\omega_{j \rightarrow i}| < \infty,$$

assim, garantindo a existência de uma cadeia estacionária.

Dito isso, podemos definir o modelo neural a ser considerado neste trabalho. Para

cada neurônio $i \in I$ em cada instante de tempo $t \in \mathbb{Z}$ definimos o potencial de membrana do neurônio i no instante de tempo $t - 1$ da seguinte maneira

$$V_{t-1}^{(i)} = \begin{cases} 0, & \text{se } X_{t-1}^{(i)} = 1, \\ \sum_{j=1}^N \omega_{j \rightarrow i} \sum_{s=L_t^{(i)}+1}^{t-1} X_s^{(j)}, & \text{caso contrário,} \end{cases}$$

em que $\omega_{j \rightarrow i} \in \mathbb{R}$ representa o peso sináptico que o neurônio $j \in I$ exerce sobre o neurônio i e $L_t^{(i)}$ é o instante em que ocorre o último disparo do neurônio $i \in I$, antes do instante de tempo $t \in \mathbb{Z}$, isto é,

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

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 $V_t^{(i)} = 0$. Dessa forma, para cada neurônio $i \in I$, podemos reescrever (3.1) da seguinte maneira

$$P\left(X_t^{(i)} = 1 \mid \mathbf{X}_{-\infty}^{t-1} = \mathbf{x}_{-\infty}^{t-1}\right) = \phi\left(\left(1 - x_{t-1}^{(i)}\right) \left(\sum_{j=1}^N \omega_{j \rightarrow i} \sum_{s=L_t^{(i)}+1}^{t-1} x_s^{(j)}\right)\right).$$

Podemos observar que o potencial de membrana acumulado do neurônio $i \in I$ depende do momento mais recente $L_t^{(i)}$ em que esse neurônio disparou. Em outras palavras, quanto maior for $L_t^{(i)}$, maior será a probabilidade dele disparar no instante de tempo t . Ainda, iremos adotar que $L_t^{(i)} \geq t - k$, em que k é um número inteiro positivo que representa o maior comprimento de memória com todos os neurônios estocásticos da rede.

As funções sigmóides são não lineares e estritamente crescentes em seus respectivos domínios. Em um contexto neurobiológico, as funções sigmóides captam a relação de aumento da probabilidade do neurônio disparar à medida que o potencial de membrana cresce. Sendo assim, inspirados pelos modelos de neurociência computacional em [Jolivet et al. \(2006\)](#), vamos fixar a taxa de disparo neuronal ϕ como uma função logística, que é um exemplo comum de função sigmóide, além de ter um conjunto imagem no intervalo aberto $]0, 1[$, tornando-a plausível para descrever o comportamento de probabilidades. Dessa forma, a função de ligação ou taxa de disparo é representada como $\phi : \mathbb{R} \rightarrow]0, 1[$, tal que

$$\phi(x) = \frac{\exp(x)}{1 + \exp(x)},$$

para todo neurônio $i \in I$.

Com isso, podemos reescrever a Equação (3.1) com base na função de taxa de disparo logística da seguinte maneira

$$P\left(X_t^{(i)} = 1 \mid \mathbf{X}_{-\infty}^{t-1} = \mathbf{x}_{-\infty}^{t-1}\right) = \frac{\exp(V_{t-1}^{(i)})}{1 + \exp(V_{t-1}^{(i)})}. \quad (3.2)$$

3.3 Grafo de conectividade

Um grafo é uma estrutura matemática que consiste em um conjunto de vértices ligados por um conjunto de arestas, ou seja, pontos ligados por linhas. O seu objetivo principal é representar de forma pictórica as conexões entre elementos de um determinado estudo.

Além disso, todo grafo é determinado por uma matriz de adjacência, que é uma representação tabular das ligações dos vértices em estudo. Uma matriz de adjacência é composta por $i \times i$ entradas, em que i é o número de vértices, que podem ser binárias ou inteiras. Portanto, quando as entradas da matriz são binárias marcamos como 0 a posição (i, j) quando não há ligação entre i e j , e 1 caso contrário. Por outro lado, quando as entradas são números inteiros elas representam o peso ou força da interação entre os vértices, de acordo com a indexação da matriz.

Definição 3.3 (Grafo) *Um grafo G é composto por um par ordenado (V, E) , em que V são os objetos em análise, representados por vértices, e E são as conexões que ligam os objetos em análise, representadas por um conjunto de arestas, i.e., $E \subset V \times V$.*

Além disso, os grafos podem ser classificados nas categorias orientados e não-orientados, dependendo de suas características. Os grafos orientados possuem uma direção específica de ligação dos objetos. Logo, suas arestas representam caminhos unidirecionais, ou seja, um vértice A é relacionado ao vértice B , mas não necessariamente B é relacionado ao A . De modo geral, O grafo G é orientado quando o conjunto E de arestas estabelece uma relação direcional entre os elementos do conjunto V de vértices. Já para grafos não-orientados, a relação estabelecida pelas arestas é simétrica e bidirecional, sem uma ordem específica de ligação e, além disso, a matriz de adjacência é binária e simétrica em relação à diagonal principal.

No contexto neurobiológico, se I é um conjunto finito de neurônios e F uma região finita amostrável, que é um subconjunto de I , então podemos utilizar um grafo de co-

nectividade para representar as interações dos neurônios de F . Primeiramente, vamos exemplificar a representação das interações dos neurônios por meio de um grafo não-orientado.

Exemplo 3.4 *Suponha uma região amostrável $F = \{1, 2, 3, 4\}$ composta por 4 neurônios. Nesse caso, F é considerado o conjunto de vértices de um grafo G , e $F \subset I$, sendo I um conjunto finito de neurônios. Agora, vamos definir o conjunto E , que representa as relações entre os neurônios em F , ou seja, E são as arestas do grafo G .*

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

Anteriormente, definimos que quando não há uma ordem específica de ligação das arestas entre os vértices o grafo é classificado como não-orientado, como é o caso de G . Sendo assim, a representação pictórica de G é dada da seguinte maneira:

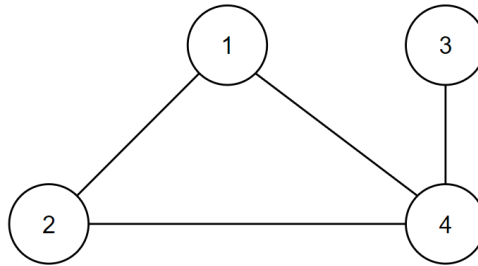


Figura 3.1: Representação pictórica do grafo $G = (F, E)$, em que $F = \{1, 2, 3, 4\}$ e $E = \{(1, 2), (1, 4), (2, 4), (3, 4)\}$.

Como dissemos anteriormente, o grafo do exemplo 3.4 não estabelece uma ordem específica para as ligações e, por isso, a matriz de adjacência que determina o grafo de conectividade é dada por

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

,logo, W é binária e simétrica em relação à diagonal principal.

Entretanto, o grafo do Exemplo 3.4 não é uma representação adequada das interações entre os neurônios, uma vez que não levamos em consideração os efeitos inibitórios e excitatórios que foram comentados na Seção 3.2. Para corrigirmos a omissão dessas características, podemos optar por um grafo orientado e atribuir pesos às conexões entre os vértices, i.e., às conexões entre os neurônios. Como para cada par de neurônios $i \in I$ e $j \in I$ estão associados os pesos sinápticos $\omega_{i \rightarrow j}$ e $\omega_{j \rightarrow i}$, podemos utilizá-los para representar os pesos a serem atribuídos às arestas (i, j) e (j, i) , respectivamente. Essa abordagem permite fortalecer ou enfraquecer as interações entre os vértices, tornando-a mais adequada para representar as conexões neuronais. De modo geral, vamos determinar um grafo de conectividade orientado utilizando uma matriz de adjacência com entradas não binárias, em que cada posição (i, j) da matriz é o peso sináptico associado à interação das células i e j . Além disso, no contexto neurobiológico, um grafo de conectividade orientado não possui ligações unidirecionais, uma vez que toda interação entre dois neurônios é estritamente bidirecional, entretanto, a matriz de adjacência não é necessariamente simétrica em relação à diagonal principal.

Seja $G = (V, E)$ um grafo orientado com pesos. Podemos utilizar uma matriz real W de ordem $|V| \times |V|$ tal que o elemento $w_{i \rightarrow j}$ que está na linha i e na coluna j da matriz W , representa a influência que o vértice i possui sobre o vértice j , para todo $i, j \in V$. Nesse sentido, dizemos que W é a matriz de pesos do grafo G . Note que se $\omega_{i \rightarrow j} = 0$, então os vértices i e j não estão conectados e, conseqüentemente, devemos ter $\omega_{j \rightarrow i} = 0$. Neste trabalho, cada entrada da matriz W representa o peso sináptico atribuído à interação do neurônio pré-sináptico (nas linhas) com o neurônio pós-sináptico (nas colunas).

Exemplo 3.5 *Suponha uma região amostrável $F = \{1, 2, 3, 4\}$ composta por 4 neurônios. Nesse caso, F é considerado o conjunto de vértices de um grafo G , e $F \subset I$, sendo I um conjunto finito de neurônios. Agora, vamos definir o conjunto E , que representa as relações entre os neurônios em F , ou seja, E são as arestas do grafo G .*

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

Para o grafo $G = (V, E)$, vamos associar uma matriz de pesos W dada por

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

Como dito anteriormente, cada entrada da matriz W representa o efeito do neurônio pré-sináptico (nas linhas) no neurônio pós-sináptico (nas colunas). Sendo assim, a representação pictórica do grafo ordenado ponderado G é dada por

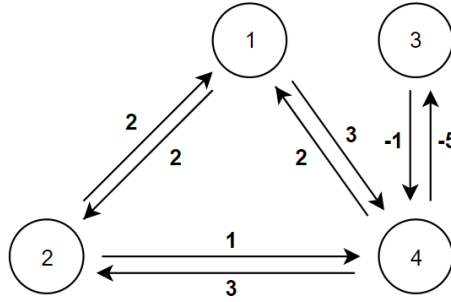


Figura 3.2: Representação pictórica do grafo orientado com pesos $G = (F, E)$, em que $F = \{1, 2, 3, 4\}$ e $E = \{(1, 2), (1, 4), (2, 4), (3, 4)\}$ e matriz de pesos W .

Como vimos anteriormente, no contexto neurobiológico todas as ligações são bidirecionais, o que é refletido na figura acima. Além disso, cada conexão é expressa por meio do seu peso sináptico, que pode ter efeito excitatório ou inibitório no neurônio pós-sináptico.

No Exemplo 3.5, ao utilizarmos um grafo orientado podemos notar que, de modo geral, os neurônios não necessariamente sofrem influência no potencial de membrana diretamente de todas as outras células da rede, como é o caso do vértice 3 que é influenciado apenas pela sinapse do neurônio 4. Assim, vamos chamar de **vizinhança de interação** o conjunto de todos os neurônios que estão diretamente conectados à célula em observação por uma sinapse.

Definição 3.6 (Vizinhança de interação) Dada uma região amostrável finita F de neurônios, com um grafo de conectividade $G = (F, E)$ e uma matriz de pesos sinápticos

W , a vizinhança de interação de um neurônio $i \in F$ é representada pelo conjunto $\mathcal{V}_i$, que é definido por

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

Desse modo, se avaliarmos novamente o grafo orientado G do exemplo 3.5, podemos definir a vizinhança de interação para cada neurônio $i \in F$, $F = \{1, 2, 3, 4\}$. Portanto, a partir da definição 3.6, a vizinhança no neurônio 1 engloba apenas as células 2 e 4, já que essas são as únicas interações diretas que o vértice 1 recebe. Ao fazermos esse mesmo processo sucessivamente para todos os neurônios $i \in F$ teremos $\mathcal{V}_1 = \{2, 4\}$; $\mathcal{V}_2 = \{1, 4\}$; $\mathcal{V}_3 = \{4\}$; $\mathcal{V}_4 = \{1, 2, 3\}$.

3.4 Estimação dos pesos sinápticos e do grafo de conectividade

Como vimos na Seção 3.3, um grafo de conectividade orientado G para j células neuronais é determinado a partir da matriz de adjacência W , em que as entradas $\omega_{j \rightarrow i}$ representam os pesos sinápticos da interação do neurônio j com o neurônio i , para $j \neq i$. Além disso, vimos que a vizinha de interação é composta pelos j neurônios da rede em que o peso sináptico $\omega_{j \rightarrow i} \neq 0$, $j \neq i$. Portanto, o grafo de conectividade orientado é, em suma, a representação pictórica da vizinhança de interação. Desse modo, estimar a vizinhança de interação de uma rede de neurônios implica na estimação do grafo de conectividade, o qual é determinado pela matriz de pesos sinápticos W . Em outras palavras, ao estimarmos os pesos sinápticos que compõem a matriz de pesos ou matriz de adjacência, podemos inferir sobre a conectividade dos neurônios da rede, o que, por sua vez, culmina na estimação do grafo de conectividade orientado e vizinhança de interação. Dessa forma, inferir a matriz de pesos sinápticos é equivalente a estimar a vizinhança de interação (Ost e Reynaud-Bouret, 2020).

3.4.1 Método da máxima verossimilhança com e sem regularização dos parâmetros

Nesse sentido, para estimarmos os coeficientes da matriz de pesos sinápticos, utilizaremos o método de estimação por máxima-verossimilhança com e sem regularização ℓ_1 . Logo, escreveremos uma função log-verossimilhança associada ao vetor de pesos $W^{(i)}$ para cada neurônio da rede. A utilização da log-verossimilhança como método para a estimação da matriz de pesos sinápticos e, conseqüentemente, do grafo de conectividade é justificada pelo fato de que, no instante de tempo T , os neurônios disparam de forma independente ao outro quando condicionados ao mesmo histórico passado da rede.

Sejam N e T números inteiros positivos e considere um cenário onde temos um conjunto finito de neurônios I e uma região amostrável finita $F \subset I$, contendo N neurônios. Essa região é observada ao longo de um intervalo de tempo, que é discretizado em T janelas temporais. Dessa forma, dada a configuração $\mathbf{x}_1^T := (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T) \in \{0, 1\}^{N \times T}$, derivada da cadeia estocástica $\mathbf{X}$. Nosso objetivo é utilizar a configuração $\mathbf{x}_1^T$ para estimar a vizinhança de interação $\mathcal{V}_i$ para qualquer neurônio $i \in F$. Entretanto, como discutimos anteriormente, podemos estimar as vizinhanças $\mathcal{V}_i, i \in F$, a partir da estimação da matriz de pesos sinápticos

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

Considerando para cada neurônio $i \in F$ que o processo $\mathbf{X}^{(i)}$ que modela sua atividade neuronal é uma cadeia de alcance variável finita com profundidade no máximo k , podemos escrever a função de verossimilhança associada ao problema, condicionada ao estados iniciais $\mathbf{x}_{-k+1}^0$, como sendo

$$\begin{aligned} \mathcal{L}(W, \mathbf{x}_1^T) &:= \prod_{t=1}^T \prod_{i=1}^N P\left(X_t^{(i)} = x_t^{(i)} \mid \mathbf{X}_{-\infty}^{t-1} = \mathbf{x}_{-\infty}^{t-1}\right) \\ &:= \prod_{t=1}^T \prod_{i=1}^N \phi(V_{t-1}^{(i)})^{x_t^{(i)}} \left(1 - \phi(V_{t-1}^{(i)})\right)^{1-x_t^{(i)}}. \end{aligned}$$

Consequentemente, a função de log-verossimilhança associada é

$$\begin{aligned}\ell(W, \mathbf{x}_1^T) &= \sum_{i=1}^N \sum_{t=1}^T \left[x_t^{(i)} \log \left(\frac{\phi(V_{t-1}^{(i)})}{1 - \phi(V_{t-1}^{(i)})} \right) + \log \left(1 - \phi(V_{t-1}^{(i)}) \right) \right] \\ &= \sum_{i=1}^N \ell^{(i)}(W^{(i)}, \mathbf{x}_1^T),\end{aligned}$$

em que para cada neurônio $i \in F$, a função de log-verossimilhança associada ao vetor de pesos sinápticos $\mathbf{w}^{(i)} = (\omega_{1 \rightarrow i}, \omega_{2 \rightarrow i}, \dots, \omega_{N \rightarrow i})^\top$, dada a configuração $\mathbf{x}_1^T$, é de tal forma que

$$\begin{aligned}\ell(W^{(i)}, \mathbf{x}_1^T) &:= \sum_{t=1}^T \left[x_t^{(i)} \log \left(\frac{\phi(V_{t-1}^{(i)})}{1 - \phi(V_{t-1}^{(i)})} \right) + \log \left(1 - \phi(V_{t-1}^{(i)}) \right) \right] \\ &:= \sum_{t=1}^T \left[x_t^{(i)} \left(V_{t-1}^{(i)} \right) - \log \left(1 + \exp \left(V_{t-1}^{(i)} \right) \right) \right],\end{aligned}$$

em que $V_{t-1}^{(i)}$ é o potencial de membrana do neurônio i no instante de tempo $t - 1$, i.e.,

$$V_{t-1}^{(i)} = \begin{cases} 0, & \text{se } X_{t-1}^{(i)} = 1, \\ \sum_{j=1}^N \omega_{j \rightarrow i} \sum_{s=L_t^{(i)}+1}^{t-1} X_s^{(j)}, & \text{caso contrário.} \end{cases}$$

Com isso, podemos realizar a estimação dos parâmetros do modelo via função de log-verossimilhança. Entretanto, para inferirmos sobre a conectividade da cadeia neuronal vamos considerar o disparo ou não disparo dos neurônios da cadeia $\mathbf{X}$ em cada instante de tempo t como covariáveis do nosso problema. Sendo assim, o número de covariáveis é, em geral, ainda maior que a quantidade de observações, o que afeta diretamente na esparsidade do nosso modelo. Além disso, a grande quantidade de parâmetros prejudica a estimação via função de log-verossimilhança, já que o método depende da inversão de uma matriz de dados, que com o aumento da dimensionalidade não possui posto completo. Portanto, neste trabalho, estudaremos, em um primeiro momento, a estimação do grafo de conectividade a partir do método máxima verossimilhança considerando apenas microcircuitos, ou seja, redes com poucos neurônios, garantindo que não haja um grande número de parâmetros a serem estimados e, assim, reduzindo a dimensionalidade do pro-

blema. Além disso, o nosso modelo de rede neuronal considera neurônios estocásticos de memória com alcance variável, porém finita, logo, são um caso particular das cadeias de Markov em que algumas probabilidades de transição são iguais. Desse modo, no cenário de microcircuitos também vamos variar o tamanho das amostras e explorar diferentes valores k (tamanho máximo da memória do neurônio) para avaliarmos os impactos no poder preditivo do modelo gerado. Além disso, também aplicaremos a metodologia da máxima verossimilhança junto ao método de seleção que propomos, o qual é baseado na distância euclidiana entre o vetor das probabilidades preditas (ver Subseção 3.4.2) nas mesmas condições de tamanho de amostra, cenário e ordem da memória para compararmos o poder preditivo com o do modelo gerado a partir da máxima verossimilhança sem regularização. Por fim, vale ressaltar que a estimação via máxima verossimilhança é um método não enviesado, independentemente das condições em que tenha sido aplicado.

Posteriormente, vamos corrigir o problema de alta dimensionalidade empregando uma regularização ℓ_1 nos coeficientes do modelo, com o propósito de escolher as conexões neuronais que formarão o grafo de conectividade, assim como feito por Tibshirani (1996) em um contexto de regressão linear. Esse método é conhecido como LASSO (*Least Absolute Shrinkage and Selection Operator*) e seu objetivo principal é encontrar os coeficientes dos preditores que minimizam a soma dos quadrados dos resíduos, de forma semelhante a como fazemos na regressão linear comum, mas com uma restrição adicional. Neste caso, temos

$$\left(\hat{W}^{(i)}\right) := \arg \min_{W^{(i)}} \left(-\ell \left(W^{(i)}, \mathbf{x}_1^T \right) + \lambda_i \sum_{j=1}^N |\omega_{j \rightarrow i}| \right), \quad (3.7)$$

em que $\ell \left(W^{(i)}, \mathbf{x}_1^T \right)$ é a função de log-verossimilhança, $\omega_{j \rightarrow i}$ são os pesos sinápticos. No método LASSO, a restrição adicional que comentamos é construída a partir soma dos valores absolutos dos coeficientes multiplicada por um parâmetro de regularização, dado por λ_i na Equação (3.7), impondo uma restrição no valor dos coeficientes e fazendo com que sejam significativamente reduzidos, chegando a zero em alguns casos. Além disso, o λ_i é escolhido por validação-cruzada e busca minimizar o risco de predição, logo, selecionamos um λ_i que gere equilíbrio entre a tendência do modelo e a sensibilidade a variações nos dados, ou seja, a constante λ_i é escolhida de modo que equilibre o viés e a variância. Nesse cenário, conforme a expressão (3.7), o nosso objetivo é minimizar a log-verossimilhança negatizada sob a norma ℓ_1 dos pesos sinápticos $\omega_{j \rightarrow i}$, multiplicados pelo parâmetro de

regularização λ , para determinarmos os estimadores. Desse modo, com a regularização ℓ_1 do método de máxima-verossimilhança, conseguimos resolver o problema de alta dimensionalidade. Sendo assim, uma vez contornada a questão da grande quantidade de parâmetros, vamos utilizar o LASSO para estimar os coeficientes do modelo para a atividade neuronal em cenários mais abrangentes em comparação aos que utilizaremos para a estimação via máxima-verossimilhança, ou seja, teremos um número maior de neurônios na rede. Além disso, assim como no caso dos microcircuitos, vamos aplicar o modelo em diferentes tamanhos de amostra e valores de k (tamanho máximo da memória do neurônio). Ainda, para fins comparativos, também aplicaremos a estimação por máxima verossimilhança junto ao método de seleção de variáveis que propomos nesse trabalho (ver Subseção 3.4.2) nas mesmas condições de experimento.

3.4.2 Método proposto

Seja F uma região finita amostrável composta por N neurônios cuja atividade é registrada ao longo de um intervalo de tempo finito discretizado em T janelas temporais. Dada a configuração $\mathbf{x}_1^T := (\mathbf{x}_1, \dots, \mathbf{x}_T) \in \{0, 1\}^{N \times T}$ da região F , queremos estimar a vizinhança de interação $\mathcal{V}_i$ para cada neurônio $i \in F$. O objetivo do método de seleção estatística que vamos propor nesta seção é identificar o conjunto $\mathcal{V}_i$ a partir dos dados de uma maneira consistente.

A grosso modo, no procedimento de seleção estatística que consideramos, observamos a atividade da rede dentro de uma região amostral finita durante um intervalo de tempo finito. Para cada neurônio i na região amostrada, estimamos a sua probabilidade de disparo dada a atividade de todos os outros neurônios desde o último disparo do neurônio i . Para cada neurônio $j \neq i$, introduzimos uma medida de sensibilidade dessa probabilidade de disparo condicional com respeito a ausência do neurônio pré-sináptico j . Se esta medida de sensibilidade é estatisticamente pequena, concluímos que o neurônio j não pertence à vizinhança de interação do neurônio i . Na sequência, definimos rigorosamente esse procedimento estatístico.

Para cada neurônio $i \in F$, vamos definir a probabilidade estimada do neurônio i disparar no instante de tempo t , dada a atividade dos neurônios pertencentes ao subconjunto G de F por

$$\hat{P}_t^{(i)}(G) := \phi \left((1 - x_{t-1}^{(i)}) \left(\sum_{j \in G} \hat{\omega}_{j \rightarrow i} \sum_{s=L_t^{(i)}+1}^{t-1} x_s^{(i)} \right) \right),$$

em que $\hat{\omega}_{j \rightarrow i}$ é a estimativa de máxima verossimilhança do peso sináptico entre o neurônio j e o neurônio i . Essa probabilidade estimada é um preditor da atividade do neurônio i no instante t quando influenciada pela atividade de todos os neurônios pertencentes a G . O vetor $\hat{\mathbf{P}}_G^{(i)} := (\hat{P}_1^{(i)}(G), \hat{P}_2^{(i)}(G), \dots, \hat{P}_T^{(i)}(G))$ é, portanto, o vetor das probabilidades de disparo preditas ao longo das T janelas temporais.

Em particular, quando $G = F - \{j\}$, i.e., o conjunto de todos neurônio de F exceto o neurônio j , podemos definir uma medida de sensibilidade $d : [0, 1]^T \times [0, 1]^T \rightarrow [0, +\infty[$ tal que

$$d(\hat{\mathbf{P}}_F^{(i)}, \hat{\mathbf{P}}_G^{(i)}) := \left\| \hat{\mathbf{P}}_F^{(i)} - \hat{\mathbf{P}}_G^{(i)} \right\|_{\ell^2},$$

em que $\|\cdot\|_{\ell^2}$ denota a norma Euclidiana. Nesse contexto, dizemos que o neurônio j é um vizinho do neurônio i quando $d(\hat{\mathbf{P}}_F^{(i)}, \hat{\mathbf{P}}_G^{(i)}) > \epsilon$, para algum ponto de corte $\epsilon > 0$ fixado. Portanto, a vizinhança estimada do neurônio $i \in F$ é dada por

$$\hat{\mathcal{V}}_i := \left\{ j \in F - \{i\} : d(\hat{\mathbf{P}}_F^{(i)}, \hat{\mathbf{P}}_G^{(i)}) > \epsilon \right\}.$$

Por definição, se os neurônios $i \in F$ e $j \in F$ não são vizinhos, então os pesos sinápticos entre eles são $\omega_{i \rightarrow j} = \omega_{j \rightarrow i} = 0$. Desde que o método da máxima verossimilhança seja consistente em estimar os pesos sinápticos, temos, para amostras suficientemente grandes, $\hat{\omega}_{i \rightarrow j} \approx \hat{\omega}_{j \rightarrow i} \approx 0$. Portanto, para um ponto de corte ϵ adequado, temos $i \notin \hat{\mathcal{V}}_j$ e $j \notin \hat{\mathcal{V}}_i$, o que faz do método de seleção estatística proposto consistente. É essa consistência que vamos explorar no estudo de simulação proposto no próximo capítulo.

3.5 Predição da atividade neuronal

A predição da atividade neuronal desempenha um papel importante na compreensão dos mecanismos complexos que governam o cérebro humano. Ao antecipar padrões de atividade neuronal, os pesquisadores conseguem melhorar a sua compreensão a cerca dos processos cognitivos, tomada de decisões e até mesmo distúrbios neurológicos. Neste trabalho, a predição da atividade neuronal dar-se-á a partir dos preditores obtidos pelos métodos de estimação descritos na seção anterior.

Dado um conjunto de treino $\mathbf{x}_1^T = (\mathbf{x}_1, \dots, \mathbf{x}_T)$, $\mathbf{x}_i \in \{0, 1\}^N$, gerado a partir do modelo neuronal definido na Seção 3.2, desejamos prever sequencialmente os próximos S valores do conjunto de teste $\mathbf{x}_{T+1}^{T+S} = (\mathbf{x}_{T+1}, \dots, \mathbf{x}_{T+S})$. Para cada $t = 1, \dots, S$, a

predição de $\mathbf{x}_{T+i}$, condicionada a todo passado $\mathbf{x}_1^{T+i-1}$, é dada pela distribuição condicional $\hat{P}_t(\mathbf{x} | \mathbf{x}_1^{T+t-1})$, $\mathbf{x} \in \{0, 1\}^N$, em que $\hat{P}_t$ é a probabilidade estimada no passo t do conjunto de teste.

Uma das suposições adotadas neste trabalho é que, condicionado a todo passado, os neurônios disparam independentemente. Nesse sentido, podemos escrever, para cada $t = 1, 2, \dots, S$,

$$\hat{P}_t(\mathbf{x} | \mathbf{x}_1^{T+i-1}) = \prod_{i=1}^N \hat{P}_t(x^{(i)} | \mathbf{x}_1^{T+t-1}).$$

Assim, a performance do preditor resultante pode ser avaliada a partir de medidas baseadas na matriz de confusão.

As medidas baseadas na matriz de confusão consideradas neste trabalho são:

- **Sensibilidade.** A sensibilidade (S) mede a capacidade do modelo em identificar corretamente os casos positivos entre todos os casos verdadeiramente positivos.
- **Especificidade.** A Especificidade (E) mede a proporção de casos negativos que foram corretamente identificados pelo modelo em relação ao total de casos negativos verdadeiros.
- **Acurácia.** A Acurácia (ACC) mede a proporção de previsões corretas feitas pelo modelo em relação ao número total de amostras.
- **Valor preditivo positivo.** O Valor Preditivo Positivo (VPP) mede a proporção de casos positivos previstos corretamente em relação ao total de casos classificados como positivos pelo modelo.
- **Valor preditivo negativo.** O Valor Preditivo Negativo (VPN) mede a proporção de casos negativos que foram corretamente identificados pelo modelo em relação ao total de casos classificados como negativos pelo modelo.
- **G-média:** A G-média é a média geométrica da sensibilidade e da especificidade.
- **Coeficiente de correlação de Mathews.** O coeficiente de correlação de Mathews (MCC) é a correlação entre o valor observado na amostra de teste e os valores preditos pelo classificador, a partir da utilização dos valores encontrados na matriz de confusão.

Capítulo 4

Aplicação Computacional

Registros da atividade neuronal indicam que os disparos elétricos dos neurônios podem ocorrer espontaneamente e de maneira irregular, inclusive apresentando 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). Essas observações empíricas apontam para a presença de ruídos nos sinais elétricos, sugerindo a necessidade de uma estrutura estocástica na descrição matemática dos fenômenos neurais. Neste trabalho, a atividade de cada neurônio é, portanto, representada por um processo estocástico a tempo discreto, cujas variáveis aleatórias indicam se houve ou não um disparo em um dado instante de tempo. Propomos a adoção de um modelo do tipo Galves e Löcherbach (2016) para descrever a dinâmica neural ao longo do tempo. Por meio desse modelo, geramos computacionalmente uma rede de neurônios estocásticos com base em premissas neurobiologicamente plausíveis e realizamos a estimação do grafo de conectividade a partir do método da máxima verossimilhança.

Mesmo em um microcircuito com um número reduzido de neurônios, o grafo de conectividade pode ser determinado por um grande número de parâmetros. Modelos com um grande número de parâmetros livres podem descrever uma ampla variedade de fenômenos, mas a concordância com os dados disponíveis não garante a qualidade do modelo. Isso pode indicar apenas que o modelo se ajusta bem aos dados existentes, mas falha em generalizar para novas situações. Para abordar essa questão, neste trabalho, exploramos dois métodos de seleção de modelo. O primeiro envolve a aplicação de uma regularização ℓ_1 nos coeficientes do modelo, um método conhecido na literatura estatística. Até onde sabemos, nenhum estudo robusto de simulação foi conduzido considerando esse método para o modelo neural específico abordado aqui. O segundo método é uma contribuição

original desta monografia.

A avaliação da performance do método de estimação proposto é realizada por meio de réplicas de simulações de Monte Carlo. Utilizamos essas simulações para comparar nossas estimativas com os valores dos parâmetros fixados por meio do erro quadrático médio empírico. Nesse contexto, evidenciamos a consistência do estimador proposto, assim como a eficácia dos métodos de seleção de variáveis. Por outro lado, avaliamos a performance dos preditores obtidos a partir do ajuste do modelo proposto sob os diferentes contextos de seleção de variáveis por meio de medidas obtidas a partir da matriz de confusão tais como o coeficiente de correlação de Matthews (MCC), a G-Média, a sensibilidade, a especificidade, o valor preditivo positivo (*VPP*) e o valor preditivo negativo (*VPN*). Os resultados obtidos indicam que, em cenários de microcircuitos, o método proposto se sobressai em termos de precisão na previsão de valores, visto que inclui seleção de variáveis, porém sem introduzir viés nas estimativas. Por outro lado, em macrocircuitos, sua performance é comparável à abordagem de máxima verossimilhança com regularização ℓ_1 dos parâmetros.

4.1 Hipóteses e cenários

Na neurociência, existem diversas hipóteses sobre o armazenamento e uso da informação pelos neurônios, assim como o impacto dos estímulos externos em sua capacidade de trocar informações. Nesse sentido, conduzimos o estudo de simulação considerando sete cenários distintos. Cada um estabelecido a partir de princípios neurobiológicos distintos, na tentativa de levar em consideração a complexidade e a variabilidade inerentes ao comportamento neuronal. Dessa forma, conseguimos avaliar a robustez das conclusões obtidas, diminuindo o risco de simplificações excessivas que poderiam comprometer a validade e a aplicabilidade dos resultados simulados em um contexto de dados eletrofisiológicos.

Neste trabalho, as amostras das redes neuronais foram obtidas a partir da simulação de modelos do tipo Galves e Löcherbach. As redes se diferem em número de neurônios, em relação à esparsidade (número de neurônios conectados) e quanto ao balanceamento e desbalanceamento (número de neurônios excitatórios e inibitórios). Nesse sentido, temos os seguintes cenários:

- **Cenário 1.** A matriz de pesos sinápticos fixada foi construída de forma que dentre as 20 conexões possíveis entre neurônios na rede, temos a seguinte distribuição: 10%

das conexões estão desconectadas, ou seja, 2 neurônios não possuem conexão entre si. Além disso, 70% das conexões são inibitórias (14 das 20 conexões), enquanto 20% são excitatórias (4 das 20 conexões). A matriz de pesos sinápticos, nesse caso, é dada por

	Neurônio 1	Neurônio 2	Neurônio 3	Neurônio 4	Neurônio 5
Neurônio 1	0	0	-1	-1	-1
Neurônio 2	0	0	-1	-1	-1
Neurônio 3	-1	-1	0	4	4
Neurônio 4	-1	-1	4	0	-1
Neurônio 5	-1	-1	4	-1	0

Figura 4.1: Matriz de pesos sinápticos W para o cenário 1.

O cenário 1 é considerado patológico, uma vez que, geralmente, observamos uma proporção de 4 neurônios excitatórios para cada neurônio inibitório, como é observado no córtex.

- **Cenário 2.** A matriz de pesos sinápticos fixada foi construída de forma que dentre as 20 possíveis conexões entre neurônios na rede, consideramos 10% desconectados (2 neurônios sem conexão entre si), 50% de conexões inibitórias (10 das 20 conexões) e 40% de conexões excitatórias (8 das 20 conexões). A matriz de pesos sinápticos, nesse caso, é dada por

	Neurônio 1	Neurônio 2	Neurônio 3	Neurônio 4	Neurônio 5
Neurônio 1	0	0	4	-1	-1
Neurônio 2	0	0	-1	-1	-1
Neurônio 3	4	-1	0	4	4
Neurônio 4	-1	-1	4	0	4
Neurônio 5	-1	-1	4	4	0

Figura 4.2: Matriz de pesos sinápticos W para o cenário 2.

O cenário 2 é também considerado patológico, entretanto, aumentamos razoavelmente a quantidade de neurônios excitatórios na rede quando comparado ao cenário 1.

- **Cenário 3.** A matriz de pesos sinápticos fixada foi construída de forma que dentre as 20 conexões possíveis entre os neurônios na rede, contemplamos 10% como desconectadas (2 neurônios sem conexão entre si), 30% como conexões inibitórias (6 das 20 conexões) e 60% como conexões excitatórias (12 das 20 conexões). A matriz de pesos sinápticos, nesse caso, é dada por

	Neurônio 1	Neurônio 2	Neurônio 3	Neurônio 4	Neurônio 5
Neurônio 1	0	0	4	4	4
Neurônio 2	0	0	-1	-1	-1
Neurônio 3	4	-1	0	4	4
Neurônio 4	4	-1	4	0	4
Neurônio 5	4	-1	4	4	0

Figura 4.3: Matriz de pesos sinápticos W para o cenário 3.

O cenário 3 apresenta uma configuração mais próxima das proporções de neurônios excitatórios e inibitórios observadas em redes neuronais do córtex.

- **Cenário 4.** A matriz de pesos sinápticos fixada foi construída de forma que dentre as 20 possíveis conexões entre neurônios distintos da rede, consideramos 10% como desconectados (2 neurônios sem conexão entre si), 20% como conexões inibitórias (4 das 20 conexões) e 70% como conexões excitatórias (14 das 20 conexões). A matriz de pesos sinápticos, nesse caso, é dada por

	Neurônio 1	Neurônio 2	Neurônio 3	Neurônio 4	Neurônio 5
Neurônio 1	0	0	4	4	4
Neurônio 2	0	0	4	-1	-1
Neurônio 3	4	4	0	4	4
Neurônio 4	4	-1	4	0	4
Neurônio 5	4	-1	4	4	0

Figura 4.4: Matriz de pesos sinápticos W para o cenário 4.

O cenário 4 é o que possui as proporções de neurônios excitatórios e inibitórios mais próximas daquelas observadas em redes neuronais reais.

- **Cenário 5.** A matriz de pesos sinápticos fixada foi construída de forma que dentre as 20 possíveis conexões entre neurônios distintos da rede, consideramos 50% como

desconectados (10 das 20 conexões), 10% como conexões inibitórias (2 das 20 conexões) e 40% como conexões excitatórias (8 das 20 conexões). A matriz de pesos sinápticos, nesse caso, é dada por

	Neurônio 1	Neurônio 2	Neurônio 3	Neurônio 4	Neurônio 5
Neurônio 1	0	0	0	0	0
Neurônio 2	0	0	0	4	-1
Neurônio 3	0	0	0	4	4
Neurônio 4	0	4	4	0	4
Neurônio 5	0	-1	4	4	0

Figura 4.5: Matriz de pesos sinápticos W para o cenário 5.

O cenário 5 é patológico uma vez que considera rede com muito mais desconexões do que conexões.

- **Cenário 6.** A matriz de pesos sinápticos fixada foi construída de forma que dentre as 20 possíveis conexões (grafo de conectividade orientado) entre neurônios distintos da rede, consideramos 90% como desconectados (18 das 20 conexões) e 10% como conexões excitatórias (1 das 20 conexões). A matriz de pesos sinápticos, nesse caso, é dada por

	Neurônio 1	Neurônio 2	Neurônio 3	Neurônio 4	Neurônio 5
Neurônio 1	0	0	0	0	0
Neurônio 2	0	0	0	0	0
Neurônio 3	0	0	0	0	0
Neurônio 4	0	0	0	0	4
Neurônio 5	0	0	0	4	0

Figura 4.6: Matriz de pesos sinápticos W para o cenário 6.

O cenário 6 é ainda mais patológico que o cenário 5, uma vez que aumentamos o número de desconexões.

- **Cenário 7.** Nesse caso, fixamos um macrocircuito com 20 neurônios, sendo no total 380 parâmetros a serem estimados. Para isso, vamos considerar 60% das 380 possíveis conexões como desconexões e manter a proporção de 4 neurônios excitatórios para cada neurônio inibitório para as conexões restantes.

4.2 Simulação

O estudo da conectividade neuronal por meio de simulações é uma abordagem prática que facilita a compreensão das conexões sinápticas em condições controladas e reproduzíveis, sem a necessidade de depender exclusivamente de amostras biológicas. Além disso, a partir de simulações é possível realizar experimentos que podem ser impraticáveis na realidade, além de testar hipóteses e cenários específicos. Isso oferece uma abordagem flexível e adaptável para explorar o funcionamento das conexões sinápticas em diferentes contextos e condições experimentais.

As simulações foram feitas em R, versão 8.16, e o código foi disponibilizado publicamente no github do autor ¹ juntamente com a documentação para uso futuro da comunidade científica. Nesse contexto, o primeiro passo é gerar uma amostra dos disparos neuronais de uma rede de neurônios. Dada a amostra gerada, realizamos a estimação do grafo de conectividade neuronal a partir do método de máxima verossimilhança considerando dois métodos de seleção de modelos: com regularização ℓ_1 dos coeficientes e o que propomos na Seção 3.4.2. Contudo, considerar apenas uma amostra não nos fornece informação suficiente a respeito do desempenho dos métodos propostos 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.

4.2.1 Geração da atividade neuronal

Em todos os cenários, vamos avaliar se conforme aumentamos o tamanho da amostra, as estimativas dos pesos sinápticos tornam-se mais próximas dos valores fixados. Em outras palavras, estamos interessados em estudar, computacionalmente, a consistência (em média) dos estimadores propostos nesta monografia. Nesse sentido, fixamos os seguintes tamanhos de amostra: $T = 120$, $T = 1.200$, $T = 12.000$ e $T = 24.000$. Além disso, vamos estudar o poder preditivo do modelo neuronal quando ajustado considerando os métodos de seleção propostos. Para realizar esse estudo, fixamos $T = 30.000$.

Neste trabalho, vamos estudar microcircuitos com 5 neurônios e um macrocircuito com 20 neurônios de modo que a função de ativação utilizada seja a taxa de disparo logística

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

ϕ . A simulação da atividade dos neurônios sintéticos acontece a partir da comparação de números pseudo-aleatórios gerados a partir de uma distribuição uniforme no intervalo $[0, 1]$ com as probabilidades de disparo que definem o modelo neural que estamos adotando. Nesse sentido, temos o seguinte pseudo-algoritmo para geração das redes neuronais.

Algoritmo. Pseudo-algoritmo para a geração das redes 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

Analisando o pseudo-algoritmo, notamos que para inicializá-lo é necessário definir a atividade inicial dos neurônios. Os neurônios corticais levam, em média, de 1 a 3 milissegundos (ms) para disparar. Nesse sentido, na prática, em geral, é fixado um tamanho de bin de 0,30ms (ver Softky Koch (1993)). Logo, vamos assumir que estamos trabalhando com neurônios corticais e fixar os quatro primeiros instantes iniciais dos neurônios das redes.

4.2.2 Estimação dos pesos sinápticos e do grafo de conectividade

Uma vez gerada as amostras das atividades dos neurônios da rede, podemos utiliza-las para realizar a estimação dos pesos sinápticos e, conseqüentemente, do grafo de conectividade. Neste estudo, adotamos três diferentes métodos de estimação baseados no método da máxima verossimilhança: , sendo eles: (i) sem considerar regularização dos parâmetros, (ii) considerando regularização ℓ_1 dos coeficientes do modelo e (iii) utilizando o método de seleção que propomos, o qual é baseado na distância euclidiana entre o vetor das probabilidades preditas. Cada uma das metodologias está explicada e detalhada Capítulo 3.

Todos os métodos seguem o mesmo princípio, ou seja, para cada neurônio $i \in F$ e para cada tamanho amostral T fixado, consideramos o vetor de respostas, a matriz de covariáveis e o vetor de parâmetros, dados, respectivamente, por

$$y_T^{(i)} = \begin{bmatrix} x_1^{(i)} \\ x_2^{(i)} \\ \vdots \\ x_T^{(i-1)} \end{bmatrix}, \quad A_T^{(i)} = \begin{bmatrix} a_1^{(1)} & a_1^{(2)} & \cdots & a_1^{(N)} \\ a_2^{(1)} & a_2^{(2)} & \cdots & a_2^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ a_T^{(1)} & a_T^{(2)} & \cdots & a_T^{(N)} \end{bmatrix}, \quad W_{\rightarrow i} = \begin{bmatrix} W_{1 \rightarrow i} \\ W_{2 \rightarrow i} \\ \vdots \\ W_{N \rightarrow i} \end{bmatrix},$$

em que

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

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

Assim, para cada neurônio $i \in F$, temos uma matriz de covariáveis $A_T^{(i)}$, pois $a_t^{(i)}$ varia conforme o neurônio i considerado. Em uma rede com N neurônios, temos N matrizes de dados distintas serão utilizadas em N modelos independentes. Esses modelos são independentes, pois essa é uma das suposições que fizemos ao definir o modelo neuronal adotado nesta monografia.

Os códigos foram desenvolvidos exclusivamente em linguagem R, utilizando a biblioteca `glmnet` para o ajuste do modelo. Na abordagem com regularização ℓ_1 , aplicamos a função `cv.glmnet` do pacote, que, além de produzir estimativas, realiza internamente a validação cruzada para encontrar o valor ótimo de λ . Para o modelo sem regularização, escolhemos a função `glm` com a especificação da família "binomial" nos parâmetros, assegurando que a modelagem não selecione variáveis e, simultaneamente, realize a estimação

por máxima verossimilhança.

No contexto do método proposto, para implementarmos a seleção de variáveis conforme detalhado em 3.4.2, desenvolvemos uma função específica que identifica os neurônios não pertencentes à vizinhança de todas as células da rede. Com essa função, compilamos os resultados de saída e modificamos a matriz de dados utilizada para o ajuste, incorporando as vizinhanças pré-selecionadas. Posteriormente, aplicamos o `glm` da mesma forma que no método não regularizado para obtermos $\hat{W}$.

O pseudo-algoritmo básico para nossa função de seleção de conexões é apresentado abaixo:

Algoritmo. Pseudo-algoritmo para a encontrar neurônios fora da vizinhança.

- **Entrada:** Índice do neurônio, matriz de probabilidades de disparo calculadas e matriz resposta binária.
- **Saída:** Neurônios que não fazem parte da vizinhança de i

Função(fora_vizinhanca_interna(i , $\mathbf{matriz}_i$, y_i))

- Para cada y_i e $\mathbf{matriz}_i$, ajuste o modelo completo em relação a i
- Armazene os valores ajustados
- Crie um dicionário vazio

Para j em $1 : n$

Se $j \neq i$

$\mathbf{matriz_sem_j} \leftarrow \mathbf{matriz}_i[:, -\{j\}]$

- Ajuste o modelo defasado
 - Armazene as probabilidades do defasado
 - Calcule a distância dos valores ajustados em relação ao modelo completo
- $\mathbf{dic_distancias}[j] \leftarrow \text{dist}$

Fim se

Fim para

Fim função

- Verifique valores em que as distâncias são menores que um valor ϵ
- Retorne as chaves dos valores verificados como verdadeiro

Para definirmos o valor de ϵ que delimita a vizinhança neuronal, conduzimos um experimento em todos os cenários predefinidos. Em cada configuração de matriz de pesos sinápticos determinada, calculamos e armazenamos a distância entre os modelos defasados e completos. Posteriormente, analisamos esses dados para identificar um valor que, de maneira geral, serviria como ponto de corte na definição das conexões sinápticas. Os resultados mostraram que, em grande parte dos testes, a distância entre os modelos ajustados era inferior a aproximadamente 0,6 para neurônios desconectados. Portanto, para otimizarmos a qualidade da seleção de variáveis, optamos por fixar ϵ em 0,6.

Além disso, uma 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.

4.2.3 Predição da atividade neuronal

A capacidade de um modelo prever a atividade neuronal não apenas oferece uma visão mais profunda da organização funcional do cérebro, mas também tem implicações significativas em campos como a neurociência computacional, permitindo a elaboração de algoritmos eficientes capazes de tornar mais compreensível o que aconteceria com a rede em cenários difíceis, e algumas vezes impossíveis, de serem reproduzidos na vida real.

Os experimentos conduzidos para fins de predição são baseados em amostras de tamanho 30.000. Inicialmente, as amostras foram divididas em dois subconjuntos: o conjunto de treino e o conjunto de teste. A composição desses subconjuntos foi considerada a partir de três níveis de proporções: (i) 10% para treino e 90% para teste, (ii) 50% para treino e 50% para teste e (iii) 90% para treino e 10% para teste. Uma vez ajustado o modelo em cada um dos conjuntos de treino, a sua performance na predição da atividade futura da rede dar-se-á por medidas baseadas na matriz de confusão: acurácia, sensibilidade, especificidade, valor preditivo positivo, valor preditivo negativo, G-média e coeficiente de correlação de Mathews.

A seguir, apresentamos o pseudo-algoritmo para a predição de uma rede neuronal.

Algoritmo. Pseudo-algoritmo para a execução do processo de predição.

- **Entrada:** matriz simulada X referente ao conjunto de treino, índice do neurônio e instante de tempo.

- **Saída:** Matriz de confusão calculada com base nas probabilidades de disparo estimadas.

-**Gerar** a cadeia X com o ajuste no potencial de membrana.

-**Separar** a cadeia X em treino e teste.

-**Estimar** $\hat{W}$ a partir da amostra de treino.

Função(Potencial_membrana_com_matriz_estimada(X, i, t))

potencial $\leftarrow 0$

ultimoX $\leftarrow$ **último disparo**(**matriz_x**(X, i), 1, t)

Se (*ultimoX* = t)

potencial $\leftarrow \log \left(\frac{0,05}{0,95} \right)$

Caso contrário

pesosX $\leftarrow$ **Linhas menos i**(X, i)[, (*ultimoX* + 1) : t]

potencial $\leftarrow \sum \left(\frac{1}{2^{t-\text{ultimoX}}} \times \text{Soma das linhas}(\text{pesosX}) \times \text{Linhas menos i}(\hat{w}, i)[, i] \right)$

Fim se

Retorne *potencial*

Fim função

-**Calcular** a probabilidade estimada para cada tempo t do teste.

-**Binarizar** os resultados a partir de um ponto de corte.

-**importar pacote caret**().

-**Calcular** a matriz de confusão.

Na próxima seção, descrevemos os resultados obtidos com o estudo de simulação proposto.

4.3 Resultados

Nesta seção, apresentamos uma discussão a cerca dos resultados dos métodos de estimação propostos no Capítulo 3. Apresentamos as tabelas e gráficos dos resultados da estimação dos pesos sinápticos no Apêndice A, da estimação do grafo de conectividade no Apêndice B e da predição da atividade da rede neuronal no Apêndice C. Assim, permitimos uma leitura mais fluida e focada nas discussões que gostaríamos de tratar nesta monografia.

Em cada cenário especificado na Seção 4.1, foram geradas 100 réplicas Monte Carlo para cada tamanho amostral. Para cada réplica, os pesos sinápticos e o grafo de conectividade foram estimados pelos três métodos de estimação considerados neste trabalho. A avaliação da eficácia desses métodos na estimação dos pesos sinápticos baseou-se no erro quadrático médio empírico para as redes com 5 neurônios e na distância euclidiana média entre as matrizes estimadas e a matriz original de pesos sinápticos para a rede com 20 neurônios. Enquanto isso, a avaliação do desempenho dos métodos na estimação do grafo de conectividade foi realizada por meio da análise das proporções de conexões sinápticas corretamente identificadas, bem como da proporção de conexões excitatórias e inibitórias estimadas corretamente nos seis primeiros cenários. Na rede com 20 neurônios, calculamos a média de acertos para o tipo de conexão em 100 amostras. No que tange a detecção de conexões sinápticas, utilizamos um nível de precisão de 20%, i.e, se $\left| \hat{W}_{j \rightarrow i}^{(T)} \right| \geq 0,2$, tanto nos problemas de alta quanto nos de baixa dimensionalidade, então consideramos que os neurônios i e j estão conectados. Do ponto de vista da predição, a avaliação da performance é feita por meio de medidas baseadas na matriz de confusão.

Por meio da análise das tabelas, observamos que em todos cenários, inclusive no que contém 20 neurônios, os erros quadráticos médios empíricos, calculados a partir de 100 réplicas de Monte Carlo, tendem a diminuir e se aproximar de zero na medida que os tamanhos amostrais crescem, não havendo sequer um caso em que o erro não diminuiu conforme o tamanho amostral aumentou, o que é um indicativo de que os três métodos adotados neste trabalho são consistentes na estimação dos pesos sinápticos para o modelo neuronal do tipo Galves e Löcherbach que estamos considerando.

Uma análise comparativa dos três métodos nos microcircuitos, mostra que independente da proporção de neurônios excitatórios e inibitórios na rede e de sua esparsidade, os métodos encontram mais dificuldades no processo de estimação dos pesos sinápticos quando o tamanho da amostra é pequeno. Todavia, essa dificuldade é mitigada conforme aumenta-se o tamanho amostral. Com $T = 24000$ janelas temporais de 0,30 milissegundos, observamos que todos os métodos apresentam erros quadráticos médios empíricos similares, indicando que, na prática, onde o tamanho amostral, em geral, é bem maior, os métodos em consideração terão um bom desempenho na estimação dos pesos sinápticos.

No cenário 7, tanto o método proposto quanto o método da máxima verossimilhança com regularização ℓ_1 dos parâmetros demonstram um desempenho significativamente superior na convergência dos estimadores dos pesos sinápticos para os verdadeiros valores

dos parâmetros em comparação com o método de máxima verossimilhança sem seleção de variáveis. No entanto, apesar dos resultados menos precisos, a máxima verossimilhança sem regularização também reduz a distância entre as matrizes estimadas e a original com o aumento do tamanho da amostra.

A partir desse estudo de simulação, conjecturamos que os métodos de estimação adotados produzem estimativas consistentes dos parâmetros que representam os pesos sinápticos no modelo neuronal adotado. Em outras palavras, temos razões para acreditar que a seguinte afirmação seja verdadeira:

Afirmção 4.1 *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 de um modelo do tipo Galves e Löcherbach, isto é,*

$$P\left(X_t^{(i)} = 1 \mid \mathbf{X}_{-\infty}^{t-1} = \mathbf{x}_{-\infty}^{t-1}\right) = \phi\left((1 - x_{t-1}^{(i)}) \sum_{j \in I} \omega_{j \rightarrow i} \sum_{s=L_{t-1}^{(i)}+1}^{t-1} x_s^{(j)}\right).$$

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 $\hat{W}_{j \rightarrow i}^{(T)}$ um estimador do peso sináptico $\omega_{j \rightarrow i}$, obtido por algum dos métodos adotados, então

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

Com a análise dos gráficos do Apêndice A, há um indicativo de que a convergência para zero do erro quadrático médio empírico dos estimadores dos pesos sinápticos tende a decair exponencialmente. Do ponto de vista matemático, conjecturamos que seja possível também fixarmos resultados teóricos sobre o limite da probabilidade de concentração dos estimadores em torno dos verdadeiros valores dos parâmetros.

Os gráficos do B, no caso dos microcircuitos (cenário 1 a 6) exibem a proporção de vezes em que a conexão sináptica foi identificada corretamente pelos métodos 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 e desconexões acertadas apresentam 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 que o tamanho da amostra aumenta.

Além disso, a tabela da média de acerto de conexões no problema de alta dimensionalidade, também disponível no Apêndice B, apresenta valores com um comportamento semelhante ao dos microcircuitos testados. Apesar de estar em escalas semelhantes para os dois métodos regularizados e valores consideravelmente menores para a máxima verossimilhança sem seleção de variáveis, o aumento do tamanho de amostra diminui a distância média entre as matrizes estimadas e originais e, conseqüentemente, a média de conexões e desconexões identificadas corretamente aumenta. Portanto, temos um indicativo de que os métodos adotados neste trabalho também são consistentes na estimação do grafo de conectividade tanto em problemas de baixa dimensionalidade quanto em alta dimensionalidade.

Em particular, é possível observar que, em todos os cenários, o método de seleção estatística proposto possui uma performance superior aos demais métodos na identificação correta de desconexões e possui um desempenho similar na identificação de conexões. Além disso, é importante destacar que conforme aumenta-se a proporção de neurônios excitatórios na rede, mais dificuldade o método LASSO possui em estimar o grafo de conectividade; e quanto mais esparsa é a rede, mais dificuldade o método da máxima verossimilhança sem regularização dos parâmetros possui em identificar desconexões. Assim, concluímos que metodologia proposta é a mais robusta tanto em termos do desbalanceamento quanto da esparsidade da rede.

Assim, temos razões para acreditar que a seguinte afirmação é verdadeira:

Afirmação 4.2 *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,

$$P\left(X_t^{(i)} = 1 \mid \mathbf{X}_{-\infty}^{-1} = \mathbf{x}_{-\infty}^{-1}\right) = \phi\left((1 - x_{t-1}^{(i)}) \sum_{j \in I} \omega_{j \rightarrow i} \sum_{s=L_{t-1}^{(i)}+1}^{t-1} x_s^{(j)}\right).$$

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 $\hat{W}_{j \rightarrow i}^{(T)}$ o estimador do peso sináptico $\omega_{j \rightarrow i}$, obtido por algum dos métodos adotados, e

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

o estimador da vizinhança $\mathcal{V}_i$ 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

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

No que diz respeito o poder preditivo dos métodos propostos nos cenários com 5 neurônios, notamos, pela análise das tabelas do Apêndice C, que as três metodologias são robustas quanto ao aprendizado. Mesmo com uma proporção de treino relativamente baixa, equivalente a apenas 10% (3.000 observações), todos os métodos apresentam uma acurácia próxima ou acima de 70% em todos os cenários.

Em particular, observamos que o método proposto possui o maior coeficiente de correlação de Mathews (MCC) em todos os cenário de baixa dimensionalidade cujo treino é conduzido com 90% da rede. Destacamos o MCC, pois ele não é afetado pelo desbalanceamento das classes, uma vez que leva em conta todos os aspectos da matriz de confusão e, por conta disso, avalia o desempenho do classificador em ambas as classes. Assim, de acordo com o MCC, concluímos que o método proposto possui um desempenho ligeiramente superior aos outros métodos adotados.

Em relação ao cenário de alta dimensionalidade, ao contrário das redes com 5 neurônios,

o método proposto e a máxima verossimilhança com regularização ℓ_1 produzem resultados próximos entre si. No entanto, esses resultados são significativamente maiores do que os obtidos pela metodologia de máxima verossimilhança sem seleção de variáveis. Essa disparidade é atribuída à grande quantidade de parâmetros a serem estimados, o que impacta diretamente na performance das variância dos estimadores. De qualquer forma, tanto o LASSO quanto o método proposto indicam robustez em relação à previsão de valores na rede nessas circunstâncias. Mesmo com apenas 10% do conjunto de dados utilizado para o treino, as medidas de acurácia para ambos os métodos são de aproximadamente 95%. Além disso, enquanto o MCC para a verossimilhança sem seleção de variáveis atinge no máximo 41,97% dentro dos balanceamentos testados, os outros dois métodos resultam em coeficientes de Matthews que oscilam em torno de 91%, indicando uma concordância quase perfeita com os dados.

Assim, de maneira comparativa, evidenciamos empiricamente que o método proposto, ao não apresentar uma característica enviesada nas estimativas e, simultaneamente, realizar uma seleção de variáveis, produz valores ligeiramente mais precisos em relação aos demais abordados nesta pesquisa em problemas que envolvem a estimação de poucos parâmetros. Ao testarmos o método em um macrocircuito, sua performance, tanto na estimação da matriz sináptica e no grafo de conectividade quanto na previsão dos valores da rede, assemelha-se à do LASSO, que é uma técnica amplamente conhecida e utilizada em problemas de alta dimensionalidade. Considerando esses resultados, acreditamos que tanto o método proposto quanto a máxima verossimilhança com regularização ℓ_1 podem ser mais relevantes em situações práticas, especialmente em problemas reais em que o número de neurônios na rede tende a ser elevado como veremos no capítulo seguinte.

Capítulo 5

Aplicação em dados eletrofisiológicos

Nesta fase do trabalho, aplicamos os métodos de estimação estudados no Capítulo 4 em um conjunto de dados eletrofisiológicos obtidos de ratos e camundongos durante a realização de atividades livres. Os registros foram feitos em diversas áreas do cérebro como hipocampo, tálamo e córtex cerebral. Todavia, nesse trabalho, optamos por trabalhar apenas com as células do córtex. A justificativa para essa escolha é que durante o estudo de simulação descrito no capítulo anterior, as escolhas de alguns parâmetros foram feitas mediante considerações neurobiológicas envolvendo células dessa região cerebral. A base de dados, disponibilizados por Buzsaki Lab Petersen & Hernandez (2018), contém informações relativas aos tempos de disparo, em segundos, de 34 neurônios.

5.1 Pré-processamento dos dados

Como discutimos anteriormente, o banco de dados que utilizamos é composto 34 vetores, os quais registram os momentos em que ocorreram disparos para cada neurônio específico durante o processo de monitoramento do EEG, e não representam diretamente a atividade neuronal da rede. Portanto, antes de avançarmos para o treinamento do modelo, é essencial realizar um tratamento dessas informações, transformando as marcações de tempo em vetores que indicam se, em um instante t específico, houve ou não um disparo para um determinado neurônio i , de maneira semelhante ao que foi empregado nos estudos simulados. Essa transformação foi concretizada através da aplicação de uma função de discretização que, de forma simplificada, converte os dados obtidos no monitoramento em uma matriz binária (0 ou 1). Nas linhas dessa matriz, temos a representação das células da rede, sendo que as colunas representam os intervalos de tempo e a informação

correspondente sobre se houve ou não um disparo naquele instante.

A entrada da função de discretização, além da base de dados, requer o tamanho do bin, para o qual calculamos o tempo médio de disparo de todos os neurônios disponíveis. Dessa forma, inicialmente, o processo gera uma matriz composta exclusivamente por zeros. Em seguida, com base nas informações dos tempos de disparo e no tamanho do intervalo escolhido (bin), os instantes (ou colunas) nos quais cada neurônio das linhas disparou são substituídos por 1. Feito isso, ao final do processo obtivemos uma amostra $T = 7524$ para todos os 34 neurônios.

Após o tratamento das informações da base, procedemos à estimação e, posteriormente, ao processo de predição. Destacamos que para realizar a previsão utilizamos quatro diferentes proporções de treino e teste para avaliar a robustez de cada um dos métodos aplicados, sendo elas: (i) 40% para treino e 90% para teste, (ii) 60% para treino e 40% para teste, (iii) 75% para treino e 25% para teste, e (iv) 90% para treino e 10% para teste.

5.2 Estimação dos pesos sinápticos

Com as informações da base de dados devidamente processadas, o processo de estimação dos pesos sinápticos é conduzido de maneira análoga aos estudos realizados nos cenários simulados descritos no Capítulo 4.

O método de estimação por máxima verossimilhança sem regularização dos parâmetros, resultou em uma matriz que, com base na amostra observada, temos um indicativo de que neurônios 32 e 31 são os que possuem a conexão excitatória mais forte, pois as estimativas $\hat{W}_{32 \rightarrow 31} = 24.57628$ e $\hat{W}_{31 \rightarrow 32} = 31.70749$ são as maiores. Por outro lado, a conexão inibitória mais significativa ocorre entre os neurônios 28 e 32, em que $\hat{W}_{28 \rightarrow 32} = -35.8874$.

No caso do método de máxima verossimilhança com regularização ℓ_1 dos parâmetros, destacamos que a conexão excitatória mais expressiva ocorre entre os neurônios 20 e 12, pois $\hat{W}_{20 \rightarrow 12} = 7.0717$. Além disso, a conexão inibitória mais forte ocorre entre os neurônios 17 e 12, em que $\hat{W}_{17 \rightarrow 12} = -3.0083$.

No que diz respeito ao método proposto, observa-se que ele estimou valores mais próximos de zero. Nesse contexto, a conexão excitatória mais proeminente ocorre entre os neurônios 12 e 32, em que $\hat{W}_{12 \rightarrow 32} = 2.03142$. Além disso, a conexão inibitória mais forte ocorre entre os neurônios 20 e 29 em que $\hat{W}_{20 \rightarrow 29} = -1.03193$.

5.3 Estimação do grafo de conectividade

Para avaliar o grafo de conectividade estimado, utilizamos novamente a medida de precisão fixada em $\epsilon = 0,2$, com o objetivo de considerar também como desconectados os neurônios em que as estimativas são próximas de 0.

Tabela 5.1: Proporção dos tipos de conexões sinápticas detectadas por cada método.

Proporção identificada	
EMV + LASSO	
Excitatórios	0.1634
Inibitórios	0.0501
Desconectados	0.7863
EMV + Método proposto	
Excitatórios	0.1660
Inibitórios	0.0513
Desconectados	0.7827
EMV	
Excitatórios	0.2944
Inibitórios	0.3139
Desconectados	0.3917

A diferença mais notória ocorre na proporção de desconexões estimadas entre a metodologia sem regularização e as outras duas. Ambos os métodos que empregam seleção de variáveis estimam praticamente o dobro de desconexões em comparação com o método de máxima verossimilhança sem regularização dos parâmetros. No entanto, dado que a matriz de pesos sinápticos original é desconhecida nesse caso, qualquer análise comparativa nesse sentido não seria conclusiva em relação à qualidade das estimativas produzidas por cada método. Em todo caso, vale destacar que o método proposto tem uma performance parecida com o LASSO na estimação do grafo de conectividade.

5.4 Previsão da atividade da rede

Nesta etapa, vamos avaliar a qualidade dos ajustes por meio de uma tabela com a formatação semelhante às que analisamos no Capítulo 4. Portanto, os métodos estão divididos nos três blocos principais, enquanto as linhas da tabela representam a proporção de treino e teste utilizada, respectivamente. As colunas, por sua vez, contêm as medidas relativas à matriz de confusão.

Como esperado, de acordo com a Tabela 5.2, o modelo de máxima verossimilhança sem regularização dos parâmetros não atinge um ponto satisfatório de ajuste aos dados em nenhuma das proporções testadas. Apesar das métricas de Acurácia flutuarem em torno

Tabela 5.2: Métricas da matriz de confusão obtida na previsão do comportamento da rede no banco de dados reais com 34 neurônios monitorados por EEG.

	ACC	S	E	VPP	VPN	G-Média	MCC
EMV + LASSO							
40/60	0.7131	0.9526	0.3522	0.6890	0.8316	0.7569	0.4083
60/40	0.7279	0.9310	0.4352	0.7038	0.8140	0.7568	0.4354
75/25	0.7332	0.9382	0.3940	0.7191	0.7941	0.6284	0.4321
90/10	0.7397	0.9340	0.4605	0.7134	0.8291	0.7690	0.4625
EMV + Método proposto							
40/60	0.7132	0.9524	0.3527	0.6891	0.8310	0.7567	0.4077
60/40	0.7278	0.9301	0.4363	0.7040	0.8123	0.7562	0.4349
75/25	0.7239	0.9250	0.4323	0.7027	0.7990	0.7492	0.4233
90/10	0.7392	0.9299	0.4651	0.7142	0.8219	0.7661	0.4601
EMV							
40/60	0.6124	0.8678	0.2567	0.6375	0.5631	0.5991	0.0995
60/40	0.6022	0.8540	0.2393	0.6181	0.5322	0.5735	0.1184
75/25	0.6139	0.8603	0.2065	0.6419	0.4719	0.4732	0.1151
90/10	0.5917	0.8704	0.1910	0.6073	0.5063	0.5545	0.0835

de 60%, as demais medidas indicam um ajuste inadequado do modelo. Por exemplo, a especificidade apresenta seu valor máximo nas proporções estudadas para essa metodologia de 23,93%. Enquanto isso, tanto o valor preditivo positivo quanto o negativo situam-se em torno de 60% e 50%, respectivamente. Além disso, os coeficientes de Matthews, nesse caso, são muito próximos de zero, indicando que a classificação de disparos determinada pela modelagem se aproxima de uma escolha aleatória.

Por outro lado, a máxima verossimilhança com regularização ℓ_1 dos parâmetros mantém suas métricas de acurácia em torno de 73%, praticamente 13 pontos percentuais acima do que observamos para o EMV sem seleção de variáveis. Além disso, todas as outras medidas situam-se em valores maiores, destacando-se o MCC, que, em média, é de 43%, indicando uma concordância satisfatória do ajuste com os dados. Esse valor representa um aumento de mais de 30% em relação à metodologia sem regularização.

Contudo, os resultados do método proposto demonstram empiricamente sua eficácia e qualidade na seleção de variáveis e ajuste do modelo em uma situação real. Nesse contexto, com 34 neurônios na rede, as medidas da matriz de confusão para a metodologia se equiparam aos resultados obtidos com o método da máxima verossimilhança com regularização ℓ_1 dos parâmetros, um método de modelagem amplamente conhecido e utilizado em problemas de alta dimensionalidade. As métricas de acurácia ficam em torno de 73% em todas as proporções, assim como o coeficiente de Matthews, que atinge, no caso do balanceamento 90/10, 46% de concordância com os dados.

Além disso, tanto o método de máxima verossimilhança com regularização ℓ_1 quanto

o método proposto mostram-se técnicas robustas no problema com diversos parâmetros. Mesmo com uma baixa volumetria de dados no conjunto de treino, como no caso da proporção 40/60, as estimativas já apresentam resultados satisfatórios.

Capítulo 6

Considerações finais

De modo geral, apresentamos nesse trabalho, nos capítulos iniciais, todos os conceitos teóricos que constituem o alicerce para os nossos objetivos principais, ou seja, a predição de redes de neurônios estocásticos com memória de alcance variável, a estimação da matriz de pesos sinápticos e do grafo de conectividade. Com isso, no Capítulo 3, discutimos como a incorporação da memória de alcance variável pode melhorar a precisão da predição da atividade neuronal. Em seguida, encontramos a função de log-verossimilhança do nosso modelo e, a partir daí, propusemos três modos diferentes para a estimação baseados no método da máxima verossimilhança, sendo eles: (i) sem considerar regularização dos parâmetros, (ii) considerando regularização ℓ_1 dos coeficientes do modelo e (iii) utilizando o método de seleção que propomos, o qual é baseado na distância euclidiana entre o vetor das probabilidades preditas, do qual também discutimos os detalhes.

Dessa forma, para realizar a predição de redes de neurônios estocásticos com memória de alcance variável, desenvolvemos algoritmos computacionais para simular diversos cenários da rede. Esses cenários abrangeram tanto macro quanto microcircuitos, com diferentes tamanhos de T (tamanho de amostra), todos explorados e detalhados no Capítulo 4. Subsequentemente, analisamos os resultados da aplicação computacional, avaliando e comparando a performance dos modelos ajustados pelas metodologias de estimação discutidas. Ao final, constatamos a eficiência e robustez do método proposto em todas as condições de experimento testadas. Essa eficácia se manifesta tanto na estimação da matriz de pesos sinápticos e do grafo de conectividade quanto na previsão dos valores da rede.

No Capítulo 5, aplicamos os métodos de estimação discutidos a uma base de dados reais proveniente de registros por EEG de diversas áreas do cérebro, como hipocampo, tálamo e córtex cerebral. Essa base de dados é publicamente disponibilizada pelo Buzsáki

Lab ([Petersen *et al.*, 2020](#)). Com isso, reforçamos os indícios de robustez e qualidade dos métodos nesse contexto, especialmente o método proposto, que apresenta um desempenho muito semelhante ao LASSO, uma técnica amplamente conhecida e utilizada em situações reais de alta dimensionalidade.

Nos próximos passos, almejamos estabelecer teoricamente a consistência dos métodos na geração de estimativas condizentes ao contexto, complementando os resultados empíricos já obtidos. Além disso, como uma perspectiva futura, planejamos explorar a viabilidade de desenvolver um estimador Bayesiano para a matriz de pesos sinápticos da rede. Pretendemos também avaliar sua aplicação em cenários diversos e a precisão dos possíveis valores preditos para a atividade neuronal da rede.

Referências Bibliográficas

- Baccelli, F. e Taillefumier, T. (2019). Replica-mean-field limits for intensity-based neural networks. *SIAM Journal on Applied Dynamical Systems*, **18**(4), 1756–1797.
- Baccelli, F., Davydov, M. e Taillefumier, T. (2022). Replica-mean-field limits of fragmentation-interaction-aggregation processes. *Journal of Applied Probability*, **59**(1), 38–59.
- 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.
- Bejerano, G. e Yona, G. (2001). Variations on probabilistic suffix trees: statistical modeling and prediction of protein families. *Bioinformatics*, **17**(1), 23–43.
- 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), 1–22.
- 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.
- 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.
- De Masi, A., Galves, A., Löcherbach, E. e Presutti, E. (2015). Hydrodynamic limit for interacting neurons. *Journal of Statistical Physics*, **158**, 866–902.
- Duarte, A. e Ost, G. (2014). A model for neural activity in the absence of external stimuli. *arXiv preprint arXiv:1410.6086*.

- Duarte, A., Ost, G. e Rodríguez, A. A. (2015). Hydrodynamic limit for spatially structured interacting neurons. *Journal of Statistical Physics*, **161**, 1163–1202.
- Fournier, N. e Löcherbach, E. (2016). On a toy model of interacting neurons.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, **11**(2), 127–138.
- 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.
- Galves, A., Galves, C., Garcia, N. e Peixoto, C. (2004). Correlates of rhythm in written texts of brazilian and european portuguese. *Preprint. Available as Technical Report*, **8**(08).
- Gerstner, W. e Brette, R. (2009). Adaptive exponential integrate-and-fire model. *Scholarpedia*, **4**(6), 8427.
- 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.
- Hodara, P. e Löcherbach, E. (2017). Hawkes processes with variable length memory and an infinite number of components. *Advances in Applied Probability*, **49**(1), 84–107.
- Hodgkin, A. L. e Huxley, A. F. (1952). The components of membrane conductance in the giant axon of loligo. *The Journal of Physiology*, **116**(4), 473.
- 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.

- London, M., Shcreibman, A. e Segev, I. (2002). Estimating information theoretic quantities of spike-trains using the context tree weighting algorithm. *Nature Neuroscience*, **5**.
- Mood, A. M., Graybill, F. A. e Boes, D. C. (1974). *Introduction to the Theory of Statistics 1974*. McGraw-Hill Kogakusha.
- 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.
- Petersen, P. C., Hernandez, M. e Buzsáki, G. (2020). The buzsaki lab databank – public electrophysiological datasets from awake animals. <http://doi.org/10.5281/zenodo.4307883>.
- Rissanen, J. (1983). A universal data compression system. *IEEE Transactions on information theory*, **29**(5), 656–664.
- Robert, P. e Touboul, J. (2016). On the dynamics of random neuronal networks. *Journal of Statistical Physics*, **165**, 545–584.
- Sporns, O., Tononi, G. e Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Computational Biology*, **1**(4), e42.
- 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.
- y Cajal, S. R. (1888). *Estructura de los centros nerviosos de las aves*.
- Yaginuma, K. (2016). A stochastic system with infinite interacting components to model the time evolution of the membrane potentials of a population of neurons. *Journal of Statistical Physics*, **163**(3), 642–658.

Apêndice A

Tabelas e gráficos com os resultados da estimação da matriz de pesos sinápticos

A.1 Cenário 1

Tabela A.1: Erro quadrático médio empírico calculado para o cenário 1 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T=100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	3.6701	0.1546	0.0220	0.0111
W3→1	-1	2.7237	0.1942	0.0199	0.0080
W4→1	-1	2.5916	0.1610	0.0145	0.0074
W5→1	-1	3.8121	0.02233	0.0125	0.0087
W1→2	0	4.1121	0.1762	0.0164	0.0081
W3→2	-1	2.6885	0.1992	0.0153	0.0093
W4→2	-1	3.1697	0.1736	0.0273	0.0143
W5→2	-1	3.6270	0.1610	0.0170	0.0089
W1→3	-1	1736.715	0.02844	0.0341	0.0178
W2→3	-1	1543.117	0.02308	0.0324	0.0160
W4→3	4	2114.304	0.3531	0.0344	0.0162
W5→3	4	1982.059	0.3083	0.0335	0.0150
W1→4	-1	3.5616	0.02257	0.0152	0.0076
W2→4	-1	3.2840	0.02281	0.0214	0.0096
W3→4	4	3.6718	0.02790	0.0175	0.0093
W5→4	-1	2.3489	0.1628	0.0184	0.0074
W1→5	-1	5.3108	0.02358	0.0191	0.0096
W2→5	-1	8.1915	0.02446	0.0198	0.0089
W3→5	4	10.9789	0.02316	0.0206	0.0113
W4→5	-1	3.2957	0.02277	0.0159	0.0070

Tabela A.2: Erro quadrático médio empírico calculado para o cenário 1 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T= 100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	0.9312	0.0822	0.0078	0.0026
W3→1	-1	0.7416	0.1811	0.0197	0.0067
W4→1	-1	0.9035	0.1557	0.0148	0.0091
W5→1	-1	1.1661	0.1960	0.0127	0.0098
W1→2	0	1.2825	0.1064	0.0059	0.0016
W3→2	-1	0.9753	0.1777	0.0155	0.0085
W4→2	-1	0.9365	0.1587	0.0277	0.0135
W5→2	-1	1.0402	0.1567	0.0159	0.0105
W1→3	-1	21.3503	0.3000	0.0392	0.0238
W2→3	-1	5.8243	0.02501	0.0299	0.0194
W4→3	4	10.7602	0.3782	0.0436	0.0236
W5→3	4	8.7889	0.3755	0.0410	0.0242
W1→4	-1	1.4885	0.02283	0.0170	0.0103
W2→4	-1	1.4896	0.02204	0.0220	0.0115
W3→4	4	4.7329	0.02947	0.0304	0.0219
W5→4	-1	0.8632	0.1398	0.0191	0.0091
W1→5	-1	1.9326	0.02210	0.0211	0.0109
W2→5	-1	2.6802	0.02175	0.0194	0.0116
W3→5	4	7.0263	0.02400	0.0270	0.0212
W4→5	-1	1.3271	0.02229	0.0173	0.0095

Tabela A.3: Erro quadrático médio empírico calculado para o cenário 1 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T= 100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	2.2516	0.0863	0.0156	0.0081
W3→1	-1	2.2007	0.3123	0.0197	0.0075
W4→1	-1	1.9519	0.02503	0.0145	0.0073
W5→1	-1	2.6430	0.3045	0.0124	0.0088
W1→2	0	2.5155	0.1149	0.0098	0.0051
W3→2	-1	1.9436	0.3233	0.0151	0.0088
W4→2	-1	2.6842	0.3024	0.0276	0.0141
W5→2	-1	2.8535	0.0286	0.0165	0.0092
W1→3	-1	1545.532	0.5255	0.0341	0.0178
W2→3	-1	817.280	0.4122	0.0324	0.0160
W4→3	4	1967.855	0.3842	0.0344	0.0162
W5→3	4	1847.689	0.3265	0.0335	0.0150
W1→4	-1	3.1563	0.3160	0.0152	0.0076
W2→4	-1	2.2207	0.3016	0.0214	0.0096
W3→4	4	3.2820	0.3008	0.0175	0.0093
W5→4	-1	2.1796	0.02290	0.0184	0.0074
W1→5	-1	4.0780	0.3615	0.0191	0.0096
W2→5	-1	7.6836	0.3096	0.0198	0.0089
W3→5	4	10.02207	0.02384	0.0206	0.0113
W4→5	-1	2.8822	0.3183	0.0159	0.0070

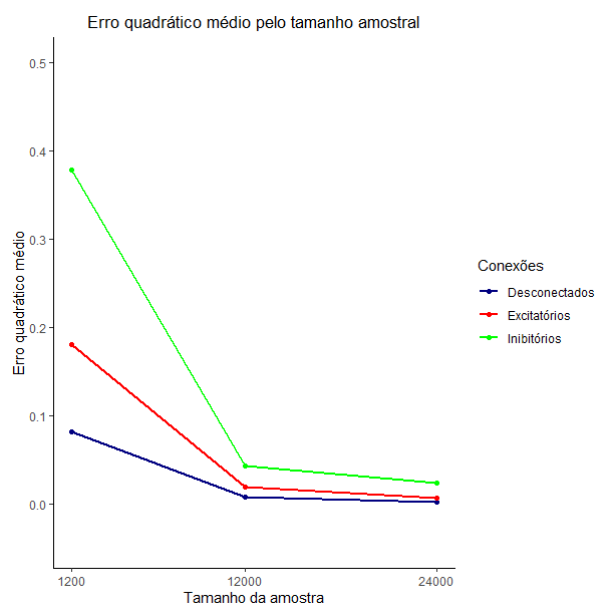


Figura A.1: Gráfico dos valores do erro quadrático médio empírico no cenário 1 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.

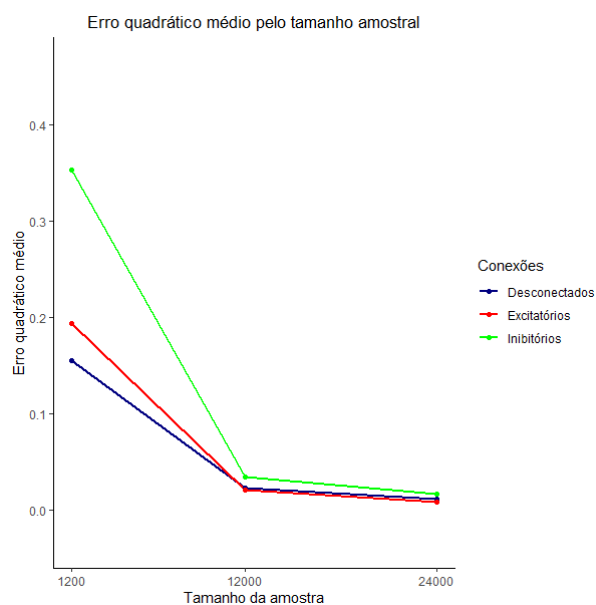


Figura A.2: Gráfico dos valores do erro quadrático médio empírico no cenário 1 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.

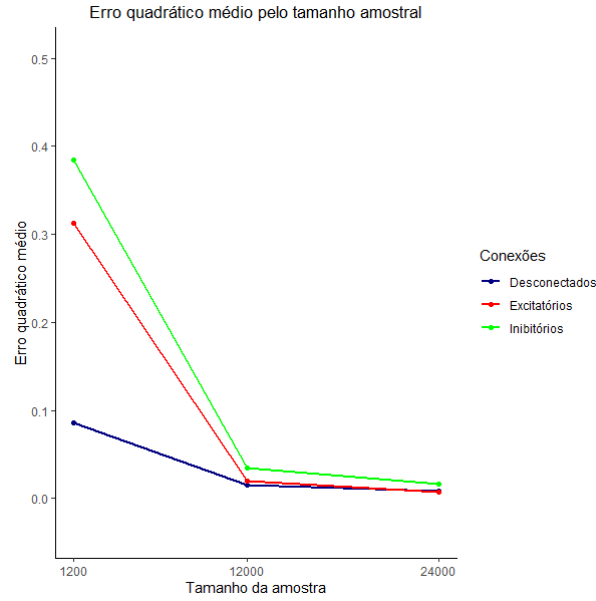


Figura A.3: Gráfico dos valores do erro quadrático médio empírico no cenário 1 em relação ao tamanho da amostra para o método de estimação proposto.

A.2 Cenário 2

Tabela A.4: Erro quadrático médio empírico calculado para o cenário 2 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T=100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	9.7026	0.1102	0.0109	0.0065
W3→1	4	23.6294	0.02493	0.0305	0.0236
W4→1	-1	2.8524	0.1767	0.0200	0.0135
W5→1	-1	1.9000	0.02339	0.0175	0.0133
W1→2	0	1.2754	0.1198	0.0075	0.0026
W3→2	-1	1.6006	0.02165	0.0202	0.0090
W4→2	-1	1.2644	0.1682	0.0284	0.0125
W5→2	-1	1.0766	0.1531	0.0223	0.0107
W1→3	4	89.0400	0.5418	0.0724	0.0511
W2→3	-1	17.7619	0.8404	0.1010	0.0680
W4→3	4	112.9805	0.6336	0.0860	0.0442
W5→3	4	92.2510	0.8656	0.0684	0.0443
W1→4	-1	3.1755	0.3424	0.0368	0.0217
W2→4	-1	3.3831	0.4061	0.0382	0.0309
W3→4	4	7.6155	0.4632	0.0393	0.0288
W5→4	4	5.7123	0.4561	0.0513	0.0277
W1→5	-1	25.0958	0.02961	0.0366	0.0242
W2→5	-1	43.7468	0.3837	0.0371	0.0194
W3→5	4	29.4169	0.3976	0.0431	0.0277
W4→5	4	25.8728	0.4926	0.0456	0.0290

Tabela A.5: Erro quadrático médio empírico calculado para o cenário 2 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	T= 100	T=1200	T=12000	T=24000
W2→1	0	0.7765	0.0822	0.0220	0.0137
W3→1	4	0.6903	0.1811	0.0199	0.0120
W4→1	-1	0.7490	0.1557	0.0145	0.0117
W5→1	-1	0.9305	0.1960	0.0125	0.0099
W1→2	0	0.7005	0.1064	0.0164	0.0083
W3→2	-1	1.0749	0.1777	0.0153	0.0094
W4→2	-1	1.4025	0.1587	0.0273	0.0125
W5→2	-1	0.9856	0.1567	0.0170	0.0103
W1→3	4	2.5198	0.3000	0.0341	0.0440
W2→3	-1	2.8165	0.02501	0.0324	0.0333
W4→3	4	4.2311	0.3782	0.0344	0.0314
W5→3	4	54.1112	0.3755	0.0335	0.0310
W1→4	-1	1.2547	0.02283	0.0152	0.0115
W2→4	-1	1.0876	0.02204	0.0214	0.0241
W3→4	4	3.7857	0.02947	0.0175	0.0161
W5→4	4	1.0884	0.1398	0.0184	0.0156
W1→5	-1	1.8345	0.02210	0.0191	0.0177
W2→5	-1	1.7491	0.02175	0.0198	0.0127
W3→5	4	4.3346	0.02400	0.0206	0.0149
W4→5	4	1.5781	0.02229	0.0159	0.0199

Tabela A.6: Erro quadrático médio empírico calculado para o cenário 2 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	T= 100	T=1200	T=12000	T=24000
W2→1	0	145.8163	0.0740	0.0137	0.0105
W3→1	4	254.6410	0.02152	0.0186	0.0120
W4→1	-1	15.7535	0.02603	0.0173	0.0117
W5→1	-1	69.6214	0.3198	0.0160	0.0097
W1→2	0	2.5560	0.1476	0.0105	0.0041
W3→2	-1	3.5206	0.4195	0.0201	0.0088
W4→2	-1	2.9751	0.3123	0.0284	0.0122
W5→2	-1	3.0849	0.02315	0.0231	0.0101
W1→3	4	4251.933	0.6092	0.0722	0.0440
W2→3	-1	2011.696	0.9305	0.0741	0.0333
W4→3	4	3062.532	0.6967	0.0747	0.0314
W5→3	4	6825.733	0.8329	0.0525	0.0310
W1→4	-1	1959.9841	0.5354	0.0296	0.0115
W2→4	-1	327.6859	0.6343	0.0345	0.0241
W3→4	4	5750.0652	0.4087	0.0294	0.0161
W5→4	4	2512.6241	0.4378	0.0443	0.0156
W1→5	-1	3125.712	0.4883	0.0294	0.0177
W2→5	-1	3671.282	0.6346	0.0309	0.0127
W3→5	4	2778.160	0.3753	0.0316	0.0149
W4→5	4	8131.369	0.4264	0.0326	0.0199

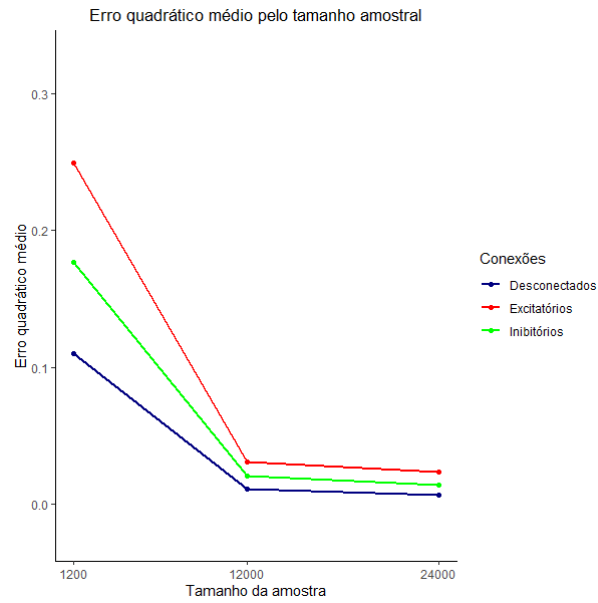


Figura A.4: Gráfico dos valores do erro quadrático médio empírico no cenário 2 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.

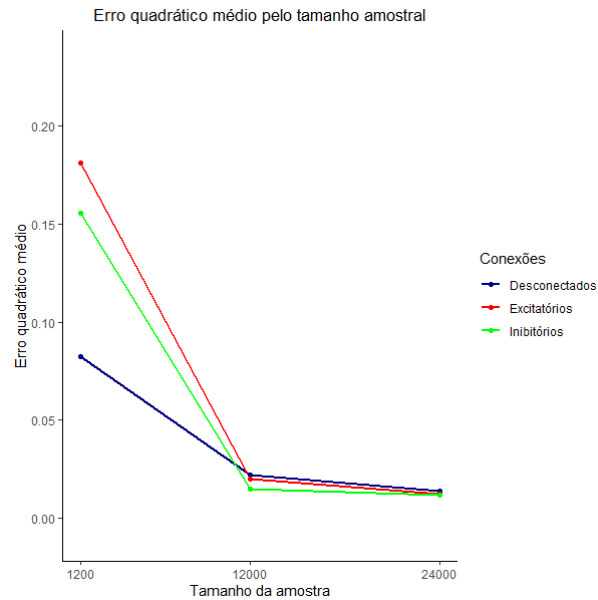


Figura A.5: Gráfico dos valores do erro quadrático médio empírico no cenário 2 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.

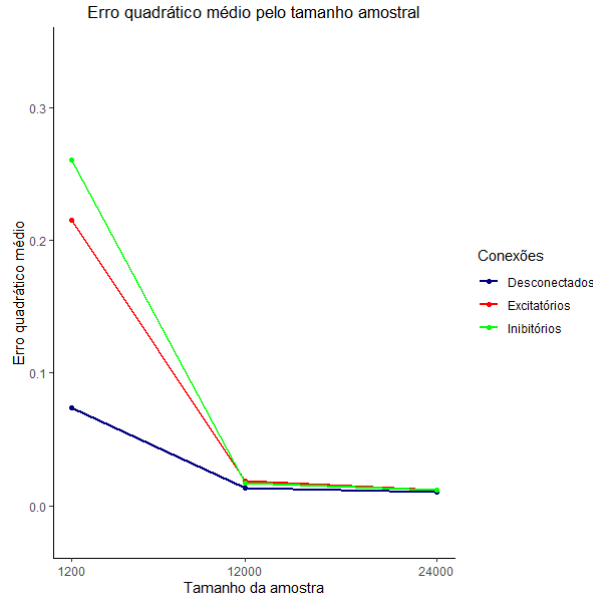


Figura A.6: Gráfico dos valores do erro quadrático médio empírico no cenário 2 em relação ao tamanho da amostra para o método de estimação proposto.

A.3 Cenário 3

Tabela A.7: Erro quadrático médio empírico calculado para o cenário 3 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T= 100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	69.4421	0.4155	0.0558	0.0316
W3→1	4	114.4161	1.0134	0.0863	0.0572
W4→1	4	130.4472	0.6421	0.0880	0.0419
W5→1	4	118.4643	0.7196	0.0996	0.0561
W1→2	0	2.5003	0.1465	0.0107	0.0034
W3→2	-1	1.4981	0.0247	0.0216	0.0100
W4→2	-1	1.8632	0.1904	0.0266	0.0137
W5→2	-1	2.1349	0.1865	0.0240	0.0117
W1→3	4	91.1726	0.6698	0.0721	0.0495
W2→3	-1	31.0342	1.0150	0.1077	0.0739
W4→3	4	125.8384	0.7896	0.0873	0.0480
W5→3	4	121.6533	1.1288	0.0824	0.0465
W1→4	4	111.1674	0.7113	0.0924	0.0471
W2→4	-1	24.1519	0.9865	0.1351	0.0986
W3→4	4	74.8083	0.7763	0.0906	0.0518
W5→4	4	82.8123	0.7210	0.1001	0.0481
W1→5	4	73.4467	0.6225	0.0745	0.0369
W2→5	-1	21.9856	0.7909	0.1208	0.0663
W3→5	4	83.7794	0.7427	0.0944	0.0445
W4→5	4	74.2888	0.7958	0.0907	0.0504

Tabela A.8: Erro quadrático médio empírico calculado para o cenário 3 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	T= 100	T=1200	T=12000	T=24000
W2→1	0	3173.385	0.8676	0.0955	0.0483
W3→1	4	4535.935	1.5054	0.0734	0.0454
W4→1	4	3921.198	0.8097	0.0813	0.0381
W5→1	4	3354.237	0.9336	0.0806	0.0420
W1→2	0	9.8652	0.0223	0.0218	0.0095
W3→2	-1	5.4845	0.0275	0.0215	0.0094
W4→2	-1	6.0551	0.0293	0.0261	0.0143
W5→2	-1	7.6035	0.0272	0.0254	0.0113
W1→3	4	4877.564	0.7240	0.0670	0.0429
W2→3	-1	2657.557	0.9554	0.0756	0.0387
W4→3	4	4103.504	0.8530	0.0853	0.0357
W5→3	4	2526.497	1.1805	0.0671	0.0325
W1→4	4	5472.337	0.7954	0.0913	0.0449
W2→4	-1	1580.000	0.9753	0.0843	0.0547
W3→4	4	2906.708	0.6948	0.0850	0.0388
W5→4	4	2375.311	0.8531	0.0809	0.0342
W1→5	4	4691.977	0.6724	0.0614	0.0306
W2→5	-1	2605.110	0.6659	0.0755	0.0303
W3→5	4	8408.114	0.7486	0.0865	0.0306
W4→5	4	5054.726	0.8382	0.0780	0.0403

Tabela A.9: Erro quadrático médio empírico calculado para o cenário 3 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	T= 100	T=1200	T=12000	T=24000
W2→1	0	2462.881	0.0336	0.0237	0.0236
W3→1	4	3623.131	1.4370	0.0715	0.0453
W4→1	4	3341.537	0.7504	0.0781	0.0370
W5→1	4	2651.064	0.8526	0.0740	0.0406
W1→2	0	8.0208	0.1564	0.0111	0.0049
W3→2	-1	3.8276	0.4658	0.0205	0.0088
W4→2	-1	4.3129	0.3719	0.0259	0.0139
W5→2	-1	6.5028	0.3579	0.0245	0.0110
W1→3	4	4318.107	0.7608	0.0670	0.0429
W2→3	-1	2244.249	1.1735	0.0947	0.0387
W4→3	4	3804.152	0.8398	0.0877	0.0357
W5→3	4	2182.737	1.1829	0.0687	0.0325
W1→4	4	4422.9009	0.8121	0.0913	0.0449
W2→4	-1	770.8124	1.0506	0.0843	0.0547
W3→4	4	2922.6710	0.7432	0.0850	0.0388
W5→4	4	2028.9649	0.8132	0.0809	0.0342
W1→5	4	3682.363	0.6501	0.0620	0.0306
W2→5	-1	2173.432	0.9720	0.0997	0.0303
W3→5	4	7555.557	0.7848	0.0896	0.0306
W4→5	4	3650.072	0.8412	0.0787	0.0403

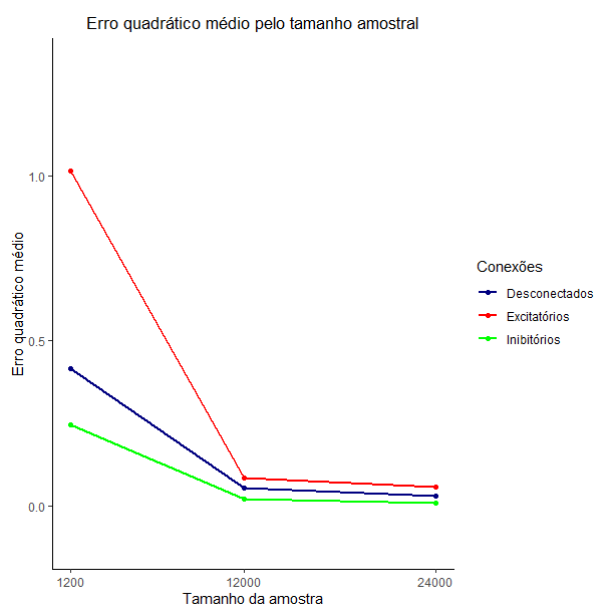


Figura A.7: Gráfico dos valores do erro quadrático médio empírico no cenário 3 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.

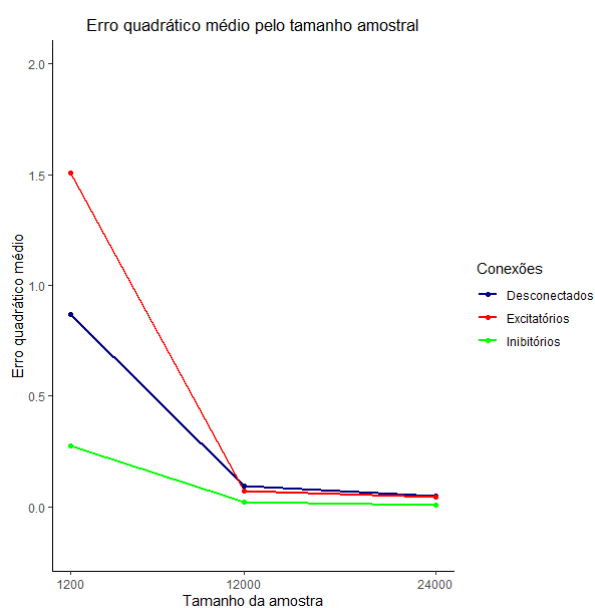


Figura A.8: Gráfico dos valores do erro quadrático médio empírico no cenário 3 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.

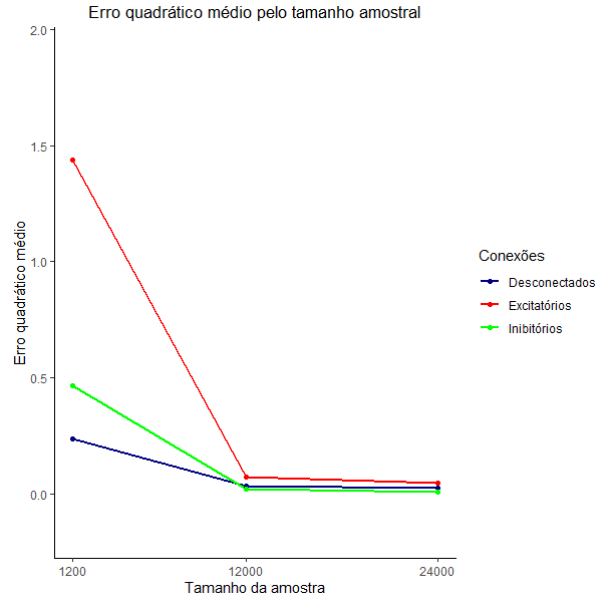


Figura A.9: Gráfico dos valores do erro quadrático médio empírico no cenário 3 em relação ao tamanho da amostra para o método de estimação proposto.

A.4 Cenário 4

Tabela A.10: Erro quadrático médio empírico calculado para o cenário 4 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	T= 100	T=1200	T=12000	T=24000
W2→1	0	19.6191	0.4050	0.0532	0.0316
W3→1	4	36.2392	1.0099	0.0947	0.0572
W4→1	4	33.8552	0.7693	0.1106	0.0419
W5→1	4	28.3994	0.8494	0.1048	0.0561
W1→2	0	0.5173	0.1588	0.0087	0.0034
W3→2	4	5.0397	0.02169	0.0331	0.0100
W4→2	-1	1.2514	0.02453	0.0202	0.0137
W5→2	-1	1.5055	0.02265	0.0234	0.0117
W1→3	4	79.3552	1.0053	0.1489	0.0495
W2→3	4	46.5661	2.4134	0.02158	0.0739
W4→3	4	104.3181	1.7863	0.1311	0.0480
W5→3	4	98.9977	2.4885	0.1535	0.0465
W1→4	4	84.7478	0.7939	0.1069	0.0471
W2→4	-1	26.8750	0.7818	0.1323	0.0986
W3→4	4	82.8876	0.8224	0.0833	0.0518
W5→4	4	76.9335	0.7757	0.1104	0.0481
W1→5	4	97.4153	0.6677	0.0834	0.0369
W2→5	-1	25.1856	0.9179	0.1034	0.0663
W3→5	4	135.9016	0.7154	0.0840	0.0445
W4→5	4	90.3131	0.9430	0.1013	0.0504

Tabela A.11: Erro quadrático médio empírico calculado para o cenário 4 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T= 100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	1017.627	0.8112	0.0796	0.0415
W3→1	4	4057.952	1.5331	0.0796	0.0499
W4→1	4	2566.125	0.9817	0.1003	0.0449
W5→1	4	4469.885	1.0620	0.0852	0.0419
W1→2	0	3.2477	0.02210	0.0185	0.0111
W3→2	4	5.9610	0.02069	0.0221	0.0086
W4→2	-1	4.0445	0.02847	0.0212	0.0120
W5→2	-1	5.3281	0.02362	0.0208	0.0084
W1→3	4	2207.664	10.657	0.1695	0.0803
W2→3	4	2014.377	164.50	0.02207	0.1106
W4→3	4	2203.589	83.161	0.1488	0.0814
W5→3	4	2253.408	136.01	0.1633	0.0644
W1→4	4	2655.825	0.8445	0.0902	0.0410
W2→4	-1	1728.898	0.6534	0.0704	0.0423
W3→4	4	4648.485	0.6808	0.0736	0.0398
W5→4	4	2697.288	0.8213	0.0809	0.0338
W1→5	4	3462.5343	0.6775	0.0672	0.0274
W2→5	-1	835.8945	0.7442	0.0681	0.0256
W3→5	4	3678.2526	0.7207	0.0722	0.0284
W4→5	4	3283.9610	0.8814	0.0854	0.0420

Tabela A.12: Erro quadrático médio empírico calculado para o cenário 4 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T= 100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	574.5436	0.4106	0.0437	0.0238
W3→1	4	3478.8476	1.4541	0.0754	0.0511
W4→1	4	2104.3649	0.8324	0.0995	0.0393
W5→1	4	3857.9215	0.9903	0.0777	0.0385
W1→2	0	1.2272	0.1779	0.0102	0.0076
W3→2	4	5.6265	0.02052	0.0215	0.0091
W4→2	-1	3.3585	0.3814	0.0196	0.0111
W5→2	-1	3.9951	0.3815	0.0206	0.0077
W1→3	4	1327.0292	10.7394	0.1695	0.0803
W2→3	4	678.6436	165.5148	0.02207	0.1106
W4→3	4	1363.1691	84.6336	0.1488	0.0814
W5→3	4	1569.6173	136.2342	0.1633	0.0644
W1→4	4	1744.194	0.8587	0.0955	0.0410
W2→4	-1	1128.693	0.9219	0.0881	0.0423
W3→4	4	3624.747	0.7087	0.0749	0.0398
W5→4	4	2286.756	0.8442	0.0833	0.0338
W1→5	4	2189.2051	0.6986	0.0682	0.0274
W2→5	-1	509.6191	0.9073	0.0719	0.0256
W3→5	4	2882.3238	0.7331	0.0731	0.0284
W4→5	4	2477.7593	0.8796	0.0842	0.0420

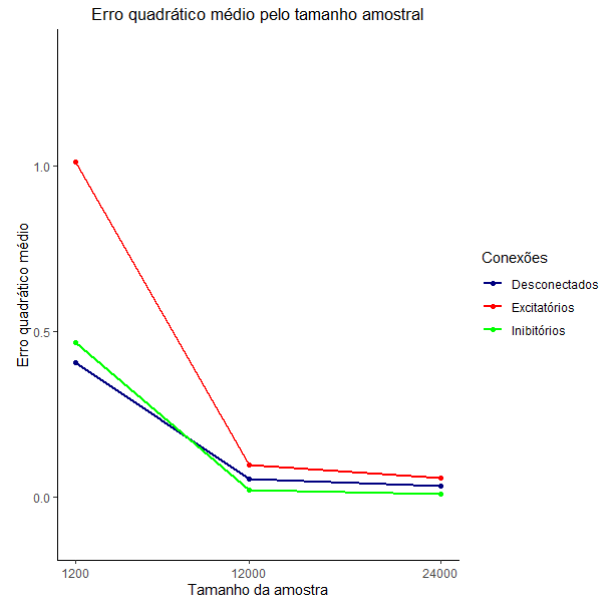


Figura A.10: Gráfico dos valores do erro quadrático médio empírico no cenário 4 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.

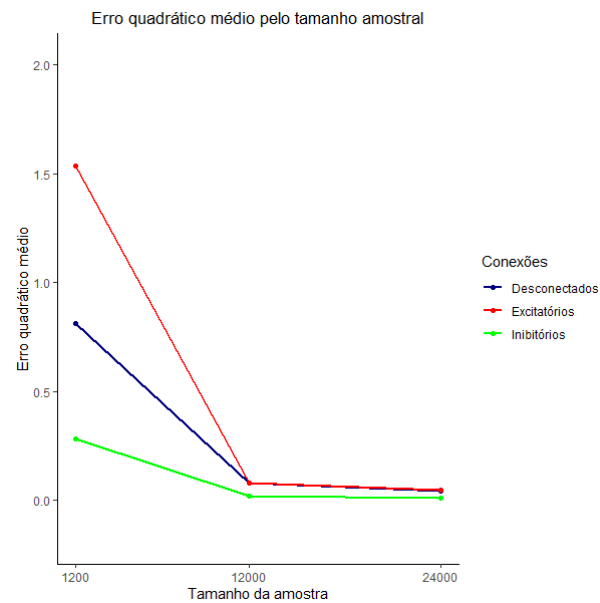


Figura A.11: Gráfico dos valores do erro quadrático médio empírico no cenário 4 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.

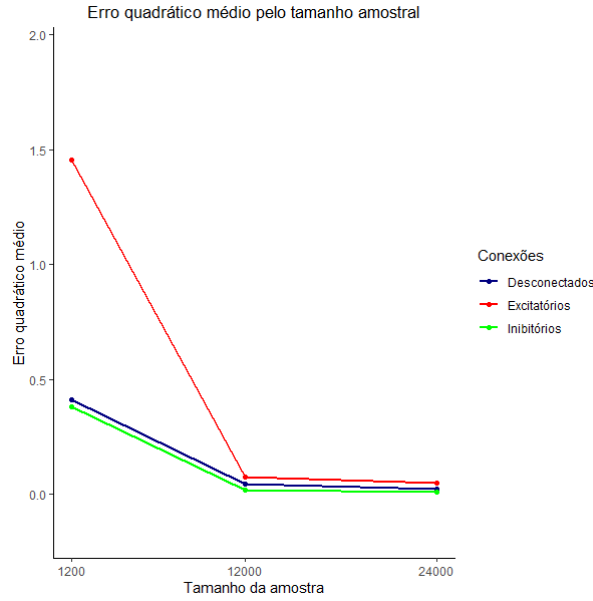


Figura A.12: Gráfico dos valores do erro quadrático médio empírico no cenário 4 em relação ao tamanho da amostra para o método de estimação proposto.

A.5 Cenário 5

Tabela A.13: Erro quadrático médio empírico calculado para o cenário 5 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	T= 100	T=1200	T=12000	T=24000
W2→1	0	0.0238	0.0316	0.0285	0.0021
W3→1	0	0.5254	0.0572	0.0371	0.0017
W4→1	0	0.3520	0.0419	0.0267	0.0010
W5→1	0	0.5143	0.0561	0.0444	0.0016
W1→2	0	1.4455	0.1202	0.0095	0.0034
W3→2	0	1.4844	0.1329	0.0100	0.0072
W4→2	4	4.3257	0.3056	0.0332	0.0137
W5→2	-1	2.0083	0.3051	0.0247	0.0117
W1→3	0	52.6717	0.02023	0.0495	0.0161
W2→3	0	17.7742	0.1923	0.0739	0.0089
W4→3	4	118.0492	0.4283	0.0494	0.0480
W5→3	4	92.0437	0.5933	0.0638	0.0465
W1→4	0	51.4558	0.4391	0.0471	0.0454
W2→4	4	61.1021	0.7376	0.0986	0.0801
W3→4	4	23.1272	0.5684	0.0774	0.0518
W5→4	4	106.1549	0.6104	0.0892	0.0481
W1→5	0	4.3541	0.02066	0.0369	0.0182
W2→5	-1	2.1968	0.4661	0.0663	0.0416
W3→5	4	26.5907	0.4470	0.0504	0.0445
W4→5	4	12.0703	0.5628	0.0510	0.0504

Tabela A.14: Erro quadrático médio empírico calculado para o cenário 5 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T= 100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	2.4415	0.1152	0.0149	0.0097
W3→1	0	2.9034	0.1676	0.0168	0.0079
W4→1	0	2.7635	0.1271	0.0130	0.0063
W5→1	0	3.0429	0.1758	0.0162	0.0109
W1→2	0	7.7017	0.1991	0.0190	0.0080
W3→2	0	7.8641	0.0219	0.0174	0.0086
W4→2	4	11.0834	0.0251	0.0231	0.0124
W5→2	-1	10.5326	0.0277	0.0186	0.0090
W1→3	0	1509.9589	0.3921	0.0344	0.0234
W2→3	0	650.3596	0.3598	0.0222	0.0165
W4→3	4	2689.5174	0.4603	0.0396	0.0165
W5→3	4	2227.7432	0.5122	0.0486	0.0195
W1→4	0	5999.072	0.7119	0.0695	0.0311
W2→4	4	6421.033	0.7976	0.0886	0.0396
W3→4	4	2682.781	0.7227	0.0698	0.0399
W5→4	4	7182.662	0.9032	0.0806	0.0349
W1→5	0	1240.1033	0.3247	0.0341	0.0186
W2→5	-1	659.1444	0.3387	0.0305	0.0127
W3→5	4	1738.7614	0.3794	0.0391	0.0169
W4→5	4	1605.0232	0.4333	0.0400	0.0208

Tabela A.15: Erro quadrático médio empírico calculado para o cenário 5 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T= 100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	1.0836	0.0814	0.0238	0.0107
W3→1	0	1.4666	0.0932	0.0511	0.0100
W4→1	0	1.6444	0.0659	0.0393	0.0076
W5→1	0	1.7722	0.1189	0.0385	0.0097
W1→2	0	6.2314	0.1193	0.0101	0.0076
W3→2	0	5.1188	0.1378	0.0091	0.0082
W4→2	4	9.4342	0.0210	0.0192	0.0111
W5→2	-1	8.0158	0.3722	0.0208	0.0077
W1→3	0	1318.9727	0.1744	0.0803	0.0121
W2→3	0	512.2219	0.1357	0.1106	0.0061
W4→3	4	2313.4039	0.3784	0.0814	0.0374
W5→3	4	2110.02188	0.5154	0.0644	0.0441
W1→4	0	5216.043	0.3491	0.0410	0.0390
W2→4	4	5810.093	0.7371	0.0922	0.0423
W3→4	4	2137.837	0.6925	0.0650	0.0398
W5→4	4	6445.341	0.8528	0.0754	0.0338
W1→5	0	1151.8437	0.1687	0.0274	0.0151
W2→5	-1	580.9882	0.5916	0.0300	0.0256
W3→5	4	1130.1963	0.3969	0.0386	0.0284
W4→5	4	1474.3901	0.4685	0.0420	0.0391

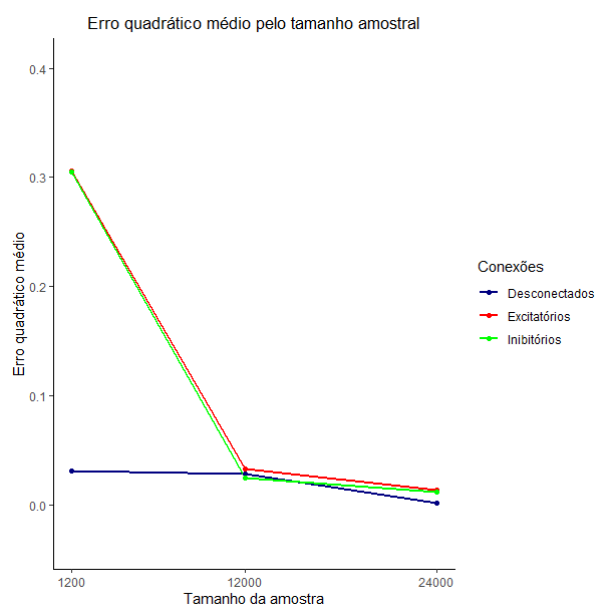


Figura A.13: Gráfico dos valores do erro quadrático médio empírico no cenário 5 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.

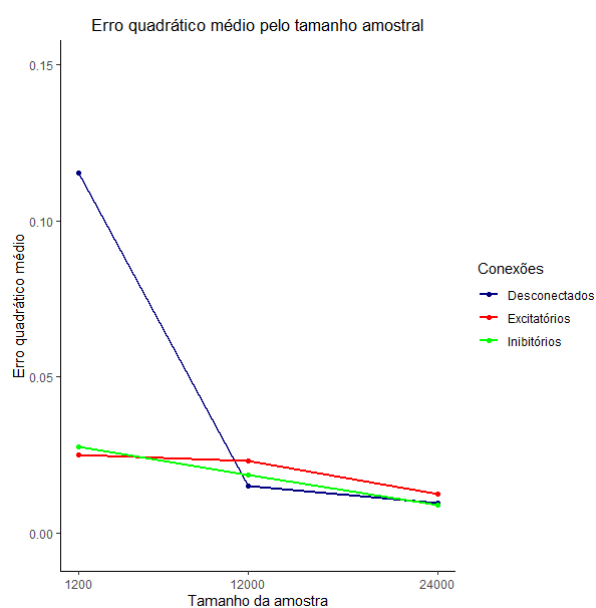


Figura A.14: Gráfico dos valores do erro quadrático médio empírico no cenário 5 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.

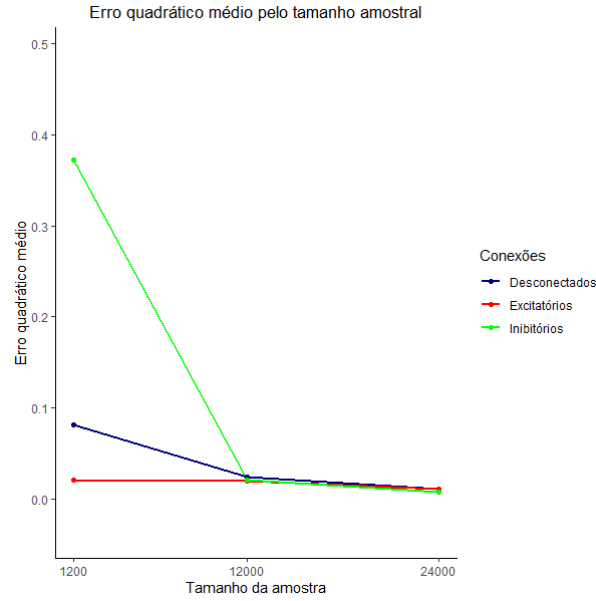


Figura A.15: Gráfico dos valores do erro quadrático médio empírico no cenário 5 em relação ao tamanho da amostra para o método de estimação proposto.

A.6 Cenário 6

Tabela A.16: Erro quadrático médio empírico calculado para o cenário 6 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança com LASSO. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T= 100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	0.2389	0.0285	0.0015	0.0013
W3→1	0	0.5254	0.0371	0.0036	0.0032
W4→1	0	0.3520	0.0267	0.0013	0.0010
W5→1	0	0.5143	0.0444	0.0013	0.0009
W1→2	0	1.4455	0.1202	0.0010	0.0008
W3→2	0	1.4844	0.1329	0.0020	0.0016
W4→2	0	4.3257	0.3056	0.0019	0.0015
W5→2	0	2.0083	0.3051	0.0013	0.0010
W1→3	0	52.6717	0.2023	0.0022	0.0017
W2→3	0	17.7742	0.1923	0.0003	0.0001
W4→3	0	118.0492	0.4283	0.0006	0.0005
W5→3	0	92.0437	0.5933	0.0014	0.0010
W1→4	0	51.4558	0.4391	0.0052	0.0014
W2→4	0	61.1021	0.7376	0.0053	0.0025
W3→4	0	23.1272	0.5684	0.0065	0.0016
W5→4	4	106.1549	0.6104	0.0369	0.0230
W1→5	0	4.3541	0.2066	0.0069	0.0014
W2→5	0	2.1968	0.4661	0.0062	0.0011
W3→5	0	26.5907	0.4470	0.0037	0.0014
W4→5	4	12.0703	0.5628	0.0341	0.0227

Tabela A.17: Erro quadrático médio empírico calculado para o cenário 6 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método da máxima verossimilhança sem regularização de parâmetros. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T=100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	2.1725	0.1252	0.0134	0.0091
W3→1	0	2.0791	0.1673	0.0173	0.0068
W4→1	0	2.0520	0.1031	0.0123	0.0065
W5→1	0	2.2953	0.1476	0.0165	0.0091
W1→2	0	3.0986	0.1282	0.0104	0.0056
W3→2	0	2.4261	0.1569	0.0141	0.0078
W4→2	0	3.2322	0.1109	0.0149	0.0079
W5→2	0	2.4817	0.1349	0.0120	0.0064
W1→3	0	2.4642	0.1433	0.0164	0.0067
W2→3	0	2.4232	0.1233	0.0096	0.0050
W4→3	0	2.5007	0.1357	0.0123	0.0072
W5→3	0	1.9786	0.1374	0.0130	0.0068
W1→4	0	32.9673	0.2008	0.0190	0.0082
W2→4	0	39.5284	0.2132	0.0179	0.0123
W3→4	0	27.9801	0.2165	0.0183	0.0101
W5→4	4	47.3706	0.1864	0.0214	0.0102
W1→5	0	3.6728	0.2036	0.0210	0.0096
W2→5	0	5.8746	0.1873	0.0189	0.0082
W3→5	0	5.6926	0.1826	0.0169	0.0087
W4→5	4	143.9530	0.2139	0.0170	0.0131

Tabela A.18: Erro quadrático médio empírico calculado para o cenário 6 a partir de 100 estimativas da matriz de conectividade neuronal de uma rede com 5 neurônios, utilizando o método de estimação do grafo de conectividade proposto neste trabalho. Os cálculos foram realizados considerando 4 diferentes tamanhos amostrais: $T=120$, $T=1200$, $T=12000$, $T=24000$.

Parâmetros	Simulado	$T=100$	$T=1200$	$T=12000$	$T=24000$
W2→1	0	1.3317	0.0809	0.0087	0.0057
W3→1	0	0.8919	0.1218	0.0134	0.0034
W4→1	0	1.1408	0.0503	0.0073	0.0041
W5→1	0	1.3009	0.0924	0.0111	0.0061
W1→2	0	2.0817	0.0582	0.0063	0.0025
W3→2	0	1.2019	0.1057	0.0087	0.0051
W4→2	0	2.0386	0.0631	0.0091	0.0054
W5→2	0	1.4565	0.0706	0.0072	0.0043
W1→3	0	1.5303	0.0838	0.0101	0.0034
W2→3	0	1.5842	0.0620	0.0039	0.0026
W4→3	0	1.3108	0.0707	0.0056	0.0038
W5→3	0	0.9843	0.0830	0.0088	0.0036
W1→4	0	33.2096	0.1177	0.0103	0.0041
W2→4	0	1.4594	0.1278	0.0086	0.0076
W3→4	0	27.5591	0.1113	0.0077	0.0060
W5→4	4	47.4958	0.1635	0.0195	0.0099
W1→5	0	1.7384	0.1248	0.0105	0.0050
W2→5	0	3.4374	0.1028	0.0133	0.0036
W3→5	0	3.8582	0.0970	0.0063	0.0040
W4→5	4	157.6461	0.1811	0.0177	0.0111

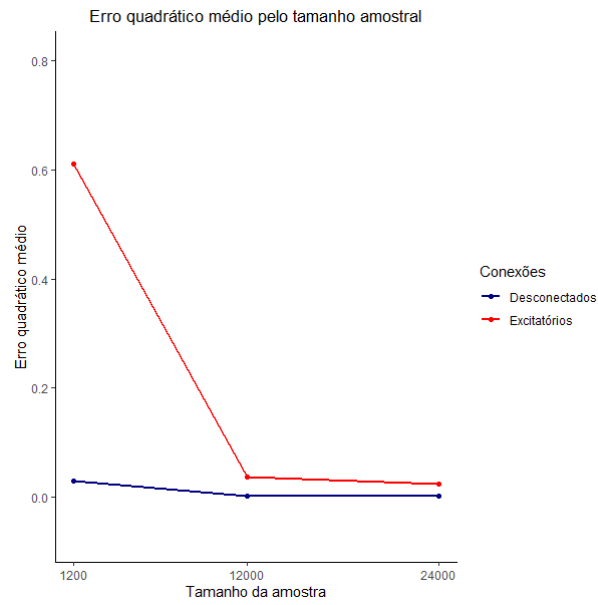


Figura A.16: Gráfico dos valores do erro quadrático médio empírico no cenário 6 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança com regularização ℓ_1 dos parâmetros.

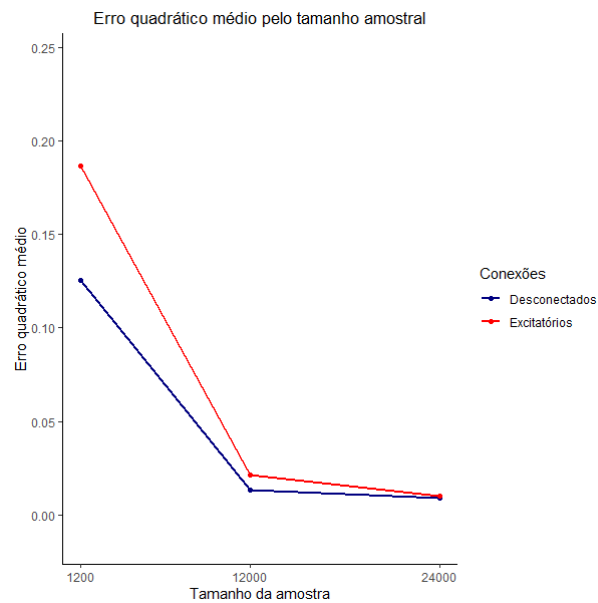


Figura A.17: Gráfico dos valores do erro quadrático médio empírico no cenário 6 em relação ao tamanho da amostra para o método de estimação por máxima verossimilhança sem regularização.

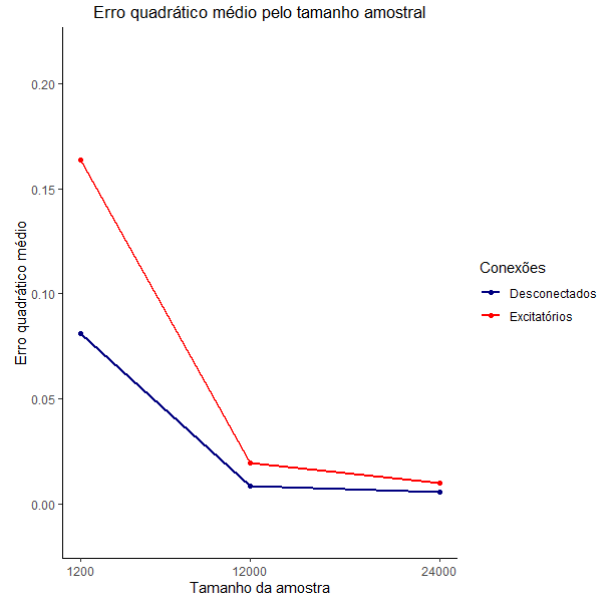


Figura A.18: Gráfico dos valores do erro quadrático médio empírico no cenário 6 em relação ao tamanho da amostra para o método de estimação proposto.

A.7 Cenário 7

Tabela A.19: Distância euclidiana média entre a matriz estimada e a matriz de pesos sinápticos original tamanho de T nos métodos abordados para o cenário 7.

	T=100	T=1200	T=12000	T=24000
EMV + LASSO				
Distância média	11.1808	8.5472	4.6757	2.0446
EMV + Método proposto				
Distância média	12.1567	9.4732	5.1278	3.4941
EMV				
Distância média	58.5056	45.3047	37.3666	36.5527

Apêndice B

Gráficos com os resultados da estimação do grafo de conectividade

B.1 Cenário 1



Figura B.1: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 4 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.



Figura B.2: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 1 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.



Figura B.3: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 1 para o método de do grafo de conectividade inspirado proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.

B.2 Cenário 2



Figura B.4: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 2 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.



Figura B.5: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 2 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.



Figura B.6: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 2 para o método de do grafo de conectividade inspirado proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.

B.3 Cenário 3

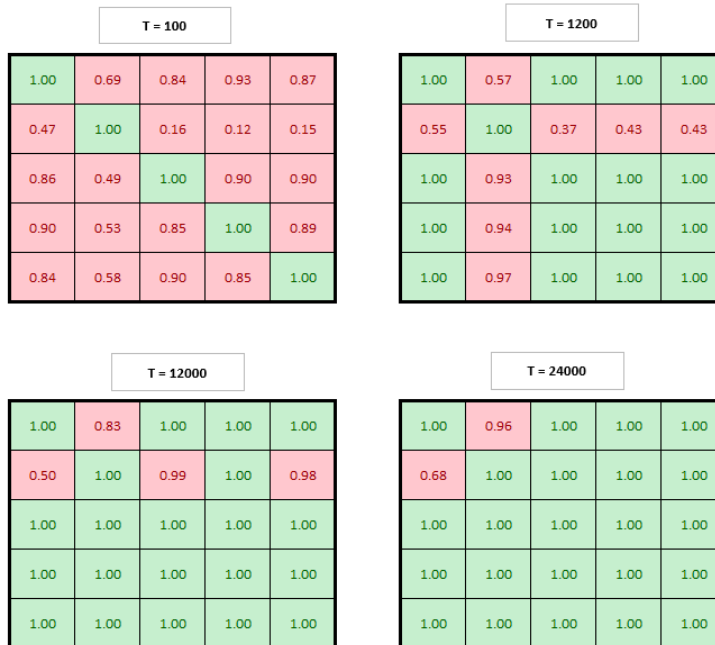


Figura B.7: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 3 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.

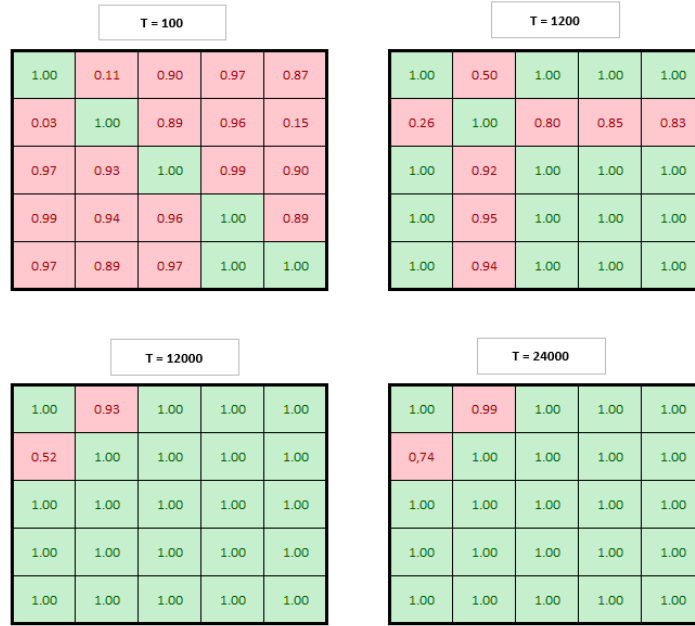


Figura B.8: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 3 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.

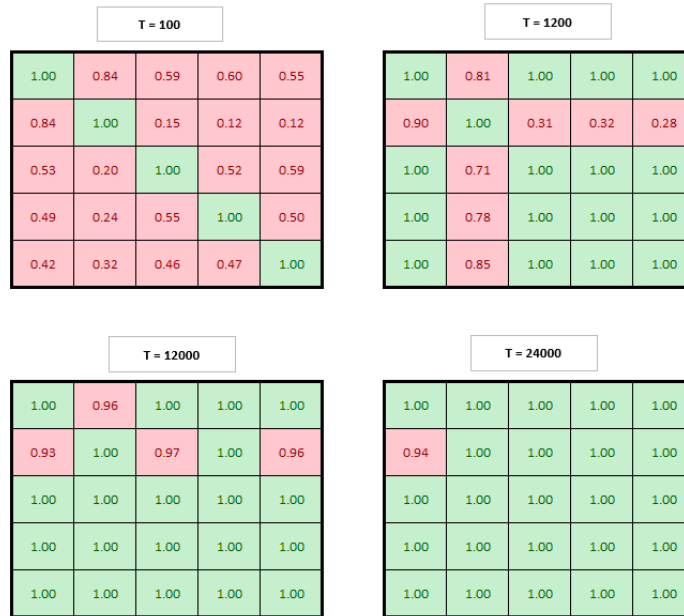


Figura B.9: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 3 para o método de do grafo de conectividade proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.

B.4 Cenário 4



Figura B.10: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 4 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.



Figura B.11: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 4 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.



Figura B.12: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 4 para o método de do grafo de conectividade inspirado proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.

B.5 Cenário 5



Figura B.13: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 5 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.



Figura B.14: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 5 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.



Figura B.15: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 5 para o método de do grafo de conectividade inspirado proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.

B.6 Cenário 6



Figura B.16: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 6 para o método de estimação do grafo de conectividade por máxima verossimilhança com regularização LASSO. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.



Figura B.17: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 6 para o método de estimação do grafo de conectividade por máxima verossimilhança sem regularização. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.



Figura B.18: Ilustração do grafo de conectividade por meio da proporção de acertos da conexão no cenário 6 para o método de do grafo de conectividade inspirado proposto. Código de cor: Verde - 100% na proporção de acertos da conexão, Vermelho - caso contrário. Considerando a medida de precisão $\epsilon = 0.2$.

B.7 Cenário 7

Tabela B.1: Média de acertos de conexões sinápticas nas 100 réplicas simuladas no cenário 7 por todos os métodos com medida de precisão $\epsilon = 0, 2$.

	T=100	T=1200	T=12000	T=24000
EMV + LASSO				
Conexões	76.566	112.600	133.550	147.002
Desconexões	134.800	133.733	158.033	207.132
EMV + Método proposto				
Conexões	75.954	107.983	133.956	145.866
Desconexões	121.652	123.633	148.166	205.133
EMV				
Conexões	144.016	144.216	144.116	144.483
Desconexões	3.483	13.666	62.35	86.266

Apêndice C

Tabelas com os resultados da predição da atividade da rede neuronal

C.1 Cenário 1

Tabela C.1: Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 1 para todos os métodos empregados no estudo com variações da proporção de treino e teste.

	ACC	S	E	VPP	VPN	G-Média	MCC
EMV + LASSO							
10/90	0.7058	0.5681	0.8641	0.5609	0.7831	0.7006	0.4669
50/50	0.6995	0.5566	0.8894	0.5577	0.7787	0.7035	0.4656
90/10	0.6920	0.5626	0.8681	0.5487	0.7832	0.6988	0.4506
EMV + Método proposto							
10/90	0.6987	0.8709	0.6048	0.6145	0.9059	0.7065	0.4522
50/50	0.7085	0.8849	0.5909	0.6217	0.9009	0.7076	0.4702
90/10	0.6987	0.8893	0.5796	0.6286	0.9046	0.7031	0.4592
EMV							
10/90	0.6956	0.8963	0.5792	0.5526	0.8899	0.7075	0.4630
50/50	0.7033	0.8893	0.5867	0.5564	0.8981	0.7079	0.4665
90/10	0.6917	0.8944	0.5753	0.5475	0.9011	0.7037	0.4527

C.2 Cenário 2

Tabela C.2: Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 2 para todos os métodos empregados no estudo com variações da proporção de treino e teste.

	ACC	S	E	VPP	VPN	G-Média	MCC
EMV + LASSO							
10/90	0.8051	0.7175	0.9032	0.9173	0.7020	0.8050	0.6399
50/50	0.8053	0.7143	0.9059	0.9180	0.7002	0.8044	0.6404
90/10	0.8055	0.7160	0.9016	0.9187	0.7110	0.8034	0.6413
EMV + Método proposto							
10/90	0.8059	0.9258	0.7204	0.7009	0.9193	0.8166	0.6425
50/50	0.8060	0.9325	0.7148	0.7069	0.9146	0.8164	0.6438
90/10	0.8064	0.9389	0.7143	0.7082	0.9199	0.8189	0.6527
EMV							
10/90	0.8054	0.9333	0.7111	0.7046	0.9163	0.8146	0.6421
50/50	0.8056	0.89325	0.7137	0.7069	0.9146	0.7984	0.6438
90/10	0.8059	0.9325	0.7158	0.7082	0.9195	0.8169	0.6509

C.3 Cenário 3

Tabela C.3: Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 3 para todos os métodos empregados no estudo com variações da proporção de treino e teste.

	ACC	S	E	VPP	VPN	G-Média	MCC
EMV + LASSO							
10/90	0.8612	0.7963	0.9421	0.9613	0.9464	0.8661	0.8627
50/50	0.8629	0.7917	0.9403	0.9461	0.8686	0.8628	0.8625
90/10	0.8721	0.7948	0.9402	0.9449	0.8691	0.8644	0.8646
EMV + Método proposto							
10/90	0.8602	0.9386	0.7953	0.7914	0.9400	0.8639	0.8626
50/50	0.8640	0.9410	0.7930	0.7903	0.9413	0.8638	0.8645
90/10	0.8799	0.9435	0.7950	0.7920	0.9435	0.8660	0.8651
EMV							
10/90	0.8592	0.9420	0.7906	0.7882	0.9428	0.8629	0.8620
50/50	0.8638	0.9408	0.7925	0.7900	0.9416	0.8634	0.8641
90/10	0.8762	0.9425	0.7947	0.7917	0.9435	0.8654	0.8649

C.4 Cenário 4

Tabela C.4: Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 4 para todos os métodos empregados no estudo com variações da proporção de treino e teste.

	ACC	S	E	VPP	VPN	G-Média	MCC
EMV + LASSO							
10/90	0.9169	0.9262	0.9169	0.9121	0.9303	0.9211	0.8340
50/50	0.9139	0.9287	0.9121	0.9070	0.9322	0.9198	0.8325
90/10	0.9176	0.9265	0.9117	0.9067	0.9305	0.9185	0.8367
EMV + Método proposto							
10/90	0.9156	0.9431	0.8900	0.8887	0.9438	0.9158	0.8337
50/50	0.9153	0.9427	0.8880	0.8867	0.9433	0.9145	0.8330
90/10	0.9201	0.9453	0.8842	0.8832	0.9458	0.9139	0.8399
EMV							
10/90	0.9151	0.9443	0.8880	0.8870	0.9448	0.9154	0.8331
50/50	0.9141	0.9434	0.8869	0.8859	0.9439	0.9144	0.8310
90/10	0.9188	0.9453	0.8842	0.8832	0.9458	0.9139	0.8382

C.5 Cenário 5

Tabela C.5: Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 5 para todos os métodos empregados no estudo com variações da proporção de treino e teste.

	ACC	S	E	VPP	VPN	G-Média	MCC
EMV + LASSO							
10/90	0.8451	0.7612	0.9198	0.8842	0.8272	0.8367	0.6960
50/50	0.8483	0.7659	0.9183	0.8837	0.8287	0.8386	0.6981
90/10	0.8487	0.7706	0.9076	0.8696	0.8318	0.8362	0.6866
EMV + Método proposto							
10/90	0.8491	0.9356	0.7723	0.7678	0.9372	0.8529	0.7075
50/50	0.8501	0.9383	0.7753	0.7719	0.9394	0.8534	0.7125
90/10	0.8507	0.9374	0.7688	0.7643	0.9388	0.8489	0.7046
EMV							
10/90	0.8443	0.9384	0.7704	0.7668	0.9395	0.8487	0.7034
50/50	0.8496	0.9379	0.7747	0.7705	0.9389	0.8524	0.7005
90/10	0.8484	0.9374	0.7679	0.7638	0.9379	0.8484	0.6898

C.6 Cenário 6

Tabela C.6: Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 6 para todos os métodos empregados no estudo com variações da proporção de treino e teste.

	ACC	S	E	VPP	VPN	G-Média	MCC
EMV + LASSO							
10/90	0.7206	0.4338	0.8967	0.7199	0.7214	0.6236	0.3819
50/50	0.7190	0.3776	0.9286	0.7646	0.7084	0.5921	0.3805
90/10	0.7189	0.4843	0.9862	0.6836	0.7319	0.6910	0.3796
EMV + Método proposto							
10/90	0.7211	0.9023	0.6094	0.5856	0.9100	0.7415	0.5039
50/50	0.7222	0.9014	0.6122	0.5880	0.9109	0.7428	0.5057
90/10	0.7224	0.8984	0.6145	0.5881	0.9120	0.7430	0.5044
EMV							
10/90	0.7180	0.9012	0.6071	0.5812	0.9087	0.7396	0.5019
50/50	0.7211	0.9010	0.6109	0.5845	0.9101	0.7419	0.5028
90/10	0.7209	0.8968	0.6134	0.5833	0.9194	0.7416	0.5021

C.7 Cenário 7

Tabela C.7: Métricas da matriz de confusão obtida na previsão do comportamento da rede no cenário 7 para todos os métodos empregados no estudo com variações da proporção de treino e teste.

	ACC	S	E	VPP	VPN	G-Média	MCC
EMV + LASSO							
10/90	0.9132	0.9254	0.9777	0.9521	0.9243	0.9511	0.9088
50/50	0.9589	0.9450	0.9732	0.9729	0.9455	0.9591	0.9182
90/10	0.9557	0.9416	0.9699	0.9695	0.9424	0.9558	0.9117
EMV + Método proposto							
10/90	0.8977	0.8844	0.9139	0.9458	0.9266	0.8994	0.9065
50/50	0.9585	0.9481	0.9690	0.9695	0.9482	0.9585	0.9171
90/10	0.9554	0.9439	0.9671	0.9667	0.9444	0.9555	0.9110
EMV							
10/90	0.6272	0.7598	0.3363	0.6139	0.5889	0.5054	0.4038
50/50	0.6361	0.7607	0.3456	0.6302	0.5911	0.5127	0.4197
90/10	0.6348	0.7589	0.3423	0.6286	0.5950	0.5096	0.4123