

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Programa de Pós-Graduação em Física

Codificação de Sequências Temporais em Padrões Espaciais em Redes Neurais

Nathália Mariath Cristimann

Dissertação de mestrado realizada sob orientação do prof. Dr. Marco Aurélio Pires Idiart e apresentada ao Instituto de Física da UFRGS como requisito parcial para obtenção do título de Mestre em Física

Porto Alegre - Brasil
2019

CIP - Catalogação na Publicação

Mariath Cristimann, Nathália
Codificação de Sequências Temporais em Padrões
Espaciais em Redes Neurais / Nathália Mariath
Cristimann. -- 2020.
82 f.
Orientador: Marco Aurélio Pires Idiart.

Dissertação (Mestrado) -- Universidade Federal do
Rio Grande do Sul, Instituto de Física, Programa de
Pós-Graduação em Física, Porto Alegre, BR-RS, 2020.

1. memória de trabalho. 2. redes aleatórias. 3.
comunicação síncrona-assíncrona. I. Pires Idiart,
Marco Aurélio, orient. II. Título.

Agradecimentos

Agradeço ao meu orientador, prof. Dr. Marco Aurélio Pires Idiart, e aos orientadores com quem trabalhei durante a graduação, prof. Dr Carlos Felipe Lardizabal (Instituto de Matemática) e prof. Dr. Luis Gustavo Pereira (Instituto de Física). Agradeço aos meus familiares, colegas e amigos. Em especial, agradeço aos amigos Giovani, Murilo e Lucas pelas conversas e reflexões que ocorreram recentemente.

Resumo Simplificado

Neurônios presentes nos cérebros de mamíferos recebem milhares de sinais de entrada. Essas células, que são responsáveis por representar e transmitir informações, geram sequências de disparos em diversos padrões temporais, e há evidências de que diferentes redes neuronais cerebrais podem ter formas distintas de manter informações. Dois referenciais teóricos são frequentemente utilizados para modelar a função de memória no cérebro: buffers de memória de trabalho e redes atratoras recorrentes. No presente trabalho, é proposto um modelo de comunicação entre essas diferentes formas de processamento e armazenamento de informação. O modelo consiste em uma estrutura neural composta por dois buffers de memória de trabalho e uma rede recorrente, que faz a tradução de uma sequência de elementos para um padrão espacial. Quando as conexões entre os neurônios da rede recorrente são tais que as ordens de todos os pares de elementos da sequência de entrada são comparadas, o processo de codificação resulta em uma sequência de saída idêntica à sequência de entrada. Entretanto, o mesmo pode não ocorrer quando a rede recorrente possui arquitetura aleatória. Nesse trabalho, iremos estudar o comportamento do modelo proposto para diferentes conectividades da rede recorrente, considerando redes aleatórias construídas de acordo com quatro modelos distintos.

Palavras-chave: modelos de redes, buffers de memória de trabalho, comunicação síncrona-assíncrona

Resumo

Há evidências de que diferentes redes neuronais cerebrais podem ter formas distintas de manter informações, tanto em termos de mecanismo quanto de codificação. Em particular, quando se modela a função de memória no cérebro, dois referenciais teóricos são frequentemente usados: redes atratoras recorrentes e buffers de memória de trabalho baseados na biestabilidade. Neste trabalho, propomos estudar o acoplamento funcional entre diferentes mecanismos de armazenamento e processamento de informação, focalizando o caso especial de uma arquitetura neural composta de dois buffers de memória de trabalho e uma rede recorrente (RNN) que é capaz de manter memórias de longo prazo como atratores. A sequência temporal que chega do buffer de entrada é armazenada como um padrão espacial na RNN, e depois decodificada como um padrão temporal no buffer de saída. Analisamos a questão que diz respeito a possibilidade de uma estrutura de rede aleatória na RNN ser suficiente para garantir a transferência de informação entre os dois buffers. Exploramos quatro modelos de conectividade aleatória: Erdős-Rényi (ER), Watts-Strogatz (WS), Newman-Watts-Strogatz (NWS) e Barabási-Albert (BA). Usando como métrica para o erro de codificação a distância de edição entre as sequências de entrada e saída, mostramos que os modelos de conectividade que correspondem a redes com propriedades de pequeno-mundo são mais eficientes do que os outros modelos.

Palavras-chave: modelos de redes, buffers de memória de trabalho, comunicação síncrona-assíncrona

Abstract

There are evidences that different brain networks may have distinct forms of holding information, both in terms of mechanism and coding. In particular, when modeling memory function in the brain, two theoretical frameworks have been used: recurrent attractor networks and bistability based working memory buffers. In this work we propose a mechanism using inhibitory competition that provides a satisfactory functional coupling between such different forms of information storage and processing. We focus in the simpler case of a neural architecture comprised of two working memory buffers that interact via a recurrent neural network (RNN) that is capable of holding long term memories as attractors. The temporal sequence coming from the input buffer is stored as a spatial pattern in the RNN, and subsequently decoded as a temporal pattern in the output buffer. We address the question of whether a random network structure in RNN could be sufficient to guarantee information transfer between the two buffers. We explore four models of random connectivity: Erdős-Rényi (ER), Watts-Strogatz (WS), Newman-Watts-Strogatz (NWS) and Barabási-Albert (BA). Using as a metric for the encoding error the edit distance between the output and input sequences, we show that the connectivity models which correspond to networks that have small-world properties are more efficient than the other models.

Keywords: network models, working memory buffers, asynchronous to synchronous communication

APOIO DE FINANCIAMENTO CAPES

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior –Brasil (CAPES).

Sumário

1	Introdução	9
2	O Sistema Nervoso	11
2.1	O Neurônio	12
2.2	O Potencial de Ação	13
2.3	A Sinapse	14
2.4	Memória	15
2.4.1	Memória de Longa Duração - Modelo de Hopfield	17
2.4.2	Memória de Trabalho - Modelo Lisman-Idiart	21
3	Metodologia	24
3.1	Modelos de Neurônios	24
3.1.1	Modelo Integra-e-dispara	24
3.1.2	Modelo Taxa de Disparo	28
3.2	Medidas de Redes	33
3.3	Modelos de Redes Aleatórias	38
3.3.1	Erdős-Rényi	38
3.3.2	Pequeno Mundo	40
3.3.3	Livre de Escala	44
3.4	Medidas de similaridade	48
4	Modelo computacional para codificação de sequências temporais	51
4.1	Modelo com Neurônios Binários	52
4.1.1	Rede Ordenada	52
4.1.2	Rede Aleatória	55
4.2	Modelo com Neurônios Taxa de Disparo	58
5	Resultados	61
5.1	Modelo com Neurônios Binários	61
5.1.1	Perfil de ativação dos neurônios da rede recorrente	61

5.1.2	Arquitetura da rede recorrente	65
5.1.3	Tamanho da rede recorrente	68
5.2	Modelo com Neurônios Taxa de Disparo	70
6	Conclusões	73
	Apêndice A Modelo de Hodgkin-Huxley	75

1 Introdução

O armazenamento e a recuperação de sequências de informações são processos fundamentais para percepção e cognição. Visto que diversos tipos de estímulos de entrada que chegam ao cérebro envolvem sequências de informação na qual a ordem é importante, como a fala e a música, são necessários processos de tradução de informação que garantem não somente a identidade dos elementos de entrada, como também a ordem em que ocorrem. Além disso, processos de comunicação internos no cérebro também podem envolver sequências temporais [30, 36].

Existem evidências da utilização de buffers baseados em um esquema de multiplexação temporal pela memória de trabalho no cérebro [26, 52]. Esse esquema é organizado por oscilações teta (com frequência entre 4 e 7Hz) e gama (30-90Hz), de forma que representações de até ~ 7 itens em uma ordem definida podem ser codificadas por múltiplos ciclos gama acoplados em um ciclo teta [38, 40]. Nesse tipo de codificação temporal, uma sequência de itens é representada por diferentes padrões de disparos síncronos, cada um deles em uma fase diferente da oscilação teta. A proporção entre os períodos das duas oscilações em questão determinam aproximadamente a capacidade máxima de armazenamento do buffer.

Redes atratoras recorrentes armazenam informações nas conexões sinápticas e a memória é uma propriedade de rede. Por outro lado, os buffers de memória de trabalho podem se basear em propriedades de célula única e ter uma capacidade de armazenamento muito menor que pode ser contornada por um código de multiplexação como o código temporal teta-gama. Além disso, embora seja provável que redes recorrentes apresentem um estado assíncrono irregular [10], o mesmo pode não ser verdade para os buffers de memória de trabalho do tipo teta-gama, em que a sincronia é um recurso essencial. Mas no final, se ambas as redes estiverem presentes no cérebro, elas precisam se comunicar para trocar informações. Nesse trabalho é proposto um mecanismo cuja estrutura consiste em dois buffers de memória de trabalho e uma rede recorrente atratora, que pode proporcionar o acoplamento funcional entre

esses diferentes mecanismos de armazenamento e processamento de informação. O presente trabalho também tem como objetivo a análise da eficiência do modelo para diferentes conectividades da rede recorrente. Para isso, será definida uma medida que quantifica eficiência da codificação (seção 3.4), e considerados quatro modelos de redes aleatórias: Erdős-Rényi (ER), Watts-Strogatz (WS), Newman-Watts-Strogatz (NWS) e Barabási-Albert (BA).

No capítulo 2, faremos uma breve revisão de neurociência e discutiremos conceitos básicos que serão usados nesta dissertação. No capítulo 3, serão apresentados dois modelos de neurônios (integra-e-dispara e taxa de disparo), e um resumo sobre redes aleatórias. Nesse capítulo também serão apresentados quatro modelos de redes aleatórias, e algumas medidas de similaridade que serão utilizadas nas seções seguintes. No capítulo 4, apresentaremos o modelo computacional com dois tipos de neurônios: binários (seção 4.1) e taxas de disparo (seção 4.2). No capítulo 5, apresentaremos os resultados obtidos das simulações.

2 O Sistema Nervoso

A neurociência tem como objetivo a compreensão do funcionamento do sistema nervoso. Porquanto a atividade cerebral afeta o comportamento, experimentos comportamentais também são usados para avaliar capacidades e limitações da função cerebral [13]. Além disso, modelos computacionais que buscam reproduzir algumas propriedades do cérebro podem auxiliar no entendimento dessas propriedades e de como elas surgem. No que segue, será apresentado um breve resumo sobre conceitos importantes na neurociência; posteriormente serão apresentados dois modelos computacionais de memória.

O sistema nervoso é dividido em Sistema Nervoso Central (SNC) e Sistema Nervoso Periférico (SNP). O sistema nervoso central é constituído pelo encéfalo e pela medula espinhal, e as estruturas que compõem o sistema nervoso periférico são os nervos e os gânglios. A figura 1 mostra uma representação dessa classificação do sistema nervoso.

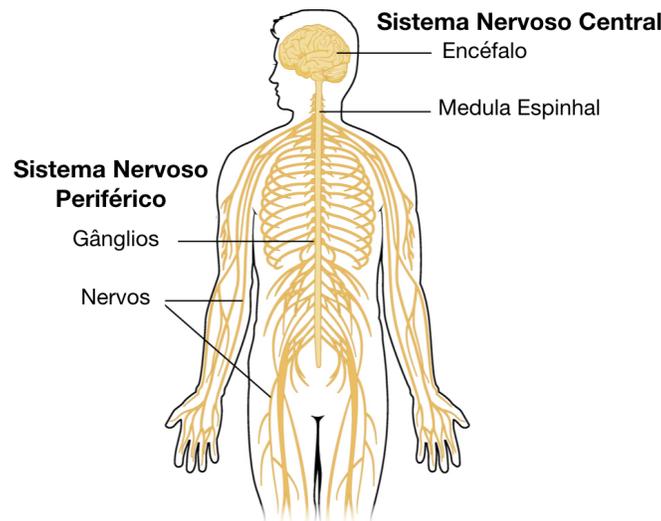


Figura 1: Sistema nervoso. **Fonte:** OpenStax Anatomy and Physiology, 2016 ¹

¹Adaptado. OpenStax - <https://cnx.org/contents/FPtK1zmh@8.25:fEI3C8Ot@10/Preface>, CC BY 4.0.

Todas as partes das estruturas que constituem o SNC são envolvidas por estruturas ósseas e por três membranas de tecido conjuntivo denominadas meninges. O crânio reveste e protege o encéfalo, e a medula espinhal é revestida e protegida pela coluna vertebral. O encéfalo é formado pelas seguintes estruturas: cérebro, cerebelo, tálamo, hipotálamo e bulbo. O SNP, cuja função é conectar o SNC com as outras partes do corpo humano, é dividido em duas partes: SNP somático e SNP visceral. O SNP somático é constituído de nervos espinhais que inervam pele, músculos e articulações que são controlados voluntariamente. O SNP visceral, também chamado de SNP involuntário, consiste de neurônios que inervam vasos sanguíneos, glândulas e órgãos internos.

2.1 O Neurônio

As células do sistema nervoso podem ser classificadas em duas categorias: neurônios (ou células nervosas), e glia (ou células de suporte). Apesar de ambos os tipos de célula existirem em aproximadamente mesma quantidade no cérebro humano adulto (mais de 85 bilhões cada) [13], são os neurônios as células responsáveis pela maioria de funções únicas do cérebro. A figura 2 apresenta uma representação gráfica de um neurônio e suas partes. O neurônio é constituído por quatro regiões: soma (ou corpo neuronal), dendritos, axônio e terminais axonais. A solução aquosa presente dentro do corpo neuronal que contém diversos solutos, como íons inorgânicos e moléculas orgânicas, é chamada citosol. No citosol encontram-se diversas estruturas chamadas organelas, que são as mesmas encontradas em outras células animais. Uma dessas estruturas revestidas por membrana é o núcleo - dentro do qual encontram-se os cromossomos, que contém material genético. O conjunto de estruturas formado pelo citosol e organelas, exceto o núcleo, é chamado citoplasma. Os limites da célula são definidos pela membrana neuronal, que apresenta estrutura de bicamada fosfolipídica.

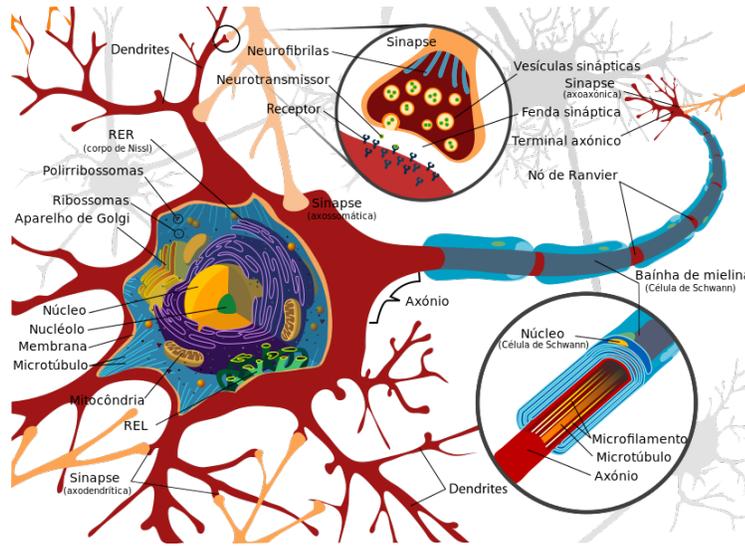


Figura 2: Neurônios. Fonte: RUIZ, 2007. ²

2.2 O Potencial de Ação

Canais iônicos são formados por proteínas na membrana neuronal, e permitem que íons como sódio (Na^+) e potássio (K^+) se movam para dentro e fora da célula. Além disso, proteínas presentes na membrana podem se agrupar para formar estruturas chamadas bombas de íons, que realizam a passagem de íons contra os gradientes de concentração. O potencial de membrana é a diferença de potencial (voltagem) entre as regiões interna e externa da célula. O potencial de equilíbrio ou potencial de equilíbrio iônico é a diferença de potencial na qual o fluxo líquido de um determinado íon através da membrana é nulo. O potencial de membrana de repouso de um neurônio típico é aproximadamente -70mV , e os potenciais de equilíbrio são dados pela equação de Nernst

$$E_{ion} = \frac{RT}{zF} \ln \frac{[ion]_{out}}{[ion]_{in}} \quad (1)$$

²Imagem de domínio público, disponível em <https://commons.wikimedia.org/w/index.php?curid=2683622>

onde z é a carga do íon, R é a constante dos gases, T é a temperatura absoluta, F é a constante de Faraday e $[ion]_{out}$ e $[ion]_{in}$ são as concentrações do íon fora e dentro da célula, respectivamente.

O *potencial de ação* (também chamado *spike* ou impulso nervoso) é uma inversão que ocorre no potencial de membrana que se propaga pelo axônio. Canais de sódio dependentes de voltagem, que permanecem fechados quando o valor potencial está próximo do potencial de repouso da membrana, se abrem quando o potencial atinge um determinado limite, chamado limiar de disparo, ou apenas limiar (*threshold*), causando uma despolarização do potencial de membrana. Esses canais então se fecham, e canais de sódio dependentes de voltagem são ativados, causando a repolarização. O potencial de membrana diminui até um valor mais baixo do que o potencial de repouso da membrana, de modo que há um período chamado pós-hiperpolarização que ocorre antes da restauração do potencial de repouso da membrana. A figura 3 mostra a representação gráfica de um potencial de ação. Estímulos fracos - que não resultam em um aumento no potencial de membrana tal que seu valor alcance o limiar de disparo (representado pela linha hachurada na figura 3) - não são capazes de gerar potenciais de ação.

2.3 A Sinapse

O termo *sinapse* é utilizado para descrever a zona especializada de comunicação entre um neurônio e outro neurônio ou outra célula, e *transmissão sináptica* é o processo de transmissão de informação entre neurônios. Há duas formas de sinapses: a sinapse elétrica e a sinapse química. A sinapse elétrica é feita por meio de junções comunicantes (*gap junction*), que são regiões constituídas por estruturas proteicas especializadas. Nesse tipo de sinapse, a informação é transmitida instantaneamente, com transferência direta de corrente elétrica entre as células pré- e pós-sináptica. Na sinapse química, não há continuidade estrutural entre os neurônios, diferentemente das sinapses elétricas, e a transmissão sináptica é mediada pela liberação de substân-

cias químicas chamadas *neurotransmissores*.

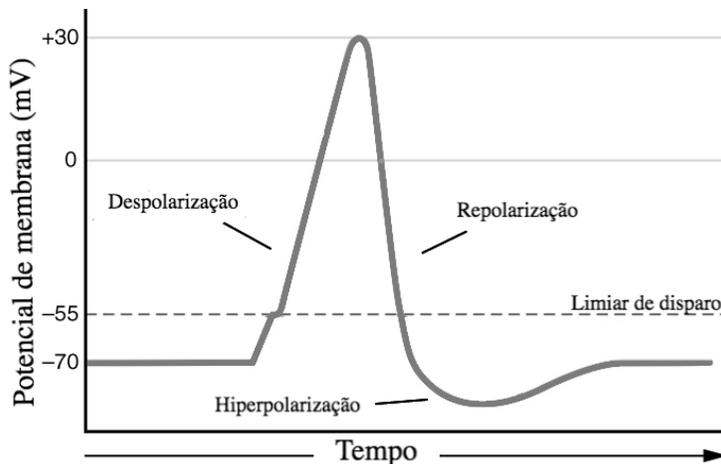


Figura 3: Imagem ilustrativa de um potencial de ação. **Fonte:** OpenStax, 2016. ³

2.4 Memória

Memória é a formação, conservação e evocação de informações [16]. Algumas distinções referentes a determinados aspectos da memória podem ser feitas, levando a classificações úteis do ponto de vista da neurociência. Uma distinção relevante é quanto ao conteúdo: uma memória pode ser classificada como *declarativa* (também chamada de *explícita*) ou *não-declarativa* (*implícita*). Memórias declarativas são aquelas que geralmente pensamos quando nos referimos a memórias - são as memórias de fatos e eventos. Essas memórias podem ser *episódicas* (caráter autobiográfico) ou *semânticas* (fatos e informações genéricas). Memórias episódicas são referentes a experiências de vida, e por isso também são chamadas de memórias *autobiográficas*. Memória não-declarativa é um tipo de memória que não é evocada explícita ou

³Adaptado. OpenStax - <https://cnx.org/contents/FPtK1zmh@8.25:fEI3C8Ot@10/Preface>, CC BY 4.0, <https://commons.wikimedia.org/w/index.php?curid=30147928>.

intencionalmente. As memórias implícitas podem ser divididas nas seguintes categorias: *procedurais*, *priming* e *condicionamento*. Memórias procedurais são referentes a capacidades e habilidades motoras e sensoriais; alguns autores utilizam o termo *priming* para a classificação de memórias que são evocadas através de dicas (por exemplo, quando expostos a um fragmento de som, recordamos uma música inteira). A figura 4 mostra uma representação gráfica das classificações de memória.

Há dois tipos de condicionamento: *condicionamento clássico* e *condicionamento operante* (ou *instrumental*). O condicionamento clássico foi descoberto e caracterizado pelo fisiologista russo Ivan Pavlov no século 19. Esse tipo de condicionamento envolve a associação de um estímulo que causa uma resposta mensurável (chamado estímulo incondicionado) com um estímulo que não causa essa mesma resposta (estímulo condicionado). No famoso experimento de Pavlov, os cães recebem carne após ouvir um sino. Depois de diversas repetições dessa situação, o som do sino é o suficiente para fazer os cães salivarem. Nesse experimento, receber a carne corresponde ao estímulo incondicionado descrito acima, que causa como resposta a salivação dos cães, e o sino é o estímulo condicionado - que os cães passam a associar ao recebimento da comida. No condicionamento operante, um determinado comportamento é reforçado ou punido, resultando em seu fortalecimento ou extinção. É o caso, por exemplo, de um rato que recebe comida se realizar determinada atividade.

Quanto a sua duração, as memórias podem ser divididas em *memórias de curto prazo* (duração entre 1 e 6 horas) e *memória de longo prazo* (duração entre dias e décadas) [13]. Além disso, há um outro tipo de memória: a *memória de trabalho*⁴, que é responsável por manter informações relevantes para uma determinada tarefa, e sua duração é da ordem de segundos. A memória de trabalho é geralmente estudada por meio de experimentos que medem o *span de dígito* de um indivíduo (quantidade máxima de números que uma pessoa é capaz de repetir, após ouvi-los em uma ordem aleatória). Em comparação com a memória de curto prazo, a memória de trabalho apresenta, além de duração mais curta, capacidade mais limitada (7 ± 2

⁴O termo *memória de trabalho* foi introduzido por Miller, Galanter e Pribram em 1960 no livro 'Plans and the structure of behaviour' [48], e utilizado por Atkinson e Shiffrin em 1968 [6]. O termo é utilizado muitas vezes como sinônimo de memória de curto prazo, embora a maioria dos autores considerem distintas essas duas formas de memória. [18]

itens, conforme sugerido por Miller em 1956 [40]). Nas duas seções seguintes, serão apresentados modelos computacionais de dois tipos de memória: memória de longo prazo (modelo de Hopfield) e memória de trabalho (modelo Lisman-Idiart).

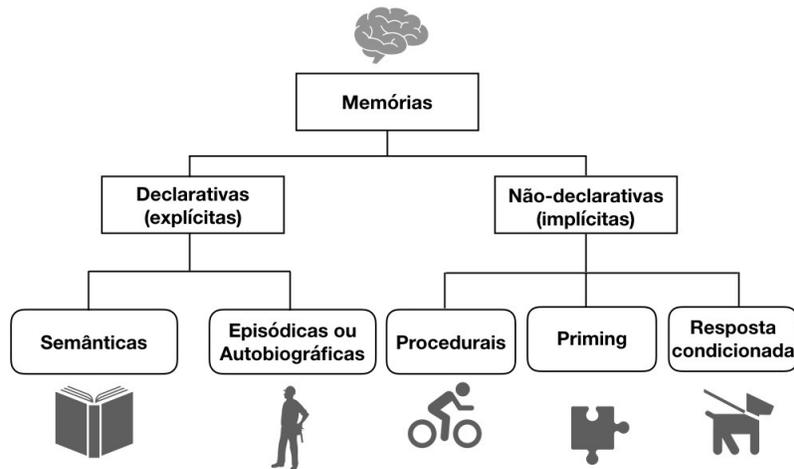


Figura 4: Representação esquemática dos tipos de memória. **Fonte:** Autora

2.4.1 Memória de Longa Duração - Modelo de Hopfield

Redes de Hopfield - introduzidas por Hopfield em 1982 [28] e também descritas por Little em 1974 [39] - funcionam como sistemas de memória associativa, e constituem um modelo que tem como objetivo a compreensão da memória humana. Em memórias associativas, diferente de memórias RAM (*random access memory*), a recuperação da memória é baseada em conteúdo, e não em endereço. Por esse motivo, memórias associativas também são chamadas de *memórias endereçadas por conteúdo* (CAM) [17]. A função de um sistema de memória associativa é recuperar padrões de disparos neurais previamente aprendidos, a partir de padrões de entrada incompletos.

Uma rede de Hopfield é composta por N neurônios binários, i.e., unidades que

possuem somente dois estados: "ligado", se a variável de estado do neurônio i assume o valor $S_i = +1$, e "desligado", se assume o valor $S_i = -1$. Cada par de neurônios i, j possui uma conexão descrita pelo peso de conectividade w_{ij} . Essas conexões possuem, em geral, as seguintes restrições: $w_{ii} = 0, \forall i$ (uma unidade não se conecta com ela mesma), e $w_{ij} = w_{ji}$ (as conexões são simétricas). A primeira restrição implica uma diagonal principal com elementos nulos na matriz de conexão W cujos elementos são dados pelos pesos w_{ij} . A segunda restrição resulta em uma matriz W simétrica. A dinâmica da rede é dada por

$$S_i = \text{sgn}\left(\sum_j w_{ij} S_j - \theta_i\right) \quad (2)$$

onde θ_i é o limiar da unidade i , e a função $\text{sgn}(x)$ é

$$\text{sgn}(x) = \begin{cases} 1, & \text{se } x \geq 0 \\ -1, & \text{se } x < 0 \end{cases} \quad (3)$$

As atualizações especificadas por (2) podem ser realizadas de duas maneiras:

- **Síncrona:** Todas as unidades são atualizadas simultaneamente;
- **Assíncrona:** As unidades são atualizadas uma por vez.

O primeiro modelo requer um relógio central (ou marcapasso) para o sistema a fim de manter a sincronização. Esse modelo é considerado, pela maioria dos autores, como menos realista. O modelo assíncrono seria, portanto, o mais apropriado tanto para o cérebro, como para redes artificiais [25]. A atualização assíncrona pode ser feita de duas formas: na primeira, uma unidade é selecionada aleatoriamente a cada passo de tempo, e atualizada conforme a equação (2); na segunda, cada unidade é atualizada conforme (2) independentemente com uma probabilidade constante por unidade de tempo.

Uma rede de Hopfield pode ser treinada (ou seja, os pesos sinápticos podem ser especificados) utilizando a regra de Hebb ⁵

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^m \xi_i^\mu \xi_j^\mu \quad (4)$$

onde ξ_i^μ representa o i -ésimo bit do padrão μ , e m é o número de padrões. Um modelo de memória associativa com unidades binárias e dinâmica assíncrona que utiliza a regra de Hebb (4) é geralmente chamado de modelo de Hopfield.

Uma importante contribuição do modelo proposto por Hopfield foi a introdução da ideia de energia na teoria de redes neurais. Um valor escalar chamado de energia E é associado aos estados de uma rede de Hopfield:

$$E = -\frac{1}{2} \sum_{i,j} w_{ij} S_i S_j + \sum_i \theta_i S_i \quad (5)$$

Vamos mostrar que a quantidade dada por (5) ou permanece constante, ou decai com a atualização dos valores das unidades da rede e é, portanto, uma função de Lyapunov. Para conexões simétricas e considerando $\theta_i = 0$, $\forall i$, a equação (5) pode ser escrita como

$$E = C - \sum_{(ij)} w_{ij} S_i S_j \quad (6)$$

onde (ij) representa todos os pares distintos ij , e C é uma constante dada pelos termos $i = j$ do somatório. Consideremos que, após a atualização dada por (2), um neurônio i teve seu valor modificado de S_i no tempo t para $S'_i = -S_i$ em $t + \Delta t$, enquanto todos os outros neurônios mantêm seus valores. A variação da energia é

⁵Introduzida por Donald Hebb em 1949 no livro "A Organização do Comportamento" [24], a teoria de Hebb busca explicar o aprendizado associativo (ou aprendizado de Hebb), no qual a ativação simultânea de células resulta no aumento da força sináptica entre essas células. Em redes neurais artificiais, a regra de Hebb é utilizada para determinar alterações nos pesos sinápticos das conexões entre os neurônios da rede.

$$\begin{aligned}
E' - E &= - \sum_{j \neq i} w_{ij} S'_i S_j + \sum_{j \neq i} w_{ij} S_i S_j \\
&= 2S_i \sum_{j \neq i} w_{ij} S_j \\
&= 2S_i \sum_j w_{ij} S_j - 2w_{ii}
\end{aligned} \tag{7}$$

O primeiro termo é negativo, pois $S'_i = \text{sgn}(\sum_j w_{ij} S_j) = -S_i$. Assim, para o caso analisado aqui, no qual uma unidade não se conecta com ela mesma ($w_{ii} = 0$), concluimos que $E' - E < 0$.

A inicialização da rede é feita atribuindo aos valores das unidades o padrão inicial desejado. Após repetidas atualizações, a rede irá convergir para um estado que é mínimo local da função de energia (5). Esse estado é um estado de rede, ou seja, é definido pelo conjunto de todos os estados dos neurônios individuais. Como o modelo de Hopfield é binário, o estado da rede ocupa os vértices de um hipercubo de N dimensões.

Redes de Hopfield são *redes atratoras*, pois sua dinâmica converge para estados atratores (i.e., estados com energia mínima). Ou seja, a presença das conexões sinápticas faz com que algumas configurações neurais (quais neurônios estão ligados e quais estão desligados) sejam preferidas. Então, se a rede começa de um estado qualquer, a sua dinâmica - que simula a descida do gradiente de energia - vai levar o estado da rede para um destes estados preferidos de energia mínima. O sistema é, portanto, um sistema dinâmico que apresenta atratores do tipo ponto fixo, e esses pontos fixos são interpretados como as memórias armazenadas na rede.

Apesar da importância do modelo de Hopfield para a área de Redes Neurais por ser o modelo que introduziu a ideia das redes recorrentes atratoras na biologia, hoje em dia ele é considerado um modelo clássico mas ligeiramente desatualizado. A maior crítica a esse modelo é justamente de ele considerar estados estereotipados de disparo (neurônios que estariam simplesmente "ligados" ou "desligados") e também a hipótese das conexões simétricas. Na biologia, diferente da física, a interação é

dada por contatos sinápticos que podem ou não existir nas duas direções. Hopfield considerou a simetria justamente para fazer o modelo mais tratável usando conceitos de física estatística. De fato, se a condição de simetria for abandonada, o modelo de rede neural resultante não poderá mais ser descrito por uma função de energia e passará, entretanto, a apresentar uma riqueza muito maior de comportamentos que aproximam o modelo do comportamento real das redes neuronais no cérebro.

2.4.2 Memória de Trabalho - Modelo Lisman-Idiart

Desde a formulação da teoria de que a ativação correlacionada de neurônios leva a mudanças na força da conectividade neuronal, em 1949 [24], a análise da atividade sincronizada no cérebro está presente em diversos estudos sobre memória. Um conceito importante nesses estudos é que a sincronização da atividade neuronal é um dos mecanismos de armazenamento de informação no cérebro. Um conjunto de neurônios pode sincronizar sua atividade em frequências específicas, de forma que são produzidas oscilações, e estas podem ser medidas por métodos não invasivos, como electroencefalograma (EEG)⁶ e magnetoencefalograma (MEG). Em seres humanos, as frequências podem variar entre 0.05Hz e 500Hz, e são classificadas por letras gregas, de acordo com o alcance (*delta*: < 4Hz, *teta*: 4-7Hz, *alfa*: 8-13Hz, *beta*: 15-30Hz, *gama*: 30-90Hz) [13]. Estudos apontam uma relação entre oscilações neurais e diversos processos cognitivos em humanos. Em particular, resultados de experimentos de EEG e MEG evidenciam um papel funcional das oscilações gama e teta na memória [29].

Há dois aspectos importantes no que diz respeito à diferença entre memórias de curto e de longo prazo: duração e capacidade. Experimentos fisiológicos sugerem que a memória de longo prazo é armazenada por meio de mudanças nas forças das conexões entre os neurônios, enquanto a memória de trabalho depende de atividade

⁶O electroencefalograma foi descrito pela primeira vez pelo psiquiatra austríaco Hans Berger, em 1929, e hoje é utilizado para auxílio do diagnóstico de certas condições neurológicas e em estudos sobre sono e processos cognitivos.

neuronal persistente [20, 22]. Além disso, há evidências de que essa atividade neuronal persistente, presente na memória de trabalho, pode ter caráter oscilatório [49].

O modelo proposto por Lisman e Idiart [38] considera a memória de trabalho como sendo uma memória de curtíssimo prazo (duração da ordem de segundos) que em princípio não depende de sinapses. A memória é mantida ativamente utilizando um mecanismo de célula única, baseado em propriedades conhecidas de neurônios hipocampais e corticais: um aumento na excitabilidade da membrana causa a manutenção do disparo das células [5]. A figura 5 mostra a representação do modelo.

Em condições normais, uma célula piramidal do hipocampo sofre uma pós-hiperpolarização após emitir um potencial de ação. Entretanto, na presença de neurotransmissores como acetilcolina e serotonina, o disparo da célula induz uma *pós-despolarização* (do inglês *afterdepolarization*, ADP), que resulta em um aumento transiente na excitabilidade da membrana. Esse ADP possui duração suficiente para armazenar informações entre ciclos de oscilações teta-alfa (entre 5 e 12 Hz). Portanto, se o ADP desencadeado em um ciclo promove o disparo das células no próximo ciclo, esse processo poderia se repetir (o ADP seria renovado a cada ciclo de oscilações), de modo que os disparos seriam mantidos.

Cada neurônio (figura 5A) recebe, além da entrada excitatória que carrega a informação a ser armazenada, uma entrada que gera uma oscilação sub-limiar de baixa frequência. As simulações mostram que um neurônio entra em modo de disparo repetitivo após sua excitação por um breve sinal de entrada. Outro aspecto importante levado em consideração na construção do modelo está relacionado aos resultados de experimentos realizados por Sternberg [52], nos quais os participantes eram expostos a uma lista de itens, e depois a um item de teste, e deviam indicar se o item estava ou não presente na lista. Para cada item adicionado à lista, o tempo de reação (tempo que o participante levava para apertar um botão indicando a existência de um item de teste na lista) teve um acréscimo de $\sim 38ms$. Esse incremento corresponde a um ciclo de uma oscilação beta-gama.

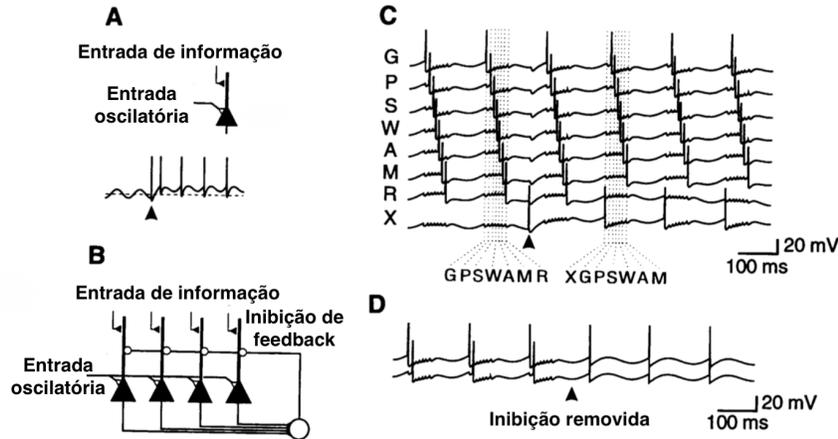


Figura 5: (A) O neurônio recebe duas correntes: uma corrente informacional supra-limiar, e uma corrente sub-limiar que induz o potencial de membrana a oscilar com frequência teta. Simulações mostram o potencial de membrana antes e depois da corrente informacional (indicada com a ponta de uma flecha). (B) Rede na qual células excitatórias piramidais estão conectadas a um neurônio inibitório que produz inibição de feedback nas células piramidais. (C) A rede é capaz de manter os disparos e a fase correta de sete grupos de células que estão ativas durante diferentes subciclos da oscilação teta. A capacidade limitada de memória da rede é demonstrada pela sua incapacidade de armazenar oito memórias. (D) Se a inibição de feedback é removida, a informação de fase é rapidamente perdida. **Fonte:** Adaptado de LISMAN & IDIART, 1995 [38] ⁷

Cada item é representado pelas células que disparam durante um determinado ciclo da oscilação gama. O primeiro item é representado pelos neurônios que disparam no primeiro ciclo gama, o segundo item, pelos neurônios que disparam no segundo ciclo gama, e assim sucessivamente. O número de itens que pode ser armazenado pela rede é determinado pelo número de subciclos que cabe dentro de um ciclo de baixa frequência. Outra característica importante do modelo Lisman-Idiart é a inibição de feedback gerada por neurônios inibitórios (figuras 5B, 5D), necessária para garantir a existência dos subciclos. Oscilações teta-gama acopladas foram observadas em diversas regiões do cérebro [7,53], e estudos sugerem que o esquema temporal organizado por esse acoplamento das oscilações descrito pode funcionar como um código mais geral utilizado pelo cérebro para o armazenamento de múltiplos itens ordenados [37].

⁷Imagem utilizada com permissão

3 Metodologia

3.1 Modelos de Neurônios

Com o objetivo de relacionar resultados de simulações com descobertas experimentais, faz-se necessário empregar modelos de neurônios que apresentem a dinâmica de neurônios reais em resposta a estímulos. Em 1952, Hodgkin e Huxley propuseram um modelo que consiste em modelar de forma explícita a dinâmica das condutâncias iônicas (apêndice A). A vantagem desse modelo e de modelos similares, chamados *modelos baseados em condutância*, é que as variáveis utilizadas possuem interpretação direta em termos de variáveis que são medidas experimentalmente. Por outro lado, o grande número de variáveis existentes nesses modelos e a relação não-linear entre elas tornam difícil a análise matemática desse tipo de modelo. Além disso, o grande número de variáveis torna as simulações desses modelos caras do ponto de vista computacional. A fim de evitar essas desvantagens, modelos de neurônios mais simples dos pontos de vista matemático e computacional foram propostos. No que segue, serão apresentados dois desses modelos que são mais simples e muito utilizados em simulações.

3.1.1 Modelo Integra-e-dispara

Nos modelos do tipo integra-e-dispara, não existe uma tentativa de descrever o formato dos potenciais de ação. Como os potenciais de ação de um determinado neurônio possuem todos aproximadamente o mesmo formato, a informação transmitida pelo neurônio deve estar não na forma, mas na presença ou ausência dos potenciais de ação. Assim, nesses modelos os potenciais de ação são interpretados como “eventos”.

O modelo de neurônio conhecido como *integra-e-dispara* foi proposto pela primeira vez por Lapicque, em 1907 [35]. O circuito básico de um modelo desse tipo consiste em um capacitor C em paralelo com um resistor R , com corrente total I . A corrente pode ser dividida em duas componentes: $I = I_R + I_{cap}$. O primeiro termo é a corrente resistiva, que passa pelo resistor R , e é dada pela lei de Ohm $I_R = V/R$, onde V é a voltagem no resistor, e R é a resistência. O segundo termo, I_{cap} , é a corrente que carrega o capacitor. A partir da definição de capacidade $C = q/V$, onde q é a carga e V a voltagem, a corrente capacitiva é dada por $I_{cap} = CdV/dt$. Assim, a corrente total é

$$I = \frac{V(t)}{R} + C \frac{dV}{dt} \quad (8)$$

onde $V(t)$ representa o potencial de membrana no tempo t . Multiplicando a equação (8) por R e introduzindo a constante $\tau_m = RC$, chamada constante de tempo da membrana, temos

$$\tau_m \frac{dV}{dt} = -V(t) + RI(t) \quad (9)$$

Quando a voltagem atinge um determinado valor limiar V_{th} , assume-se que o neurônio emite um disparo. Quando ocorre um disparo, a voltagem é instantaneamente restabelecida para um valor de repouso V_r . A equação (9) representa a dinâmica do modelo conhecido como *integra-e-dispara com vazamento*, cuja sigla em inglês é LIF (*leaky integrate-and-fire*). Pode ser incorporado ao modelo um período refratário absoluto τ_{abs} , no qual, após o potencial de membrana atingir V_{th} , não existe a possibilidade de formação de novos disparos. Nesse caso, se V atinge o limiar num tempo $t = t^{(f)}$, a dinâmica (9) é interrompida durante o tempo τ_{abs} , e recomeça, então, no tempo $t^{(f)} + \tau_{abs}$. Nos exemplos que seguem, serão analisados dois tipos de estímulo de entrada: corrente constante, e corrente variável no tempo.

Exemplo 3.1. Corrente constante

Vamos considerar o caso em que o estímulo de entrada é uma corrente constante

$I = I_0$. Assumiremos, por simplicidade, $V_r = 0$ e $t_0 = 0$. A solução da equação (9) é

$$V(t) = RI_0[1 - \exp(-\frac{t}{\tau_m})] \quad (10)$$

o valor assintótico do potencial de membrana é RI_0 . Se esse valor é menor do que o limiar V_{th} , não pode haver a geração de potenciais de ação. Se $RI_0 > V_{th}$, então o neurônio dispara periodicamente. Considerando $V(0) = V_r = 0$, o tempo $t = t^{(1)}$ em que o neurônio dispara pela primeira vez pode ser obtido da equação (10). Observando que $V(t^{(1)}) = V_{th}$, obtemos

$$t^{(1)} = \tau_m \ln\left(\frac{RI_0}{RI_0 - V_{th}}\right) \quad (11)$$

Após o disparo, o potencial de membrana retorna ao valor de repouso, e o processo de integração recomeça. Se o estímulo de entrada permanece constante, o próximo disparo irá ocorrer após um intervalo de duração T , que também é dado pela equação (12). Portanto, a frequência de disparos (ou *taxa de disparos*, definida como $f = 1/T$) é dada por

$$f = T^{-1} = [\tau_m \ln\left(\frac{RI_0}{RI_0 - V_{th}}\right)]^{-1} \quad (12)$$

Se um período refratário absoluto τ_{abs} for adicionado ao modelo, o período entre os disparos passa a ser

$$T = \tau_{abs} + \tau_m \ln\left(\frac{RI_0}{RI_0 - V_{th}}\right) \quad (13)$$

e a frequência nesse caso é dada por

$$f = [\tau_{abs} + \tau_m \ln\left(\frac{RI_0}{RI_0 - V_{th}}\right)]^{-1} \quad (14)$$

A figura 6 mostra a simulação de um neurônio do tipo integra-e-dispara com vazamento (LIF) estimulado por uma corrente constante.

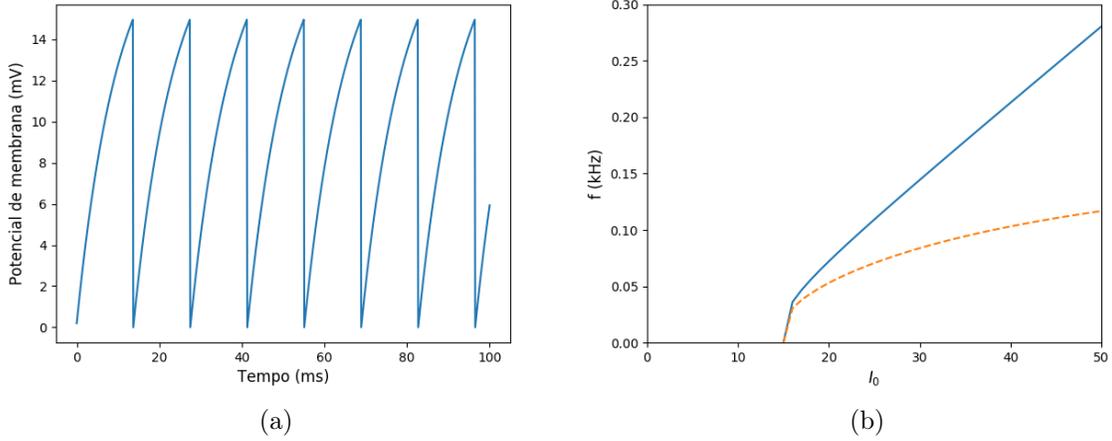


Figura 6: Simulação de um neurônio LIF estimulado por uma corrente constante. **(a)** potencial de membrana para um neurônio sem período refratário absoluto. **(b)** A curva frequência contra corrente (curva $f - I$) para um neurônio com $\tau_{abs} = 0ms$ (linha sólida) e $\tau_{abs} = 5ms$ (linha hachurada). Foram utilizados os seguintes parâmetros: $RI_0 = 20mV$, $\tau_m = 10ms$, $V_r = 0$, $V_{th} = 15mV$. **Fonte:** Autora

Exemplo 3.2. Corrente variável

Para o caso em que a corrente de entrada $I(t)$ varia no tempo, a solução da equação (9) é

$$V(t) = V_r \exp\left(-\frac{t-t_0}{\tau_m}\right) + \frac{R}{\tau_m} \int_0^{t-t_0} \exp\left(-\frac{s}{\tau_m}\right) I(t-s) ds \quad (15)$$

É fácil ver que a equação (10) é um caso particular de (15), em que a corrente é constante. Quando o potencial alcança o limiar V_{th} , ele é instantaneamente restabelecido para o valor de repouso V_r . A equação (15) descreve a dinâmica do potencial $V(t)$ entre os disparos do neurônio. A figura 7 mostra a simulação de um neurônio do tipo integra-e-dispara com vazamento para o caso em que a corrente de entrada $I(t)$ não é constante.

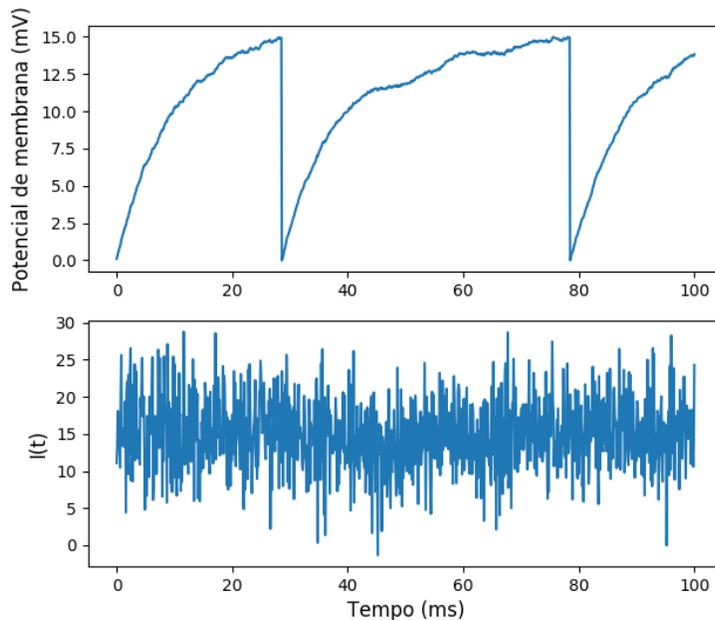


Figura 7: Simulação de um neurônio LIF estimulado por uma corrente variável. O gráfico inferior mostra a corrente $I(t)$. Nesse exemplo, foi utilizada uma corrente cujos valores a cada passo de tempo obedecem uma distribuição Gaussiana com média 15 e desvio padrão 5. O valor da resistência é $R = 1m\Omega$, de forma que a corrente e o potencial de membrana possuem a mesma unidade. O gráfico superior mostra o comportamento do potencial de membrana $V(t)$ em resposta à $I(t)$. Foram utilizados os seguintes parâmetros: $\tau_m = 10ms$, $V_r = 0$, $V_{th} = 15mV$. **Fonte:** Autora

3.1.2 Modelo Taxa de Disparo

Em algumas regiões do cérebro, neurônios estão organizados em populações de unidades que apresentam propriedades semelhantes [33]. Devido ao grande número de neurônios nessas populações, é possível descrever a dinâmica da rede em termos da atividade média da população, em vez de modelar os potenciais de ação individualmente. Modelos *taxa de disparo*, que descrevem a atividade média de um conjunto de neurônios, constituem uma abordagem interessante para o estudo de grandes redes neurais devido a eficiência de simulação e análise matemática relativamente simples.

Desde o trabalho de Wilson e Cowan [56], publicado em 1972, diversos modelos taxa de disparo foram sugeridos.

Para construir um modelo do tipo taxa de disparo, devemos primeiro determinar como a taxa de disparo de um neurônio se relaciona com as taxas de disparo das entradas que ele recebe. Consideremos um neurônio que recebe N_u entradas sinápticas $b = 1, 2, \dots, N_u$, conforme representado na figura 8. As taxas de entrada são representadas pelo vetor \mathbf{u} , de N_u componentes, onde u_b é a taxa de disparo da entrada b . Os pesos sinápticos são representados pelo vetor \mathbf{w} , composto de N_u componentes w_b , que determinam a amplitude da corrente sináptica. Para sinapses excitatórias, $w_b > 0$, e para sinapses inibitórias, $w_b < 0$. Se um potencial de ação chega na entrada b no tempo $t = 0$, a corrente sináptica gerada no soma do neurônio pós-sináptico no tempo t é escrita como $I_b(t) = w_b K_s(t)$, onde $K_s(t) \geq 0$ é chamado núcleo sináptico e descreve a corrente sináptica em resposta a um potencial de ação pré-sináptico que ocorre no tempo $t = 0$.

Assumindo que os potenciais de ação em uma sinapse são independentes, a corrente sináptica total na entrada b no tempo t devido à sequência de potenciais de ação pré-sinápticos emitidos nos tempo t_i é dada por

$$I_b(t) = w_b \sum_{t_i < t} K_s(t - t_i) = w_b \int_{-\infty}^t d\tau K_s(t - \tau) \rho_b(\tau) \quad (16)$$

onde $\rho_b(\tau) = \sum_i \delta(\tau - t_i)$, chamada função de resposta neural, descreve a sequência de potenciais de ação emitidos pelo neurônio pré-sináptico b . Assumindo que as contribuições de diferentes sinapses somam linearmente, a corrente sináptica total é

$$I_s = \sum_b w_b \int_{-\infty}^t d\tau K_s(t - \tau) \rho_b(\tau) \quad (17)$$

No modelo taxa de disparo se assume, devido ao grande número de sinapses de entrada, que a função de resposta neural $\rho_b(\tau)$ pode ser substituída pela taxa de

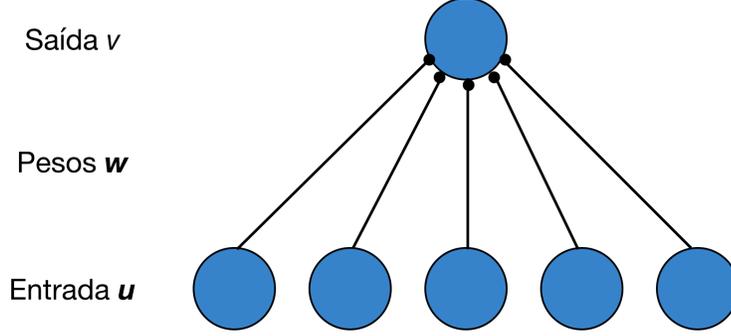


Figura 8: Sinais de entrada recebidos por um único neurônio. As taxas de entrada dadas pelo vetor \mathbf{u} levam o neurônio a uma taxa de disparo de saída v , através de pesos sinápticos que são determinados pelo vetor \mathbf{w} . **Fonte:** Autora

disparo do neurônio. Portanto, podemos escrever

$$I_s = \sum_b w_b \int_{-\infty}^t d\tau K_s(t - \tau) u_b(\tau) \quad (18)$$

O núcleo sináptico mais utilizado nesse tipo de modelo possui a forma exponencial

$$K_s(t) = \frac{1}{\tau_s} \exp(-t/\tau_s) \quad (19)$$

Utilizando (19) e derivando a equação (18) com respeito a t , podemos descrever I_s por uma equação diferencial

$$\tau_s \frac{dI_s}{dt} = -I_s + \sum_{b=1}^{N_u} w_b u_b = -I_s + \mathbf{w} \cdot \mathbf{u} \quad (20)$$

Essa equação determina a corrente sináptica que entra no soma de um neurônio pós-sináptico em termos das taxas de disparo dos neurônios pré-sinápticos. Se a corrente sináptica for constante, podemos escrever a taxa de disparo do neurônio pós-sináptico como $v = F(I_s)$, onde F , chamada função de ativação, é a taxa de disparo de curso estável em função da corrente de entrada. Geralmente, F é tomada como uma função de saturação como a função sigmoideal. Portanto, para entradas

independentes do tempo, a relação $v = F(I_s)$ junto com a equação (20) descrevem o modelo. Para entradas dependentes do tempo, a hipótese mais simples é $v = F(I_s(t))$, de forma que a dinâmica do modelo taxa de disparo é dada pelas equações

$$\tau_s \frac{dI_s}{dt} = -I_s + \mathbf{w} \cdot \mathbf{u}, \quad v = F(I_s) \quad (21)$$

Uma formulação alternativa do modelo taxa de disparo pode ser construída assumindo que as taxas de disparo de um neurônio não são afetadas instantaneamente por mudanças na corrente sináptica total. A taxa de disparo dependente do tempo, nesse caso, é descrita por

$$\tau_r \frac{dv}{dt} = -v + F(I_s(t)) \quad (22)$$

A constante τ_r determina, para corrente constante, a rapidez com a qual a taxa de disparo se aproxima do seu valor estável, e para corrente variável, o quão próximo v pode ser afetada por flutuações rápidas de $I_s(t)$. Essa nova formulação do modelo consiste nas equações (20) e (22). Se uma dessas equações atinge o equilíbrio muito mais rápido do que a outra, o par de equações pode ser reduzido a uma única equação. Se $\tau_r \gg \tau_s$, a substituição $I_s = \mathbf{w} \cdot \mathbf{u}$ na equação (22) pode ser feita:

$$\tau_r \frac{dv}{dt} = -v + F(\mathbf{w} \cdot \mathbf{u}) \quad (23)$$

A figura 9 mostra exemplos de redes com dois tipos de conectividades: *feedforward* e recorrente. A rede *feedforward* ilustrada na figura é composta por N_v unidades de saída com taxas v_a ($a = 1, 2, \dots, N_v$), representadas pelo vetor \mathbf{v} , e N_u unidades de entrada com taxas dadas pelas componentes do vetor \mathbf{u} . As componentes W_{ab} da matriz \mathbf{W} representam a força da conexão sináptica da unidade b até a unidade a .

Usando a formulação da equação (23), as taxas de disparo das unidades de saída são dadas por

$$\tau_r \frac{d\mathbf{v}}{dt} = -\mathbf{v} + \mathbf{F}(\mathbf{W} \cdot \mathbf{u}) \quad (24)$$

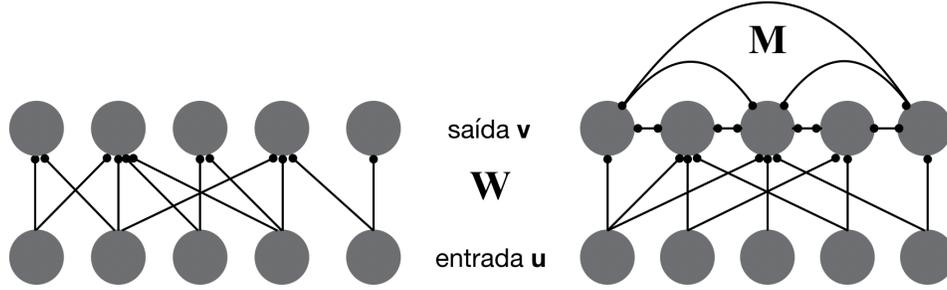


Figura 9: Exemplos de rede *feedforward* (esquerda) e recorrente (direita). Ambas as redes possuem taxas de entrada \mathbf{u} e de saída \mathbf{v} , e matriz de conexão sináptica feedforward \mathbf{W} . Além disso, a rede recorrente possui também uma matriz de conexão sináptica recorrente \mathbf{M} . **Fonte:** Autora

Essa equação pode ser escrita em termos das componentes dos vetores como

$$\tau_r \frac{dv_a}{dt} = -v_a + F\left(\sum_{b=1}^{N_u} W_{ab}u_b\right) \quad (25)$$

onde $\sum_b W_{ab}u_b = \mathbf{W} \cdot \mathbf{u}$, e $\mathbf{F}(\mathbf{W} \cdot \mathbf{u})$ representa o vetor com componentes $F(\sum_b W_{ab}u_b)$ para $a = 1, 2, \dots, N_v$.

A rede recorrente, ilustrada à direita da figura 9, também possui duas camadas de neurônios cujas taxas são dadas por \mathbf{u} e \mathbf{v} . Nesse caso, os neurônios da camada externa também possuem conexões entre si, e essas conexões são descritas pela matriz \mathbf{M} . O elemento $M_{aa'}$ representa a força da conexão sináptica da unidade a' até a unidade a . As taxas de disparo para essa rede são dadas por

$$\tau_r \frac{d\mathbf{v}}{dt} = -\mathbf{v} + \mathbf{F}(\mathbf{h} + \mathbf{M} \cdot \mathbf{v}) \quad (26)$$

onde $\mathbf{h} = \mathbf{W} \cdot \mathbf{u}$. Os neurônios são geralmente classificados como ou excitatórios, ou inibitórios. Portanto, é conveniente descrever os neurônios excitatórios e inibitórios separadamente.

3.2 Medidas de Redes

Nesta seção, será apresentado um breve resumo de teoria de grafos. Serão abordados alguns conceitos importantes, como o coeficiente de clustering e o comprimento de caminho médio, e esses conceitos serão utilizados na seção seguinte (3.3), na construção de modelos específicos de redes aleatórias.

Uma rede (objeto que consiste em um conjunto de vértices (ou nodos) e arestas (ou links)) é representada por um Grafo $G = (V, E)$, onde V é um conjunto de nodos e E é um conjunto de arestas [45]. Uma aresta é um subconjunto formado por dois elementos de V . Uma rede *direcionada* é uma rede na qual cada aresta possui uma direção (ou seja, vai de um vértice até outro), de forma que E é, nesse caso, um conjunto de pares *ordenados* de vértices. Uma rede é denominada *simples* se não possui loops (arestas que conectam um nodo a ele mesmo) e multi-arestas (quando há mais de uma aresta entre o mesmo par de nodos).

Considere um grafo constituído de N vértices indexados por $i = 1, \dots, N$, e m arestas. Uma rede direcionada pode ser representada matematicamente pela *matriz de adjacência* \mathbf{A} , cujos elementos A_{ij} são dados por

$$A_{ij} = \begin{cases} 1, & \text{se existe uma aresta de } j \text{ até } i \\ 0, & \text{caso contrário} \end{cases} \quad (27)$$

onde $i, j \in [1, N]$. Note que a direção das arestas é do segundo índice até o primeiro.

⁸ Para redes não-direcionadas, a primeira frase do lado direito da equação (27) é substituída por "se existe uma aresta entre i e j ".

O *grau* k_i de um vértice i é o número de arestas conectadas a ele. Para redes não-direcionadas, esse valor é $k_i = \sum_{j=1}^N A_{ij}$. Nesse caso, como cada aresta possui duas terminações, a soma dos graus dos N vértices do grafo é igual ao dobro do número de arestas m , ou seja, $\sum_{i=1}^N k_i = 2m$. Portanto, o *grau médio* de uma rede

⁸No caso de redes neurais, os índices i e j da matriz de adjacência (também chamada matriz de conexão) representam neurônios pós- e pré-sinápticos, respectivamente.

não-direcionada é

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i = \frac{2m}{N} \quad (28)$$

Para redes direcionadas, são definidos dois graus para cada vértice i : grau interno k_i^{in} e grau externo k_i^{out} , dados por

$$k_i^{in} = \sum_{j=1}^N A_{ij}, \quad k_i^{out} = \sum_{j=1}^N A_{ji} \quad (29)$$

Nesse caso, a soma de cada um dos graus dos N vértices do grafo é igual ao número de arestas, ou seja, $\sum_{i=1}^N k_i^{in} = \sum_{i=1}^N k_i^{out} = m$. Portanto, os graus médios $\langle k_{in} \rangle$ e $\langle k_{out} \rangle$ são iguais:

$$\langle k_{in} \rangle = \frac{1}{N} \sum_{i=1}^N k_i^{in} = \frac{1}{N} \sum_{i=1}^N k_i^{out} = \langle k_{out} \rangle = \frac{m}{N} \quad (30)$$

A notação $\langle k_{in} \rangle = \langle k_{out} \rangle \equiv \langle k \rangle$ é usada para especificar o grau médio de uma rede direcionada.

A *distribuição de grau* p_k é a probabilidade de um determinado nodo em uma rede possuir grau k , e para uma rede com N nodos pode ser escrita como

$$p_k = \frac{N_k}{N} \quad (31)$$

onde N_k é o número de nodos com grau k . O grau médio de uma rede pode ser escrito em termos de p_k como

$$\langle k \rangle = \sum_{k=0}^{\infty} k p_k \quad (32)$$

O *coeficiente de clustering* (ou coeficiente de agrupamento) é uma medida que quantifica a tendência de vértices de um grafo se agruparem [54]. Para um vértice i , o coeficiente de clustering local é definido como

$$C_i = \frac{\text{número de pares de vizinhos de } i \text{ que são conectados}}{\text{número de pares de vizinhos de } i} \quad (33)$$

C_i representa a probabilidade de que dois vértices conectados a i estejam conectados entre si. Como o número de pares de vizinhos de um vértice i com grau k_i é $k_i(k_i - 1)/2$, o coeficiente de clustering local é dado por

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \quad (34)$$

onde E_i é o número de conexões entre os k_i vizinhos do vértice i , e $C_i \equiv 0$ para nodos com grau $k_i = 0, 1$. O coeficiente de clustering médio para uma rede composta por N vértices é dado por

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^N C_i \quad (35)$$

Coefficiente de clustering global, também chamado *transitividade*, mede a fração de triângulos existentes em um grafo. Possíveis triângulos (ou seja, duas arestas conectadas ao mesmo vértice) são identificados como "triplets".

$$C = \frac{3 \times \text{número de triângulos}}{\text{número de triplets}} \quad (36)$$

O fator 3 no numerador é devido ao fato de que cada triângulo corresponde a 3 triplets.

O *caminho mínimo* (também chamado *distância*) $d_{i,j}$ entre dois nodos i, j de um grafo é o *caminho* (sequência de vértices conectados por uma sequência de arestas) com o menor número de arestas. O *diâmetro* de uma rede, d_{max} , é o maior caminho mínimo na rede. O *comprimento de caminho médio* $\langle d \rangle = \ell$ é definido como a distância média entre todos os pares de nodos de uma rede. Para uma rede direcionada com N nodos, esse comprimento é dado por

$$\langle d \rangle = \frac{1}{N(N-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^N d_{i,j} \quad (37)$$

Assortatividade é uma medida utilizada para quantificar a preferência dos nodos de uma rede em se conectar com outros nodos que possuem características semel-

hantes. Em geral, essa característica em questão é o grau dos nodos. O *grau restante* é o número de arestas que possui um vértice além da aresta que está sendo considerada (ou seja, é igual a uma unidade a menos do que o grau do vértice). Proposto por Newman [42, 43], o *coeficiente de assortatividade* ou *coeficiente de correlação de grau* é definido como o coeficiente de correlação de Pearson entre os graus de nodos conectados

$$r = \frac{1}{\sigma_q^2} \sum_{jk} jk(e_{jk} - q_j q_k) \quad (38)$$

onde

$$\sum_{jk} e_{jk} = 1, \quad \sum_j e_{jk} = q_k \quad (39)$$

e a distribuição normalizada q_k do grau restante é dada por

$$q_k = \frac{(k+1)p_{k+1}}{\sum_j j p_j} \quad (40)$$

e $\sigma_q^2 = \sum_k k^2 q_k - [\sum_k k q_k]^2$ é a variância da distribuição q_k .

O coeficiente de assortatividade varia entre -1 e 1: $r < 0$ corresponde a redes desassortativas, $r > 0$ a redes assortativas, e $r = 0$ a redes neutras.

As tabelas 1 e 2 mostram valores calculados das quantidades mencionadas nessa seção para algumas redes reais. Muitos tipos de sistemas podem ser representados por redes, e uma quantidade significativa dessas redes não possui estrutura regular. Grafos aleatórios têm sido extensivamente utilizados como modelos de diversos tipos de redes reais como a Internet, redes sociais e biológicas. Nas seções 3.3.1-3.3.3 são expostos 3 modelos importantes de redes aleatórias.

Rede	N	$\langle k \rangle$	ℓ	C	Referência
World Wide Web	153127	35.21	3.1	0.1078	[1]
Colaboração entre atores	225226	61	3.65	0.79	[54]
Colaboração autoral (matemática)	70975	3.9	9.5	0.59	[8]
Colaboração autoral (neurociência)	209293	11.5	6	0.76	[8]
Rede elétrica (EUA)	4941	2.67	18.7	0.08	[54]
Rede neural (<i>C. Elegans</i>)	282	14	2.65	0.28	[54]

Tabela 1: Tamanho, grau médio, comprimento de caminho médio e coeficiente de clustering calculados para alguns exemplos de redes reais.

Rede	N	r	Referência
World Wide Web	269504	-0.067	[2, 54]
Colaboração entre atores	449913	0.208	[4]
Colaboração autoral (matemática)	253339	0.120	[23]
Colaboração autoral (biologia)	1520251	0.127	[41]
Colaboração autoral (física)	52909	0.363	[41]
Rede elétrica (EUA)	4941	-0.003	[54]
Rede neural (<i>C. Elegans</i>)	307	-0.226	[54, 55]
Interação de proteínas	2115	-0.156	[31]
Rede metabólica	765	-0.240	[32]

Tabela 2: Coeficiente de assortatividade de redes reais.

3.3 Modelos de Redes Aleatórias

3.3.1 Erdős-Rényi

Os dois modelos de redes aleatórias mais frequentemente utilizados são $G(N, m)$ (introduzido em 1959 por Erdős e Rényi [19]) e $G(N, p)$ (introduzido por Gilbert [21]). O primeiro consiste em N nodos com m arestas aleatoriamente distribuídas. No segundo modelo, cada par dentre os N nodos indexados é conectado com probabilidade p . Devido à relevância do trabalho de Erdős e Rényi em teoria de grafos aleatórios, ambos os modelos são conhecidos como modelos de Erdős-Rényi.

No modelo $G(N, m)$, as arestas são colocadas em m dentre os $N(N - 1)/2$ possíveis pares de vértices do grafo, de forma que existe um total de $\binom{N(N-1)}{m}$ grafos com N vértices e m arestas. Nesse caso, a probabilidade de se obter um determinado grafo com m arestas e N nodos é $1/\binom{N(N-1)}{m}$.

No modelo $G(N, p)$, o número total de arestas é uma variável aleatória com valor esperado $\mathbb{E}(m) = p\frac{N(N-1)}{2}$. Se G_0 é um determinado grafo com N nodos e m arestas, a probabilidade de se obter G_0 pelo processo de construção de grafos mencionado anteriormente é $p^m(1 - p)^{\frac{N(N-1)}{2} - m}$.

Enquanto para o modelo $G(N, m)$ o grau médio de um vértice do grafo é dado por $\langle k \rangle = 2m/N$, para o modelo $G(N, p)$ esse valor é $\langle k \rangle = 2\mathbb{E}(m)/N = p(N - 1)$. Os dois modelos são equivalentes, pela lei dos grandes números, fazendo $p\binom{N}{2} = m$ no limite $N \rightarrow \infty$. Nesse trabalho, será utilizado o modelo $G(N, p)$ para construção de redes aleatórias as quais serão denominadas redes do tipo Erdős-Rényi (ER).

A probabilidade de que um nodo i tenha exatamente k arestas em uma rede aleatória construída de acordo com o modelo $G(N, p)$ com N nodos e probabilidade de conexão p é dada pela multiplicação de três termos: a probabilidade de que k arestas estejam presentes, p^k , a probabilidade de que as restantes possíveis arestas estejam ausentes, $(1 - p)^{N-1-k}$, e o número de maneiras de selecionar k arestas entre $N - 1$ possíveis arestas. Portanto, a distribuição de grau $p_k = P(k_i = k) = P(k)$ é

dada pela distribuição binomial com parâmetros $N - 1$ e p :

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k} \quad (41)$$

e o grau médio $\langle k \rangle$ é dado por

$$\langle k \rangle = \sum_{k=0}^{N-1} k p_k = (N-1)p \quad (42)$$

Para grandes valores de N , a distribuição binomial pode ser aproximada pela distribuição de Poisson,

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \quad (43)$$

onde, de acordo com (42), $\langle k \rangle = p(N-1) \simeq pN$. O valor esperado do número de nodos X_k com grau k é $\mathbb{E}(X_k) = NP(k_i = k)$. A figura 10 mostra a distribuição de grau de uma rede aleatória construída de acordo com o modelo $G(N, p)$ para uma rede de tamanho $N = 10000$ e probabilidade de conexão $p = 0.0015$.

O comprimento de caminho médio $\langle d \rangle \equiv \ell$ para redes do tipo ER (redes construídas de acordo com o modelo $G(N, p)$) varia com o logaritmo do tamanho da rede [3]:

$$\ell_{ER} \sim \frac{\ln N}{\ln \langle k \rangle} \quad (44)$$

No modelo $G(N, p)$, a probabilidade de que dois vizinhos de um determinado nodo sejam conectados entre si é a mesma probabilidade de que quaisquer dois nodos da rede sejam conectados. Portanto, o coeficiente de clustering para esse tipo de rede é dado por

$$C_{ER} = p \simeq \frac{\langle k \rangle}{N} \quad (45)$$

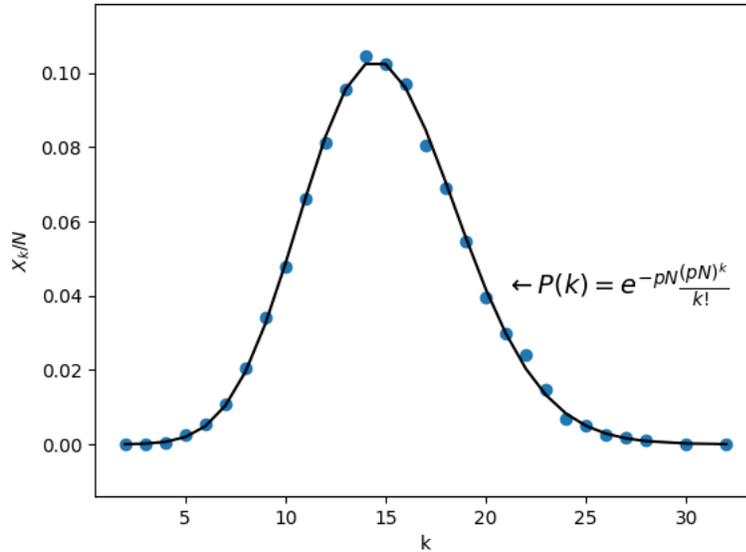


Figura 10: Distribuição de grau de um grafo aleatório construído de acordo com o modelo $G(N, p)$. Foi gerada uma rede não-direcionada com $N = 10000$ nodos e probabilidade de conexão $p = 0.0015$, e calculado o número de nodos X_k com grau k . O gráfico compara X_k/N com a distribuição de Poisson dada por (43) com $\langle k \rangle = pN$. **Fonte:** Autora

3.3.2 Pequeno Mundo

As redes aleatórias descritas na seção anterior são caracterizadas por valores pequenos de ambos comprimento de caminho médio ℓ e coeficiente de clustering C . Diversas redes reais apresentam, entretanto, elevados valores de coeficientes de clustering associados a pequenos ℓ . Watts e Strogatz [54] propuseram um modelo de redes que apresentam essas características, e cuja arquitetura está entre a arquitetura de uma rede regular e a de uma rede aleatória.

O processo de construção de redes do modelo Watts-Strogatz (WS) é o seguinte:

1. Começamos de uma rede de N nodos com estrutura regular e periódica, onde

cada nodo se conecta com os $2K$ nodos mais próximos (K para cada lado);

2. Cada aresta é rearranjada aleatoriamente com probabilidade β (de maneira que a formação de loops e multi-arestas não é permitida).

Para $\beta = 0$, a rede possui estrutura regular e o comprimento de caminho médio varia com o tamanho da rede, $\ell(N, 0) \sim N/4K$. Além disso, o coeficiente de clustering é $C(0) = \frac{3(K-1)}{2(2K-1)}$, e tende a $3/4$ no limite para grandes valores de K . Para $\beta = 1$, o comprimento de caminho médio varia com o logaritmo do número de nodos da rede, $\ell(N, 1) \sim \ln(N)/\ln(2K - 1)$, e coeficiente de clustering é $C(1) \simeq 2K/N$ [11, 44]. A figura 11 mostra exemplos de redes do tipo pequeno mundo construídas com diferentes valores de β .

Um resultado importante foi obtido por Newman, Moore e Watts [46]:

$$\ell = \frac{N}{K} f(NK\beta) \quad (46)$$

onde

$$f(x) = \frac{1}{2\sqrt{x^2 + 2x}} \tanh^{-1} \frac{x}{\sqrt{x^2 + 2x}} \quad (47)$$

A solução (46) é exata nos limites $N \rightarrow \infty$, com $NK\beta \rightarrow \infty$ ou $NK\beta \rightarrow 0$.

Barrat e Weigt [11] mostraram que o coeficiente de clustering de uma rede do tipo WS pode ser escrito como

$$C = \frac{3(K-1)}{2(2K-1)} (1-\beta)^3 \quad (48)$$

As figuras 12 e 13 mostram o comportamento dos coeficientes de clustering e dos comprimentos de caminho médio para redes de tamanho $N = 1000$ em função do parâmetro β . Além disso, a figura 12b também mostra esses valores para redes de tamanhos $N = 2000$ e $N = 5000$.

Um modelo similar ao modelo WS foi proposto por Newman e Watts [47]. Nesse

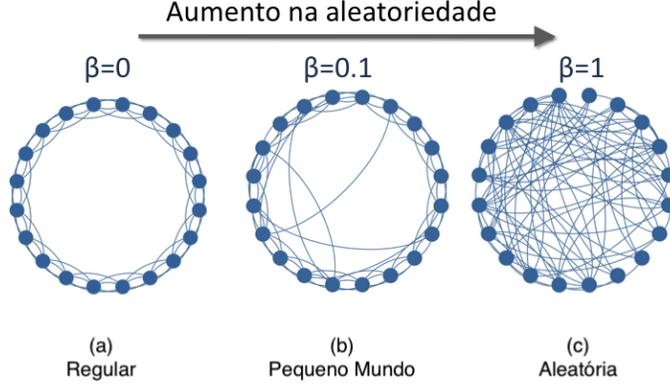


Figura 11: Modelo Watts-Strogatz, com parâmetros $N = 20$ e $K = 2$: comparação entre 3 diferentes valores de β . (a) Para $\beta = 0$, a rede possui estrutura regular onde cada nodo é conectado com os 4 nodos mais próximos (dois à direita e dois à esquerda). (b) Com probabilidade $\beta \in (0, 1)$, cada conexão é refeita (de um nodo $i = 1, \dots, N$ até um nodo j aleatoriamente selecionado). (c) Para $\beta = 1$, todas as conexões foram refeitas, e a rede passa a ter comportamento similar ao de uma rede aleatória do tipo Erdős-Rényi. **Fonte:** Autora

modelo, chamado Newman-Watts-Strogatz (NWS), para cada aresta da rede, uma aresta aleatória é *criada* (em vez de rearranjada) com probabilidade β . A figura 14 mostra exemplos de redes geradas de acordo com os dois modelos mencionados. Para uma rede do tipo NWS, o coeficiente de clustering é dado por [44]

$$C = \frac{3(K-1)}{2(2K-1) + 4K\beta(\beta+2)} \quad (49)$$

A distribuição de grau do modelo WS para $k \geq K/2$ segue [11]

$$p_k = \sum_{n=0}^{f(k,K)} \binom{K}{n} (1-\beta)^n \beta^{K-n} \frac{(\beta K)^{k-K-n}}{(k-K-n)!} e^{-\beta K} \quad (50)$$

onde $f(k, K) = \min(k - K, K)$. Essa distribuição é similar à distribuição do modelo ER, e possui um pico em $k = 2K$. A principal limitação dos modelos WS e NWS é o fato de que eles produzem distribuição de grau diferente da observada em redes reais.

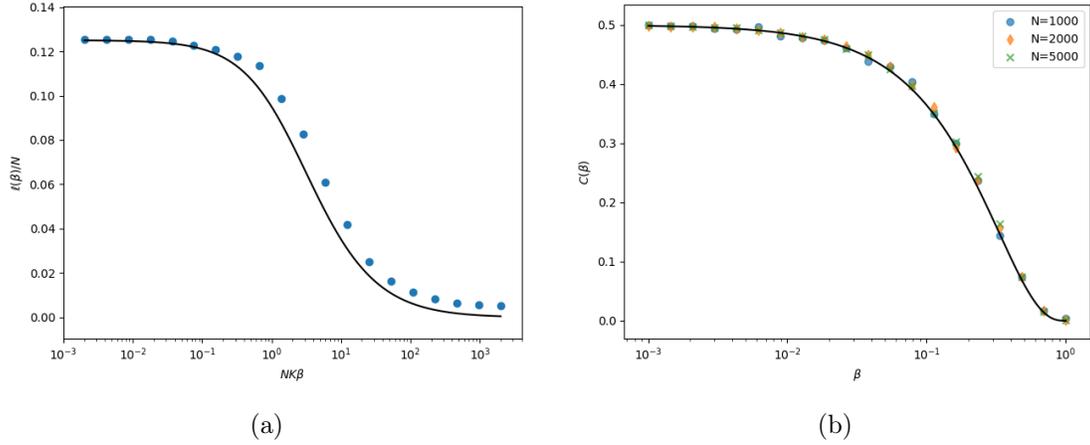


Figura 12: **(a)** Comprimento de caminho médio ℓ , normalizado pelo tamanho da rede, em função de $NK\beta$ para uma rede do tipo WS com parâmetros $K = 2$ e $N = 1000$. A curva é dada pela equação (46). **(b)** Coeficiente de clustering $C(\beta)$ em função de β para uma rede do tipo WS com $K = 2$ e $N = 1000, 2000, 5000$. A curva é dada pela equação (48). **Fonte:** Autora

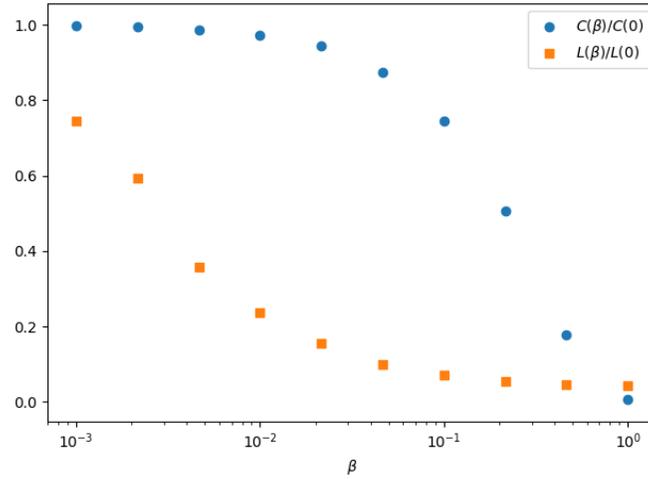
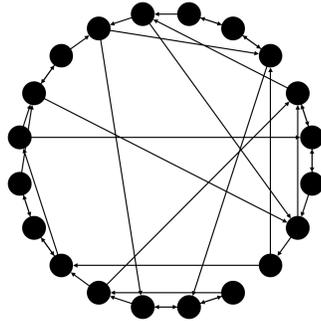
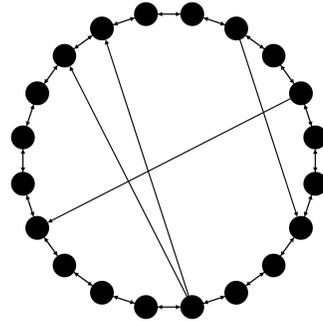


Figura 13: Comprimento de caminho médio $L(\beta) \equiv \ell(\beta)$ e coeficiente de clustering $C(\beta)$, normalizados pelos valores $L(0)$ e $C(0)$. Os dados mostrados no gráfico são a média entre 20 redes diferentes do tipo WS geradas pelo processo descrito no início desta seção, com parâmetros $N = 1000$ e $K = 2$. **Fonte:** Autora



(a) Modelo Watts-Strogatz



(b) Modelo Newman-Watts-Strogatz

Figura 14: Redes aleatórias do tipo pequeno mundo, com parâmetros $N = 20$, $K = 1$, $\beta = 0.3$ (Watts-Strogatz), $\beta = 0.1$ (Newman-Watts-Strogatz). **Fonte:** Autora

3.3.3 Livre de Escala

Muitas redes reais apresentam uma distribuição de grau que não segue distribuição de Poisson (43), apresentada na seção 3.3.1. Uma rede livre de escala é uma rede cuja distribuição de grau p_k segue uma lei de potência

$$p_k \sim k^{-\gamma} \quad (51)$$

onde γ é chamado expoente de grau.

Muitas redes reais apresentam duas características importantes que as diferem das redes aleatórias já mencionadas e têm papel importante na forma da distribuição de grau de uma rede:

1. Crescimento: diversas redes reais são o resultado de um processo contínuo de crescimento, ou seja, o número de nodos N da rede aumenta com o tempo. O tamanho da rede, portanto, não é fixo como nos modelos ER, WS e NWS.
2. Preferential-attachment: em alguns casos, os novos nodos adicionados à rede tendem a se conectar com os nodos que possuem um maior número de conexões.

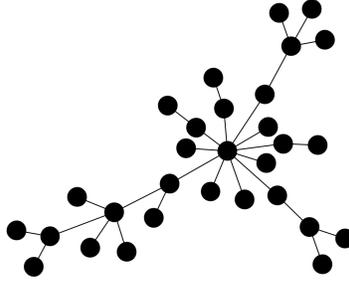


Figura 15: Exemplo de rede aleatória não-direcionada construída de acordo com o modelo de Barabási-Albert, com $N = 30$ e $m_0 = m = 1$. **Fonte:** Autora

Barabási propôs [9] em 1999 um modelo que apresenta as características citadas acima. Nesse modelo, conhecido como modelo de Barabási-Albert (BA), m_0 nodos são conectados arbitrariamente, com no mínimo uma conexão para cada nodo. A figura 15 mostra um exemplo de uma rede de tamanho $N = 30$ construída de acordo com o modelo proposto por Barabási. A cada intervalo de tempo, um novo nodo é conectado por $m \leq m_0$ arestas a m nodos existentes, com uma probabilidade que depende do grau desses nodos. A probabilidade $\Pi(k)$ de que a conexão seja feita com um nodo i é

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \quad (52)$$

onde a soma no denominador é feita sobre todos os nodos existentes na rede no momento em que o novo nodo é adicionado.

Após t passos de tempo, a rede gerada possui $N = m_0 + t$ nodos e $m_0 + mt$ arestas. O processo descrito acima resulta em uma rede livre de escala cuja distribuição de grau obedece (51) com expoente de grau $\gamma = 3$. Ou seja, o modelo de Barabási-Albert possui distribuição de grau dada por

$$p_k \sim k^{-3} \quad (53)$$

e essa distribuição não depende do valor de m nem do tamanho da rede. A figura 16 mostra a distribuição de grau de uma rede do modelo BA de tamanho $N = 100000$.

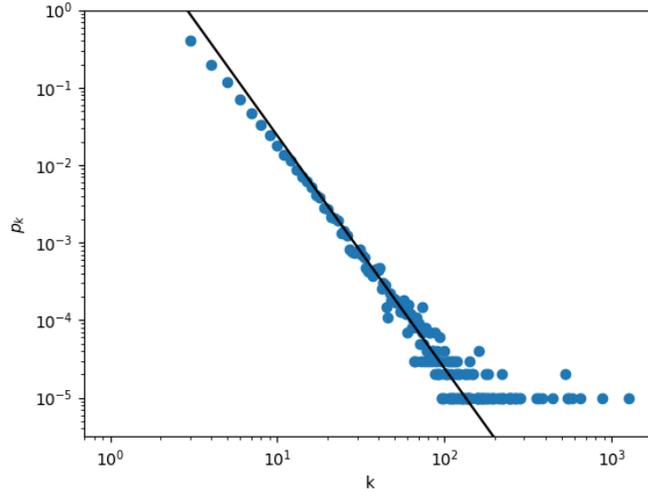


Figura 16: Distribuição de grau de uma rede do modelo BA com $N = 100000$ e $m_0 = m = 3$. A reta possui inclinação $\gamma = 3$. **Fonte:** Autora

O valor esperado do coeficiente de clustering do modelo BA segue [15]

$$C_{BA} \sim \frac{(\ln N)^2}{N} \quad (54)$$

O comprimento de caminho médio desse modelo é dado por

$$\ell_{BA} \sim \frac{\ln N}{\ln \ln N} \quad (55)$$

A equação (55) mostra que ℓ_{BA} cresce mais lentamente do que $\ln N$, de forma que os comprimentos médios para redes do modelo BA são menores do que os comprimentos observados em redes de mesmo tamanho do modelo ER, para grandes valores de N . A figura 17a mostra uma comparação entre os coeficientes de clustering de redes aleatórias do modelo BA com redes do tipo ER, e a figura 17b mostra a dependência do comprimento de caminho médio de redes do modelo BA com o tamanho da rede. A figura 18 mostra exemplos de redes construídas de acordo com os quatro modelos analisados.

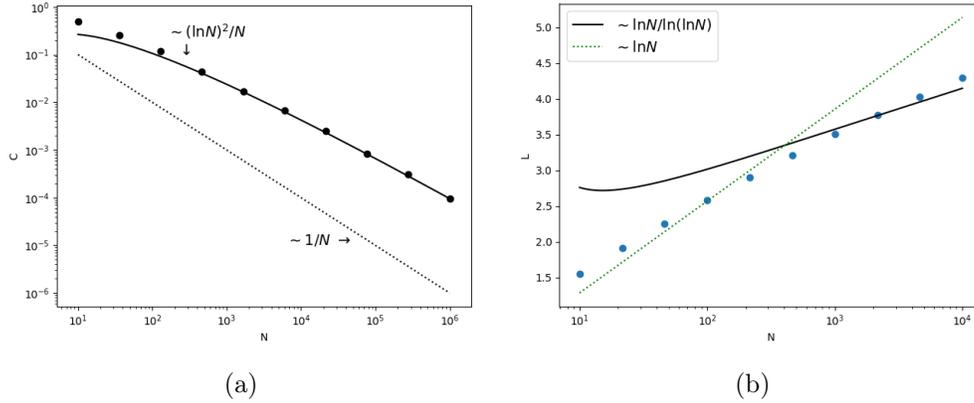


Figura 17: **(a)** Coeficiente de clustering em função do número de nodos da rede do modelo BA (média de 10 redes com $m = m_0 = 4$). A linha hachurada mostra o comportamento para redes do tipo ER ($\sim 1/N$). A linha sólida é dada por $C = \frac{m}{4} \frac{(\ln N)^2}{N}$ [34]. **(b)** Comprimento de caminho médio em função do número de nodos da rede do modelo BA. A linha hachurada mostra o comportamento para redes do tipo ER, e a linha sólida é dada pela equação (55). O gráfico mostra a média dos valores para 10 redes com $m = m_0 = 3$. **Fonte:** Autora

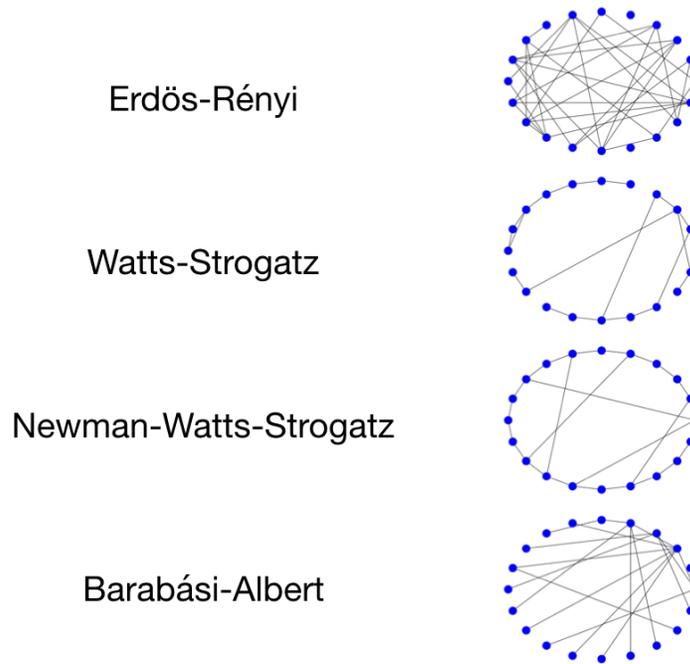


Figura 18: 4 modelos de redes aleatórias: ER, WS, NWS e BA. **Fonte:** Autora

3.4 Medidas de similaridade

O processo de codificação/decodificação analisado nesse trabalho envolve três conjuntos de dados: a sequência de entrada determinada pelo buffer, o padrão espacial gerado pela rede recorrente, e a sequência de saída obtida da decodificação. Quando uma rede aleatória é utilizada como rede recorrente, é possível que ocorra a situação em que diferentes sequências de entrada dão origem a um mesmo padrão espacial - e se espera que, quanto mais eficiente for a rede, com menos frequência isso irá ocorrer. Dessa forma, algumas medidas de similaridades conhecidas podem ser úteis para analisar as sequências e padrões gerados pela rede em termos do quão semelhantes eles são entre si. Além disso, essas medidas também podem ser utilizadas na definição de uma medida que quantifica a eficiência da rede no processo de codificação.

Semelhança de cosseno (cosine similarity) é uma medida de similaridade entre dois vetores \mathbf{u}, \mathbf{v} não nulos definidos em um espaço vetorial com produto interno. Essa medida considera apenas a orientação dos vetores, e é definida por

$$similarity = \cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} = \frac{\sum_{i=1}^n u_i v_i}{\sqrt{\sum_{i=1}^n u_i^2} \sqrt{\sum_{i=1}^n v_i^2}} \quad (56)$$

onde u_i e v_i , $i = 1, \dots, n$ são as componentes dos vetores \mathbf{u} e \mathbf{v} , respectivamente. Essa medida - a qual chamaremos *redundância* - será usada neste trabalho para comparação entre os padrões espaciais. Um valor de redundância igual a 1 corresponde a padrões iguais.

Distância de Kendall tau é uma medida que indica o número de discordâncias emparelhadas entre duas listas. Como é equivalente ao número de trocas realizadas pelo algoritmo bubble sort para transformar uma lista na outra, essa medida também é chamada distância de bubble sort. A distância de Kendall tau $K(\tau_1, \tau_2)$ varia entre 0 (se as listas τ_1 e τ_2 são idênticas) e $n(n-1)/2$ (se uma lista é o inverso da outra), onde n é o tamanho da lista. Um valor nulo da distância de Kendall tau normalizada implica listas iguais, e $K(\tau_1, \tau_2) = 1$ implica discordância máxima entre as listas

consideradas.

Distância de Levenshtein (também conhecida como distância de edição) é uma medida de similaridade utilizada em teoria da informação, linguística e ciência da computação para medir a diferença entre duas sequências. Sua vantagem em relação às outras medidas é o fato de que se pode comparar duas listas de tamanhos diferentes. A distância é dada pelo número mínimo de edições (inserções, exclusões e substituições) necessárias para transformar uma sequência na outra, e possui como limite inferior a diferença entre os tamanhos das sequências, e como limite superior o tamanho da maior sequência. Essa medida de similaridade será usada neste trabalho para comparações entre as sequências de entrada. O procedimento para calcular a distância de Levenshtein $Lev(s, t)$ entre duas sequências s e t de tamanhos m e n , respectivamente, é o seguinte:

1. Construir uma matriz nula D de tamanho $(m + 1) \times (n + 1)$;
2. Atribuir à primeira linha da matriz D os elementos $(0, 1, \dots, n)$ e à primeira coluna os elementos $(0, 1, \dots, m)$;
3. Para j de 1 a n , e i de 1 a m : se $s[i] = t[j]$ (ou seja, se o i -ésimo elemento da sequência s é igual ao j -ésimo elemento da sequência t), atribui-se o valor 0 à constante "*custo*". Caso contrário, atribui-se o valor 1 à constante "*custo*".
4. O elemento $D[i, j]$ da matriz é o mínimo entre:
 - $D[i - 1, j] + 1$ (o elemento acima mais um)
 - $D[i, j - 1] + 1$ (o elemento à esquerda mais um)
 - $D[i - 1, j - 1] + custo$ (o elemento da diagonal superior esquerda mais o custo)
5. A distância de Levenshtein é dada pelo elemento $D[m, n]$ da matriz.

A fim de determinar qual modelo de rede aleatória é o mais adequado para a codificação de padrões temporais em padrões espaciais, torna-se necessária a introdução de uma medida que considera o erro na codificação, ou seja, a diferença entre

a sequência de entrada e a sequência obtida na decodificação do padrão espacial gerado pela rede. Nesse trabalho, o erro de codificação é definido como a distância de Levenshtein entre a sequência de entrada e a sequência obtida da decodificação descrita acima:

$$erro = Lev(input, output) \quad (57)$$

onde *input* e *output* são as sequências de entrada e saída, respectivamente.

4 Modelo computacional para codificação de sequências temporais

O modelo proposto no presente trabalho consiste em uma estrutura neural composta por dois buffers de memória de trabalho e uma rede atratora recorrente de memória associativa. Estudaremos em um primeiro momento um modelo estereotipado com neurônios binários. A figura 19 mostra um esquema da arquitetura do modelo de codificação proposto. Na seção seguinte (4.1), será estudado o modelo com neurônios binários. Serão consideradas duas formas de conectividade na rede recorrente: na primeira (seção 4.1.1), a matriz de conexão da rede é construída de forma que todos os pares de elementos da sequência de entrada são comparados; na segunda (seção 4.1.2), a conexão dos neurônios da rede recorrente é feita de maneira aleatória. Na seção 4.2, a rede recorrente será construída com neurônios do tipo taxa de disparo, e a decodificação será feita com neurônios integra-e-dispara.

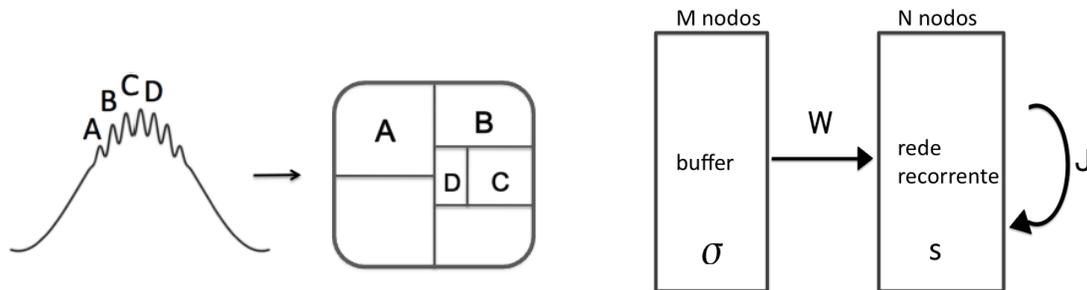


Figura 19: Uma sequência "ABCD" é recebida pelo buffer na forma de uma oscilação do tipo tetagama, e codificada em um padrão espacial no qual a área de cada polígono está relacionada à ordem dos elementos na sequência de entrada (esquerda). Esquema da arquitetura neural proposta para tradução de sequências temporais em padrões espaciais (direita). **Fonte:** Autora

4.1 Modelo com Neurônios Binários

4.1.1 Rede Ordenada

A ordem temporal de uma sequência de entrada pode ser facilmente codificada em termos de ordem de subsequências menores dos elementos da sequência. Se "ABCD" é uma sequência de entrada, então as seguintes observações independentes são verdadeiras: $t_A < t_B$, $t_B < t_C$ e $t_C < t_D$, onde t_X é o tempo de ocorrência do elemento X . Um mecanismo que compara a ordem de dois elementos de uma sequência pode ser construído utilizando dois neurônios (ou grupo de neurônios) que se inibem mutuamente. Considere duas células da rede recorrente sensíveis aos elementos A e B denominadas 1 e 2, respectivamente. Elas se inibem mutuamente, de maneira que se uma célula está ativa, a outra não pode ser ativada. Portanto, a determinação de qual célula estará ativa depende somente da ordem em que os padrões A e B são apresentados para a rede. A ativação da célula 1 implica uma das seguintes situações: ou A precede B na sequência de entrada, ou dentre esses dois elementos, somente A está presente na sequência de entrada.

Vamos considerar uma rede que compara, dois a dois, todas as possíveis combinações de elementos de uma sequência de entrada. Se M é o número de elementos do buffer, representados por $\sigma_i, i = 0, \dots, M - 1$, a rede recorrente deverá ser composta por $N = M^2$ células, com M delas sensíveis a cada um dos σ . Os neurônios da rede recorrente possuem seus estados no tempo t representados por $s_i(t), i = 0, \dots, N - 1$. Os neurônios que recebem a informação vinda do buffer são células excitatórias piramidais. No entanto aqui, para simplificar, não modelaremos explicitamente os neurônios que mediarão a inibição entre estas células, considerando, dessa forma, que as células principais se inibem diretamente.

Os neurônios σ conectam-se com os neurônios s via matriz sináptica W , que é $N \times M$, e os neurônios da rede recorrente conectam-se entre si via matrix J , de dimensão $N \times N$. A figura 20 mostra uma representação da arquitetura da rede para

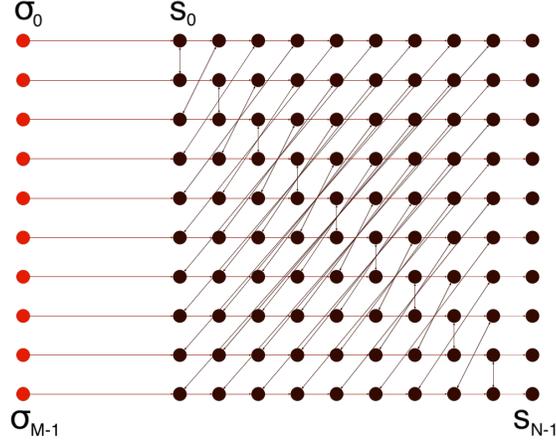


Figura 20: Arquitetura da rede ordenada, com $M = 10$. A primeira coluna à esquerda é composta pelos neurônios que representam os elementos de uma sequência de entrada, e os $N = 10^2$ neurônios da rede recorrente estão representados a partir da segunda coluna. **Fonte:** Autora

$M = 10$. A matriz de conexão W , para um valor de M arbitrário, é dada por

$$W_{ij} = \begin{cases} 1, & \text{se } j \cdot M \leq i \leq (j + 1) \cdot M - 1 \\ 0, & \text{caso contrário} \end{cases} \quad (58)$$

onde $0 \leq j \leq M - 1$ e $0 \leq i \leq N - 1$. A matriz W cujos elementos são dados por (58) estabelece conexões entre os neurônios σ e os neurônios s de forma que cada um dos neurônios do buffer se conecta com M neurônios diferentes da rede recorrente. O primeiro neurônio do buffer, σ_0 , se conecta com os neurônios s_k da rede recorrente com índices $k \in [0, M - 1]$. O segundo neurônio, σ_1 , se conecta com os neurônios s_k com índices $k \in [M, (2M) - 1]$, e assim sucessivamente.

A matriz de conexão J para essa mesma rede, com M inteiro arbitrário, pode ser construída definindo-se uma matriz A de tamanho $M \times M$, cujos elementos da i -ésima linha e j -ésima coluna são dados pelos índices $x = 0, 1, \dots, N - 1$ das células s_x da rede recorrente:

$$A_{ij} = i \cdot M + j \quad (59)$$

onde $0 \leq i, j \leq M - 1$. A matriz J pode ser escrita como

$$J_{kl} = \begin{cases} 1, & \text{se } \begin{cases} k = A_{ij}, l = A_{j+1,i}, \text{ ou} \\ l = A_{ij}, k = A_{j+1,i} \end{cases}, & A_{j+1,i} > A_{i,j}, i \neq (j+1), j < (M-1) \\ 0, & \text{caso contrário} \end{cases} \quad (60)$$

onde $0 \leq k, l \leq N - 1$, e $N = M^2$. Essa rede - a qual chamaremos *rede ordenada* -, construída com as matrizes de conexão W e J descritas acima, possibilita a comparação de todos os elementos da sequência de entrada, dois a dois.

Exemplo 4.1. Para $M = 3$, as matrizes de conexão W e J dadas por (58) e (60) são

$$W = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}; \quad J = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

As células σ do buffer são acionadas em uma dada sequência. Ou seja, cada célula é ativada em um tempo discreto, de forma que podemos dizer que

$$\sigma_i(t) = \delta(i, o_t) \quad (61)$$

onde $\delta(x, y)$ é a função Delta de Kronecker e o vetor $\mathbf{o} = (o_1, o_2, \dots, o_L)$, com $L \leq M$, é uma dada sequência de ativação dos neurônios σ .

A dinâmica dos neurônios binários presentes na rede recorrente é dada por

$$s_i(t) = \begin{cases} 1 & \text{se } s_i(t-1) = 0 \text{ e } W_{ij}\sigma_j(t-1) > 0 \\ 1 & \text{se } s_i(t-1) = 1 \\ 0 & \text{se } s_i(t-1) = 0 \text{ e } W_{ij}\sigma_j(t-1) = 0 \\ -1 & \text{se } s_i(t-1) = 0 \text{ e } W_{ij}\sigma_j(t-1) = 0 \text{ e } \sum_{j \neq i} J_{ij}s_j(t-1) < 0 \\ -1 & \text{se } s_i(t-1) = -1 \end{cases} \quad (62)$$

podendo então ter três estados: ativo (1), quiescente (0) e inibido (-1). Como neurônios inibidos são também quiescentes para análises da informação mútua entre padrões, consideraremos os neurônios inibidos como quiescentes.

Quando a rede recebe uma sequência temporal de padrões, ela faz a conversão em um padrão espacial da seguinte maneira: o primeiro elemento da sequência de entrada ativa todas as M células com as quais possui ligação; o segundo elemento ativa $M-1$ células, e assim sucessivamente. A decodificação pode ser feita a partir da matriz de conexão W , que indica a quais neurônios da rede recorrente os elementos da sequência de entrada estão conectados, e dos estados finais dos neurônios da rede recorrente. Portanto, o padrão espacial, determinado pelo estado final dos neurônios da rede recorrente, informa quais padrões estão presentes na sequência de entrada, e a ordem em que eles aparecem.

4.1.2 Rede Aleatória

Consideraremos novamente que as células do buffer são representadas por $\sigma_i(t)$ com $i = 0, \dots, M-1$, e os neurônios da rede recorrente por $s_i(t)$ com $i = 0, \dots, N-1$. Os neurônios σ conectam-se com os neurônios s via matriz sináptica W , que é $N \times M$. Neste trabalho, vamos considerar duas formas de construção da matriz W . A primeira forma consiste em conexões binárias com probabilidade q de ocorrer, ou seja,

$$P(W_{ij}) = q \delta(W_{ij} - 1) + (1 - q) \delta(W_{ij}) \quad (63)$$

Nesse caso, cada neurônio do buffer se conecta com os neurônios da rede recorrente com probabilidade q . Portanto, a quantidade de conexões de cada neurônio σ varia. Além disso, podem ocorrer sobreposições de conexões (dois neurônios σ podem se conectar com o mesmo neurônio s). Chamaremos uma matriz W construída dessa forma de *matriz de conexão aleatória*.

Na segunda forma de construir a matriz W , o número de conexões é fixo, de modo que cada um dos M neurônios σ se conecta com exatamente N/M neurônios s . Nesse caso, não há sobreposições de conexões, e os elementos da matriz W são dados por

$$W_{ij} = \begin{cases} 1, & \text{se } i \in [j \frac{N}{M}, (j+1) \frac{N}{M}) \\ 0, & \text{caso contrário} \end{cases} \quad (64)$$

onde os índices i e j (que representam os neurônios pós- e pré-sinápticos) variam de 0 a $N-1$ e de 0 a $M-1$, respectivamente.

Devido ao fato de que os índices dos nodos em redes do tipo pequeno mundo determinam sua posição na rede, faz-se necessário utilizar uma outra ordem de índices na construção da matriz de conexão W . Essa configuração pode ser obtida da seguinte maneira: primeiro, os neurônios s_i são ordenados aleatoriamente (a ordem é uma permutação dos índices $i \in [0, N-1]$). Depois, conecta-se o primeiro neurônio do buffer com os N/M primeiros neurônios s dessa sequência, o segundo com os N/M seguintes, e assim sucessivamente. Seja $n = (n_0, n_1, n_2, \dots, n_{N-1})$ uma permutação dos índices i . Os elementos de W serão dados por

$$W_{n_i j} = \begin{cases} 1, & \text{se } i \in [j \frac{N}{M}, (j+1) \frac{N}{M}) \\ 0, & \text{caso contrário} \end{cases} \quad (65)$$

No presente trabalho, uma matriz construída de acordo com (65) será chamada de *matriz de conexão seletiva*.

Os neurônios da rede recorrente conectam-se entre si via matrix J , que é $N \times N$. Essas conexões são também binárias, mas negativas (no caso de conexões inibitórias). Se a rede recorrente é do tipo Erdős-Rényi (seção 3.3.1), as conexões ocorrem com

uma probabilidade p , que em princípio é diferente de q , ou seja

$$P(J_{ij}) = p \delta(J_{ij} + 1) + (1 - p) \delta(J_{ij}) \quad (66)$$

A dinâmica dos neurônios da rede recorrente é dada pela equação (62), que é válida para o caso em que todas as conexões são inibitórias. Se as conexões da rede recorrente podem ser inibitórias ou excitatórias, o neurônio s_i pode ser ativado ou inibido no tempo t , de acordo com os sinais dos termos $s_j(t - 1)$ e $\sum_{j \neq i} J_{ij}$, de modo que se o neurônio s_j está ativo no tempo $t - 1$, então o neurônio s_i ficará ativo no tempo t se há uma conexão excitatória de s_j até s_i , e inativo, se a conexão é inibitória.

Na seção anterior, vimos que quando a rede recorrente é construída de forma que todos os possíveis pares de elementos da sequência de entrada são comparados, o primeiro elemento ativa todas as M células da rede recorrente com as quais possui ligação; o segundo elemento ativa $M - 1$ células, e assim sucessivamente. Entretanto, quando a rede recorrente possui arquitetura aleatória, isso não ocorre. Nesse caso, se a matriz de conexão W for seletiva (ou seja, construída de acordo com (65)), o primeiro elemento da sequência de entrada irá ativar M neurônios da rede recorrente, mas o elemento seguinte não irá necessariamente ativar $M - 1$ neurônios, podendo ativar um número menor de células. Similarmente, os neurônios seguintes poderão ativar um número de neurônios que não é fixo, e pode variar entre 0 e $M - 1$ neurônios (se as conexões da rede recorrente são todas inibitórias). Portanto, a sequência de saída dada pela decodificação do padrão espacial poderá ser diferente da sequência de entrada. Isso implica que, no caso de redes recorrentes com arquitetura aleatória, pode haver um erro não nulo no processo de codificação (seção 3.4). Referente à arquitetura da rede recorrente, serão considerados neste trabalho os seguintes modelos de redes aleatórias: Erdős-Rényi, pequeno mundo e livre de escala, abordados na seção 3.3.

4.2 Modelo com Neurônios Taxa de Disparo

Consideremos o mecanismo descrito anteriormente para a comparação da ordem de elementos de uma sequência, porém com a inclusão de interneurônios inibitórios para mediar a inibição entre os neurônios (ou grupo de neurônios) principais, como representado na figura 21. Os grupos de neurônios (*neural pool*) serão representados por um modelo taxa de disparo, no qual a atividade representa a atividade coletiva média dos neurônios do grupo. Interações entre neurônios dentro de um mesmo grupo são representadas por auto-acoplamento, ou seja, conexões de feedback de uma unidade para ela mesma.

A taxa de disparo do grupo de neurônios que representa o estímulo A e que compete com o estímulo B (o grupo de neurônios AB) é dada por

$$\tau \frac{d}{dt} r_{AB}(t) = -r_{AB}(t) + F[W_e r_{AB}(t) - W_i r_{BA}^i(t) + I_A(t) - \theta] \quad (67)$$

onde $r_{BA}^i(t)$ é a taxa de disparo da inibição que vem do grupo de neurônios BA

$$\tau \frac{d}{dt} r_{BA}^i(t) = -r_{BA}(t) + F[W_{ei} r_{BA}(t) - \theta] \quad (68)$$

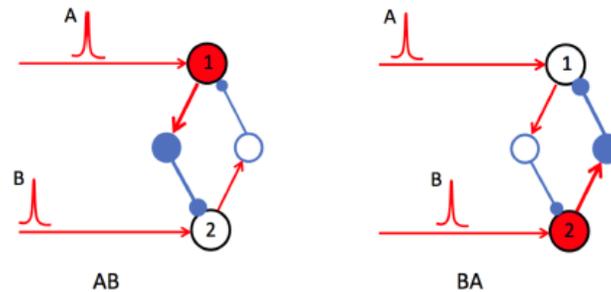


Figura 21: Dois grupos de neurônios biestáveis 1 e 2 são mutuamente inibidos por grupos de interneurônios inibitórios. A ordem dos estímulos de entrada (representados como A e B) determina qual dos grupos de células permanecerá ativo. **Fonte:** IDIART, 2019 ⁹

⁹Imagem utilizada com permissão

com a função de transferência F dada por

$$F[r] = \begin{cases} 0, & \text{se } r < 0 \\ r, & \text{se } 0 < r < r_{max} \\ r_{max}, & \text{se } r > r_{max} \end{cases} \quad (69)$$

A função de entrada é $I_A(t) = I_0 \cdot I(t - t_A)$, onde

$$I(t) = H(t)(e^{-\frac{t}{\tau_1}} - e^{-\frac{t}{\tau_2}}) \quad (70)$$

A decodificação é feita utilizando um modelo integra-e-dispara com cada neurônio ligado a um padrão de acordo com

$$\tau_s \frac{d}{dt} \nu_A(t) = -\nu_A(t) + \alpha \tanh(h_A(t)) + \beta \sin(\omega t) - \nu_A^i(t) \quad (71)$$

$$\text{if } \nu_A(t) > \nu_{thres} \quad \nu_A(t) \leftarrow \nu_{rest}$$

onde $h_A(t) = \sum_i r_{Ai}(t)$ é a soma de todas as taxas de disparos dos grupos de neurônios ativados pelo padrão A. Existe uma inibição de feedback $\nu_A^i(t)$ seguindo cada potencial de ação que previne que a mesma célula dispare novamente no mesmo período teta, dada por

$$\tau_{si} \frac{d}{dt} \nu_A^i(t) = -\nu_A^i(t) + \nu_0^i \sum_k \delta(t - t_A^k) \quad (72)$$

onde t_A^k é o tempo em que ocorre o k -ésimo disparo no neurônio sensível ao padrão A. Os parâmetros usados na simulação, como quantidades adimensionais, são $\tau = 0.1$, $W_e = 5.0$, $W_i = -0.095$, $W_{ei} = 4.0$, $I_0 = 6.0$, $\theta = 1.0$, $r_{max} = 10.0$, $\tau_1 = 0.03$ e $\tau_2 = 0.08$; $\tau_s = 1.0$, $\alpha = 3.0$, $\beta = 10.0$, $\omega = 6\pi$, $\nu_{rest} = 0$ e $\nu_{thres} = 10.0$; $\tau_s = 10.0$ e $\nu_0^i = -10.0$.

Consideramos que cada um dos grupos de neurônios presentes no buffer de memória pode estar conectado a um determinado número de regiões da rede recorrente. Essas regiões são constituídas de grupos de neurônios biestáveis. Se o padrão A está presente na sequência de entrada determinada pelo buffer, então os neurônios

das regiões da rede recorrente que possuem conexão com ele irão disparar caso não tenham sido inibidos anteriormente devido ao disparo de neurônios da rede recorrente que estão conectados a outro elemento do buffer. Dessa forma, cada elemento da sequência de entrada possui um determinado grupo de "alvos" na rede recorrente (neurônios com os quais estão conectados), e devido a inibição mútua, o número de alvos que um elemento irá ativar depende da posição dele na sequência de entrada.

Se a arquitetura da rede recorrente é tal que todos os possíveis pares de elementos da sequência de entrada possuem sua ordem comparada, ela é capaz de fazer, utilizando o mecanismo descrito, a tradução da sequência em um padrão espacial. Nesse caso, a sequência de saída obtida do processo de decodificação descrito acima será sempre idêntica à sequência de entrada. Entretanto, conforme mencionado na seção anterior, essa situação não ocorre no caso de redes recorrentes aleatórias.

Quando as conexões da rede recorrente são construídas de maneira aleatória, existe a possibilidade de que nem todos os pares de elementos da sequência de entrada sejam comparados. Portanto, nesse caso é possível que o processo de codificação leve a uma sequência de saída (obtida da decodificação do padrão espacial) diferente da sequência de entrada, ou seja, pode haver um erro não nulo no processo de codificação (seção 3.4). É esperado, entretanto, que quanto maior o tamanho da rede recorrente, mais eficiente se torne o processo de codificação.

5 Resultados

5.1 Modelo com Neurônios Binários

Serão analisados na presente seção os resultados obtidos para o modelo com neurônios binários cuja dinâmica é dada por (62). Os resultados serão divididos em três partes: na seção 5.1.1, serão estudados os perfis de ativação dos neurônios da rede recorrente, ou seja, a forma do padrão espacial obtido da codificação da sequência temporal recebida pela rede. Na seção 5.1.2, estudaremos a influência da arquitetura da rede recorrente na eficiência da codificação, considerando os quatro modelos de redes aleatórias referidos na seção 3.3. Na seção 5.1.3, será analisada a relação entre o erro na codificação e o tamanho da rede recorrente.

5.1.1 Perfil de ativação dos neurônios da rede recorrente

Uma sequência temporal é determinada pela ordem de ativação dos neurônios σ , de maneira que um padrão de entrada é dado por um vetor $\mathbf{o} = (o_1, o_2, \dots, o_L)$, onde $L \leq M$ é o número de neurônios ativados dentre os M neurônios do buffer. A dinâmica da rede recorrente gera um padrão $(s_0(t_L), s_1(t_L), \dots, s_{N-1}(t_L))$, onde t_L é o tempo de ativação do L -ésimo neurônio σ . A decodificação é feita utilizando o fato de que o primeiro elemento da sequência de entrada ativa um maior número de células da rede recorrente do que o segundo, e assim sucessivamente. O processo pode ser realizado da seguinte maneira: a partir da matriz de conexão W , soma-se o número de neurônios ativos s_i , $i = 0, \dots, N - 1$ conectados a σ_j (ou seja, tais que $W_{ij} > 0$), para cada $j \in [0, M - 1]$. Depois, coloca-se em ordem decrescente esse número de neurônios ativos. Os índices referentes a cada um desses valores determina a ordem dos elementos na sequência decodificada. Por exemplo, considerando uma

rede com dinâmica dada por (62), a soma dos neurônios ativos s_i ligados a σ_j é dada por $S(j) = \sum_{i=0}^{N-1} W_{ij}s_i$. Se $S(0) > S(2) > S(1)$, então a decodificação resulta na sequência $(\sigma_0, \sigma_2, \sigma_1)$. Se essa sequência é diferente da sequência de entrada, inferimos que há um erro não nulo na codificação.

No caso em que a rede recorrente possui arquitetura que compara todas as possíveis combinações de elementos do buffer (seção 4.1.1), a sequência do número de neurônios s ativos referentes a cada neurônio j do buffer terá a forma $(M, M - 1, M - 2, \dots, 1)$ se o número de elementos na sequência de entrada é igual ao número de neurônios no buffer ($L = M$). Por outro lado, se $L < M$, então a sequência mencionada terá a forma $(M, M - 1, M - 2, \dots, M - (L - 1), 0, \dots, 0)$, onde a quantidade de zeros (referentes aos elementos que não estão presentes na sequência de entrada) é dada por $M - L$. Devido à arquitetura das redes, isso não ocorre para o caso em que a rede recorrente possui estrutura aleatória (seção 4.1.2).

Na construção da rede recorrente com arquitetura aleatória, consideramos duas formas da matriz de conexão W . Na primeira forma, chamada matriz W aleatória e dada pela equação (63), podem haver sobreposições de conexões. Na segunda forma, chamada matriz W seletiva e dada por (65), essas sobreposições não ocorrem, conforme explicado na seção 4.1.2. Essa diferença entre as duas construções da matriz em questão resulta em diferentes comportamentos quanto ao perfil de ativação dos neurônios da rede recorrente.

No caso da matriz W aleatória, existe a possibilidade de neurônios σ do buffer que não estão presentes na sequência de entrada (quando $L < M$) estarem presentes na sequência de saída gerada pela rede. Por exemplo, para uma sequência de entrada "ABC", uma rede aleatória com matriz de conexão W aleatória poderia gerar a sequência de saída "ABD", apesar de o padrão "D", presente no buffer, não fazer parte da sequência em questão. Por outro lado, essa situação não ocorre quando a matriz W for do tipo seletiva, devido ao fato de não existirem sobreposições de conexões (considerando o caso em que todas as conexões da rede recorrente são inibitórias).

Nos gráficos da figura 22, os neurônios do buffer estão representados no eixo horizontal, e a altura da barra representa o número de neurônios s conectados a

cada um dos σ que estão ativos ao final da dinâmica temporal. No caso da rede ordenada (figura 22a), o primeiro neurônio da sequência temporal ativa todos os 10 neurônios da rede recorrente conectados a ele. O segundo neurônio ativa 9, e assim sucessivamente, até o último neurônio presente na sequência de entrada. Se um neurônio σ_k do buffer não está presente na sequência de entrada, ele não é ligado, portanto nenhum dos neurônios da rede recorrente que estão conectados a ele é ligado. Assim, as barras de todos os neurônios σ_k com $k \geq 4$ têm altura zero, no caso da rede ordenada.

Uma situação semelhante ocorre quando a rede recorrente é aleatória, e a matriz W é do tipo seletiva (figura 22c): para $k \geq 4$, nenhum dos neurônios s que estão conectados a σ_k está ativo ao final da dinâmica temporal. Nesse caso, o primeiro neurônio da sequência ativa os 10 neurônios da rede recorrente a ele conectados (devido à estrutura da matriz de conexão W). Entretanto, os próximos neurônios do buffer presentes na sequência temporal não ativam necessariamente 9, 8 e 7 neurônios da rede recorrente, devido à estrutura matriz de conexão aleatória J . Nesse caso, existe a possibilidade de a sequência de saída gerada pela rede recorrente aleatória ser diferente da sequência de entrada.

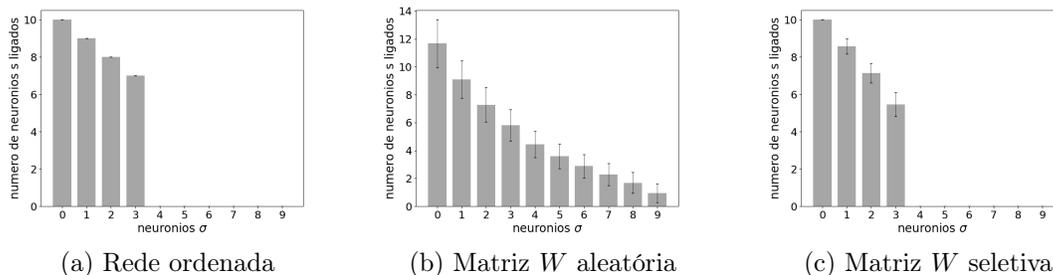
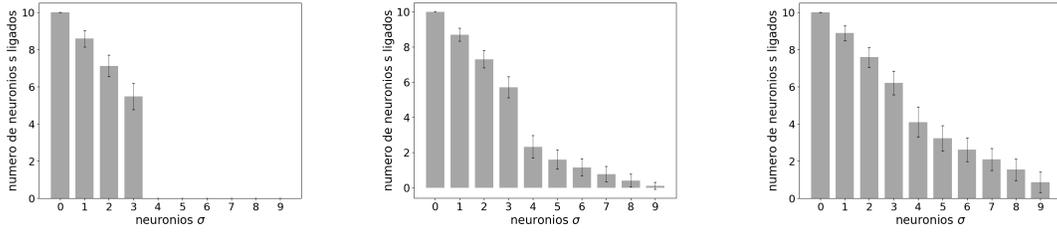


Figura 22: Número de neurônios da rede recorrente ativos ao final da dinâmica temporal (eixo vertical) conectados com cada um dos neurônios do buffer (eixo horizontal), para 3 diferentes modelos de redes: **(a)** Rede ordenada; **(b)** Rede recorrente aleatória do tipo ER com matriz de conexão W aleatória, com probabilidade de conexão $q = 0.1$; **(c)** Rede recorrente aleatória do tipo ER com matriz de conexão W seletiva. Foram utilizados os seguintes parâmetros: $L = 4$, $M = 10$, $N = 100$, $p = 0.02$ (seção 4.1.2), e todas as conexões das redes recorrentes são inibitórias. Para as redes aleatórias, foram geradas 400 redes e consideradas 20 sequências para cada uma delas. **Fonte:** Autora

Se a matriz W é do tipo aleatória (figura 22b), a situação descrita acima não ocorre. Nesse caso, neurônios s que são ligados pelos neurônios σ podem possuir conexões com outros neurônios do buffer, mesmo que estes não estejam presentes na sequência de entrada. Portanto, as alturas das barras referentes aos neurônios σ_k podem ser não nulas $\forall k \in [0, M - 1]$ em gráficos como aqueles representados na figura 22. Além disso, como os mesmos neurônios s podem ser contados mais de uma vez (pois podem estar conectados com mais de um neurônio do buffer, devido à estrutura da matriz de conexão W aleatória), as alturas das barras no gráfico são, em geral, maiores do que nos dois casos abordados anteriormente.

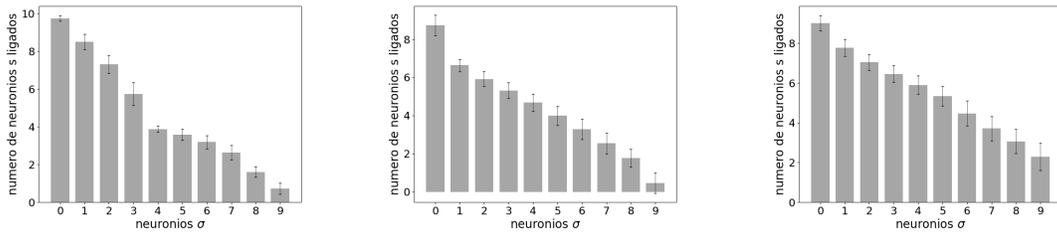
A diferença no perfil de ativação dos neurônios da rede recorrente para redes aleatórias com os dois tipos de matriz de conexão W mencionados é mostrada na figura 22. Como o processo de decodificação é baseado no número de neurônios s ativos, a análise da figura 22 sugere que o uso da matriz W do tipo seletiva resulta em um menor erro na codificação quando comparado à matriz W aleatória.

As figuras 23 e 24 mostram o perfil de ativação dos neurônios da rede recorrente para redes aleatórias com diferentes quantidades de conexões excitatórias e matrizes de conexão W seletiva e aleatória, respectivamente. A comparação dessas figuras com o gráfico da figura 22a (referente à rede ordenada) sugere que, dentre as configurações consideradas, aquela que apresentaria menor erro na decodificação é o caso da matriz de conexão W do tipo seletiva, e rede recorrente composta somente de conexões inibitórias.



(a) 0% conexões excitatórias (b) 20% conexões excitatórias (c) 40% conexões excitatórias

Figura 23: Número de neurônios da rede recorrente ativos ao final da dinâmica temporal (eixo vertical) conectados com cada um dos neurônios do buffer (eixo horizontal), para diferentes quantidades de conexões excitatórias: **(a)** 0%; **(b)** 20%; **(c)** 40%. Em todos os casos, as redes recorrentes são do tipo ER com matriz de conexão W seletiva, e parâmetros $L = 4$, $M = 10$, $N = 400$, $p = 2/N$. Foram geradas 400 redes e consideradas 20 sequências para cada uma delas. **Fonte:** Autora



(a) 0% conexões excitatórias (b) 20% conexões excitatórias (c) 40% conexões excitatórias

Figura 24: Número de neurônios da rede recorrente ativos ao final da dinâmica temporal (eixo vertical) conectados com cada um dos neurônios do buffer (eixo horizontal), para diferentes quantidades de conexões excitatórias: **(a)** 0%; **(b)** 20%; **(c)** 40%. Em todos os casos, as redes recorrentes são do tipo ER com matriz de conexão W aleatória com probabilidade de conexão $q = 0.1$, e parâmetros $L = 4$, $M = 10$, $N = 400$, $p = 2/N$. Foram geradas 400 redes e consideradas 20 sequências para cada uma delas. **Fonte:** Autora

5.1.2 Arquitetura da rede recorrente

Redes aleatórias podem ser obtidas a partir de diferentes algoritmos de construção, conforme abordado na seção 3.3. Portanto, um aspecto importante que deve ser observado na rede que será utilizada é a sua arquitetura. Nesse sentido, são

relevantes as seguintes questões: de que maneira a arquitetura da rede recorrente influencia o erro médio do processo de codificação? Existe um modelo de rede aleatória que se mostra mais eficiente que os outros? A fim de abordar essas questões, a distância de Levenshtein (seção 3.4) será utilizada como medida para quantificar a eficiência do processo de codificação descrito.

A figura 25 mostra os erros médios para quatro modelos de redes aleatórias: Erdős-Rényi, Watts-Strogatz, Newman-Watts-Strogatz e Barabási-Albert. Para a comparação, foram consideradas redes de mesmo tamanho e com valores de graus médios similares. Os gráficos dessa figura indicam que, para os parâmetros utilizados, as redes do tipo WS e NWS podem ser consideradas mais eficientes do que os outros dois modelos no processo de codificação. Além disso, a figura 25b também mostra a comparação entre as redes de conexão W dos dois tipos citados na seção 4.1.2, e fica evidente a superioridade da matriz W do tipo seletiva (65) em termos de eficiência na codificação.

Redes aleatórias dos tipos ER, WS e NWS possuem um parâmetro importante referente à probabilidade de conexão entre os nodos das redes: p (ER) e β (WS e NWS). O gráfico da figura 26 mostra como esses parâmetros influenciam o erro na codificação para redes de tamanho $N = 400$, considerando valores de p e β entre 10^{-6} e 1. Os dados representados no gráfico mostram que, mesmo para os valores do parâmetro p que apresentam menor erro, o erro na codificação ainda é maior para redes do tipo ER em comparação com redes do tipo WS e NWS, para quase todo valor de β . Apesar de existir um intervalo de valores p e β no qual o contrário ocorre (ou seja, redes do tipo ER seriam mais eficientes), esse intervalo é pequeno, e uma rede do tipo WS ou NWS com parâmetro $\beta \approx 1$ não possui as características de uma rede do tipo pequeno mundo, conforme explicado na seção 3.3.2.

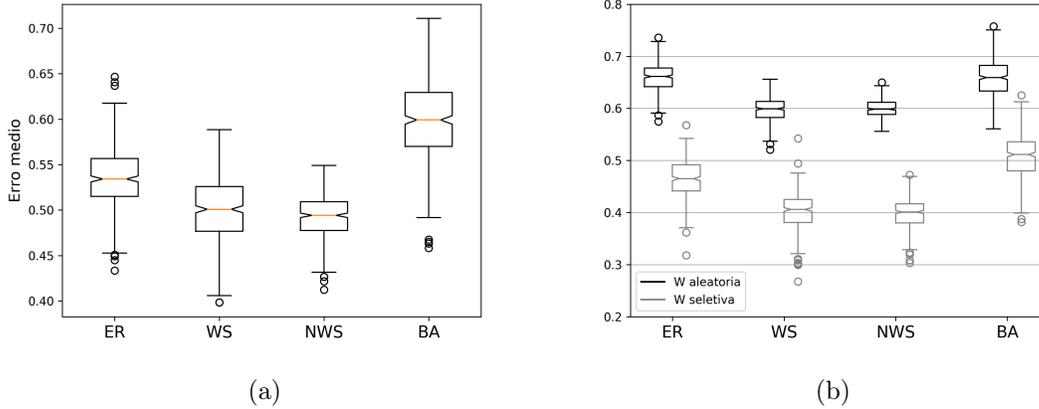


Figura 25: Influência do modelo de rede aleatória no erro da codificação. **(a)** Erro médio normalizado para quatro modelos de redes aleatórias de tamanho $N = 100$: Erdős-Rényi (ER), Watts-Strogatz (WS), Newman-Watts-Strogatz (NWS), e Barabási-Albert (BA), com matriz de conexão W do tipo seletiva. **(b)** Erro médio normalizado para redes dos quatro modelos citados, de tamanho $N = 200$, e com os dois tipos de matriz de conexão W : aleatória (preto) e seletiva (cinza). Em ambos os casos, foram utilizados os seguintes parâmetros: $L = 6$, $M = 10$, $p = 0.01$, $\beta = 0.1$, $K = 1$, $m = 1$. Foram consideradas 400 redes para cada modelo, e utilizadas 200 seqüências de entrada diferentes para o cálculo do erro. **Fonte:** Autora

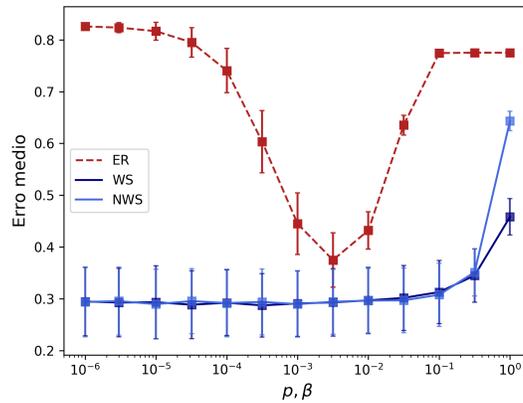


Figura 26: Erro médio normalizado para três modelos de redes aleatórias (ER, WS e NWS), para diferentes valores dos parâmetros de conexão p e β . Foram utilizados os seguintes parâmetros: $L = 6$, $M = 10$, $N = 400$, $K = 1$, e matriz W do tipo seletiva. Foram consideradas 400 redes para cada modelo, e utilizadas 20 seqüências de entrada diferentes para o cálculo do erro. **Fonte:** Autora.

5.1.3 Tamanho da rede recorrente

No modelo de rede descrito na seção 4.1.1, onde todos os pares de elementos da sequência de entrada são comparados em relação à sua ordem na sequência, a rede recorrente deve ter tamanho $N = M^2$. Entretanto, quando consideramos redes recorrentes de arquitetura aleatória, essa condição não é necessária, de modo que o modelo pode ser construído com redes de tamanho $N \geq M^2$. Devido à arquitetura aleatória da rede, existe a possibilidade de que nem todos os pares de elementos da sequência de entrada tenham suas ordens comparadas, enquanto alguns pares podem ser comparados mais de uma vez. Um maior número de neurônios na rede recorrente implica uma maior probabilidade de que todos os pares de elementos sejam de fato comparados. Assim, é esperado que redes recorrentes maiores estejam relacionadas a um menor erro de codificação.

Os gráficos da figura 27 indicam que o erro diminui exponencialmente com o aumento do tamanho da rede recorrente. Foi realizada regressão linear utilizando o métodos dos mínimos quadrados para avaliar a relação entre o tamanho da rede e o erro de codificação médio. Os dados da regressão linear estão resumidos na tabela 3. Comparando esses dados para os modelos de redes aleatórias citados (ER, WS, NWS e BA), observamos que os coeficientes angulares referentes a redes dos tipo WS e NWS são maiores em módulo do que os referentes aos outros dois modelos. Os gráficos da figura 27 e os dados da tabela 3 mostram que, de fato, um aumento no tamanho das redes recorrentes implica uma redução do erro (e, portanto, aumento na eficiência) no processo de codificação de sequências temporais.

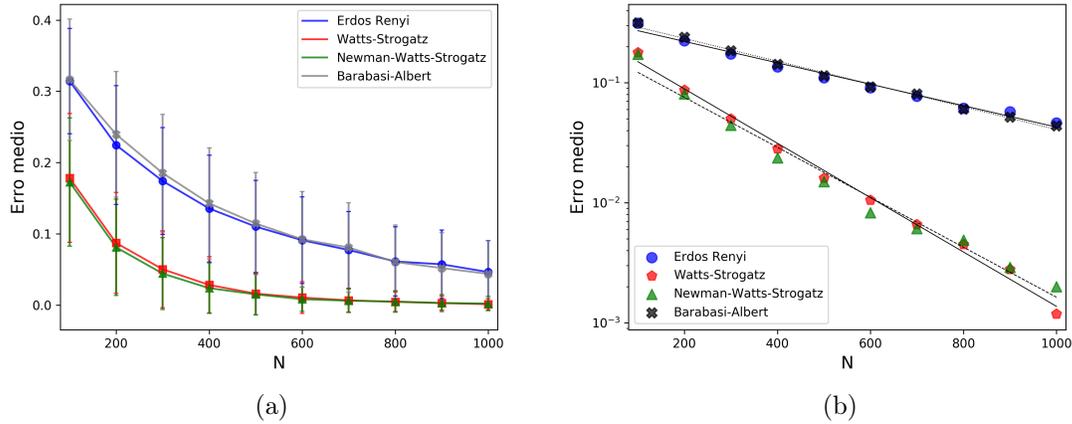


Figura 27: Influência do tamanho da rede recorrente no erro médio da codificação: **(a)** Erro médio normalizado para quatro modelos de redes aleatórias: Erdős-Rényi (ER), Watts-Strogatz (WS), Newman-Watts-Strogatz (NWS), e Barabási-Albert (BA), com matriz de conexão W do tipo seletiva, para diferentes tamanhos de rede. **(b)** Escala logarítmica. Os parâmetros das quatro retas (duas sólidas e duas hachuradas) foram obtidos a partir do método dos mínimos quadrados de regressão linear. Foram utilizados os seguintes parâmetros: $L = 4$, $M = 5$, $q = 2/N$, $\beta = 2/N$, $K = 1$, $m = 1$. Foram consideradas 400 redes para cada modelo, e utilizadas 20 seqüências de entrada diferentes para o cálculo do erro. **Fonte:** Autora

	Coefficiente angular	Intersecção	r	r ²	P
ER	-0.0020	0.3332	-0.9919	0.9839	<.001
WS	-0.0052	0.2524	-0.9968	0.9936	<.001
NWS	-0.0048	0.1971	-0.9903	0.9807	<.001
BA	-0.0022	0.3619	-0.9967	0.9934	<.001

Tabela 3: Regressão linear (ver figura 27b)

5.2 Modelo com Neurônios Taxa de Disparo

Serão analisados na presente seção os resultados obtidos para o modelo com neurônios taxa de disparo, definido na seção 4.2. Quando forem consideradas redes recorrentes com arquitetura aleatória, os interneurônios que mediarão as inibições entre os neurônios da rede não serão modelados explicitamente, de forma que consideraremos que as células principais se inibem diretamente. Representaremos o buffer de memória de trabalho pelas entradas que ele fornece à rede recorrente, e consideraremos que 4 de 5 padrões são mantidos no buffer.

A figura 28a mostra a atividade neural dada por (70) de 5 células do buffer de memória de trabalho, cada uma representando um determinado padrão. Na figura 28b, são mostradas as atividades de 20 grupos de neurônios na rede recorrente. A figura 28c mostra o potencial de membrana de 5 neurônios do buffer, e mostra como a ordem decodificada é construída a partir da atividade na rede recorrente, de forma que corresponde à sequência de entrada original da figura 28a.

Quando a rede recorrente possui arquitetura ordenada, de forma que as ordens

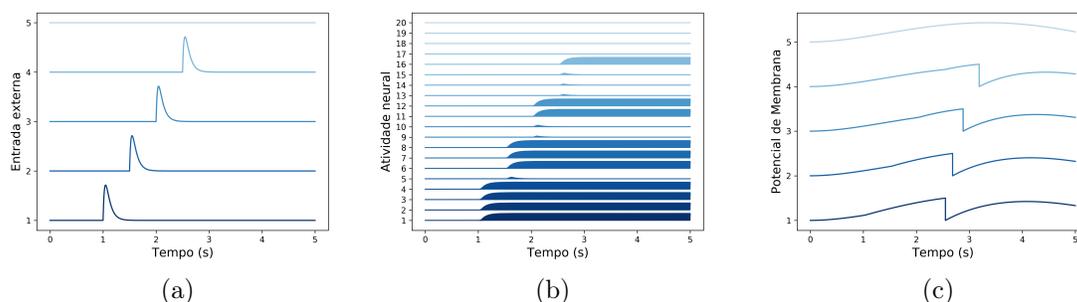


Figura 28: **(a)** Os traços indicam os sinais de entrada que chegam do buffer. Os quatro padrões são ativados em uma ordem específica, cada um em diferentes tempos. Nas simulações, todas as quantidades são mostradas em unidades adimensionais. Cada linha do gráfico corresponde a um grupo de neurônios diferente. **(b)** Atividade na rede recorrente. Cada linha indica quando um grupo de neurônios é ativado. O primeiro padrão da sequência temporal é capaz de ativar 4 grupos de neurônios alvos. O segundo padrão, ativa somente 3, e assim sucessivamente. **(c)** Decodificação realizada pelo segundo buffer de memória de trabalho. Os neurônios disparam de acordo com a atividade média relativa a cada padrão na rede recorrente. **Fonte:** Autora

de todos os padrões da sequência de entrada são comparadas, a sequência de saída é sempre igual à sequência de entrada. Nesse caso, o “perfil de ativação” dos neurônios possui uma estrutura bem definida, similarmente ao que foi descrito na seção 5.1.1. Entretanto, aqui, o padrão espacial será determinado pelas taxas de disparo dos neurônios da rede recorrente ao final da dinâmica temporal, de acordo com (67). A taxa de disparo média dos neurônios s que estão conectados ao primeiro neurônio da sequência temporal, σ_0 , será maior do que a taxa do segundo neurônio s , e assim sucessivamente. Esse fato pode ser observado no gráfico da figura 29a: o tamanho das barras correspondentes aos elementos da sequência de entrada se encontra em ordem decrescente, da esquerda para a direita. Além disso, a ausência de barras de erro no gráfico 29a indica que a ordem dos elementos da sequência de saída é sempre igual a da sequência de entrada, ou seja, o erro é nulo no caso de redes recorrentes com arquitetura ordenada. No caso de redes recorrentes com arquitetura aleatória, a situação descrita acima não ocorre (figura 29b). Similarmente ao caso do modelo com neurônios binários, quando a rede recorrente é aleatória, é possível que haja um erro na codificação.

A figura 30a mostra os erros médios na codificação para quatro tipos de redes

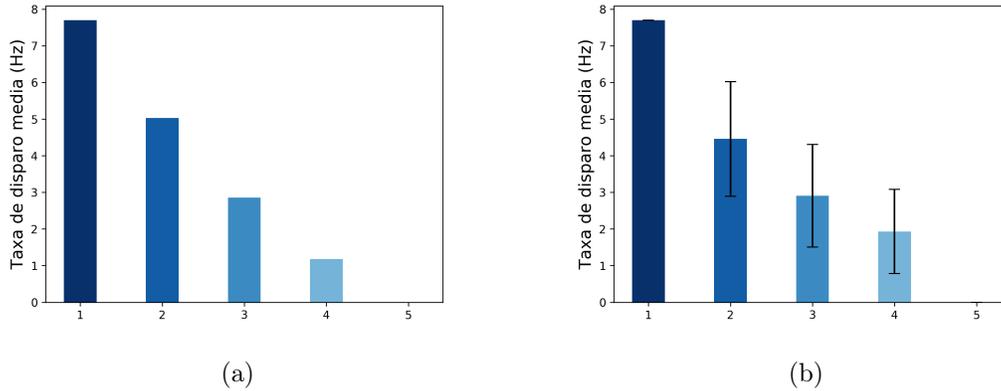


Figura 29: Taxa de disparo média (eixo vertical) correspondendo a cada um dos 5 padrões de entrada (eixo horizontal) para dois modelos de redes recorrentes de mesmo tamanho ($N=20$): **(a)** rede ordenada tal que as ordens de todos os padrões da sequência de entrada são comparadas; **(b)** Erdős-Rényi (com probabilidade de conexão $p=0.1$). 400 redes foram consideradas no cálculo da taxa de disparo média. **Fonte:** Autora

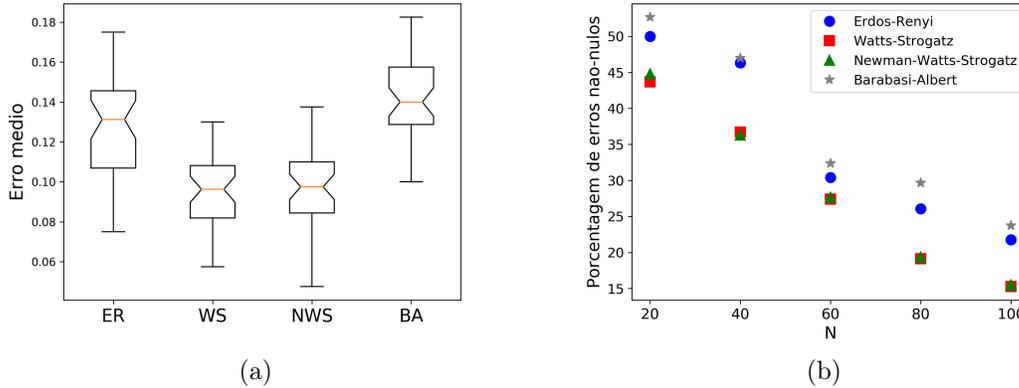


Figura 30: **(a)** Erro médio para quatro modelos de redes aleatórias: ER, WS, NWS e BA, com tamanho $N = 80$. 400 redes e 20 diferentes seqüências de entrada para cada rede foram utilizadas no cálculo do erro. **(b)** Porcentagem de erros não nulos para redes com tamanhos entre $N = 20$ e $N = 100$. Para o cálculo da porcentagem, foram utilizadas 40 amostras constituídas de 40 redes e 20 diferentes seqüências de entrada. Foram utilizados os seguintes parâmetros: $p = 2/N$, $\beta = 2/N$, $K = 1$, $m = 1$, e os valores das constantes são aqueles mencionados na seção 4.2. **Fonte:** Autora

aleatórias. Redes recorrentes do tipo WS e NWS apresentam erros médios menores do que redes do tipo ER e BA. Portanto, os dois primeiros modelos mencionados apresentam maior eficiência no processo de codificação analisado. Para a comparação, foram utilizadas rede de graus médios semelhantes. A figura 30b mostra que o erro médio diminui com o aumento do tamanho da rede recorrente, similarmente ao caso do modelo com neurônios binários (figura 27). Além disso, o mesmo comportamento observado no gráfico 30a é observado para todos os valores de N considerados: redes BA apresentam a pior eficiência (maior erro médio), ER apresentam a segunda pior eficiência, e as redes do tipo WS e NWS, que correspondem a redes que apresentam propriedades de redes pequeno mundo, são as mais eficientes no processo de codificação de seqüências temporais. Simulações foram escritas em C e Python. Integrações numéricas foram realizadas com o método de Euler, e $\Delta t = 0.0001$.

6 Conclusões

Existem evidências da utilização de buffers baseados em um esquema de multiplexação temporal pela memória de trabalho no cérebro, organizado por oscilações teta e gama [26, 38, 52]. Nesse tipo de codificação temporal, sequências são representadas por diferentes padrões de disparos síncronos, cada um deles em uma fase diferente da oscilação teta. Esses buffers de memória de trabalho se baseiam em propriedades de célula única, e possuem capacidade de armazenamento limitada. Por outro lado, redes atratoras recorrentes armazenam informações nas conexões sinápticas, e a memória é uma propriedade da rede.

Nesse trabalho, foi apresentado um modelo para comunicação neural entre um buffer de memória de trabalho, que utiliza o acoplamento de oscilações teta-gama para armazenar informações, e uma rede recorrente atratora, que armazena informações como sequências espaciais nas conexões sinápticas. O modelo foi construído com neurônios binários (seção 4.1), e neurônios taxa de disparo (seção 4.2). Na construção da rede recorrente, foram considerados quatro modelos de conectividades aleatórias. A análise dos resultados das simulações nos permite concluir que redes do tipo Watts-Strogatz e Newman-Watts-Strogatz, que apresentam propriedades de redes do tipo pequeno mundo, são as mais eficientes para o processo de codificação apresentado, dentre os modelos de redes aleatórias analisadas.

Muitas redes complexas observadas na natureza apresentam topologia de pequeno mundo, que é caracterizada por um elevado coeficiente de clustering e baixo comprimento de caminho médio [14]. O primeiro sistema nervoso a ser formalmente descrito como uma rede do tipo pequeno mundo foi o do nematoide *C. Elegans*, em escala microscópica [54]. Diversos estudos recentes com a utilização de ressonância magnética funcional e eletroencefalografia em humanos reportam redes cerebrais com topologia de pequeno mundo [12, 51, 57]. Simard *et al* [50] mostraram que redes neurais com topologia de pequeno mundo permitem maiores taxas de processamento de informação e aprendizado do que outros tipos de grafos aleatórios. Dessa forma,

os resultados obtidos aqui quanto a relação da arquitetura da rede e eficiência no processo de codificação de sequências temporais parecem estar de acordo com resultados descritos na literatura.

No presente trabalho, não consideramos ruído sináptico (ou seja, onde se consideraria que os neurônios disparariam usando uma distribuição de probabilidade). Foi feito, entretanto, o estudo da aleatoriedade das conexões no modelo proposto. O estudo desses modelos simples de codificação-decodificação podem ser importantes para o desenvolvimento de circuitos analógicos eletrônicos utilizados para imitar arquiteturas neurobiológicas presentes no sistema nervoso, chamados circuitos neuromórficos.

Em termos de trabalhos futuros que poderiam ser desenvolvidos a partir desse trabalho, podemos citar o uso de modelos de neurônios integra-e-dispara na construção da rede recorrente. Esse tipo de modelo de neurônio é muito mais realista do que o modelo taxa de disparo, e permite um estudo mais relevante da importância do ruído no modelo de codificação e decodificação.

Apêndice A Modelo de Hodgkin-Huxley

Hodgkin e Huxley propuseram [27], em 1952, um modelo que descreve a formação e propagação de potenciais de ação. Esse modelo (também chamado de modelo baseado em condutância) considera 3 correntes: uma corrente de potássio (I_K), uma corrente de sódio (I_{Na}), e uma corrente ôhmica de vazamento (I_L), devida em maior parte a íons de cloro. As equações de Hodgkin-Huxley são

$$\begin{aligned}
 C \frac{dV}{dt} &= I - \overbrace{\bar{g}_K n^4 (V - E_K)}^{I_K} - \overbrace{\bar{g}_{Na} m^3 h (V - E_{Na})}^{I_{Na}} - \overbrace{g_L (V - E_L)}^{I_L} \\
 \frac{dn}{dt} &= \alpha_n(V)(1 - n) - \beta_n(V)n \\
 \frac{dm}{dt} &= \alpha_m(V)(1 - m) - \beta_m(V)m \\
 \frac{dh}{dt} &= \alpha_h(V)(1 - h) - \beta_h(V)h
 \end{aligned} \tag{A.1}$$

onde

$$\begin{aligned}
 \alpha_n(V) &= 0.01 \frac{10 - V}{\exp(\frac{10-V}{10}) - 1}, & \beta_n &= 0.125 \exp(\frac{-V}{80}), \\
 \alpha_m(V) &= 0.1 \frac{25 - V}{\exp(\frac{25-V}{10}) - 1}, & \beta_m &= 4 \exp(\frac{-V}{18}), \\
 \alpha_h(V) &= 0.07 \exp(\frac{-V}{20}), & \beta_h &= \frac{1}{\exp(\frac{30-V}{10}) + 1}.
 \end{aligned} \tag{A.2}$$

Esses parâmetros correspondem a um deslocamento de aproximadamente $65mV$ no potencial de membrana [27], de forma que o potencial de repouso é $V \approx 0$. Nesse caso, os potenciais de equilíbrio iônico são $E_K = -12mV$, $E_{Na} = 120mV$, $E_L = 10.6mV$. Valores típicos das condutâncias máximas são $\bar{g}_K = 36mS/cm^2$, $\bar{g}_{Na} = 120mS/cm^2$ e $g_L = 0.3mS/cm^2$. $C = 1\mu F/cm^2$ é a capacitância da mem-

brana neuronal, e I é a corrente aplicada. n , m e h são quantidades adimensionais entre 0 e 1 relacionadas a ativação do canal de potássio, ativação do canal de sódio, e desativação do canal de sódio, respectivamente. A figura 31 mostra a dinâmica do potencial da membrana V e das variáveis m , n , e h .

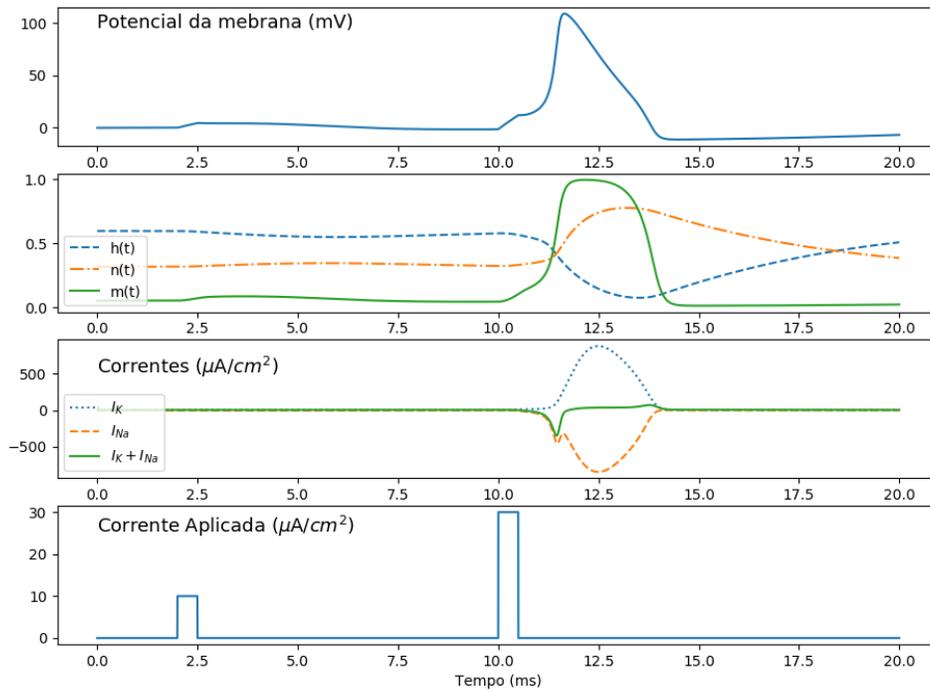


Figura 31: Dinâmica do potencial da membrana V e das variáveis m , n , e h do modelo de Hodgkin-Huxley. **Fonte:** Autora

Referências

- [1] L. A. Adamic. The small world web. In *Proceedings of the Third European Conference on Research and Advanced Technology for Digital Libraries, ECDL '99*, pages 443–452, London, UK, UK, 1999. Springer-Verlag.
- [2] R. Albert, H. Jeong, and A. L. Barabási. Diameter of the world-wide web. *Nature*, 401:130–, September 1999.
- [3] Réka Albert and Albert lászló Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 2002.
- [4] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. Classes of small-world networks. *PNAS*, 97(21), 2000.
- [5] R. Andrade. Cell excitation enhances muscarinic cholinergic responses in rat association cortex. *Brain Research*, 548:81–93, 1991.
- [6] R. C. Atkinson and R. M. Shiffrin. Human memory: A proposed system and its control processes. In *K. W. Spence and J. T. Spence (Eds.), The Psychology of learning and motivation: Advances in research and theory (vol. 2).*, pages 89 – 105, 1968.
- [7] N. Axmacher, M. M. Henseler, O. Jensen, I. Weinreich, C. E. Elger, and J. Fell. Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, 107(7):3228–3233, 2010.
- [8] A. L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311(3-4):590 – 614, 2002.
- [9] A. L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 1999.

- [10] F. Barbieri and N. Brunel. Can attractor network models account for the statistics of firing during persistent activity in prefrontal cortex? *Frontiers in Neuroscience*, 2(1), 114–122., 2008.
- [11] A. Barrat and M. Weigt. On the properties of small-world network models. *Eur. Phys. J. B*, 13:547, 2000.
- [12] D. S. Bassett and E. Bullmore. Small world brain networks. *Neuroscientist*, 12(6):512–523, 2006.
- [13] Mark Bear, Barry Connors, and Michael Paradiso. *Neuroscience: Exploring the brain (4th edition)*. Wolters Kluwer, 01 2015.
- [14] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(4):175 – 308, 2006.
- [15] B. Bollobás and O.M. Riordan. Mathematical results on scale-free random graphs. In *Handbook of Graphs and Networks*, editors, *S. Bormholdt and A. G. Schuster*. Wiley, 2003.
- [16] J. B. Brewer, J. D. E. Gabrieli, A. R. Preston, C. J. Vaidya, and A. C. Rosen. *Textbook of Clinical Neurology*, chapter 5. Saunders Elsevier, 3 edition, 2007.
- [17] P. Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Computational neuroscience. MIT Press, 2001.
- [18] A. Diamond. Executive functions. *Annual review of psychology*, 64:135–168, 2013.
- [19] P. Erdős and A. Rényi. On random graphs i. *Publicationes Mathematicae Debrecen*, 6, 1959.
- [20] J. M. Fuster and G. E. Alexander. Neuron activity related to short-term memory. *Science*, 173 3997:652–4, 1971.

- [21] E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 1959.
- [22] P. Goldman-Rakic. Cellular basis of working memory. *Neuron*, 14(3):477–85, 1995.
- [23] J. Grossman and P. Ion. On a portion of the well-known collaboration graph. *Congressus Numerantium*, 108, 08 1995.
- [24] D. O. Hebb. *The organization of behavior: A neuropsychological theory*. Wiley, 1949.
- [25] J. Hertz, A. Krogh, and R. G. Palmer. *Introduction to the Theory of Neural Computation*. Addison-Wesley Longman Publishing Co., Inc., 1991.
- [26] A. C. Heusser, D. Poeppel, Y. Ezzyat, and L. Davachi. Episodic sequence memory is supported by a theta-gamma phase code. *Nature neuroscience*, 19(10):1374–1380, 2016.
- [27] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol.*, 117(4):500-44, 1952 Aug 28.
- [28] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79(8):2554–2558, April 1982.
- [29] O. Jensen and L. L. Colgin. Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences*, 11:267–269, 2007.
- [30] O. Jensen and J. E. Lisman. Hippocampal ca3 region predicts memory sequences: accounting for the phase precession of place cells. *Learning and Memory*, 3:279–87, 09 1996.
- [31] H. Jeong, S.P. Mason, A.-L. Barabási, and Z.N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411, 2001.

- [32] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabási. The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654, 2000.
- [33] E. R. Kandel, J. H. Schwarz, and T. M. Jessel. *Principles of Neural Science*. McGraw-Hill, 4. edition, 2000.
- [34] K. Klemm and V. M. Eguíluz. Growing scale-free networks with small world behavior. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 65:057102, 06 2002.
- [35] L. Lapique. Considérations préambules sur la nature du phénomène par lequel l'électricité excite les nerfs. *Journal de Physiologie et de Pathologie Générale*, 9:565–578, 1907.
- [36] A. K. Lee and M. A. Wilson. Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron*, 36:1183–1194, 2002.
- [37] J. Lisman. The theta/gamma discrete phase code occurring during the hippocampal phase precession may be a more general brain coding scheme. *Hippocampus*, 15:913–22, 01 2005.
- [38] J. E. Lisman and M. A. P. Idiart. Storage of 7 ± 2 short-term memories in oscillatory subcycles. *Science*, 267(5203):1512–1515, 1995.
- [39] W. A. Little. The existence of persistent states in the brain. *Math. Biosci.*, 19(1-2):101–120, 1974.
- [40] G. A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63(2):81–97, March 1956.
- [41] M. E. J. Newman. The structure of scientific collaboration networks. *PNAS*, 98(2):404–409, 2001.
- [42] M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89, 2002.

- [43] M. E. J. Newman. Mixing patterns in networks. *Phys. Rev. E*, 2003.
- [44] M. E. J. Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.
- [45] M. E. J. Newman. *Networks: an introduction*. Oxford University Press, Oxford; New York, 2010.
- [46] M. E. J. Newman, C. Moore, and D. J. Watts. Mean-Field Solution of the Small-World Network Model. *Phys. Rev. Lett.*, 84:3201–3204, 2000.
- [47] M. E. J. Newman and D. J. Watts. Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4-6):341–346, 1999.
- [48] K. H. Pribram, G. A. Miller, and E. Galanter. *Plans and the structure of behaviour*. New York: Holt, Rinehart and Winston., 1960.
- [49] S. Raghavachari, M. J. Kahana, D. S. Rizzuto, J. B. Caplan, M. P. Kirschen, B. Bourgeois, J. R. Madsen, and J. E. Lisman. Gating of human theta oscillations by a working memory task. *J Neurosci.*, 1:21(9):3175–83, 2001.
- [50] D. Simard, L. Nadeau, and H. Kröger. Fastest learning in small-world neural networks. *Physics Letters A*, 336(1):8–15, Feb 2005.
- [51] O. Sporns, D. R. Chialvo, M. Kaiser, and C. Hilgetag. Organization, development and function of complex brain networks. *Trends in Cognitive Sciences*, 8(9):418 – 425, 2004.
- [52] S. Sternberg. High-speed scanning in human memory. *Science*, 153:652–654, 08 1966.
- [53] A. B. Tort, R. W. Komorowski, J. R. Manns, N. J. Kopell, and H. Eichenbaum. Theta-gamma coupling increases during the learning of item-context associations. *Proceedings of the National Academy of Sciences of the United States of America*, 106(49):20942–20947, 2009.

- [54] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, pages 440–442, 1998.
- [55] J. G. White, E. Southgate, J. N. Thomson, and S. Brenner. The structure of the nervous system of the nematode *c. elegans*. *Philosophical transactions Royal Society London*, 314:1–340, 1986.
- [56] H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12:1–24, 1972.
- [57] S. Yu, D. Huang, W. Singer, and D. Nikoloc. A small world of neuronal synchrony. *Cerebral Cortex*, 18(12):2891–2901, 2008.