IDENTIFICAÇÃO DE SISTEMAS NEURAIIS COM REDES BAYESIANAS DINÂMICAS E TRANSFERÊNCIA DE ENTROPIA

Fernando Pasquini Santos

Tese apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Doutor em Ciências no Programa de Engenharia Elétrica

ORIENTADOR: Prof. Dr. Carlos Dias Maciel
ÁREA DE CONCENTRAÇÃO: Sistemas Dinâmicos

Trata-se da versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

São Carlos
2017

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

S231i Santos, Fernando Pasquini
Identificação de Sistemas Neurais com redes
Bayesianas Dinâmicas e Transferência de Entropia /
Fernando Pasquini Santos; orientador Carlos Dias
Maciel. São Carlos, 2017.

Tese (Doutorado) - Programa de Pós-Graduação em
Engenharia Elétrica e Área de Concentração em Sistemas
Dinâmicos -- Escola de Engenharia de São Carlos da
Universidade de São Paulo, 2017.

1. Redes Bayesianas dinâmicas. 2. Transferência de
entropia. 3. Identificação de Sistemas. I. Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **FERNANDO PASQUINI SANTOS**.

Título da tese: "Identificação de sistemas neurais com redes bayesianas dinâmicas e transferência de entropia".

Data da defesa: 04/04/2017.

Comissão Julgadora:

Resultado:

Prof. Associado **Carlos Dias Maciel**
(Orientador)
(Escola de Engenharia de São Carlos/EESC)

APROVADO

Prof. Associado **Ricardo Zorzetto Nicoliello Vêncio**
(Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto/FFCLRP-USP)

APROVADO

Prof. Dr. **Ailton Akira Shinoda**
(Universidade Estadual Paulista "Júlio de Mesquita Filho"/UNESP – Ilha Solteira)

APROVADO

Prof. Dr. **Aparecido Augusto de Carvalho**
(Universidade Estadual Paulista "Júlio de Mesquita Filho"/UNESP – Ilha Solteira)

APROVADO

Prof. Titular **Marco Henrique Terra**
(Escola de Engenharia de São Carlos/EESC)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Associado **Luís Fernando Costa Alberto**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Luís Fernando Costa Alberto**

Dedicatória

---

*Dedico este trabalho
à minha esposa Jemima,
e aos meus pais.*

Agradeço ao Deus Trino; o Criador, Redentor e Sustentador de todas as coisas.

Agradeço à minha esposa, Jemima, por estar sempre ao meu lado neste tempo, dando-me a alegria, a segurança, o encorajamento e o aprendizado que o casamento proporciona. Agradeço pelo lar amoroso e acolhedor que ela me dá, que sempre me ensina a olhar "as aves do céu e os lírios do campo", vendo que a vida é muito mais do que meu trabalho. Em certo sentido ela tem sua parte neste trabalho, pois esteve sempre nos bastidores, ajudando com tudo o que eu não daria conta de fazer, enquanto estava, muitas vezes, ocupado demais com a pesquisa ou cansado.

Agradeço aos meus pais, Fernando e Viviani, aos meus irmãos, Daniel e Henrique, e aos meus avós, José e Cleusa, que são pontos de referência pelos quais minha vida se orienta.

Agradeço ao meu orientador, professor Carlos Dias Maciel, pela oportunidade deste trabalho, bem como por todo o apoio, dedicação e conselhos durante o mesmo. Agradeço por acreditar e valorizar meus esforços muitas vezes mais do que eu mesmo.

Agradeço à Fundação de Apoio à Pesquisa do Estado de São Paulo pela bolsa de doutorado direto recebida durante o projeto.

Agradeço a todos os colegas do Laboratório de Processamento de Sinais (EESC-USP), do Laboratório de Pesquisa em Radiofrequência e do Laboratório de Neuroimagem de Geriatria Psiquiátrica (GPN) (ambos na Universidade de Pittsburgh - EUA), pelo tempo de convivência em um ambiente sempre tão leve e amigável, e por toda as conversas e trocas de ideias. Posso estar sendo injusto ao escolher alguns nomes para citar, mas não posso deixar de deixar um agradecimento especial pela companhia e ajuda destes amigos em cada fase do trabalho: Wagner Endo, no início; Daniel Rodrigues de Lima, no meio; e Tales Santini, no final.

Agradeço ao professor Tamer S. Ibrahim, da Universidade de Pittsburgh (EUA), por me receber em seu laboratório e pela experiência de trabalho no departamento de radiologia, que contribuiu tanto para este trabalho e minha formação. Agradeço ao professor Howard J. Aizenstein e a Stephen Smagula pelas sugestões e colaboração na área de conectividade funcional. E também não posso deixar de agradecer tantas pessoas maravilhosas, cujos nomes infelizmente não cabem nestas páginas, que, durante parte do projeto, me receberam em Pittsburgh com as

casas e corações abertos; pela amizade, ajuda e pelos passeios. Sempre ficará a memória e a saudade deste tempo tão bom.

Agradeço também a todos os amigos que, mesmo muitas vezes entendendo pouco ou quase nada do que se encontra nas próximas páginas, sempre se interessaram pelo meu trabalho e me ajudaram a pensar um pouco mais sobre seu significado e escopo. Agradeço pelas conversas, ideias, sugestões de livros e até orações — tudo o que, durante este tempo, contribuiu não só para meu conhecimento técnico, mas também intelectual e pessoal.

*“Endless invention, endless experiment,
Brings knowledge of motion, but not of stillness;
Knowledge of speech, but not of silence;
Knowledge of words, and ignorance of the Word.”*
— T. S. Eliot

*“O universo da ciência é construído sobre o mundo-vivido e, se queremos pensar a própria
ciência com rigor, apreciar exatamente seu sentido e seu alcance, precisamos, primeiramente,
despertar essa experiência do mundo da qual ela é a expressão segunda.”*
— Maurice Merleau-Ponty

Conteúdo

Lista de Figuras	xv
Lista de Tabelas	xix
Lista de Símbolos	xxi
Lista de Siglas	xxiii
Resumo	xxv
Abstract	xxvii
1 Introdução	1
2 Redes Bayesianas Dinâmicas	9
2.1 Definição	9
2.2 Aprendizagem estrutural	12
2.2.1 Métodos baseados em pontuação	13
2.2.2 Métodos baseados em testes de independência condicional	15
2.2.3 Métodos híbridos	16
2.2.4 Aprendizagem em nível de grupo	16
2.3 DBNs em sinais não-estacionários	17
2.3.1 <i>Detrending</i> com Singular Spectrum Analysis	21
3 Usos de teoria de informação em DBNs	25
3.1 Definições	26
3.2 Aprendizagem de DBNs pela pontuação GlobalMIT	31
3.3 Valores de informação mútua par-a-par como estratégia para aprendizagem de redes Bayesianas com nós de mesmo tipo	33
3.4 Transferência de informação com atrasos de tempo contínuos	38

3.4.1	Redes Bayesianas Dinâmicas de Altas Ordens	40
4	Aplicações de identificação de sistemas neurais com DBNs	53
4.1	Ressonância Magnética Funcional (fMRI)	53
4.1.1	Conectividade funcional entre sub-áreas do hipocampo	56
4.1.2	Conectividade direcionada e modelos dinâmicos em fMRI	58
4.2	Identificação de sistema neuromotores de insetos	61
4.2.1	Descrição do sistema neuromotor do gafanhoto	62
4.2.2	Circuitos e conectividade neural	65
5	Materiais e métodos	71
5.1	Simulações	72
5.1.1	Aprendizagem estrutural de redes Bayesianas com nós de mesmo tipo	72
5.1.2	Aprendizagem estrutural de HO-DBNs	75
5.2	Conectividade neural entre sub-áreas do hipocampo em estado de repouso	77
5.2.1	Aquisição das imagens	79
5.2.2	Pré-processamento e aprendizagem de DBNs individuais	80
5.2.3	Pré-processamento e aprendizagem de DBNs em nível de grupo	81
5.3	Transferência de entropia atrasada no sistema neuromotor do gafanhoto	83
5.3.1	Aquisição dos sinais	83
5.3.2	Pré-processamento: Análise de Espectro Singular (SSA)	84
5.3.3	Medidas de transferência de entropia	86
5.3.4	Análise em nível de grupo: integração Bayesiana de medidas de trans- ferência de informação	87
6	Resultados e discussões	91
6.1	Simulações	91
6.1.1	Aprendizagem estrutural de redes Bayesianas com nós de mesmo tipo	91
6.1.2	Aprendizagem estrutural de HO-DBNs	96
6.2	Conectividade neural entre sub-áreas do hipocampo em estado de repouso	98
6.2.1	Aprendizagem de DBN individual com três ROIs	98
6.2.2	Aprendizagem de DBN em nível de grupo	99
6.3	Transferência de entropia atrasada no sistema neuromotor do gafanhoto	102
6.3.1	Pré-processamento com SSA	102

6.3.2	Impacto do número de amostras nas medidas de DTE	103
6.3.3	Consistência das medidas de DTE	104
6.3.4	Integração Bayesianas das medidas	107
6.3.5	Discussão	109
7	Publicações e Submissões	111
8	Conclusão e etapas futuras	113
	Bibliografia	119

Lista de Figuras

1.1	Tipos de modelos dinâmicos probabilísticos.	3
2.1	Redes de transição e <i>priori</i> de uma <i>DBN</i>	11
2.2	Diferenças entre <i>inter-timesliceDBNs</i> e <i>DBNs</i>	12
2.3	Classes de métodos de aprendizagem estrutural de redes Bayesianas	13
2.4	Fluxograma das considerações a serem tomadas na modelagem com sinais não-estacionários, juntamente com os tipos de modelo resultantes dessas considerações.	19
2.5	Ilustração para as três considerações sobre sinais não-estacionários, utilizando um sinal gravado nos neurônios do sistema neuromotor proprioceptor do gafanhoto.	20
3.1	Redes Bayesianas com nós de vários tipos de acordo com um perfil de transferência de informação.	34
3.2	A organização esperada nos valores de informação mútua de acordo com a configuração da rede entre dois nós (IIG) em uma rede Bayesiana com CPDs fixos.	37
3.3	Considerações sobre redes complexas com atrasos de tempo entre os nós.	39
3.4	Rede de transição de uma <i>HO-DBN</i>	41
3.5	Mapeamento de uma <i>HO-DBN</i> para um conjunto de listas de precedência.	47
4.1	Comparação entre imagens estruturais obtidas em equipamentos de MRI de 3T e 7T de força de campo magnético, em três ângulos.	55
4.2	Localização e morfologia do órgão femoral cordotonal (<i>FeCO</i>) em um gafanhoto.	63
4.3	Localização e morfologia do gânglio metatorácico e dos neurônios motores do membro anterior distal em um gafanhoto.	64
4.4	Topologia do sistema neuromotor do membro posterior do gafanhoto.	66

4.5	Estrutura de grafo do sistema neuromotor do gafanhoto obtida a partir de gravações de sinais realizadas por Kondoh et al. (1995) e medidas de informação mútua atrasada entre estes.	68
4.6	Grafo com as conexões do gânglio metatorácico do sistema neuromotor do gafanhoto encontrada a partir dos experimentos de medidas de atraso com informação mútua atrasada.	69
5.1	Redes Bayesianas com nós OU-ruídos utilizadas para o experimento — um exemplo de rede Bayesiana com CPDs fixos.	73
5.2	Uma descrição dos experimentos conduzidos para se checar uma separação nos valores de informação mútua entre dois nós em uma rede a partir da diretividade entre eles.	74
5.3	Quantidade de amostras necessárias para treinar <i>HO-DBNs</i> simuladas.	78
5.4	Equipamentos de MRI utilizado para aquisição dos dados.	79
5.5	Experimento conduzido por Newland e Kondoh (1997) para a coleta de sinais dos neurônios sensores, motores e interneurônios controlando a propriocepção na junção femoro-tibial do gafanhoto (e suas localizações no animal).	84
5.6	<i>Set up</i> experimental para as gravações realizadas no sistema neuromotor proprioceptor de gafanhotos.	85
5.7	Procedimento para obtenção das medidas de transferência de entropia atrasada normalizada, aplicado às gravações dos neurônios.	87
6.1	Histogramas e aproximação por <i>KDE</i> dos valores de informação mútua obtidos nas simulações de acordo com a diretividade entre dois nós.	92
6.2	Probabilidades de sucesso e erro na classificação dos valores de informação mútua em diretividade entre dois nós em uma rede Bayesiana com <i>CPDs</i> fixas.	93
6.3	Exemplos de <i>IIGs</i> com diferentes valores de diretividade nos quais os valores de informação mútua podem ser bastante parecidos (dependendo da quantidade de perdas de um nó para outro)	94
6.4	Tempo médio de convergência dos algoritmos <i>PSO</i> e <i>greedy</i> para aprendizagem estrutural de <i>HO-DBNs</i> simuladas	96
6.5	Resultado da segmentação automática realizada pelo software ASHS nas imagens obtidas.	99

6.6	Estrutura de DBN de maior pontuação com o algoritmo GlobalMIT aplicado a três nós de apenas um indivíduo.	99
6.7	Voxels da imagem de fMRI cuja correlação com <i>osseeds</i> (CA, DG e Sub) passaram do teste-T com conectividade maior do que zero.	100
6.8	Estrutura de DBN com maior pontuação no algoritmo GlobalMIT na análise com cinco ROIs e abordagem por VTS de análise em grupo.	100
6.9	Sinais do neurônio motor FETi antes e depois do pré-processamento com a técnica de remoção de tendências não-lineares e descontinuidades através de análise de espectro singular (seção 2.3.1).	103
6.10	Efeito do número de amostras no cálculo de DTE.	104
6.11	Efeito do número de amostras nas medidas de DTE.	105
6.12	Comparação das medidas de DTE (sem normalização) em três trechos uniformes de um mesmo sinal gravado no neurônio motor FETi.	106
6.13	Medidas normalizadas de DTE do estímulo de 27 Hz para o neurônio motor FETi, aplicado em diferentes gravações nos mesmos animais.	107
6.14	Medidas normalizadas de DTE do estímulo de 27 Hz para o neurônio motor FETi em todas as 23 gravações disponíveis.	107
6.15	Resultado da análise Bayesiana sobre as medidas normalizadas de transferência de entropia atrasada entre todos os experimentos entre o estímulo GWN de 27 Hz e a resposta no neurônio motor FETi . Os gráficos apresentam curvas de nível das probabilidades <i>a posteriori</i> de cada um dos dois picos (obtidos através da pontuação BIC), sendo atualizados com a inclusão de novas evidências e proporcionando uma maior densidade de probabilidade em pontos específicos. .	108
6.16	Diferença entre medidas de informação mútua atrasada e transferência de entropia atrasada aplicadas entre os mesmos sinais — do estímulo GWN de 27 Hz para o neurônio motor PFFiTi.	110

Lista de Tabelas

4.1	Visão geral de métodos de conectividade neural direcionada para modelos dinâmicos em fMRI.	61
4.2	Descrição dos neurônios motores do sistema neuromotor do gafanhoto, conforme também indicados na Figura 4.3(b) (Maciel et al., 2012)	65
4.3	Resumo das abordagens utilizadas nos estudos sobre o sistema proprioceptor na junta femoro-tibial do gafanhoto.	68
5.1	Gravações do sistema neuromotor proprioceptor do gafanhoto utilizadas para as medidas de informação mútua atrasada.	90
6.1	ROIs obtidas a partir do rotulamento das áreas ativadas na análise <i>seed-to-voxel</i>	101

Lista de Símbolos

$X_t^{(i)}$	variável aleatória de um processo X^i no tempo t 9
t	tempo 9
$\mathbb{G}$	grafo 9
$P(X)$	função de densidade de probabilidade da variável aleatória X 10
$P(X Y)$	função de densidade de probabilidade da variável aleatória X dado Y 10
$(X \perp Y)$	independência marginal entre X e Y 10
$(X \perp Y Z)$	independência condicional entre X e Y dado Z 10
$\mathbf{Pa}(X)$	pais do nó X em uma rede Bayesiana 10
$\mathbf{Nd}(X)$	não-descendentes do nó X em uma rede Bayesiana 10
τ	<i>lag time</i> ou intervalo de tempo entre amostras do processo reconstruído por <i>embedding</i> 38
λ	ordem da propriedade de Markov 40

Lista de Siglas

<i>2-TBN</i>	<i>2-timeslice Bayesian network</i>	10
<i>BDe</i>	<i>Bayesian-Dirichlet equivalent</i>	14
<i>BGe</i>	<i>Bayesian Gaussian equivalent</i>	20
<i>BIC</i>	<i>Bayesian Information Criterion</i>	15
<i>BOLD</i>	<i>Blood Oxygen Level Dependent</i>	54
<i>CPD</i>	Dist. de Probabilidade Condicional (<i>Conditional Probability Distribution</i>) ...	11
<i>DAG</i>	Grafo Direcionado Acíclico (<i>Directed Acyclic Graph</i>)	10
<i>DBN</i>	Rede Bayesiana Dinâmica (<i>Dynamic Bayesian Network</i>)	2
<i>DCM</i>	Modelagem Causal Dinâmica (<i>Dynamic Causal Modeling</i>)	6
<i>DTE</i>	Transferência de Entropia Atrasada (<i>Delayed Transfer Entropy</i>)	38
<i>DMI</i>	Informação Mútua Atrasada (<i>Delayed Mutual Information</i>)	38
<i>EPI</i>	Imagem Eco-Planar (<i>Echo-Planar Imaging</i>)	54
<i>FeCO</i>	Órgão Femoral Cordotonal (<i>Femoral Chordotonal Organ</i>)	63
<i>fMRI</i>	Ressonância Magnética Funcional (<i>Functional Magnetic Resonance Imaging</i>) ..	53
<i>GC</i>	Causalidade de Granger (<i>Granger Causality</i>)	58
<i>GWN</i>	Ruído Branco Gaussiano (<i>Gaussian White Noise</i>)	83
<i>HO-DBN</i>	Rede Bayesiana Dinâmica de Altas-Ordens (<i>High-Order Dynamic Bayesian Network</i>)	40
<i>IAAFT</i>	<i>Iterative Amplitude Adjusted Fourier Transform</i>	28
<i>IIG</i>	Grafo de Influência de Informação (<i>Information Influence Graph</i>)	36
<i>MI</i>	Informação Mútua (<i>Mutual Information</i>)	4
<i>PSO</i>	Otimização por Enxame de Partículas (<i>Particle Swarm Optimization</i>)	45
<i>ROI</i>	Região de Interesse (<i>Region of Interest</i>)	57
<i>SSA</i>	Análise de Espectro Singular (<i>Singular Spectrum Analysis</i>)	21
<i>TE</i>	Transferência de Entropia (<i>Transfer Entropy</i>)	27

Resumo

Pasquini Santos, Fernando. **Identificação de Sistemas Neurais com Redes Bayesianas Dinâmicas e Transferência de Entropia**. 2017. 168 p. Tese. Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos/SP, 2017.

Redes Bayesianas Dinâmicas (DBNs) são modelos capazes de representar um sistema dinâmico por meio de uma rede complexa que codifica as independências estatísticas condicionais entre os seus estados internos. Entre seus métodos de aprendizagem estrutural a partir de dados, o uso daqueles baseados em teoria de informação têm ganhado bastante espaço nos últimos anos, devido às suas vantagens de serem livres de modelo e permitirem uma aprendizagem *offline* a partir de medidas em múltiplas repetições do experimento. No entanto, resta uma exploração dos paralelos entre a área de aprendizagem de DBNs e aquela interessada em realizar medidas de transferência de informação entre elementos de um sistema neural, principalmente por meio de transferência de entropia (TE).

O presente trabalho busca, assim, aproximar estes dois focos de pesquisa, identificando suas equivalências e tratando de alguns dos desafios relacionados à sua implementação em identificação de sistemas neurais. Nota-se que uma das maiores dificuldades relacionadas ao uso de teoria de informação em sistemas multivariados concerne a alta dimensionalidade das funções de distribuição de probabilidade, exigindo grandes quantidades de dados observados simultaneamente. Não obstante, a aplicação de DBNs e transferência de entropia em sistemas de tempo contínuo também envolve considerações sobre a discretização dos sistemas no tempo, o que implica na necessidade de relaxamento da suposição da propriedade de Markov de primeira ordem (presente na definição de DBNs), e leva, assim, à proposta de redes Bayesianas dinâmicas de altas ordens (HO-DBNs).

Além de realizar uma revisão das principais propostas para a solução destas dificuldades,

o trabalho primeiramente propõe que, sob a suposição de um sistema com elementos se comportando de forma igual, os valores das medidas baseadas em teoria de informação com baixa dimensionalidade podem ser utilizados para a aprendizagem de estruturas de rede. Isso é mostrado a partir do uso de informação mútua par a par para a aprendizagem de redes Bayesianas simuladas com distribuições de probabilidade condicional fixas. No que concerne o uso de HO-DBNs, também se propõe um algoritmo baseado em otimização por enxame de partículas (PSO) para percorrer o espaço de busca de estruturas de HO-DBNs de forma mais eficiente.

Em seguida, duas aplicações de modelagem de DBNs com uso de teoria de informação são exploradas na área de sistemas neurais, tendo em vista a obtenção de conhecimento acerca de conectividade funcional e até uma aplicação futura em engenharia bioinspirada. Os desafios apresentados anteriormente são, assim, exemplificados, junto com algumas propostas de solução. A primeira área diz respeito à elicitación de conectividade funcional entre as sub-áreas do hipocampo, no cérebro humano, a partir de dados de ressonância magnética funcional (fMRI) de alta resolução. A partir de uma análise *seed-to-voxel* em grupo, regiões de interesse (ROIs) são identificadas e um modelo inicial de DBN é proposto, que é coerente com alguns estudos já feitos na literatura.

A segunda área de aplicação concerne a conectividade neural do sistema neuromotor do gafanhoto, a partir de gravações intracelulares de potencial sináptico em neurônios sensores, motores e interneurônios, sob estimulação com um fórceps no órgão femoral cordotonal (FeCO). Embora um modelo completo de DBN ainda não seja possível devido à ausência de gravações simultâneas suficientes, os atrasos de transferência de entropia entre o estímulo e a resposta nos neurônios motores são obtidos e integrados a partir de uma análise Bayesiana, dado também um pré-processamento com análise de espectro singular (SSA) que, ao remover a não-estacionariedade do sinal (que se deve a fatores extrínsecos ao sistema), aumentou consideravelmente a quantidade de amostras disponíveis. Tais resultados, ao ajudar a reduzir o espaço de busca de DBNs, também servem para direcionar futuros experimentos e pesquisas na área.

Palavras-chave: redes Bayesianas dinâmicas, transferência de entropia, identificação de sistemas, sistemas neuromotores, fMRI, engenharia bioinspirada, análise de espectro singular.

Abstract

*Pasquini Santos, Fernando. **Neural Systems Identification with Dynamic Bayesian Networks and Transfer Entropy**. 2017. 160p. Doctoral Thesis. São Carlos School of Engineering, University of São Paulo, São Carlos/SP - Brazil, 2017.*

Dynamic Bayesian Networks (DBNs) are models capable of representing a dynamical system by means of a complex network which codifies statistical conditional independencies between their internal states. Among their structural learning methods based on data, the use of ones based on information theory are gaining ground in recent years, due to their advantages of being model-free and permitting offline learning from multiple repetitions of an experiment. However, there still remains an exploration of the parallels between the areas of DBN structure learning and those interested in obtaining measures of information transfer between elements of neural systems, mainly through transfer entropy (TE).

Thus, the current work seeks to approximate these two foci of research by identifying some of their equivalences and challenges related to their usage in neural systems identification. It is noted that one of the main difficulties related to the use of information theory in multivariate neural systems concerns the high dimensionality of the probability distribution functions, requiring thus great quantities of data observed simultaneously. Furthermore, the application of DBNs and transfer entropy on continuous time systems also involves considerations about their discretization on time, which implies the necessity of relaxing the first order Markov property (intrinsic to the definition of DBNs), and thus leads to the proposal of high-order dynamic Bayesian networks (HO-DBNs).

Besides performing a review on the main proposals for solving these difficulties, this work first proposes that, under the supposition of a system with elements behaving in a similar way, the values of information theory based measures with low dimensions can be employed for

learning network structures. This is shown with the use of pairwise mutual information for learning simulated Bayesian networks with fixed conditional probability distributions. And concerning the use of HO-DBNs, an algorithm based on PSO is proposed in order to pass through their search space more efficiently.

Next, two applications of DBN modeling with information theory are explored in the field of neural systems, in view of obtaining knowledge about functional connectivity and even of a future application of bioinspired engineering. The challenges presented earlier are then exemplified along with some proposals of solutions. The first field regards the elicitation of functional connectivity between hippocampal subfields on the human brain based on high resolution fMRI data. Starting from a seed-to-voxel group analysis, regions of interest (ROIs) are identified and an initial DBN model is proposed, which is coherent with some studies already conducted in the literature.

The second field of application concerns the neural connectivity between the neuromotor system of the locust, based on intracellular synaptic potential recordings on sensory neurons, interneurons and motor neurons under stimulation by a forceps in the femoral chordotonal organ (FeCO). Although a complete DBN model is still not possible due to the absence of sufficient and simultaneous recordings, the transfer entropy delays between stimulus and responses on the motor neurons are obtained and integrated by a Bayesian analysis, given also a pre-processing based on Singular Spectrum Analysis (SSA) which, by removing the non-stationarity characteristics of the signal (which are due to extrinsic factors on the system), considerably increased the number of available samples for learning. Such results, by helping to reduce the search space of DBNs, also direct further experiments and studies on this field.

Keywords: *dynamic Bayesian networks, transfer entropy, system identification, neuromotor systems, fMRI, bioinspired engineering, singular spectrum analysis.*

Capítulo 1

Introdução

Modelos Dinâmicos e Analogias

Verum et factum convertuntur ("o verdadeiro e o factível são permutáveis"): desde a formulação deste princípio por Giambattista Vico, os cientistas têm cada vez mais se dedicado à proposição e obtenção de modelos, sob o pressuposto de que compreender a realidade significa reproduzi-la ou imitá-la (Dupuy e DeBevoise, 2009). Os modelos científicos, devido a sua capacidade de abstração, conseguem abstrair as características de um fenômeno ou sistema, colocando-os em uma classe de equivalência e estabelecendo particularidades a partir de parâmetros. Tais modelos tiveram grande sucesso desde o início da ciência moderna; no entanto, junto com seu avanço, a complexidade dos fenômenos que se desejava representar levantou a necessidade de se incorporar a incerteza e a probabilidade nos mesmos — seja devido ao desconhecimento das contingências presentes em muitas aplicações práticas, seja devido ao desconhecimento dos mecanismos complexos de interação entre várias partes. Como exemplo, os autores Sebastian Thrun, Wolfram Burgard e Dieter Fox, em seu livro "Probabilistic Robotics", retratam esta realidade na área da robótica móvel: "à medida em que os robôs estão saindo dos chãos de fábrica em direção a ambientes cada vez mais desestruturados, a habilidade de lidar com a incerteza é crucial para a construção de robôs bem-sucedidos (Thrun et al., 2005, p. 2).

Assim, os modelos probabilísticos têm recebido um grande foco de pesquisa dentro do paradigma científico atual, principalmente na área de modelos dinâmicos, ou seja, que buscam descrever (ou prever) o comportamento de um fenômeno ou sistema ao longo do tempo. Entre os modelos dinâmicos probabilísticos propostos recentemente, pode-se destacar os modelos

ocultos de Markov (Stratonovich, 1960) e os filtros de Kalman (Kalman, 1960). Os modelos ocultos de Markov ou HMMs (do inglês, *Hidden Markov Models*) são utilizados com sucesso em aplicações de reconhecimento de padrões e processamento de sinais (Rabiner e Juang, 1986) e representam o estado de um sistema como um processo estocástico discreto X_t (sendo t a posição no tempo), junto com as distribuições de probabilidade *a priori* $P(X_1)$, de transição de estados $P(X_t|X_{t-1}, U_t)$ e de observação $P(Y_t|X_t, U_t)$ (sendo U_t a entrada e Y_t a saída) (Murphy, 2002). Já os filtros de Kalman representam o estado com um vetor de variáveis aleatórias contínuas, porém limitam as distribuições de probabilidade *a priori*, de transição de estados e de observação a gaussianas; sendo, assim, uma representação linear. Portanto, existe um *trade-off* entre HMMs e filtros de Kalman — o primeiro não restringe as distribuições de probabilidade, mas limita o estado a uma variável, enquanto o segundo amplia para qualquer número de variáveis de estado, mas restringe as distribuições de probabilidade.

Uma das saídas encontradas para este problema reside nos esforços para a eliciação dos relacionamentos entre os próprios estados internos do sistema, isolando-os a partir de suposições de independência condicional (Pearl, 2000). Tal é o caso na área dos métodos de inferência Bayesiana multivariada (Candy, 2011) e, principalmente, de modelos gráficos capazes de representar independências entre variáveis, como as redes Bayesianas (Pearl, 2011) ou os campos aleatórios de Markov (em inglês, *Markov random fields* ou *MRFs*) (Rozanov, 1982). Isso também foi possibilitado com o próprio aumento de poder computacional para cálculos de inferência e aprendizagem destes modelos. Assim, em tempos recentes observou-se uma crescente utilização de redes Bayesianas dinâmicas, ou *DBNs* (do inglês, *Dynamic Bayesian Networks*), sendo estas redes Bayesianas aplicadas a sistemas dinâmicos e análise de séries temporais. Estas acabam-se mostrando também como uma caso geral de HMMs e filtros de Kalman; ou seja, permitindo mais de uma variável na representação de estado, e relações não-lineares entre estas (Murphy, 2002). Os algoritmos de inferência e aprendizado são possibilitados neste domínio uma vez que as próprias independências condicionais previstas no modelo ocasionam simplificações nos cálculos de probabilidade (vide seção 2.1). Tais modelos, enfim, também fazem parte de uma abordagem de redes complexas para a compreensão dos fenômenos e sistemas (Mitchell, 2009).

A Figura 1.1 apresenta os tipos de métodos de modelagem probabilística, ilustrando a ideia de *DBNs* como generalização de *HMMs* e filtros de Kalman.

Uma outra característica dos modelos científicos é que eles se valem de analogias para re-

modelos probabilísticos

a) sem estados: $\mathbf{X} \rightarrow \boxed{P(\mathbf{Y}|\mathbf{X})} \rightarrow \mathbf{Y}$

a) com estados:

i) Hidden Markov models (HMMs):

- apenas um estado
- qualquer p.d.f.

ii) Filtros de Kalman (KFs):

- múltiplos estados
- p.d.f. gaussiana (linear)

iii) Redes Bayesianas Dinâmicas (DBNs)

- múltiplos estados
- qualquer p.d.f.
- inferência Bayesiana

Figura 1.1: Tipos de modelos dinâmicos probabilísticos utilizados na engenharia. Os modelos sem estado simplesmente representam a entrada em função da saída com uma função de distribuição de probabilidade (*p.d.f.*). Já os modelos com estado incluem os modelos ocultos de Markov (*HMMs*) e os filtros de Kalman (*KFs*); cada um com suas capacidades e restrições, sendo estas últimas superadas uma vez que se considere um modelo generalizado — as redes Bayesianas dinâmicas *DBNs*.

presentar o conhecimento; e um grande exemplo disso é a teoria de informação, iniciada pelo trabalho de Claude Shannon, e que passou a enxergar os processos e fenômenos em termos de sistemas de comunicação (Cover e Thomas, 2006). Alan Turing, posteriormente, estabelece a base para o uso da analogia de sistemas de computação e processamento de informação para a compreensão de sistemas biológicos (que encontra seu auge no advento da cibernética, nas Macy Conferences, entre 1946 e 1953) (Dupuy e DeBevoise, 2009). Ele divide estes tipos de sistemas em três funções principais: armazenamento, transferência e modificação de informação (Langton, 1990), relacionados aos conceitos fundamentais de memória, comunicação e processamento, respectivamente (Lizier, 2010). Cada uma dessas funções tem dado origem a diversos tipos de medidas, de acordo com a área de aplicação Wollstadt et al. (2014), entre as quais se destaca a proposta da transferência de entropia como uma medida de transferência de informação ou TE (do inglês, *transfer entropy*) por Schreiber (2000), a qual será explorada no presente trabalho.

Uma das características mais vantajosas encontradas em abordagens baseadas em teoria de informação em sistemas envolvendo probabilidade é sua característica não-linear e livre de

modelo (Vicente et al., 2011). Por "livre de modelo" entende-se que a análise é mais guiada pelos dados, apresentando pouca ou quase nenhuma suposição acerca das equações que regem o sistema — no caso, que as medidas são realizadas diretamente a partir das distribuições de probabilidade dos dados. Por outro lado, esta vantagem também tem seu custo: ela envolve o uso de distribuições de probabilidade com altas dimensões, exigindo, por conseguinte, também altas quantidades de dados para a estimação das mesmas (Runge, 2014).

Outra vantagem encontrada nestes está no fato de que medidas como informação mútua (MI, do inglês *mutual information*) ou TE são compreendidas como medidas *funcionais*, no sentido de que mapeiam funções de distribuição de probabilidade em números reais que resumem características do sistema (Papoulis e Pillai, 2002; Wollstadt et al., 2014)). Em alguns tipos de aplicação, isso elimina a necessidade de que os dados utilizados para a análise sejam observados em um só experimento, um caso que pode ser definido como de sistemas parcialmente observados. Sob a suposição de um sistema igual em diversas repetições de experimentos, pode-se supor (dadas as devidas precauções) que as medidas funcionais sobre os dados não se alteram, e com isso os dados necessários para se treinar uma rede podem ser obtidos separadamente (Wollstadt et al., 2014).

Assim, dadas as vantagens de ambas as abordagens por DBNs e teoria de informação, diversos trabalhos têm buscado estabelecer relações entre as duas, de forma a usá-las juntamente em diversos tipos de aplicação (Campos, 2006; Lizier et al., 2013). Digno de nota é a correspondência encontrada entre medidas de transferência de informação, principalmente no caso da TE, na identificação de independências condicionais em um sistema, dadas certas adaptações nos conceitos e definições (Wollstadt et al., 2014). A revisão das principais propostas, tanto da literatura como neste trabalho, será realizada no capítulo 3.

Aplicações e Desafios em Sistemas Neurais

Os modelos dinâmicos probabilísticos, bem como o paradigma de redes complexas e de teoria de informação, têm encontrado uma grande incursão nas pesquisas em neurociência (Bielza e Larrañaga, 2014; Dimitrov et al., 2011; Sporns, 2011). Junto com eles, também encontram-se diversos métodos que se valem da análise de altos volumes de dados, de diversas origens e tipos, para a proposição de modelos e medidas — dando origem, inclusive, a várias discussões sobre seu significado, ou mesmo sua validade e reprodutibilidade (Coveney et al., 2016). O próprio uso de modelos de computação para o estudo, por exemplo, de neurociência

cognitiva ou sistemas neuromotores, exige uma reflexão acerca de suas limitações enquanto analogia e abstração, de forma a se manter os campos de investigação científica dentro de suas respectivas esferas, evitando reducionismos (Clouser, 2008).

Um dos pontos, por exemplo, onde as limitações da modelagem devem ser reconhecidas diz respeito ao conceito de causalidade, conforme discute a obra de Judea Pearl (Pearl, 2000). Valdes-Sosa et al. (2011) mostra que duas interpretações acerca de causalidade são possíveis: uma delas enxerga a causalidade como uma *precedência temporal*, ou seja, que uma causa é aquilo que precede uma consequência; enquanto a outra enxerga a causalidade como *controle*, ou seja, que a mudança em uma causa produz uma mudança nas consequências. Interpretações como essas dão origem a diferentes métodos na área de neurociência: a primeira geralmente utiliza análises baseadas em causalidade de Granger (Granger, 1988), enquanto a segunda utiliza as ferramentas de cálculo causal desenvolvidas por Pearl (2000).

No contexto de DBNs e teoria de informação aplicadas a sistemas neurais, diversos autores têm chamado atenção quanto ao perigo de confundir certas medidas de associação com causalidade (Valdes-Sosa et al., 2011; Wibral et al., 2014b; Koller e Friedman, 2009). As DBNs devem ser compreendidas como modelos gráficos probabilísticos capazes de representar apenas independências condicionais entre variáveis, e, como propõe Pearl (2000), a representação de causalidade requer o emprego de seu formalismo conhecido como *do-calculus*. Semelhantemente, Wibral et al. (2014b) chama atenção para a possibilidade de que um elemento em um sistema pode agir como causa de outro sem a existência de transferência de informação (por exemplo, suponha um som repetitivo sendo recebido e interpretado pelo sistema auditivo), levando à conclusão de que medidas de transferência de informação como a TE não podem ser consideradas como medidas de causalidade.

Assim, se o uso de modelagem probabilística de sistemas biológicos, seja com DBNs ou com transferência de informação, não implica causalidade, qual seria sua importância em aplicações de modelagem em sistemas neurais? Em primeiro lugar, nota-se que, embora uma relação estatística não implique causalidade, a descoberta de independências condicionais consiste em um primeiro passo para esta finalidade — no caso, trata-se do que a literatura costuma chamar de abordagem *exploratória* Valdes-Sosa et al. (2011). Ao encontrar independências condicionais entre variáveis, uma DBN, por exemplo, pode dirigir análises ou mesmo experimentos posteriores, tendo em vista abordagens conhecidas como confirmatórias, e que, portanto, já se valem do uso de conceitos de causalidade e *do-calculus*, como por exemplo, a modelagem cau-

sal dinâmica, ou DCM (do inglês, *dynamic causal modeling*) proposta por Smith et al. (2011).

Em segundo lugar, as técnicas mencionadas podem ajudar na descoberta de arquiteturas de controle em diversos tipos de sistema neuromotores, gerando *insights* para engenheiros e programadores (Wessnitzer e Webb, 2006). Assim, elas têm encontrando ampla utilização em aplicações de engenharia e robótica bioinspirada (Kovac, 2014), que já tem obtido resultados em várias áreas como locomoção de robôs (Buschmann et al., 2015), navegação (Srinivasan et al., 1999; Sinha, 2015), estabilização de voo (Fuller et al., 2014), visão computacional (Franceschini et al., 1992; Prabhakara et al., 2012), desenvolvimento de próteses, na área de reabilitação (Jimenez-Fabian e Verlinden, 2012; Ying et al., 2007)) e até aplicações de interação humano-máquina (principalmente na área de interfaces cérebro-máquina) (Cannan e Hu, 2011).

Por fim, também vale citar que diversas técnicas lineares têm utilizadas com sucesso em aplicações de neurociência (Sameshima e Baccalá, 1999; Schelter et al., 2006). Pode-se destacar, por exemplo, medidas de significância da seletividade neural (Ince et al., 2012) e medidas de correlação parcial direcionada (Baccalá e Sameshima, 2001; Schelter et al., 2006) — todas apresentadas sob suposições de modelos lineares ou *point-process* (Gerhard et al., 2013). No entanto, em vários outros casos é notável a necessidade de se modelar comportamento não-linear dos sistemas, e assim, os modelos de DBN e medidas de transferência de informação surgem como uma alternativa.

Objetivo e Estrutura do Trabalho

O presente trabalho tem como objetivo realizar um estudo das técnicas de identificação de sistemas com DBNs baseadas em medidas de teoria de informação, explorando alguns dos desafios concernentes à sua integração e aplicação em dois tipos de dados obtidos de sistemas neurais. Em termos específicos, os objetivos são:

- Realizar uma revisão das diversas formas nas quais os métodos de teoria de informação têm sido empregados juntamente com modelos de DBN, incluindo a proposta de redes Bayesianas de altas ordens (HO-DBNs) para sistemas com atrasos de tempo diferentes;
- Mostrar, através de simulações e experimentos, as vantagens e desvantagens do uso de teoria de informação com DBNs, incluindo um método de aprendizagem de HO-DBNs e outro para redução de dimensionalidade dada a suposição de nós iguais;

- Explorar esta integração de métodos em duas aplicações práticas, com resultados importantes para as áreas de neurociência cognitiva e engenharia bioinspirada; sendo elas:
  - O estudo de conectividade funcional entre sub-áreas do hipocampo humano, a partir de imagens de fMRI, incluindo o pré-processamento com segmentação automática destas sub-áreas, aprendizagem de uma DBN por meio de uma pontuação baseada em teoria de informação, e uma análise em nível de grupo dos resultados obtidos, e
  - a análise de conectividade no sistema neuromotor do gafanhoto, a partir de sinais de potencial sináptico, incluindo o pré-processamento para remoção de tendências complexas e não-lineares no sinal baseada em Análise de Espectro Singular (SSA), seguida de medidas de transferência de entropia e uma análise em nível de grupo baseada em um método Bayesiano.

O trabalho se divide como se segue: o capítulo 2 introduz as DBNs e seus métodos de aprendizagem estrutural e análise em grupo, além de considerações sobre seu uso em sistemas não-estacionários. Em seguida, o capítulo 3 detalha e discute a utilização de técnicas de teoria de informação junto com o paradigma de aprendizagem de DBNs, realizando uma revisão das medidas de transferência de informação, a apresentação de uma pontuação de DBNs baseada em informação mútua condicional, a proposta de redes Bayesianas com nós do mesmo tipo como uma simplificação no uso da informação mútua condicional e, por fim, uma consideração sobre sistemas com diferentes atrasos de tempo, relaxando a propriedade de Markov de primeira ordem das DBNs. Então, as duas aplicações são apresentadas e discutidas no capítulo 4: a identificação de conectividade funcional entre as sub-áreas do hipocampo humano, a partir de imagens de fMRI, e a identificação de conexões no sistema neuromotor do gafanhoto, a partir de gravações intracelulares de respostas sinápticas de neurônios motores sob um sinal de excitação nos neurônios sensores. Tais aplicações ilustram os dois tipos de vantagem mencionados anteriormente decorrentes da modelagem probabilística de sistemas biológicos; a saber, a abordagem exploratória na descoberta de conectividade neural e *insights* para engenharia bioinspirada na área de locomoção de robôs.

Dada a apresentação e revisão dos conceitos fundamentais, o capítulo 5 apresenta os experimentos realizados tendo em vista o uso de DBNs com medidas de transferência de informação e sua aplicação na identificação de sistemas neurais. Dados simulados foram usados para a avaliação de um algoritmo de aprendizagem de redes Bayesianas dinâmicas de altas ordens

(HO-DBNs), importantes para a utilização conjunta com métodos de transferência de informação, e para a avaliação de um algoritmo de aprendizagem de redes sob uma suposição de nós do mesmo tipo, capaz de reduzir a exigência de altas dimensões nas distribuições de probabilidade e observações simultâneas do sistema (seção 5.1). Em seguida, os métodos utilizados para o estudo das aplicações discutidas no capítulo 4 são apresentados. Em primeiro, busca-se analisar a conectividade funcional entre sub-áreas do hipocampo humano através de um pré-processamento com segmentação automática das sub-áreas, aprendizagem de uma DBN por meio de uma pontuação baseada em teoria de informação, e uma análise em nível de grupo dos resultados obtidos (seção 5.2). A seguir, conduz-se uma análise de conectividade no sistema neuromotor do gafanhoto, com um pré-processamento para remoção de uma tendência complexa e não-linear baseada em Análise de Espectro Singular (SSA), medidas de transferência de entropia e uma análise em nível de grupo baseada em um método Bayesiano. Os resultados são, por fim, apresentados no capítulo 6.

As publicações e submissões realizadas durante o presente trabalho são listadas no capítulo 7. Por fim, o capítulo 8 apresenta as principais conclusões do trabalho e direções para pesquisas futuras.

Capítulo 2

Redes Bayesianas Dinâmicas

Conforme afirmado na introdução, as redes Bayesianas dinâmicas podem ser entendidas como modelos gráficos probabilísticos aplicados a sistemas dinâmicos, capazes de generalizar tanto os modelos ocultos de Markov (Stratonovich, 1960), uma vez que permitem múltiplos estados internos, como os filtros de Kalman (Kalman, 1960), uma vez que ampliam-nos para o caso não-linear ao permitirem qualquer forma de representação das probabilidades condicionais.¹

Neste capítulo, será realizada uma revisão dos conceitos básicos envolvidos na definição (seção 2.1) e aprendizagem (seção 2.2) de modelos de rede Bayesiana dinâmica, seguida de uma consideração sobre seu emprego em sistemas não-estacionários (seção 2.3).

2.1 Definição

Seja $\mathbf{X} = \{X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(N)}\}$ um conjunto de N processos de tamanho T em um sistema dinâmico, onde $X_t^{(i)}$ é uma variável aleatória que consiste em uma realização do processo $X^{(i)}$ no tempo t . Assume-se também que $X_{t_1:t_2}^{(i)}$ é o conjunto de variáveis $X^{(i)}$ do tempo t_1 a t_2 , e $\mathbf{X}_{t_1:t_2}$ é o conjunto de todas as variáveis do sistema, do tempo t_1 a t_2 . P é a distribuição de probabilidade conjunta do sistema.

Define-se também $\mathbb{G}_{T_r}$, intitulado rede de transição, um grafo cujos nós são as variáveis

---

¹Os filtros de Kalman estendidos e filtros de Kalman *unscented* (Wan e Van Der Merwe, 2000) são conhecidos como versões do filtro de Kalman para sistemas não-lineares, e são aplicados com sucesso em vários estudos, como (Chen e Yu, 2014; Al-Hussein e Haldar, 2016). Porém, sendo sua a ideia utilizar derivadas parciais em torno de um ponto (caso estendido) ou transformações nas amostras (caso *unscented*), suas funções de distribuição de probabilidade continuam restritas a um modelo gaussiano linear.

em dois instantes genéricos de tempo, $t - 1$ e t , com arestas não permitidas entre nós em $t - 1$ e saindo dos nós em t para $t - 1$ (este tipo de grafo também é chamado de *2-TBN* — *2-timeslice Bayesian network* (Koller e Friedman, 2009)). $\mathbb{G}_{Pr}$, por sua vez, é chamado rede de *priori*, e consiste em um DAG (grafo acíclico direcionado, ou *directed acyclic graph*) cujos nós são as variáveis no instante de tempo inicial (Friedman et al., 1998). A Figura 2.1 apresenta um exemplo de rede de transição e rede de *priori*, com a extensão dos grafos em uma rede Bayesiana comum. Define-se uma Rede Bayesiana Dinâmica como o trio $(\mathbb{G}_{Tr}, \mathbb{G}_{Pr}, P)$ que satisfaz as seguintes condições:

1. Propriedade de Markov de Primeira Ordem: para cada instante t , os valores futuros são independentes dos valores passados dado o presente; formalmente, $(\mathbf{X}_{t+1} \perp \mathbf{X}_{0:t-1} | \mathbf{X}_t)$.
2. Estacionariedade / Homogeneidade: $P(\mathbf{X}_t | \mathbf{X}_{t-1})$ tem a mesma distribuição de probabilidade para cada instante t .
3. Condição Causal de Markov: para cada variável $X^{(i)}$ na rede de *priori*, qualquer nó $X^{(i)}$ é independente de seus não-descendentes dado os seus pais; formalmente, $(X_t^{(i)} \perp \mathbf{Nd}(X_t^{(i)}) | \mathbf{Pa}(X_t^{(i)}))$ (onde $\mathbf{Nd}(X_t^{(i)})$ são os não-descendentes de $X_t^{(i)}$, e $\mathbf{Pa}(X_t^{(i)})$ são os pais de $X_t^{(i)}$).

É importante não confundir a propriedade de Markov de primeira ordem, relacionada à independência no tempo, com a condição causal de Markov, relacionada à independência entre as variáveis nas redes de transição e *priori*. Nota-se também que a definição de *DBNs* amplia a capacidade de descrição das redes Bayesianas, permitindo a descrição de ciclos e circuitos de malha fechada — situações que só poderiam ser consideradas a partir de uma perspectiva temporal, pois é uma impossibilidade lógica que haja um ciclo de dependências estatísticas entre as variáveis de um sistema em um mesmo instante de tempo.

Com a propriedade de Markov de primeira ordem e a suposição de estacionariedade, é possível decompor a distribuição de probabilidade conjunta das variáveis do sistema da seguinte forma:

$$P(\mathbf{X}_{0:t}) = P(\mathbf{X}_0) \prod_{t=1}^T P(\mathbf{X}_t | \mathbf{X}_{t-1}) \quad (2.1)$$

Os grafos da rede de transição e rede *prioris* da *DBN* representam independências condicionais na distribuição de probabilidade $P(\mathbf{X}_t | \mathbf{X}_{t-1})$. Assim, pode-se ainda decompor os termos

Figura 2.1: Exemplo de rede Bayesiana dinâmica com três nós $\{X_1, X_2, X_3\}$. (a) contém as redes de priori e transição; e (b) mostra a rede Bayesiana equivalente, estendida no tempo (Friedman et al., 1998).

$P(\mathbf{X}_t|\mathbf{X}_{t-1})$ como:

$$P(\mathbf{X}_t|\mathbf{X}_{t-1}) = \prod_{i=1}^N P(X_t^{(i)}|\mathbf{Pa}(X_t^{(i)})) \quad (2.2)$$

onde $\mathbf{Pa}(X_t^{(i)})$ são os pais do nó $X_t^{(i)}$ nos grafos da rede de transição e *prioris*. Estas decomposições são propriedades úteis para inferências estatísticas com *DBNs*, e provam que o propósito dos grafos de rede de transição e *prioris* é indicar as independências condicionais entre os nós do sistema — que é a própria vantagem do uso de métodos de redes Bayesianas ou mesmo modelos gráficos probabilísticos (Koller e Friedman, 2009). Uma rede Bayesiana ou *DBN*, portanto, pode ser especificada pelo grafo que define a estrutura da rede e as funções de distribuição de probabilidade condicionais ou *CPDs* (do inglês *conditional probability distributions*) $P(X^{(i)}|\mathbf{Pa}(X^{(i)}))$, $\forall i$.

É digno de nota que muitos autores utilizam em seus trabalhos *DBNs* cujas redes de transição permitem arestas apenas saindo dos nós em $t - 1$ até os nós em t (ou seja, apenas arestas *inter-timeslice*), e não entre nós no mesmo instante de tempo (Xuan et al., 2012b) (*intra-timeslice*). Assim, os autores usam uma representação de grafo simples, sem os dois instantes de tempo genérico (*2-TBN*), supondo implicitamente que toda aresta sai de um nó no instante $t - 1$ e vai para um nó em t . Este é o caso da maioria dos trabalhos recentes que exploram extensões de *DBNs* em domínios não-estacionários, visando a aplicação em redes de regulação de genes (como Jia e Huan (2010); Robinson e Hartemink (2010) e Grzegorzczuk e Husmeier (2011)) e fluxo de informação neural (como Smith et al. (2006)). Nesses casos, referir-se-á aos modelos utilizados como *inter-timeslice DBNs*.

Esta consideração é importante, pois atribui diferentes significados aos modelos e pode ser

Figura 2.2: Diferença entre *DBNs* consideradas na literatura. (a) é um exemplo de *inter-timeslice DBN*, ou seja, que permite arestas apenas saindo dos nós em $t - 1$ até os nós em t , e não entre nós no mesmo instante de tempo. Alguns autores como Smith et al. (2006) e Robinson e Hartemink (2010) admitem uma representação equivalente, apontada em (b), que consiste em um grafo direcionado cíclico. (c) é um exemplo de *DBN* segundo as definições originais, com arestas *intra* e *inter-timeslice*.

crucial na implementação de métodos de inferência e aprendizado dos mesmos. Os modelos *inter-timeslice* são mais próximos de casos reais, e lidam com situações puramente causais, onde o presente depende apenas do passado (Daly et al., 2009). Além disso, o grafo da rede de transição deste modelo é naturalmente ordenado e não necessariamente acíclico (o que exclui a necessidade de checar por ciclos nos algoritmos de aprendizado, proporcionando um melhor desempenho computacional). Por outro lado, os modelos que permitem arestas *intra* e *inter-timeslice*, embora mais complexos, são capazes de lidar com relacionamentos entre sinais sem a necessidade de uma causa direta presente no modelo (como no caso de uma leitura de sensor em um sistema robótico) — o que leva à conclusão que esses modelos também podem ajudar a encontrar variáveis ocultas em um sistema. A Figura 2.2 apresenta em (a) um modelo *inter-timeslice*, com sua representação equivalente em (b), adotada por alguns autores, e em (c) um modelo *intra* e *inter-timeslice*.

2.2 Aprendizagem estrutural

Uma vez que o conhecimento da estrutura e parâmetros de uma *DBN* nem sempre está disponível por parte de um especialista, a área de pesquisa em redes Bayesianas conta com vários métodos, baseados em aprendizagem de máquina, tendo em vista duas finalidades: a aprendizagem de parâmetros e a aprendizagem estrutural (Koller e Friedman, 2009). A aprendizagem de parâmetros é a que já conta com a estrutura de grafo da rede, e busca obter apenas

Figura 2.3: Classes de métodos de aprendizagem estrutural de redes Bayesianas.

as funções de distribuição de probabilidade para cada variável, de acordo com os nós pais e filhos descritos pelo grafo; já a aprendizagem estrutural busca obter a melhor ou as melhores estruturas de grafo da rede a partir dos dados.

A aprendizagem estrutural pode ser dividida em dois grupos de métodos: métodos baseados em testes de independência condicional e métodos baseados em pontuação. O primeiro consiste em realizar uma série de cálculos que pretendem elicitar as independências entre variáveis do sistema, e assim, construir a rede — para isso, por exemplo, são usadas medidas como desvio da hipótese nula ou informação mútua. Já o segundo grupo pode ser visto como um problema de seleção de modelo, e utiliza uma métrica que avalia o quão bem o modelo se adequa aos dados (ou seja, a pontuação). A Figura 2.3 resume as classes de métodos utilizados para o aprendizado estrutural de redes Bayesianas e *DBNs*. Nas seguintes seções, serão detalhadas estas duas classes de métodos, apontando-se suas vantagens e desvantagens.

2.2.1 Métodos baseados em pontuação

A aprendizagem estrutural de redes Bayesianas através de pontuação pode também ser definida como a tarefa de encontrar um grafo ou um conjunto de grafos que melhor se adequam ao conjunto de dados de treinamento D . Isso é feito através do cálculo da *posteriori* dos grafos dado o conjunto de dados, $P(G|D)$ (onde G é uma estrutura de grafo possível no espaço de grafos). Utilizando o teorema de Bayes, pode-se reescrever a expressão:

$$P(G|D) \propto P(D|G)P(G) \quad (2.3)$$

onde $P(G)$ é o conhecimento *a priori* sobre a estrutura do grafo, e precisa ser considerado para garantir um sentido inferencial ao resultado buscado; ou seja, o quão bem um grafo se adequa

aos dados e não o contrário (que conferiria um sentido dedutivo, apenas). A verossimilhança ou *likelihood* dos dados $P(D|G)$, por sua vez, pode ser calculada de diversas maneiras, através de várias suposições. Vários autores apresentam diversas propostas para realizar este cálculo, e no presente trabalho, destaca-se a proposta de Heckerman et al. (1995). Os autores mostram que, para dados discretos, considerando que as amostras são multinomiais, os parâmetros utilizados para as distribuições de probabilidade condicionais dos nós são modulares e independentes entre si, e as *prioris* seguem uma distribuição de Dirichlet, o *likelihood* pode ser calculado pela seguinte fórmula:

$$P(D|G) \propto \prod_{i=1}^N \prod_{j=1}^{q_i} \left(\frac{\Gamma(N'_{ij})}{\Gamma(N_{ij} + N'_{ij})} \times \prod_{k=1}^{r_i} \frac{\Gamma(N_{ijk} + N'_{ijk})}{\Gamma(N'_{ijk})} \right) \quad (2.4)$$

onde N é o número de nós no grafo, r_i é o número de estados discretos da variável X_i , q_i é o número de combinações de todos os estados discretos dos nós pais de X_i , e Γ é a função gama. N_{ijk} é uma estatística suficiente para as ocorrências, no conjunto de dados, da variável X_i com valor k e pais com valores j ; N'_{ijk} são hiperparâmetros ou estatísticas suficientes *a priori*, que são dadas a partir do conhecimento *a priori da rede*. Esta é chamada de métrica *Bayesian-Dirichlet (BD)*. No caso de não se possuir este conhecimento, escolhe-se *a priori* não informativa fazendo todos os valores N'_{ijk} iguais a 1; esta é chamada de métrica *Bayesian-Dirichlet equivalent (BDe)* (Heckerman et al., 1995).

Cooper e Herskovits (1992) mostram que é possível simplificar estes cálculos aplicando logaritmo a esta aproximação de $P(G|D)$, definindo a pontuação $K2$:

$$K2(D : G) = \sum_{i=1}^N \sum_{j=1}^{q_i} \left(\log \frac{(r_i - 1)!}{(N_{ij} + r_i + 1)!} + \sum_{k=1}^{r_i} \log(N_{ijk}!) \right) + \log(P(G)) \quad (2.5)$$

Uma vez com uma função de pontuação, pode-se escolher entre buscar a rede com a melhor pontuação do espaço de busca, ou uma estimativa *a posteriori* das redes com as melhores pontuações (Robinson, 1977). No primeiro caso, tem-se um método de busca, enquanto no segundo, um método de amostragem (ou também definido como *BMA – Bayesian Model Averaging* (Koller e Friedman, 2009)). Obviamente os métodos de busca são muito mais rápidos que os de amostragem, no entanto, ocultam informações que podem ser úteis na decisão de um modelo por parte de um especialista. Os métodos de busca usam buscas exatas ou heurísticas (esta última no caso da busca exata se tornar intratável computacionalmente, devido a muitas variáveis), enquanto os de amostragem baseiam-se em métodos de *MCMC (Markov Chain*

Monte Carlo) para o cálculo de probabilidades *a posteriori* no espaço de busca (Madigan et al., 1995).

Dentre as pontuações propostas para a aprendizagem estrutural de *DBNs*, destaca-se o conhecido trabalho de Friedman et al. (1998), que define pontuações baseadas em *BIC* (*Bayesian Information Criterion*) e *BDe*, ambas para a rede de transição como para a rede de *prioris*. Também é importante notar que, neste caso, algumas restrições devem ser impostas nos espaço de modelos possíveis: os *DAGs* não podem possuir arestas saindo dos nós no instante t aos nós no instante $t - 1$, nem arestas entre os nós no instante $t - 1$ (Pérez-Ariza et al., 2012). Esta constatação leva também à possibilidade de simplificação na função de pontuação em *DBNs*, uma vez que, nestes modelos, apenas os nós no instante t possuem pais. Assim, a estatística suficiente para os nós que não pertencem ao instante t será a mesma, independentemente da estrutura do grafo; e como a tarefa de aprendizagem estrutural consiste em obter a melhor ou melhores pontuações de grafo, os nós podem ser retirados da função de pontuação, tornando os cálculos mais simples e melhorando o desempenho computacional.

2.2.2 Métodos baseados em testes de independência condicional

O objetivo dos métodos de aprendizagem estrutural baseados em testes de independência condicional é descobrir as independências condicionais entre os nós do sistema através de medidas de desvio de hipótese nula ou informação mútua e, assim, propor a estrutura de rede que mais se adeque a essas independências. A desvantagem desse tipo de método, conforme argumenta Koller e Friedman (2009, p. 785), encontra-se no fato de que, caso haja um erro ao encontrar-se uma independência, toda a estrutura de rede proposta é comprometida. Não obstante, Villanueva Talavera (2012) também ressalta que "a confiabilidade dos testes estatísticos cai rapidamente à medida que a ordem do conjunto aumenta", levando a um aumento exponencial do número de amostras necessárias para se manter a mesma confiabilidade). Os métodos baseados em pontuação, por sua vez, não possuem este problema. Por outro lado, a vantagem dos métodos baseados em testes de independência condicional está naquilo que já foi observado na introdução: trata-se de medidas funcionais sobre os dados e, que, por isso, não exigem que o sistema seja observado completamente em um só experimento para ser modelado.

Vários algoritmos também são propostos para se obter a estrutura da rede a partir dos testes de independência condicional ou teoria de informação. Entre eles pode-se citar o *Sparse Candidate Algorithm* por Friedman et al. (1999), o *max-min hill-climbing* por Tsamardinos

et al. (2006), e recentemente, o trabalho de Kojima et al. (2010), que formaliza o conceito de superestrutura de uma rede Bayesiana — um grafo que, a partir dos testes de independência condicional, codifica todas as restrições estruturais na rede Bayesiana que se busca encontrar (ou seja, arestas que não são permitidas no resultado final). O recente trabalho de Villanueva e Maciel (2014) também propôs alguns métodos eficientes para se obter essas superestruturas. Por fim, todos estes algoritmos têm como pano de fundo alguns algoritmos de aprendizagem de rede tradicionais como o SGS (nomeado de acordo com os criadores, Spirtes, Glymour e Scheines) (Spirtes et al., 1990) e o PC (Spirtes e Glymour, 1991).

2.2.3 Métodos híbridos

As vantagens e desvantagens em ambas as classes de métodos de aprendizado estrutural de redes Bayesianas também levou, naturalmente, ao desenvolvimento de métodos híbridos. Isso é feito normalmente utilizando-se os testes de independência condicional para se obter a superestrutura da rede, e em seguida utilizando-se pontuações para encontrar uma rede, dadas as restrições estruturais. Este é o caso, por exemplo, dos trabalhos de Singh e Valtorta (1995), Wang et al. (2007) e Villanueva Talavera (2012) (com este último realizando uma revisão mais detalhada de todos estes métodos).

Outra interessante forma de unir os dois tipos de métodos de aprendizagem é a proposta de uma pontuação baseada em medidas de informação mútua (MI), no trabalho de Campos (2006), chamada MIT (para redes Bayesianas estáticas), com sua adaptação para DBNs intitulada GlobalMIT (Vinh et al., 2011b). Tal proposta consegue garantir, assim, a vantagem de ambos os métodos, e será detalhada no próximo capítulo (3), tratando do uso de teoria de informação no contexto de DBNs.

2.2.4 Aprendizagem em nível de grupo

Em muitos tipos de aplicação de DBNs — principalmente em biomedicina (Ide et al., 2014b) —, várias repetições de gravações dos dados são realizadas em situações ou indivíduos diferentes a fim de se obter uma melhor confirmação do modelo, ou mesmo investigar alguma característica comum ou diferente entre os mesmos. Com isso, torna-se necessário o desenvolvimento de abordagens que analisem dados de diferentes experimentos, indicando o grau de certeza do modelo e as possíveis semelhanças ou diferenças. Tal tipo de procedimento é

chamado de análise em nível de grupo (no inglês, *group level analysis*).

Li et al. (2008) apresenta uma revisão dos principais métodos de análise em nível de grupo, classificando-os em três grupos. O autor também prova que nenhuma das abordagens é universalmente superior a outra, e que portanto, sua escolha depende do tipo de aplicação buscada. São elas:

- Indivíduo virtual típico (*virtual-typical-subject* ou VTS); que reúne ou obtém a média dos dados do grupo, tratando-os como se fossem obtidos de um único indivíduo hipotético virtual;
- Estrutura individual (*individual structure* ou IS): obtém uma estrutura de rede diferente para cada indivíduo, e então busca características em comum entre elas; e
- Estrutura comum (*common structure* ou CS): impõe a mesma estrutura de rede a todos os indivíduos, mas permite que os parâmetros variem entre os mesmos.

2.3 DBNs em sinais não-estacionários

Um dos desafios encontrados na aplicação de modelos de DBN concerne a classe de sistemas com comportamento não-estacionário. Um processo estocástico é dito estacionário se suas propriedades estatísticas permanecem invariantes ao deslocamento temporal (Papoulis e Pillai, 2002). Tal não é o caso em muitas gravações de sinais neurais — um exemplo disso é o estudo de Smith et al. (2006), sobre o fluxo de informação neural nos cérebros de aves canoras (no inglês, songbirds). Quando um pássaro desta espécie nasce, ele não é capaz de proferir nenhum som, mas à medida em que ouve o som de outros pássaros da mesma espécie, a informação neural se propaga e este aprende a cantar conforme o ambiente inserido. Neste caso há uma variação, no decorrer do tempo, nas características do sistema neural. Assim, este capítulo apresentará um resumo das recentes considerações, definições e técnicas têm sido sugeridas para se lidar com isso, tanto na área de medidas de transferência de informação em neurociência, como na de modelagem com *DBNs* (uma vez que ambas contam com a suposição de estacionariedade nos modelos).

A princípio, a forma mais comum de se lidar com sinais não-estacionários é repetir um experimento várias vezes e realizar-se medidas nos dados em mesmos instantes de tempo (métodos de *ensemble*, por exemplo, são eficientes neste objetivo (Gómez-Herrero et al., 2015)).

No entanto, tais situações podem ser difíceis em alguns sistemas nos quais se dispõe de poucos experimentos e amostras. Por esta razão, pode-se tentar fazer algumas suposições sobre o sistema, como, por exemplo, a consideração de sistema estacionário em partes e ergódico — ou seja, cujas amostras em intervalos fixos de tempo podem ser usadas para inferir as propriedades estatísticas.

Assim, pode-se fazer algumas perguntas sobre sistemas que geram sinais não-estacionários, que guiarão as decisões em direção a uma solução de modelagem:

- O mecanismo que produz a não-estacionariedade nos sinais faz parte do comportamento que se deseja modelar; ou seja, é um mecanismo exógeno ou endógeno ao sistema? No caso de mecanismos endógenos, há uma grande dificuldade caso o sistema seja grande e pouco observável, além da necessidade de muitas amostras, e sincronizadas entre si, para a reconstrução da dinâmica interna do sistema.
- Caso o mecanismo que produz a não-estacionariedade seja exógeno, deve-se avaliar se esta estacionariedade fará parte do modelo buscado. A não-estacionariedade nos sinais pode ser desconsiderada caso seja resultante de ruído ou outros efeitos do experimento que não são inerentes ao sistema — neste caso, pode-se usar técnicas de remoção de ruído e tendências nos dados, e em seguida obter-se o modelo a partir dos sinais estacionários resultantes.
- Caso a estacionariedade deva ser incorporada ao modelo, porém ainda permaneça exógena ao sistema, deve-se então perguntar: como são as variações das propriedades estatísticas dos sinais no tempo: poucas e abruptas ou muitas e graduais? No primeiro caso pode-se entender o processo não-estacionário como um conjunto de processos estacionários menores separados por *change-points* (ou seja, instantes de tempo nos quais as propriedades do sistema mudam). Assim, uma vez também com um tempo suficiente entre estes *change-points*, pode-se inclusive garantir a *ergodicidade* do sistema. Já no segundo caso a situação torna-se mais complicada — mudanças graduais indicam a presença de memória no sistema, e uma modelagem desses sinais exigiria diversas repetições do experimento, sincronizadas no tempo, para conseguir-se inferir as propriedades estatísticas.

A Figura 2.4 ilustra as três perguntas apresentadas acima e as considerações sobre modelos, métodos e dados exigidos a partir de suas respostas. Os sinais gravados no sistema neuromotor

Figura 2.4: Fluxograma das considerações a serem tomadas na modelagem com sinais não-estacionários, juntamente com os tipos de modelo resultantes dessas considerações.

proprioceptor do gafanhoto apresentados na Figura 2.5 constituem uma boa ilustração para as três considerações.

Pode-se também entender os principais modelos de *DBNs* não-estacionárias propostos na literatura conforme as considerações acima. No caso, tais modelos se aplicam a situações onde a não-estacionariedade é exógena ao sistema, e esta faz parte do sistema que se deseja modelar (ou seja, não é ruído). Portanto, pode-se agrupar os modelos em duas classes: os modelos baseados em *change-points*, como nos trabalhos de Robinson e Hartemink (2010), Grzegorzcyk e Husmeier (2011) e Lebre et al. (2010), e os modelos baseados em parâmetros variando em função do tempo, como no trabalho de Song et al. (2009) (neste caso, porém, os autores consideram processos lineares e autoregressivos). Já os trabalhos de Grzegorzcyk et al. (2008) e Ko et al. (2009) apresentam outro tipo de solução, introduzindo variáveis latentes no modelo Bayesiano, que influenciam as *CPDs* dos outros nós (ou seja, como se o tempo fosse modelado como uma variável do sistema). Esta última abordagem também é interessante pois transforma um mecanismo exógeno ao sistema em um endógeno, através de métodos de aprendizagem de redes a partir dos dados.

Figura 2.5: Ilustração para as três considerações sobre sinais não-estacionários, utilizando um sinal gravado nos neurônios do sistema neuromotor proprioceptor do gafanhoto. O sinal superior é o estímulo no *apódema* do animal — um ruído branco gaussiano de 200 Hz, e o inferior é a resposta em um interneurônio *spiking*. O caso (1) ilustra uma situação de um mecanismo endógeno no sistema, onde a presença do estímulo ocasiona uma inibição gradual do trem de *spikes*. O caso (2) ilustra a presença de uma não-estacionariedade irrelevante ao modelo, uma vez que é uma flutuação resultante da introdução do eletrodo no neurônio do animal. Já o caso (3) ilustra um sinal não-estacionário que pode ser dividido em trechos estacionários ergódicos.

Algumas outras considerações podem ser notadas nos modelos propostos:

1. **Escopo dos *change-points*:** Nos trabalhos que consideram *change-points*, estes podem ser considerados como presentes em todo o sistema (trabalhos de Robinson e Hartemink (2010) e Grzegorzczuk et al. (2008)), tal como específico de cada nó na rede (trabalhos de Grzegorzczuk e Husmeier (2011), Ko et al. (2009) e Lebre et al. (2010)).
2. **Mudanças de parâmetros ou estrutura:** A não-estacionariedade dos modelos pode ser interpretada como uma mudança apenas nos parâmetros da rede (ou seja, que controlam as distribuições de probabilidade condicional) (caso de Grzegorzczuk et al. (2008), Grzegorzczuk e Husmeier (2011)), ou como mudanças tanto na estrutura como nos parâmetros da rede (todos os outros trabalhos citados). Grzegorzczuk e Husmeier (2011) argumenta que, no caso de redes de regulação de genes, um modelo que permite mudança de estruturas pode apresentar uma super flexibilidade, o que para séries temporais curtas, leva inevitavelmente a um *overfitting* dos dados ou uma inflação da incerteza da inferência.
3. **Pontuação e forma dos dados:** Alguns autores desenvolvem pontuações para avaliar os modelos de *DBNs* não-estacionários. Os trabalhos de Robinson e Hartemink (2010) e Grzegorzczuk e Husmeier (2011) são complementares pois apresentam, respectivamente, uma pontuação baseada em *BDe* e uma baseada em *BGe* (*Bayesian Gaussian equivalent*),

que é semelhante à pontuação *BDe*, porém adaptada para domínios de dados contínuos (Geiger e Heckerman, 1994). Ko et al. (2009), por sua vez, apresenta uma pontuação baseada em *BIC*. Somente o trabalho de Robinson e Hartemink (2010) lida com dados discretos.

Também importante notar que os modelos propostos são direcionados primariamente à modelagem de redes de regulação de genes, com exceção do trabalho de (Smith et al., 2006). Não obstante, todos eles são, conforme indicado anteriormente (seção 2.1), modelos *inter-timeslice*, ou seja, que permitem arestas apenas saindo dos nós em $t - 1$ até os nós em t .

2.3.1 *Detrending* com Singular Spectrum Analysis

Os sinais que serão analisados no presente trabalho (seções 5.2 e 5.3) apresentam características de não-estacionariedade que são quase sempre consideradas como irrelevantes ao modelo (ou seja, com um mecanismo exógeno, como ruído ou capacitâncias parasitas no momento da aquisição) (Dewhirst et al., 2013; Ide et al., 2014a). No entanto, acredita-se que métodos de *detrending* linear podem não remover completamente a não-estacionariedade de muitos desses sinais, uma vez que podem apresentar características não-lineares e/ou descontinuidades (vide, por exemplo, o sinal de potencial sináptico no neurônio motor do gafanhoto, na Figura 6.9). A aplicação de filtros passa alta, como os discutidos por Barnett e Seth (2011), também nem sempre é uma opção viável, devido à necessidade da especificação de parâmetros, como frequência de corte (entre outros), que podem ser arbitrários e variar de sinal a sinal. Ademais, o estudo por Florin et al. (2010) sugere que a filtragem de dados neurais distorce a ordem da série temporal (devido à resposta de fase), e que isso pode influenciar medidas baseadas em causalidade de Granger e de transferência de entropia atrasada (esta última sendo uma versão não-linear da causalidade de Granger, e será apresentada na seção 3.1).

A análise por *Singular Spectrum Analysis* (SSA) surge como uma abordagem livre de modelo viável para remoção de tendências não-lineares e complexas (incluindo descontinuidades) em séries temporais (Golyandina e Zhigljavsky, 2013). O método decompõe a série em uma soma de componentes independentes e interpretáveis como tendências que variam devagar, componentes oscilatórios e ruído não-estrutural (Hassani, 2007) e permite sua reconstrução de acordo com as componentes desejadas. No caso, as tendências podem ser removidas reconstruindo o sinal sem a componente associada ao maior autovalor da matriz de *embedding* do sinal (a ser definida a seguir). Isso se baseia em Palus e Novotná (1998), que mostraram que

a decomposição por SSA em componentes ortogonais com diferentes propriedades dinâmicas possui os maiores autovalores relacionados aos modos lentos, enquanto os menores autovalores estão relacionados às altas frequências. Assim, o maior autovalor costuma ser associado às tendências complexas e descontinuidades no sinal.

O algoritmo de SSA é descrito da seguinte forma (Golyandina e Zhigljavsky, 2013): considere-se uma série temporal $Y_T = (y_1, \dots, y_T)$. A matriz de *embedding* de Y_T é obtida através da expressão:

$$\mathbf{X} = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_K \\ y_2 & y_3 & y_4 & \cdots & y_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \cdots & y_T \end{pmatrix}$$

onde $K = T - L + 1$ e L são os tamanhos de janela considerados para o cálculo ($L \leq T/2$). É importante notar que a matriz $\mathbf{X}$ é uma matriz de Hankel, com elementos iguais nas anti-diagonais, o que permite a sua decomposição em valores singulares (SVD) da matriz $\mathbf{X}\mathbf{X}'$, de forma que esta é representada na forma $\mathbf{X}\mathbf{X}' = P\Lambda P'$, onde $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_L)$ é a matriz diagonal dos autovalores de $\mathbf{X}\mathbf{X}'$ em ordem decrescente de magnitude ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L$) e $P = (P_1, P_2, \dots, P_L)$ é a matriz ortogonal correspondente com os autovetores de $\mathbf{X}\mathbf{X}'$. Também define-se d como:

$$d = \max(i, \text{tal que } \lambda_i > 0) = \text{rank}\mathbf{X}$$

(Embora, com dados reais, geralmente se tem $d = \min\{L, K\}$). Os componentes principais (PCs) V_i ($i = 1, \dots, d$) da matriz de *embedding* são então obtidos pela expressão $V_i = \mathbf{X}'P_i/\sqrt{\lambda_i}$ e, assim, podem ser representados como:

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \dots + \mathbf{X}_d$$

where $\mathbf{X}_i = \sqrt{\lambda_i}P_iV_i'$ ($i = 1, \dots, d$).

Estes componentes principais podem ser escolhidos de acordo com as propriedades desejadas e projetados de volta para as coordenadas originais da série temporal, um procedimento intitulado *reconstrução de componentes* (*component reconstruction*). Isto é feito selecionando-

se e particionando os índices $i = 1, \dots, d$ em subconjuntos disjuntos $I_1, I_2, \dots, I_m$. Assim, para um dado subconjunto $I = \{i_1, i_2, \dots, i_L\}$ a matriz resultante $\mathbf{X}_I$ é definida como $\mathbf{X}_I = \mathbf{X}_{i_1} + \mathbf{X}_{i_2} + \dots + \mathbf{X}_{i_L}$. Isto também mostra que a decomposição em valores singulares pode ser representada como $\mathbf{X} = \mathbf{X}_{I_1} + \mathbf{X}_{I_2} + \dots + \mathbf{X}_{I_m}$.

O procedimento de média diagonal (*diagonal averaging*) é, então, aplicado à matriz $\mathbf{X}_{I_k}$, resultando na série temporal reconstruída $\tilde{\mathbf{X}}^{(k)} = (\tilde{x}_1^{(k)}, \tilde{x}_2^{(k)}, \dots, \tilde{x}_N^{(k)})$. Dessa forma, a série original é decomposta em uma soma de m sub-séries reconstruídas:

$$x_n = \sum_{k=1}^m \tilde{x}_n^{(k)} (n = 1, 2, \dots, N) \quad (2.6)$$

Com isso, a série temporal é reconstruída de acordo com as componentes desejadas — e tal será o caso na remoção de tendências nas análises dos sinais de fMRI e de potencial sináptico nas seções 5.2.2 e 5.3.2.

Capítulo 3

Usos de teoria de informação em DBNs

Dada a revisão conceitual de DBNs realizada no capítulo 2, este capítulo apresenta as principais formas nas quais o uso de medidas de transferência de informação têm auxiliado a identificação de sistemas com DBNs, bem como considerações acerca da integração destas duas áreas. Antes disso, porém, uma revisão sobre medidas baseadas em teoria de informação é apresentada na seção 3.1, juntamente com uma discussão sobre os desafios de aplicação de teoria de informação em sistemas multivariados e o problema da "maldição da dimensionalidade" em sistemas com muitos nós e/ou parcialmente observados. Em seguida, apresenta-se uma descrição da pontuação GlobalMIT, proposta por (Vinh et al., 2011b), para a aprendizagem de DBNs com informação mútua condicional. A seção 3.3 apresentará a suposição de redes Bayesianas com nós de mesmo tipo como uma simplificação importante para a aprendizagem de BNs e DBNs com teoria de informação em sistemas grandes e/ou parcialmente observados. Por fim, a presença de diferentes atrasos de tempo em um sistema, já obtidos por meio de medidas como informação mútua atrasada e transferência de entropia, exige uma consideração sobre sua representação e aprendizagem em DBNs. Para tal, a proposta de redes Bayesianas de altas ordens (HO-DBNs) é apresentada e discutida, juntamente com um algoritmo baseado em PSO para percorrer seu espaço de busca. Este será o conteúdo da seção 3.4

3.1 Definições

Seja $\mathbf{X} = \{X, Y, Z, \dots\}$ um conjunto de processos estocásticos estacionários e $\mathbf{X}_t, X_t, Y_t, \dots$ os seus valores no discretizados nos instantes de tempo t . Os alfabetos (valores possíveis) das variáveis aleatórias do processo são os conjuntos $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ e suas funções de massa de probabilidade são (por fins de conveniência) $p(x), p(y), p(z), \dots$. Define-se a *entropia de Shannon* H de uma variável aleatória X_t como (Cover e Thomas, 2006):

$$H(X_t) = - \sum_{\forall x \in \mathcal{X}} p(x) \log p(x) \quad (3.1)$$

(sendo o logaritmo de base dois e a entropia expressa em bits). Esta última medida define a quantidade média de informação, em bits, necessária para descrever a variável aleatória X_t . Da mesma forma, a *entropia de Shannon condicional* $H(X_t|Y_t)$ define a quantidade de informação, em bits, necessária para descrever X_t uma vez que se conheça Y_t , sendo calculada por:

$$H(X_t|Y_t) = - \sum_{\substack{\forall x \in \mathcal{X} \\ \forall y \in \mathcal{Y}}} p(x, y) \log p(x|y) \quad (3.2)$$

em que $p(x, y)$ é a função de massa de probabilidade conjunta das variáveis aleatórias X_t e Y_t . Assim, segue que a *informação mútua* I entre X_t e Y_t é dada pela diferença entre a entropia de X_t e a entropia de X_t condicional a Y_t — significando, assim, a redução da incerteza de X_t devido ao conhecimento de Y_t e vice-versa (pela propriedade de simetria):

$$I(X_t; Y_t) = H(X_t) - H(X_t|Y_t) = H(Y_t) - H(Y_t|X_t) = \sum_{\substack{\forall x \in \mathcal{X} \\ \forall y \in \mathcal{Y}}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (3.3)$$

Por fim, a *informação mútua condicional* (ou *parcial*, quando entendida em relação à correlação parcial (Frenzel e Pompe, 2007)) $I(X_t, Y_t|Z_t)$ define a redução da incerteza de X_t

devido ao conhecimento de Y_t quando Z_t é dado:

$$I(X_t; Y_t | Z_t) = H(X_t | Z_t) - H(X_t | Y_t, Z_t) = \sum_{\substack{\forall x \in \mathcal{X} \\ \forall y \in \mathcal{Y} \\ \forall z \in \mathcal{Z}}} p(x, y, z) \log \frac{p(x, y | z)}{p(x | z) p(y | z)} \quad (3.4)$$

Dadas as medidas acima, pode-se começar a definir as medidas de transferência de informação utilizadas na área de neurociência computacional, ou mesmo em aplicações de sistemas dinâmicos em geral (Wibral et al. (2014b) realiza uma revisão das propostas). A principal e mais usada delas é a transferência de entropia (TE, ou *transfer entropy*), proposta de formas independentes por Schreiber (2000) e Paluš et al. (2001). Esta é definida como a medida da "quantidade de informação que o estado passado de um processo X contém sobre uma futura observação de um processo Y dado que já se sabe o estado passado do processo X " (Wibral et al., 2014a). Assim, em termos de informação mútua condicional, tem-se:

$$TE(X \rightarrow Y) = I(\mathbf{X}^-, Y^+ | \mathbf{Y}^-) \quad (3.5)$$

onde $\mathbf{X}^-$ e $\mathbf{Y}^-$ são o estado passado dos processos X e Y . Esta definição é, obviamente, ainda indefinida — resta saber como se interpretar e representar os estados passados de X e Y . Schreiber (2000) considera uma propriedade generalizada de Markov, onde o valor presente de Y depende apenas de k estados passados de Y e de l estados passados de X , possibilitando o cálculo da TE como:

$$TE(X \rightarrow Y) = \sum_{\substack{\forall x \in \mathcal{X} \\ \forall y \in \mathcal{Y}}} p(y_n, \mathbf{y}_n^{(k)}, \mathbf{x}_n^{(l)}) \log \frac{p(y_n | \mathbf{y}_n^{(k)}, \mathbf{x}_n^{(l)})}{p(y_n | \mathbf{y}_n^{(k)})} \quad (3.6)$$

onde $\mathbf{y}_n^{(k)} = \{y_{n-1}, \dots, y_{n-k}\}$ e $\mathbf{x}_n^{(l)} = \{x_{n-1}, \dots, x_{n-l}\}$ (considerados em tempo discreto).

Assim, o cálculo pode ser feito a partir dos dados, estimando-se as probabilidades exigidas na equação acima (o que também pode exigir a suposição de ergodicidade no sistema, para o uso de amostras no tempo; e caso isso não seja possível, deve-se usar diversas repetições do experimento (Wollstadt et al., 2014)). Os autores Wibral et al. (2014a), por sua vez, propõem a *estimação* da TE a partir de uma reconstrução do espaço de estados a partir do teorema de *delay embedding* de Takens. O problema com esta abordagem, no entanto, conforme os próprios

autores apontam, é que a presença de ruído colorido nas amostras pode provocar desvios e tendências na reconstrução dos estados (Kantz e Schreiber, 2004), e diante da falta de uma abordagem analítica a este problema, observa-se apenas que a abordagem é "*consideravelmente robusta sob ruído branco Gaussiano*" e que "simulações com outros ruídos típicos em sinais neurofisiológicos *parecem* garantidas" Wibral et al. (2014a, ênfase do autor).

Também é importante notar que medidas sobre dados obtidos de processos nunca revelarão independências completas, uma vez que os dados são finitos. Medidas de transferência de informação normalmente apresentam valores diferentes de zero mesmo que não exista qualquer conexão entre dois processos em um sistema — ou seja, há um *bias* no estimador (Wibral et al., 2014b). Várias técnicas foram desenvolvidas para a correção do bias do estimador da TE — por exemplo, a correção do tipo Miller e Madow (Wibral et al., 2014b) — no entanto, para fins de testes de significância das medidas, a ferramenta mais comum é a geração de *surrogate data* (Schreiber e Schmitz, 2000), ou seja, séries temporais simuladas que buscam representar o máximo possível as características dos processos reais, porém sem qualquer transferência de informação, a fim de fornecer uma linha de base para as medidas. No caso de aplicações em neurociência, nota-se que a coerência entre os sinais acontece na sincronização das fases (Yang et al., 2013), portanto o trabalho de Endo (2014) recentemente aponta que uma forma interessante para a geração de *surrogate data* para sinais neurais é através do método de embaralhamento de fases proposto por (Schreiber e Schmitz, 2000), chamado também de *IAAFT* (*Iterative Amplitude Adjusted Fourier Transform*), que gera sinais com as mesmas amplitudes e distribuição de densidade espectral de potência (Venema et al., 2006), porém com os componentes de fase fora de ordem (eliminando, assim, a transferência de informação). Uma outra alternativa mais simples para a geração de *surrogates* envolve simplesmente o embaralhamento das amostras da série temporal, destruindo, assim, a estrutural de precedência temporal (Wibral et al., 2014b). Dolan e Spano (2001), no entanto, observa que, uma vez que este método preserva as mesmas amplitudes e frequências de amostras dos dados originais, ele pode não prover a variabilidade nos dados necessária para a estimação das distribuições de probabilidade (e, assim, não capturar o *bias*).

Informação multivariada e a "maldição da dimensionalidade"

As definições acima levantam a questão sobre o uso de medidas de teoria de informação em sistemas multivariados — sabe-se que Shannon, originalmente, trabalhou caracterizando a

informação mútua apenas entre dois processos (Cover e Thomas, 2006), e pouco foi explorado sobre a informação entre três ou mais. Diante disso, vários autores propuseram medidas como correlação total (Watanabe, 1960), multiinformação (Studený e Vejnarová, 1998) e informação de interação ou co-informação (McGill, 1954). No entanto, conforme ressaltou (Williams e Beer, 2010), estas medidas pouco indicam algo sobre uma *estrutura* da informação, de forma a permitir interpretações significativas sobre as interações entre processos. A informação de interação, por exemplo, pode assumir até mesmo valores negativos — uma propriedade interessante, a qual a literatura tem buscado interpretar, mas que no entanto, é ao mesmo tempo considerada não muito útil (Cover e Thomas, 2006) e pouco intuitiva (Williams e Beer, 2010). Assim, o recente trabalho de Williams e Beer (2010) lançou luz neste problema ao propor uma decomposição não-negativa da informação multivariada, em uma malha de contribuições de acordo com propriedades definidas como *sinergia*, *redundância* e *informação única*. E embora recentemente Rauh et al. (2014)) tenha apontado que esta decomposição apresenta algumas falhas, ela ainda permanece com uma excelente base de interpretação para informação multivariada e, conforme mostra a literatura recente, funciona de forma adequada em casos de três variáveis (Timme et al., 2014; Bertschinger et al., 2014; Rauh et al., 2014).

Os conceitos de sinergia, redundância e informação única são úteis uma vez que orientam todos os cuidados tomados na proposição de algoritmos para a reconstrução correta de grafos a partir de medidas de teoria de informação. Medidas par-a-par não consideram a sinergia e a redundância entre processos, e por este motivo podem adicionar ou remover arestas de um grafo de forma inadequada. Os trabalhos de Wibral et al. (2014a); Lizier et al. (2013) citam tais problemas, os quais podem ser classificados conforme a lista a seguir:

- *Informação redundante*: trata-se de uma mesma informação sendo transmitida ou recebida por dois ou mais processos, a qual pode produzir:

Efeitos de cascata: por exemplo, um processo A transfere informação para um processo B que, por sua vez, transfere para outro processo C. Caso parte da informação de B para C também seja a informação passada de A para B, uma medida par-a-par de transferência de informação de A para C revelaria resultados positivos, mesmo que uma aresta direta entre A e C não exista.

Efeitos de controlador comum: por exemplo, um processo A transfere informação para dois processos B e C. Caso a informação transferida por A (ou parte dela) seja a mesma, uma medida par-a-par de transferência de informação entre B e C também

revelaria resultados positivos, embora não exista uma aresta direta entre B e C.

- *Informação sinérgica*: trata-se da informação sendo transmitida ou recebida somente quando considera-se dois ou mais processos conjuntamente — ou seja, as contribuições de informação por dois ou mais processos não pode ser considerada como uma simples soma ou mesmo união das informações individuais. Por exemplo, dois processos A e B transferem informação para um processo C. Caso exista sinergia entre A e B na relação com C, medidas par-a-par de transferência de informação de A para C e de B para C revelariam um resultado igual a zero em ambos os casos, enquanto que existem, de fato, arestas diretas de A para C e de B para C.

Nota-se que, nos algoritmos de reconstrução de grafos a partir de medidas par-a-par, a informação redundante entre dois processos pode ocasionar a *adição de arestas inexistentes*, enquanto que a informação sinérgica pode ocasionar a *remoção de arestas existentes*.

Assim, por estes motivos, os algoritmos de aprendizagem estrutural de redes Bayesianas e DBNs baseada em teoria da informação (seção 3.2, a seguir), ou mesmo os recentes trabalhos em medidas de transferência de informação, têm ressaltado a importância da *informação mútua condicional* como a forma de capturar as reais conexões entre os processos do sistema (Faes et al., 2011; Lizier et al., 2013; Runge, 2014). No entanto, com isso surge um desafio devido ao aumento das dimensões das funções de distribuição de probabilidade, exigindo também um maior número de amostras para sua estimação e uma maior capacidade computacional — um problema chamado de "*maldição da dimensionalidade*" (Scott e Sain, 2005). Exemplo disso é a proposta de Lizier et al. (2008) de uma medida de transferência de entropia condicional, a qual, dependendo do tamanho do sistema, pode precisar de estimação de probabilidades com dimensionalidades extremamente altas:

$$TE(X \rightarrow Y|Z) = I(\mathbf{X}^-, Y^+ | \mathbf{Y}^-, Z^-) \quad (3.7)$$

Tal situação também é o caso nos sistemas chamados de parcialmente observados, nos quais os cientistas não conseguem coletar amostras para cada variável em relação a cada outra (que é o caso de várias aplicações de neurociência), nos quais, portanto, é impossível uma análise da informação multivariada a partir de medidas de informação mútua condicional ou decomposição em elementos de sinergia, redundância e informação única contendo todos os nós. Tais sistemas também podem ser grandes e complexos, possuindo de centenas a milhões

de elementos, de forma que é impensável considerar cada um deles individualmente como uma variável interna de um modelo. E este também é o caso da aplicação que se busca neste trabalho, apresentada na seção 5.3, na qual se dispõe de gravações em pares ou, no máximo, trios de eletrodos, em um sistema com vários neurônios os quais sequer são identificados individualmente durante os experimentos.

Diante destes problemas, várias soluções têm sido propostas, como a limitação no número de pais para cada nó (Campos, 2006), suposição de grafos esparsos (Runge, 2014), inferências aproximadas (Faes et al., 2011; Stramaglia et al., 2012), e até a proposição de medidas de informação locais (Lizier, 2014). Algumas soluções serão revistas nas seções seguintes (3.3 e 3.4).

3.2 Aprendizagem de DBNs pela pontuação GlobalMIT

Na seção 2.2.3 foi citada a proposta de aprendizagem de redes Bayesianas por meio de uma pontuação baseada em informação mútua intitulada MIT, proposta por Campos (2006), e que é capaz de aliar vantagens de métodos baseados em pontuação e de testes de independência condicional. Esta métrica de pontuação avalia o quão bem uma rede se adequa aos dados observados de acordo com a "informação mútua total compartilhada entre cada nó e seus pais, penalizada por um termo que quantifica o grau de significância estatística desta informação compartilhada" (Xuan et al., 2012a), e foi mostrada como competindo no mesmo nível de outras pontuações (em eficiência e performance) como BDe e BIC. (Campos, 2006).

Dado um grafo candidato G , com N nós cujos conjunto de estados discretos é $\{r_1, \dots, r_n\}$ (ou seja, n estados discretos), e um conjunto de dados D com N_D observações, a pontuação MIT é definida como (Campos, 2006):

$$S_{MIT}(G : D) = \sum_{\substack{i=1 \\ \mathbf{Pa}(X^{(i)}) \neq \emptyset}}^N \{2N_D \cdot I(X^{(i)}, \mathbf{Pa}(X^{(i)})) - \sum_{j=1}^{s_i} \chi_{\alpha, l_{i\sigma(j)}}\} \quad (3.8)$$

onde $I(X^{(i)}, \mathbf{Pa}(X^{(i)}))$ é a informação mútua entre o nó $X^{(i)}$ e seus pais, e $\chi_{\alpha, l_{i\sigma(j)}}$ é o valor tal que $p(\chi^2(l_{ij}) \leq \chi_{\alpha, l_{ij}}) = \alpha$ (ou seja, a distribuição qui-quadrado a um nível de significância

igual a $1 - \alpha$, e o termo $l_{i\sigma_i(j)}$ definido como:

$$l_{i\sigma_i(j)} = \begin{cases} (r_i - 1)(r_{i\sigma_i(j)} - 1) \prod_{k=1}^{j-1} r_{i\sigma_i(k)}, & j = 2, \dots, s_i \\ (r_i - 1)(r_{i\sigma_i(j)} - 1), & j = 1 \end{cases} \quad (3.9)$$

onde $\sigma_i = \{\sigma_i(1), \dots, \sigma_i(s_i)\}$ é qualquer permutação do conjunto de índices $\{1, \dots, s_i\}$ de $\mathbf{Pa}(X^{(i)})$, com a primeira variável possuindo o maior número de estados; a segunda, o segundo maior, e assim por diante. Nota-se que a equação 3.8 apresenta um cálculo de informação mútua condicional envolvendo vários nós do sistema, o que pode apresentar um problema quanto à dimensionalidade e quantidade de dados exigida para a estimação das probabilidades – que geralmente é resolvido ao se limitar o número de pais possíveis para um determinado nó. Algumas possíveis soluções para este problema serão tratados nas próximas seções.

Vinh et al. (2011b) adapta esta pontuação para o caso de DBNs *inter-timeslice*, e propõe um algoritmo de busca operando em tempo polinomial. O algoritmo GlobalMIT investiga, para cada nó, cada conjunto de pais potenciais, com cardinalidade crescente, até que a solução globalmente ótima seja encontrada. Isso garante uma complexidade de tempo polinomial no pior caso, dada a suposição de DBNs *inter-timeslice* (eliminando, assim, a necessidade de se checar a aciclicidade dos grafos candidatos) e também cujas variáveis tenham o mesmo número de estados discretos (Vinh et al., 2011a). GlobalMIT é também um algoritmo determinístico, e portanto tem vantagem sobre outros métodos de otimização global estocástica, uma vez que não corre o risco de ficar preso em mínimos locais (Vinh et al., 2011b). O algoritmo também usa a seguinte propriedade de decomposição da informação mútua:

$$I(X^{(i)}, \mathbf{Pa}(X^{(i)}) \cup X^{(j)}) = I(X^{(i)}, \mathbf{Pa}(X^{(i)})) + I(X^{(i)}, X^{(j)} | \mathbf{Pa}(X^{(i)})) \quad (3.10)$$

que permite o cálculo incremental da informação mútua, com valores intermediários mantidos em cache para prevenir cálculos redundantes e, assim, garantir uma melhor eficiência.

O algoritmo e função de pontuação GlobalMIT é implementado e disponibilizado como uma ferramenta de Matlab/C++ (Vinh et al., 2011a), e será empregado neste trabalho para a aprendizagem de DBNs representando a conectividade funcional entre sub-áreas do hipocampo (seções 5.2 e 6.2).

3.3 Valores de informação mútua par-a-par como estratégia para aprendizagem de redes Bayesianas com nós de mesmo tipo

A seção 3.1 mostrou que, embora o uso de medidas de teoria de informação na aprendizagem de DBNs tenha suas vantagens, seu uso em sistemas grandes ou parcialmente observados pode ser difícil, uma vez que muitas das abordagens exigem o cálculo de informação mútua condicional com vários nós envolvidos, o que exige altas quantidades de dados disponíveis e todos observados simultaneamente. As pontuações MIT e GlobalMIT, por exemplo (seção 3.2), resolvem este problema com uma suposição de um número máximo de pais para cada nó da rede (que pode estar disponível pelo conhecimento do especialista).

Uma consideração importante sobre estes métodos, entretanto, é que eles seguem uma filosofia de "maximização", em que as medidas de teoria de informação são apenas observadas em relação a um valor de *threshold*, utilizado para identificar dependência ou independência nas variáveis. Nada é perguntado, entretanto, sobre *como os próprios valores das medidas de teoria de informação podem refletir algo sobre a estrutura da rede*. Obviamente há uma razão para isso: não há como saber como as variáveis estão relacionadas em termos de teoria de informação; ou seja, quanta informação está sendo transmitida de um nó para outro, de forma que qualquer valor poderia ser esperado nessas medidas. Isso está ilustrado na Figura 3.1(a).

No entanto, o que ocorreria caso se considerasse uma rede na qual as variáveis representam processos com comportamentos iguais ou similares — ou, mais especificamente, com um mesmo "perfil de transferência de informação"? Este é o caso da rede exemplificada na Figura 3.1(b,c). Neste caso, poder-se-ia esperar que os valores (ou faixas de valores) de medidas de teoria de informação refletissem características da estrutura da rede — uma ideia apoiada pelo fato de que uma perda fixa de informação entre nós conectados diretamente também ocasionaria perdas fixas entre nós conectados indiretamente (sendo estas últimas menores, de acordo com o teorema da desigualdade do processamento de dados (Cover e Thomas, 2006)). Assim, neste caso, pode-se inferir o "nível de diretividade" entre duas variáveis apenas pelo valor de uma medida de teoria de informação, como informação mútua ou transferência de entropia. E esta informação, por sua vez, pode ser muitíssimo valiosa em uma tarefa de se propor a estrutura da rede — a começar da possibilidade de reduções no espaço de busca, até a própria proposição da estrutura completa.

A evidência experimental já apoia esse tipo de raciocínio, que acontece exatamente nos

Figura 3.1: O quadro tradicional de redes Bayesianas propõe redes com nós representando variáveis de diferentes tipos e comportamentos, como indicado em (a). No entanto, pode ser perguntado quais simplificações seriam possíveis nos algoritmos de aprendizado, e mesmo de inferência, caso se considere sistemas com nós interagindo de formas similares — especificamente, com um mesmo "perfil de transferência de informação" (como em (b)). Pode-se até mesmo evitar-se uma restrição tão grande e considerar redes com um número limitado de tipos, conhecidos *a priori*, aos quais se conhece este perfil de transferência de informação, como indicado em (c) (por exemplo, treinaria-se a rede a partir do conhecimento *a priori* das relações triângulo-triângulo, círculo-círculo e triângulo-círculo).

domínios onde as variáveis modeladas possuem comportamentos semelhantes — o recente trabalho por Endo et al. (2015) sugere, com confirmação da literatura de biologia, padrões de conectividade sináptica na rede de neurônios proprioceptiva do gafanhoto utilizando apenas medidas par-a-par entre os neurônios. O raciocínio é o mesmo indicado acima: valores baixos de informação mútua (embora também suportados pelo valor de atraso encontrado) podem indicar relações mais indiretas entre os nós. O trabalho de Patnaik et al. (2011), por sua vez, inclusive propõe um caso específico de *DBN* intitulado rede dinâmica excitatória (*EDN – Excitatory Dynamic Network*), e impõe algumas restrições nas funções de probabilidade condicional dos nós em questão — ou seja, considera as variáveis como sendo de um tipo específico para introduzir simplificações nos algoritmos de aprendizado. Assim, nota-se que a ideia poderia ser bastante aplicada em neurociência, onde de fato muito tem sido perguntado sobre o comportamento de neurônios em termos de teoria de informação (Gurney, 2001), mas também pode ser ampliada para outras áreas que envolvam redes com nós similares — como redes de distribuição de energia (Chien et al., 2002), redes de transporte (Xie e Levinson, 2009), redes de comunicação (Monge e Contractor, 2003) e quaisquer outros sistemas de rede em geral.

Assim, a subseção a seguir discutirá como pode-se fixar estes perfis de transferência de informação e utilizá-los para inferir o nível de diretividade entre as variáveis de uma rede. O estudo, no entanto, será realizado dentro do paradigma de redes Bayesianas estáticas (e não *DBNs*); porém o raciocínio pode ser prontamente expandido para este caso.

Redes Bayesianas com CPDs Fixos

É necessário certo cuidado ao se pensar em uma forma de considerar um perfil fixo de transferência de informação nas variáveis de uma rede Bayesiana. Afinal, embora se possa esperar que perdas fixas de informação entre duas variáveis conectadas indiretamente produzam também perdas fixas de informação entre variáveis conectadas diretamente, simplesmente assumir um valor fixo de informação mútua entre dois nós diretamente conectados não seria suficiente para se calcular a informação mútua entre todos os outros nós. Por exemplo, considere-se o caso de uma cadeia de Markov denotada por $X \rightarrow Y \rightarrow Z$. Embora possa-se encontrar valores iguais na informação passada de X para Y e de Y para Z , a informação de X para Z pode diferir de muitas formas, devido ao valor de $I(X, Z|Y)$ (Frenzel e Pompe, 2007) (o qual, em um caso de limitações em observações simultâneas, pode não estar disponível). A informação mútua não é uma função injetora; ou seja, as CPDs dos nós da rede não podem ser obtidos somente a partir dos valores de informação mútua, uma vez que a medida pode ser entendida como uma *média ponderada* (do logaritmo da probabilidade conjunta sobre as probabilidades marginais — $MI(X, Y) = \sum_{x,y} p(x, y) \log \frac{p(x,y)}{p(x)p(y)}$).

Outra dificuldade surge na caracterização dos efeitos da informação sendo transmitida em nós com múltiplos pais. Os trabalhos recentes por Williams e Beer (2010) e Rauh et al. (2014) sugeriram formas de realizar uma decomposição de informação multivariada, mas ainda permanece aberta a questão da interpretação dos subcomponentes e seus cálculos — informação única, informação redundante (ou compartilhada) e informação sinérgica (ou complementar) — à luz de propriedades como simetria forte e monotonicidade direita (embora Bertschinger et al. (2013) tenha obtido sucesso no caso de apenas três variáveis). Assim, havendo uma forma de se especificar e fixar estes subcomponentes de informação multivariada, esta seria uma boa interpretação para um perfil de transferência de informação em uma rede com nós iguais. No entanto, em qualquer caso, o grande obstáculo para medidas de teoria de informação com orden baixas (par-a-par, trio-a-trio, etc.) permanece no caso de informação sinérgica — que não poderia ser capturada por essas medidas uma vez que está presente apenas quando se considera vários nós simultaneamente (por exemplo, em uma porta lógica XOR (Bertschinger et al., 2014)). Neste último caso, portanto, resta apenas descartar esta hipótese com base no conhecimento do domínio (e, de fato, muitos domínios não apresentam sinergias com muitas variáveis simultâneas).

Assim, por causa das dificuldades acima, é suficiente para o momento, e por fins de sim-

plicidade, considerar uma rede Bayesiana com nós iguais como uma *rede na qual cada nó tem o mesmo tipo de variável discreta e a mesma distribuição de probabilidade condicional (CPD)*. Uma mesma distribuição de probabilidade, no entanto, requer que cada variável tenha o mesmo número de pais para cada variável — mas isso pode ser resolvido introduzindo um ruído uniforme como pai, em variáveis com menos pais do que os previstos pelo CPD (outra forma seria considerar um CPD diferente para cada número de pais, conforme será sugerido na discussão). Desta forma, pode-se garantir uma perda fixa de informação de um nó a outro e, portanto, também perdas fixas entre nós indiretamente conectados.

Valores de informação mútua agrupados de acordo com a diretividade

Uma vez fixados os CPDs em uma rede Bayesiana, é esperado que quanto mais diretamente conectados estejam dois nós, maior será a informação mútua entre eles. Assim, pode-se esperar que os valores de informação mútua estejam organizados em algumas faixas, de acordo com as conexões mais diretas entre os nós, e pequenas flutuações devido às conexões mais indiretas. A Figura 3.2 ilustra este conceito.

Nota-se, assim, que uma estrutura de rede "entre" dois nós influencia o valor da informação mútua. Pode-se formalizar isso da seguinte forma: define-se um *grafo de influência de informação (IIG, do inglês information influence graph)* para duas variáveis X e Y como *um subgrafo de uma rede Bayesiana B_S no qual cada nó não é d -separado de ambos X e Y , considerando-se também que nenhum deles é controlado* (uma explanação dos conceitos do cálculo causal de Pearl está disponível em Hayduk et al. (2003)). *IIGs* são, portanto, as diferentes configurações de rede entre dois nós que proporcionam diferentes valores de informação mútua.

Estes valores são, como notado, caracterizados pela conexão mais direta entre os nós, com pequenas variações entre nós mais indiretos. Portanto, formaliza-se também a *diretividade* entre dois nós X e Y como *o tamanho do menor caminho não-direcionado entre os dois nós em seu grafo de influência de informação (IIG)*.

Assim, uma vez que exista uma separação entre os valores de informação mútua com base na diretividade entre dois nós, pode-se inferir esta última utilizando-se algum algoritmo de reconhecimento de padrões, que será aplicado e testado nas seções 5.1.1 e 6.1.1. Tal situação é interessante em sistemas com muitos nós ou parcialmente observados nos quais as medidas multivariadas de teoria de informação não são possíveis.

Figura 3.2: A organização esperada nos valores de informação mútua de acordo com a configuração da rede entre dois nós (IIG) em uma rede Bayesiana com CPDs fixos. Nós mais diretamente conectados devem ter valores mais altos de informação mútua, e pequenas flutuações nos valores aparecem devido às conexões mais indiretas.

3.4 Transferência de informação com atrasos de tempo contínuos

As interações entre elementos de uma rede complexa que acontecem através de processos físicos ocorrem dentro de intervalos de tempo contínuo. Em uma rede complexa de neurônios, por exemplo, um neurônio pode transmitir sinais a outros diretamente conectados a eles em um intervalo que varia de 5 a 10 ms, conforme demonstra Kimpo et al. (2003). Com isso, torna-se necessário, para a aplicação de medidas de transferência de informação dinâmicas como a TE, uma discretização no tempo através de métodos de reconstrução por atrasos de tempo (Kantz e Schreiber, 2004) ou subamostragem (Barnett e Seth, 2017). Tal procedimento também é importante para o uso de amostras no tempo para estimação das probabilidades.

Assim, pode-se definir τ como o número de amostras entre os instantes de tempo do processo reconstruído (também chamado de *lag time* (Kantz e Schreiber, 2004)), e redefinir os vetores $\mathbf{y}_n^{(k)}$ e $\mathbf{x}_n^{(l)}$ da equação 3.6 como $\mathbf{y}_n^{(k)} = \{y_{n-\tau}, \dots, y_{n-k\tau}\}$ e $\mathbf{x}_n^{(l)} = \{x_{n-\tau}, \dots, x_{n-l\tau}\}$. O valor τ geralmente é obtido através de duas formas: pelo primeiro mínimo na função de auto informação mútua do processo em função do tempo (ou seja, a informação mútua entre um processo e ele mesmo atrasado), ou pelo cruzamento em zero da função de autocorrelação do processo (Kantz e Schreiber, 2004).

As medidas de transferência de informação apresentadas anteriormente também são utilizadas para identificar os atrasos de tempo entre dois processos ou mais (Wilmer et al., 2012; Pampu et al., 2013; Endo et al., 2015), recebendo os nomes informação mútua atrasada (DMI, do inglês *delayed mutual information*) e transferência de entropia atrasada (DTE, do inglês *delayed transfer entropy*). Estes atrasos, no entanto, exigem a suposição de que a transferência de informação entre dois processos ocorre em função de apenas um estado passado do processo de origem, ou seja, sem sinergia entre estes estados passados. Formalmente, isso implica que o valor l na equação 3.6 para o cálculo da TE seja igual a 1.

Assim, com base na ideia de pico de coerência máxima (Silchenko et al., 2010), o atraso δ , no qual ocorrerá uma sincronização das fases dos sinais X e Y , pode ser obtido, para o caso da DMI, pela expressão $\delta = \arg \max_{\text{beta}} (I(X_{t-\text{beta}}; Y_t))$. No caso da TE, supondo-se $l = 1$,

Figura 3.3: Considerações sobre redes complexas com atrasos de tempo entre os nós. (a) ilustra uma situação onde o valor do atraso obtido entre dois nós elimina a possibilidade de erro devido a um efeito em cascata — um atraso de 30 ms já indica uma relação indireta entre dois nós. Já a situação (b) mostra que nem todos os casos podem ser resolvidos com isso, uma vez que é possível encontrar ligações indiretas em intervalos menores ou mesmo iguais aos de relações diretas. Por fim, o caso (c) é um exemplo de rede onde dois nós em diferentes atrasos (α e β) contribuem sinergisticamente na transferência de informação a um terceiro nó, mostrando que medidas par a par ainda não são suficientes para uma reconstrução completa da rede caso este tipo de situação seja possível. Este caso ainda levanta a hipótese de usarem-se hipergrafos direcionados para se representar um sistema complexo de neurônios.

chega-se à seguinte expressão para a DTE:

$$\delta = \arg \max_{\beta} (TE(X \rightarrow Y, \beta)) = \sum_{\substack{\forall x \in \mathcal{X} \\ \forall y \in \mathcal{Y}}} p(y_n, \mathbf{y}_n^{(k)}, x_{n-\beta}) \log \frac{p(y_n | \mathbf{y}_n^{(k)}, x_{n-\beta})}{p(y_n | \mathbf{y}_n^{(k)})} \quad (3.11)$$

Com isso, no caso de sistemas de Markov de primeira ordem, como os das DBNs, os atrasos encontrados por meio da DMI e DTE podem auxiliar na descoberta de estruturas de rede, inclusive superando algumas das limitações da "maldição da dimensionalidade". Esta é a ideia nos trabalhos de Wibral et al. (2013) e Wibral et al. (2012), que notam que as ligações indiretas entre os nós podem ser facilmente diferenciadas das ligações diretas se detectadas em intervalos de tempo maiores do que o menor encontrado no sistema (e garantido pela propriedade de Markov de primeira ordem). Assim, as medidas de teoria de informação atrasadas par a par ou trio a trio podem ser para se reconstruir o grafo, superando-se, assim, o problema da maldição da dimensionalidade ou mesmo a exigência de todos os dados observados simultaneamente. A Figura 3.3(a) ilustra este caso. Wibral et al. (2012) propõe, enfim, um algoritmo baseado em programação dinâmica para a obtenção de grafos, considerando que as somas dos intervalos de tempo entre conexões diretas entre os nós devem se igualar aos intervalos de tempo encontrados em ligações indiretas.

Alguns cuidados, no entanto, precisam ser tomados, pois em alguns casos tal diferenciação entre ligações diretas ou indiretas pode ser comprometida. O caso ilustrado na Figura 3.3(b) é um exemplo disso: pode existir uma ligação indireta entre dois nós com o mesmo atraso encontrado nas ligações diretas. Não obstante, a consideração de atrasos de tempo ainda não garante o uso de apenas medidas par a par entre os nós — é possível ainda existirem nós em diferentes atrasos contribuindo *sinergisticamente* para transferir informação a outros nós; um efeito que só pode ser considerado a partir de medidas de informação mútua condicional. Estas contribuições podem ser melhor representadas através de hipergrafos direcionados (Gallo et al., 1993); ou combinadas em uma só *CPD*. No entanto, caso se considere que o sistema não possui sinergias deste tipo, o procedimento acima fica justificado.

3.4.1 Redes Bayesianas Dinâmicas de Altas Ordens

Quando um sistema passa pelo processo de *delay reconstruction* tendo em vista a discretização de seus atrasos no tempo, é muito comum que a suposição da propriedade de Markov de primeira ordem, na definição original de *DBNs* (seção 2.1), seja violada; ou seja, não é mais possível supor que os nós em um instante de tempo específico dependem apenas dos outros nós da rede em instantes de tempo imediatamente anteriores (Friedman et al., 1998). Por esta razão, os autores trabalhando com uso de *DBNs* nas áreas de neurociência têm apontado a necessidade de se generalizar as *DBNs* aumentando a propriedade de Markov das *DBNs* para ordens mais altas, como Smith et al. (2006), Daly et al. (2009) e Xuan et al. (2012b). A este tipo de modelo dá-se o nome Redes Bayesianas Dinâmicas de Altas Ordens (*HO-DBNs*, do inglês *High-Order Dynamic Bayesian Networks*) (Li e Ngom, 2013). O desenvolvimento de definições e métodos para *HO-DBNs* é, portanto, crucial para uma modelagem probabilística de sistemas complexos de neurônios.

Definições

Define-se uma *HO-DBN* aumentando a rede de transição de forma que contenha nós não apenas nos tempos t e $t - 1$ (*2-TBN*) (considerado, agora, um sistema discretizado no tempo), mas agora em $t - k$, para $k = 0, 1, \dots, \lambda$, onde λ é a ordem da propriedade de Markov considerada. Assim, uma *DBN* de ordem λ contém uma rede de transição na forma de uma λ -TBN e uma rede de *priori* que representam as independências condicionais da distribuição $P(\mathbf{X}_t | \mathbf{X}_{t-1}, \mathbf{X}_{t-2}, \dots, \mathbf{X}_{t-\lambda})$. Esta expressão também pode ser decomposta na expressão 2.2.

Figura 3.4: Uma rede de transição para uma Rede Bayesiana Dinâmica de ordem λ , que contém N variáveis. Sendo um modelo de *template* para cada instante genérico t em uma série temporal, a representação admite que apenas os nós no tempo t tenham pais. Não obstante, estamos considerando para esta definição apenas arestas *inter-timeslice* (como todos os outros autores consideram).

Ressalta-se também que todos os casos de *HO-DBNs* tratados pela literatura são do tipo cujas redes de transição permitem arestas somente saindo dos nós em $t - 1$ até os nós em t , ou seja, *inter-timeslice*. A Figura 3.4 ilustra a rede de transição de uma *DBN* (*inter-timeslice*) de ordem λ .

A aprendizagem estrutural de *HO-DBNs* traz consigo vários desafios. Um deles concerne o tamanho da rede a ser obtida: em um sistema com N variáveis, uma *HO-DBN* de ordem λ precisa lidar com redes de transição contendo $N \times (\lambda + 1)$ nós; ou seja, uma camada de nós para o tempo atual t e para os tempos passados $t - 1$ a $t - \lambda$. Sabe-se que o número de estruturas possíveis para uma rede Bayesiana cresce super-exponencialmente com o número de nós do grafo (Robinson, 1977) — o que dificulta a tarefa de aprendizado. No entanto, algumas restrições podem ser aplicadas neste caso: primeiramente, a restrição de precedência temporal, na qual arestas saindo de nós em tempos futuros não podem chegar em nós de tempos passados, e em seguida, nota-se também que caso considere-se os processos markovianos (somente os processos, e não o sistema como um todo), também pode-se excluir arestas saindo de tempos em um mesmo nó — ou seja, o estado presente de um processo depende apenas de seu estado passado imediato e o passado de outros processos. Esta última consideração também justifica o fato dos cálculos de transferência de entropia serem realizados apenas em função do passado imediato do processo de destino, variando-se apenas os passados do processo de origem (ou seja, a TE como $TE(X \rightarrow Y) = MI(\mathbf{X}_{t-n}, Y_t | \mathbf{Y}_{t-1})$, variando-se n).

Neste ponto também é interessante notar que os algoritmos desenvolvidos na área de neurociência para a descoberta de redes neurais efetivas se assemelham bastante às técnicas já desenvolvidas na área de redes Bayesianas, embora não recebam o nome de "aprendizagem estrutural de redes Bayesianas dinâmicas de altas ordens". Há, portanto, um ponto de convergência entre as duas áreas — as medidas de teoria de informação desenvolvidas, como MI condicional e TE, podem ser interpretadas exatamente como os testes de independência condicional em um aprendizado estrutural de *DBNs*. Nota-se também que, embora estes métodos utilizem a notação de MI condicional, o conceito de TE também está presente, uma vez que ele mesmo é um tipo de MI condicional em sistemas dinâmicos (vide as equações 3.5 e 3.7). Portanto, nas próximas seções serão expostos os principais métodos desenvolvidos nessas áreas para a aprendizagem estrutural de *HO-DBNs*.

Aprendizado estrutural baseado no algoritmo PC

Na seção 2.2.3 foram indicados alguns algoritmos para se obter a estrutura da rede a partir dos testes de independência condicional ou mesmo medidas de teoria de informação. A maioria deles se baseia em algoritmos tradicionais de obtenção de redes, como o SGS (Spirtes et al., 1990) e o PC (Spirtes e Glymour, 1991). Os trabalho de Runge (2014), por exemplo, se utiliza desses algoritmos.

A ideia consiste em escolher-se os pais para cada nó da rede através de testes de independência baseados em informação mútua condicional (no caso, uma vez que estas são feitas no decorrer do tempo, costuma-se referir-se a estas medidas de transferência de entropia). Os testes começam com medidas em baixas ordens, adicionando todas as arestas encontradas nestes, e à medida que aumenta a ordem, vai removendo as arestas espúrias (devido à redundância). O Algoritmo 1 apresenta o procedimento.

Infelizmente, quando observa-se este tipo de algoritmo à luz dos efeitos de sinergia possíveis em um grafo, surge uma falha grave: o presente algoritmo, ao excluir logo no início algumas arestas do grafo a partir de medidas de baixas ordens (sem o condicionamento de outras variáveis), acaba não deixando espaço para a introdução de arestas que representam efeitos de sinergia. Por exemplo, é possível que dois nós em uma rede transfram informação sinergisticamente para um terceiro nó, porém medidas individuais entre os nós de origem e destino podem não capturar informação alguma (caso de uma porta lógica XOR). Assim, pode-se dizer que os procedimentos de testes de independência condicional que iniciam com medidas par a

---

Algoritmo 1: Algoritmo *PC* para aprendizagem estrutural de *HO-DBNs*.

Input: Todos os processos do sistema $\mathbf{X}$, máxima ordem da *HO-DBN* resultante λ_{max} , número inicial de variáveis utilizadas no condicionamento da MI n_0 , limiar de significância ϵ

para cada $Y_t \in \mathbf{X}$ **faça**

- Inicializar os pais preliminares de Y_t (Pa_{Y_t}):
 $Pa_{Y_t} = \{X_{t-\lambda} : X \in \mathbf{X}, 0 < \lambda \leq \lambda_{max}, I(X_{t-\lambda}; Y_t) > \epsilon\};$
- $n \leftarrow n_0;$
- repita**
  - para cada** combinação $\tilde{Pa}_{Y_t}^{n,i}$ de n nós pertencentes a Pa_{Y_t} **e cada**
 $X_{t-\lambda} \in Pa_{Y_t}$ **faça**
    - se** $I(X_{t-\lambda}; Y_t | \tilde{Pa}_{Y_t}^{n,i}) < \epsilon$ **então** remover $X_{t-\lambda}$ de Pa_{Y_t} ;
  - fim**
- até** $|Pa_{Y_t}| \leq n$, caso contrário, $n \leftarrow n + 1$;

fim

---

par — também adotados nos algoritmos de SGS e PC — foram propostos tendo em vista fugir da introdução de arestas não-existent, decorrentes de redundância, porém com isso também abriram a possibilidade de remoção de arestas existentes, decorrentes de sinergia. Isso não invalida os métodos completamente; no entanto, ao utilizá-los, é preciso sempre pressupor que não existe sinergia no sistema em questão.

Aprendizado estrutural *greedy*

Outras formas de aprendizagem estrutural, no entanto, foram propostas além das baseadas em algoritmos como o *PC*. As propostas por Faes et al. (2011) e Lizier et al. (2013) resolvem os problemas mencionados no algoritmo acima, uma vez que permitem tanto a adição como remoção de arestas no grafo, a partir das medidas de informação mútua condicional. Outra grande contribuição envolve a adição de primeiramente as arestas que providenciam o maior ganho de informação, e a subsequente remoção de arestas que providenciam a maior perda de informação — o que consiste em uma estratégia mais interessante para se percorrer o espaço de busca. O procedimento está descrito no Algoritmo 2.

Diversas simplificações podem ser propostas para este algoritmo, a começar da consideração de contribuições indiretas por atrasos de tempo, conforme indicado anteriormente. Não obstante, o método desenvolvido é uma estratégia *greedy* — outros métodos de otimização podem ser propostos para a busca do grafo, entre os quais um deles será apresentado na próxima seção.

Algoritmo 2: Algoritmo *enhanced iterative greedy* para aprendizagem estrutural de *HO-DBNs*.

Input: Todos os processos do sistema $\mathbf{X}$, máxima ordem da *HO-DBN* resultante λ_{max} , limiar de significância ϵ , número máximo de considerações simultâneas na adição n_{max}

para cada $Y \in \mathbf{X}$ **faça**

    Inicializar os pais preliminares de Y : $P_Y = \emptyset$;

para n de 1 a n_{max} **faça**

repita /* fase de adição de arestas */

$\tilde{\mathbf{X}}^{n,i} = \{X_{t-\lambda} : X \in \mathbf{X} \setminus P_Y, 0 < \lambda \leq \lambda_{max}, |\tilde{\mathbf{X}}^{n,i}| = n\}$

            /* combinações de tamanho n de possíveis pais de Y */

$i_{max} = \arg \max_i I(\tilde{\mathbf{X}}^{n,i}, Y_t | P_Y)$;

se $I(\tilde{\mathbf{X}}^{n,i_{max}}, Y_t | P_Y) > \epsilon$ **então** adicionar $\tilde{\mathbf{X}}^{n,i_{max}}$ a P_Y ;

até nenhum $\tilde{\mathbf{X}}^{n,i}$ poder ser encontrado ou adicionado;

fim

repita /* fase de remoção de arestas */

$\tilde{X}^i = X_{t-\lambda}$, tal que $X_{t-\lambda} \in P_Y$ /* possíveis não-pais de Y ,

        porém adicionados na fase anterior */

$i_{min} = \arg \min_i I(\tilde{X}^i, Y_t | P_Y \setminus \tilde{X}^i)$;

se $I(\tilde{X}^{i_{min}}, Y_t | P_Y \setminus \tilde{X}^i) < \epsilon$ **então** remover $\tilde{X}^{i_{min}}$ de P_Y ;

até nenhum $\tilde{X}^i$ poder ser encontrado ou removido;

fim

Percorrendo o espaço de busca de HO-DBNs com PSO

Embora o algoritmo por Lizier et al. (2013) apresente uma estratégia *greedy* para a elicitação de um grafo, pode-se começar a pensar em melhores alternativas de se percorrer o espaço de busca. Nas áreas de aprendizagem estrutural de redes Bayesianas baseada em heurísticas, recentemente vários métodos de otimização têm sido adaptados — como os evolutivos (Villanueva e Maciel, 2010, 2014), baseados em *simulated annealing* (Ye et al., 2008), otimização por colônia de formigas (*ACO*) (Daly e Shen, 2009) e por enxame de partículas (*PSO*) (Sahin e Devasia, 2007). No entanto, em todos estes casos as heurísticas são utilizadas para a otimização de uma função de pontuação, que utiliza diretamente os dados fornecidos para o treinamento, e não medidas de transferência de informação como informação mútua e transferência de entropia. Não obstante, nenhum desses métodos foi adaptado para o espaço de busca de *HO-DBNs*.

Assim, durante o presente trabalho foi desenvolvido um método heurístico baseado em *PSO* para estruturas de *HO-DBNs*. O método ainda se baseia em funções de pontuação genéricas (nada foi considerado sobre como incorporar medidas de transferência de informação em uma pontuação), no entanto, já consistiu em um primeiro passo em direção a uma solução. A escolha do *PSO* foi feita devido à sua simplicidade e poucos parâmetros de otimização necessários para atingir bons resultados.

Otimização por Enxame de Partículas ou *PSO* (*Particle Swarm Optimization*) é uma técnica de otimização bioinspirada proposta por Kennedy et al. (1995), e se baseia na filosofia da inteligência de enxames. Sua proposta consiste em simular o comportamento coletivo de pássaros ou cardumes definindo uma população de partículas ou indivíduos, com posições e velocidades, que se movem em um espaço de busca, de acordo com (a) sua melhor posição encontrada (melhor posição local), (b) a melhor posição encontrada por seus vizinhos (melhor posição global) e (c) um fator de perturbação aleatório na sua trajetória. A cada passo do algoritmo, uma função de *score* ou *fitness* é aplicada em cada partícula, definindo-se as melhores posições locais e globais. Assim, cada uma tem sua velocidade e posição atualizada conforme as expressões:

$$v_i(t+1) = \omega v_i(t) + \psi_1 R_1(x_{gi} - x_i(t)) + \psi_2 R_2(x_{bi} - x_i(t)) \quad (3.12)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \cdot 1 \quad (3.13)$$

onde v_i e x_i são a velocidade e posição da partícula i , x_{gi} é a melhor posição global, e x_{bi}

é a melhor posição local. R_1 e R_2 são números aleatórios tirados do intervalo $[-0.5, 1.5]$ — um fator de perturbação aleatório usado para explorar melhor o espaço de busca. Finalmente, ψ_1 e ψ_2 são, respectivamente, os parâmetros social e cognitivo, que definem a quantidade de influência que uma partícula recebe dos máximos global e local. ω é um fator inercial, usado para reduzir a velocidade que uma partícula já possui anteriormente. Estas últimas três variáveis são dadas *a priori*, sendo parâmetros de controle do algoritmo. As posições são atualizadas com $v_i(t + 1)$ multiplicado por um tempo discreto igual a 1.

A grande dificuldade a ser tratada na adaptação de *PSO* para um espaço de busca de *DAGs* é o fato deste espaço ser discreto e, portanto, não diferenciável, o que dificulta uma interpretação de velocidade e inércia para as partículas (Kennedy e Eberhart, 1997). Assim, as adaptações são geralmente feitas seguindo uma de duas vertentes: uma utiliza-se de algoritmos discretos de *PSO* de propósito geral já estabelecidos na literatura, como *PSO* Binário (Nezamabadi-pour et al., 2008), adaptando a representação em *DAG* para a representação exigida por estes algoritmos, enquanto a outra consiste em adaptações dos conceitos e operações realizadas em *PSO*, como posição, velocidade e influência local/global, a uma busca no espaço de *DAGs* ou representações de *DAGs*. O trabalho de Sahin e Devasia (2007) segue a primeira vertente e muda o espaço de busca para um espaço contínuo e multidimensional onde cada variável pode estar no intervalo $[0, 1]$; sendo o valor destas variáveis o *likelihood* de uma certa característica discreta no modelo — no caso, a presença ou não de uma aresta no grafo. Por outro lado, os trabalhos de Du et al. (2005) Heng et al. (2006) seguem a segunda vertente, e embora exijam a reformulação do algoritmo como um todo, também permanecem mais abertos para melhorias específicas no domínio, e deixam mais explícitos os mecanismos de inteligência de enxames pretendidos no método. Especificamente, o trabalho de Du et al. (2005) representa um *DAG* como uma matriz $n \times n$, com n sendo o número de nós da rede, e define operações entre matrizes. Por sua vez, Heng et al. (2006) utiliza uma representação por listas dinâmicas de adjacências, e define também as operações necessárias. Este último caso foi o adotado no presente trabalho.

Assim, a descrição do algoritmo começa definido o que se entende por posições, velocidades e inércia de partículas no espaço de busca de *HO-DBNs*.

Define-se uma *unidade de precedência*, $\mathbf{PU}_{t-j}^{(i)}$, como uma lista que contém cada variável do sistema que, em um instante anterior $t - j$, é pai de uma variável $X^{(i)}$. Uma *lista de precedência*, $\mathbf{PL}_{t-j}$, para um dado instante $t - j$, é uma lista que contém uma unidade de

Representação resultante: $\{ \{ \{X^{(1)}, X^{(3)}\}, \{X^{(1)}, X^{(2)}\}, \{\emptyset\} \}, \{ \{X^{(2)}\}, \{\emptyset\}, \{X^{(2)}, X^{(3)}\} \} \}$

Figura 3.5: Mapeamento de uma *HO-DBN* em um conjunto de listas de precedência. Esta é a representação adotada para uma posição no espaço de busca do algoritmo de *PSO* proposto.

precedência para cada variável do sistema — e, assim, representa todos os pais em um instante $t - j$. Finalmente, define-se uma posição no espaço de busca - que é, afinal, uma representação de *HO-DBN* como um conjunto de listas de precedência, contendo uma para cada instante anterior previsto na ordem de Markov da *HO-DBN*. A Figura 3.5 ilustra esta representação.

Dado que o espaço de busca é discreto e, por isso, não diferenciável, não existe uma forma simples de definir o conceito de velocidade. Adota-se o conceito introduzido em por Heng et al. (2006), baseado em listas de troca. Uma lista de troca é uma lista de precedência cujas unidades de precedência admitem sinais positivos ou negativos. Isso significa que um $-X^{(i)}$ em uma unidade precedência dentro da lista de troca de $X^{(j)}$, no tempo $t - \lambda$, representa a remoção de uma aresta saindo de $X_{t-\lambda}^{(i)}$ e chegando em $X^{(j)}$. Um $+X_i$, por sua vez, representa a adição de uma aresta. Assim, define-se a velocidade de uma partícula como um conjunto de listas de trocas, com a mesma ordem e número de variáveis do modelo desejado.

Também define-se o módulo de uma velocidade como a soma do tamanho de todas as unidades de precedência presentes na velocidade; formalmente, $\|Velocidade\| = \sum_{i=1}^N \sum_{j=1}^{\lambda} \|\mathbf{PU}_{t-j}^{(i)}\|$, onde $\|\mathbf{PU}_{t-j}^{(i)}\|$ é o tamanho da unidade de precedência de uma variável $X^{(i)}$ no instante $t - j$ (em uma *HO-DBN* de ordem λ e N variáveis).

O algoritmo de *PSO* envolve várias operações com posições e velocidades. Portanto, além de definir representações para posições e velocidades no espaço de busca, precisa-se também

definir o que significam e como são realizadas operações com essas representações. As operações são definidas a seguir:

1. **Adição de posição com velocidade:** Seja x uma posição e v uma velocidade. A posição resultante da operação $x' = x + v$ é obtida aplicando adições e remoções em x conforme indicadas em cada unidade de precedência das listas de troca de v . Por exemplo, em uma *HO-DBN* de ordem 2 e 3 variáveis, tem-se:

$$\begin{aligned} x &= \{ \{ \{X^{(1)}, X^{(3)}\}, \{X^{(2)}\}, \{X^{(1)}\} \}, \{ \{X^{(2)}\}, \{\emptyset\}, \{X^{(2)}, X^{(3)}\} \} \} \\ v &= \{ \{ \{+X^{(1)}, -X^{(3)}\}, \{+X^{(1)}\}, \{\emptyset\} \}, \{ \{-X^{(2)}, -X^{(3)}\}, \{\emptyset\}, \{+X^{(1)}\} \} \} \\ x' = x + v &= \{ \{ \{X^{(1)}\}, \{X^{(1)}, X^{(2)}\}, \{X^{(1)}\} \}, \{ \{\emptyset\}, \{\emptyset\}, \{X^{(1)}, X^{(2)}, X^{(3)}\} \} \} \end{aligned}$$

2. **Subtração de posição com posição:** Sejam x_1 e x_2 duas posições no espaço de busca. Esta operação retorna uma velocidade $v = x_2 - x_1$ tal que $x_1 + v$ seja x_2 . Por exemplo, em uma *HO-DBN* de ordem 2 e 3 variáveis, tem-se:

$$\begin{aligned} x_1 &= \{ \{ \{X^{(1)}, X^{(3)}\}, \{X^{(2)}\}, \{X^{(1)}\} \}, \{ \{X^{(2)}\}, \{\emptyset\}, \{\emptyset\} \} \} \\ x_2 &= \{ \{ \{X^{(1)}, X^{(2)}\}, \{X^{(2)}\}, \{\emptyset\} \}, \{ \{X^{(2)}, X^{(3)}\}, \{\emptyset\}, \{X^{(2)}, X^{(3)}\} \} \} \\ v = x_2 - x_1 &= \{ \{ \{+X^{(2)}, -X^{(3)}\}, \{\emptyset\}, \{-X^{(1)}\} \}, \{ \{-X^{(3)}\}, \{\emptyset\}, \{+X^{(2)}, +X^{(3)}\} \} \} \end{aligned}$$

3. **Adição de velocidade com velocidade:** Seja $v' = v_1 + v_2$ a soma das duas velocidades v_1 e v_2 . v' tem suas unidades de precedência como o resultado da concatenação das unidades de precedência de v_1 e v_2 , contraídas para retirar redundâncias — ou seja, se a concatenação resulta na unidade de precedência $\{+X^{(1)}, +X^{(2)}, -X^{(2)}, -X^{(3)}\}$, esta é contraída para $\{+X^{(1)}, -X^{(3)}\}$.

4. **Multiplicação de constante com velocidade:** Seja α um número real e v uma velocidade — no caso do algoritmo de PSO, α é o fator de perturbação aleatório que multiplica as velocidades. Pode-se dividir a operação $v' = \alpha v$ em três casos:

- (a) $\alpha=0$: v' tem todas as unidades de precedência vazias.
- (b) $\alpha>0$: Arestas são removidas ou adicionadas aleatoriamente até que $\|v'\| = \lfloor \alpha \|v\| \rfloor$.

Por exemplo, em uma *HO-DBN* de ordem 2 e 3 variáveis, tem-se a velocidade:

$$v = \{ \{ \{ +X^{(1)}, -X^{(3)} \}, \{ +X^{(1)} \}, \{ \emptyset \} \}, \{ \{ -X^{(2)}, -X^{(3)} \}, \{ \emptyset \}, \{ +X^{(1)} \} \} \}$$

Em um caso onde $\alpha=1.333$, precisa-se adicionar arestas até que o módulo seja $\lfloor \alpha \|v\| \rfloor = \lfloor 1.333 * 6 \rfloor = 8$, ou seja, $8 - 6 = 2$ arestas aleatórias.

- (c) $\alpha < 0$: Inverte-se os sinais de cada elemento das listas de precedência de v , e multiplica-se o resultado por $-\alpha$, conforme o último caso.

Percebe-se que esta última operação não é determinística. Isso não significa um problema se assumirmos que a constante α é usada para introduzir uma perturbação aleatória nas velocidades. O conceito é similar ao introduzido no trabalho de Espezua et al. (2014), sobre operadores em algoritmos genéticos, onde a randomização não modifica apenas o módulo do vetor de direção, mas também o projeta em outras dimensões do espaço de busca, otimizando, assim, a exploração do espaço.

Uma vez que espaços de busca discretos não são diferenciáveis, não há como conceber uma partícula se movendo no mesmo em uma linha reta, e por isso, também não há como conceber velocidades constantes e o conceito inércia. Isso se torna ainda mais claro no conceito de velocidade definido anteriormente: uma vez que a posição da partícula é atualizada com sua velocidade, não há sentido em manter esta última, uma vez que ela consiste apenas em uma direção para a posição atual.

Assim, propõe-se a seguinte solução para este problema: cada partícula não mantém sua presente velocidade v a cada iteração, mas adota uma nova velocidade v' aleatória com o mesmo módulo de v multiplicado pelo fator de inércia ω , ou seja, $\|v'\| = \lfloor \omega \|v\| \rfloor$. Esta solução baseia-se no trabalho de Espezua et al. (2014), que utiliza a ideia de projetar um vetor de direção em outras dimensões do espaço de busca, conseguindo assim uma aleatoriedade adicional no algoritmo. Similarmente, o presente trabalho aproveita o conceito de inércia das partículas para introduzir esta aleatoriedade, resultando, assim, em uma melhor exploração do espaço de busca.

Finalmente, o algoritmo de *PSO* proposto, incluindo as inicializações, está descrito no Algoritmo 3. Destaca-se que, na presente abordagem, a melhor posição global é definida como a melhor posição em toda a população, uma vez que não há uma forma simples de definir ou mesmo encontrar partículas vizinhas em um espaço de busca discreto. Seguindo também

recomendações propostas no trabalho de Robinson e Hartemink (2010), o algoritmo opera com restrições: de velocidade máxima e de máxima quantidade de pais em um nó — fatores que evitarão que o algoritmo opere em extremos do espaço de busca e possa, assim, convergir mais rápido para a solução.

As seções 5.1.2 e 6.1.2 apresentarão as simulações realizadas para se verificar a corretude e eficiência do algoritmo.

Algoritmo 3: Algoritmo de *PSO* para aprendizagem estrutural de *HO-DBNs*.

Input: Função de pontuação ou *fitness* $Score()$ (no caso, a pontuação K2 para *HO-DBNs* proposta na seção 3.4.1), tamanho da população S , fator de inércia ω , parâmetro social ψ_1 , parâmetro cognitivo ψ_2 , número de épocas Γ , máxima velocidade v_{max} e número máximo de pais para um nó p_{max} .

inicialização: Gerar S partículas com posições x_i e velocidades v_{i+1} aleatórias;

```
/* Define as melhores posições locais  $x_{bi}$  */
para cada  $i$  faça  $x_{bi} \leftarrow x_i$ ;
/* Define a melhor posição global  $x_g$  */
 $x_g \leftarrow \max(Score(x_{bi}), \forall i)$ ;
repita /* Loop principal do algoritmo */
  para cada  $i$  faça
     $I \leftarrow Inertia(\omega, v_{i+1})$  /* Calcula a contribuição da inércia */
     $R_1 \leftarrow Random(-0.5, 1.5)$ ;
     $R_2 \leftarrow Random(-0.5, 1.5)$ ;
    /* Atualiza as velocidades das partículas de acordo
       com a inércia e as melhores posições locais e
       globais. */
     $v_{i+1} \leftarrow I + \psi_1 R_1(x_g - x_i) + \psi_2 R_2(x_{bi} - x_i)$ ;
    /* Verifica a restrição de máxima velocidade, se for
       maior, reduz */
    se  $\|v_{i+1}\| > v_{max}$  então
      |  $v_{i+1} \leftarrow v_{i+1} * \frac{v_{max}}{\|v_{i+1}\|}$ 
    fim
    /* Atualiza a posição com a nova velocidade */
     $x_i \leftarrow x_i + v_{i+1} \cdot 1$ ;
    /* Verifica a restrição de qtd. máxima de pais */
    enquanto  $Pais_{X^{(j)}}(x_i) > p_{max}$  para qualquer  $X^{(j)}$  faça
      | Remove um pai do nó  $X^{(j)}$  aleatoriamente;
    fim
    /* Atualiza a melhor pos. local */
    se  $Score(x_i) > Score(x_{bi})$  então
      |  $x_{bi} \leftarrow x_i$ ;
    fim
  fim
  /* Atualiza a melhor pos. global */
   $x_g \leftarrow \max(Score(x_{bi}), \forall i)$ ;
até  $\Gamma$ ;
```

Capítulo 4

Aplicações de identificação de sistemas neurais com DBNs

Tendo visto o atual estado da arte em DBNs e sua integração com teoria de informação, o presente capítulo apresentará duas possíveis áreas de aplicação, as quais serão exploradas no presente trabalho: a descoberta de conectividade funcional entre as sub-áreas do hipocampo humano a partir de ressonância magnética funcional (fMRI) e a identificação do sistemas neuromotores do gafanhoto a partir de gravações intracelulares de potenciais sinápticos. Tais aplicações serão situadas dentro de um escopo mais amplo; a saber, o de estudos de conectividade com fMRI (4.2) e de sistemas neuromotores (4.2), incluindo uma revisão da literatura e uma discussão das principais questões relacionadas. Com isso, as aplicações específicas são descritas (seções 4.1.1 e 4.2.1) e considerações são feitas acerca de seu estudo com as técnicas descritas no presente trabalho.

4.1 Ressonância Magnética Funcional (fMRI)

Desde sua invenção em 1973 por Paul C. Lauterbur (Lauterbur, 1973), a técnica de imagem por ressonância magnética (MRI) tem sido uma das mais utilizadas para diagnóstico médico e estudos sobre anatomia e função do cérebro humano, com uma estimativa de mais de 25.000 *scanners* em uso ao redor do mundo (Rinck, 2001). Avanços recentes nesta área incluem o desenvolvimento de novos tipos de bobinas e sequências (ou seja, configurações de pulsos e gradientes de campo magnético que resultam em diferentes tipos de imagens) e a solução de

desafios resultantes do aumento da força do campo magnético, entre os quais se incluem os campos não-homogêneos, o aumento de artefatos resultantes de susceptibilidade magnética, e limitações de taxa de absorção específica (van der Kolk et al., 2013b). O aumento do campo magnético, no entanto, apresenta vantagens que são de grande interesse para a comunidade científica, uma vez que é capaz de gerar imagens com altas resoluções dentro de tempos aceitáveis de aquisição, e com melhores relações de sinal-ruído (SNR) e contraste-ruído (CNR). Neste contexto, destaca-se o emprego dos *scanners* de MRI de campo ultra-alto, tais como o de força de campo de 7T, no contexto de pesquisas clínicas, dados seus potenciais e capacidades demonstradas para o diagnóstico clínico e uma grande variedade de estudos. (Kraff et al., 2015; van der Kolk et al., 2013a). A Figura 4.1 demonstra como este tipo de equipamento é capaz de alcançar uma alta resolução espacial em imagens *in-vivo* do sistema nervoso central, quando visto em comparação com um equipamento de 3T. Além disso, o surgimento das bobinas de radiofrequência multicanal e as técnicas de excitação de multibanda também consistem em um passo importante, contribuindo para a obtenção de imagens de alta resolução com aquisição rápida, devido à sua capacidade de paralelização (Sohn, 2013).

Tais avanços são especialmente interessantes para aplicações de ressonância magnética funcional (fMRI), uma vez que aumenta a resolução temporal das imagens (Feinberg e Setsompop, 2013). O fMRI têm ocupado um importante lugar nos estudos de conectividade cerebral (Zheng e Rajapakse, 2006; Stephan e Friston, 2010), ao lado de outras técnicas comuns como eletroencefalograma (EEG) e magnetoencefalograma (MEG) (Wang et al., 2016; Haufe et al., 2011; Mutlu e Aviyente, 2009; Smith et al., 2006). Sua vantagem se encontra em sua alta resolução espacial, permitindo a investigação da integração de várias áreas do cérebro (MEG e EEG, por outro lado, formam sua contraparte devido à alta resolução temporal, além de consistirem em métodos menos invasivos se comparados ao fMRI).

As técnicas de fMRI também têm sido recentemente melhoradas com sequências de imagem eco-planar (*echo planar imaging* ou EPI) devido à sua alta sensibilidade ao contraste dependente de nível de oxigenação do sangue (*blood oxygenation level dependent* ou BOLD) (Feinberg e Setsompop, 2013), que é associado à atividade cerebral nos estudos de fMRI através da modelagem de uma resposta hemodinâmica (Friston et al., 1994). As sequências imagem eco-planar foram primeiramente propostas por Mansfield (1977) e são capazes de obter, rapidamente, várias linhas de dados de imagem, depois de uma única excitação de radiofrequência consistindo em uma série de ecos de gradiente de campo magnético (Poustchi-Amin

Figura 4.1: Comparação entre imagens estruturais obtidas em equipamentos de MRI de 3T (a,b,c) e 7T (d,e,f) de força de campo magnético, em três ângulos. As imagens foram adquiridas em voluntários no departamento de Radiologia da Universidade de Pittsburgh (EUA).

et al., 2001). Como exemplo, no estudo realizado por Moeller et al. (2010a), que usa EPI juntamente com técnicas de excitação multi-banda (também chamada de *simultaneous multi-slice* ou SMS), obteve-se imagens com resolução de $2 \times 2 \times 2mm^3$, com tempo de repetição de pulso de $1.25s$ e um campo de visão de $256 \times 256 \times 176mm^3$ e, em seguida, imagens com resolução de $1 \times 1 \times 2mm^3$, com tempo de repetição de pulso de $1.5s$ e campo de visão de $192 \times 172 \times 176mm^3$. Assim, em uma avaliação recente, Preibisch et al. (2015) conclui que "dado o *hardware* apropriado, uma sequência de EPI multi-banda com uma aceleração total igual a 4 é capaz de reduzir significativamente o tempo de aquisição e claramente aumenta a sensibilidade à detecção de redes de conectividade cerebral em estado de repouso", o que indica que o EPI multi-banda é um "novo padrão para fMRI capaz de melhorar nosso entendimento de dinâmica de redes em cérebros tanto saudáveis como doentes.

4.1.1 Conectividade funcional entre sub-áreas do hipocampo

Uma das aplicações importantes da aquisição de imagens por ressonância magnética tem sido o estudo do hipocampo humano (Bartsch, 2012; den Heijer et al., 2012). O hipocampo é uma estrutura localizada no lobo medial temporal e tem sido associado com as funções de navegação espacial, emoção, aprendizado e memória (Jeneson e Squire, 2012). Pesquisas clínicas têm investigado como mudanças em sua forma e tamanho podem estar relacionadas à doença de Alzheimer (Frisoni, 2001; Mueller et al., 2005; Mortimer et al., 2013), esquizofrenia, (Csernansky et al., 1998; Heckers, 2001; Harrison, 2004) e epilepsia (Bernhardt et al., 2015).

Pesquisas recentes também têm revelado que o hipocampo apresenta especializações funcionais ao longo de seu eixo longitudinal (Poppenk et al., 2013). Um estudo de conectividade funcional descobriu que a parte anterior do hipocampo (aHP) possui uma maior correlação funcional com o córtex peririnal (que se correlaciona a uma rede nas regiões temporal anterior e frontal), enquanto que a parte posterior do hipocampo (pHP) está funcionalmente correlacionada ao córtex parahipocampal (que se correlaciona com uma rede na região temporal posterior-medial, parietal e occipital) (Libby et al., 2012). Outros estudos têm sido consistentes com esta descoberta; por exemplo, Adnan et al. (2016) mostra uma maior conectividade do aHP com o córtex temporal e orbital, enquanto o pHP mostra uma maior conectividade com o córtex medial e parietal lateral. Abordagens baseadas puramente em dados, as quais não dependem de segmentações estruturais *a priori*, também confirmam padrões distintos de conectividade ao longo do hipocampo. Certo estudo utiliza análise de componentes independentes (ICA) para dados de fMRI no hipocampo em pacientes em estado de repouso (Blessing et al., 2016). A análise revela componentes que são discretos ao longo do eixo longitudinal, com padrões de conectividade distintos em relação ao resto do cérebro. Com base neste fato e outra literatura adicional (conforme a revisão por Poppenk et al. (2013)), várias hipóteses têm sido levantadas quanto à natureza desta especialização ao longo do eixo longitudinal do hipocampo.

Pesquisas como essas apoiam o argumento de Maruszak e Thuret (2014) de que uma melhor compreensão das diferentes funções exercidas pelas sub-áreas do hipocampo é um passo crítico em direção ao entendimento neurobiológico do funcionamento do cérebro, e como disfunções podem surgir nele. Tal empreendimento se torna apenas possível com os avanços em equipamentos de campo magnético ultra-alto, como o de 7T, capazes de adquirir imagens de alta resolução apresentando partes do cérebro a um altíssimo nível de detalhe e precisão

(Van Leemput et al., 2009; Duyn, 2012).

Os avanços em segmentação automática também facilitam a extração de regiões de interesse (ROIs, do inglês, *regions of interest*) e a disponibilização de mais dados para análise, uma vez que a segmentação manual destas regiões pode ser trabalhosa, consumindo muito tempo e ainda contar com o risco de erros e divergências devido a diferentes protocolos de segmentação (Yushkevich et al., 2015a). Os algoritmos de segmentação automática são implementados de várias formas, entre as quais se destacam os modelos deformáveis e as baseadas em atlas Dill et al. (2015). No caso específico das sub-áreas do hipocampo, no entanto, apenas abordagens baseadas em atlas foram estudadas e disponibilizadas como ferramentas. Yushkevich et al. (2015b) utiliza uma técnica de fusão de rótulos (ou seja, marcações que identificam as áreas em uma imagem) de múltiplos atlas em imagens T1 e T2-*weighted* intitulada ASHS (*Automatic Segmentation Hippocampal Subfields*). Um trabalho recente também publicou um atlas com imagens obtidas em sistemas de 7T no estudo PREDICT-MR, e disponibilizou-o para utilização com a ferramenta ASHS (Wisse et al., 2014). Pipitone et al. (2014), por sua vez, desenvolveram um procedimento de votação de múltiplos-atlas baseada em bibliotecas padrão formadas pelo conjunto de todos os indivíduos de um experimento, e intitularam-no MAGeT (*Multiple Automatically Generated Templates*). Finalmente, Iglesias et al. (2015) propõem um método de atlas probabilístico baseado primariamente em MRI ex-vivo de ultra-alta resolução (voxel igual a 0.17 mm isotrópico) e secundariamente em imagens T1 de tamanho de voxel igual a 1 mm isotrópico obtidas em sistemas de 3T. O método é disponibilizado como uma funcionalidade do software Freesurfer 6.

Dada uma avaliação experimental destes três métodos de segmentação automática, a conclusão é que a qualidade dos resultados em cada método depende quase que exclusivamente dos tipos e qualidades de imagem utilizadas como atlas e como entrada para segmentação. Yushkevich et al. (2015b) nota que imagens T1 apresentam um contraste melhor para a localização do hipocampo e algumas de suas subestruturas em relação ao resto do cérebro, enquanto que imagens T2-*weighted* apresentam um melhor contraste para o delineamento destas subestruturas. Para imagens obtidas em sistemas de 7T, a escolha mais natural no momento seria o atlas de 7T disponibilizado para uso com ASHS, uma vez que é um dos poucos atlas de 7T públicos até o momento. No entanto, ASHS requer ambas imagens T1 e T2-*weighted* como entrada, o que pode não estar disponível em vários estudos. Assim, outros métodos se tornam opções viáveis, e embora imagens obtidas em sistemas de 3T e 7T apresentem algumas diferenças de

contraste, segmentações aceitáveis ainda podem ser obtidas em imagens de 7T utilizando-se atlas de 3T. Particularmente, o método implementado no Freesurfer 6 por Iglesias et al. (2015) é interessante para imagens de 7T uma vez que utiliza imagens ex-vivo de alta resolução em seu atlas probabilístico.

4.1.2 Conectividade direcionada e modelos dinâmicos em fMRI

A elicitación de redes de conectividade neural, uma área em crescente desenvolvimento dentro do paradigma de sistemas complexos, tem refletido a mudança na concepção de funcionalidade do cérebro: enquanto, antes, uma função cerebral era vista como localizada em diferentes áreas (ou seja, a abordagem de segregação funcional), hoje ela é vista como a integração complexa de diversas áreas, produzida um comportamento emergente (ou seja, a abordagem de integração funcional) (Friston, 2011; Sporns, 2011). No entanto, é natural que, dados os diversos paradigmas científicos, muitas interpretações surjam quanto ao que seria uma explicação verdadeira, precisa e útil de funcionalidade cerebral, bem como o significado dos modelos propostos e dos resultados obtidos a partir da análise dos dados. Assim, várias opiniões ainda dividem a comunidade científica de neurociência, promovendo várias discussões, importantes e necessárias, no decorrer dos anos (Ramsey et al., 2010; Valdes-Sosa et al., 2011). O que se segue é uma apresentação das diferentes formas de se dividir e abordar as técnicas de conectividade neural em fMRI.

Uma das mais importantes distinções é a estabelecida por Friston (2011), entre conectividade funcional e efetiva. Ele define a primeira como "dependências estatísticas entre eventos neurofisiológicos remotos", e a última como "a influência que um sistema neural exerce sobre outro, seja em um nível sináptico ou de população". Isso implica que a conectividade funcional consiste em uma abordagem exploratória e guiada por dados, enquanto a conectividade efetiva é uma abordagem mais confirmatória e guiada por modelos. Exemplos de técnicas de conectividade funcional são as medidas de correlação estatísticas e modelos gráficos probabilísticos (redes Bayesianas e campos aleatórios de Markov), enquanto de conectividade efetiva são a modelagem causal dinâmica (*dynamic causal modeling* ou DCM), a modelagem por equações estruturais (*structural equation modeling* ou SEM) e a modelagem por causalidade de Granger (*Granger causality* ou GC)¹. Outra forma de dividir as técnicas é proposta no trabalho de

---

¹A GC, no entanto, está aberta a discussões; Friston (2011), por exemplo, argumenta que ela não pode ser considerada um método de conectividade efetiva uma vez que reparametriza o modelo em função de parâmetros de autoregressão, que é diferente de um acoplamento causal (vide Friston (2011, pp. 19-20))

(Lohmann et al., 2013), agrupando-as em análises focadas na descoberta de padrões espaciais de conectividade, tais como as medidas de associação, análise de componentes independentes (*independent component analysis* ou ICA) e análise de padrão multivariada (*multivariate pattern analysis* ou MVPA), e análises dentro do paradigma de redes complexas, tais como as baseadas em topologia de redes, redes Bayesianas e DCM.

Um outro foco de interesse dos métodos de análise de conectividade concerne a busca de conectividade direcionada e de modelos dinâmicos. A conectividade direcionada se refere aos modelos de rede cujas arestas não apenas indicam a presença ou intensidade de uma conexão, mas também sua direção. Isto se torna importante à medida que as pesquisas se voltam para a dinâmica dos sistemas neurais; como por exemplo, na busca de redes cíclicas e mecanismos de *feedback* (conforme explicado na seção 2.1). Este é o caso de métodos como DCM, proposto por Smith et al. (2011), mostrando uma vantagem sobre métodos como *Structural Causal Modeling* (SCM) e redes Bayesianas tradicionais por não se limitarem a modelos de grafo acíclicos. Assim, também junto com a proposta de DCM, outras pesquisas têm integrado abordagens de causalidade de Granger como SCM ao converterem grafos cíclicos em acíclicos usando nós em diferentes instantes de tempo — uma solução que Razi e Friston (2016) define como uma formulação em tempo discreto de DCMs e que, afinal, o próprio modelo de DBN (porém com a suposição de gaussianidade, no caso de GC), a ser detalhado a seguir. Valdes-Sosa et al. (2011) também propõe o conceito de modelagem causal estrutural dinâmica como síntese das abordagens propostas na literatura concernentes à análise de conectividade efetiva em modelos dinâmicos com representação de grafo.

Ao se considerar o estudo de conectividade entre as sub-áreas do hipocampo, uma análise de conectividade funcional é a melhor opção, uma vez que pouco conhecimento está disponível sobre este sistema (ou seja, uma abordagem exploratória, ao invés de confirmatória, é necessária). Ademais, os estudos atuais ainda não investigaram a direção das conexões ao considerarem atrasos de tempo entre os elementos do sistema. Assim, acredita-se que os dois métodos mais promissores para avançar o conhecimento sobre conectividade entre sub-áreas do hipocampo são as medidas de transferência de informação e DBNs, que correspondem, segundo a divisão de Lohmann et al. (2013), a abordagens baseadas em padrões espaciais e redes complexas, respectivamente.

Medidas de TE foram aplicadas a dados de fMRI por Hinrichs et al. (2006), Chai et al. (2009) e Lizier et al. (2011). Lizier et al. (2011) propõe uma extensão para a transferência

de entropia chamada *transferência de entropia multivariada*, que busca estimar a transferência de informação entre regiões do cérebro constituídas por várias séries temporais (ou seja, vários voxels). Os resultados deste método aplicado a um estudo de coordenação visual-motora mostraram-se promissores, apresentando "uma estrutura hierárquica com regiões conhecidas por planejamento de movimentos direcionando regiões visuais e motoras" (Lizier et al., 2011). No entanto, os autores também notam que "a baixa resolução temporal do fMRI coloca limitações na interpretação dos resultados da TE", e que "pode existir confusão em medidas de TE entre sinais de diferentes regiões quando se trata de interpretar estas ligações diretamente no nível neural. O atraso hemodinâmico pode diferir entre diferentes regiões corticais em até 1 segundo, o que pode, no pior dos casos, inverter a sucessão temporal uma vez que se compare a TE a partir de dados de BOLD e a partir da atividade neural real.

De fato, este é um problema já discutido nas aplicações de causalidade de Granger (sendo a transferência de entropia a sua versão não-linear, conforme afirmado anteriormente). Lohmann et al. (2013) apresenta uma revisão desta controvérsia, notando o argumento de Smith et al. (2011) de que "a variabilidade da resposta BOLD faz a interpretação de efeitos temporais sujeita a erros", uma vez que "diferenças inter-regionais na função de resposta hemodinâmica podem, de fato, existirem devido à vasculatura, ao invés de efeitos neurais" (Lohmann et al., 2013). Com isso, diversos estudos de simulação foram feitos tendo em vista investigar a escala destes atrasos de tempo entre regiões, incluindo a proposta de extensões para os métodos de GC que incluam uma função de resposta hemodinâmica explícita para considerar a variabilidade entre as regiões (Rodrigues e Andrade, 2014). Finalmente, o estudo por Seth et al. (2013) mostrou que a GC é robusta a uma grande variedade de mudanças na resposta hemodinâmica, no entanto, ela ainda está sujeita a erros na presença de muito ruído ou de *downsampling*. Por isso, Lohmann et al. (2013) sugere uma atitude de cautela em relação a abordagens baseadas em GC (e, portanto, TE).

Ide et al. (2014b) and Bielza e Larrañaga (2014) realizam uma revisão do uso da abordagem de redes Bayesianas em fMRI, incluindo comentários sobre DBNs. Crucial para esta abordagem é a escolha de métodos para a seleção de ROIs a serem usadas no modelo — as abordagens variam deste métodos baseados em dados, como *independent cluster analysis* (ICA), até métodos baseados em *seeds*, como de interação psicofisiológica (PPI ou *psychophysiological interaction* e baseados em ROIs, como o uso de análises de localização funcional ou atlas anatômicos do cérebro. As abordagens também são usadas em conjunto em vários

Tabela 4.1: Visão geral de métodos de conectividade neural direcionada para modelos dinâmicos em fMRI, de acordo com as distinções propostas por Lohmann et al. (2013) e Friston (2011). Uma característica adicional a ser notada é que os métodos de conectividade funcional apresentados são livres de modelo, enquanto os de conectividade assumem, a princípio (exceto suas extensões), modelos lineares gaussianos.

	Conectividade Efetiva	Conectividade Funcional
Padrões Espaciais	Causalidade de Granger (GC)	Transferência de Entropia (TE)
Redes Complexas	Modelagem Causal Dinâmica (DCM)	Rede Bayesiana Dinâmica (DBN)

casos (Ide et al., 2014b). As DBNs, no entanto, devem ser consideradas como uma abordagem de conectividade funcional uma vez que investiga relações estatísticas entre os sinais (Friston, 2011).

Na aplicação de DBNs em fMRI, os modelos utilizados representam ou dados discretos, com tabelas de distribuição de probabilidade (como implementado por Rajapakse e Zhou (2007) e Burge et al. (2009)) ou dados contínuos, através da suposição de gaussianidade (como em Zhang et al. (2005), Chen et al. (2012) e Wu et al. (2014)). A vantagem do caso discreto é de que os sistemas não-lineares podem ser modelados, mas com o custo da redução do sinal original a níveis discretos e a necessidade de mais amostras à medida que o número de níveis discretos aumenta. Já no caso contínuo, o sistema é considerado linear, porém os dados não precisam ser discretizados.

Em resumo, quando se trata de modelos dinâmicos com conectividade direcionada, a Tabela 4.1 agrupa os diferentes tipos de abordagem descritas de acordo com as distinções feitas por Lohmann et al. (2013) e Friston (2011).

4.2 Identificação de sistema neuromotores de insetos

O estudo dos sistemas neuromotores dentro do contexto de neurociência envolve a compreensão de como partes do sistema nervoso de um organismo codificam uma grande quantidade de informações sensoriais (do ambiente e do próprio corpo) e traduzem-nas em diversos comportamentos dos membros (Burrows, 1996). Tal conhecimento não é apenas interessante do ponto de vista biológico, mas também conta com diversas aplicações em engenharia e robótica bioinspirada. Como exemplo, um pouco de reflexão acerca da locomoção por patas ou pernas revela que esta é uma forma altamente versátil e flexível, e que pode ser buscada para robôs, no futuro — Buschmann et al. (2015) argumentam que, embora o emprego de rodas em carros

ou robôs permita altas velocidades em terrenos planos, a locomoção com pernas provê estabilidade e um movimento eficiente em grande variedade de ambientes, sendo estes muitas vezes complexos e imprevisíveis. Não obstante, conforme observado na introdução, estes modelos também pode orientar o desenvolvimento de próteses, na área de reabilitação (Jimenez-Fabian e Verlinden, 2012; Ying et al., 2007)) e aplicações de interação humano-máquina (Cannan e Hu, 2011).

Assim, uma grande diversidade de estudos tem se focado em descobrir e entender as arquiteturas de controle responsáveis tanto pela integração multimodal dos sensores como pela codificação em movimentos com tantos graus de liberdade, presentes nos sistemas neuromotores (Wessnitzer e Webb, 2006). Para isso, os insetos tem sido amplamente usados nestes estudos devido à simplicidade de seus sistemas neurais (se comparados aos de um mamífero), tal como o fato destes serem distribuídos ao longo do corpo do animal (Burrows, 1996), facilitando tanto a compreensão como a condução de experimentos de obtenção de sinais neurais para posterior análise. A investigação pode ser feita em vários níveis de abstração, e estudos revelaram uma divisão do sistema nervoso, relativamente identificável, em subsistemas responsáveis por funções específicas no controle do comportamento motor: em geral, as partes anteriores do sistema nervoso, geralmente chamadas de sistema nervoso central, são responsáveis por receber e processar informações do ambiente através de visão, audição, tato e quimioceptores, além dos estados internos do sistema, enviando a seguir comandos a sistemas de mais baixo nível responsáveis por gerar comportamentos motores como iniciação e manutenção do passo, manutenção de postura e equilíbrio, e alterações de velocidade e direção (Buschmann et al., 2015). Estes últimos sistemas, por sua vez, enviam comandos aos controladores locais dos membros, que por sua vez também retornam informações sensoriais que informam a carga e força exercida nos membros, como também sua posição, velocidade e aceleração (Fouad et al., 2002).

4.2.1 Descrição do sistema neuromotor do gafanhoto

Entre os estudos focados no controle local de membros, observa-se a recente atenção dirigida ao estudo da pata traseira dos gafanhotos da espécie (*Schistocerca gregaria*) e o papel da propriocepção no controle dos reflexos de seus membros. A propriocepção ou cinestesia consiste nas informações posicionais que são encaminhadas para um sistema nervoso central através de receptores encontrados nos músculos, tendões, ligamentos ou articulações; e que

Figura 4.2: Localização e morfologia do órgão femoral cordotonal (*FeCO*) em um gafanhoto. (a) indica a região dos neurônios sensores, que detectam os movimentos femorotibiais, e (b) indica estrutura neural do *FeCO* englobando o apódema, os dentritos e o axônio. [Retirada de Burrows (1987) e Endo (2014)].

tem como função tornar o organismo consciente de seus movimentos (Fouad et al., 2002). No caso dos membros de um gafanhoto, tal função se dá através das *sensilas campaniformes* e o órgão femoral cordotonal, ou *FeCO* (*Femoral Chordotonal Organ*) (Burrows, 1996). As sensilas campaniformes assemelham-se a discos ovais planos, encontrados em todo o corpo do inseto, e capturam flexões no exoesqueleto. Já o *FeCO* identifica movimentos de flexão e extensão na junção femorotibial do membro anterior distal do inseto (Burrows, 1996), podendo estar relacionada aos movimentos nos tendões e ligamentos, também denominados *apódemas* (Vidal-Gadea et al., 2010).

A localização e morfologia do *FeCO* está descrita na Figura 4.2.

O estudo de Burrows (1996) sugere uma classificação dos neurônios encontrados nos sistemas neuromotores do gafanhoto. Esta se baseia tanto na forma dos sinais gravados como na função que desempenham dentro do sistema, sendo:

- Neurônios sensores: são os neurônios responsáveis por detectar características do ambiente e dos próprios músculos do animal e transmitir estímulos ao sistema nervoso central, sinalizando estas informações. No caso do sistema neuromotor proprioceptor, os neurônios sensores encontram-se nas sensilas campaniformes e no *FeCO*.
- Neurônios motores: são os neurônios que recebem estímulos de outros neurônios e as-

Figura 4.3: Localização e morfologia do gânglio metatorácico e dos neurônios motores do membro anterior distal dentro do sistema nervoso (em escala) de um gafanhoto do deserto (*Schistocerca gregaria*) [Adaptada de Pflüger e Duch (2011) e Burrows (1996)].

sim excitam os músculos do animal, produzindo movimento. Estes são bem mais simples de identificar uma vez que ficam inervados no músculo a que se relacionam (Burrows, 1987). A partir disso, pode-se subdividir os neurônios ligados a um músculo de acordo com o tipo de contração que causal ou a posição das células no sistema nervoso central (Burrows, 1996, apud Endo, 2014). A Figura 4.2 relaciona os neurônios motores encontrados no membro posterior do gafanhoto.

- Interneurônios: no caminho dos neurônios sensores para os motores existem vários outros neurônios responsáveis por realizar o processamento das informações, traduzindo os sinais recebidos dos neurônios sensores (e de outras partes do sistema nervoso) em sinais enviados aos neurônios motores, que produzirão os movimentos. Pode-se classificar estes neurônios de acordo com o tipo de sinal que transmitem, sendo eles os interneurônios *spiking*, ou seja, que transmitem sinais em pulsos (*spikes*), e os interneurônios *non-spiking*, que operam de forma analógica, sem pulsos.

Tanto os neurônios motores como os interneurônios *spiking* e *nonspiking* localizam-se no gânglio metatorácico do gafanhoto, conforme ilustrado na Figura 4.3.

Tabela 4.2: Descrição dos neurônios motores do sistema neuromotor do gafanhoto, conforme também indicados na Figura 4.3(b) (Maciel et al., 2012)

Movimento no joelho:	
ASFITi	Neurônio motor flexor anterior lento da tíbia
AFFITi	Neurônio motor flexor anterior rápido da tíbia
AIFITi	Neurônio motor flexor anterior intermediário da tíbia
PSFITi	Neurônio motor flexor posterior lento da tíbia
PFFITi	Neurônio motor flexor posterior rápido da tíbia
PIFITi	Neurônio motor flexor posterior intermediário da tíbia
SETi	Neurônio motor extensor lento da tíbia
FETi	Neurônio motor extensor rápido da tíbia
Movimento do tarso:	
SLeTa	Neurônio motor levantador lento do tarso
FLeTa	Neurônio motor levantador rápido do tarso
SDTa	Neurônio motor depressor lento do tarso
FDTa	Neurônio motor depressor rápido do tarso
Movimento da garra:	
SRU	Neurônio motor retrator lento da garra
FRU	Neurônio motor retrator rápido da garra
Movimento do trocânter:	
FLeTr	Neurônio motor levantador rápido do trocânter
FDTr	Neurônio motor depressor rápido do trocânter

4.2.2 Circuitos e conectividade neural

O trabalho de Burrows (1996) propõe uma topologia genérica para o sistema neuromotor proprioceptor do membro posterior do gafanhoto, a qual está na Figura 4.4. No entanto, tais representações ainda são genéricas, ou seja, não representam os caminhos neurais para os neurônios específicos como os do *FeCO* ou os neurônios motores apresentados na Tabela 4.2. Por esta razão, (Kondoh et al., 1995) realizou uma série de gravações intraneurais do potencial sináptico nos neurônios no sistema proprioceptor, as quais foram analisadas em diversos trabalhos.

Kondoh et al. (1995) inicialmente realizou gravações intracelulares nos neurônios sensoriais do *FeCO*, responsável pelo controle proprioceptivo na junta femoro-tibial ("joelho") do gafanhoto, e após aplicar uma modelagem baseada em kernel de Wiener (Watanabe e Stark, 1975), descobriu que tais neurônios codificam a posição, velocidade e aceleração dos movimentos da tíbia em relação a esta junta. Em seguida, o trabalho de (Vidal-Gadea et al., 2010) modelou as respostas dos interneurônios no sistema nervoso central responsáveis pela propriocepção utilizando identificação de sistemas baseada em kernel de Wiener com estimação de parâmetros com validação cruzada. (Dewhirst et al., 2013) observaram que, embora tais abor-

Figura 4.4: Topologia do sistema neuromotor do membro posterior do gafanhoto. Observa-se os interneurônios *spiking* realizando conexões inibitórias com os *nonspiking*, e estes, por sua vez, se conectam aos neurônios motores. [Retirada de Burrows (1987) e Endo (2014)]

dagens ofereceram melhores estimativas de respostas do que as dos estudos anteriores, estas ainda estão sujeitas a limitações devido à dinâmicas de adaptação e não-linearidades. Os autores então propõem uma melhoria no método com uma estimação de parâmetros para modelos de Volterra com complexidade reduzida por meio de funções base de Laguerre, aplicando-os a uma população de interneurônios locais *non-spiking* responsáveis pelo controle pré-motor dos movimentos da pata traseira.

Nos casos citados acima, a estratégia empregada para a compreensão dos sistemas neuromotores é identificação de sistemas, nos quais busca-se melhor adequar um tipo de modelo aos dados observados e, em seguida, interpretar seus parâmetros, ligando-os à funcionalidade do sistema representado (Marmarelis, 2004) ou a seu esquema de codificação (Brown et al., 2004). As abordagens também são caracterizadas por empregar modelos baseados apenas em entrada e saída ("caixas-pretas", entre estímulo e resposta), sem considerar estados internos do sistema. Neste sentido, o recente trabalho por (Meruelo et al., 2016) demonstra como a inclusão de estados internos no modelo, através de redes neurais artificiais, pode melhorar o modelo. Vale também notar que o uso de modelos probabilísticos com estados internos, como os modelos ocultos de Markov (Stratonovich, 1960) e os filtros de Kalman (Kalman, 1960) (especificamente, os *unscented Kalman filters*, versões não-lineares dos filtros de Kalman (Wan e Van Der Merwe, 2000)) também são uma possibilidade para melhorar este tipo de modelagem.

Ressalta-se, no entanto, que outro tipo de abordagem desenvolvida recentemente busca compreender sistemas neurais a partir de modelos de redes complexas (Mitchell, 2009; Bettencourt et al., 2008; van Steen, 2010). A abordagem também mostra-se mais interessante do que

o uso de redes neurais artificiais uma vez que os estados internos do modelo são conhecidos e identificados com a própria estrutura interna do sistema modelado.

O trabalho de (Maciel et al., 2012), por exemplo, realizou medidas de informação mútua atrasada entre os sinais gravados e encontrou a estrutura de grafo apresentada na Figura 4.5, que também agrupa os sinais *nonspiking* em duas categorias, de acordo com os diferentes atrasos de tempo encontrados entre eles e o sinal de excitação no FeCO. Endo et al. (2015), por sua vez, utiliza um método de *clustering* para agrupar os neurônios *spiking* e *nonspiking* do sistema de acordo com camadas, a partir destes diferentes atrasos e dos valores de informação mútua atrasada. Os resultados obtidos confirmam a literatura de biologia: Burrows (1996) afirma que há dois padrões de caminho nos interneurônios *spiking* e três nos *nonspiking*.

A partir destes resultados, portanto, propõe-se uma estrutura do circuito neural do sistema neuromotor proprioceptor que especifica ainda mais os resultados obtidos por Maciel et al. (2012). Tal estrutura encontra-se na Figura 4.6. Conseguiu-se, assim, uma descrição mais detalhada do sistema, embora que ainda resume o comportamento de um grande número de neurônios (os quais seria impossível de mapear individualmente).

Tais resultados são importantes para a aplicação de medidas de DTE, a ser apresentada na parte experimental (seções 5.3 e 6.3) e, futuramente, uma modelagem mais precisa do sistema com uma *DBN*.

Por fim, a Tabela 4.3 resume, assim, das abordagens utilizadas nos estudos sobre o sistema proprioceptor na junta femoro-tibial do gafanhoto.

Figura 4.5: Estrutura de grafo do sistema neuromotor do gafanhoto obtida a partir de gravações de sinais realizadas por Kondoh et al. (1995) e medidas de informação mútua atrasada entre estes. Os valores nas arestas indicam o tempo médio dos atrasos entre os nós, em ms. As arestas em vermelho indicam conexões nas quais ainda não há dados. Linhas sólidas são o conhecimento já dado da fisiologia, e linhas pontilhadas indicam os caminhos inferidos pelas medições sobre os dados. [Retirada de Maciel et al. (2012)]

Tabela 4.3: Resumo das abordagens utilizadas nos estudos sobre o sistema proprioceptor na junta femoro-tibial do gafanhoto. No caso das análises baseadas em identificação de sistemas, busca-se melhor adequar um tipo de modelo aos dados observados e, em seguida, interpretar seus parâmetros, ligando-os à funcionalidade do sistema representado (Marmarelis, 2004) ou a seu esquema de codificação (Brown et al., 2004). Já na abordagem de obtenção de redes efetivas, busca-se a estrutura interna do sistema através de medidas de transferência de informação entre os elementos neurais dos quais obtêm-se gravações. Por fim, a abordagem baseada em redes Bayesianas dinâmicas implementa ambas as abordagens uma vez que utiliza-se de uma estrutura de rede para inferir a resposta do sistema a partir da entrada.

Tipos de Abordagens	
Identificação de Sistemas	Redes Efetivas com Transferência de Informação
Modelos de Wiener-Volterra (Kondoh et al., 1995) (Vidal-Gadea et al., 2010; Dewhirst et al., 2013)	Informação mútua atrasada (Maciel et al., 2012; Endo et al., 2013)
Redes Neurais Artificiais (Meruelo et al., 2016)	<i>Clustering</i> dos atrasos para identificação de caminhos neurais (Endo et al., 2015)
Redes Bayesianas Dinâmicas (ambas) - proposta atual	

Figura 4.6: Grafo com as conexões do gânglio metatorácico do sistema neuromotor do gafanhoto encontrada a partir dos experimentos de medidas de atraso com informação mútua atrasada. Os sinais 200Hz e 27Hz são os geradores de excitação utilizados no experimento, o nó *Sensory* é o neurônio sensor encontrado no *FeCO*, os interneurônios são apresentados a partir dos grupos encontrados no *clustering*, e os neurônios motores encontram-se divididos em flexores e extensores. [Retirada de Endo (2014)].

Capítulo 5

Materiais e métodos

Após apresentar os desafios concernentes à modelagem de sistemas neurais com DBNs, tanto do ponto de vista conceitual como em aplicações, o presente capítulo passará a descrever os materiais e métodos experimentais utilizados durante o trabalho. Estes incluem, primeiramente, simulações e algoritmos para a aprendizagem de DBNs em sistemas parcialmente observados com medidas de teoria de informação, bem como um método de busca no espaço de HO-DBNs baseado em PSO. Em seguida, passa-se aos experimentos cuja base teórica foi estabelecida no capítulo anterior (4): a modelagem por DBNs da conectividade funcional entre sub-áreas do hipocampo (seção 5.2) e a realização de medidas de transferência de entropia em gravações de potencial sináptico no sistema neuromotor do gafanhoto (seção 5.3).

Durante o trabalho buscou-se, sempre que possível, utilizar-se a linguagem *Python*¹, versão 2.7.5, em ambiente *Linux*, com a distribuição Ubuntu versão 13.10. Também através do software *iPython* muitos experimentos foram conduzidos (Pérez e Granger, 2007), incluindo a execução paralela de alguns algoritmos. As únicas exceções ocorreram na geração de conjuntos de dados a partir de redes Bayesianas simuladas, na seção 5.1.1, e na utilização do software de pré-processamento SPM12 e a implementação do algoritmo GlobalMIT, na seção 5.2, também em Matlab.

---

¹Disponível em: <http://www.python.org/>

5.1 Simulações

5.1.1 Aprendizagem estrutural de redes Bayesianas com nós de mesmo tipo

Na seção 3.3 também foi sugerida uma simplificação nos algoritmos de aprendizagem estrutural capaz de lidar tanto com a "maldição da dimensionalidade" devido a altas ordens em funções de distribuição de probabilidade conjunta quanto com conjuntos de dados onde as variáveis não são todas observadas simultaneamente. Isso pode ser feito uma vez que se considerem redes onde nós têm comportamentos iguais ou semelhantes — conceito que foi interpretado como *CPDs* (distribuições de probabilidade condicional) fixas para os nós. Isso sugere que os valores de informação mútua sejam separados de acordo com a diretividade entre duas variáveis, o que permite uma inferência desta última característica a partir da primeira.

Assim, realizaram-se alguns experimentos para checar em quais casos estas suposições são válidas, e para isso usou-se dados de redes Bayesianas simuladas com nós representando portas lógicas OU-ruidosas (*Noisy-OR*). Assim, a informação mútua entre dois nós é calculada (sendo a diretividade um valor dado), e uma análise baseada em teoria de decisão Bayesiana é conduzida para avaliar-se a eficácia de uma classificação da diretividade entre dois nós com base nos valores de informação mútua.

Os algoritmos foram executados em paralelo, através do ambiente iPython, em um cluster consistindo em oito máquinas com Ubuntu Server 14.04.1 LTS, conectadas em uma rede local Ethernet de 1Gbps, com *CPUs* Intel(R) Core(TM) i7-3770 e 8Gb de RAM (em média).

Redes Bayesianas simuladas com nós OU-ruidosos

O dados simulados são gerados a partir de redes Bayesianas cujos nós representam portas lógicas OU-ruidosas. Isso significa que os *CPDs* são tais que o valor da saída tem alta probabilidade no resultado esperado em uma operação de OU lógico nas entradas, e baixa probabilidade nos outros resultados possíveis (representando um erro na operação). A Figura 5.1(a) ilustra o caso de um *CPD* em uma rede com variáveis discretizadas com três *bits* e permitindo, no máximo, dois pais. A razão para se escolher uma porta lógica OU está no fato desta não envolver sinergia nos valores de entrada, o que não seria capturado por medidas par-a-par. As *CPDs* são geradas de acordo com o número máximo de pais permitidos para cada nó, e como indicado na seção 3.3, um ruído uniforme é introduzido como pai de um nó que, na estrutura de rede, tem

Figura 5.1: Redes Bayesianas com nós OU-ruídos utilizada para o experimento — um exemplo de rede Bayesiana com CPDs fixos. (a) ilustra o cálculo de um *CPD* para nós com no máximo dois pais, três *bits* de discretização e uma entropia de base de $[0.7, 0.3]$. (b) é um exemplo de rede aleatória gerada para este *CPD* fixo, ressaltando o fato de que um nó com menos de dois pais recebe ruído uniforme como pais.

menos pais do que o máximo permitido. Isso é indicado na Figura 5.1(b).

Com diferentes configurações de *bits* de discretização, número máximo de pais e entropia de base utilizada para os CPDs, são gerados DAGs aleatórios, nos quais escolhem-se dois nós aleatoriamente e calcula-se o tamanho do menor caminho não-direcionado entre os dois nós no seu grafo de influência de informação (*IIG*) — definido na seção 3.3 como a diretividade entre dois nós. Em seguida, dados são gerados para esta rede (excluindo os nós que não têm influência; ou seja, d-separados dos dois nós em questão) e a informação mútua entre os nós é calculada. A Figure 5.2 mostra o procedimento usado para o experimento.

A razão para o uso de dados simulados com redes aleatórias se deve aos fatos de que 1) o número de *IIGs* pode ser enorme, e cresce rapidamente com o valor de diretividade entre dois nós e o número máximo de pais permitidos, tornando infactível a avaliação de cada um deles; e 2) por causa do número de nós envolvidos em alguns *IIGs*, um cálculo em forma fechada dos valores de informação mútua a partir dos parâmetros da *CPD* é ainda desconhecido (e poderia consistir uma futura direção de pesquisa).

Os parâmetros são variados da seguinte forma: 50 redes Bayesianas aleatórias com 15 nós são geradas para cada valor de diretividade entre dois nós aleatórios, variando de 1 a 4, para as seguintes entropias base para os *CPDs*: $(0.6, 0.4)$, $(0.7, 0.3)$, $(0.8, 0.2)$ e $(0.9, 0.1)$, e para as seguintes configurações de números de *bits de discretização* (r) e máximo número de pais nos *CPDs* (m_p), respectivamente: $(2, 2)$, $(2, 3)$, $(2, 4)$, $(3, 2)$, $(3, 3)$ e $(4, 2)$. O número de amostras (N_s) geradas para cada rede é dada pela equação $N_s = r^{m_p} * 100$.

Figura 5.2: Uma descrição dos experimentos conduzidos para se checar uma separação nos valores de informação mútua entre dois nós em uma rede a partir da diretividade entre eles. Como o espaço de possíveis *IIGs* é muito grande, e um cálculo direto de informação mútua entre dois nós baseado apenas nos parâmetros dos *CPDs* ainda é desconhecido, decidiu-se utilizar dados simulados e redes aleatórias no experimento.

Teoria de decisão Bayesiana

Utiliza-se teoria de decisão Bayesiana para se classificar os valores de informação mútua de acordo com a diretividade nos *IIGs*. Seguindo a explanação por Duda et al. (2001), defina-se ω como uma variável descrevendo uma classe à qual deseja-se associar uma observação x , e ω_i como as classes possíveis ($i=1, \dots, n$ sendo n o número de classes). Assim, tem-se $P(\omega_i)$, a *probabilidade a priori*, $P(x|\omega_i)$, a *verossimilhança* e $P(x)$ o *fator de normalização*. Então, de acordo com a teoria de decisão Bayesiana, para uma observação x , tende-se a atribuir a classe na qual a probabilidade *a posteriori* $P(\omega_i|x)$ é a maior. Esta probabilidade, por sua vez, é calculada através de fórmula de Bayes, em que:

$$P(\omega_i|x) = \frac{P(x|\omega_i)P(\omega_i)}{P(x)}$$

Pode-se também estimar as probabilidades de erro ao se decidir por uma classe errada em uma observação. Estas são as probabilidades *a posteriori* $P(\omega_i|x)$ quando a decisão por x é diferente de ω_i . Todas estas, no entanto, são probabilidades condicionadas a uma única observação x ; de forma que para se estimar o erro ou acerto médio para uma dada classe, integra-se a probabilidade *a posteriori* probabilities em todos os valores que x pode assumir.

No presente experimento, as observações x são o logaritmo do valor de informação mútua encontrado entre dois nós. A razão para a aplicação do logaritmo é que, ao olhar-se para a distribuição dos valores de informação mútua, nota-se que esta ocorre exponencialmente de acordo com o valor de diretividade entre os nós. A aplicação do logaritmo, portanto, garante melhores resultados na estimação das distribuições de probabilidade.

As observações podem assumir quatro classes, com cada uma indicando a diretividade

entre dois nós, variando de 1 a 4. Estas são definidas como w_1 , w_2 , w_3 and w_4 . Em seguida obtém-se a verossimilhança $P(x|\omega_i)$ a partir de uma estimativa de densidade kernel (*KDE*, do inglês *kernel density estimation* nas observações de cada classe. O fator de normalização $P(x)$ é calculado da mesma forma (mas, neste caso, considerando-se as observações em todas as classes). Consideram-se também *prioris* não informativas, ou seja, $P(\omega_i)=0.25, \forall i$. Por fim, calculam-se as probabilidades *a posteriori* com a fórmula de Bayes formula e as probabilidades média de sucesso e erro para cada classe são obtidas, sendo apresentadas na seção de resultados (6.1.1).

5.1.2 Aprendizagem estrutural de HO-DBNs

Na seção 3.4.1 foi sugerida uma estratégia de busca no espaço de *HO-DBNs* baseada em otimização por enxame de partículas (*PSO*). Na presente seção, portanto, serão descritos os experimentos realizados para checar a validade e eficiência do algoritmo proposto.

Para a realização de testes com os algoritmos desenvolvidos, houve a necessidade de gerar conjuntos de dados amostrados a partir de redes Bayesianas simuladas. Como não foi possível encontrar uma biblioteca capaz de realizar esta tarefa em Python, utilizou-se o *software MATLAB* em ambiente *Windows 8*, juntamente com a biblioteca *Bayes Network Toolbox* (*BNT*) (Murphy, 2001). *BNT* é uma biblioteca que provê ferramentas para o trabalho com redes Bayesianas de diferentes tipos, tal como algoritmos de aprendizagem, inferência e amostragem. Embora a biblioteca apresente ferramentas para redes Bayesianas dinâmicas, ela não suporta uma representação direta de *HO-DBN*, no entanto, isto não prejudicou a tarefa de gerar dados amostrados, uma vez que bastou acrescentar o instante de tempo de um nó ao seu nome, e evitar arestas proibidas no modelo. Assim, uma rotina foi criada para a geração dos dados de *HO-DBNs* simuladas, e também paralelizada para rodar em quatro *cores* através do *MATLAB*.

O algoritmo de aprendizado estrutural de *HO-DBNs* baseado em *PSO*, descrito na seção 3.4.1, foi implementado e testado com o apoio das bibliotecas *PEBL* (Shah e Woolf, 2009) e *DEAP* Fortin et al. (2012). *PEBL*, ou *Python Environment for Bayesian Learning* provê classes, estruturas de dados, algoritmos e utilidades em geral para a aprendizagem estrutural de redes Bayesianas. *DEAP*, ou *Distributed Evolutionary Algorithms in Python*, embora tendo sido desenvolvido primariamente para algoritmos evolutivos, inclui ferramentas úteis para o desenvolvimento de quaisquer outros algoritmos bioinspirados.

A implementação foi realizada estendendo a classe *Learner*, já implementada na biblioteca

PEBL. Assim, cria-se a classe *HODBNPSOLearner*, com novos parâmetros de configuração que refletem os parâmetros de controle do algoritmo *PSO*, e a redefinição de funções de inicialização da classe e execução do algoritmo. Na implementação do algoritmo de *PSO*, algumas utilidades da biblioteca *DEAP* foram inseridas para estimativas de valor máximo e mínimo obtidos na função de pontuação. Além da classe *Learner*, também houve uma extensão da classe *Evaluator*, utilizada para o cálculo da pontuação da rede. A função de pontuação utilizada na biblioteca é a métrica *K2*, descrita na equação 2.5; e uma adaptação foi feita para reduzir a quantidade de cálculos considerando apenas os nós no instante de tempo t (já que são os únicos que podem ter pais, em uma *DBN*).

Para fins de comparação, também foi implementado um algoritmo de aprendizagem estrutural de *HO-DBNs* com uma estratégia *greedy*. O algoritmo começa com uma rede aleatória e a cada iteração, propõe a adição ou remoção de uma aresta, que é aceita caso a pontuação da nova rede seja maior que da antiga. Para evitar os máximos locais, o algoritmo também verifica a quantidade de iterações sem melhoria na pontuação, e quando esta atinge um valor limite, a rede atual é trocada por uma nova rede aleatória. A razão do uso de apenas um algoritmo *greedy* também se deve à ausência de outras abordagens de aprendizagem de máquina disponíveis na literatura para o tipo de modelo em questão.

Como *HO-DBNs* são ainda propostas recentes na literatura, não é possível encontrar conjuntos de dados já estabelecidos para realizar *benchmarks* e testar os algoritmos implementados para performance e corretude. Assim, é proposta a geração de alguns conjuntos de dados a partir de *HO-DBNs* (*intra-timeslice*) simuladas. Iterando-se o número de variáveis, ordens de Markov da rede, e número de *bits* nas variáveis discretas, foram geradas quatro redes para cada configuração possível: 1) uma rede completamente conectada (ou seja, arestas saindo de todos os nós em instantes de tempo passados para todos os nós no instante atual); 2) uma rede sem arestas; 3) uma rede com poucas arestas (30% do número máximo de arestas possíveis) e 4) uma rede com muitas arestas (70% do número máximo de arestas possíveis).

É importante notar que à medida que o número de *bits* nas variáveis discretas cresce, o tamanho do conjunto de dados requerido para realizar o treinamento torna-se muito grande, levando a casos onde é impossível realizar o treinamento, devido a limites de memória ou mesmo tempo de processamento. O tamanho do conjunto de dados requerido também cresce com o aumento do número de variáveis e da ordem de Markov, e leva a complicações no treinamento. Para estimar o número de amostras N_S no conjunto de dados necessárias para o

treinamento, temos, para um sistema discretizado em d bits, N variáveis, e ordem τ :

$$\eta \times d^{N(\tau+1)} \quad (5.1)$$

onde η é um fator usado para fornecer uma margem de probabilidade para cada possível amostra — e no presente trabalho, é adotada como 10.

A Figura 5.3 apresenta a quantidade de amostras necessárias para uma rede simulada com cada combinação de número de variáveis e ordem de Markov, junto com uma estimativa do tamanho ocupado pelo conjunto de dados. No presente trabalho, devido às limitações de memória e tempo (pois o algoritmo que gera os dados demora para executar), resolveu-se realizar os testes com redes simuladas cujos conjuntos de dados de treinamento se limitassem a 150MB. Dessa forma, embora não existam experimentos com muitas variáveis e ordens muito altas, ainda é possível fazer uma estimativa do tempo requerido, por extrapolação, para o treinamento de redes maiores.

O algoritmo de *PSO* executou, em todos os casos, com um tamanho de população igual a 100, com 80 gerações, máxima velocidade permitida igual a 3, e os parâmetros social e cognitivo iguais a 1. Para cada caso de teste, o algoritmo *greedy* executou com 4000 iterações. Devido ao fato que os métodos envolvem aleatoriedade, repetiu-se cada execução 35 vezes, medindo-se o tempo para a convergência do algoritmo até às soluções esperadas e calculando a média destes tempos. Os experimentos também requerem bastante tempo para execução, de forma que foram executados apenas os casos com discretização em 2 bits.

5.2 Conectividade neural entre sub-áreas do hipocampo em estado de repouso

A seguir encontram-se os métodos para a obtenção de conectividade funcional entre sub-áreas do hipocampo através de DBNs. Dois tipos de análise são realizadas: a primeira obtém uma DBN com apenas 3 ROIs para indivíduos separados, enquanto a segunda utiliza ROIs de diversas partes do cérebro, obtidas a partir de uma análise de correlação *seed to voxel*, e em seguida obtém uma estrutura de DBN mais completa a partir de uma análise em grupo de sujeito típico virtual (VTS).

(a) Discretização com 2 bits

(b) Discretização com 3 bits

Figura 5.3: Quantidade de amostras e tamanho em *bytes* requerido para os conjuntos de dados de treinamento das *HO-DBNs* simuladas (em escala logarítmica). N é o número de variáveis do sistema. Para o presente trabalho, limitou-se o tamanho dos conjuntos de dados em 150MB — indicado pela linha horizontal —, utilizando-se apenas os casos abaixo da linha para testes e simulação.

5.2.1 Aquisição das imagens

A sequência de pulso EPI multibanda acelerada desenvolvida no Centro de Pesquisa em Ressonância Magnética da Universidade de Minnesota (EUA) (Moeller et al., 2010b) foi instalada e utilizada no sistema de MRI humano de 7T, no Departamento de Radiologia da Universidade de Pittsburgh (EUA). A aquisição é feita com uma bobina de 20-a-8 canais de transmissão e 32 canais ajustáveis de recepção desenvolvida por Ibrahim et al. (2013). O equipamento é mostrado na Figura 5.4a, e em experimentos iniciais, foi capaz de obter imagens com resolução de até 1,5mm isotrópico (ou seja, cubos cujos lados têm tamanho de 1,5mm).

(a) 7T MRI Scanner (Magnetom, Siemens Medical Systems)

(b) bobina de radiofrequência de 20-a-8 canais de transmissão e 32 canais ajustáveis de recepção

Figura 5.4: Equipamentos de MRI utilizado para aquisição dos dados no Departamento de Radiologia da Universidade de Pittsburgh (EUA). (a) demonstra o sistema de MRI humano de 7T (Magnetom, Siemens Medical Systems, Erlangen, Germany), e (b) a bobina de radiofrequência de 20-a-8 canais de transmissão e 32 canais ajustáveis de recepção (Ibrahim et al., 2013))

Imagens de fMRI 4D de todo o cérebro (excluindo o cerebelo), em orientação axial, whole-brain (excluding the cerebellum) são obtidas por meio de uma sequência EPI de 155 instantes de tempo e parâmetros: tamanho de voxel de 2,3mm isotrópico, 42 lâminas entrelaçadas, $TR/TE=2500/20 \text{ ms}^2$, ângulo de *flip* de 70° , $FoV^3=1546 \times 1546$, matriz de aquisição de 96×96 e tempo de aquisição igual a 6 min e 27 s. 20 pacientes adultos, na faixa dos 50 e 60 anos, diagnosticados com depressão de início tardio e voluntários no projeto de pesquisa NIH 1R01MH111265-01, são escaneados enquanto permanecem em repouso com os olhos abertos, ouvindo música por meio fones de ouvido (para impedir que entrem em estado de sono), com as luzes apagadas.

Também são obtidas imagens T1 de todo o cérebro, em orientação axial, por meio de

²TR é o tempo de repetição (*repetition time*) e TE, o tempo de eco (*echo time*)

³Ou campo de visão (*field of view*)

uma sequência *Magnetization Prepared Rapid Gradient Echo* (MPRAGE) com orientação 3D, tamanho de voxel de 0.7mm isotrópico e parâmetros TR/TI⁴/TE=3000/1200/2.47 ms, ângulo de *flip* de 6°, FoV=192 × 220, matriz de aquisição 280 × 320 e tempo de aquisição igual a 10 min e 31 s.

Finalmente, para apenas 3 dos 20 pacientes (uma vez que a sequência foi implementada apenas posteriormente), imagens T2-*weighted* de todo o cérebro, em orientação axial, foram obtidas com uma sequência *Turbo Spin Echo* (TSE) 3D com tamanho de voxel 0.37mm × 0.37mm × 1.5mm, 32 lâminas entrelaçadas, TR/TE=7900/56 ms, ângulo de *flip* de 130°, FoV=165 × 189, e tempo de aquisição igual a 5 min e 25 s.

Conforme ressaltado na seção 4.1.1, o software ASHS, proposto por (Yushkevich et al., 2015b) e o atlas de 7T de (Wisse et al., 2016), consistem, até o presente momento, na opção mais viável para segmentação automática de sub-áreas do hipocampo com MRI de 7T. No entanto, uma vez que estes métodos exigem o uso de imagens T2-*weighted*, as quais estão disponíveis apenas para 3 dos 20 pacientes, eles são utilizados apenas para uma análise inicial feita como pacientes individuais, a fim de se checar a viabilidade e plausibilidade do trabalho. Em seguida, a análise em nível de grupo é feita com imagens de todos os pacientes através do método de segmentação automática oferecido no software Freesurfer 6, proposto por Iglesias et al. (2015)

5.2.2 Pré-processamento e aprendizagem de DBNs individuais

Apenas três ROIs são consideradas no modelo inicial: o giro dentado (DG), as quatro áreas do *cornu ammonis* (CA1-4) (considerados juntamente) e o córtex entorinal (ERC) de apenas um hemisfério do cérebro. A razão para o uso de apenas estas áreas se deve ao fato da necessidade de altas quantidades de dados necessárias à medida que o número de nós de uma DBN aumenta, e no presente caso, uma vez que apenas um indivíduo é considerado, apenas 155 amostras (ou seja, os instantes de tempo) estão disponíveis.

O pré-processamento e alinhamento das imagens é realizado com o software SPM12 (Penny et al., 2011). As imagens de fMRI tiveram as lâminas corrigidas de acordo com o tempo (*slice-time correction*) (com a lâmina do meio utilizada como referência para a interpolação), corrigidas com relação a movimentos (a imagem média foi usada como referência, com a métrica

---

⁴TI é o tempo de inversão (*inversion time*)

de similaridade baseada em informação mútua e interpolação com B-splines de quarta ordem) e suavizadas com um *kernel* Gaussiano (com largura à meia altura igual a 8mm). As imagens T2-weighted, juntamente com as ROIs obtidas pela segmentação automática, são alinhadas e reamostradas de acordo com a imagem de fMRI. Séries temporais são então extraídas dos voxels de acordo com as ROIs determinadas pela segmentação, e uma normalização é aplicada, com o valor médio da ROI em cada instante de tempo subtraído de cada série temporal e a amplitude variando de -1 a 1. Por fim, a média destas séries temporais é obtida, produzindo uma única série temporal para cada ROI, a ser utilizada como dado para a DBN (Ide et al., 2014a).

Estas séries temporais também geralmente apresentam não-estacionariedade e tendências de nível DC, impedindo a utilização dos instantes de tempo como amostras para a aprendizagem de DBNs (ou seja, a suposição de ergodicidade). Usualmente, estas tendências são consideradas irrelevantes para o modelo e são removidas por métodos de *detrending* linear (Ide et al., 2014a). No entanto, é importante considerar que tais tendências podem apresentar comportamentos não-lineares e descontinuidades, de forma a impedir que o sinal resultante fique estacionário. Assim, o método de *detrending* por SSA descrito na seção 2.3.1 é aplicado em cada série temporal para os ROIs.

Com isso, o algoritmo GlobalMIT para aprendizagem de DBNs discretas, apresentado na seção 3.2, e que se encontra implementado e publicado como um pacote de Matlab, é aplicado e a estrutura com a maior pontuação é obtida. Os resultados e discussão são apresentados na seção 6.2.1.

5.2.3 Pré-processamento e aprendizagem de DBNs em nível de grupo

A segunda análise expande a primeira ao incluir mais ROIs na estrutura de rede de DBN. As imagens estruturais T1 de todos os 20 indivíduos são segmentadas pelo software *Freesurfer*, obtendo os voxels para as ROIs correspondentes ao giro dentado (DG), *cornu ammonis* (CA) e subículo (Sub) dos hipocampos de cada hemisfério do cérebro. Estas então serão usadas como *seeds* para a análise em grupo da correlação com todos os outros voxels do cérebro (análise *seed-to-voxel*) indicando as áreas principais de ativação a serem usadas como ROIs adicionais no modelo de DBN.

A análise *seed-to-voxel* é feita da seguinte forma, utilizando o software SPM12: primeiramente, obtêm-se os voxels correspondentes à massa cinzenta e branca e fluido cerebrospinal

do cérebro através de uma segmentação da imagem T1, baseada em seis *priors* espaciais, resultando em um campo de deformação que pode ser usado para normalizar qualquer imagem alinhada com esta imagem T1. Também é feita uma correção de *bias* (regularização de luz) para a remoção de artefatos. Usando esta segmentação, é criada uma máscara de volume intracranial (ICV) e, com ela, uma imagem com o crânio removido. Nas imagens de fMRI, o mesmo procedimento de pré-processamento descrito na seção anterior é utilizado (*slice-time correction*, correção de movimento e suavização com um *kernel* Gaussiano).

As imagens estruturais com o crânio removidos são então alinhadas às imagens funcionais (fMRI), utilizando-se uma métrica de similaridade baseada em informação mútua e interpolação com B-splines de quarta ordem. Aplica-se então um *wavelet despiking* na imagem funcional, de acordo com um método previamente estabelecido (Patel et al., 2014), para a remoção de ruído. Então, séries temporais são extraídas da massa branca e fluido cerebrospinal.

Para cada ROI correspondente ao DG, CA e Sub de cada hemisfério, uma análise de componentes principais é realizada para gerar um coeficiente de correlação entre suas séries temporais e as séries temporais de cada outro voxel do cérebro (apenas da massa branca e fluido cerebrospinal). Isso resulta em seis mapas de conectividade para cada indivíduo. Estes mapas são então alinhados entre si, gerando um mapa de conectividade média entre todos os indivíduos. Um teste-T de única amostra é então utilizado para identificar os voxels nos quais a conectividade é significativamente maior do que zero. Isso se dá através de uma abordagem de teste por permutação (com 5000 perturbações aleatórias e a formação de *cluster* no limiar de $p < 0.005$). As análises são significantes a nível de *clusters* com $p < 0.05$.

Com isso, os voxels com maior conectividade com DG, CA e Sub são extraídos e alinhados com um atlas, obtendo rótulos para cada uma das regiões encontradas. Isso fornece as várias ROIs a serem usadas para o modelo de DBN, a ser obtido com o algoritmo GlobalMIT, após o mesmo procedimento descrito na seção anterior para a obtenção das séries temporais. Além disso, uma vez que a inclusão de mais nós requer o uso de mais dados, uma análise em grupo baseada em um indivíduo típico virtual (VTS) é utilizada, reunindo dados de todos os indivíduos para a obtenção de um só modelo de rede.

5.3 Transferência de entropia atrasada no sistema neuromotor do gafanhoto

A metodologia descrita a seguir foi desenvolvida tendo como objetivo a elicitção de conexões par a par entre as diferentes partes do sistema neuromotor do gafanhoto (descrito na seção 4.2.1) a partir de gravações intracelulares. Estas últimas apresentam duas dificuldades — suas características não-estacionárias e a ausência de todos os sinais gravados simultaneamente (no caso, cada experimento se limita a apenas quatro canais de gravação). Com isso em vista, aplicou-se um método de remoção de tendências não-lineares e descontinuidades baseado em análise de espectro singular (SSA), os quais estão presentes no sinal devido ao *setup* experimental, que introduz uma corrente DC que varia à medida que o eletrodo se move pelo corpo da célula (seção 2.3.1). Assim, medidas de transferência de entropia atrasada puderam ser aplicadas a trechos maiores dos sinais gravados — anteriormente, em Endo et al. (2015), apenas trechos pequenos do sinal, considerados estacionários, eram selecionados. A estas medidas foi aplicada também uma normalização, com a finalidade de comparar e integrar os resultados em diferentes experimentos e animais. Finalmente, um método de análise Bayesiana também foi desenvolvido para esta integração das medidas em diferentes experimentos, capaz de prover uma probabilidade *a posteriori* da presença de picos de transferência de informação entre dois sinais dado as várias evidências experimentais.

5.3.1 Aquisição dos sinais

As gravações foram feitas no trabalho de Kondoh et al. (1995), obtidas imobilizando-se os gafanhotos com uma massa de modelagem e deixando o apódema do *FeCO* exposto através de uma pequena abertura no fêmur anterior distal do inseto (Kondoh et al., 1995). O gânglio metatorácico é exposto com uma abertura no tórax ventral, removendo as bolsas de ar e o tecido conectivo; e em seguida microeletrodos de resistência DC de 50 a 80M Ω são conduzidos através da bainha para dentro dos processos dos neurônios *spiking* e *nonspiking*, nas regiões mostradas na Figura 4.3(b). O *FeCO* é então segurado por um fórceps que vibra com um sinal de ruído branco gaussiano (*Gaussian White Noise* — GWN) de 200 Hz, gerado através da filtragem de uma sequência binária pseudoaleatória limitada em banda a 27 Hz por um filtro passa-baixas de Chebyshev de quarta ordem (em outros experimentos, o sinal é limitado a 58 Hz). Os estímulos e as respostas dos neurônios são então armazenadas em fita magnética, e

Figura 5.5: Experimento conduzido por Newland e Kondoh (1997) para a coleta de sinais dos neurônios sensores, motores e interneurônios controlando a propriocepção na junção femoro-tibial do gafanhoto (e suas localizações no animal). Um sinal de ruído branco gaussiano de 27 Hz, submetido a um filtro passa-baixas de Chebyshev de 4ª ordem, balança um fórceps preso no apódema do órgão femoral cordotonal (FeCO). A resposta é gravada por microeletrodos introduzidos no corpo celular dos neurônios sensores, motores e interneurônios presentes no gânglio metatorácico.

amostradas a uma taxa de 24 kHz para um computador (Kondoh et al., 1995). Cada gravação apresenta no máximo quatro sinais simultâneos gravados no experimento. A Figura 5.5 ilustra o procedimento realizado para a obtenção dos dados e a Figura 5.6 apresenta uma foto do *setup* experimental.

Cada gravação possui no máximo 4 canais simultâneos. A Tabela 5.1 indica os canais em cada gravação utilizados para os experimentos.

5.3.2 Pré-processamento: Análise de Espectro Singular (SSA)

Uma característica comum em circuitos neurais e neurônios individuais é que eles geralmente se adaptam a uma estimulação repetitiva ou constante (Prescott e Sejnowski, 2008). Esta adaptação intrínseca se deve a muitas funções, incluindo a estabilização do controle dos membros (Kittmann, 1997), detecção de coincidências (Cook et al., 2003) e separação de sinais de diferentes escalas de tempo (Benda et al., 2005). Esta adaptação em uma escala de tempo de segundos, também já foi mostrada que ocorre nos mesmos circuitos de controle do gafanhoto estudados, em nível de neurônios sensoriais (Newland, 1991), interneurônios (Angarita-Jaimes

Figura 5.6: *Set up* experimental para as gravações realizadas no trabalho de Kondoh et al. (1995), no sistema neuromotor proprioceptor de gafanhotos.

et al., 2012) e neurônios motores (Burrows, 1996). Dewhirst et al. (2013), no entanto, mostrou que a dinâmica principal do sistema permanece inalterada durante uma estimulação repetitiva, e a maior parte desta adaptação ocorre durante os primeiros segundos do estímulo. Assim, seguindo-se as considerações apresentadas na seção 2.3, pode-se supor que os sinais que se deseja analisar apresentam não-estacionariedades somente devido a artefatos introduzidos pelo *setup* experimental, e a dinâmica intrínseca do sistema permanece estacionária durante a maior parte do sinal. O método de SSA descrito na seção 2.3.1, portanto, foi utilizado para a remoção de tendências e descontinuidades no sinal.

O algoritmo de SSA é aplicado aos canais de gravação dos sinais em cada experimento, e os sinais são reconstruídos desconsiderando-se o maior autovalor, o qual, sendo o menor componente oscilatório, pode ser entendido como contendo o nível DC, as tendências complexas e as descontinuidades da série temporal — os quais, no presente caso, são os artefatos experimentais. Para isso, utiliza-se o software SSA-MTM Toolkit (Vautard et al., 1992), com os parâmetros de tamanho de janela igual a 2000 amostras, decompondo o sinal em 64 componentes principais e reconstruindo-o com 63 (ou seja, excluindo o menor componente oscilatório). Os parâmetros são escolhidos a partir das limitações computacionais impostas pelo algoritmo e software utilizados.

5.3.3 Medidas de transferência de entropia

Para a implementação da medida de transferência de entropia dada na equação 3.6, nota-se que não há garantia que uma estimação de probabilidades com base no teorema de Takens (Wibral et al., 2014a) funcionaria para gravações intracelulares como as atuais, uma vez que ela pode envolver outros tipos de ruído além do branco gaussiano. Portanto, na presente análise realiza-se uma estimação direta das funções de distribuição de probabilidade a partir de histogramas, aproveitando a disponibilidade de um grande número de amostras nos sinais gravados. O número de *bins* ou níveis de discretização utilizados foi de 32, em todos os experimentos, baseado no tamanho médio das séries temporais de 320.000 pontos. No caso da discretização no tempo, o método do primeiro mínimo da função de auto informação mútua foi utilizado para a obtenção do *time lag* τ (vide seção 3.4 e Kantz e Schreiber (2004)).

A equação 3.11 é, assim, implementada fixando-se também $k=1$, sob a suposição de que o sistema modelado é Markoviano de primeira ordem. Com isso, esta equação para a obtenção dos atrasos de tempo entre dois sinais toma a seguinte forma:

$$\delta = \arg \max_{\beta} (TE(X \rightarrow Y, \beta)) = \sum_{\substack{\forall x \in \mathcal{X} \\ \forall y \in \mathcal{Y}}} p(y_n, y_{n-\tau}, x_{n-\beta}) \log \frac{p(y_n | y_{n-\tau}, x_{n-\beta})}{p(y_n | y_{n-\tau})} \quad (5.2)$$

As medidas então são feitas variando-se o atraso β da *DTE* de -105 ms a 105 ms (valores máximos estipulados para a presença de um atraso entre elementos do sistema).

10 sinais de *surrogate data* também foram gerados para cada sinal (vide seção 3.1), utilizando-se o método *IAAFT*, a fim de analisar-se a significância das medidas, e também devido à sua melhor capacidade de capturar o *bias* (conforme exposto na seção 3.1. O método foi implementado com auxílio da biblioteca FFTW3 (Frigo e Johnson, 1998), que aumentou a performance do código consideravelmente.

Os valores de transferência de entropia, no entanto, variam de acordo com o experimento e o número de amostras empregado, o que influencia as funções de densidade de probabilidade. Portanto, um procedimento de normalização de medidas deve ser aplicado de forma a permitir uma comparação dos resultados entre os experimentos e, assim, permitir sua integração. Gourevitch e Eggermont (2007) propõe que a normalização das medidas deve ser realizada pela diferença entre a transferência de entropia no dado original e a média das transferências de entropia com os dados de *surrogate*, dividida pela entropia do valor presente da série Y dado o

seu passado, ou seja, $H(Y_n|Y_{n-\tau})$:

$$H(Y_n|Y_{n-\tau}) = - \sum_{\substack{\forall x \in \mathcal{X} \\ \forall y \in \mathcal{Y}}} p(y_n, y_{n-\tau}) \log \frac{p(y_n, y_{n-\tau})}{p(y_{n-\tau})} \quad (5.3)$$

e, assim, obtém-se a medida normalizada de transferência de entropia atrasada (Gourévitch e Eggermont, 2007):

$$\widehat{TE}(X \rightarrow Y, \beta) = \frac{TE(X \rightarrow Y, \beta) - TE(X \rightarrow Y_{surrogate}, \beta)}{H(Y_n|Y_{n-\tau})} \in [0, 1] \quad (5.4)$$

O processo completo para a obtenção das medidas normalizadas de transferência de informação é apresentado na Figura 5.7.

Figura 5.7: Procedimento para obtenção das medidas de transferência de entropia atrasada normalizada, aplicado às gravações dos neurônios. O pré-processamento com SSA é aplicado para a remoção da não-estacionariedade e, em seguida, as medidas de transferência de entropia atrasada são aplicadas, de 0 a 105 ms, entre o estímulo de 27Hz de ruído gaussiano branco (GWN) e os neurônios motores (no caso em exemplo, o FETi) e entre 10 dados de *surrogate* gerados a partir do estímulo, através do algoritmo IAAFT, e os mesmos neurônios motores. O número de surrogates é escolhido arbitrariamente, considerando-se que um número maior impactaria consideravelmente o tempo de execução. Por fim, a normalização é aplicada de acordo com a equação 5.4.

5.3.4 Análise em nível de grupo: integração Bayesiana de medidas de transferência de informação

Conforme notado na introdução, uma das vantagens do uso de medidas de teoria de informação está no fato de mapear o espaço de funções de distribuição de probabilidade para um espaço de números reais — ou seja, são medidas também chamadas de *funcionais* (Wollstadt et al., 2014)). Assim, sob a suposição de que o sistema modelado permanece inalterado em

diferentes amostras e indivíduos, tais medidas podem ser usadas para fornecer um modelo geral. Gómez-Herrero et al. (2015), por exemplo, usou esta vantagem para calcular medidas de transferência de informação em sinais considerados como cicloestacionários e, assim, vencer a limitação de não-estacionariedade. No caso de gravações intracelulares, no entanto, algumas dificuldades surgem, e métodos como de análise de grupo (Li et al., 2008), por exemplo, não são aplicáveis uma vez que buscam obter redes nas quais todos os nós presentes são gravados simultaneamente. Assim, o foco deve ser movido para redes obtidas a partir medidas de transferência de informação, e a integração dos resultados deve ser feita nesse nível. Propõe-se, assim, um método Bayesiano de integração de medidas de transferência de entropia atrasada — movendo a análise de um paradigma frequentista para um Bayesiano, no qual pode-se quantificar a certeza acerca de certos picos de máxima transferência de informação dadas várias repetições do experimento.

Seja n o número de conexões (determinadas pela transferência de informação) esperadas entre dois sinais. Este número pode ser obtido a partir de modelos de mistura de gaussianas 2D aplicado aos picos de transferência de entropia atrasada encontrados em cada experimento. Variando-se o número de gaussianas em cada modelo e utilizando uma pontuação BIC (Žak-Szatkowska e Bogdan, 2011) para avaliá-los, pode-se definir n a partir do modelo com a menor pontuação BIC.

Sejam então $\mathbf{c}_1, \dots, \mathbf{c}_n$ tuplas de variáveis aleatórias representando o atraso de tempo e o valor de transferência entropia atrasada em cada uma das n conexões entre os dois sinais. A evidência dos dados é então utilizada para atualizar probabilidades *a posteriori* para cada uma destas conexões. Os picos obtidos através das medidas nos dados são então separados como evidência para cada conexão de acordo com a gaussiana à qual eles pertencem no modelo de mistura de gaussianas utilizado anteriormente, e assim definem-se $\mathbf{p}_{i_1}, \dots, \mathbf{p}_{i_k}$ como tuplas contendo os atrasos de tempo e valor de transferência de entropia atrasado obtidos nos dados, sendo elas os k picos utilizados como evidência para a conexão i . Assim, a probabilidade *a posteriori* buscada para cada conexão é representada como $P(\mathbf{c}_i | \mathbf{p}_{i_1}, \dots, \mathbf{p}_{i_k})$.

Estas probabilidades, por sua vez, podem ser decompostas pela regra de Bayes, relacionando a verossimilhança com as probabilidades *a priori* (Gelman et al., 2014):

$$P(\mathbf{c}_i | \mathbf{p}_1, \dots, \mathbf{p}_k) = \frac{P(\mathbf{p}_{i_1}, \dots, \mathbf{p}_{i_k} | \mathbf{c}_i) \times P(\mathbf{c}_i)}{P(\mathbf{p}_{i_1}, \dots, \mathbf{p}_{i_k})} \quad (5.5)$$

e, considerando-se os experimentos independentes entre si, tem-se:

$$P(\mathbf{c}_i | \mathbf{p}_1, \dots, \mathbf{p}_k) = \frac{\prod_{j=1}^k P(\mathbf{p}_{i_j} | \mathbf{c}_i) \times P(\mathbf{c}_i)}{\prod_{j=1}^k P(\mathbf{p}_{i_j})} \quad (5.6)$$

Resta, portanto, definir como pode-se representar a verossimilhança $P(\mathbf{p}_{i_j} | \mathbf{c}_i), j=1, \dots, k$ e a probabilidade *a priori* (o denominador $P \prod_{j=1}^k P(\mathbf{p}_{i_j})$ é um fator de normalização e não precisa ser calculado uma vez que se considere que a soma de todas as probabilidades equivale a 1). A distribuição *a priori* é utilizada para codificar o conhecimento de especialista, e por razões de metodologia, será adotada como uma distribuição uniforme. As verossimilhanças $P(\mathbf{p}_{i_j} | \mathbf{c}_i), j=1, \dots, k$, por sua vez, são modeladas como distribuições normais multivariadas com média $\boldsymbol{\mu}_{i_j}$ no pico $\mathbf{p}_{i_j}$ e matriz de covariância $\boldsymbol{\Sigma} = \begin{pmatrix} 15 \text{ ms} & 0 \\ 0 & 0.15 \text{ bits/s} \end{pmatrix}$. Estes últimos valores de covariância são escolhidos de forma a garantir uma probabilidade de 90% em um pico dada uma média de 25 repetições do experimento em um caso bem estabelecido. Assim, elas determinam um modelo padrão a ser estendido e utilizado com medidas em outros neurônios.

Tabela 5.1: Gravações do sistema neuromotor proprioceptor do gafanhoto utilizadas para as medidas de informação mútua atrasada. Os canais marcados com 'X' se referem a gravações incertas sobre o sistema; os números 200, 58 e 27 são os sinais de ruído branco gaussiano que excitam o sistema, em Hertz; FCO se refere a uma gravação nos neurônios sensores do *FeCO*; Flexor se refere a um neurônio motor flexor não especificado e os outros rótulos indicam outros neurônios motores, conforme indicados na tabela 4.2

ID	Canal 1	Canal 2	Canal 3	Canal 3	Trecho (em s)
403	X	FETi	200	58	21,8453 s
404	X	FETi	200	58	43,6906 s
405	X	FETi	200	58	10,9227 s
407	X	FETi	200	58	21,8453 s
410	X	FETi	200	58	43,6906 s
433	X	FETi	200	58	10,9227 s
435	X	FETi	200	58	21,8453 s
436	X	FETi	200	58	21,8453 s
437	X	FETi	200	58	21,8453 s
438	X	FETi	200	58	21,8453 s
783	X	FCO	27	200	21,8453 s
785	X	FCO	27	200	43,6907 s
1201	Nonspiker	PIFITi	27	200	21,8453 s
1206	X	PIFITi	58	200	21,8453 s
1207	X	PIFITi	27	200	21,8453 s
1311	Nonspiker	AFIFITi	58	200	10,9227 s
1410	Nonspiker	AFFFFITi	SETi	27	5,4613 s
1413	Nonspiker	AFFFFITi	SETi	200	10,9227 s
1417	Flexor	SETi	PSFITi	200	10,9227 s
1419	Nonspiker	SETi	PIFITi	200	5,4613 s
1420	Nonspiker	SETi	PIFITi	200	10,9227 s
1422	Nonspiker	SETi	PIFITi	200	10,9227 s
1424	Nonspiker	PIFITi	PFFITi	200	10,9227 s
1425	Nonspiker	PIFITi	PFFITi	200	10,9227 s
1427	FETi	PIFITi	PFFITi	200	10,9227 s
1429	Flexor	PIFITi	PFFITi	200	10,9227 s
1430	Flexor	PIFITi	PFFITi	200	10,9227 s
1434	Nonspiker	AI ou AFFITi	PSFITi	27	2,7307 s
1445	Nonspiker	SETi	58	X	10,9227 s
1505	Nonspiker	SETi	PIFITi	X	10,9227 s
1508	Nonspiker	SETi	PIFITi	200	10,9227 s
1510	Nonspiker	SETi	PIFITi	200	21,8453 s

Capítulo 6

Resultados e discussões

6.1 Simulações

6.1.1 Aprendizagem estrutural de redes Bayesianas com nós de mesmo tipo

A Figura 6.1 apresenta os diferentes histogramas para as medidas de informação mútua de acordo com a diretividade entre os nós, e as funções de distribuição de probabilidade obtidas através de estimação de densidade *kernel* (*KDE*). Em alguns casos há uma separação clara nas medidas de acordo com a diretividade, enquanto que em outros não. Nota-se que a separação é total, nos casos de diretividade 1 e 2, quando a informação passada diretamente de um nó para outro é pequena (porém as diretividades de 3 e 4 são confundidas). Isso ocorre tanto pelo aumento da entropia base usada para o cálculo das *CPDs* quanto pelo aumento do número de *bits* de discretização (que, no presente caso, influenciam no cálculo da *CPD*). Por outro lado, quando a informação passada diretamente é alta, as diferenças entre as diretividades 2, 3 e 4 são mais perceptíveis, embora ainda com uma probabilidade de erro. A Figura 6.2 mostra o resultado da análise de teoria de decisão Bayesiana, indicando as probabilidades de sucesso e erro na classificação de valores de informação mútua em diretividade entre os nós.

Em seguida, checam-se os *IIGs* nos quais os valores de informação mútua não são tão claramente classificáveis de acordo com a diretividade. A Figura 6.3 mostra um exemplo de dois *IIGs* com diferentes diretividades porém com valores de informação mútua bastante similares. A explicação para esta similaridade encontra-se na estrutura dos grafos: em 6.3(a), os nós D e E tem diretividade de 3, entretanto, muita informação é compartilhada pelo ancestral comum A. Se, portanto, as perdas de A para D e de A para E forem pequenas, a informação mútua

Figura 6.1: Histogramas e aproximação por *KDE* dos valores de informação mútua obtidos nas simulações de acordo com a diretividade entre dois nós. r é o número de *bits* de discretização e m_p é o máximo número de pais permitidos na rede (também utilizado para o cálculo dos CPDs)

Figura 6.2: Probabilidades de sucesso e erro na classificação dos valores de informação mútua em diretividade entre dois nós em uma rede Bayesiana com *CPDs* fixas. As linhas se referem à probabilidade de sucesso em cada valor de diretividade (de 1 a 4) e as probabilidades de erro de classificação em outro valor. Já as colunas se referem a diferentes configurações de número de *bits* (r) e número máximo de pais permitidos (m_p). Cada gráfico também é feito contra as diferentes entropias base usadas para o cálculo das *CPDs*.

Figura 6.3: Exemplos de *IIGs* com diferentes valores de diretividade nos quais os valores de informação mútua podem ser bastante parecidos (dependendo da quantidade de perdas de um nó para outro). Em (a) (diretividade igual a 3), o ancestral A dos nós D e E podem transmitir uma quantidade de informação de forma que a informação mútua seja quase igual à encontrada entre os nós B e C em (b) (que têm diretividade igual a 2).

será praticamente a mesma de um caso de diretividade igual a 2, como em 6.3(b) (de B para C). Assim, como mostra a Figura 6.1, os valores de informação mútua são mais separáveis nos casos onde as perdas de informação são maiores (ou seja, menos informação mútua entre dois nós conectados diretamente). Isso apresenta um *trade-off*, pois ao passo que a diferenciação das diretvidades com valores altos (3 e 4, por exemplo), por outro lado permite uma excelente classificação nas diretvidades de valores baixos (1 e 2) — e que, de qualquer forma, já são boas informações sobre a estrutura do grafo e podem gerar boas reduções do espaço de busca em um método de aprendizagem estrutural.

Discussão

Embora os experimentos tenham sido feitos em redes estáticas e casos bastante específicos, os experimentos realizados são um primeiro passo em uma exploração sobre que tipo de conhecimento pode ser encontrado nos valores de teoria de informação uma vez que se considerem redes nas quais os nós são do mesmo tipo ou apresentem o mesmo comportamento. A grande vantagem disso encontra-se na superação do problema de gravações de sinais onde todos os nós não são observados simultaneamente. Com redes Bayesianas simuladas com *CPDs* fixas, foi mostrado que, em casos nos quais a informação passada de um nó para outro é suficientemente pequeno, a diretividade (ou seja, o tamanho do menor caminho não-direcionado) entre duas variáveis pode ser inferida e utilizada para reduzir o espaço de busca de redes.

Poderia-se perguntar, no entanto, se as simplificações adotadas refletem, de fato, os casos reais. Para esta pergunta pode-se fazer três considerações:

1. Resultados experimentais, como no trabalho de Endo et al. (2015), suportam esta linha de raciocínio, na qual pode-se reconstruir uma rede baseada apenas nos valores de medidas

de teoria de informação entre as variáveis gravadas;

2. A metodologia desenvolvida também pode ser aplicada a qualquer interpretação que se der a nós similares em uma rede — inclusive em extensões do conceito de *CPDs* fixas. Redes simuladas podem ser geradas conforme estas interpretações, e a separação nos valores de informação mútua de acordo com a diretividade (ou mesmo outras características que se deseja procurar) pode ser checada a partir da estimação de densidade *kernel* e a análise por teoria de decisão Bayesiana.
3. As extensões possíveis ao conceito de *CPDs* fixas podem ser:
  - (a) Ao invés de se introduzir ruído uniforme nas variáveis com menos pais do que os previstos na *CPDs*, diferentes *CPDs* poderiam ser propostas para cada número de pais possíveis em uma variável; ou
  - (b) Para redes nas quais se consideram classes ou tipos de nós, como na Figura 3.1(c), pode-se também atribuir um *CPD* diferente para cada classe; ou
  - (c) Se uma decomposição de informação multivariada pode ser conduzida, como as propostas em Williams e Beer (2010); Rauh et al. (2014), pode-se também fixar os componentes de informação obtidos e calcular-se as *CPDs* de acordo com a estrutura da rede que se tem em foco.

A razão de usar-se nestes experimentos uma medida de teoria de informação par-a-par como a informação mútua é por simplicidade — poderia-se também pensar em descoberta de conhecimento sobre a estrutura da rede com medidas trio-a-trio ou mesmo medidas de ordens mais altas, até porque medidas par-a-par não são capazes de capturar efeitos de sinergia nas redes. O método desenvolvido, portanto, poderia ser generalizado para não apenas encontrar-se a diretividade entre dois nós, mas até mesmo o que costuma-se chamar de *network motifs*, baseados nessas medidas de teoria de informação com ordens mais altas. Estes *network motifs*, por sua vez, podem representar reduções ainda maiores no espaço de busca, uma vez que são entendidos como "simples blocos de construção de redes complexas"(Milo et al., 2002).

Por fim, também vale notar que estes experimentos sugerem uma mudança na ordem convencional de aprendizagem de redes Bayesianas — eles primeiro buscam elicitare os parâmetros da rede (*CPDs*), e só depois disso buscam inferir a estrutura, baseado no que se espera desses parâmetros e como eles se refletem em medidas funcionais sobre os dados, como informação mútua. Tal mudança pode ser favorável em domínios como neurociência, por exemplo, no qual

Figura 6.4: Tempo médio de convergência, em segundos e escala logarítmica, dos dos algoritmos *PSO* e *greedy* para aprendizagem estrutural de *HO-DBNs* simuladas. Os dados são discretizados em 2 bins. As linhas representam n , o número de variáveis no sistema, e as colunas representam τ , as ordens de Markov. As porcentagens nas barras indicam o número aproximado de arestas na rede simulada (100% indica todas as arestas possíveis na *HO-DBN*).

um entendimento de transferência de informação a um nível local de neurônios tem sido cada vez mais buscado (Gurney, 2001; Di Maio, 2008; McDonnell et al., 2011). Acredita-se que este tipo de inferência pode ser bastante aplicada neste domínio, no qual as redes Bayesianas têm sido bastante utilizadas, porém sem considerações a partir do fato que se está lidando com redes cujos nós são iguais. A única abordagem neste sentido que se conhece até o momento é a das redes Bayesianas Orientadas a Objeto, por Koller e Pfeffer (1997), porém ainda sem considerações sobre redução do espaço de busca a partir dessas considerações, ainda mais em situações nas quais o sistema não é totalmente e simultaneamente observado.

6.1.2 Aprendizagem estrutural de HO-DBNs

O algoritmo de *PSO* convergiu para as soluções desejadas em todos os casos, e a Figura 6.4 apresenta, em escala logarítmica, o tempo médio de convergência para os algoritmos *greedy* e *PSO*, de acordo com a ordem de Markov da *HO-DBN*.

Observa-se que, exceto nos casos com poucas arestas, o algoritmo de *PSO* supera o algoritmo *greedy* em tempo de convergência, e à medida que o número de variáveis ou a ordem de Markov cresce, a diferença se torna ainda maior (lembrando que os gráficos estão em escala logarítmica). Assim, os resultados indicam que há um ganho exponencial de desempenho no algoritmo de *PSO* no casos com muitas variáveis e ordens altas. Observa-se também, com

referência às diferenças de desempenho nas configurações com nenhum, poucas, muitas e todas arestas na rede, que nos dois algoritmos o tempo de convergência cresce à medida que o número de arestas aumenta.

Discussão

Observando a comparação de desempenho entre o algoritmo de *PSO* e o *greedy*, pode-se supor que *PSO* dá uma abordagem coletiva a um problema *greedy* de apenas uma partícula percorrendo o espaço de busca — e, assim, isso explica o porquê do algoritmo convergir mais rápido para a solução. Esta abordagem coletiva, juntamente com as adaptações propostas para os conceitos de inércia e contribuições dos parâmetros cognitivo e social, que adicionam aleatoriedade ao algoritmo, providenciam uma boa exploração do espaço de busca, em pouco tempo. Além disso, uma vez considerada a simplicidade das representações propostas e a necessidade de poucos parâmetros de controle, a abordagem de *PSO* prova-se como um método adequado para aprendizagem de estruturas com muitas variáveis e ordens altas. No entanto, ainda permanece um grande problema acerca do tamanho dos conjuntos de dados requeridos para este treinamento — o que deve direcionar a pesquisa em métodos que lidam com dados faltantes.

Assim, com uma discussão sobre os tipos de adaptações de *PSO* para a aprendizagem de estruturas de *DAGs*, e as considerações sobre posições e velocidades em um espaço de busca discreto, o presente trabalho serviu para levantar ideias e princípios importantes a serem considerados na adaptação de abordagens de inteligência de enxames para problemas de otimização discretos e em espaços de grafos — e principalmente, na tarefa de aprendizagem de redes a partir de dados. O trabalho também levantou mais um exemplo da ideia apresentada no trabalho de Espezuza et al. (2014), que propõe que a introdução de aleatoriedade adicional em um método de otimização — que, no caso atual, consistiu em propor novas velocidades aleatórias que mantivessem a inércia — é uma consideração útil para os métodos de otimização discreta.

Embora as reinterpretações de movimento, velocidade e inércia propostas no trabalho sejam apresentadas como melhorias específicas para o domínio de *HO-DBNs*, estas ainda podem ser consideradas uma abordagem geral para o caso de otimização em quaisquer problemas que possam ser representados como listas de características (que, no caso presente, são as listas de causalidade). Este também é o caso do trabalho de Heng et al. (2006), que, em última análise, não se aplica só ao problema de aprendizagem de redes Bayesianas, mas também a um problema geral de otimização em um espaço de listas de características. Isso significa que embora

a solução proposta não adapta as representações para algoritmos gerais de *PSO* discreto, estabelecidos na literatura, como *PSO* Binário, ainda há uma adaptação de representações de redes para um espaço de busca de listas de causalidade, com as quais o método de *PSO* foi desenvolvido. Assim, o domínio de problema de otimização em um espaço de *DAGs*, permanece sem uma adaptação direta em *PSO*, e esta consiste em um caminho interessante para pesquisa posterior. Pode-se, por exemplo, desenvolver formas de estruturar e ordenar o espaço de busca de *DAGs*, simulando um espaço contínuo, de formas eficientes e condizentes com o problema de otimização. Neste caso, também, uma abordagem baseada em probabilidades e *likelihoods* de arestas também é desejável, como nos trabalhos de Du et al. (2005) e Sahin e Devasia (2007).

Finalmente, considera-se que embora o método conte com valores muito pequenos de *bins* para a discretização — 2 e 3 *bins* —, isso acaba sendo razoável e suficiente para muitas aplicações de modelagem de circuitos neurais, considerando que a maioria dos neurônios operam através de pulsos de potencial elétrico — também chamados *spikes*. O que acaba sendo o mais importante na modelagem da dinâmica de um circuito neural são os atrasos presentes entre os nós, que são responsáveis pelo comportamento emergente da rede. Temporalidade é o fator crucial nesses sistemas, conforme aponta Arthur e Boahen (2006) — e quanto esta passa a ser entendida, irá desencadear avanços importantes na modelagem de circuitos neurais e aplicação posterior em áreas como biorrobótica, interação humano-máquina, entre outras.

6.2 Conectividade neural entre sub-áreas do hipocampo em estado de repouso

6.2.1 Aprendizagem de DBN individual com três ROIs

A segmentação automática das sub-áreas do hipocampo foi aplicada às imagens T1 e T2-*weighted*, e a Figura 6.5 mostra o resultado para um dos indivíduos. Após o pré-processamento, as séries temporais das três ROIs (DG, CA1-4 e ERC) são extraídas da imagem de fMRI. Então, a técnica de *detrending* por SSA é aplicada — as séries temporais são decompostas em 50 componentes, e 2 componentes associados aos maiores autovalores da matriz de trajetória são removidos. Por fim, os dados são discretizados para 5 bins ($d=5$) e o algoritmo GlobalMIT é aplicado para encontrar a estrutura de DBN discreto. Após a busca no espaço de estruturas, o grafo com maior pontuação foi o que está representado na Figura 6.6, com nível de significância igual a 0.05 ($\alpha=0.95$, segundo a equação 3.8).

Figura 6.5: Resultado da segmentação automática realizada pelo software ASHS (Yushkevich et al., 2015b) nas imagens obtidas. Os rótulos para cada sub-área são providenciados em alinhamento com a imagem T2-weighted, que, no caso, possui tamanho de voxel $0.37 \times 0.37 \times 1.5\text{mm}$, mas estes são, logo depois, alinhados com as imagens funcionais para a extração das séries temporais das ROIs e aprendizagem da DBN.

Figura 6.6: Estrutura de DBN de maior pontuação com o algoritmo GlobalMIT aplicado a três nós de apenas um indivíduo. CA é o *cornu ammonis*; DG, o giro dentado e ERC, o córtex entorinal.

A estrutura de rede encontrada revela uma conexão mútua entre CA e DG e uma conexão direcionada do DG para o ERC. A mediação entre CA e ERC por meio do DG pode indicar porque o DG foi descoberto como mais conectado funcionalmente aos córtices parahipocampal e cingulado do que o CA, no trabalho de (de Flores et al., 2015) (dada, também, a consideração em Libby et al. (2012) de que as conexões entre as sub-áreas do hipocampo e os córtices parahipocampal e cingulado são mediadas pelo ERC).

6.2.2 Aprendizagem de DBN em nível de grupo

A Figura 6.7 apresenta a análise *seed-to-voxel* de conectividade funcional, com os voxels marcados sendo aqueles que passaram no teste-T com conectividade maior do que zero. Em geral, observa-se uma conectividade extensiva com o lobo medial temporal em ambos os hemisférios. No entanto, as sub-áreas do hipocampo direito tendem à conectividade exclusiva com o lobo temporal intra-medial, enquanto as sub-áreas do hipocampo esquerdo se conectam mais com áreas fora do lobo medial temporal (ou seja, as regiões cinguladas). A conectividade também varia de acordo com a sub-área: o CA esquerdo é a única subárea a apresentar conec-

Figura 6.7: Voxels da imagem de fMRI cuja correlação com osseeds (CA, DG e Sub) passaram do teste-T com conectividade significativamente maior do que zero.

Figura 6.8: Estrutura de DBN com maior pontuação no algoritmo GlobalMIT na análise com cinco ROIs e abordagem por VTS de análise em grupo. CA é o *cornu ammonis*; DG, o giro dentado; SUB, o subículo; PHG, o giro parahipocampal e AMY, a amígdala.

tividade com o córtex cingulado anterior. Já a conectividade com o DG e o Sub se sobrepõem.

Através do alinhamento destes resultados com um atlas anatômico, as ROIs obtidas para posterior análise são as apresentadas na Tabela 6.1.

A aprendizagem estrutural é então aplicada para 5 ROIs — CA, DG, SUB, amígdala (AMY) e o giro parahipocampal (PHG) do hemisfério direito — ao se considerar uma abordagem de indivíduo típico virtual (VTS). A rede com maior pontuação encontrada pelo algoritmo GlobalMIT é a apresentada na Figura 6.8. O resultado é consistente com a rede obtida anteriormente, com apenas três nós (Figura 6.6), uma vez que o DG e o CA ainda permanecem mutuamente conectados.

Tabela 6.1: ROIs obtidas a partir do rotulamento das áreas ativadas na análise *seed-to-voxel*, a serem usadas para a aprendizagem de uma DBN.

Regiões de Interesse (ROIs) (quase todas incluindo os hemisférios direito e esquerdo, com indicação no caso contrário)
Amígdala
Sulco Calcarino
Núcleo caudado (apenas esquerdo)
Cerebelo
Cíngulo Anterior
Cíngulo Médio
Cíngulo Posterior
Córtex Orbifrontal Inferior
Córtex Orbifrontal Médio
Córtex Orbifrontal Superior
Giro Fusiforme
Giro de Heschl (apenas esquerdo)
Córtex Insular
Córtex Lingual
Córtex Olfatório (apenas esquerdo)
Giro Parahipocampal
Pecúneo
Putâmen (apenas esquerdo)
Opérculo Rolândico
Giro Temporal Inferior
Giro Temporal Médio
Giro Temporal Superior
Pólo Temporal Médio
Pólo Temporal Superior
Tálamo
Vermis Cerebelar

6.3 Transferência de entropia atrasada no sistema neuromotor do gafanhoto

A metodologia descrita na seção 5.3 foi aplicada — primeiro, o pré-processamento com SSA para a remoção da não-estacionariedade; em seguida, as medidas de transferência de entropia atrasada foram aplicadas entre os sinais de estímulo (ruído branco gaussiano de 27 Hz) e a resposta do neurônio motor FETi (4.2), revelando picos de máxima transferência de informação, os quais indicam os atrasos de tempo entre os sinais. Estes atrasos e valores de transferência de entropia são então reunidos de vários experimentos envolvendo os mesmos nós e integrados através de uma análise Bayesiana, capaz de quantificar a certeza destes valores dado a evidência dos dados.

6.3.1 Pré-processamento com SSA

A remoção das tendências não-lineares e descontinuidades dos sinais do FETi através de SSA permitiu a utilização de um maior número de amostras dos sinais empregados nas medidas, uma vez que antes, como em Endo et al. (2015), pequenos trechos dos sinais, considerados estacionários em partes, eram usados. Assim, a estimação de distribuições de densidade de probabilidade a partir foi melhorada. Alguns dos sinais resultantes do pré-processamento são apresentados na Figura 6.9. No total, 216 sinais obtidos dos neurônios motores foram analisados, com tamanho variando de 17.8 s a 123.7 s. Baseado no teste Wald-Wolfowitz de estacionariedade (Grazzini, 2012), descobre-se que 36.6% dos dados eram estacionários antes do pré-processamento com SSA; já com este procedimento, 95.8% dos dados tornaram-se estacionários e usáveis para a análise de DTE. Dados instáveis foram removidos para análises posteriores.

O resultado da aplicação de medidas de DTE (do estímulo de 27 Hz para o sinal do FETi) também confirma que a não-estacionariedade presente nos sinais também se deve apenas aos artefatos experimentais, uma vez que, depois da sua remoção, os mesmos resultados são obtidos para sinais diferentes (vide Figura 6.12).

Figura 6.9: Sinais do neurônio motor FETi antes e depois do pré-processamento com a técnica de remoção de tendências não-lineares e descontinuidades através de análise de espectro singular (seção 2.3.1). A aplicação da DTE entre o estímulo e a resposta do neurônio revela picos em aproximadamente os mesmos atrasos de tempo.

6.3.2 Impacto do número de amostras nas medidas de DTE

Em seguida, o efeito do número de amostras na estimação de funções de densidade de probabilidade e no cálculo da DTE é testado.

Uma gravação pré-processada do neurônio motor é utilizada, e sua análise de DTE revela dois picos distintos na associação entre o estímulo de GWN no FeCO e a resposta sináptica (Figura 6.10). A gravação original contém 1.2M amostras (traço 12) na Figura 6.10, e tem seu tamanho variado de 100k até 1.2M em incrementos de 100k (traços 1-12). Os valores absolutos de DTE são calculados, no entanto, estes podem incluir falsos positivos e, por isso, conduz-se um teste de significância com o cálculo da DTE em dados de *surrogate*, que são indicados na Figura apenas para os traços 1, 2 e 12 (por fins de clareza). Fica claro pela Figura 6.10 que uma redução do tamanho do sinal aumenta consideravelmente o nível de ruído (compare-se, por exemplo os traços 1 e 12).

Sinais com diferentes números de amostras, decorrentes de gravações de três neurônios motores — FETi, PSFITi e AIFITi — também foram escolhidos e medidas de DTE foram aplicadas entre o estímulo de ruído gaussiano branco de 27 Hz e estes sinais. Sinais de *surrogate* também foram gerados para o teste de significância. Observa-se, assim, que, à medida que o número de amostras cresce, os valores médios de DTE nos dados de *surrogate* são menores (ou seja, o nível de ruído), dando uma significância estatística maior aos picos de máxima DTE

Figura 6.10: Efeito do número de amostras no cálculo de DTE. A análise de uma gravação do neurônio motor PSF1Ti tem seu tamanho usado de 100k (traço 1) até 1.2M amostras (traço 2) e rotulado de 1 a 2 (traços azuis). A análise da DTE revela dois picos distintos. Dados de *surrogate*, representando o nível de ruído, foram usados para o teste de significância. Nota-se que tanto o nível de significância como de ruído diminuem com o aumento de amostras, devido ao uso de 32 *bins* de discretização.

nos dados originais (Figura 6.11). Como exemplo, com os sinais de 27 Hz para AIF1Ti com 189,260 amostras, a diferença entre a DTE no primeiro pico e a DTE máxima encontrada nos surrogates é igual a 0.246 bits, enquanto que, com 343.027 amostras (ou seja, o dobro), a diferença sobe para 0.324 bits, e com 757.043 amostras (o quádruplo), a diferença sobe para 0.300 bits.

6.3.3 Consistência das medidas de DTE

Para o estabelecimento da DTE seguida de um pré-processamento de SSA como um método válido para sinais intracelulares de potencial sináptico em insetos, vários testes também são feitos verificando sua consistência em três casos: (1) consistência em um mesmo neurônio de um único animal no decorrer de apenas uma gravação; (2) consistência entre múltiplas gravações de um mesmo neurônio do mesmo animal e (3) consistência entre múltiplas gravações do mesmo neurônio em animais diferentes.

Figura 6.11: Efeito do número de amostras nas medidas de DTE. Cada linha corresponde ao mesmo sinal com diferentes números de amostras utilizados. A primeira apresenta a DTE do estímulo de 27 Hz com o FETi, a segunda, com o PSFITi, e a terceira, com o AIFITi. À medida que o número de amostras aumenta, os valores médios de DTE nos dados de surrogate diminuem, e portanto o nível de significância dos picos encontrados nos dados originais é maior. O uso do pré-processamento com SSA permite este aumento, uma vez que, ao remover tendências não-lineares e descontinuidades do sinal, disponibiliza um maior número de amostras.

Figura 6.12: Comparação das medidas de DTE (sem normalização) em três trechos uniformes de um mesmo sinal gravado no neurônio motor FETi, de aproximadamente 16,6 s, mostrando picos consistentes com média $13.60 \pm 0.41\text{ms}$ para o primeiro pico e $31.97 \pm 1.13\text{ms}$ para o segundo. Comprova-se assim que, após o pré-processamento com SSA, o sinal resultante pode ser considerado estacionário.

No decorrer de apenas uma gravação

Sinais com longa duração são divididos em segmentos de três partes e medidas de DTE foram aplicadas a cada um deles separadamente, em relação ao estímulo de 27 Hz. Como exemplo, a Figura 6.12 mostra o procedimento realizado em uma gravação do neurônio motor FETi, com aproximadamente 1.2M amostras (50 s). A comparação dos resultados mostra pequenas diferenças entre os picos de máxima DTE encontrados, com uma média de $13.60 \pm 0.41\text{ms}$ (média $\pm$ desvio padrão). para o primeiro pico e uma média de $31.97 \pm 1.13\text{ms}$ para o segundo. A porcentagem de picos com desvio padrão normalizado menor que 5% é de 88%, enquanto que a 10% é igual a 98%. Com isso nota-se a eficácia do método de *detrending* por SSA (além de se provar que as gravações utilizadas dizem respeito a apenas um neurônio).

No mesmo animal e entre diferentes animais

Para a determinação da consistência das medidas em repetições do mesmo experimento, realizados em intervalos de no máximo 5 minutos, no mesmo animal, utilizam-se gravações de neurônio motor FETi enquanto o apódema é estimulado pelo sinal GWN de 27 Hz. A Figura 6.13 apresenta resultados de três animais diferentes, cada um contendo, respectivamente, 6, 6 e 4 repetições do experimento, e também revela pouca variação nos picos de atraso.

Finalmente, ao se comparar gravações entre diferentes animais, uma maior variabilidade é

Figura 6.13: Medidas normalizadas de DTE do estímulo de 27 Hz para o neurônio motor FETi, aplicado em diferentes gravações nos mesmos animais. Cada gráfico contém resultados encontrados para cada um dos 3 animais empregados, com o primeiro contendo 6 gravações, o segundo, 6, e o terceiro, 4.

Figura 6.14: Medidas normalizadas de DTE do estímulo de 27 Hz para o neurônio motor FETi em todas as 23 gravações disponíveis. O atraso médio encontrado no primeiro pico é de $16.21 \pm 1.36 ms$ e, no segundo, 38.08 ± 2.64 .

encontrada, porém ainda suficientemente pequena — o atraso médio no primeiro pico é $16.21 \pm 1.36 ms$ e, no segundo, 38.08 ± 2.64 (Figura 6.14). Pode-se observar, assim, que o perfil de transferência de informação entre o estímulo e a resposta do FETi é preservado entre diferentes gravações e animais utilizados nas mesmos.

6.3.4 Integração Bayesianas das medidas

Os atrasos médios encontrados, no entanto, não indicam o nível de certeza que se tem nelas, dadas as evidências disponíveis nas gravações. Portanto, a análise Bayesiana descrita na seção 5.3.4 é aplicada aos valores de DTE e atraso de tempo encontrados em cada gravação, de forma a quantificar esta certeza. Para a obtenção do número de picos, a menor pontuação

Figura 6.15: Resultado da análise Bayesiana sobre as medidas normalizadas de transferência de entropia atrasada entre todos os experimentos entre o estímulo GWN de 27 Hz e a resposta no neurônio motor FETi . Os gráficos apresentam curvas de nível das probabilidades *a posteriori* de cada um dos dois picos (obtidos através da pontuação BIC), sendo atualizados com a inclusão de novas evidências e proporcionando uma maior densidade de probabilidade em pontos específicos.

BIC para o modelo 2D de mistura de gaussianas é encontrada para $n=2$ e, assim, assume-se que dois picos de máxima transferência de informação estão presentes do estímulo para a resposta do FETi. Em seguida, o modelo descrito na equação 5.6 é utilizado para atualizar as probabilidades *a posteriori* destes dois picos. A Figura 6.15 apresenta gráficos em curva de nível destas probabilidades sendo atualizadas com os resultados encontrados em cada experimento. No primeiro gráfico, com 5 experimentos sendo considerados, ao se integrar uma elipse de raio $(10ms, 0.05)$ (em cada eixo, respectivamente, com o primeiro representando o atraso de tempo e o segundo a DTE normalizada) em torno da máxima probabilidade *a posteriori*, obtém-se uma probabilidade de 49.03% em volta de um centro em $(16.91ms, 0.132)$ para o primeiro pico, e uma probabilidade de 47.79% em volta de um centro em $(39.46ms, 0.011)$ para o segundo (observa-se que, ao se integrar uma elipse maior, a probabilidade aumenta, mas a precisão no valor esperado no pico é reduzida). Com 14 experimentos, no segundo gráfico, a mesma probabilidade sobe para 78.27% em volta de um centro em $(16.21ms, 0.127)$ para o primeiro pico e para 76.35% em torno de um centro em $(38.76, 0.019)$ para o segundo. Finalmente, com todos os experimentos utilizados como evidência (23 no total), a probabilidade sobe para 91.70% em torno de um centro em $(16.21ms, 0.119)$ para o primeiro pico e para 90.25% em torno de um centro em $(38.76, 0.020)$ para o segundo.

6.3.5 Discussão

Ao se aplicar a análise de espectro singular (SSA) para a remoção da não-estacionariedade nos dados, o número de amostras disponíveis para as medidas de DTE aumentou consideravelmente (de 36.6% para 95.8% dos dados), se comparado aos resultados em Endo et al. (2015), no qual foram considerados apenas pequenos trechos estacionários dos sinais. Tal resultado é importante para este tipo de aplicação — Hlaváčková-Schindler et al. (2007) e Lee et al. (2012), notam que a estimação da TE em sinais como estes é complicada devido aos sinais cujos tamanhos são limitados devido às restrições experimentais e/ou não-estacionariedade. Com uma maior disponibilidade de amostras, portanto, as medidas de DTE também puderam ser aplicadas. Os testes realizados em diferentes trechos do sinal depois do pré-processamento confirmaram que esta não-estacionariedade se deve apenas a fatores externos impostos no sinal pelo *setup* experimental, e não refletem características que deseja-se incluir em um modelo. O uso de SSA ainda contou com a vantagem de eliminar tendências não-lineares e descontinuidades sem a necessidade de se especificar diferentes parâmetros para cada gravação (o que seria o caso com filtros passa-altas, por exemplo).

Uma comparação realizada entre as medidas de DMI (empregadas em Endo et al. (2015)) e DTE também revelou as diferenças significativas entre as duas. Isto é esperado, uma vez que a DTE condiciona a entropia entre dois processos aos estados passados e, assim, captura a sinergia entre os estados presentes e passados, encontrando, assim, valores maiores de transferência de informação (Marinazzo et al., 2014, , vide também a seção 3.1). Além disso, Wibral et al. (2014a) argumenta que o condicionamento ao passado também é necessário para se diferenciar a transferência de informação do armazenamento de informação — ou seja, a exclusão do passado capturaria a memória do sistema, confundindo-a com uma transferência. Tal situação é o caso da Figura 6.16, que apresenta ambas as medidas realizadas nos mesmos sinais — entre o estímulo de 27 Hz e o neurônio motor PFFITi —, e onde pode-se observar que a informação mútua não apresenta picos definidos e possui valores significativos (acima do nível de ruído dos dados de *surrogate*) em grandes atrasos de tempo, onde pode ser compreendido como um efeito de memória. Os valores de transferência de entropia, por sua vez, apresentam um pico mais definido, o que pode resultar do fato de que a medida considera o efeito de sinergia entre o valor atrasado do sinal de estímulo (27 Hz) e o passado do sinal de resposta (PFFITi). Não obstante, outra vantagem da DTE sobre a DMI encontra-se no fato de que a DTE é uma medida assimétrica (Kaiser e Schreiber, 2002) e, assim, pode representar melhor a direcionalidade da

transferência de informação.

Figura 6.16: Diferença entre medidas de informação mútua atrasada e transferência de entropia atrasada aplicadas entre os mesmos sinais — do estímulo GWN de 27 Hz para o neurônio motor PFFITi. A transferência de entropia atrasada revela picos mais definidos de transferência de informação — algo que pode resultar da sinergia entre o valor atrasado estímulo e o passado do sinal de resposta —, ao passo que a informação mútua atrasada não, além de apresentar um possível efeito de memória, dados os valores significativos que persistem à medida que aumenta o valor de atraso de tempo.

Por fim, a análise Bayesiana também se mostrou importante para quantificar a certeza acerca dos atrasos de tempo encontrados nas diferentes gravações, dada um conhecimento *a priori*. Isso permite um processo de aprendizagem *offline*, que pode conduzir novos experimentos específicos, cujos resultados podem ser incorporados para aumentar a certeza — algo que é de muita importância no caso do tipo de experimento com gravações intracelulares, que são difíceis e caras de se obter.

Capítulo 7

Publicações e Submissões

As publicações e submissões a seguir estão relacionadas ao presente trabalho, sendo desenvolvidas no decorrer do mesmo.

- **Trabalho Publicado em Periódico:** ENDO, W.; SANTOS, F. P.; SIMPSON, D.; MACIEL, C. D.; NEWLAND, P. L. (2015). Delayed mutual information infers patterns of synaptic connectivity in a proprioceptive neural network. *Journal of Computational Neuroscience*, 38(2):427–438 (sobre o uso de medidas de informação mútua atrasada e *clustering* dos resultados para a obtenção de conectividade neural no sistema neuromotor do gafanhoto);
- **Trabalho Publicado e Apresentado em Congresso:** (conferência IEEE Biosignals & Biorobotics (ISSNIP), ocorrida em 26 a 28 de Maio de 2014 em Salvador-BA): SANTOS, F. P.; MACIEL, C. D. (2014). A PSO approach for learning transition structures of higher-order dynamic bayesian networks. *Em: Biosignals and Biorobotics Conference (2014): Biosignals and Robotics for Better and Safer Living (BRC), 5th ISSNIP-IEEE*, pp. 1–6, (sobre o algoritmo baseado em PSO para a busca de soluções no espaço de HO-DBNs);
- **Trabalho Publicado em Congresso:** SANTOS, F. P.; SMAGULA, S. S.; KARIM, H.; SANTINI, T. S.; AIZENSTEIN, H. J.; IBRAHIM, T. S.; MACIEL, C. D. (2017). Dynamic Bayesian Network Modeling of Hippocampal Subfields Connectivity with 7T fMRI: A Case Study. *Em: 10th International Conference on Bio-inspired Systems and Signal Processing (Biosignals 2017)*, no prelo (aplicação da pontuação GlobalMIT para a

aprendizagem de uma DBN descrevendo a conectividade funcional entre sub-áreas do hipocampo);

- **Trabalho Publicado em Congresso:** LIMA, D. R.; SANTOS, F. P.; MACIEL, C. D. (2016) Network Structural Reconstruction based on Delayed Transfer Entropy and Sythetic Data. *Em: XXI Congresso Brasileiro de Automática (CBA2016)* (uso de medidas de transferência de entropia atrasada (DTE) para obtenção de estruturas de redes simuladas);
- **Trabalho Submetido a Periódico:** SANTOS, F. P.; ENDO, W.; MACIEL, C. D. (2016). Using pairwise information-theoretic measures to find structures of Bayesian networks with nodes of the same type. *Engineering Applications of Artificial Intelligence* (sobre aprendizagem estrutural de redes Bayesianas com nós de mesmo tipo);
- **Trabalho Submetido a Periódico:** SANTOS, F. P.; MACIEL, C. D.; NEWLAND, P. L. (2017). Pre-processing and Transfer Entropy Measures in Motor Neurons Controlling Limb Movements. *Journal of Computational Neuroscience* (sobre o pré-processamento dos sinais intracelulares com SSA e medidas de transferência de informação entre sinais de excitação e de resposta nos neurônios motores do gafanhoto).

Capítulo 8

Conclusão e etapas futuras

Neste estudo, buscou-se explorar algumas das diversas formas nas quais os métodos de teoria de informação tem sido empregados juntamente com modelos de DBN, apresentando-se duas áreas de aplicação — a saber, análise de conectividade funcional entre sub-áreas do hipocampo e conectividade no sistema neuromotor do gafanhoto — nas quais já se têm notado que uma aproximação pode ser de grande valor para a análise de conectividade em sistemas neurais e utilização de modelos em engenharia bioinspirada.

Entre os pontos nos quais as técnicas de teoria de informação convergem com DBNs, o *insight* principal mostra que o uso de medidas de MI condicional (e TE condicional) em sistemas dinâmicos com atrasos, como no trabalho de Runge (2014) e Lizier et al. (2013), corresponde à busca de um modelo HO-DBN através de métodos de independência condicional; com a diferença da vantagem das técnicas de teoria de informação consistirem em medidas funcionais, que podem assumir um mesmo sistema sendo observado em diversas repetições de experimento e configurações de observação parcial do sistema e, assim, utilizarem vários dados para a aprendizagem (isso, por sua vez, também exige supor, ou mesmo confirmar, que as medidas permanecem consistentes em diferentes experimentos e indivíduos sendo observados, conforme a análise na seção 6.3.3). Com isso, faz sentido pensar, conforme Vinh et al. (2011b) já sugeriu, na futura proposta de um método de pontuação de DBNs utilizando a TE ou TE condicional. Afinal, deve-se também notar que o termo de penalização da pontuação Global-MIT (seção 3.2) consistindo em um tipo de correção de *bias*, já usado na área de medidas de transferência de entropia com *surrogate data*. A percepção desta convergência entre DBNs e medidas baseadas em teoria de informação deve permitir mais *insights* na proposta de soluções em pesquisas futuras.

Quando se trata da aplicação destas técnicas em dados obtidos de sistemas neurais, vários desafios surgiram, entre os quais um dos principais é o da presença de *atrasos contínuos e que diferem, entre os diferentes elementos do sistema* (conforme apresentado na seção 3.4). Os atrasos são obtidos com medidas de DMI e DTE, e exigem técnicas como de discretização no tempo por *delay reconstruction* e suposições de markovianeidade no sistema, o que também aumenta a dimensão das funções de densidade de probabilidade das DBNs (também requeridas para medidas de teoria de informação), e, conseqüentemente, a exigência de amostras para as medidas e aprendizagem de DBNs. No caso de diferentes atrasos de tempo entre os elementos, também torna-se necessária a extensão para HO-DBNs, estudadas durante o trabalho com a proposição de um método de busca de estruturas de rede baseado em PSO. Algumas propostas de solução para a "maldição da dimensionalidade" também foram apresentadas no trabalho, junto com a proposta e estudo do uso de medidas de teoria de informação par a par para a inferência do quão diretamente um nó pode estar conectado a outro em uma rede, dada a suposição de que os elementos desta rede se comportam de forma igual, o que foi exemplificado, no trabalho, por redes Bayesianas com CPDs fixos. Por fim, a técnica de *detrending* por SSA apresentada, e utilizada nas aplicações, também mostrou-se uma ferramenta útil para quando os sistemas apresentam mecanismos de não-estacionariedade exógenas ao sistema (quase sempre associadas a ruído ou limitações no *setup* experimental), aumentando o número de amostras disponíveis para a estimação das probabilidades a partir da suposição de ergodicidade do sinal.

Em última análise, no entanto, a avaliação das suposições de markovianeidade, bem como da discretização no tempo e da disponibilidade de amostras para obtenção do modelo, exige considerações específicas para cada área de aplicação. Considerações sobre o que incluir no modelo também são importantes, principalmente na definição dos nós da rede, que podem indicar diferentes elementos do sistema e/ou regiões de interesse em diversos níveis de abstração (desde elementos individuais como neurônios até sistemas neurais inteiros). Disto, afinal, depende a própria utilidade do modelo em explicar algum comportamento ou sistema assim, prover conhecimento científico válido na área de aplicação. Por isso, também é importante se ter uma ideia de onde se quer chegar antes da obtenção dos dados sua análise (pré-processamento e modelagem).

Tendo isso em vista, o trabalho explorou duas destas áreas de aplicação, apresentando uma revisão, metodologia, resultados e discussões sobre o uso de DBNs e teoria de informação em seus contextos. Os parágrafos seguintes prosseguirão a algumas notas finais sobre os resultados

encontrados e sugestões de etapas futuras.

Conectividade neural entre sub-áreas do hipocampo em estado de repouso

Os avanços em MRI de campo magnético ultra-alto têm aberto espaço para investigações mais profundas do sistema nervoso central humano, conforme se observa na presente aplicação em análise de conectividade entre as sub-áreas do hipocampo utilizando-se imagens de fMRI (bem como imagens estruturais do tipo T1 e T2-*weighted*) com resoluções nunca antes alcançadas. Não obstante, as técnicas de segmentação automática destes subcampos também têm contribuído para aumentar a quantidade de dados disponíveis para análise, devido à facilidade e precisão de seus resultados.

O presente trabalho procurou demonstrar a viabilidade dos estudos de conectividade funcional nas sub-áreas do hipocampo apresentando as DBNs como uma ferramenta útil para este propósito. Estudos anteriores exploram a conectividade entre estas sub-áreas a partir de modelos e medidas estáticas como correlação estatística ou redes Bayesianas estáticas (Libby et al., 2012; de Flores et al., 2015), que não capturam a dinâmica temporal das séries (e, no entanto, já chegam a resultados importantes). Os modelos obtidos neste trabalho, no entanto, são capazes de identificar a direção das conexões e diferenciar conexões diretas das indiretas, sem a introdução de uma suposição de linearidade no sistema — todas estas consistindo em vantagens do uso de DBNs. Trabalhos futuros buscarão melhorar e mesmo interpretar os resultados obtidos com as DBNs. Como exemplo, uma investigação das conformidades e variações entre DBNs obtidas em indivíduos (abordagens IS e CS em análises em nível de grupo, vide seção 2.2.4) pode oferecer resultados importantes para estudos funcionais e clínicos (Bielza e Larrañaga, 2014) sobre o hipocampo e o lobo medial temporal, como a identificação de estágios iniciais da progressão da doença de Alzheimer (Das et al., 2013).

As limitações na quantidade de dados são um fator crítico nos estudos de conectividade funcional em fMRI com DBNs e teoria de informação, que possuem uma resolução temporal muito baixa, se comparada a outros métodos de imagem. Isso explica porque uma abordagem com medidas de transferência de informação (como a TE) não foi explorada durante o trabalho. Não obstante, a discussão em curso acerca da validade de medidas baseadas em causalidade de Granger, como a TE (apresentada na seção 4.1.2), em dados de fMRI exige uma cautela adicional quanto à utilização destes métodos, conforme notada pelo próprio trabalho de Lizier et al. (2011). Seth et al. (2013) mostrou que a GC é robusta a uma grande variedade de mu-

danças na resposta hemodinâmica, no entanto, alerta que ela está sujeita a erros na presença de muito ruído ou de *downsampling* — que liga-se aos casos também já discutidos neste trabalho: escolheu-se não aplicar os métodos de estimação da TE desenvolvidos por Wibral et al. (2014a) devido à possibilidade de desvios e tendências na reconstrução dos estados causados por ruído colorido, além do fato de que a baixa resolução temporal pode dificultar a obtenção do *time lag* τ e mesmo os atrasos de tempo entre as amostras.

Uma das formas recentemente propostas para superar as restrições de dados limitados para a aprendizagem de redes é o uso de técnicas de aprendizagem multi-modais; ou seja, que utilizam dados de diversos tipos e fontes para a redução do espaço de busca de modelos. Schulz et al. (2004), por exemplo, realiza uma análise de conectividade com imagens tanto de fMRI como de magnetoencefalograma (MEG); Rykhlevskaia et al. (2008) apresenta uma revisão dos métodos integrando fMRI com ressonância magnética estrutural, e Zhu et al. (2014) uma revisão dos métodos integrando fMRI com imagens de difusão de tensor (*diffusion tensor imaging* ou DTI). No caso de redes Bayesianas, Iyer et al. (2013), por exemplo, utiliza o DTI como uma matriz de base para reduzir o espaço de busca de redes Bayesianas. *Insights* deste tipo podem ser incorporados em pesquisas futuras, uma vez que se disponha das imagens necessárias.

Transferência de entropia atrasada no sistema neuromotor do gafanhoto

A metodologia empregada nos sinais neurais do gafanhoto explorou o uso da DTE normalizada entre um sinal GWN de 27 Hz excitando o FeCO e suas respostas sinápticas nos neurônios motores. Além disso, um método Bayesiano é utilizado para integrar os atrasos obtidos em diversas gravações e repetições de experimentos e, assim, generalizar o conhecimento sobre transferência de informação (e, consequentemente, dos caminhos neurais e da conectividade efetiva) no sistema neuromotor proprioceptor do gafanhoto.

Os resultados obtidos indicam direções para futuros experimentos e análises. O trabalho de Maciel et al. (2012), por exemplo, sugere que os atrasos encontrados entre os neurônios motores na tíbia podem fornecer *insights* importantes sobre a forma como posição, velocidade e aceleração são conduzidas a músculos rápidos (PFFITi, AFFITi, FETi), intermediários (PIFITi, AIFITi) e lentos (PSFITi, ASFITi, SETi). Não obstante, nota-se o trabalho de Endo et al. (2015), que encontrou três atrasos de tempo distintos a partir da DMI entre a excitação de 27 Hz no FeCO e as respostas nos interneurônios, e argumentou que os atrasos menores muito provavelmente se devam a conexões excitatórias monossinápticas aferentes entre os dois (Bur-

rows, 1988). Os resultados obtidos com a DTE também mostram atrasos similares a tempos de propagação conhecidos para interneurônios *spiking* e *nonspiking* e neurônios motores da tibia (Burrows, 1987, 1988), sugerindo, assim, que a maior entrada excitatória no FETi ocorre devido a uma entrada monossináptica. Com isso, a análise sugere que ainda há conectividade funcional a ser revelada por métodos fisiológicos, morfológicos e de análise de dados.

Observa-se, porém, que a metodologia desenvolvida limitou-se apenas a medidas par a par, e a razão disso encontra-se no fato de que, à medida que as gravações disponíveis foram sendo analisadas, não se encontrou gravações suficientes com alguns dos nós desejados no modelo de rede gravados simultaneamente; ou seja, não há um recobrimento suficiente da rede com gravações contendo três ou mais canais. As gravações limitam-se apenas a quatro canais, com um ou dois dele consistindo apenas no sinal de excitação e, não obstante, muitos dos canais gravados consistem em interneurônios, os quais, dada a divisão por *clustering* proposta por Endo et al. (2015), resultam em poucas gravações disponíveis e dificilmente gravadas simultaneamente — ou seja, a não identificabilidade destes nós no momento da gravação torna bastante difícil a disponibilidade destes nós serem gravados com os outros elementos da rede como os neurônios motores ou sensores. Obviamente, os atrasos encontrados nas medidas par a par não correspondem necessariamente a conexões diretas entre, porém já ajudam a propor um modelo inicial, reduzindo o espaço de busca de redes, conforme indicado nas seções 3.3 e 3.4. O posterior refinamento exigiria, portanto, medidas de transferência de informação condicionais, conforme os trabalhos de Lizier et al. (2013), Runge et al. (2012) e Marinazzo et al. (2014). Além disso, para uma futura proposta de um modelo de DBN para este sistema, os atrasos devem ser bem conhecidos para passar por sua discretização no tempo e serem integrados em uma estrutura de HO-DBN. Tal modelo consiste em um etapa posterior para o trabalho, notando-se também que a aprendizagem de parâmetros do mesmo pode contar com aplicações de engenharia bioinspirada ou mesmo interfaces humano-máquina.

Dadas as conclusões e futuras etapas apresentadas acima, nota-se, por fim, que a obtenção de modelos de conectividade funcional, embora não representando o conceito de causalidade (Valdes-Sosa et al., 2011, , vide também a Introdução), pode ser interpretada, conforme nota Williams e Beer (2011), como a obtenção de um sistema com máxima *controlabilidade*: medidas de transferência de informação, podem ser entendidas como a capacidade de levar o sistema de um estado X para um estado X' através de um processo C — uma medida como a

TE, por exemplo, representa exatamente isso: $I(\mathbf{X}', C|\mathbf{X})$. Uma DBN ou HO-DBN, portanto, na medida que codifica estas medidas de transferência de informação representaria o sistema de maior controlabilidade encontrado para os dados. Chega-se novamente, então, à ideia de Vico citada na introdução: *verum et factum convertuntur*, ou também, numa inversão da máxima Baconiana: "poder é conhecimento".

Bibliografia

- Adnan, A., Barnett, A., Moayedi, M., McCormick, C., Cohn, M., e McAndrews, M. P. (2016). Distinct hippocampal functional networks revealed by tractography-based parcellation. *Brain Structure and Function*, 221(6):2999–3012.
- Al-Hussein, A. e Haldar, A. (2016). Unscented kalman filter with unknown input and weighted global iteration for health assessment of large structural systems. *Structural Control and Health Monitoring*, 23(1):156–175.
- Angarita-Jaimes, N., Dewhirst, O. P., Simpson, D. M., Kondoh, Y., Allen, R., e Newland, P. L. (2012). The dynamics of analogue signalling in local networks controlling limb movement. *European Journal of Neuroscience*, 36(9):3269–3282.
- Arthur, J. V. e Boahen, K. (2006). Learning in silicon: Timing is everything. *Advances in neural information processing systems*, 18:75.
- Baccalá, L. A. e Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biological cybernetics*, 84(6):463–474.
- Barnett, L. e Seth, A. K. (2011). Behaviour of granger causality under filtering: theoretical invariance and practical application. *Journal of neuroscience methods*, 201(2):404–419.
- Barnett, L. e Seth, A. K. (2017). Detectability of granger causality for subsampled continuous-time neurophysiological processes. *Journal of Neuroscience Methods*, 275:93–121.
- Bartsch, T. (2012). *The clinical neurobiology of the hippocampus: An integrative view*, vol. 151. Oxford University Press.
- Benda, J., Longtin, A., e Maler, L. (2005). Spike-frequency adaptation separates transient communication signals from background oscillations. *Journal of Neuroscience*, 25(9):2312–2321.

- Bernhardt, B. C., Hong, S.-J., Bernasconi, A., e Bernasconi, N. (2015). Magnetic resonance imaging pattern learning in temporal lobe epilepsy: classification and prognostics. *Annals of neurology*, 77(3):436–446.
- Bertschinger, N., Rauh, J., Olbrich, E., e Jost, J. (2013). Shared information—new insights and problems in decomposing information in complex systems. In *Proceedings of the European Conference on Complex Systems 2012*, pp. 251–269. Springer.
- Bertschinger, N., Rauh, J., Olbrich, E., Jost, J., e Ay, N. (2014). Quantifying unique information. *Entropy*, 16(4):2161–2183.
- Bettencourt, L. M., Gintautas, V., e Ham, M. I. (2008). Identification of functional information subgraphs in complex networks. *Physical review letters*, 100(23):238701.
- Bielza, C. e Larrañaga, P. (2014). Bayesian networks in neuroscience: a survey. *Frontiers in computational neuroscience*, 8:131.
- Blessing, E. M., Beissner, F., Schumann, A., Brünner, F., e Bär, K.-J. (2016). A data-driven approach to mapping cortical and subcortical intrinsic functional connectivity along the longitudinal hippocampal axis. *Human brain mapping*, 37(2):462–476.
- Brown, E. N., Kass, R. E., e Mitra, P. P. (2004). Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature neuroscience*, 7(5):456–461.
- Burge, J., Lane, T., Link, H., Qiu, S., e Clark, V. P. (2009). Discrete dynamic bayesian network analysis of fmri data. *Human brain mapping*, 30(1):122–137.
- Burrows, M. (1987). Parallel processing of proprioceptive signals by spiking local interneurons and motor neurons in the locust. *Journal of Neuroscience*, 7:1064–1080.
- Burrows, M. (1988). Responses of spiking local interneurons in the locust to proprioceptive signals from the femoral chordotonal organ. *Journal of Comparative Physiology A*, 164(2):207–217.
- Burrows, M. (1996). *The neurobiology of an insect brain*. Oxford University Press.
- Buschmann, T., Ewald, A., von Twickel, A., e Büschges, A. (2015). Controlling legs for locomotion — insights from robotics and neurobiology. *Bioinspiration & biomimetics*, 10(4):041001.

- Campos, L. M. d. (2006). A scoring function for learning bayesian networks based on mutual information and conditional independence tests. *Journal of Machine Learning Research*, 7(Oct):2149–2187.
- Candy, J. V. (2011). *Bayesian signal processing: Classical, modern and particle filtering methods*, vol. 54. John Wiley & Sons.
- Cannan, J. e Hu, H. (2011). Human-machine interaction (hmi): A survey. *University of Essex*.
- Chai, B., Walther, D., Beck, D., e Fei-Fei, L. (2009). Exploring functional connectivities of the human brain using multivariate information analysis. In *Advances in neural information processing systems*, pp. 270–278.
- Chen, K. e Yu, J. (2014). Short-term wind speed prediction using an unscented kalman filter based state-space support vector regression approach. *Applied Energy*, 113:690–705.
- Chen, R., Resnick, S. M., Davatzikos, C., e Herskovits, E. H. (2012). Dynamic bayesian network modeling for longitudinal brain morphometry. *NeuroImage*, 59(3):2330 – 2338.
- Chien, C.-F., Chen, S.-L., e Lin, Y.-S. (2002). Using bayesian network for fault location on distribution feeder. *Power Delivery, IEEE Transactions on*, 17(3):785–793.
- Clouser, R. (2008). A blue print for a non-reductionist theory of reality. *Transdisciplinarity in Science and Religion*, pp. 27–58.
- Cook, D. L., Schwindt, P. C., Grande, L. A., e Spain, W. J. (2003). Synaptic depression in the localization of sound. *Nature*, 421(6918):66–70.
- Cooper, G. F. e Herskovits, E. (1992). A bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9:309–347.
- Coveney, P. V., Dougherty, E. R., e Highfield, R. R. (2016). Big data need big theory too. *Phil. Trans. R. Soc. A*, 374(2080):20160153.
- Cover, T. M. e Thomas, J. A. (2006). *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience.
- Csernansky, J. G., Joshi, S., Wang, L., Haller, J. W., Gado, M., Miller, J. P., Grenander, U., e Miller, M. I. (1998). Hippocampal morphometry in schizophrenia by high dimensional brain mapping. *Proceedings of the National Academy of Sciences*, 95(19):11406–11411.

- Daly, R., Edwards, K., e O'Neill, J. (2009). Using higher-order dynamic Bayesian networks to model periodic data from the circadian clock of *Arabidopsis thaliana*. *Pattern Recognition in ...*, pp. 67–78.
- Daly, R. e Shen, Q. (2009). Learning bayesian network equivalence classes with ant colony optimization. *Journal of Artificial Intelligence Research*, 35(1):391.
- Das, S. R., Pluta, J., Mancuso, L., Kliot, D., Orozco, S., Dickerson, B. C., Yushkevich, P. A., e Wolk, D. A. (2013). Increased functional connectivity within medial temporal lobe in mild cognitive impairment. *Hippocampus*, 23(1):1–6.
- de Flores, R., Mutlu, J., Bejanin, A., Tomadesso, C., Landeau, B., Mézenge, F., de La Sayette, V., Eustache, F., e Chételat, G. (2015). Intrinsic connectivity of hippocampal subfields in normal elderly and its disturbance in mci patients. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 11(7):P61.
- den Heijer, T., van der Lijn, F., Vernooij, M. W., de Groot, M., Koudstaal, P. J., van der Lugt, A., Krestin, G. P., Hofman, A., Niessen, W. J., e Breteler, M. M. (2012). Structural and diffusion mri measures of the hippocampus and memory performance. *Neuroimage*, 63(4):1782–1789.
- Dewhirst, O. P., Angarita-Jaimes, N., Simpson, D. M., Allen, R., e Newland, P. L. (2013). A system identification analysis of neural adaptation dynamics and nonlinear responses in the local reflex control of locust hind limbs. *Journal of computational neuroscience*, 34(1):39–58.
- Di Maio, V. (2008). Regulation of information passing by synaptic transmission: a short review. *Brain research*, 1225:26–38.
- Dill, V., Franco, A. R., e Pinho, M. S. (2015). Automated methods for hippocampus segmentation: the evolution and a review of the state of the art. *Neuroinformatics*, 13(2):133–150.
- Dimitrov, A. G., Lazar, A. A., e Victor, J. D. (2011). Information theory in neuroscience. *Journal of computational neuroscience*, 30(1):1–5.
- Dolan, K. T. e Spano, M. L. (2001). Surrogate for nonlinear time series analysis. *Physical Review E*, 64(4):046128.
- Du, T., Zhang, S., e Wang, Z. (2005). Efficient learning bayesian networks using pso. In *International Conference on Computational and Information Science*, pp. 151–156. Springer.

- Duda, R., Hart, P., e Stork, D. (2001). *Pattern classification*. Pattern Classification and Scene Analysis: Pattern Classification. Wiley.
- Dupuy, J. e DeBevoise, M. (2009). *On the Origins of Cognitive Science: The Mechanization of the Mind*. A Bradford Book. Mit Press.
- Duyn, J. H. (2012). The future of ultra-high field mri and fmri for study of the human brain. *Neuroimage*, 62(2):1241–1248.
- Endo, W. (2014). *Modelagem de Circuitos Neurais do Sistema Neuromotor de Insetos com o Uso de Transferência de Informação*. Tese de Doutorado, Escola de Engenharia de São Carlos.
- Endo, W., Maciel, C. D., Simpson, D. M., e Newland, P. L. (2013). Multiple pathways network description of motor neurons in locusts: An information sensitivity analysis. In *Biosignals and Biorobotics Conference (BRC), 2013 ISSNIP*, pp. 1–6. IEEE.
- Endo, W., Santos, F. P., Simpson, D., Maciel, C. D., e Newland, P. L. (2015). Delayed mutual information infers patterns of synaptic connectivity in a proprioceptive neural network. *Journal of computational neuroscience*, 38(2):427–438.
- Espezua, S., Villanueva, E., e Maciel, C. D. (2014). Towards an efficient genetic algorithm optimizer for sequential projection pursuit. *Neurocomputing*, 123:40–48.
- Faes, L., Nollo, G., e Porta, A. (2011). Information-based detection of nonlinear granger causality in multivariate processes via a nonuniform embedding technique. *Physical Review E*, 83(5):051112.
- Feinberg, D. A. e Setsompop, K. (2013). Ultra-fast mri of the human brain with simultaneous multi-slice imaging. *Journal of magnetic resonance*, 229:90–100.
- Florin, E., Gross, J., Pfeifer, J., Fink, G. R., e Timmermann, L. (2010). The effect of filtering on granger causality based multivariate causality measures. *Neuroimage*, 50(2):577–588.
- Fortin, F.-A., De Rainville, F.-M., Gardner, M.-A., Parizeau, M., e Gagné, C. (2012). DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research*, 13:2171–2175.
- Fouad, K., Fischer, H., e Büschges, A. (2002). Comparative psychology of motor systems. *Handbook of psychology*.

- Franceschini, N., Pichon, J.-M., Blanes, C., e Brady, J. (1992). From insect vision to robot vision [and discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 337(1281):283–294.
- Frenzel, S. e Pompe, B. (2007). Partial mutual information for coupling analysis of multivariate time series. *Physical review letters*, 99(20):204101.
- Friedman, N., Murphy, K., e Russell, S. (1998). Learning the structure of dynamic probabilistic networks. *Proceedings of the Fourteenth . . .*
- Friedman, N., Nachman, I., e Peér, D. (1999). Learning bayesian network structure from massive datasets: the «sparse candidate» algorithm. In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, pp. 206–215. Morgan Kaufmann Publishers Inc.
- Frigo, M. e Johnson, S. G. (1998). Fftw: An adaptive software architecture for the fft. In *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, vol. 3, pp. 1381–1384. IEEE.
- Frisoni, G. (2001). Structural imaging in the clinical diagnosis of alzheimer’s disease: problems and tools. *Journal of Neurology, Neurosurgery & Psychiatry*, 70(6):711–718.
- Friston, K., Jezzard, P., e Turner, R. (1994). Analysis of functional MRI time-series. *Human Brain Mapping*, 1:153–171.
- Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain connectivity*, 1(1):13–36.
- Fuller, S. B., Karpelson, M., Censi, A., Ma, K. Y., e Wood, R. J. (2014). Controlling free flight of a robotic fly using an onboard vision sensor inspired by insect ocelli. *Journal of The Royal Society Interface*, 11(97).
- Gallo, G., Longo, G., Pallottino, S., e Nguyen, S. (1993). Directed hypergraphs and applications. *Discrete applied mathematics*, 42(2):177–201.
- Geiger, D. e Heckerman, D. (1994). Learning gaussian networks. In *Proceedings of the Tenth international conference on Uncertainty in artificial intelligence*, pp. 235–243. Morgan Kaufmann Publishers Inc.
- Gelman, A., Carlin, J. B., Stern, H. S., e Rubin, D. B. (2014). *Bayesian data analysis*, vol. 2. Taylor & Francis.

- Gerhard, F., Kispersky, T., Gutierrez, G. J., Marder, E., Kramer, M., e Eden, U. (2013). Successful reconstruction of a physiological circuit with known connectivity from spiking activity alone.
- Golyandina, N. e Zhigljavsky, A. (2013). *Singular Spectrum Analysis for time series*. Springer Science & Business Media.
- Gómez-Herrero, G., Wu, W., Rutanen, K., Soriano, M. C., Pipa, G., e Vicente, R. (2015). Assessing coupling dynamics from an ensemble of time series. *Entropy*, 17(4):1958–1970.
- Gourévitch, B. e Eggermont, J. J. (2007). Evaluating information transfer between auditory cortical neurons. *Journal of Neurophysiology*, 97(3):2533–2543.
- Granger, C. W. (1988). Some recent development in a concept of causality. *Journal of econometrics*, 39(1-2):199–211.
- Grazzini, J. (2012). Analysis of the emergent properties: stationarity and ergodicity. *Journal of Artificial Societies and Social Simulation*, 15(2):7.
- Grzegorzcyk, M. e Husmeier, D. (2011). Non-homogeneous dynamic bayesian networks for continuous data. *Mach. Learn.*, 83(3):355–419.
- Grzegorzcyk, M., Husmeier, D., Edwards, K. D., Ghazal, P., e Millar, A. J. (2008). Modelling non-stationary gene regulatory processes with a non-homogeneous bayesian network and the allocation sampler. *Bioinformatics*, 24(18):2071–2078.
- Gurney, K. N. (2001). Information processing in dendrites: I. input pattern generalisation. *Neural networks*, 14(8):991–1004.
- Harrison, P. J. (2004). The hippocampus in schizophrenia: a review of the neuropathological evidence and its pathophysiological implications. *Psychopharmacology*, 174(1):151–162.
- Hassani, H. (2007). Singular spectrum analysis: Methodology and comparison. *Journal of Data Science*, 5:239–257.
- Haufe, S., Nikulin, V., e Nolte, G. (2011). Identifying brain effective connectivity patterns from eeg: performance of granger causality, dtf, pdc and psi on simulated data. *BMC Neuroscience*, 12(Suppl 1):P141.

- Hayduk, L., Cummings, G., Stratkotter, R., Nimmo, M., Grygoryev, K., Dosman, D., Gillespie, M., Pazderka-Robinson, H., e Boadu, K. (2003). Pearl's d-separation: One more step into causal thinking. *Structural Equation Modeling: A Multidisciplinary Journal*, 10(2):289–311.
- Heckerman, D., Geiger, D., e Chickering, D. M. (1995). Learning bayesian networks: The combination of knowledge and statistical data. vol. 20, pp. 197–243. Springer.
- Heckers, S. (2001). Neuroimaging studies of the hippocampus in schizophrenia. *Hippocampus*, 11(5):520–528.
- Heng, X. C., Qin, Z., Wang, X. H., e Shao, L. P. (2006). Research on learning bayesian networks by particle swarm optimization. *Information Technology Journal*, 5: 540-545.
- Hinrichs, H., Heinze, H., e Schoenfeld, M. (2006). Causal visual interactions as revealed by an information theoretic measure and fmri. *NeuroImage*, 31(3):1051 – 1060.
- Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M., e Bhattacharya, J. (2007). Causality detection based on information-theoretic approaches in time series analysis. *Physics Reports*, 441(1):1–46.
- Ibrahim, T. S., Zhao, Y., Krishnamurthy, N., Raval, S. B., Zhao, T., Wood, S., e Kim, J. (2013). 20-to-8 channel tx array with 32-channel adjustable receive-only insert for 7t head imaging. In *The 21th ISMRM Annual Meeting*, p. 4408.
- Ide, J. S., Zhang, S., e Chiang-shan, R. L. (2014a). Bayesian network models in brain functional connectivity analysis. *International Journal of Approximate Reasoning*, 55(1):23–35.
- Ide, J. S., Zhang, S., e shan R. Li, C. (2014b). Bayesian network models in brain functional connectivity analysis. *International Journal of Approximate Reasoning*, 55(1, Part 1):23 – 35. Applications of Bayesian Networks.
- Iglesias, J. E., Augustinack, J. C., Nguyen, K., Player, C. M., Player, A., Wright, M., Roy, N., Frosch, M. P., McKee, A. C., Wald, L. L., et al. (2015). A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution mri: application to adaptive segmentation of in vivo mri. *Neuroimage*, 115:117–137.
- Ince, R. A., Mazzoni, A., Bartels, A., Logothetis, N. K., e Panzeri, S. (2012). A novel test to determine the significance of neural selectivity to single and multiple potentially correlated

- stimulus features. *Journal of Neuroscience Methods*, 210(1):49 – 65. Special Issue on Computational Neuroscience.
- Iyer, S. P., Shafran, I., Grayson, D., Gates, K., Nigg, J. T., e Fair, D. A. (2013). Inferring functional connectivity in mri using bayesian network structure learning with a modified pc algorithm. *Neuroimage*, 75:165–175.
- Jeneson, A. e Squire, L. R. (2012). Working memory, long-term memory, and medial temporal lobe function. *Learning & Memory*, 19(1):15–25.
- Jia, Y. e Huan, J. (2010). Constructing non-stationary dynamic bayesian networks with a flexible lag choosing mechanism. *BMC bioinformatics*, 11(Suppl 6):S27.
- Jimenez-Fabian, R. e Verlinden, O. (2012). Review of control algorithms for robotic ankle systems in lower-limb orthoses, prostheses, and exoskeletons. *Medical engineering & physics*, 34(4):397–408.
- Kaiser, A. e Schreiber, T. (2002). Information transfer in continuous processes. *Physica D: Nonlinear Phenomena*, 166(1):43–62.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45.
- Kantz, H. e Schreiber, T. (2004). *Nonlinear Time Series Analysis*. Cambridge nonlinear science series. Cambridge University Press.
- Kennedy, J., Eberhart, R., et al. (1995). Particle swarm optimization. In *Proceedings of IEEE international conference on neural networks*, vol. 4, pp. 1942–1948. Perth, Australia.
- Kennedy, J. e Eberhart, R. C. (1997). A discrete binary version of the particle swarm algorithm. In *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, vol. 5, pp. 4104–4108 vol.5.
- Kimpo, R. R., Theunissen, F. E., e Doupe, A. J. (2003). Propagation of correlated activity through multiple stages of a neural circuit. *The Journal of neuroscience*, 23(13):5750–5761.
- Kittmann, R. (1997). Neural mechanisms of adaptive gain control in a joint control loop: muscle force and motoneuronal activity. *Journal of experimental biology*, 200(9):1383–1402.

- Ko, Y., Zhai, C., e Rodriguez-Zas, S. (2009). Inference of gene pathways using mixture bayesian networks. *BMC Systems Biology*, 3(1):54.
- Kojima, K., Perrier, E., Imoto, S., e Miyano, S. (2010). Optimal search on clustered structural constraint for learning bayesian network structure. *The Journal of Machine Learning Research*, 11:285–310.
- Koller, D. e Friedman, N. (2009). *Probabilistic graphical models: principles and techniques*. MIT press.
- Koller, D. e Pfeffer, A. (1997). Object-oriented bayesian networks. In *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence*, pp. 302–313. Morgan Kaufmann Publishers Inc.
- Kondoh, Y., Okuma, J., e Newland, P. L. (1995). Dynamics of neurons controlling movements of a locust hind leg: Wiener kernel analysis of the responses of proprioceptive afferents. *Journal of Neurophysiology*, 73(5):1829–1842.
- Kovac, M. (2014). The Bioinspiration Design Paradigm: A Perspective for Soft Robotics. *Soft Robotics*, 1(1):28–37.
- Kraff, O., Fischer, A., Nagel, A. M., Mönninghoff, C., e Ladd, M. E. (2015). Mri at 7 tesla and above: Demonstrated and potential capabilities. *Journal of Magnetic Resonance Imaging*, 41(1):13–33.
- Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Phys. D*, 42(1-3):12–37.
- Lauterbur, P. (1973). Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature*, 242(5394):190–191.
- Lebre, S., Becq, J., Devaux, F., Stumpf, M., e Lelandais, G. (2010). Statistical inference of the time-varying structure of gene-regulation networks. *BMC Systems Biology*, 4(1):130.
- Lee, J., Nemati, S., Silva, I., Edwards, B. A., Butler, J. P., e Malhotra, A. (2012). Transfer entropy estimation and directional coupling change detection in biomedical time series. *Biomedical engineering online*, 11(1):19.
- Li, J., Wang, Z. J., Palmer, S. J., e McKeown, M. J. (2008). Dynamic bayesian network modeling of fmri: a comparison of group-analysis methods. *Neuroimage*, 41(2):398–407.

- Li, Y. e Ngom, A. (2013). The max-min high-order dynamic Bayesian network learning for identifying gene regulatory networks from time-series microarray data. *2013 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pp. 83–90.
- Libby, L. A., Ekstrom, A. D., Ragland, J. D., e Ranganath, C. (2012). Differential connectivity of perirhinal and parahippocampal cortices within human hippocampal subregions revealed by high-resolution functional imaging. *The Journal of Neuroscience*, 32(19):6550–6560.
- Lizier, J., Rubinov, M., et al. (2013). Multivariate construction of effective computational networks from observational data. In *Workshop on Information Processing in Cognition (IPCog-2013)*.
- Lizier, J. T. (2010). *The local information dynamics of distributed computation in complex systems*. Tese de Doutorado, School of Information Technologies, The University of Sydney.
- Lizier, J. T. (2014). Measuring the dynamics of information processing on a local scale in time and space. In *Directed Information Measures in Neuroscience*, pp. 161–193. Springer.
- Lizier, J. T., Heinzle, J., Horstmann, A., Haynes, J.-D., e Prokopenko, M. (2011). Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fmri connectivity. *Journal of computational neuroscience*, 30(1):85–107.
- Lizier, J. T., Prokopenko, M., e Zomaya, A. Y. (2008). Local information transfer as a spatio-temporal filter for complex systems. *Physical Review E*, 77(2):026110.
- Lohmann, G., Stelzer, J., Neumann, J., Ay, N., e Turner, R. (2013). “more is different” in functional magnetic resonance imaging: a review of recent data analysis techniques. *Brain connectivity*, 3(3):223–239.
- Maciel, C. D., Simpson, D. M., e Newland, P. L. (2012). Inference about multiple pathways in motor control limb in locust. In *Biosignals*, pp. 69–75.
- Madigan, D., York, J., e Allard, D. (1995). Bayesian graphical models for discrete data. *International Statistical Review/Revue Internationale de Statistique*, pp. 215–232.
- Mansfield, P. (1977). Multi-planar image formation using nmr spin echoes. *Journal of Physics C: Solid State Physics*, 10(3):L55.

- Marinazzo, D., Wu, G., Pellicoro, M., e Stramaglia, S. (2014). Information transfer in the brain: Insights from a unified approach. In *Directed Information Measures in Neuroscience*, pp. 87–110. Springer.
- Marmarelis, V. Z. (2004). *Nonlinear dynamic modeling of physiological systems*, vol. 10. John Wiley & Sons.
- Maruszak, A. e Thuret, S. (2014). Why looking at the whole hippocampus is not enough — a critical role for anteroposterior axis, subfield and activation analyses to enhance predictive value of hippocampal changes for alzheimer’s disease diagnosis. *Frontiers in Cellular Neuroscience*, 8:95.
- McDonnell, M. D., Ikeda, S., e Manton, J. H. (2011). An introductory review of information theory in the context of computational neuroscience. *Biological cybernetics*, 105(1):55–70.
- McGill, W. (1954). Multivariate information transmission. *Psychometrika*, 19(2):97–116.
- Meruelo, A. C., Simpson, D. M., Veres, S. M., e Newland, P. L. (2016). Improved system identification using artificial neural networks and analysis of individual differences in responses of an identified neuron. *Neural Networks*, 75:56 – 65.
- Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., e Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827.
- Mitchell, M. (2009). *Complexity: A guided tour*. Oxford University Press.
- Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., e Uğurbil, K. (2010a). Multiband multislice ge-epi at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fmri. *Magnetic Resonance in Medicine*, 63(5):1144–1153.
- Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., e Uğurbil, K. (2010b). Multiband multislice ge-epi at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fmri. *Magnetic Resonance in Medicine*, 63(5):1144–1153.
- Monge, P. R. e Contractor, N. S. (2003). *Theories of communication networks*. Oxford University Press.

- Mortimer, A. M., Likeman, M., e Lewis, T. T. (2013). Neuroimaging in dementia: a practical guide. *Practical neurology*, 13(2):92–103.
- Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W., Trojanowski, J. Q., Toga, A. W., e Beckett, L. (2005). Ways toward an early diagnosis in alzheimer’s disease: the alzheimer’s disease neuroimaging initiative (adni). *Alzheimer’s & Dementia*, 1(1):55–66.
- Murphy, K. P. (2001). The bayes net toolbox for matlab. *Computing Science and Statistics*, 33:2001.
- Murphy, K. P. (2002). *Dynamic bayesian networks: representation, inference and learning*. Tese de Doutorado, University of California, Berkeley.
- Mutlu, A. Y. e Aviyente, S. (2009). Inferring effective connectivity in the brain from eeg time series using dynamic bayesian networks. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pp. 4739–4742.
- Newland, P. L. (1991). Morphology and somatotopic organisation of the central projections of afferents from tactile hairs on the hind leg of the locust. *Journal of Comparative Neurology*, 312(4):493–508.
- Newland, P. L. e Kondoh, Y. (1997). Dynamics of neurons controlling movements of a locust hind leg ii. flexor tibiae motor neurons. *Journal of Neurophysiology*, 77(4):1731–1746.
- Nezamabadi-pour, H., Rostami-Shahrabaki, M., e Maghfoori-Farsangi, M. (2008). Binary particle swarm optimization: challenges and new solutions. *CSI J Comput Sci Eng*, 6(1):21–32.
- Paluš, M., Komárek, V., Hrnčíř, Z., e Štěrbová, K. (2001). Synchronization as adjustment of information rates: detection from bivariate time series. *Physical Review E*, 63(4):046211.
- Palus, M. e Novotná, D. (1998). Detecting modes with nontrivial dynamics embedded in colored noise: Enhanced monte carlo ssa and the case of climate oscillations. *Physics Letters A*, 248(2):191–202.
- Pampu, N., Vicente, R., Muresan, R., Priesemann, V., Siebenhuhner, F., e Wibral, M. (2013). Transfer entropy as a tool for reconstructing interaction delays in neural signals. In *Signals, Circuits and Systems (ISSCS), 2013 International Symposium on*, pp. 1–4.

- Papoulis, A. e Pillai, S. U. (2002). *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education.
- Patel, A. X., Kundu, P., Rubinov, M., Jones, P. S., Vértes, P. E., Ersche, K. D., Suckling, J., e Bullmore, E. T. (2014). A wavelet method for modeling and despiking motion artifacts from resting-state fmri time series. *Neuroimage*, 95:287–304.
- Patnaik, D., Laxman, S., e Ramakrishnan, N. (2011). Discovering excitatory relationships using dynamic bayesian networks. *Knowledge and information systems*, 29(2):273–303.
- Pearl, J. (2000). *Causality: models, reasoning and inference*, vol. 29. Cambridge Univ Press.
- Pearl, J. (2011). Bayesian networks. *Department of Statistics, UCLA*.
- Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., e Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Academic press.
- Pérez, F. e Granger, B. E. (2007). IPython: a system for interactive scientific computing. *Computing in Science and Engineering*, 9(3):21–29.
- Pérez-Ariza, C. B., Nicholson, A. E., Korb, K. B., Mascaro, S., e Hu, C. H. (2012). Causal discovery of dynamic bayesian networks. In *Australasian Joint Conference on Artificial Intelligence*, pp. 902–913. Springer.
- Pflüger, H.-J. e Duch, C. (2011). Dynamic neural control of insect muscle metabolism related to motor behavior. *Physiology*, 26(4):293–303.
- Pipitone, J., Park, M. T. M., Winterburn, J., Lett, T. A., Lerch, J. P., Pruessner, J. C., Lepage, M., Voineskos, A. N., Chakravarty, M. M., Initiative, A. D. N., et al. (2014). Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. *Neuroimage*, 101:494–512.
- Poppenk, J., Evensmoen, H. R., Moscovitch, M., e Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends in cognitive sciences*, 17(5):230–240.
- Poustchi-Amin, M., Mirowitz, S. A., Brown, J. J., McKinstry, R. C., e Li, T. (2001). Principles and applications of echo-planar imaging: A review for the general radiologist 1. *Radiographics*, 21(3):767–779.
- Prabhakara, R. S., Wright, C. H., e Barrett, S. F. (2012). Motion detection: A biomimetic vision sensor versus a ccd camera sensor. *Sensors Journal, IEEE*, 12(2):298–307.

- Preibisch, C., Bührer, M., Riedl, V., et al. (2015). Evaluation of multiband epi acquisitions for resting state fmri. *PloS one*, 10(9):e0136961.
- Prescott, S. A. e Sejnowski, T. J. (2008). Spike-rate coding and spike-time coding are affected oppositely by different adaptation mechanisms. *Journal of Neuroscience*, 28(50):13649–13661.
- Rabiner, L. e Juang, B.-H. (1986). An introduction to hidden markov models. *ASSP Magazine, IEEE*, 3(1):4–16.
- Rajapakse, J. C. e Zhou, J. (2007). Learning effective brain connectivity with dynamic bayesian networks. *Neuroimage*, 37(3):749–760.
- Ramsey, J. D., Hanson, S. J., Hanson, C., Halchenko, Y. O., Poldrack, R. A., e Glymour, C. (2010). Six problems for causal inference from fmri. *neuroimage*, 49(2):1545–1558.
- Rauh, J., Bertschinger, N., Olbrich, E., e Jost, J. (2014). Reconsidering unique information: Towards a multivariate information decomposition. In *Information Theory (ISIT), 2014 IEEE International Symposium on*, pp. 2232–2236. IEEE.
- Razi, A. e Friston, K. J. (2016). The connected brain: Causality, models, and intrinsic dynamics. *IEEE Signal Processing Magazine*, 33(3):14–35.
- Rinck, P. (2001). *Magnetic Resonance in Medicine*. Basic Textbook of the European Magnetic Resonance Forum. Wiley.
- Robinson, J. W. e Hartemink, A. J. (2010). Learning non-stationary dynamic bayesian networks. *Journal of Machine Learning Research*, 11(Dec):3647–3680.
- Robinson, R. W. (1977). Counting unlabeled acyclic digraphs. In Little, C. (Ed.), *Combinatorial Mathematics V*, vol. 622 of *Lecture Notes in Mathematics*, chapter 2, pp. 28–43. Springer Berlin Heidelberg.
- Rodrigues, J. e Andrade, A. (2014). Lag-based effective connectivity applied to fmri: a simulation study highlighting dependence on experimental parameters and formulation. *NeuroImage*, 89:358–377.
- Rozanov, Y. A. (1982). Markov random fields. In *Markov Random Fields*, pp. 55–102. Springer.

- Runge, J. (2014). *Detecting and quantifying causality from time series of complex systems*. Tese de Doutorado, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät.
- Runge, J., Heitzig, J., Petoukhov, V., e Kurths, J. (2012). Escaping the curse of dimensionality in estimating multivariate transfer entropy. *Phys. Rev. Lett.*, 108:258701.
- Rykhlevskaia, E., Gratton, G., e Fabiani, M. (2008). Combining structural and functional neuroimaging data for studying brain connectivity: a review. *Psychophysiology*, 45(2):173–187.
- Sahin, F. e Devasia, A. (2007). Distributed particle swarm optimization for structural Bayesian network learning. *Swarm Intelligence, Focus on Ant and Particle Swarm Optimization*.
- Sameshima, K. e BaccalÃ¡, L. A. (1999). Using partial directed coherence to describe neuronal ensemble interactions. *Journal of Neuroscience Methods*, 94(1):93 – 103.
- Schelter, B., Winterhalder, M., Eichler, M., Peifer, M., Hellwig, B., Guschlbauer, B., Lücking, C. H., Dahlhaus, R., e Timmer, J. (2006). Testing for directed influences among neural signals using partial directed coherence. *Journal of neuroscience methods*, 152(1):210–219.
- Schreiber, T. (2000). Measuring information transfer. *Physical review letters*, 85(2):461.
- Schreiber, T. e Schmitz, A. (2000). Surrogate time series. *Physica D: Nonlinear Phenomena*, 142(3):346–382.
- Schulz, M., Chau, W., Graham, S. J., McIntosh, A. R., Ross, B., Ishii, R., e Pantev, C. (2004). An integrative meg–fmri study of the primary somatosensory cortex using cross-modal correspondence analysis. *NeuroImage*, 22(1):120–133.
- Scott, D. W. e Sain, S. R. (2005). 9-multidimensional density estimation. *Handbook of statistics*, 24:229–261.
- Seth, A. K., Chorley, P., e Barnett, L. C. (2013). Granger causality analysis of fmri {BOLD} signals is invariant to hemodynamic convolution but not downsampling. *NeuroImage*, 65:540 – 555.
- Shah, A. e Woolf, P. (2009). Python environment for bayesian learning: Inferring the structure of bayesian networks from knowledge and data. *Journal of Machine Learning Research*, 10(Feb):159–162.

- Silchenko, A. N., Adamchic, I., Pawelczyk, N., Hauptmann, C., Maarouf, M., Sturm, V., e Tass, P. A. (2010). Data-driven approach to the estimation of connectivity and time delays in the coupling of interacting neuronal subsystems. *Journal of Neuroscience Methods*, 191(1):32 – 44.
- Singh, M. e Valtorta, M. (1995). Construction of bayesian network structures from data: a brief survey and an efficient algorithm. *International journal of approximate reasoning*, 12(2):111–131.
- Sinha, A. (2015). *Bio-mimetic navigation in a dynamic environment*. Tese de Doutorado, University of Technology, Sydney.
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., e Woolrich, M. W. (2011). Network modelling methods for fmri. *Neuroimage*, 54(2):875–891.
- Smith, V. A., Yu, J., Smulders, T. V., Hartemink, A. J., e Jarvis, E. D. (2006). Computational inference of neural information flow networks. *PLoS computational biology*, 2(11):e161.
- Sohn, S.-M. (2013). *RF and electronic design perspective on ultra-high field MRI systems*. Tese de Doutorado, University of Minnesota.
- Song, L., Kolar, M., e Xing, E. P. (2009). Time-varying dynamic bayesian networks. In *Advances in neural information processing systems*, pp. 1732–1740.
- Spirtes, P. e Glymour, C. (1991). An algorithm for fast recovery of sparse causal graphs. *Social science computer review*, 9(1):62–72.
- Spirtes, P., Glymour, C. N., e Spirtes, P. (1990). Causal structure among measured variables preserved with unmeasured variables. Technical report, Carnegie Mellon University.
- Sporns, O. (2011). *Networks of the Brain*. MIT Press.
- Srinivasan, M. V., Chahl, J. S., Weber, K., Venkatesh, S., Nagle, M. G., e Zhang, S.-W. (1999). Robot navigation inspired by principles of insect vision. *Robotics and Autonomous Systems*, 26(2):203–216.
- Stephan, K. E. e Friston, K. J. (2010). Analyzing effective connectivity with functional magnetic resonance imaging. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(3):446–459.

- Stramaglia, S., Wu, G.-R., Pellicoro, M., e Marinazzo, D. (2012). Expanding the transfer entropy to identify information circuits in complex systems. *Physical Review E*, 86(6):066211.
- Stratonovich, R. (1960). Conditional markov processes. *Theory of Probability & Its Applications*, 5(2):156–178.
- Studený, M. e Vejnarová, J. (1998). The multiinformation function as a tool for measuring stochastic dependence. In *Learning in graphical models*, pp. 261–297. Springer.
- Thrun, S., Burgard, W., e Fox, D. (2005). *Probabilistic robotics*. MIT press.
- Timme, N., Alford, W., Flecker, B., e Beggs, J. M. (2014). Synergy, redundancy, and multivariate information measures: an experimentalist’s perspective. *Journal of computational neuroscience*, 36(2):119–140.
- Tsamardinos, I., Brown, L., e Aliferis, C. (2006). The max-min hill-climbing bayesian network structure learning algorithm. *Machine Learning*, 65(1):31–78.
- Valdes-Sosa, P. A., Roebroek, A., Daunizeau, J., e Friston, K. (2011). Effective connectivity: influence, causality and biophysical modeling. *Neuroimage*, 58(2):339–361.
- van der Kolk, A. G., Hendrikse, J., Zwanenburg, J. J., Visser, F., e Luijten, P. R. (2013a). Clinical applications of 7 t {MRI} in the brain. *European Journal of Radiology*, 82(5):708 – 718.
- van der Kolk, A. G., Hendrikse, J., Zwanenburg, J. J., Visser, F., e Luijten, P. R. (2013b). Clinical applications of 7t mri in the brain. *European journal of radiology*, 82(5):708–718.
- Van Leemput, K., Bakkour, A., Benner, T., Wiggins, G., Wald, L. L., Augustinack, J., Dickerson, B. C., Golland, P., e Fischl, B. (2009). Automated segmentation of hippocampal subfields from ultra-high resolution in vivo mri. *Hippocampus*, 19(6):549–557.
- van Steen, M. (2010). *Graph Theory and Complex Networks: An Introduction*. On Demand Publishing, LLC-Create Space.
- Vautard, R., Yiou, P., e Ghil, M. (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D: Nonlinear Phenomena*, 58(1):95–126.
- Venema, V., Ament, F., e Simmer, C. (2006). A stochastic iterative amplitude adjusted fourier transform algorithm with improved accuracy. *Nonlinear Processes in Geophysics*, 13(3):321–328.

- Vicente, R., Wibral, M., Lindner, M., e Pipa, G. (2011). Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *Journal of computational neuroscience*, 30(1):45–67.
- Vidal-Gadea, A. G., Jing, X., Simpson, D., Dewhirst, O. P., Kondoh, Y., Allen, R., e Newland, P. L. (2010). Coding characteristics of spiking local interneurons during imposed limb movements in the locust. *Journal of neurophysiology*, 103(2):603.
- Villanueva, E. e Maciel, C. D. (2010). Modeling associations between genetic markers using bayesian networks. *Bioinformatics*, 26(18):i632–i637.
- Villanueva, E. e Maciel, C. D. (2014). Efficient methods for learning bayesian network superstructures. *Neurocomputing*, 123(0):3 – 12.
- Villanueva Talavera, E. R. (2012). *Sistema evolutivo eficiente para aprendizagem estrutural de redes Bayesianas*. Tese de Doutorado, Escola de Engenharia de São Carlos.
- Vinh, N. X., Chetty, M., Coppel, R., e Wangikar, P. P. (2011a). Globalmit: learning globally optimal dynamic bayesian network with the mutual information test criterion. *Bioinformatics*, 27(19):2765–2766.
- Vinh, N. X., Chetty, M., Coppel, R., e Wangikar, P. P. (2011b). Polynomial time algorithm for learning globally optimal dynamic bayesian network. In *International Conference on Neural Information Processing*, pp. 719–729. Springer.
- Wan, E. A. e Van Der Merwe, R. (2000). The unscented kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pp. 153–158. IEEE.
- Wang, C., Xu, J., Zhao, S., e Lou, W. (2016). Graph theoretical analysis of eeg effective connectivity in vascular dementia patients during a visual oddball task. *Clinical Neurophysiology*, 127(1):324–334.
- Wang, M., Chen, Z., e Cloutier, S. (2007). A hybrid bayesian network learning method for constructing gene networks. *Computational biology and chemistry*, 31(5):361–372.
- Watanabe, A. e Stark, L. (1975). Kernel method for nonlinear analysis: Identification of a biological control system. *Mathematical Biosciences*, 27(1):99–108.

- Watanabe, S. (1960). Information theoretical analysis of multivariate correlation. *IBM Journal of research and development*, 4(1):66–82.
- Wessnitzer, J. e Webb, B. (2006). Multimodal sensory integration in insects - towards insect brain control architectures. *Bioinspiration & Biomimetics*, 1(3):63.
- Wibral, M., Pampu, N., Priesemann, V., Siebenhühner, F., Seiwert, H., Lindner, M., Lizier, J. T., e Vicente, R. (2013). Measuring information-transfer delays. *PloS one*, 8(2):e55809.
- Wibral, M., Vicente, R., e Lindner, M. (2014a). Transfer entropy in neuroscience. In Wibral, M., Vicente, R., e Lizier, J. T. (Eds.), *Directed Information Measures in Neuroscience, Understanding Complex Systems*, pp. 3–36. Springer Berlin Heidelberg.
- Wibral, M., Vicente, R., e Lizier, J. T. (2014b). *Directed information measures in neuroscience*. Springer.
- Wibral, M., Wollstadt, P., Meyer, U., Pampu, N., Priesemann, V., e Vicente, R. (2012). Revisiting wiener’s principle of causality: interaction-delay reconstruction using transfer entropy and multivariate analysis on delay-weighted graphs. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 3676–3679.
- Williams, P. L. e Beer, R. D. (2010). Nonnegative decomposition of multivariate information. *arXiv preprint arXiv:1004.2515*.
- Williams, P. L. e Beer, R. D. (2011). Generalized measures of information transfer. *arXiv preprint arXiv:1102.1507*.
- Wilmer, A., de Lussanet, M., e Lappe, M. (2012). Time-delayed mutual information of the phase as a measure of functional connectivity. *PLoS ONE*, 7(9):e44633.
- Wisse, L., Kuijf, H., Honingh, A., Wang, H., Pluta, J., Das, S., Wolk, D., Zwanenburg, J., Yushkevich, P., e Geerlings, M. (2016). Automated hippocampal subfield segmentation at 7t mri. *American Journal of Neuroradiology*, 37(6):1050–1057.
- Wisse, L. E., Biessels, G. J., Heringa, S. M., Kuijf, H. J., Luijten, P. R., Geerlings, M. I., Group, U. V. C. I. V. S., et al. (2014). Hippocampal subfield volumes at 7t in early alzheimer’s disease and normal aging. *Neurobiology of aging*, 35(9):2039–2045.
- Wollstadt, P., Martínez-Zarzuela, M., Vicente, R., Díaz-Pernas, F. J., e Wibral, M. (2014). Efficient transfer entropy analysis of non-stationary neural time series. *PloS one*, 9(7):e102833.

- Wu, X., Wen, X., Li, J., e Yao, L. (2014). A new dynamic bayesian network approach for determining effective connectivity from fmri data. *Neural Computing and Applications*, 24(1):91–97.
- Xie, F. e Levinson, D. (2009). Modeling the growth of transportation networks: A comprehensive review. *Networks and Spatial Economics*, 9(3):291–307.
- Xuan, N., Chetty, M., Coppel, R., e Wangikar, P. P. (2012a). Gene regulatory network modeling via global optimization of high-order dynamic bayesian network. *BMC bioinformatics*, 13(1):131.
- Xuan, N. V., Chetty, M., Coppel, R., e Wangikar, P. P. (2012b). Gene regulatory network modeling via global optimization of high-order dynamic Bayesian network. *BMC bioinformatics*, 13:131.
- Yang, C., Jeannès, R. L. B., Faucon, G., e Shu, H. (2013). Detecting information flow direction in multivariate linear and nonlinear models. *Signal Processing*, 93(1):304–312.
- Ye, S., Cai, H., e Sun, R. (2008). An algorithm for bayesian networks structure learning based on simulated annealing with mdl restriction. In *Natural Computation, 2008. ICNC '08. Fourth International Conference on*, vol. 3, pp. 72–76.
- Ying, C., Jia-fan, Z., Can-jun, Y., e Bin, N. (2007). Design and hybrid control of the pneumatic force-feedback systems for arm-exoskeleton by using on/off valve. *Mechatronics*, 17(6):325 – 335.
- Yushkevich, P. A., Amaral, R. S., Augustinack, J. C., Bender, A. R., Bernstein, J. D., Boccardi, M., Bocchetta, M., Burggren, A. C., Carr, V. A., Chakravarty, M. M., ChÃ©telat, G., Daugherty, A. M., Davachi, L., Ding, S.-L., Ekstrom, A., Geerlings, M. I., Hassan, A., Huang, Y., Iglesias, J. E., Joie, R. L., Kerchner, G. A., LaRocque, K. F., Libby, L. A., Malykhin, N., Mueller, S. G., Olsen, R. K., Palombo, D. J., Parekh, M. B., Pluta, J. B., Preston, A. R., Pruessner, J. C., Ranganath, C., Raz, N., Schlichting, M. L., Schoemaker, D., Singh, S., Stark, C. E., Suthana, N., Tompary, A., Turowski, M. M., Leemput, K. V., Wagner, A. D., Wang, L., Winterburn, J. L., Wisse, L. E., Yassa, M. A., e Zeineh, M. M. (2015a). Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo mri: Towards a harmonized segmentation protocol. *NeuroImage*, 111:526 – 541.

- Yushkevich, P. A., Pluta, J. B., Wang, H., Xie, L., Ding, S.-L., Gertje, E. C., Mancuso, L., Kliot, D., Das, S. R., e Wolk, D. A. (2015b). Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. *Human brain mapping*, 36(1):258–287.
- Żak-Szatkowska, M. e Bogdan, M. (2011). Modified versions of the bayesian information criterion for sparse generalized linear models. *Computational Statistics & Data Analysis*, 55(11):2908–2924.
- Zhang, L., Samaras, D., Alia-Klein, N., Volkow, N., e Goldstein, R. (2005). Modeling neuronal interactivity using dynamic bayesian networks. In *Advances in neural information processing systems*, pp. 1593–1600.
- Zheng, X. e Rajapakse, J. C. (2006). Learning functional structure from fmr images. *Neuroimage*, 31(4):1601–1613.
- Zhu, D., Zhang, T., Jiang, X., Hu, X., Chen, H., Yang, N., Lv, J., Han, J., Guo, L., e Liu, T. (2014). Fusing DTI and fMRI data: a survey of methods and applications. *Neuroimage*, 102:184–191.