**Universidade de São Paulo
Instituto de Física de São Carlos**

César Henrique Comin

**A relação entre estrutura e função em redes
complexas observada localmente**

São Carlos
2012

César Henrique Comin

A relação entre estrutura e função em redes complexas observada localmente

Dissertação apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para a obtenção do título de Mestre em Física.

Área de Concentração: Física Aplicada

Opção: Física Computacional

Orientador: Luciano da Fontoura Costa

Versão original

São Carlos
2012

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

Ficha catalográfica elaborada pelo Serviço de Biblioteca e Informação do IFSC,
com os dados fornecidos pelo(a) autor(a)

Comin, César

A relação entre estrutura e função em redes
complexas observada localmente / César Comin;
orientador Luciano Costa -- São Carlos, 2012.
118 p.

Dissertação (Mestrado - Programa de Pós-Graduação em
Física Aplicada Computacional) -- Instituto de Física
de São Carlos, Universidade de São Paulo, 2012.

1. Redes complexas. 2. integra-e-dispara. 3.
Diferenciação dinâmica. 4. Dinâmica neuronal. 5.
Comportamento local. I. Costa, Luciano, orient. II.
Titulo.

FOLHA DE APROVAÇÃO

Cesar Henrique Comin

Dissertação apresentada ao Instituto de Física de São Carlos da Universidade de São Paulo para obtenção do título de Mestre em Ciências.
Área de Concentração: Física Aplicada -
Opção: Física Computacional.

Aprovado(a) em: 14.02.2012

Comissão Julgadora

Prof(a). Dr(a). Luciano da Fontoura Costa

Instituição: IFSC/USP

Prof(a). Dr(a). Alneu de Andrade Lopes

Instituição: ICMC/USP

Prof(a). Dr(a). Claudio José Magon

Instituição: IFSC/USP

AGRADECIMENTOS

Agradeço a minha família, pela paciência e confiança.

Aos meus professores ao longo de toda a vida, por serem responsáveis pelo conhecimento que possuo hoje. Especialmente aos professores do IFSC, cuja dedicação à ciência inspira qualquer aluno.

Aos meus companheiros de fiscomp, pelos momentos de alegria e discussões ao longo da graduação.

Ao João, cujas discussões sobre redes e exemplo de profissional me fizeram o pesquisador que sou hoje.

Ao Thomas, pelas longas conversas sobre o sentido da vida.

Ao professor Luciano, que me levou ao caminho da ciência.

Aos companheiros de grupo, pelas frutíferas discussões sobre redes.

A CAPES, pelo apoio financeiro.

A Richard Feynman que, indiretamente, me ensinou o que é ser um cientista.

“The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead, at each level of complexity entirely new properties appear, and the understanding of the new behaviors requires research which I think is as fundamental in its nature as any other.”

— PHILIP WARREN ANDERSON

RESUMO

COMIN, C. H. *A relação entre estrutura e função em redes complexas observada localmente*. 2012. 118 p. Dissertação (Mestrado em ciência) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2012.

O estudo de redes complexas tem despertado muita atenção nos últimos anos, principalmente pela sua capacidade de permitir a análise dos mais diversificados sistemas através de um mesmo conjunto de ferramentas matemáticas e computacionais. Até pouco tempo a ênfase nessa área era sobre o estudo das propriedades estruturais e sua influência em características globais da dinâmica ocorrendo sobre o sistema. Recentemente foi-se percebendo a riqueza de comportamentos que podemos estudar ao olharmos para grupos de vértices e as interações que ocorrem entre esses grupos, de forma que cada vez mais se torna necessária a criação de uma forma sistemática de quantificar, a nível de nó, como uma dinâmica é influenciada (ou diferenciada) pela estrutura do sistema. Apresentamos nesse trabalho um primeiro passo nessa direção, no qual definimos uma medida que, baseada em características dinâmicas obtidas em diversas condições iniciais, busca comparar o comportamento dos nós presentes no sistema. Através dessa medida encontramos a alta capacidade de modelos de rede geográficas de produzir, por simples flutuações estatísticas, grupos de nós de dinâmica muito distinta em relação ao demais, um fato que não ocorre para uma rede definida de forma semelhante mas não geográfica. Nessas redes conseguimos verificar também se um nó de topologia local distinta dos demais consegue possuir uma dinâmica diferenciada, o que encontramos foi que mesmo em regiões da rede extremamente regulares existe a formação de grupos dinâmicos devido ao efeito da borda dessa região. Saindo de modelos de redes para apresentar aplicações práticas do método, encontramos na rede neuronal do verme *Caenorhabditis elegans* um conjunto de nós de alta diferenciação dinâmica, que representam os interneurônios do cordão ventral do verme. Adicionalmente encontramos na rede cortical do macaco da família *Cercopithecidae* uma divisão em relação a regularidade dos sinais dinâmicos, o que indica a presença de comunidades funcionais nessa rede. Além da metodologia de diferenciação desenvolvemos ainda uma ferramenta que busca encontrar de forma totalmente automatizada as características dinâmicas mais relevantes do sistema em estudo, que foi capaz de representar medidas dinâmicas tradicionalmente utilizadas na área. Todos esses resultados abrem caminho para diversas vertentes de estudo, em especial citamos a influência de uma borda irregular no interior de uma região regular, o estudo da rede *Caenorhabditis elegans* com dinâmicas neuronais mais precisas e a aplicação sistemática de dinâmicas para encontrar a divisão funcional de comunidades em redes direcionadas, um tema que apresenta resultados promissores.

Palavras-chave: Redes complexas. Integra-e-dispara. Diferenciação dinâmica. Dinâmica neuronal. Comportamento local.

ABSTRACT

COMIN, C. H. *The relationship between structure and function in complex networks observed locally*. 2012. 118 p. Dissertação (Mestrado em ciência) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2012.

The study of complex networks has drawn much attention over the last years, mainly by virtue of its potential to characterize the most diverse systems through the same mathematical and computational tools. Not long ago the emphasis on this field mostly focused on the effects of the structural properties on the global behavior of a dynamical process taking place in the system. Recently, some studies started to unveil the richness of interactions that occur between groups of nodes when we look at the small scale of interactions occurring in the network. Such findings call for a new systematic methodology to quantify, at node level, how a dynamics is being influenced (or differentiated) by the structure of the underlying system. Here we present a first step towards this direction, in which we define a new measurement that, based on dynamical characteristics obtained for a series of initial conditions, compares the dynamical behavior of the nodes present in the system. Through this measurement we find the high capacity of networks generated by geographical models to exhibit, by means of statistical fluctuations, groups of nodes with very distinct dynamics compared to the rest of the network, a behavior that does not occur for a similar non-geographical network. We also verify if a large topological differentiation of a node necessarily reflects on its dynamics. We find that even in very regular regions of the network the nodes tend to form dynamical groups influenced by the border effects. In addition to the network models used, we present practical applications of the methodology by using the neuronal network of the nematode *Caenorhabditis elegans*, where we show that the interneurons of the ventral cord presents a very large dynamical differentiation when compared to the rest of the network. We also analyze the cortical network of the *Cercopithecidae* monkey by means of signal communication complexity, finding that it contains two well-defined functional communities. Besides the differentiation measurement, we also presents an useful mechanism for automatically obtaining the relevant dynamical characteristics of the nodes, which showed promising results by obtaining traditional measurements of the area with little effort. All these results pave the way to a range of different studies, of which we highlight the influence of an irregular border on the dynamics taking place inside a regular network region, the study of the neurons of *Caenorhabditis elegans* using more robust neuronal dynamics and the systematic application of different dynamics in order to find the functional community division of directed networks.

Keywords: Complex networks. Integrate-and-fire. Dynamical differentiation. Neuronal dynamics. Local behavior.

LISTA DE FIGURAS

Figura 1.1 - Visão geral do método proposto. (a) Uma dinâmica é aplicada sobre a rede, (b) o sinal obtido ao longo do tempo para cada nó é caracterizado por um conjunto de medidas dinâmicas (exemplificado na figura para duas medidas, f^1 e f^2 , aplicadas sobre os nós vermelho e azul), (c) após uma realização cada nó possui uma posição bem definida no espaço de medidas, ao exercitarmos diversas realizações podemos ter a formação de (d) grupos bem definidos ou (e) nenhuma distinção aparente.	28
Figura 2.1 - Densidade de probabilidade de duas variáveis aleatórias com (a) pouca variância e (b) larga variância.	45
Figura 2.2 - Exemplo de significância estatística, nos três gráficos a distância euclidiana entre a média dos grupos é sempre a mesma, mas os casos (a) e (b) possuem uma distância mais significativa do que o caso (c).	46
Figura 2.3 - Exemplo de aglomeração de dados artificiais. (a) 6 amostras são inseridas em um espaço bidimensional através de duas medidas f^1 e f^2 , (b) aplicando um método de aglomeração obtemos o dendrograma dos dados. Como a distância necessária para unir os dois últimos grupos é muito maior do que em qualquer outra união, a melhor divisão das amostras é em dois grupos (c).	49
Figura 2.4 - Exemplo da aplicação do PCA na redução de dimensionalidade de dados.	52
Figura 3.1 - Exemplo de aplicação do método. (a) Executamos 12 realizações de uma dinâmica na rede e (b) projetamos os sinais de cada nó em um espaço definido por duas medidas dinâmicas. (c) Obtemos a matriz de distâncias estatísticas entre os pontos realizados de cada nó, (d) a média de cada linha da matriz da origem aos valores de α . (e) Por fim construímos o histograma de α	56
Figura 4.1 - Patamar de ativação da dinâmica integra-e-dispara em função do grau médio da rede, $\langle k \rangle$, após um tempo muito longo.	62
Figura 4.2 - Patamar de ativação da dinâmica integra-e-dispara ao considerarmos três diferentes tamanhos de rede.	63
Figura 4.3 - Patamar de ativação da dinâmica integra-e-dispara em função da média dos limiares sorteados para os nós de acordo com uma distribuição de Poisson.	64
Figura 4.4 - Patamar de ativação da dinâmica integra-e-dispara em função da média dos limiares sorteados para os nós de acordo com uma distribuição de Poisson. A rede utilizada foi gerada com o modelo BA.	65
Figura 4.5 - Rede geográfica utilizada, cada nó foi colorido de acordo com o seu grau.	66

Figura 4.6 - Taxa média ao longo de diversas realizações do integra-e-dispara para (a) $T_I = 9$, (b) $T_I = 10$ e (c) $T_I = 11$	67
Figura 4.7 - Entropia média ao longo das mesmas realizações e limiares da Figura 4.6.	67
Figura 4.8 - Valores (a) mínimos e (b) máximos obtidos para a taxa em 1000 realizações da dinâmica com $T_I = 10$	68
Figura 4.9 - Valores (a) mínimos e (b) máximos obtidos para a entropia em 1000 realizações da dinâmica com $T_I = 10$	69
Figura 4.10 - Valores de α obtidos ao considerarmos a (a) taxa, (b) entropia e (c) maior intervalo de disparo.	70
Figura 4.11 - Valores de α obtidos ao considerarmos a combinação de (a) taxa, entropia e maior intervalo e (b) taxa e entropia.	71
Figura 4.12 - Histograma de α relativos a (a) taxa, (b) entropia e (c) maior intervalo de disparo.	71
Figura 4.13 - Rede geográfica com os nós coloridos de acordo com os picos da Figura 4.12, para as medidas de (a) taxa, (b) entropia e (c) maior intervalo de disparo.	72
Figura 4.14 - Histogramas das medidas estruturas da rede geográfica utilizada, em vermelho está mostrada a fração de nós que pertencem ao respectivo pico do histograma de α_r	74
Figura 4.15 - α médio para cada grau k da rede, juntamente com o desvio padrão dos valores. As curvas foram sobrepostas ao histograma de graus da rede.	75
Figura 4.16 - Histograma de α relativos a combinação de (a) taxa, entropia e maior intervalo e (b) taxa e entropia.	76
Figura 4.17 - Histogramas de α relativos a medida de taxa para três modelos de rede.	77
Figura 4.18 - Histograma de α relativos a (a) taxa, (b) entropia e (c) maior intervalo para a rede <i>C. elegans</i>	78
Figura 4.19 - Histograma dos graus da rede <i>C. elegans</i> , os valores correspondentes aos interneurônios do cordão ventral foram coloridos em vermelho.	79
Figura 4.20 - Dendrograma da rede geográfica com região mais densa.	82
Figura 4.21 - Grupos formados na rede geográfica.	82
Figura 4.22 - Rede utilizada para a aglomeração dinâmica.	83
Figura 4.23 - Dendrograma obtido para as taxas dos nós na rede da Figura 4.22.	83
Figura 4.24 - Grupos formados na aglomeração da taxa para a rede geográfica.	84
Figura 4.25 - Grupos formados na aglomeração da entropia para a rede geográfica.	85
Figura 4.26 - Dendrograma da aglomeração dinâmica para a medida de entropia na rede cortical.	87

Figura 4.27 - Visualização da rede utilizando o algoritmo de Fruchterman-Reingold, os nós foram coloridos de acordo com (a) grupos formados pela aglomeração da entropia (b) comunidade identificada pelo algoritmo de Reichardt-Bornholdt.	87
Figura 4.28 - Histogramas de α_f relativos a frequência efetiva dos osciladores de Kuramoto, diversos acoplamentos são considerados. Na segunda linha o subgráfico mostra a ampliação do gráfico principal.	89
Figura 4.29 - Sincronização dos osciladores de Kuramoto. Mostramos o (a) nível de sincronia dos nós, (b) número de grupos de frequência efetiva semelhante e (c) tamanho do maior grupo formado, considerando diversos acoplamentos.	91
Figura 4.30 - (a) Contagem (ou histograma não normalizado) do valor absoluto das frequências efetivas da dinâmica de Kuramoto. (b) <i>Scatter-plot</i> entre o valor absoluto das frequências naturais e a centralidade de autovetor na mesma dinâmica. Foram consideradas 100 realizações em 100 redes distintas para um acoplamento $\lambda = 0.1$	92
Figura 4.31 - Projeção 3D da dinâmica integra-e-dispara utilizando o método de PCA. Adaptado de (1).	93
Figura 4.32 - Projeções 2D do PCA para a dinâmica integra-e-dispara. (a) e (b) nós projetados coloridos de acordo com o grau, (c) e (d) histogramas da densidade de nós em cada projeção.	94
Figura 4.33 - Exemplos de sinais de alguns nós projetados ao longo de PCA3. Adaptado de (1).	95
Figura 4.34 - Explicação da medida PCA3. (a) Nós coloridos de acordo com a medida de taxa e (b) <i>scatter-plot</i> entre taxa e PCA3. O coeficiente de correlação de Pearson, γ , é mostrado na Figura.	95
Figura 4.35 - Explicação da medida PCA1. (a) Nós coloridos de acordo com a medida s_{p-i} e (b) <i>scatter-plot</i> entre essa medida e PCA1.	96
Figura 4.36 - Explicação da medida PCA2. (a) Nós coloridos de acordo com a medida s_{pi-ip} e (b) <i>scatter-plot</i> entre essa medida e PCA2.	97
Figura A.1 - Ativação dos nós na dinâmica integra-e-dispara ao longo do tempo. A rede utilizada segue o modelo de Waxman e os parâmetros dinâmicos são os mesmos utilizados no texto principal.	114
Figura A.2 - Média e variação das três medidas consideradas para diferentes janelas utilizadas no cálculo. Novamente os parâmetros são os mesmos utilizados no texto principal para a rede geográfica.	115
Figura A.3 - Ativação dos nós na dinâmica integra-e-dispara ao longo do tempo para o modelo ER e os mesmos parâmetros dinâmicos utilizados no texto principal.	117
Figura A.4 - Média e variação das três medidas consideradas para diferentes janelas utilizadas no cálculo. As redes seguem o modelo ER.	118

LISTA DE TABELAS

Tabela 4.1 - Valores calculados para os interneurônios do cordão ventral da <i>C. elegans</i> , que obtiveram alta diferenciação dinâmica em relação à taxa. . . .	80
--	----

LISTA DE ABREVIATURAS

<i>ER</i>	<i>Erdős e Rényi</i>
<i>BA</i>	<i>Barabási e Albert</i>
<i>WS</i>	<i>Watts e Strogatz</i>
<i>PCA</i>	<i>Principal Component Analysis (análise por componentes principais)</i>
<i>ISE</i>	<i>Inter-Spike Entropy (entropia de intervalos)</i>
<i>Evcnt</i>	<i>Centralidade de autovetor</i>

LISTA DE SÍMBOLOS

N	Número de nós da rede
$G(V, E)$	Grafo definido pelo conjunto dos vértices V conectados pelas arestas do conjunto E
$G(V, p)$	Grafo definido pelo conjunto dos vértices V conectados de acordo com uma probabilidade p
A	Matriz de adjacência utilizada para representar o grafo
k_i	Grau do vértice i
P_k	Distribuição de grau
$\delta_{i,j}$	Delta de Kronecker
$\langle x \rangle$	Valor médio dos valores observados da variável aleatória X
C_i	Coefficiente de agrupamento (clustering) do nó i
$k_i^{(nm)}$	Grau médio dos nós vizinhos de i
T_l	Limiar de disparo de um neurônio
$R(t)$	Fração de nós disparando na dinâmica integra-e-dispara no instante t
s_i	Estado de disparo de um neurônio
V_i^t	Potencial do neurônio i no instante t
ω_i	Frequência natural de um oscilador na dinâmica de Kuramoto
θ_i	Fase de um oscilador na dinâmica de Kuramoto
λ	Acoplamento entre os osciladores na dinâmica de Kuramoto
R_k	Nível de sincronização dos osciladores na dinâmica de Kuramoto
r_i	Taxa média de disparos de um neurônio
T_{sim}	Tempo total de simulação da dinâmica
T_{est}	Tempo longo o suficiente definido após o transiente da dinâmica
I^0	Máximo intervalo de tempo que um neurônio fica sem disparar
$\tilde{\omega}_i$	Frequência efetiva de um oscilador na dinâmica de Kuramoto

$d(r, s)$	Distância euclidiana entre os elementos r e s
t	Distância estatística de Student
h	Distância estatística de Hotelling
$d_x(r, s)$	Distância normalizada
$\mathfrak{R}$	Número total de realizações da dinâmica
f	Medida dinâmica
Ξ	Matriz de distâncias dinâmicas entre os nós
α	Diferenciação dinâmica média de um nó
Λ	Matriz contendo os sinais dinâmicos dos nós ao longo do tempo
Σ	Matriz de covariância de um conjunto de variáveis aleatórias
Q	Modularidade de uma rede
γ	Coefficiente de correlação de Pearson
$\hat{x}$	Versão estandardizada da variável aleatória x
σ	Desvio padrão de uma variável aleatória
R_p	Patamar de ativação da dinâmica integra-e-dispara após um longo tempo

SUMÁRIO

1	Introdução	p. 25
2	Materiais	p. 31
2.1	Redes complexas (estrutura)	p. 31
2.1.1	Definições	p. 32
2.1.2	Modelos de redes aleatórias	p. 33
2.1.3	Redes reais	p. 36
2.1.4	Caracterização estrutural	p. 36
2.2	Dinâmica	p. 38
2.2.1	Integra-e-dispara	p. 38
2.2.2	Kuramoto	p. 41
2.2.3	Medidas dinâmicas	p. 42
2.3	Distâncias	p. 44
2.4	Aglomerção hierárquica	p. 47
2.5	Análise por componentes principais	p. 50
3	Métodos	p. 53
3.1	Discriminando a dinâmica entre os nós	p. 53
3.2	Encontrando medidas dinâmicas relevantes com o PCA	p. 57
4	Resultados e discussões	p. 61
4.1	Patamar do estado assíncrono	p. 61
4.2	Quantificando a discriminação dos nós	p. 65
4.2.1	Análise preliminar	p. 65
4.2.2	Cálculo de α	p. 68
4.2.3	Aglomerção dinâmica	p. 81
4.2.4	Osciladores de Kuramoto	p. 88
4.3	Aplicação do PCA	p. 92

5 Conclusões	p. 99
REFERÊNCIAS	p. 103
Apêndice A – Tempo para a estabilização da dinâmica	p. 113

1 Introdução

Ao estudarmos um sistema dinâmico composto por mais de um elemento, devemos considerar as possíveis interações que podem ocorrer entre os mesmos. Estudos tradicionais de sistemas clássicos consideram que a topologia de interação entre as partículas ocorre de forma totalmente regular, como por exemplo na forma de um reticulado bidimensional no qual cada partícula possui quatro vizinhos em posições relativas bem definidas (2). Infelizmente (ou felizmente), estruturas regulares são na realidade minoria na natureza, são conhecidos inúmeros exemplos de sistemas cuja topologia heterogênea surge nas mais variadas formas (3, 4). Na última década passou-se a dar grande atenção ao estudo dos mecanismos de criação dessas topologias (5, 6), esse interesse súbito teve um grande empurrão com os já clássicos trabalhos de Strogatz et al. (7) e Barabási et al. (8), ambos criaram modelos de geração de estruturas que permitiram estudar de forma analítica e experimental dois fenômenos muito observados na natureza: o primeiro é o fenômeno de mundo pequeno conhecido especialmente no escopo de interações sociais devido ao experimento de Milgran (9) que encontrou que em média cada habitante dos Estados Unidos deve possuir um grau de separação de no máximo 6, sendo que um grau de separação 2 simboliza a distância social entre um sujeito e o amigo do amigo desse sujeito. O segundo é o surgimento de leis de potência em muitos sistemas observados (10–13), que rivaliza diretamente com a distribuição Gaussiana tão estudada por físicos e matemáticos. Ainda no contexto de redes sociais, uma lei de potência representa o fato de que existe um grande número de pessoas com poucos contatos sociais, mas existem certas pessoas que conseguem manter uma larga rede de contatos, uma consequência direta do fenômeno “o rico fica mais rico” (14). O boom que ocorreu na área de redes complexas foi facilitado pelo já muito estudado campo de teoria de grafos, que em 1960 já possuía um modelo de criação de grafos totalmente aleatórios elaborado por Erdős e Rényi (conhecido como modelo ER) (15), que foi um dos primeiros tratamentos probabilísticos nessa área.

Recentemente o foco dos trabalhos em rede tem mudado gradativamente de uma abordagem puramente topológica para a análise de dinâmicas ocorrendo *sobre* a rede, e seu relacionamento com a estrutura (16–18). Com isso nos últimos anos as mais diversas dinâmicas vem sendo estudadas na área de redes complexas (4, 19–21), esses estudos geralmente estão preocupados com dois aspectos: (a) o comportamento de alguma estatística global da dinâmica, como por exemplo, o disparo geral dos neurônios passando por uma dinâmica neuronal (22, 23), a porcentagem de indivíduos infectados em uma epidemia (19, 24) ou ainda o grau de sincronização dos osciladores de Kuramoto (18, 25) e (b) o relacionamento das características dinâmicas com a topologia, em geral representada por medidas estruturais da rede (e.g. grau, betweenness, clustering, etc) (16, 26). A primeira abordagem possui a vantagem de permitir resultados analíticos através de métodos como teoria de campo médio (27) ou equação mestra (19, 28), mas possui o efeito indesejado de desconsiderar a possibilidade da existência de grupos dinâmicos na rede, isto é, grupos de nós em diferentes regiões da rede com dinâmicas nitidamente distintas, mas que em média apresentam o comportamento global esperado (29). A segunda abordagem requer a utilização de um conjunto de medidas para caracterizar a topologia, o que para muitos problemas se torna uma escolha arbitrária, como exemplo usamos o problema de medir a importância relativa de um nó na rede (30, 31), em que podemos utilizar a medida de grau, grau médio dos vizinhos, closeness (32), betweenness (33), centralidade de autovetor (4), k -shell (34) e muitas outras definidas na literatura (26), sendo que essas medidas em geral não são fortemente correlacionadas entre si como seria o esperado (34, 35), e portanto cada uma que for utilizada no estudo da dinâmica pode apresentar resultados distintos.

Outro aspecto que também é pouco considerado é a influência das condições iniciais no comportamento microscópico dos nós, ou seja, se mesmo para diferentes condições iniciais o sistema atinge as mesmas estatísticas globais (19), é possível que essas condições mudem drasticamente a dinâmica individual de alguns nós (36, 37), elevando ainda mais a necessidade de estudarmos o sistema a nível local. Como exemplo citamos:

Epidemia iniciando em várias regiões da rede: Qual a influência da condição inicial e da topologia para o tempo de primeira infecção de cada nó?

Ataque em redes: Os nós que são mais afetados tiveram o ataque ocorrendo muito próximo a eles ou o efeito foi devido a sua topologia?

Comportamento anômalo no córtex: Devido as largas flutuações dinâmicas desse sistema, um único experimento é insuficiente para diagnosticar sinais irregulares.

Estrutura variável: Redes cuja estrutura varia ao longo do tempo devem receber especial atenção. Podemos medir grupos possuindo diferentes variações da dinâmica ao longo de diversas realizações.

Nesse trabalho apresentamos uma metodologia que visa tratar os problemas observados através do uso de medidas dinâmicas específicas, tendo como principais objetivos:

- Estudar a dinâmica na rede a nível de nó
- Separar a influência da condição inicial e da topologia na dinâmica
- Estudar o papel da topologia sem a necessidade de medidas topológicas
- Revelar o grau de diferenciação dinâmico dos nós

Na realidade, a análise dinâmica a nível de nó está fortemente ligada com o estudo da influência da condição inicial no sistema. Para entendermos essa relação, recorreremos a Figura 1.1 que mostra a ideia geral do método. Realizamos uma única dinâmica sobre a rede (Figura 1.1a) que dá origem a um sinal dinâmico para cada nó (exemplificado na Figura 1.1b para os nós coloridos em vermelho e azul). Calculamos então medidas dinâmicas $f^1, f^2, \dots, f^n$ que buscam extrair as principais características desses sinais, de forma a criar um espaço de medidas (Figura 1.1c) contendo os nós a serem estudados. A Figura 1.1(c) mostra duas medidas dinâmicas hipotéticas, f^1 e f^2 , obtidas para dois nós distintos da rede (em azul e vermelho) em uma única realização da dinâmica. Neste gráfico vemos que os nós estão localizados em regiões distintas do espaço, essa distinção pode ter sido causada por quatro fatores:

1. Condição inicial
2. Topologia do nó

3. Dinâmica particular do nó

4. Estocasticidade da dinâmica

Os casos 1 e 2 já foram motivados neste capítulo, o caso 3 refere-se ao comportamento dinâmico individual de cada nó, isto é, se dois nós isolados possuem comportamentos dinâmicos distintos, ao estudarmos eles na rede precisamos levar em conta essa distinção, um exemplo básico seria a presença de indivíduos mais propensos que outros a receber uma doença na dinâmica de epidemia (38). O caso 4 refere-se a dinâmicas estocásticas (39), cuja evolução da dinâmica depende de probabilidades de transição entre estados, criando flutuações que influenciam na dispersão dos pontos no espaço de medidas. O fato é que *podemos facilmente isolar cada um dos casos, de forma a estudá-los separadamente*. No presente trabalho utilizamos dinâmicas determinísticas em que, em geral, todos os nós possuem isoladamente o mesmo comportamento dinâmico, com a finalidade de estudar apenas a influência da topologia e da condição inicial. Uma curta seção é utilizada para introduzir a influência de diferentes parâmetros dinâmicos.

Figura 1.1 – Visão geral do método proposto. (a) Uma dinâmica é aplicada sobre a rede, (b) o sinal obtido ao longo do tempo para cada nó é caracterizado por um conjunto de medidas dinâmicas (exemplificado na figura para duas medidas, f^1 e f^2 , aplicadas sobre os nós vermelho e azul), (c) após uma realização cada nó possui uma posição bem definida no espaço de medidas, ao exercitarmos diversas realizações podemos ter a formação de (d) grupos bem definidos ou (e) nenhuma distinção aparente.

Retornando a Figura 1.1(c), está claro que se queremos identificar o que potencialmente influenciou a dinâmica dos nós precisamos aplicar mais realizações da dinâmica sobre a rede, o caso ideal seria exercitar ao longo dessas realizações todas as condições iniciais possíveis que o sistema pode assumir, como isso é impraticável de se realizar optamos por utilizar condições iniciais escolhidas aleatoriamente de acordo com alguma regra, que será explicada na apresentação dos resultados. As novas realizações podem dar origem a dois casos distintos: (a) os pontos relativos a cada nó formam dois conjuntos bem definidos, caso que é mostrado na Figura 1.1(d), onde podemos ver que cada nó possui uma variação intrínseca causada pelas diferentes condições iniciais exercitadas, mas essa variação está restrita a determinadas regiões do gráfico caracterizadas pela topologia particular de cada nó, e portanto se queremos comparar os valores obtidos entre os nós devemos primeiro eliminar a influência da condição inicial. (b) O outro caso, que é mostrado na Figura 1.1(e) é quando as projeções dos dois nós se tornam indistinguíveis. A única conclusão que podemos chegar é que ambos os nós sentiram a topologia da mesma forma.

Com isso em mente criamos uma metodologia que busca responder as seguintes perguntas:

- Qual o efeito das condições iniciais na dispersão da dinâmica?
- A dinâmica de um determinado nó está sendo diferenciada pela topologia em relação aos demais nós da rede?
- De que forma grupos de dinâmicas semelhantes estão sendo formados?
- Quais são as condições para o surgimento de *outliers* dinâmicos (nós com dinâmica muito diferente)?

Para atacar esse problema a primeira abordagem que vem à mente é a de relacionar medidas topológicas e dinâmicas dos nós e verificar se diferentes topologias dão origens a diferentes comportamentos dinâmicos. Infelizmente essa metodologia requer a utilização de medidas topológicas apropriadas para o problema (26), pois certas medidas podem estar influenciando fortemente o comportamento, enquanto outras podem não ter relação alguma com a dinâmica,

ou ainda essas medidas podem ser impossíveis de serem obtidas, em que o exemplo mais estudado é o de redes corticais (40, 41). Pensando nisso nosso método tenta responder as perguntas propostas olhando apenas para a dinâmica dos nós, sem se preocupar com a topologia do sistema.

Adicionalmente apresentamos um mecanismo que busca obter de forma totalmente automatizada as medidas dinâmicas relevantes. A ideia é utilizar a análise por componentes principais (42) para projetar o espaço multidimensional da dinâmica em um novo espaço de duas ou três dimensões, com a finalidade de que esse novo espaço represente de maneira razoável as principais características da dinâmica. O único parâmetro de todo o processo é justamente o grau de representatividade da projeção, o que possibilita a sua aplicação para qualquer dinâmica que se deseja estudar, com a ressalva de que encontramos que a representação é baixa para dinâmicas pouco correlacionadas no tempo.

No Capítulo 2 apresentamos as ferramentas básicas que serão utilizadas ao longo da nossa metodologia. Essas ferramentas abrangem o estudo de estruturas, oferecendo mecanismos para a sua criação e caracterização, a definição de dinâmicas que serão aplicadas sobre essas estruturas e o estudo de aglomeração hierárquica, que busca encontrar grupos característicos em um conjunto de dados. O Capítulo 3 apresenta o método que utilizaremos para calcular a diferença dinâmica entre os nós, bem como a técnica de detecção automática de medidas. Os resultados são apresentados ao longo do Capítulo 4 e finalmente nossas considerações finais são descritas no Capítulo 5.

2 Materiais

2.1 Redes complexas (estrutura)

Como já observamos, existem muitos comportamentos na natureza que se desenvolvem sobre uma estrutura irregular, de difícil tratamento matemático, a essa estrutura geral damos o nome de grafo. Tradicionalmente atribui-se a Euler o primeiro trabalho na área de grafos (1736) (43), que surgiu devido a seguinte pergunta: É possível atravessar todas as sete pontes da cidade de Königsberg (antiga Prússia) passando por cada ponte apenas uma única vez? A ideia de Euler foi que os únicos elementos relevantes para o problema eram as terminações das pontes, e portanto todo o sistema poderia ser representado por um elemento abstrato chamado vértice interligado por arestas. Essa representação abstrata do sistema é que viria a ser a grande contribuição da teoria dos grafos, já que as mesmas ideias tratadas em uma área do conhecimento poderiam ser aplicadas em qualquer outra como física, matemática, computação, sistemas de informação, biologia e ciências sociais. Desde Euler, diversos novos problemas surgiram na área, como o problema do cavalo de xadrez (44), o estudo de árvores (45, 46), enumeração (47), coloração (48) e muitos outros (49). Mas todos esses problemas levavam a ideia do grafo como uma estrutura fixa, ou seja, estudava-se que características uma dada configuração de conexões possuía quando aplicadas em um problema, foi somente com Paul Erdős e Alfred Rényi em 1960 (15) que surgiu a visão probabilística e dinâmica de um grafo, dada pelo comportamento de não só uma realização de conexões, mas um conjunto delas (chamado ensemble) obedecendo uma determinada regra de criação.

2.1.1 Definições

Simbolicamente, representamos um grafo na forma de um par ordenado $G(V, E)$ que define um conjunto V de vértices (também chamados de nós) relacionados a um conjunto E de conexões (também conhecidas como arestas). Tipicamente representamos um grafo na forma de uma matriz de adjacência A na qual A_{ij} é o elemento da linha i e coluna j da matriz e $A_{ij} = 1$ se o vértice i está conectado ao vértice j ou $A_{ij} = 0$ caso contrário, portanto essa matriz possui dimensão $N \times N$ onde N é o número de nós do grafo. As arestas do grafo podem ser direcionadas, ou seja, uma certa quantidade que flui de um nó i para um nó j não pode fluir de j para i , nesse caso $A_{ij} \neq A_{ji}$, podemos ter ainda arestas que possuem peso, caso em que a matriz A não é mais binária e A_{ij} passa a representar o peso da conexão entre i e j . O grau k_i de um nó i é definido como o número de arestas que estão conectadas a esse nó, matematicamente

$$k_i = \sum_{j=1}^N A_{ij}. \quad (2.1)$$

Podemos também definir para uma rede direcionada o grau de saída k_i^{out} e o grau de entrada k_i^{in} na forma

$$k_i^{out} = \sum_{j=1}^N A_{ij} \quad (2.2)$$

$$k_i^{in} = \sum_{j=1}^N A_{ji} \quad (2.3)$$

onde definimos que $A_{ij} = 1$ se existe uma conexão saindo do nó i e chegando ao nó j .

Outro conceito importante é a distribuição de grau P_k de uma rede, que nada mais é do que a probabilidade de ao escolhermos um nó aleatório, esse nó possua um grau k , ou seja

$$P_k = \frac{\sum_i \delta_{k, k_i}}{N} \quad (2.4)$$

onde $\delta_{i,j} = 1$ se $i = j$ e $\delta_{i,j} = 0$ caso contrário, e k_i representa o grau do nó i . A definição para as distribuições de grau de entrada e saída é imediata.

2.1.2 Modelos de redes aleatórias

Modelo ER: O modelo proposto por Erdős e Rényi (15), conhecido como modelo ER e já citado anteriormente, é um caso específico de uma grande classe de grafos que possuem conexões aleatórias mas que obedecem uma determinada regra para a sua criação. No caso do modelo ER construímos uma rede onde os únicos parâmetros fixos são o número N de nós e o número m de arestas, onde distribuímos estas arestas uniformemente entre os $N(N-1)/2$ pares de nós possíveis. Uma visão alternativa e equivalente dessa construção é representada por $G(V, p)$ que ao invés de fixar o número de arestas, fixa a probabilidade de conexão entre cada par de nós, de forma que $\langle m \rangle = N(N-1) * p/2$ onde $\langle x \rangle$ simboliza o valor médio da variável aleatória X . No limite de N muito grande os dois modelos se tornam idênticos.

A distribuição de graus de uma rede ER é dada por (4)

$$P_k = \binom{N-1}{k} p^k (1-p)^{N-1-k} \quad (2.5)$$

mas devido ao fato de em geral utilizarmos redes esparsas, isto é, utilizamos $\langle k \rangle \ll N$, podemos fazer a clássica aproximação da equação 2.5 para uma distribuição de Poisson

$$P_k = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle} \quad (2.6)$$

Atualmente o modelo ER é, em geral, utilizado apenas como uma forma de comparação em relação a modelos mais robustos, pois a estrutura gerada pelo modelo não possui uma rica interpretação fenomenológica, sendo meramente uma estrutura gerada de maneira randômica. A trivialidade do modelo se reflete nas dinâmicas aplicadas sobre as redes geradas por ele, que em geral não apresentam um rico comportamento.

Redes livre de escala: A distribuição Gaussiana, que pode ser obtida a partir da equação 2.5 no limite contínuo (50), é largamente utilizada para explicar diversos fenômenos físicos, particularmente devido ao Teorema Central do Limite (51) que pode ser usado para definir um mecanismo que descreve o comportamento de variáveis macroscópicas a partir de interações microscópicas. Mas existe uma outra distribuição estatística muito comum na natureza (12), cuja forma é uma lei de potência $P_x \approx x^\gamma$, que do ponto de vista estatístico possui duas diferenças fundamentais em relação a Gaussiana: a) O decaimento da probabilidade de encontrar x muito grande não é mais exponencial, o que cria largas flutuações nos valores realizados de X (52, 53) e b) essa distribuição é livre de escala, ou seja, olhando em diferentes escalas de x a distribuição não varia (12). Do ponto de vista físico uma lei de potência pode representar o já citado fenômeno “o rico fica mais rico” no qual certos valores altos de x surgem simplesmente por terem alguma preferência ou aptidão natural (54). Barabási et al. (8) publicou um dos primeiros trabalhos sobre essa distribuição na área de redes, ele mostrou que dois simples mecanismos conseguem criar uma rede cuja distribuição de grau deve ser uma lei de potência. O primeiro é chamado de crescimento, no qual a rede começa com m_0 nós e a cada iteração é adicionado um novo nó que fará m_c novas conexões com a rede. O segundo conhecido como escolha preferencial define que cada uma das conexões do vértice adicionado se conecta com a rede de acordo com uma probabilidade proporcional ao grau de cada vértice já presente na rede. Esse processo dá origem a uma rede cuja distribuição de grau pode ser escrita na forma (8, 55, 56):

$$P_k \approx am_c^2 k^{-3} \quad (2.7)$$

onde a é uma constante. Portanto a rede gerada possui uma distribuição de grau em lei de potência com expoente muito próximo de 3 (costuma-se definir o expoente sem o sinal). Esse procedimento geralmente é chamado de modelo BA.

Redes geográficas: Redes no mundo real em geral estão inseridas em um contexto no qual cada nó possui uma posição espacial bem determinada (57), essa informação é desconsiderada

nos modelos clássicos apresentados até aqui, o fato é que dependendo de que conceito queremos estudar em um sistema é preciso considerar que arestas possuem um custo associado para conectar nós que estão muito distantes geograficamente. Essa restrição espacial possui um efeito dramático nas características estruturais das redes geradas (57), e portanto é de suma importância que ela seja analisada quando queremos fazer um paralelo com redes reais. Para se ter ideia do impacto que essa restrição pode causar, em 1998 Duncan Watts chegou a afirmar que uma rede geográfica não poderia possuir características de mundo pequeno (58), pois essa rede deveria possuir apenas uma única escala espacial, ao contrário da ideia de mundo pequeno que requer duas escalas (não necessariamente espaciais) intra e inter grupos (59). O fato é que Watts utilizou um modelo geográfico não muito apropriado para essa análise, e mais tarde surgiram novos modelos para tratar o problema (60, 61). Uma visão geral sobre os principais modelos geográficos atuais e algum tratamento analítico pode ser visto em (57). Nesse trabalho nós optamos por utilizar um dos modelos mais clássicos da área, conhecido como modelo de Waxman (62). Começamos com uma região espacial quadrada de largura unitária (1x1) e sorteamos aleatoriamente com probabilidade uniforme os valores no intervalo $[0,1]$ das posições x e y dos N nós dentro desse grid, então para cada par (i, j) de nós verificamos se eles serão conectados por uma aresta de acordo com uma probabilidade dada por

$$P(i, j) = \beta e^{-d(i,j)/d_0} \quad (2.8)$$

onde $d(i, j)$ é a distância espacial euclidiana entre os referidos nós, β controla a densidade de conexões e d_0 define o comprimento típico das conexões da rede. Podemos ver que a rede gerada deve possuir muitas conexões de curto alcance e ocasionais conexões de longo alcance, que aqui chamaremos de conexões globais.

2.1.3 Redes reais

Além dos modelos aleatórios, utilizaremos duas redes reais para motivar o tipo de informação que podemos obter ao utilizarmos o método, são elas:

Rede cortical: Essa rede descreve a conectividade em larga escala de um hemisfério do córtex do macaco do gênero *Macaca* da família *Cercopithecidae*, onde os vértices representam áreas corticais e as arestas correspondem a tratos córtico-corticais, que conectam essas diferentes áreas. A rede possui 95 regiões corticais mapeadas, sendo que o maior componente conectado é formado por 85 nós com um grau médio $\langle k \rangle = 27.7$, usaremos esse componente na análise. A rede foi obtida de (63). Em (64) essa rede é estudada utilizando as ferramentas tradicionais da área de redes complexas.

Rede neuronal: A segunda rede descreve o sistema neuronal do verme *Caenorhabditis elegans*, sendo que cada vértice representa um neurônio e ocorre uma aresta entre dois vértices se eles possuem alguma forma direta de comunicação. A base original, compilada por Chen et al. (65, 66), foi obtida de (67). Essa rede possui 279 nós e $\langle k \rangle = 22.4$.

2.1.4 Caracterização estrutural

Para melhor entendermos a relação entre a estrutura e função em redes, precisamos utilizar algumas medidas que buscam caracterizar a estrutura local ou global em que um nó está inserido. Dentre as várias medidas largamente utilizadas em redes (26), optamos por utilizar além do grau, k , as seguintes medidas:

Coeficiente de agrupamento local (clustering): Dados dois nós i e j que estão conectados a um nó l , o coeficiente de agrupamento de l , que daqui em diante chamaremos de clustering, fornece a probabilidade de que i e j estejam conectados entre si. Também podemos interpretá-lo como sendo a densidade de conexões existentes entre os vizinhos do nó l . O clustering C_i relativo ao nó i é calculado pela razão (4)

$$C_i = \frac{(\text{número de pares de nós vizinhos de } i \text{ que estão conectados})}{(\text{número de pares de nós vizinhos de } i)} \quad (2.9)$$

ou seja, percorremos todos os vizinhos do nó i contando quantos pares desses vizinhos estão conectados entre si, e dividimos o valor obtido pelo número total de conexões possíveis, dado por $k_i(k_i - 1)/2$, onde k_i é o grau do nó i .

Grau médio dos vizinhos: Essa medida quantifica a média de conexões que um nó faz não somente com os primeiros vizinhos, mas também com os vizinhos dos vizinhos, podendo ser utilizada para medir correlações de grau na rede (68). Sua definição para um determinado nó i , cujo símbolo denotaremos por $k_i^{(nm)}$, é simplesmente calcular a média dos graus dos vizinhos desse nó

$$k_i^{(nm)} = \frac{1}{k_i} \sum_{j \in \eta_i} k_j \quad (2.10)$$

onde η_i representa o conjunto de vizinhos do nó i .

Centralidade de autovetor: Suponha que um determinado nó possui um grau muito baixo, e por consequência é considerado um nó de baixa importância na estrutura da rede. O fato é que se esse nó estiver diretamente conectado a outro de grau muito elevado, pode ser que sua importância seja maior do que o esperado, talvez até maior do que outro nó de grau superior. A centralidade de autovetor busca quantificar essa característica. Seja q_i a pontuação (no sentido de importância) do nó i , definida como sendo proporcional a soma da pontuação dos seus vizinhos, ou seja,

$$q_i = \frac{1}{\lambda} \sum_{j=1}^N A_{ij} q_j \quad (2.11)$$

onde A_{ij} é o elemento da matriz de adjacência da rede e λ é uma constante de proporcionalidade. Reescrevemos a equação na forma vetorial, dada por (4)

$$A\vec{q} = \lambda\vec{q} \quad (2.12)$$

O autovetor, de dimensão N , associado com o maior autovalor dessa equação representa a centralidade de autovetor dos nós da rede.

2.2 Dinâmica

2.2.1 Integra-e-dispara

A tarefa de entender os fundamentos do processamento de informações pelo cérebro tem ganhado cada vez mais atenção no meio científico. Teoricamente o problema costuma ser abordado em duas frentes: 1 - descrição fenomenológica dos neurônios, de forma a criar modelos matemáticos (com ou sem intuição biológica) que reproduzam características eletrofisiológicas observadas experimentalmente *in vivo* ou *in vitro*, como adaptação, bursting, atraso na transmissão sináptica, refratoriedade e outros (69, 70), 2 - descrição do ponto de vista informativo (71), ou seja, do código neuronal utilizado para o recebimento e transmissão da informação entre neurônios. Tradicionalmente a descrição fenomenológica é aplicada em uma única célula neuronal, de forma a, dado um estímulo, reproduzir a resposta medida em um experimento prévio. Modelos como Hodgkin-Huxley (72), Hindmarsh-Rose (73), Morris-Lecar (74) e FitzHugh-Nagumo (75) proporcionam descrições de quase todos os comportamentos neuronais observados, mas são extremamente caros de serem implementados e geralmente possuem uma grande quantidade de parâmetros, o que os tornam difíceis de serem analisados. Já modelos formais como *integra e dispara* (76, 77), *integra e dispara não linear* (78) e *kernel de resposta a disparos* (tradução livre de *spike response model*) (70, 79) são utilizados para descrever os comportamentos mais comuns de neurônios do ponto de vista informativo, como taxas, correlações de intervalos entre disparos, informação condicional e comportamento coletivo (70).

Nosso propósito é estudar o papel exercido pela topologia de sinapses ao levarmos em conta um grande número de elementos excitáveis (também denominados *acoplados por disparos*) quando descritos pelo modelo *integra e dispara*, que apesar de simples proporciona uma boa descrição fenomenológica e riqueza de informação quando estudado em rede. A equação

mais geral de um nó i executando uma dinâmica integra-e-dispara é dada por (70)

$$\tau_m \frac{dV_i}{dt} = -(V_i - E_r) + RI^{ext} + \sum_{j=1}^N \sum_f G_{ij} \alpha(t - t_j^f - t_\Delta) \quad (2.13)$$

onde V_i representa o potencial da membrana do neurônio, τ_m a escala temporal e E_r o potencial de repouso quando não há estímulos. I^{ext} em geral modela a presença de correntes externas inseridas no sistema, que podem ser entendidas como estímulos captados por neurônios sensoriais ou ruídos advindos de outras regiões do córtex (80). O duplo somatório representa os sinais pós sinápticos dos vizinhos de i , onde G é a matriz de adjacência da rede com pesos nas arestas, t_j^f é o f -ésimo instante de disparo do nó j e t_Δ o delay entre o disparo do vizinho e sua chegada ao nó i . $\alpha(t)$ é a função resposta do neurônio, isto é, como o potencial de i se comporta dado que ele recebeu um sinal de j . Quando o potencial do neurônio excede um determinado valor T_l , ele emite um disparo para cada vizinho e salta para um estado hiperpolarizado, podendo passar por um período refratário no qual o seu potencial é muito difícil de ser alterado.

Em geral o comportamento coletivo da rede é analisado através do parâmetro $R(t)$ (usualmente chamado de ativação, ou atividade, da rede) definido como (71)

$$R(t) = \frac{1}{N} \sum_{i=1}^N s_i(t) \quad (2.14)$$

onde $s_i(t) = 1$ se o nó i disparou no instante t e $s_i(t) = 0$ caso contrário, ou seja, $R(t)$ fornece a fração de nós da rede que estão disparando a cada instante de tempo. Wilson e Cowan (22) foram um dos primeiros a fornecer uma análise formal de uma população de neurônios interagentes, desde então um grande número de trabalhos surgiram a fim de tratar outros aspectos da dinâmica em população como ruído (81–83) e posições geográficas (70, 84). Mais recentemente, utilizando como base os novos modelos de redes da última década vários resultados vem sendo obtidos, é comum entre esses trabalhos estudar três aspectos da dinâmica: a) a existência de estados auto sustentados (23, 85), b) estados síncronos e assíncronos (70, 86, 87) e c) avalanches (23, 88). Estados auto sustentados ocorrem quando na ausência de estímulos externos a dinâmica nunca entra no estado absorvente, isto é, $R(t) \neq 0$ para qualquer t , esse

será um dos pontos de estudo do nosso trabalho. Podemos também estudar as condições para a existência de estados síncronos e assíncronos em diversas topologias, o primeiro ocorre quando os neurônios da rede de alguma forma passam a disparar de forma sincronizada, fazendo com que R oscile ao longo do tempo, Strogatz et al. (89) formalizou de maneira simples e elegante as condições necessárias para tal comportamento. O estado assíncrono, como era de se esperar, ocorre quando os neurônios estão completamente desincronizados (70), fazendo com que R tenda a uma constante em um tempo longo. O estudo de avalanches ocorre em duas frentes bem distintas, a primeira vem da ideia de criticalidade auto organizada que desde um paper publicado em 1987 por Bak et al. (11) para tentar explicar o chamado ruído "1/f" vem fazendo muito sucesso. No contexto de redes uma ideia é estudar o tempo que a dinâmica leva para entrar no estado absorvente quando próxima da região de transição, foi mostrado em (23) que a distribuição de probabilidades desse tempo segue uma lei de potência para uma rede de mundo pequeno. A outra frente de estudos ocorre quando inserimos um estímulo externo em um único nó da rede (88), inicialmente observamos uma fase transiente na qual $R \approx 0$, ou seja, os nós disparando estão contidos em uma região no entorno do nó que recebe o estímulo. Subitamente a dinâmica se espalha por toda a rede, passando por um período de rápida transição de R , esse período caracteriza uma avalanche de ativação que pode ser estudada em função de diversos parâmetros dinâmicos e topológicos (88).

Por mais que estejamos interessados nos diversos comportamentos que a dinâmica integra-e-dispara consegue exibir, nosso objetivo nesse trabalho não é estudar essa dinâmica, e sim utilizá-la como ferramenta para ilustrar o método proposto. Para tanto utilizaremos uma versão simplificada do modelo, na qual não há estímulos externos e os parâmetros dinâmicos foram todos tomados na sua forma mais simples. Com isso escrevemos a equação dinâmica em uma forma discreta:

$$V_i^{t+1} = \begin{cases} V_i^t + \sum_{j=1}^N \sum_f A_{ij} \delta(t - t_j^f) & \text{se } V_i^t < T_i \\ \sum_{j=1}^N \sum_f A_{ij} \delta(t - t_j^f) & \text{se } V_i^t \geq T_i \end{cases} \quad (2.15)$$

onde V_i^t é o estado do nó i na iteração t e t_j^f o instante do f -ésimo disparo do nó j . O

sinal binário resultante, $s_i(t)$, que como já dito descreve os disparos do nó i , possui uma forte semelhança com o tradicional modelo de McCulloch e Pitts (90), com a diferença de que no nosso modelo o neurônio faz o papel de um somador ao longo do tempo, ao passo que um neurônio McCulloch e Pitts é um somador instantâneo. Há também a diferença fundamental de que no nosso caso não há conexões inibitórias. Versões discretas de modelos neuronais também costumam ser analisadas no contexto de mapas de evolução (map based models em inglês) (91).

2.2.2 Kuramoto

Em 1967 Arthur Winfree (92) introduziu o estudo de osciladores fracamente acoplados, que além de serem auto-sustentados eram estáveis a pequenas perturbações, isso possibilitou o estudo de osciladores não lineares através de uma redução das variáveis que precisavam ser consideradas no problema. Suas ideias fizeram surgir um grande número de trabalhos sobre o assunto (25, 93), sendo o mais importante deles muito provavelmente o modelo de Yoshiki Kuramoto (94), que definiu uma dinâmica de Winfree simplificada, descrita pela equação

$$\dot{\theta}_i = \omega_i + \frac{\lambda}{N} \sum_{j=1}^N \sin(\theta_j - \theta_i) \quad (2.16)$$

onde θ_i e ω_i são, respectivamente, a fase e a frequência natural do oscilador i , θ_j a fase do vizinho de i e λ a constante de acoplamento geral do sistema, que faz o papel de um parâmetro de ordem no qual quando $\lambda > \lambda_c$ os osciladores entram num estado sincronizado (28). Kuramoto mostrou que podemos usar as fases dos muitos osciladores para definir o centroide do sistema escrito na forma

$$R_k e^{i\psi} = \frac{1}{N} \sum_{j=1}^N e^{i\theta_j} \quad (2.17)$$

onde R_k é chamado de coerência de fase, $R_k = 1$ significa que todos os osciladores possuem a mesma fase e o sistema se comporta como um grande oscilador e $R_k = 0$ indica que não há ordem alguma, as fases dos osciladores estão distribuídas uniformemente. No limite $N \rightarrow \infty$, R_k

em função de λ sofre uma transição abrupta do estado incoerente para o sincronizado (18).

Apesar do modelo original de Kuramoto ser definido para um grafo completo, podemos escrever a equação 2.17 para um grafo qualquer na forma (18)

$$\dot{\theta}_i = \omega_i + \frac{\lambda}{\langle k \rangle} \sum_{j=1}^N A_{ij} \sin(\theta_j - \theta_i) \quad (2.18)$$

Existem inúmeros trabalhos na literatura sobre o modelo de Kuramoto aplicado nos mais diversos modelos de rede, com destaque para redes ER, de mundo pequeno, e redes livre de escala. Uma excelente revisão dessas aplicações, inclusive em redes reais, foi feita em (18). Um resultado recente obtido por Arenas et al. (95), mostrando a ocorrência de sincronização explosiva para uma distribuição de frequências proporcional ao grau de redes livre de escala promete continuar a tradição de resultados interessantes nessa dinâmica. Na maioria desses trabalhos é comum sortear aleatoriamente as fases iniciais, θ_0 , e frequências naturais, ω , dos osciladores (18), optamos por sortear as frequências de acordo com uma distribuição uniforme com valores no intervalo $[-0.5, 0.5]$, já as fases iniciais são sorteadas também com uma distribuição uniforme mas no intervalo $[-\pi, \pi]$.

2.2.3 Medidas dinâmicas

A seguir descrevemos brevemente as medidas dinâmicas que usaremos ao longo do trabalho. Essas medidas buscam caracterizar o sinal dos neurônios do ponto de vista informativo. Um excelente estudo sobre caracterização de sinais neuronais isolados é feito em (96).

Taxa: Dado um sinal binário $s_i(t)$ de um neurônio i do sistema, representando os disparos do mesmo ao longo da simulação (ou experimento), definimos a sua taxa média, r_i na forma (71)

$$r_i = \frac{1}{T_{sim} - T_{est}} \sum_{t=T_{est}}^{T_{sim}} s_i(t) \quad (2.19)$$

onde T_{sim} é o tempo de simulação e T_{est} é o tempo empírico no qual observamos que a partir dele as estatísticas globais do sistema se tornam praticamente constantes (veja apêndice A). Essa medida nos diz a fração de vezes que o neurônio i disparou no período observado, o que é uma estatística muito utilizada na análise de dinâmicas neuronais, pois acredita-se que boa parte da informação do sinal de um neurônio está contida na sua taxa de disparos (71). Observe também a diferença entre r_i e $R(t)$ definido na equação 2.14, o primeiro é uma média temporal individual de cada nó, o segundo uma média sobre a população em um dado instante de tempo.

Entropia de intervalos: Além da taxa, sinais neuronais são conhecidos por conter informação nos intervalos entre dois disparos consecutivos do neurônio, se definimos como Δ a variável aleatória associada a esses intervalos, escrevemos a entropia de intervalos ISE_i na forma

$$ISE_i = - \sum_{\Delta} P(\Delta) \log(P(\Delta)), \quad (2.20)$$

onde $P(\Delta)$ é a probabilidade de encontrarmos um intervalo Δ entre os disparos do nó i . Essa medida quantifica a complexidade do sinal analisado, isto é, quanto de informação precisamos para caracterizar esse sinal (veja (97) para uma excelente abordagem em teoria da informação). Observamos que essa entropia é relacionada com a entropia de blocos utilizada em sinais contínuos (98).

Intervalo máximo: Outra medida que será útil ao longo do trabalho é o maior intervalo de tempo, I_i^0 que um sinal binário fica seguidamente com o valor zero, que pode ser entendido como o maior período que um nó fica inativo na dinâmica. Seja $\{\Delta\}$ o conjunto de valores de intervalo realizados para um sinal $s_i(t)$, definimos

$$I_i^0 = \max\{\Delta\} \quad (2.21)$$

Frequência efetiva: Essa medida utilizaremos na dinâmica de Kuramoto. Cada oscilador da dinâmica possui uma frequência natural de oscilação, mas devido ao efeito dos vizinhos o oscilador passa a ter uma dinâmica diferenciada na qual não existe mais uma frequência característica. No entanto consideramos que quando o parâmetro R_k converge a um valor apro-

ximadamente constante a dinâmica em cada oscilador passa a ser aproximadamente periódica, e portanto definimos a frequência efetiva média relativa ao nó i na forma (99):

$$\tilde{\omega}_i = \frac{1}{T_{sim} - T_{est}} (\theta_i(T_{sim}) - \theta_i(T_{est})) \quad (2.22)$$

onde fazemos T_{sim} tão grande quanto possível e T_{est} é um tempo longo o suficiente tal que a partir dele o parâmetro R_k seja praticamente constante.

2.3 Distâncias

Para aplicarmos nosso método precisamos definir uma distância dinâmica entre os sinais projetados dos nós, nesta seção apresentamos as distâncias que serão utilizadas. Inicialmente suponha que temos um conjunto de pontos inseridos em um espaço de m dimensões, e esses pontos formam dois grupos distintos nesse espaço (veja a Figura 2.2(a) para um exemplo bidimensional). Se quisermos obter a distância entre esses dois grupos, a primeira ideia que vem a mente é a de calcular o centro de massa (ou centroide) de cada grupo e obter a distância euclidiana entre eles, dada por

$$d(r, s) = \sqrt{\sum_i^m (\langle x_{ri} \rangle - \langle x_{si} \rangle)^2} \quad (2.23)$$

onde r e s são os índices de cada grupo e $\langle x_{ri} \rangle$ representa a média dos valores dos pontos do grupo r na i -ésima dimensão. A desvantagem da distância euclidiana é que ela não leva em conta um detalhe fundamental do problema, que é a dispersão dos grupos cuja distância está sendo medida. Suponha que temos duas variáveis aleatórias unidimensionais, X_r e X_s , das quais queremos obter a distância entre os valores realizados x_r e x_s . Na Figura 2.1 mostramos o problema de utilizar apenas a diferença entre a média dessas variáveis, suponha que os valores realizados de X_r possuem a distribuição de probabilidades em azul e os de X_s a distribuição em vermelho, claramente a distância entre as médias de X_r e X_s possui uma maior *significância* na

Figura 2.1(a) do que na Figura 2.1(b), apesar de em ambas as figuras o valor dessa distância ser o mesmo. Portanto uma distância mais honesta entre os valores dos grupos r e s deve levar em conta a dispersão interna de cada grupo, dentre as várias definições presentes na literatura optamos por utilizar a distância definida por Student (cujo nome verdadeiro é Gosset) (100, 101) dada por

$$t = \frac{\langle x_r \rangle - \langle x_s \rangle}{S_{x_r - x_s}} \quad (2.24)$$

onde lembramos que x é unidimensional e $S_{x_r - x_s}$ é uma estimativa do desvio padrão equivalente de cada grupo, dado por

$$S_{x_r - x_s}^2 = \frac{\langle (x_r - \langle x_r \rangle)^2 \rangle}{n_r} + \frac{\langle (x_s - \langle x_s \rangle)^2 \rangle}{n_s} \quad (2.25)$$

em que n_r e n_s são, respectivamente, o número de valores realizados de X_r e X_s .

Figura 2.1 – Densidade de probabilidade de duas variáveis aleatórias com (a) pouca variância e (b) larga variância.

No caso multidimensional temos uma nova propriedade que devemos levar em conta. Na Figura 2.2 mostramos três exemplos de distribuições em duas dimensões, onde está claro que

a significância da distância entre as médias dos grupos na Figura 2.2(a) e (b) é praticamente a mesma, enquanto que na Figura 2.2(c) ela é bem menor. Isso quer dizer que no caso multidimensional não basta levarmos em conta as variâncias em cada eixo, temo que considerar a variância *na direção* definida pela linha que passa pelas duas médias. Uma forma de considerar esse efeito é utilizando a estatística de Hotelling (102) definida como

$$h^2 = \frac{n_r n_s}{n_r + n_s} (\langle \vec{x}_r \rangle - \langle \vec{x}_s \rangle)' \Sigma_m^{-1} (\langle \vec{x}_r \rangle - \langle \vec{x}_s \rangle) \quad (2.26)$$

onde $\langle \vec{x}_r \rangle$ define a posição do ponto médio do grupo r e Σ_m é a estimativa da matriz de covariância equivalente de cada grupo, dada por

$$\Sigma_m = \frac{n_r \Sigma_r + n_s \Sigma_s}{n_r + n_s - 2} \quad (2.27)$$

sendo que Σ_r e Σ_s são, respectivamente, a matriz de covariância do grupo r e s . Os valores de h calculados para os pontos da Figura 2.2 estão mostrados nos respectivos gráficos.

Figura 2.2 – Exemplo de significância estatística, nos três gráficos a distância euclidiana entre a média dos grupos é sempre a mesma, mas os casos (a) e (b) possuem uma distância mais significativa do que o caso (c).

Ao utilizarmos mais de uma variável aleatória, é comum normalizarmos seus valores para que possamos compará-las na mesma escala, fazemos isso através da expressão (100)

$$\hat{x} = \frac{x - \langle x \rangle}{\sigma} \quad (2.28)$$

onde σ é o desvio padrão dado por

$$\sigma = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}$$

Sempre que utilizarmos mais de uma variável na nossa análise adotaremos a versão normalizada (ou estandardizada) da mesma.

A distância que usamos nesse trabalho na verdade é uma versão híbrida entre a euclidiana e as distâncias de Student ou Hotelling. Essa modificação foi necessária porque observamos que a dinâmica de alguns nós da rede não variou nas diversas condições iniciais exercitadas, fazendo com que a variância dos pontos relacionados aos nós fosse nula, e consequentemente as distâncias estatísticas divergiram. Como não estamos interessados no valor absoluto da medida, e sim na comparação entre os valores para diferentes grupos, definimos a versão híbrida na forma

$$\tilde{d}_x(r, s) = \frac{x}{x+1} d(r, s) \quad (2.29)$$

onde $d(r, s)$ é a distância euclidiana usual e x pode ser a distância de Student ou Hotelling.

2.4 Aglomeração hierárquica

Suponha que possuímos N elementos que precisam ser divididos em C classes, $w_1, w_2, \dots, w_C$, distintas. Esses elementos podem ser objetos físicos (e.g. bola, banana, lápis), sinais medidos (e.g. onda sonora, disparos de neurônios, batimentos do coração) ou qualquer outra instância em que é possível obter um conjunto de medidas sobre eles. A esses elementos damos o nome de amostras. Em relação às classes, elas podem ser qualquer tipo de divisão que queremos realizar sobre os elementos, como separar dois tipos de frutas ou classificar um coração entre

saudável ou doente. Para tanto precisamos aplicar um conjunto de medidas $f^1, f^2, \dots, f^k$ em cada amostra de forma a inseri-la em um espaço de medidas, no qual cada dimensão está relacionada com uma das f medidas utilizadas, ou seja, cada amostra passa a ser caracterizada por um vetor de medidas. Tendo essa representação dos dados existem dezenas de métodos que podem ser utilizados para classificar as amostras (103), no nosso caso sabemos que os grupos que estamos procurando podem formar subclasses, e essas subclasses podem possuir outras subclasses e assim por diante, para obtermos toda a informação dessas várias classes presentes no problema utilizaremos a aglomeração hierárquica, que é um método particular do estudo mais geral de técnicas de clusterização não supervisionada (103, 104).

O método de aglomeração consiste em inicialmente setar que os N elementos a serem analisados estão em N classes distintas, então de acordo com algum critério de semelhança unimos as duas amostras mais similares da base, formando um grupo de 2 elementos e obtendo $N - 1$ grupos no total. Repetimos o procedimento para obter $N - 2$ grupos e assim por diante, até que seja formado apenas um único grupo contendo todos os elementos. Com a informação das distâncias entres os grupos em cada escala construímos uma representação chamada de dendrograma (veja Figura 2.3b), com a qual é possível verificar a escala ideal para separar as amostras. Um critério comum é utilizar a partição (ou corte), anterior ao maior intervalo de variação da similaridade sem que houvesse um novo grupo formado, fato que indica que foi preciso mudar muito a escala do agrupamento para juntar os grupos formados.

Ao unir as amostras (ou grupos) podemos utilizar diversos critérios para escolher a união mais indicada, estes critérios por sua vez podem ser aplicados nas mais diversas medidas de distância entre os dados (103, 104). As medidas de distância que utilizaremos serão as definidas na seção anterior, já o critério de união poderá ser a distância média ou o método de Ward (104, 105). No caso da distância média, seja $\hat{d}(\vec{x}_{ri}, \vec{x}_{sj})$ a distância entre o vetor de medidas do i -ésimo elemento da classe r e o j -ésimo elemento da classe s , definimos que a distância entre as classes r e s é dada pela média das distância entre os elementos em cada classe, escrita na forma

Figura 2.3 – Exemplo de aglomeração de dados artificiais. (a) 6 amostras são inseridas em um espaço bidimensional através de duas medidas f^1 e f^2 , (b) aplicando um método de aglomeração obtemos o dendrograma dos dados. Como a distância necessária para unir os dois últimos grupos é muito maior do que em qualquer outra união, a melhor divisão das amostras é em dois grupos (c).

$$d_m(r, s) = \frac{1}{n_r n_s} \sum_i^{n_r} \sum_j^{n_s} \hat{d}(\vec{x}_{ri}, \vec{x}_{sj}) \quad (2.30)$$

onde n_r e n_i representam, respectivamente, o número de amostras nas classes r e s .

O outro critério de agrupamento, conhecido como método de Ward (105, 106), tenta garantir que a cada aglomeração de 2 grupos a dispersão dos grupos resultantes seja a mínima possível. A dispersão de um dado grupo, r , é definida como sendo a soma das distâncias euclidianas entre cada amostra, x_{ri} , do grupo e o centro de massa, k_r , do referido grupo. A dispersão geral de uma dada partição, G , é então definida como a soma da dispersão de todos os grupos, dada por

$$W(G) = \sum_{r=1}^C \sum_{i \in G_r} \hat{d}(\vec{x}_{ri}, \vec{k}_r) \quad (2.31)$$

onde G_r representa as amostras pertencentes ao grupo r . O critério de Ward diz que ao unirmos as classes s e r para obter a partição $G_{s \cup r}$, a diferença

$$W(G_{s \cup r}) - W(G) = \frac{n_s n_r}{n_s + n_r} \hat{d}(\vec{k}_s, \vec{k}_r) \quad (2.32)$$

deve ser a mínima possível (105).

2.5 Análise por componentes principais

A análise por componentes principais (PCA) é considerada uma das técnicas de análise multivariada mais antigas que se tem registro. Introduzida por Pearson em 1901 (107) e melhor fundamentada por Hotelling em 1933 (108), essa técnica busca reduzir a dimensionalidade de um conjunto de dados através da eliminação da redundância entre as variáveis utilizadas. Para tanto projetamos as variáveis originais em uma base conveniente gerando novas variáveis que preservam ao máximo a variância dos dados, ou seja, dado um vetor de p medidas $\vec{x}$ que caracteriza um objeto, queremos encontrar $\vec{\eta}_1$ tal que $\text{var}[\vec{\eta}_1 \vec{x}]$ seja máxima. Em seguida precisamos

encontrar $\vec{\eta}_2$ descorrelacionado com $\vec{\eta}_1$ cuja variância também seja máxima, e assim por diante, até obtermos m componentes principais, onde deseja-se que $m \ll p$ de tal forma que a dimensão do problema é reduzida de p para m . O que Pearson e, independentemente, Hotteling encontraram é que a base $\vec{\eta}_1, \vec{\eta}_2, \dots, \vec{\eta}_m$ que devemos utilizar para projetarmos os dados com máxima variância é dada pela equação (42)

$$\Sigma \vec{\eta} = \lambda \vec{\eta} \quad (2.33)$$

onde Σ é a matriz de covariância obtida em diversas observações de $\vec{x}$. A equação 2.33 é uma equação típica de autovetores e autovalores, onde foi provado que λ representa a variância de $\text{var}[\vec{\eta}\vec{x}]$. Portanto se ordenarmos os autovetores $\vec{\eta}_k$ de tal forma que os correspondentes autovalores sigam $\lambda_1 > \lambda_2 > \dots > \lambda_m$ podemos escrever (42)

$$\vec{z} = E' \vec{x} \quad (2.34)$$

onde E é uma matriz na qual a coluna k é exatamente o autovetor $\vec{\eta}_k$, e E' representa a transposta dessa matriz.

Para termos uma interpretação visual do PCA, suponha que cada elemento de uma base de dados é caracterizado por duas medidas f^1 e f^2 , e que fazendo o gráfico de $f^1 \times f^2$ obtemos a Figura 2.4. Claramente f^1 e f^2 estão correlacionados, e portanto podemos definir novos eixos z^1 e z^2 de tal forma que nesses eixos a correlação entre os dados é mínima, os novos eixos são basicamente uma versão transladada e rotacionada dos eixos antigos, sendo que a transformação é dada pela matriz E definida acima.

Figura 2.4 – Exemplo da aplicação do PCA na redução de dimensionalidade de dados.

3 Métodos

3.1 Discriminando a dinâmica entre os nós

Para aplicar nossa metodologia começamos com uma rede de N nós (que será a mesma durante todo o processo) e executamos $\mathfrak{R}$ vezes uma dada dinâmica sobre essa rede, cada execução começando de uma condição inicial distinta. O ideal seria que aplicássemos todas as combinações possíveis de condições iniciais, mas como esse processo seria extremamente custoso temos que abordar o problema através do método de Monte Carlo, ou seja, para cada realização sorteamos aleatoriamente o valor inicial da dinâmica de cada nó. É importante notar que dependendo da dinâmica podemos ter um conjunto muito pequeno de condições iniciais que levem o sistema a um estado particular, de forma que a amostragem aleatória não consiga acessar esses casos, isto não é um problema para o método, pois ao utilizarmos a técnica de amostragem na verdade estamos analisando a dinâmica para o conjunto de condições iniciais impostas, isto é, estamos estudando a influência da topologia nos sinais que são de fato observados.

Tendo os sinais dinâmicos de cada nó para as $\mathfrak{R}$ realizações, precisamos de um modo de representá-los de forma a realizar a comparação entre eles. No nosso caso utilizaremos medidas dinâmicas que busquem capturar as características mais relevantes de cada sinal. Seja F uma dessas medidas, obtemos através das várias realizações da dinâmica um conjunto de valores observados $f_{i,r}$, onde i é o índice do nó e r a realização. Esses valores podem variar devido a quatro efeitos distintos que já foram citamos na introdução desse trabalho: (a) condição inicial, (b) topologia da rede, (c) dinâmica particular do nó e (d) estocasticidade da dinâmica. Nosso objetivo inicial é estudar apenas a relação entre a influência da condição inicial e da topologia, portanto utilizamos uma dinâmica não estocástica em que a equação dinâmica é idêntica para todos os nós, com isso os dois únicos efeitos que podemos observar nos valores de F são: (a)

flutuações nos valores da medida para um mesmo nó (f_i), que *devido ao fato da topologia ser estática só pode ser causado pelas variações na condição inicial*, essa variação da origem a um conjunto de valores mais prováveis de F , cuja média é geralmente utilizada para relacionar a dinâmica com a topologia da rede; (b) diferenças nos valores mais prováveis de F entre nós distintos da rede, que *devido as propriedades dinâmicas assumidas somente pode ser causado pela diferença topológica entre os nós*. O termo *diferença topológica* deve ser usado com cautela, pois a não ser em casos muito específicos em que a rede é perfeitamente simétrica (e.g. lattice com condição de contorno toroidal) a topologia de dois nós nunca é idêntica, isto é, sempre podemos encontrar uma medida estrutural que possua valores distintos para diferentes nós. Nas redes que consideraremos a topologia sempre será diferente, e portanto qualquer diferença significativa na dinâmica entre dois nós em várias realizações deve ser causada pela topologia. É importante notar que o contrário não pode ser afirmado, se a dinâmica entre dois nós aparenta ser a mesma, a topologia deles continua sendo distinta, o que aconteceu é que ambos *sentiram* a topologia da mesma forma. Um exemplo desse último caso seria uma dinâmica de passeio aleatório simples (109, 110) que depende apenas do grau de cada nó, após um tempo muito longo nós de mesmo grau adquirem a mesma característica dinâmica (chamada de atividade), ou seja, apesar de possuírem topologia geral distintas o comportamento local da dinâmica faz com que apenas o grau diferencie os nós. O fato é que para analisarmos esse caso (mesma dinâmica entre os nós) temos que recorrer a medidas estruturais.

Para quantificar a diferença dos valores obtidos para cada nó levando em conta a dispersão entre as realizações, recorreremos a medida de distância definida na equação 2.29. Através dela podemos verificar se a diferença observada entre a média da medida dinâmica F para dois nós é de fato significativa, ou seja, estamos calculando a diferença entre a dinâmica dos dois nós normalizada pela variação do efeito provocado pela condição inicial. De forma a obter o máximo de informação das diferenças dinâmicas, construímos uma matriz Ξ onde a linha i e coluna j representa a distância obtida entre os nós i e j da rede, ou seja,

$$\Xi_{ij} = \tilde{d}_x(i, j) \quad (3.1)$$

onde $\tilde{d}_x(i, j)$ é definido pela equação 2.29. Finalmente definimos nossa medida de distância dinâmica média, α , como sendo a média de cada linha dessa matriz, ou seja

$$\alpha_i = \frac{1}{N} \sum_{j=1}^N \Xi_{ij} \quad (3.2)$$

O procedimento padrão nesse momento seria calcular a significância estatística do valor observado (no mesmo molde utilizado nos testes de Student e Hotelling (111)), mas o fato é que estamos interessados na comparação das distância entre os nós, e não no seu valor absoluto. Para efetuar a comparação construímos um histograma dos valores obtidos de α , lembrando que α é relativo a uma medida dinâmica, e portanto para cada medida utilizada teremos um histograma distinto de α . Nosso objetivo com esses histogramas é responder as perguntas propostas no início do trabalho:

- Qual o efeito das condições iniciais na dispersão da dinâmica?
- A dinâmica de um determinado nó está sendo diferenciada pela topologia em relação aos demais nós da rede?
- De que forma grupos de dinâmicas semelhantes estão sendo formados?
- Quais são as condições para o surgimento de *outliers* dinâmicos (nós com dinâmica muito diferente)?

O que procuraremos é por comportamentos peculiares dos histogramas que caracterizem cada um dos casos propostos, por exemplo, é esperado que um nó com dinâmica muito diferente dos demais possua α muito grande, de forma que o histograma mostre essa diferenciação. É importante salientar que apesar do método ter sido apresentado utilizando apenas uma medida para α , nada nos impede de calcularmos as distâncias utilizando simultaneamente diversas medidas dinâmicas. Para tanto utilizamos a estatística de Hotelling, que generaliza a distância de Student para várias dimensões. A Figura 3.1 mostra uma aplicação real do modelo apresentado.

Apesar de termos tirado a média da matriz Ξ , nada nos impede de utilizar essa matriz como entrada de um algoritmo de aglomeração (seção 2.4), pois ela nada mais é do que uma rela-

Figura 3.1 – Exemplo de aplicação do método. (a) Executamos 12 realizações de uma dinâmica na rede e (b) projetamos os sinais de cada nó em um espaço definido por duas medidas dinâmicas. (c) Obtemos a matriz de distâncias estatísticas entre os pontos realizados de cada nó, (d) a média de cada linha da matriz da origem aos valores de α . (e) Por fim construímos o histograma de α .

ção de distâncias entre os nós, com ela poderíamos encontrar os diferentes grupos dinâmicos presentes na rede. Mas nós optamos por fazer um processo um pouco diferente, utilizamos a distância euclidiana entre as realizações de cada nó e recorremos ao método de Ward (106) para produzir a aglomeração. Como o método cria os grupos tentando maximizar a relação entre a covariância inter e intra grupos, o resultado acaba sendo análogo ao de utilizar a matriz Ξ , e por este ser uma forma tradicional de aglomeração optamos por utilizá-lo. Para explicar um pouco melhor o processo, seja $\vec{x}_i$ o vetor relativo ao nó i , no qual cada elemento contém o valor da medida dinâmica aplicada sobre uma realização, e portanto $\vec{x}_i$ possui dimensão $\mathfrak{R}$, a distância Euclideana entre $\vec{x}_i$ e $\vec{x}_j$ é obtida na forma usual, de forma a criarmos novamente uma matriz de distâncias que será usada na aglomeração de Ward. Como explicado na seção 2.4 o resultado obtido é um dendrograma descrevendo como ocorreu a junção entre cada grupo de nós da rede, de forma que podemos definir um corte no dendrograma para obter um determinado número de grupos de dinâmica semelhante, ou seja, os nós que terminam em um mesmo grupo tiveram a medida dinâmica muito semelhante para todas as condições iniciais utilizadas. Esses grupos fornecem valiosas informações sobre a influência topológica, podemos encontrar um grupo gigante indicando que a topologia não consegue diferenciar a dinâmica dos nós, ou diversos grupos pequenos sugerindo forte influência topológica local, esses grupos podem inclusive indicar a iminência de uma transição na dinâmica, ou seja, o comportamento dos nós está a ponto de se tornar homogêneo para toda a rede com uma pequena alteração de algum parâmetro dinâmico.

3.2 Encontrando medidas dinâmicas relevantes com o PCA

Tendo em mente o potencial do PCA em encontrar as variáveis relevantes do sistema, quando falamos em dinâmicas em redes uma abordagem praticada na literatura (42) é a de tratar cada nó i como um vetor $\vec{x}^i$, sendo que cada elemento t de $\vec{x}^i$, x_t^i , é o valor da dinâmica do nó i no instante de tempo t . Com isso se aplicamos o PCA nos dados estamos tentando representar a dinâmica em toda a rede através de um conjunto de m nós efetivos, isto é, as m novas

variáveis obtidas tentam capturar o comportamento semelhante de conjuntos de nós presentes na rede.

Aplicações importantes dessa metodologia são, do ponto de vista topológico, redes de comunidades, na qual é comum nós terem comportamentos semelhantes dentro de uma comunidade e distintos em relação aos de outras comunidades. Do ponto de vista dinâmico podemos citar, Kuramoto com grupos de frequências naturais distintas, integra-e-dispara com diferentes estímulos externos ou epidemia com indivíduos mais ou menos susceptíveis a doenças.

No presente trabalho nossa abordagem é, de certo modo, inversa a apresentada acima. Nós definimos que cada vetor de medida $\vec{x}^t$ é referente a uma iteração da dinâmica, e x_i^t corresponde a dinâmica do nó i no instante t , note a inversão de índices em relação ao método anterior e note também que $\vec{x}^t$ é basicamente a variável de estado do sistema. A ideia do método é tentar reduzir toda a dinâmica que ocorreu na rede a um pequeno número τ de variáveis de estado z , que vão caracterizar o comportamento relevante dos elementos que compõem o sistema. A grande vantagem desse método é que as variáveis obtidas são medidas que caracterizam individualmente cada nó, em outras palavras, a dinâmica que caracteriza um dado nó i é resumida pelo conjunto de valores $z_i^1, z_i^2, \dots, z_i^\tau$, isso nos possibilita estudar a relação das medidas dinâmicas e topológicas de cada nó da rede, obtendo portanto uma visão local (ou microscópica) da dinâmica ocorrendo sobre uma estrutura heterogênea.

Matematicamente, seja Λ uma matriz que armazena na linha t e coluna i o estado do nó i na iteração t , e Σ a matriz de covariância dos vetores linha de Λ . Se a nossa simulação executou em T_{sim} iterações, definimos a matriz E contendo em cada coluna j o autovetor de Σ relacionado ao j -ésimo maior autovalor, em ordem decrescente até uma determinada dimensão $m \ll T_{sim}$. Com isso as novas variáveis, $\vec{z}_i$, são dadas pela equação 2.34

$$\vec{z}_i = E' \vec{x}_i \quad (3.3)$$

Observe que para cada nó i obtemos m novas medidas. Para verificarmos a perda de informação relacionada a projeção definimos

$$g_m^{PCA} = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^m \lambda_i} \quad (3.4)$$

onde g_m^{PCA} representa a porcentagem de informação que está sendo mantida, em função dos autovalores de Σ . Para nossos propósitos faremos sempre que possível $m \leq 3$ e observaremos a perda de informação usando g_m^{PCA} .

4 Resultados e discussões

4.1 Patamar do estado assíncrono

Propomos iniciar nossa investigação da relação estrutura função olhando primeiro para uma estatística global do modelo integra-e-dispara muito tradicional no estudo de modelos neuronais em redes (71), usualmente chamada de ativação do sistema, R , e definida na equação 2.14. No nosso modelo essa ativação sempre tende a um valor constante, R_p , após um tempo muito longo (veja a Figura A.3 no apêndice A), é esse valor que estaremos olhando nas simulações. Salientamos que como esse não é o foco principal do presente trabalho, os resultados mostrados nessa seção são apenas exemplos de comportamentos que podemos observar olhando para uma estatística global do sistema, e não possuem a intenção de analisar a fundo as particularidades observadas.

Começamos nosso estudo com o modelo ER, na Figura 4.1 está mostrado o valor de R_p , tomado após um tempo muito longo, em função de $\langle k \rangle$ para diversos limiares T_l , a condição inicial utilizada foi sortear aleatoriamente para cada nó uma carga no intervalo $[0, T_l]$ com probabilidade uniforme. Para cada ponto do gráfico executamos 20 realizações da dinâmica em 20 redes distintas com $N = 10000$ e tiramos a média de R_p . Para garantir o grau médio desejado optamos por fixar o número de arestas da rede (ao invés de fixar a probabilidade de conexão), ainda sim existem pequenas flutuações nos valores obtidos de magnitude inversamente proporcional a $\langle k \rangle$, pois se torna muito difícil garantir uma rede totalmente conexa para $\langle k \rangle \approx \ln(N)$, fazendo com que o maior componente conectado não possua exatamente o grau médio desejado. Podemos ver na figura que para os diversos limiares utilizados parece existir uma transição de fase para o estado absorvente ($R_p = 0$), verificamos que para um dado limiar T_{lp} essa transição ocorre em $\langle k_c \rangle = T_{lp} - 1$. Esse resultado não traz tantas surpresas, pois o

modelo ER é conhecido por exibir transição de fase para o estado absorvente nas mais diversas dinâmicas (17, 19), o curioso é que parece existir uma segunda transição conforme aumentamos T_l , que ocorre quando $R_p \approx 0.5$. O comportamento do sistema nessa transição é o seguinte: em $R_p \approx 0.5$ metade dos nós da rede estão disparando a cada iteração para um dado grau médio $\langle k \rangle$, vale observar que os nós que disparam não são os mesmos a cada duas iterações (i.e. não há um comportamento oscilatório), ao aumentarmos um pouco o grau médio para $\langle k' \rangle$ é como se eliminássemos alguns nós de grau k tal que $k \leq \langle k \rangle$ e adicionássemos nós tal que $k \geq \langle k' \rangle$, essa pequena alteração nos graus de um número pequeno de nós (da ordem de dezenas) provoca um efeito dramático no comportamento global do sistema, fazendo com que os nós fiquem hiper excitados e passem a disparar a todo instante. Futuramente pretendemos encontrar a relação entre $\langle k \rangle$ e T_l de forma a caracterizar melhor a condição para a ocorrência dessa transição.

Figura 4.1 – Patamar de ativação da dinâmica integra-e-dispara em função do grau médio da rede, $\langle k \rangle$, após um tempo muito longo.

Outra pergunta que podemos fazer é sobre a dependência de R_p com o tamanho da rede ER, na Figura 4.2 comparamos R_p em função de $\langle k \rangle$ para três tamanhos distintos e $T_l = 10$, vemos que a média de R_p não parece depender de N , e portanto $\langle k_c \rangle$ não varia de forma significativa

com N . Um método costumeiramente utilizado para encontrar $\langle k_c \rangle$ de forma mais exata, bem como os parâmetros envolvidos na transição, é obtido através da análise de escala de tamanho finito (finite-size scaling) (2, 112), que busca extrapolar os resultados para $N \rightarrow \infty$, mas esta análise está fora do escopo desse trabalho.

Figura 4.2 – Patamar de ativação da dinâmica integra-e-dispara ao considerarmos três diferentes tamanhos de rede.

Queremos também comparar o comportamento de R_p para o modelo BA, mas infelizmente não podemos utilizar o mesmo método de fixar T_l e variar $\langle k \rangle$, pois o modelo não permite variarmos continuamente $\langle k \rangle$. Para resolvermos a questão recorreremos a um truque que consiste em sortear para cada nó i um limiar T_l^i de acordo com uma distribuição de Poisson (com a mesma definição da equação 2.6) de média $\langle T_l \rangle$, dessa forma os limiares de cada nó continuam sendo inteiros, mas podemos variar $\langle T_l \rangle$ de forma contínua. É importante salientar que não esperamos que essa nova estratégia seja equivalente a fixar um mesmo limiar para todos os nós, estamos apenas interessados em comparar os modelos de rede, entretanto, é natural inferir que o comportamento do patamar nas duas situações deva possuir algumas semelhanças. Na Figura 4.3 está mostrado o comportamento de $\langle R_p \rangle$ em função de $\langle T_l \rangle$ para o modelo ER com diferentes $\langle k \rangle$ e $N = 10000$. Nesse caso encontramos que a transição ocorre, para um dado grau médio $\langle k_p \rangle$,

apenas em $\langle T_{lc} \rangle \approx \langle k_p \rangle + 2.5$, o que se compararmos com $\langle k_c \rangle$ em função de T_{lp} do caso anterior vemos que $\langle k_p \rangle \approx \langle k_c \rangle - 1.5$, ou seja, se distribuirmos os limiares segundo uma poissoniana a dinâmica resultante possui uma maior capacidade de auto-sustentação, pois é preciso uma rede com menor grau médio para que a dinâmica entre no estado absorvente. Vemos também pela Figura 4.3 que a região $R_p \approx 0.5$ dá indícios de seguir o mesmo comportamento da Figura 4.1 conforme aumentamos $\langle k \rangle$, mas os valores de $\langle k \rangle$ para ocorrer a transição de fase devem ser muito superiores ao caso anterior.

Figura 4.3 – Patamar de ativação da dinâmica integra-e-dispara em função da média dos limiares sorteados para os nós de acordo com uma distribuição de Poisson.

Na Figura 4.4 mostramos $\langle R_p \rangle$ em função de $\langle T_l \rangle$ para a rede BA com $\langle k \rangle = 6$ e $N = 10000$, vemos que nesse caso não existe transição de fase, um resultado que já esperávamos obter, dado que nessas redes as altas flutuações na distribuição de grau dificultam o surgimento da transição (24, 52). Uma observação importante é que essa rede possui uma capacidade de auto-sustentação muito maior do que no modelo ER, o que dá a ideia de que é menos custoso para a topologia livre de escala sustentar a dinâmica integra-e-dispara em comparação com a ER, o que está de acordo com diversos trabalhos na literatura que sugerem que redes neuronais e corticais são governadas por leis de potência (40, 113, 114).

Figura 4.4 – Patamar de ativação da dinâmica integra-e-dispara em função da média dos limiares sorteados para os nós de acordo com uma distribuição de Poisson. A rede utilizada foi gerada com o modelo BA.

4.2 Quantificando a discriminação dos nós

4.2.1 Análise preliminar

Para melhor exemplificar a informação contida na nossa medida de discriminação dos nós, começamos a nossa análise utilizando o modelo de rede geográfica de Waxman, descrito na seção 2.1.2. Primeiramente utilizamos uma rede de 1000 nós com $\langle k \rangle = 10$ que está mostrada na Figura 4.5, onde as cores representam o grau k de cada nó. Devido a flutuações estatísticas, essa rede apresenta uma região mais densa na parte superior esquerda, o que será útil ao longo da nossa análise.

Executamos sobre a rede 1000 realizações da dinâmica integra-e-dispara, de forma a caracterizar a variação do comportamento dinâmico dos nós. Na Figura 4.6 mostramos a taxa média de cada nó, $\langle r_i \rangle$, obtida nessas realizações ao considerarmos três limiares, T_l , distintos. Podemos ver que em todos os casos os nós pertencentes a região de maior densidade obtiveram uma taxa média de disparos maior do que o resto da rede, um efeito causado pelo elevado grau dos

Figura 4.5 – Rede geográfica utilizada, cada nó foi colorido de acordo com o seu grau.

nós nessa região, outro ponto a ser observado é que a região mais densa também aparenta ser a que teve a dinâmica mais modificada com a mudança do limiar. Outra medida que caracteriza bem a dinâmica é a média da entropia de intervalos (equação 2.20) mostrada na Figura 4.7, que possui um alto valor próximo a borda da rede devido as maiores flutuações causadas pela baixa conectividade e um valor muito baixo no grupo mais denso, indicando que o sinal dinâmico nessa região é bem regular. Também observamos que agora o comportamento dinâmico das regiões menos densas da rede é que parecem ter sofrido uma variação significativa, enquanto que a região mais densa se "blindou" da alteração do limiar, o que exalta o fato já observado de que ao analisarmos a discriminação dinâmica dos nós devemos fazê-lo do ponto de vista de uma medida dinâmica.

Figura 4.6 – Taxa média ao longo de diversas realizações do integra-e-dispara para (a) $T_l = 9$, (b) $T_l = 10$ e (c) $T_l = 11$.

Figura 4.7 – Entropia média ao longo das mesmas realizações e limiares da Figura 4.6.

Apesar das Figuras 4.6 e 4.7 mostrarem valores bem característicos em certas regiões da rede, esses valores podem possuir grande variação entre diversas realizações, na Figura 4.8 mostramos os valores de r_i extremos que cada nó obteve quando consideramos todas as realizações para um mesmo limiar. Por exemplo, no caso da taxa vemos na Figura 4.8 que muitos nós da região mais densa apesar de possuírem uma taxa média alta, em algumas realizações obtiveram valores intermediários de taxa, apesar da atividade global da rede ter sido praticamente idêntica, ao passo que muitos outros nós não mostram nenhuma variação dinâmica. Esse comportamento também ocorre no caso da entropia (Figura 4.9), em que novamente observamos características opostas ao da taxa. Toda essa discussão foi feita para salientar novamente que a análise dinâmica a nível de nó requer a aplicação de um tratamento especial em várias realizações da dinâmica, com a finalidade de separar os efeitos da topologia e da condição inicial.

Figura 4.8 – Valores (a) mínimos e (b) máximos obtidos para a taxa em 1000 realizações da dinâmica com $T_l = 10$.

4.2.2 Cálculo de α

Modelo geográfico

Para calcularmos os α da rede geográfica apresentada utilizamos as mesmas 1000 realizações do integra-e-dispara com limiar $T_l = 10$ da seção anterior. A condição inicial foi sortear

Figura 4.9 – Valores (a) mínimos e (b) máximos obtidos para a entropia em 1000 realizações da dinâmica com $T_l = 10$.

aleatoriamente com probabilidade uniforme uma carga no intervalo $[0, T_l]$ para cada nó, feito isso deixamos a dinâmica evoluir até garantirmos experimentalmente que os parâmetros relevantes da dinâmica não possuísem uma variação significativa ao longo do tempo (veja apêndice A). Inicialmente analisamos apenas medidas dinâmicas individuais, e portanto utilizamos a distância definida na equação 2.29 baseada na distribuição Student-t para dados univariados, com ela calculamos a distância média entre as realizações de cada nó em relação a todos os demais da rede, dando origem a α . As medidas dinâmicas utilizadas para calcular α foram a taxa (α_r), entropia de intervalos (α_e) e intervalo máximo (α_i), o resultado do cálculo é mostrado na Figura 4.10. Vemos na Figura 4.10(a) que, do ponto de vista da taxa, os nós que foram mais diferenciados pela dinâmica são aqueles que possuem maior taxa (compare com a Figura 4.6), essa diferença aparenta ter sido muito grande em relação aos demais nós da rede, dando a entender que os nós de alta taxa representam *outliers* da dinâmica (nós de dinâmica muito diferenciada) vista por essa medida. Essa afirmação ganha mais força ao observarmos que essa é justamente a região mais densa de toda a estrutura da rede. Em relação a entropia (Figura 4.10b), vemos que a região bem diferenciada pela taxa agora não chega a ser um *outlier*, ou seja, apesar de ter sido novamente diferenciada a regularidade dos sinais aparenta ter sido mantida ao longo de uma larga região (veja Figura 4.7), fazendo com que em média cada nó não se distancie tanto do

resto da rede. Também tivemos alguns nós da borda sendo bem diferenciados, um resultado já esperado. A observação interessante é que justamente vários nós da borda possuíam dinâmica semelhante ao resto da rede. Ao observarmos α para o intervalo máximo, mostrado na Figura 4.10(c) vemos que os resultados são compatíveis com α_e , com a exceção de que os nós de maior taxa foram ainda menos diferenciados do resto da dinâmica.

Figura 4.10 – Valores de α obtidos ao considerarmos a (a) taxa, (b) entropia e (c) maior intervalo de disparo.

Calculamos também α considerando simultaneamente mais de uma medida dinâmica, as distância foram calculadas através da equação 2.29 considerando a estatística de Hotelling, que é a generalização do teste de Student-t para dados multivariados. Na Figura 4.11(a) mostramos α_{rei} calculado considerando simultaneamente as três medidas dinâmicas, esse resultado é mais difícil de ser interpretado isoladamente, mas comparando-o com a Figura 4.10 vemos que ele representa uma mescla dos resultados univariados, os nós de alta taxa apesar de terem sido bem diferenciados já não aparentam mais ser os principais *outliers* dinâmicos, esse comportamento é observado apenas para alguns nós da borda da rede. Testamos também α para a combinação de taxa e entropia, cujo resultado é mostrado na Figura 4.11(b), onde vemos o surgimento de uma terceira hierarquia de diferenciação da dinâmica representada pelos nós de cor azul clara envolvendo os nós da região mais densa da rede.

Toda a discussão de α que fizemos até aqui foi baseada na observação dos resultados mostrados sobre o gráfico da rede geográfica. Agora que observamos os diversos comportamentos dinâmicos podemos finalmente olhar para o histograma de α relativo às três medidas dinâmicas

Figura 4.11 – Valores de α obtidos ao considerarmos a combinação de (a) taxa, entropia e maior intervalo e (b) taxa e entropia.

Figura 4.12 – Histograma de α relativos a (a) taxa, (b) entropia e (c) maior intervalo de disparo.

Figura 4.13 – Rede geográfica com os nós coloridos de acordo com os picos da Figura 4.12, para as medidas de (a) taxa, (b) entropia e (c) maior intervalo de disparo.

utilizadas, que é mostrado na Figura 4.12. Esses gráficos mostram os mesmos resultados discutidos anteriormente, por exemplo, em relação a α_r , mostrado na Figura 4.12(a), o pico indicado pelo número 1 representa um grupo de nós que obteve pouca diferença dinâmica em relação ao resto da rede, o pico 2 revela a presença de um grupo com uma dinâmica intermediária e semelhante entre si, enquanto que o pico 3 indica a presença de *outliers* dinâmicos na rede. Na Figura 4.13(a) colorimos os nós que estão presentes nesses picos, mostrando que os nós do pico 1 estão espalhados pela rede, de forma que eles não possuem nenhuma característica topológica diferenciada, fazendo com que sua dinâmica seja semelhante à média da rede. O pico 3 refere-se aos já citados nós do centro da região densa da rede, ao passo que o pico 2 representa a maioria dos nós ao redor dessa região. O histograma de α_e (Figura 4.12b) dá indícios da presença de um grande grupo que foi diferenciado do resto da rede (pico 2), e a presença de alguns nós com diferenças dinâmicas ainda mais largas (região 3). Novamente observamos um grupo de pouca diferenciação na rede, indicado pelo pico 1. Na Figura 4.13(b) vemos que o pico 2 novamente representa a área mais densa, enquanto que a região 3 é formada pelos nós de borda que, como já observado, apresentam uma dinâmica mais irregular que o resto da rede. O histograma de α_i (Figura 4.12c) apresenta uma combinação dos resultados para α_r e α_e , como podemos ver pelas regiões coloridas na Figura 4.13(c).

Uma análise mais específica que podemos fazer é através das medidas topológicas dos nós

pertencentes a cada pico. Na Figura 4.14 exemplificamos essa análise para o caso de α_r . Para cada pico fizemos o histograma das medidas de grau, clustering, grau médio dos vizinhos (K^{nn}) e centralidade de autovetor (chamada de *evcent* na Figura) definidas na seção 2.1.4, sendo que colorimos em vermelho a porcentagem de nós pertencentes ao pico e em azul os demais nós da rede. Vemos pelos gráficos do pico 1 que nenhuma medida estrutural chegou a diferenciar os nós presentes nesse pico, o que é um forte indicativo de que, como supomos, esses nós estão distribuídos por toda a rede e não possuem nenhuma capacidade de diferenciação em relação a taxa. Para o pico 2 o resultado é bem diferente, os nós desse pico não parecem ter sido diferenciados pelo clustering, mas as outras três medidas nitidamente apresentam valores mais elevados em relação ao resto da rede, com especial atenção para o logaritmo da centralidade de autovetor, que dentre as quatro medidas é a que melhor captura o comportamento localizado dos nós do pico 2. Como o grau não é tão localizado, podemos dizer que a dinâmica integra-e-dispara não depende apenas dos vizinhos imediatos dos nós, e sim da posição global que ele se encontra na rede, um fato que torna muito difícil qualquer tratamento analítico sobre o problema. Em relação ao pico 3 vemos que novamente o clustering não parece influenciar na diferenciação, um fato que merece ser melhor estudado no futuro, já para as outras medidas vemos que em geral os nós desse pico possuem os maiores valores possíveis da medida, ou seja, eles são os nós que apresentam maior diferenciação topológica. Essa característica reflete fortemente na taxa, fazendo com que esses nós possuam taxas muito distintas do resto da rede. Observamos que essa diferenciação topológica não causa o mesmo efeito no ponto de vista da entropia, pois, como já observamos, esses nós fazem parte do pico 2 do histograma da Figura 4.12(b).

Como uma última informação em relação aos α calculados, na Figura 4.15 mostramos a média de α obtida para todos os nós de mesmo grau, sendo que cada gráfico foi sobreposto ao histograma dos graus da rede. Na Figura 4.15(a) vemos que α_r é baixo para nós de grau com valores pequenos e intermediários, ao passo que para graus acima da média os nós passam a ser bem diferenciados do resto da rede. É preciso que fique bem entendido a relevância desse resultado, pois estamos observando que apesar da distribuição de grau ser relativamente simétrica, os poucos nós da rede de grau reduzido, que podemos de certo modo chamar de

Figura 4.14 – Histogramas das medidas estruturas da rede geográfica utilizada, em vermelho está mostrada a fração de nós que pertencem ao respectivo pico do histograma de α_r .

outliers topológicos, não apresentam uma dinâmica muito distinta da maioria dos nós da rede, pelo menos ao compararmos com aqueles de grau elevado. Isso quer dizer que se queremos observar nós de sinais bem distintos nessa dinâmica, devemos dar preferência a regiões de maior densidade de conexões, e não a regiões bem esparsas. Em relação à entropia (Figura 4.15b), vemos um resultado mais usual, pois os nós de menor diferenciação dinâmica são basicamente aqueles que possuem um valor de grau intermediário, que são a grande maioria dos nós em toda a rede. Esses nós produzem uma entropia padrão que só é quebrada pelas regiões de alta (núcleo) e baixa (borda) densidade na rede, um resultado que já foi mostrado na Figura 4.13(b). Para α_i vemos um resultado inverso ao obtido para α_r , ou seja, para um nó ser bem diferenciado em relação ao intervalo máximo, ele deve possuir grau de valor baixo.

Figura 4.15 – α médio para cada grau k da rede, juntamente com o desvio padrão dos valores. As curvas foram sobrepostas ao histograma de graus da rede.

Para finalizar mostramos dois histogramas obtidos pela combinação das medidas utilizadas, na Figura 4.16(a) e (b) mostramos, respectivamente, os histogramas de α_{rei} e α_{re} , onde vemos que o resultado é semelhante ao de α_r , indicando que a taxa parece ser a medida dominante na caracterização da diferença dinâmica entre os nós. Num primeiro momento esses resultados podem parecer óbvios, pois podemos facilmente visualizar a topologia da rede em estudo. De fato, o caso das redes geográficas é justamente o mais simples de ser analisado, pois em muitas redes reais tudo que teríamos seriam os histogramas de α .

Figura 4.16 – Histograma de α relativos a combinação de (a) taxa, entropia e maior intervalo e (b) taxa e entropia.

Redes ER e BA

Vamos discutir brevemente a comparação entre os histogramas obtidos para o modelo geográfico e os modelos ER e BA. Em cada coluna da Figura 4.17 mostramos 6 histogramas de α_r obtidos em 6 redes distintas geradas de acordo com o respectivo modelo, os pontos de destaque dos gráficos são ressaltados com um círculo vermelho. Podemos ver que as redes criadas pelo modelo geográfico e BA exibem uma forte tendência de possuir grupos de nós que se diferenciam do resto da rede. No caso geográfico verificamos (utilizando o mesmo processo que gerou a Figura 4.13) que esses nós em geral estão conectados na rede, formando grupos de mesma dinâmica causados por flutuações na criação da rede. Comparando com os histogramas da rede ER vemos que apesar de ambos os modelos serem, de certa forma, totalmente aleatórios sem nenhum privilégio aparente para certos nós, o modelo geográfico apresenta flutuações mais significativas do que o ER, sendo que o último não consegue provocar o surgimento natural dos grupos dinâmicos. Em relação ao modelo BA observamos a presença de um grupo muito diferenciado, os nós que formam esses grupos nada mais são do que os hubs (nós de grau muito alto) da rede. O resultado interessante é que o esperado seria que o histograma formasse um contínuo de probabilidade até chegar aos hubs, mas o que ocorreu é que os hubs se tornaram um grupo à parte do resto da rede. O que vemos aqui é uma segregação da dinâmica, mesmo os nós de grau intermediário são considerados próximos aos muitos nós de baixo grau, criando

uma enorme diferença dinâmica entre os nós intermediários e os hubs.

Figura 4.17 – Histogramas de α relativos a medida de taxa para três modelos de rede.

Rede *C. elegans*

Para ilustrar o tipo de informação que podemos obter utilizando α , aplicamos a dinâmica integra-e-dispara sobre a rede *C. elegans* apresentada na seção 2.1.3. É importante notar que o nosso objetivo é, novamente, dar exemplos de aplicações de α , e não explicar o comportamento dos neurônios na rede, pois para isso teríamos que no mínimo considerar conexões inibitórias e o tipo de sinapse presente na comunicação. Na Figura 4.18 está mostrado os histogramas de α obtidos com o mesmo processo da seção anterior. Um resultado imediato que podemos observar é a alta diferenciação de alguns nós em relação a α_r , indicados com o círculo vermelho, sendo que essa diferenciação se parece muito com a obtida para o modelo BA na seção anterior.

Figura 4.18 – Histograma de α relativos a (a) taxa, (b) entropia e (c) maior intervalo para a rede *C. elegans*.

Com isso em mente mostramos na Figura 4.19 a distribuição de grau da rede, onde colocamos em vermelho os nós indicados pelo círculo da Figura 4.18. Vemos que de fato os nós de maior diferenciação são os que em geral possuem o maior grau na rede, mas é importante observar que mesmo alguns nós de alto grau não chegaram a ser tão diferenciados pela dinâmica, o que indica que não é apenas o grau que influencia na taxa dos nós. A realidade é que os nós mostrados em vermelho são os interneurônios do cordão ventral da *C. elegans*, essas estruturas são conhecidas por possuírem muitas sinapses (115), já que elas fazem a comunicação entre neurônios sensoriais e motores sem muita restrição nos tipos de sinais transmitidos (outras classes de interneurônios são conhecidos por receberem apenas determinados tipos de sinais).

Figura 4.19 – Histograma dos graus da rede *C. elegans*, os valores correspondentes aos interneurônios do cordão ventral foram coloridos em vermelho.

Tabela 4.1 – Valores calculados para os interneurônios do cordão ventral da *C. elegans*, que obtiveram alta diferenciação dinâmica em relação à taxa.

Neurônio	AVAL	AVAR	AVBL	AVBR	AVDL	AVDR	AVEL	AVER	PVCL	PVCR
Posição	0.13	0.13	0.15	0.15	0.16	0.16	0.13	0.14	0.82	0.82
D1	1.00	1.00	1.00	1.44	1.33	1.22	1.33	1.22	1.00	1.00
D2	1.76	1.77	1.81	1.80	1.97	1.87	1.90	2.09	2.04	2.03

Na Tabela 4.1 apresentamos os nomes tradicionais e alguns valores calculados para esses neurônios. Posição refere-se a localização espacial de cada neurônio no eixo que vai da cabeça (valor 0) até a cauda (valor 1) do verme, onde vemos que a maioria dos neurônios estão localizados na mesma região da cabeça (mais especificamente, nas proximidades do anel nervoso (116, 117)), com exceção dos neurônios PVCL e PVCR que se localizam na cauda. D1 é a distância topológica média entre esses neurônios, isto é, dado o nó i medimos quantas arestas da rede precisamos percorrer para chegar ao nó j , e calculamos a média dessa distância para todo j , relativo aos nós que foram diferenciados. $D1 = 1$ significa que o nó em questão é vizinho (recebe sinais diretamente) de todos os outros mostrados na tabela. O parâmetro D2 indica a distância topológica entre o nó em questão e todos os demais da rede, excluindo aqueles presentes no grupo diferenciado. Vemos que em todos os casos os nós desse grupo estão mais próximos entre si do que com o resto da rede (efeito em parte provocado pela alta chance de nós de grau elevado estarem conectados), ou seja, além de possuírem alto grau esses nós possuem uma boa capacidade de comunicação interna ao grupo, o que provoca uma alta capacidade de diferenciação da sua dinâmica.

Notamos também que é muito difícil definir qual a distribuição de grau dessa rede, já que ela possui poucos nós, mas a presença de nós de grau muito elevado em relação aos demais sugere um comportamento em lei de potência. Por fim, apesar de a taxa ter gerado um grupo de alta diferenciação, ao considerarmos a entropia de intervalos (Figura 4.18b) não obtivemos grupos bem definidos, o que indica não haver na rede regiões de maior ou menor homogeneidade de sinal, notamos também que α_i obteve resultados semelhantes a α_r .

4.2.3 Aglomeração dinâmica

Redes geográficas

Procedemos agora para a análise mais específica dos grupos formados pela dinâmica integrada-dispara. Inicialmente utilizamos a mesma rede da seção anterior (Figura 4.5), que possui uma pequena região de alta densidade. Na Figura 4.20 mostramos o dendrograma obtido para o caso da medida de taxa, sendo que grupos aglomerados em uma distância menor do que 100 estão coloridos para melhor visualização. Vemos que o dendrograma sugere uma forte divisão em 2 ou 3 grupos, já que a distância para formar essa quantidade de grupos é muito maior do que em qualquer outra divisão. Com isso em mente decidimos dividir os nós em 3 grupos dinâmicos, isto é, nós que de acordo com a medida dinâmica considerada foram influenciados de forma semelhante pela topologia. Na Figura 4.21(a) vemos o referido resultado para o caso da medida de taxa, onde cada grupo foi colorido com uma cor distinta. Está claro que a região mais densa da rede influenciou fortemente a criação desses grupos, o grupo em vermelho é basicamente constituído pelos nós de maior taxa da dinâmica, já o grupo em verde forma um envoltório em torno do grupo vermelho sugerindo uma transição entre a região de alto e baixo disparo na rede. Esses resultados são corroborados pela Figura 4.21(b), em que mostramos a dinâmica separada em 5 grupos distintos, vemos novamente que são formadas “cascas” de dinâmicas semelhantes em torno da região de alta densidade.

Com a finalidade de observar melhor a formação dos grupos, optamos por utilizar outra rede geográfica também gerada pelo modelo de Waxman, a diferença é que essa nova rede não possui um grupo de alta densidade dominante, de forma que a topologia gerada possui mais heterogeneidades do que a anteriormente utilizada. Como uma análise adicional optamos por inserir uma região completamente regular no centro da rede, de forma a também observar como essa região sofrerá a influência da dinâmica. A rede utilizada está mostrada na Figura 4.22. Na Figura 4.23 mostramos o dendrograma obtido sobre essa nova rede para o caso da medida de taxa, onde novamente observamos a formação de grupos bem definidos.

Na Figura 4.24(a) mostramos a rede completa dividida em 3 grupos, para mostrarmos as

Figura 4.20 – Dendrograma da rede geográfica com região mais densa.

Figura 4.21 – Grupos formados na rede geográfica.

Figura 4.22 – Rede utilizada para a aglomeração dinâmica.

Figura 4.23 – Dendrograma obtido para as taxas dos nós na rede da Figura 4.22.

conexões entre os nós impusemos adicionalmente a condição dos grupos formados estarem conectados também pela topologia, isto é, nós com mesmo comportamento dinâmico também devem estar topologicamente próximos para ganharem uma conexão. Vemos na Figura que os grupos em vermelho tendem a formar clusters bem localizados na rede, enquanto que os grupos em verde e azul percorrem praticamente toda a estrutura. Observamos também uma quebra interessante da região regular, em que o canto inferior direito parece ter sofrido uma forte influência da vizinhança externa, o suficiente para modificar sua dinâmica. Também fizemos a separação em 5 grupos, mostrada na Figura 4.24(b), onde vemos que os grupos em vermelho ficaram praticamente inalterados, indicando que de alguma forma a topologia local desses grupos faz com que a dinâmica entre eles seja muito semelhante. Vemos também que além de não existirem mais grupos da grande tamanho, a região regular central sofreu uma nova divisão, ou seja, apenas as irregularidades da vizinhança dessa região fizeram com que a mesma fosse diferenciada dinamicamente de uma forma não trivial.

Figura 4.24 – Grupos formados na aglomeração da taxa para a rede geográfica.

Adicionalmente mostramos o surgimento dos grupos em relação a entropia ao considerarmos 3 (Figura 4.25a) e 5 (Figura 4.25b) grupos, em ambos os casos vemos a presença de um grande grupo, mostrado em vermelho, que percorre praticamente toda a rede, indicando

que grande parte da rede possuiu uma dinâmica muito semelhante. Vemos também na Figura 4.25(b) a separação da região regular em diversos grupos, ou seja, a vizinhança da região de origem a diferentes entropias de intervalo, fazendo com que do ponto de vista de complexidade do sinal, a região regular apresente os mais variados comportamentos. Esse resultado deixa evidente um estudo muito interessante a ser feito na área de redes, que é a influência que a região externa provoca no interior de uma estrutura completamente regular. Podemos estudar se é possível manter uma dinâmica regular nessa estrutura mesmo com uma borda irregular, podemos verificar condições para provocar uma determinada diferenciação dinâmica dos nós ou ainda podemos estudar o problema do ponto de vista de controle em redes (118).

Figura 4.25 – Grupos formados na aglomeração da entropia para a rede geográfica.

Rede cortical

A rede real que usaremos para aplicar o método de aglomeração é a rede cortical apresentada na seção 2.1.3. Na Figura 4.26 mostramos o dendrograma obtido ao realizarmos a aglomeração da entropia de intervalos da dinâmica integra-e-dispara aplicada sobre essa rede, considerando diferentes condições iniciais. Vemos que a dinâmica apresenta uma divisão significativa em dois grupos com praticamente o mesmo número de nós, e justamente essa igualdade

sugere que a rede deve estar dividida em dois grupos estruturais distintos. O caminho natural a seguir é verificar se realmente a estrutura da rede possui essa divisão, para tanto utilizamos um algoritmo clássico de detecção de comunidades desenvolvido por Reichardt e Bornholdt (119) que faz uma analogia com o modelo de Potts da mecânica estatística (120). Com esse algoritmo é possível inicialmente separarmos a rede em um número aleatório de comunidades (ou spins), e deixar o sistema evoluir para um estado de mínima energia que representa não só a melhor divisão, mas também o número ideal de comunidades presentes na rede. O algoritmo retorna que o número ideal de comunidades é 2, e que a divisão final da rede possui uma modularidade (que quantifica a qualidade de uma dada partição topológica de uma rede, definida em (121)) $Q_e = 0.35$. Podemos também calcular a modularidade para a divisão em 2 grupos gerada pela dinâmica, o valor obtido é de $Q_d = 0.33$, ou seja, apenas observando a entropia de intervalos conseguimos obter uma partição da rede tão boa (e significativa) quanto um algoritmo que considera apenas a estrutura. Mas o mais importante é que até hoje não há um consenso de que linha seguir na identificação de comunidades em redes direcionadas, até mesmo a versão direcionada da modularidade definida por Newman (122) é criticada por não considerar alguns casos especiais que podem surgir (123). O que fizemos com a entropia foi encontrar uma divisão *funcional* em comunidades, ou seja, nós com comportamentos dinâmicos distintos são classificados em comunidades distintas, não importa se há direção ou não nas conexões. É claro que ainda fica em aberto o problema de verificar a significância da divisão.

Na Figura 4.27 mostramos a rede considerada utilizando o algoritmo de Fruchterman e Reingold (124) para construir a visualização, a Figura 4.27(a) mostra os nós coloridos de acordo com os grupos dinâmicos formados pela entropia, a Figura 4.27(b) é a divisão obtida pelo algoritmo de Reichardt e Bornholdt. Vemos que a diferença na classificação parece ocorrer apenas nos nós presentes na interface entre as duas comunidades, região na qual é difícil dizer qual seria a classificação ideal. Esse resultado revela um grande potencial de estudo nessa ideia de divisão funcional, que é basicamente dividir a rede de acordo com o processo que ocorre sobre ela. Existem poucos trabalhos firmados nessa área, e em geral eles utilizam a dinâmica de Ising (ou Potts) (119, 125), passeio aleatório tradicional (126, 127) ou sincronização (128, 129) como processo. Em trabalhos futuros pretendemos abordar esse tema em um contexto

Figura 4.26 – Dendrograma da aglomeração dinâmica para a medida de entropia na rede cortical.

envolvendo outras dinâmicas, generalizando a análise presente na literatura e nesse trabalho.

Figura 4.27 – Visualização da rede utilizando o algoritmo de Fruchterman-Reingold, os nós foram coloridos de acordo com (a) grupos formados pela aglomeração da entropia (b) comunidade identificada pelo algoritmo de Reichardt-Bornholdt.

4.2.4 Osciladores de Kuramoto

Até agora mostramos resultados apenas para a dinâmica integra-e-dispara, pois de certa forma seria redundante analisar os mesmos casos para dinâmicas distintas, visto que o objetivo do trabalho é ilustrar a nossa metodologia. Mas o fato é que no modelo integra-e-dispara que utilizamos não há como controlar efetivamente a sincronização dos nós (89), sendo que essa é uma área de estudo muito abordada em redes. O efeito da sincronização traz como consequência a oportunidade de definir o alcance da dinâmica, que como já discutido, define a relação entre a dinâmica de dois nós a uma certa distância na rede. O modelo de Kuramoto é perfeito para tal abordagem porque no caso em que os osciladores estão desacoplados podemos dizer que o alcance da dinâmica é nulo, já no caso de perfeita sincronia o alcance é infinito (ou igual ao maior caminho mínimo entre os nós). Mas é importante observar que alcance não necessariamente implica sincronia, estamos apenas olhando para um caso particular. Para estudarmos a dinâmica na rede, fixamos a frequência natural de cada nó, e para cada realização variamos apenas as fases iniciais. A frequência natural foi sorteada aleatoriamente com uma distribuição uniforme no intervalo $[-0.5, 0.5]$, já as fases também seguem uma distribuição uniforme mas são definidas no intervalo $[-\pi, \pi]$, o número de realizações foi fixado em 100 para cada acoplamento. É importante notar que ao sortearmos frequências distintas para os nós estamos criando a possibilidade deles serem diferenciados pelo parâmetro dinâmico, e não somente pela topologia ou condição inicial, como vínhamos estudando até então. Portanto nessa análise iremos supor que como as frequências são definidas em um intervalo pequeno, e sorteadas com distribuição uniforme, não são esperadas muitas diferenças de comportamento devido a esse parâmetro, no sentido que a topologia possui maior capacidade de diferenciar os nós do que pequenas variações de frequência.

A Figura 4.28 mostra para diferentes acoplamentos, λ , da dinâmica os histogramas de α_f , calculados utilizando a medida de frequência efetiva $\tilde{\omega}$ (equação 2.22), a rede utilizada foi gerada utilizando o modelo ER com $N = 1000$ e $\langle k \rangle = 10$. No caso de acoplamento nulo ($\lambda = 0$) a fase dos nós é completamente aleatória, não há como haver influência entre vizinhos, e

portanto o grupo formado para α_f baixo no gráfico é simplesmente as diferenças entre variáveis aleatórias de probabilidade uniforme, ou seja, não há nenhuma significância. Até $\lambda \leq 0.05$ ainda temos um comportamento sem muitas diferenças em relação ao caso $\lambda = 0$, ao passarmos desse valor uma rápida transição ocorre na rede e os osciladores se tornam sincronizados, de forma que para $\lambda = 0.075$ praticamente todos os nós possuem uma dinâmica muito semelhante. Mas podemos ver pela ampliação do gráfico que ainda existem alguns nós com dinâmica diferenciada dos demais, esses nós se mostram mais difíceis de sincronizar com o resto da rede, dando origem a uma dinâmica muito distinta, para todos os efeitos esses nós se tornam *outliers* dinâmicos. Para $\lambda = 0.1$ o número de nós não sincronizados é muito pequeno, praticamente insignificante, e sua dinâmica é ainda mais diferenciada do que no caso $\lambda = 0.075$. Aumentando ainda mais o acoplamento ($\lambda = 0.15$) o número de nós diferenciados se reduz a ordem de unidades, e o acoplamento é tão forte que esse nós já não conseguem se diferenciar tanto, aos poucos sua dinâmica está se tornando idêntica aos demais nós da rede.

Figura 4.28 – Histogramas de α_f relativos a frequência efetiva dos osciladores de Kuramoto, diversos acoplamentos são considerados. Na segunda linha o subgráfico mostra a ampliação do gráfico principal.

Outra forma de observar o agrupamento dos nós é olhando para a aglomeração da dinâmica,

para tanto utilizamos o mesmo procedimento de fixar as frequências naturais e variar as fases em cada realização, para cada acoplamento geramos 20 realizações da dinâmica e calculamos o dendrograma associado. Para obter significância estatística repetimos o procedimento para 20 diferentes redes geradas pelo modelo ER com $N = 1000$ e $\langle k \rangle = 10$. Com isso temos para cada rede e para cada acoplamento um dendrograma descrevendo a aglomeração dos nós, para obter os grupos formados em cada situação precisamos fixar uma distância na qual todos os dendrogramas serão “cortados”. Escolhemos uma distância tal que nas 20 redes fossem formados aproximadamente 10 grupos para o caso desacoplado, é importante notar que apesar dessa escolha ter sido arbitrária o comportamento dos resultados não variou para um grande intervalo de valores de distância testados. Tendo os grupos formados em cada rede para cada acoplamento tiramos duas estatísticas sobre eles: (a) número de grupos formados e (b) número de nós no maior grupo formado. Então para cada acoplamento tiramos a média dessas estatísticas sobre as 20 redes para obter o número médio de grupos (N_c) e tamanho médio do maior grupo formado (G_c), o resultado é mostrado na Figura 4.29. Na Figura 4.29(a) mostramos como referência o nível de sincronização global da rede (R_k), onde vemos que a transição para o estado sincronizado ocorre em $\lambda \approx 0.06$, na Figura 4.29(b) vemos que o número médio de grupos dinâmicos é praticamente constante até valores de λ próximos de 0.06, a partir desse valor temos um rápido aumento no número de grupos que, após atingir um máximo, decai vagarosamente conforme aumentamos ainda mais λ . Para interpretarmos esse comportamento olhamos para a Figura 4.29(c) que mostra o tamanho médio do maior grupo em cada acoplamento, para $\lambda < 0.06$ o maior grupo em geral é formado por 100 nós, de forma que como temos em torno de 10 grupos estatisticamente esses grupos devem ter aproximadamente o mesmo tamanho. Ao aumentarmos um pouco o acoplamento ocorre uma transição na dinâmica, de forma que os grupos se unem para formar um grande grupo abrangendo quase toda a rede, os nós que não se unem a esse grande grupo são aqueles de difícil sincronização, que se tornam *outliers* dinâmicos, formando diversos grupos de tamanho reduzido. Aumentando ainda mais o acoplamento esses *outliers* vão aos poucos se associando com o grande grupo, de forma que todos os nós de rede passem a possuir a mesma frequência efetiva. Em (29) Gardeñes et al. fizeram uma análise semelhante a aqui apresentada, mas devido a presença de alguns parâmetros não definidos no trabalho deles,

não conseguimos reproduzir os resultados a fim de comparar com os nossos.

Figura 4.29 – Sincronização dos osciladores de Kuramoto. Mostramos o (a) nível de sincronia dos nós, (b) número de grupos de frequência efetiva semelhante e (c) tamanho do maior grupo formado, considerando diversos acoplamentos.

Como uma última análise olhamos para os *outliers* presentes na dinâmica para um acoplamento $\lambda = 0.1$. A Figura 4.30(a) mostra a contagem (ou histograma não normalizado) dos valores absolutos das frequências efetivas, $abs(\tilde{\omega})$, dos nós para o acoplamento considerado, medidas em 100 realizações com 100 redes distintas e frequências naturais dos nós *distintas* em cada realização. Foi tirado o logaritmo do eixo das ordenadas para melhor visualização, portanto tenha em mente que o número de pontos com valores pequenos de $abs(\tilde{\omega})$ é muito alto. Separamos arbitrariamente nesse gráfico os nós com $abs(\tilde{\omega}) > 0.1$ (coloridos em verde) e fizemos um *scatter-plot* entre o valor absoluto da frequência *natural*, $abs(\omega)$, e a centralidade de autovetor de cada nó (equação 2.12), mostrado na Figura 4.30(b). Vemos que os nós que foram diferenciados pela dinâmica possuem uma baixa centralidade e alto valor absoluto de frequência natural, isto é, apenas uma das duas medidas seria incapaz de identificar os *outliers*,

e portanto a condição para o nó ser diferenciado não é mais apenas devido a estrutura, mas também ao parâmetro dinâmico que ele possui. Com isso vemos que a suposição que fizemos no início da seção, de que o efeito da frequência natural deve ser pequeno na diferenciação, não é válido para os nós de baixa centralidade. A abordagem correta para analisar essa dinâmica seria considerar não só a variação da condição inicial, mas também dos parâmetros dinâmicos de cada nó, pretendemos aplicar essa ideia em trabalhos futuros.

Figura 4.30 – (a) Contagem (ou histograma não normalizado) do valor absoluto das frequências efetivas da dinâmica de Kuramoto. (b) *Scatter-plot* entre o valor absoluto das frequências naturais e a centralidade de autovetor na mesma dinâmica. Foram consideradas 100 realizações em 100 redes distintas para um acoplamento $\lambda = 0.1$.

4.3 Aplicação do PCA

Iniciamos nosso método utilizando o PCA para visualizar a dinâmica integra-e-dispara (equação 2.15) sobre uma rede ER. Nosso objetivo além da visualização é obter de forma completamente automatizada medidas dinâmicas relevantes que caracterizem o comportamento dos nós. A rede ER que utilizamos possui $N = 10000$ e $\langle k \rangle = 10$. Na Figura 4.31 está mostrada a projeção 3D do PCA para $T_l = 8$, com a explicação das novas variáveis indicadas no eixo correspondente (equação 3.4), cada ponto foi colorido com o grau do nó associado.

Podemos ver que os pontos se organizaram em uma forma bem definida, que pode ser

Figura 4.31 – Projeção 3D da dinâmica integra-e-dispara utilizando o método de PCA. Adaptado de (1).

dividida em três regiões: (i) Uma parte superior cônica relativamente esparsa no centro e aparentemente densa na ponta; (ii) uma 'cintura' com uma grande quantidade de nós; e (iii) uma 'cauda' que de certa forma se parece com a parte superior da projeção. A figura apresenta uma riqueza estrutural inesperada, dado que usamos um modelo de topologia completamente aleatória e uma dinâmica idêntica para cada nó. Na Figura 4.32 estão mostradas as projeções 2D da Figura 4.31 e os correspondentes histogramas de densidade de pontos em cada região. Vemos na Figura 4.32(a) que os pontos parecem estar restritos a uma região bem definida que lembra vagamente a figura de um 'olho', possuindo uma borda que determina o máximo de espalhamento no plano $PCA1 \times PCA2$. Em 4.32(b) Está claro que o grau dos nós possui uma forte relação com o 3º eixo do PCA, fato que investigaremos mais a frente. Vemos também em 4.32(b) um certo padrão de repetição, uma estrutura cônica que parece estar presente em diferentes escalas ao longo de $PCA3$. Nas Figuras 4.32(c) e 4.32(d) vemos que as regiões de maior concentração dos pontos ocorrem nos valores máximos e mínimos de $PCA1$, fato que explica o porquê desse eixo ter sido escolhido como o de maior variação. Vemos também uma

alta densidade logo abaixo do disco central, em uma região na qual PCA1 (e PCA2 de acordo com a Figura 4.31) parece possuir uma forte restrição em seus valores.

Figura 4.32 – Projeções 2D do PCA para a dinâmica integra-e-dispara. (a) e (b) nós projetados coloridos de acordo com o grau, (c) e (d) histogramas da densidade de nós em cada projeção.

Como observado o eixo PCA3 possui uma forte relação com o grau dos nós, para investigarmos melhor esse fenômeno na Figura 4.33 mostramos alguns exemplos de sinais dinâmicos ao longo desse eixo, onde fica fortemente indicado que PCA3 está relacionado com a taxa de disparos dos nós. Com isso em mente mostramos na Figura 4.34 a relação entre PCA3 e taxa, onde vemos que essas duas variáveis possuem um coeficiente de correlação de Pearson (111) (que mede o quão linear é a relação entre duas variáveis) $\gamma = 0.99$, ou seja, de forma totalmente automática obtivemos uma medida clássica largamente utilizada no estudo de dinâmica neuronal.

Figura 4.33 – Exemplos de sinais de alguns nós projetados ao longo de PCA3. Adaptado de (1).

Figura 4.34 – Explicação da medida PCA3. (a) Nós coloridos de acordo com a medida de taxa e (b) *scatter-plot* entre taxa e PCA3. O coeficiente de correlação de Pearson, γ , é mostrado na Figura.

Em relação a PCA1 observamos durante os experimentos uma forte relação do eixo com os instantes de disparo do nó, isto é, a projeção do sinal do nó nesse eixo está relacionada com o balanço entre os disparos que ocorreram em iterações de índice par e ímpar. Uma forma de quantificar essa característica é definir duas medidas s_{par} e s_{impar} que representam para um dado nó, respectivamente, o número de disparos em iterações par e ímpar, com isso definimos uma terceira medida na forma $s_{p-i} = s_{par} - s_{impar}$. A Figura 4.35 mostra a relação dessa medida com PCA1. A interpretação que tivemos dessa medida é que ela representa a fase relativa do sinal dos nós em relação aos nós presentes nos extremos de PCA1, esses nós (que são em grande número) parecem ditar a dinâmica na rede para esse limiar, e PCA1 indica o quanto os demais nós estão com o sinal semelhante a eles.

Figura 4.35 – Explicação da medida PCA1. (a) Nós coloridos de acordo com a medida s_{p-i} e (b) *scatter-plot* entre essa medida e PCA1.

PCA2 possui uma interpretação mais complexa, e portanto não entraremos em detalhes, mas ainda podemos qualificar a medida definindo s_{pi} como sendo o número de disparos de um nó que ocorrem em instantes pares na primeira metade do sinal, mais o número de disparos em instantes ímpares na segunda metade, de forma análoga definimos s_{ip} para obter uma terceira medida dada por $s_{pi-ip} = s_{pi} - s_{ip}$. O resultado é mostrado na Figura 4.36, onde vemos mais uma vez uma forte relação entre as variáveis.

Figura 4.36 – Explicação da medida PCA2. (a) Nós coloridos de acordo com a medida s_{pi-ip} e (b) *scatter-plot* entre essa medida e PCA2.

5 Conclusões

Na literatura é comum encontrar análises de dinâmicas em redes que levam em conta apenas as estatísticas globais do sistema. Nosso primeiro objetivo foi mostrar que essa visão pode inadvertidamente desconsiderar importantes comportamentos dos nós a nível local no sistema. Para motivar tal fenômeno citamos quatro efeitos que podem influenciar a dinâmica particular de um nó: (a) condição inicial, (b) topologia, (c) dinâmica particular e (e) estocasticidade. Nosso objetivo principal foi analisar as influências (a) e (b), de forma que definimos uma medida, α , para quantificar a diferença entre a dinâmica dos vários nós da rede quando sujeita a uma série de condições iniciais. Em uma análise preliminar motivamos a necessidade dessa medida dada a grande variação que observamos na dinâmica integra-e-dispara aplicada sobre a rede geográfica. No mesmo modelo geográfico a medida revelou a constante presença de grupos de alta diferenciação causados por flutuações estatísticas da topologia local da rede, esses grupos aparecem sob diversas formas dependendo da medida dinâmica que está sendo levada em conta. Em relação à taxa verificamos que regiões de centralidade maior do que o padrão da rede possuem uma grande capacidade de diferenciação, gerando nós cuja dinâmica se tornou muito diferente do resto da rede, ao passo que nós de baixo grau não foram tão diferenciados. Vimos também o surgimento de grupos de taxa intermediária, que caracterizam as bordas das regiões mais densas da rede, um comportamento peculiar que parece não ter sido descrito na literatura. Ao considerarmos a medida de entropia de intervalos vimos a presença de grupos de borda da rede, que dessa vez foram os que tiveram a dinâmica mais diferenciada, uma observação interessante é que do ponto de vista da entropia nós de topologia distinta da média da rede são diferenciados de forma semelhante, fato que pode trazer novas ideias ao estudarmos o sistema com métodos de teoria da informação. Para a medida de intervalo máximo observamos resultados hora parecidos com a taxa, hora parecidos com a entropia.

Ao aplicarmos a medida α para outros modelos de rede pudemos observar a potencial que a medida possui para caracterizar uma estrutura de forma funcional, vimos que as flutuações que ocorrem no modelo geográfico não possuem análogo no modelo ER, um fato interessante provavelmente causado pela alta conectividade (ou baixa média dos caminhos mínimos (4)) do último. No modelo BA observamos a extrema diferenciação que os hubs adquiriram, que efetivamente separou sua dinâmica dos demais nós da rede, um fato curioso dado que existem na rede nós intermediários que deveriam formar um contínuo de comportamentos, entretanto encontramos que esses nós possuem a dinâmica mais semelhante com os nós de baixo grau do que com os hubs. Pretendemos analisar melhor esse fenômeno ao considerarmos outras dinâmicas, de forma a verificar se essa diferenciação sempre ocorre.

Após apresentar α , mostramos um processo de criação de grupos dinâmicos no mesmo espírito dessa medida. Mas que através da aglomeração hierárquica consegue estudar de forma ainda mais específica as particularidades da dinâmica. No modelo geográfico observamos a presença de cascas de ativação ao redor das regiões mais densas ao separarmos a taxa dos nós em 3 ou 5 grupos, ainda mais interessante foi observar o efeito da presença de uma região topologicamente regular nessa rede, onde mostramos que apenas as heterogeneidades da borda da região são capazes de separar a dinâmica dos nós. Esse é um fenômeno muito interessante que precisa ser melhor estudado. Em trabalhos futuros pretendemos abordar esse tema do ponto de vista de controle em redes (118), tentando encontrar como podemos controlar a dinâmica em uma região regular através da manipulação dos nós da sua borda.

Saindo dos modelos aleatórios estudamos a diferenciação que a topologia provoca na dinâmica através de duas redes reais muito utilizadas na área. O primeiro estudo foi feito com α sobre a rede do verme *Caenorhabditis elegans*, onde encontramos a presença de um grupo de 10 neurônios, todos com uma distância topológica entre si próxima de 1, cuja topologia provoca uma larga diferença nas taxas medidas. Vimos que esses nós possuem um alto grau na rede, mas como na rede existem outros nós com grau semelhante que não chegaram a ser tão diferenciados, vimos que o alto grau não é uma condição suficiente para a diferenciação nessa rede, foi preciso que os nós também estivessem bem conectados entre si. Como já citado, pre-

tendemos continuar estudando a rede desse verme através de dinâmicas mais robustas, como o integra-e-dispara generalizado definido na equação 2.13 ou outras dinâmicas neuronais apresentadas na seção 2.13. Com relação à rede cortical do macaco, um resultado particularmente interessante foi obtido. Conseguimos identificar a divisão mais provável em comunidades da rede através da medida de entropia de intervalos, e mostramos que, do ponto de vista estrutural, a divisão possui uma qualidade muito semelhante a um algoritmo clássico utilizado em redes que considera apenas a estrutura. Esse resultado abre caminho para uma série de estudos que pretendemos fazer a respeito da divisão *funcional* de uma estrutura, onde nosso objetivo final é definir uma medida de modularidade (121) dinâmica sobre uma estrutura.

Para finalizar a análise de discriminação, utilizamos a dinâmica de Kuramoto para definir o alcance que a dinâmica possui na rede, e o seu efeito na relação entre a dinâmica dos nós. Mostramos que existem alguns nós na rede ER que se mostram difíceis de sincronizar, de forma que é preciso um acoplamento muito forte para acrescentá-los ao grupo dos nós sincronizados. Através de um *scatter-plot* entre a medida dinâmica de frequência natural e a medida de centralidade de autovetor vimos que não é mais apenas a topologia que provoca a diferenciação, tornando necessário abranger ainda mais a análise para o caso de diferentes parâmetros dinâmicos de cada nó. Em trabalhos futuros pretendemos estudar além do efeito dos parâmetros, a variação dinâmica que uma dinâmica estocástica consegue exibir.

Finalmente mostramos brevemente uma metodologia baseada na análise por componentes principais que busca adquirir medidas de uma dinâmica de forma mais automatizada possível. O método foi capaz de encontrar de forma exata a medida de taxa no integra-e-dispara, que sempre foi tradicionalmente utilizada na área. Adicionalmente observamos duas medidas de interpretação mais complexa, que se relacionam com as fases que cada nó possui em relação ao maior grupo de nós de dinâmica semelhante. Já iniciamos um estudo mais completo envolvendo outras dinâmicas em (1), e pretendemos continuar a análise em outros modelos de rede.

Os resultados que observamos durante esse trabalho acabaram sendo analisados, até certo ponto, de forma não muito específica. Isso se deve ao fato de termos observado uma riqueza muito grande de comportamentos, seria muito difícil estudar em tempo hábil cada uma das ver-

tentes iniciadas aqui, mas como sempre tomamos o cuidado de mencionar, todos os resultados observados abrem caminho para novos e interessantes estudos na área de redes. Pretendemos abordar cada um deles no longo caminho que temos pela frente na missão de tentar contribuir na compreensão da intrincada relação entre a estrutura e função em redes complexas.

REFERÊNCIAS

- 1 COMIN, C. H. et al. Unveiling the relationship between structure and dynamics in complex networks. Disponível em: < http://arxiv.org/PS_cache/arxiv/pdf/1109/1109.2963v1.pdf>. Acesso em: 30 dez. 2011.

- 2 MARRO, J.; DICKMAN, R. *Nonequilibrium phase transitions in lattice models*. Cambridge: Cambridge University Press, 2005.

- 3 COSTA, L. F. et al. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. *Advances in Physics*, v. 60, n. 3, p. 329–412, 2011.

- 4 NEWMAN, M. E. J. *Networks: an introduction*. Oxford: Oxford University Press, 2010.

- 5 DOROGOVTSSEV, S. N.; MENDES, J. F. F. Evolution of networks. *Advances in Physics*, v. 51, n. 4, p. 1079–1187, 2002.

- 6 DOROGOVTSSEV, S. N.; MENDES, J. F. F. *Evolution of networks: from biological nets to the internet and WWW*. Oxford: Oxford University Press, 2003.

- 7 WATTS, D. J.; STROGATZ, S. H. Collective dynamics of small-world networks. *Nature*, v. 393, p. 440–442, 1998.

- 8 BARABÁSI, A. L.; ALBERT, R. Emergence of scaling in random networks. *Science*, v. 286, n. 5439, p. 509–512, 1999.

- 9 MILGRAN, S. The small world problem. *Psychology Today*, v. 2, p. 60–67, 1967.

- 10 PRESS, W. H. Flicker noises in astronomy and elsewhere. *Comments on Astrophysics*, v. 7, n. 4, p. 103–119, 1978.

- 11 BAK, P.; TANG, C.; WIESENFELD, K. Self-organized criticality: An explanation of the $1/f$ noise. *Physical Review Letters*, v. 59, n. 4, p. 381–384, Jul 1987.

- 12 NEWMAN, M. E. J. Power laws, pareto distributions and zipf’s law. *Contemporary Physics*, v. 46, p. 323–351, 2005.

- 13 CLAUSET, A.; SHALIZI, C. R.; NEWMAN, M. E. J. Power-law distributions in empirical data. *SIAM Review*, v. 51, n. 4, p. 661–703, 2009.
- 14 BARABÁSI, A. L. *Linked: how everything is connected to everything else and what It means*. New York: Plume, 2003.
- 15 ERDŐS, P.; RÉNYI, A. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, v. 5, p. 17–61, 1960.
- 16 BOCCALETTI, S. et al. Complex networks: structure and dynamics. *Physics Reports*, v. 424, n. 4-5, p. 175–308, 2006.
- 17 DOROGOVTSSEV, S. N.; GOLTSEV, A. V. Critical phenomena in complex networks. *Review of Modern Physics*, v. 80, n. 4, p. 1275–1335, 2008.
- 18 ARENAS, A. et al. Synchronization in complex networks. *Physics Reports*, v. 469, n. 3, p. 93–153, 2008.
- 19 BARRAT, A.; BARTHELEMY, M.; VESPIGNANI, A. *Dynamical processes on complex networks*. Cambridge: Cambridge University Press, 2008.
- 20 EASLEY, D.; KLEINBERG, J. *Networks, crowds, and markets*. Cambridge: Cambridge University Press, 2010.
- 21 JACKSON, M. O. *Social and economic networks*. Princeton: Princeton University Press, 2008.
- 22 WILSON, H. R.; COWAN, J. D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, v. 12, n. 1, p. 1–24, 1972.
- 23 ROXIN, A.; RIECKE, H.; SOLLA, S. A. Self-sustained activity in a small-world network of excitable neurons. *Physical Review Letters*, v. 92, p. 198101, May 2004.
- 24 PASTOR-SATORRAS, R.; VESPIGNANI, A. Epidemic spreading in scale-free networks. *Physical Review Letters*, v. 86, p. 3200–3203, Apr 2001.
- 25 STROGATZ, S. H. From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D*, v. 143, n. 1, p. 1–20, 2000.
- 26 COSTA, L. F. et al. Characterization of complex networks: a survey of measurements. *Advances in Physics*, v. 56, n. 1, p. 167–242, 2007.
- 27 BAXTER, R. J. *Exactly solved models in statistical mechanics*. Mineola: Dover Publications, 2008.

- 28 PIKOVSKY, A.; ROSENBLUM, M.; KURTHS, J. *Synchronization: a universal concept in nonlinear sciences*. Cambridge: Cambridge University Press, 2003.
- 29 NES, J. G.-G.; MORENO, Y.; ARENAS, A. Paths to synchronization on complex networks. *Physical Review Letters*, v. 98, p. 034101, Jan 2007.
- 30 SABIDUSSI, G. The centrality index of a graph. *Psychometrika*, v. 31, n. 4, p. 581–603, 1966.
- 31 FRIEDKIN, N. E. Theoretical foundations for centrality measurement. *The American Journal of Sociology*, v. 96, n. 6, p. 1478–1504, 1991.
- 32 BEAUCHAMP, M. A. An improved index of centrality. *Behavioral Science*, v. 10, n. 2, p. 161–163, 1965.
- 33 FREEMAN, L. C. A set of measures of centrality based on betweenness. *Sociometry*, v. 40, n. 1, p. 35–41, 1977.
- 34 KITSACK, M. et al. Identification of influential spreaders in complex networks. *Nature Physics*, v. 6, p. 888–893, 2010.
- 35 COMIN, C. H.; COSTA, L. F. Identifying the starting point of a spreading process in complex networks. *Physical Review E*, v. 84, p. 056105, Nov 2011.
- 36 ALBERT, R.; JEONG, H.; BARABÁSI, A. L. Error and attack tolerance of complex networks. *Nature*, v. 406, n. 6794, p. 378–382, 2000.
- 37 MOTTER, A. E.; LAI, Y.-C. Cascade-based attacks on complex networks. *Physical Review E*, v. 66, p. 065102, Dec 2002.
- 38 KARRER, B.; NEWMAN, M. E. J. Competing epidemics on complex networks. *Physical Review E*, v. 84, p. 036106, Sep 2011.
- 39 ROSS, S. M. *Stochastic processes*. 2nd ed. New Jersey: Wiley, 1995.
- 40 BULLMORE E.AND SPORNS, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, v. 10, n. 3, p. 186–198, 2009.
- 41 ZHOU, C. et al. Hierarchical organization unveiled by functional connectivity in complex brain networks. *Physical Review Letters*, v. 97, p. 238103, Dec 2006.
- 42 JOLLIFFE, I. T. *Principal component analysis*. 2nd ed. New York: Springer, 2002.

- 43 BIGGS, N. L.; LLOYD, E. K.; WILSON, R. J. *Graph theory 1736-1936*. Oxford: Oxford University Press, 1999.
- 44 WATKINS, J. J. *Across the board: the mathematics of chessboard problems*. Princeton: Princeton University Press, 2007.
- 45 BOLLOBÁS, B. *Random graphs*. Cambridge: Cambridge University Press, 2001.
- 46 KNUTH, D. E. *Art of computer programming*. 3rd ed. Reading: Addison-Wesley Professional, 1997.
- 47 BONA, M. *A walk through combinatorics: an introduction to enumeration and graph theory*. 2nd ed. New Jersey: World Scientific, 2006.
- 48 JENSEN, T. R.; TOFT, B. *Graph coloring problems*. New York: Wiley-Interscience, 1994.
- 49 BOLLOBÁS, B. *Modern graph theory*. New York: Springer, 1998. (Graduate texts in mathematics, 184).
- 50 SETHNA, J. P. *Statistical mechanics: entropy, order parameters and complexity*. Oxford: Oxford University Press, 2006.
- 51 REICHL, L. E. *A modern course in statistical physics*. New York: Wiley-Interscience, 1998.
- 52 BOGUNÁ, M.; SATORRAS, R. P.; VESPIGNANI, A. Absence of epidemic threshold in scale-free networks with degree correlations. *Physical Review Letters*, v. 90, p. 028701, Jan 2003.
- 53 BOGUNÁ, M.; CASTELLANO, C.; SATORRAS, R. P. Langevin approach for the dynamics of the contact process on annealed scale-free networks. *Physical Review E*, v. 79, p. 036110, Mar 2009.
- 54 BIANCONI, G.; BARABÁSI, A. L. Competition and multiscaling in evolving networks. *Europhysics Letters*, v. 54, n. 4, p. 436–442, 2001.
- 55 BARABÁSI, A. L.; ALBERT, R.; JEONG, H. Mean-field theory for scale-free random networks. *Physica A*, v. 272, p. 173–187, 1999.
- 56 DOROGOVTSSEV, S. N.; MENDES, J. F. F.; SAMUKHIN, A. N. Structure of growing networks with preferential linking. *Physical Review Letters*, v. 85, p. 4633–4636, Nov 2000.
- 57 BARTHÉLEMY, M. Spatial networks. *Physics Reports*, v. 499, n. 1-3, p. 1–101, 2011.

- 58 WATTS, D. J. Uniform spatial graphs. In: _____. *Small worlds*. Princeton: Princeton University Press, 1999. cap. 3.2.1, p. 93–99.
- 59 WATTS, D. J. *Small worlds*. Princeton: Princeton University Press, 1999.
- 60 SEN, P.; BANERJEE, K.; BISWAS, T. Phase transitions in a network with a range-dependent connection probability. *Phys. Rev. E*, v. 66, p. 037102, 2002.
- 61 PETERMANN, T.; ; PAOLO, d. L. R. Spatial small-world networks: a wiring-cost perspective. Disponível em: < http://arxiv.org/PS_cache/condmat/pdf/0501/0501420v2.pdf>. Acesso em: 30 dez. 2011.
- 62 WAXMAN, B. M. Routing of multipoint connections. *IEEE Journal on Selected Areas in Communications*, v. 6, n. 9, p. 1617–1622, 1988.
- 63 CoCoMac. Dusseldorf, 2008. Disponível em: <http://www.biological-networks.org/?page_id=25>. Acesso em: 06 jan. 2012.
- 64 KAISER, M.; HILGETAG, C. C. Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. *PLoS Computational Biology*, v. 2, n. 7, p. 805–815, 2006.
- 65 CHEN, B. L.; HALL, D. H.; CHKLOVSKII, D. B. Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences*, v. 103, n. 12, p. 4723–4728, 2006.
- 66 VARSHNEY, L. R. et al. Structural properties of the *c. elegans* neuronal network. *PLoS Computational Biology*, v. 7, n. 2, p. 1–21, 2011.
- 67 Neuronal wiring. New York, 2011. Disponível em: <<http://www.wormatlas.org/neuronalwiring.html>>. Acesso em: 06 jan. 2012.
- 68 PASTOR-SATORRAS, R.; VÁZQUEZ, A.; VESPIGNANI, A. Dynamical and correlation properties of the internet. *Physical Review Letters*, v. 87, p. 258701, Nov 2001.
- 69 IZHIKEVICH, E. M. Which model to use for cortical spiking neurons? *IEEE Transaction on Neuronal Networks*, v. 15, n. 5, p. 1063–1070, 2004.
- 70 GERSTNER, W.; KISTLER, W. M. *Spiking Neuron Models*. Cambridge: Cambridge University Press, 2002.
- 71 RIEKE, F.; WARLAND D AND STEVENINCK, R. D. R. a. W. *Spikes: exploring the neural code*. Cambridge: The MIT Press, 1999.

- 72 HODGKIN, A. L.; HUXLEY, A. F. A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. *Journal of Physiology*, v. 117, n. 4, p. 500–544, 1952.
- 73 ROSE, R. M.; HINDMARSH, J. L. The assembly of ionic currents in a thalamic neuron. i the three-dimensional model. *Proceedings of the Royal Society B*, v. 237, n. 1288, p. 267–288, 1989.
- 74 MORRIS, C.; LECAR, H. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal*, v. 35, n. 1, p. 193–213, 1981.
- 75 FITZHUGH, R. Impulses and physiological states in models of nerve membrane. *Biophysical Journal*, v. 1, n. 6, p. 445–466, 1961.
- 76 LAPICQUE, L. Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarization. *Journal de physiologie et de pathologie générale*, v. 9, n. 6, p. 620–635, 1907.
- 77 BRUNEL, N.; ROSSUM, M. C. W. V. Quantitative investigations of electrical nerve excitation treated as polarization. *Biological Cybernetics*, v. 97, n. 5, p. 341–349, 2007.
- 78 ABBOTT, L. F.; VREESWIJK, C. V. Asynchronous states in networks of pulse-coupled oscillators. *Physical Review E*, v. 48, p. 1483–1490, Aug 1993.
- 79 GERSTNER, W. Time structure of the activity in neural network models. *Physical Review E*, v. 51, p. 738–758, Jan 1995.
- 80 KINOUCHI, O.; COPELLI, M. Optimal dynamical range of excitable networks at criticality. *Nature Physics*, v. 2, n. 5, p. 348–351, 2006.
- 81 SOFTKY, W.; KOCH, C. The highly irregular firing pattern of cortical cells is inconsistent with temporal integration of random epsps. *Journal of Neuroscience*, v. 13, n. 1, p. 334–350, 1993.
- 82 SHADLEN, M. N.; NEWSOME, W. T. Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, v. 4, n. 4, p. 569–579, 1994.
- 83 MANWANI, A.; KOCH, C. Detecting and estimating signals in noisy cable structures, i: neuronal noise sources. *Neural Computation*, v. 11, n. 8, p. 1797–1829, 1999.
- 84 BRESSLOFF, P. C. Synaptically generated wave propagation in excitable neural media. *Physical Review Letters*, v. 82, p. 2979–2982, Apr 1999.

- 85 KAISER, M.; HILGETAG, C. C. Optimal hierarchical modular topologies for producing limited sustained activation of neural networks. *Frontiers in Neuroinformatics*, v. 4, n. 8, p. 1–14, 2010.
- 86 WANG, Q. et al. Synchronization transitions on small-world neuronal networks: effects of information transmission delay and rewiring probability. *Europhysics Letters*, v. 83, n. 5, p. 50008, 2008.
- 87 PERC, M. Stochastic resonance on excitable small-world networks via a pacemaker. *Physical Review E*, v. 76, p. 066203, Dec 2007.
- 88 COSTA, L. F. Avalanches of activation and spikes in neuronal complex networks. Disponível em: < http://uk.arxiv.org/PS_cache/arxiv/pdf/0802/0802.0421v1.pdf>. Acesso em: 30 dez. 2011.
- 89 MIROLLO, R. E.; STROGATZ, S. H. Synchronization of pulse-coupled biological oscillators. *Siam Journal of Applied Mathematics*, v. 50, n. 6, p. 1645–1662, 1990.
- 90 MCCULLOCH, W.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, v. 5, n. 4, p. 115–133, 1943.
- 91 IBARZ, I.; CASADO, J. M.; SANJUAN, M. A. F. Map-based models in neuronal dynamics. *Physics Reports*, v. 501, n. 1-2, p. 1–74, 2011.
- 92 WINFREE, A. Biological rhythms and the behavior of populations of coupled oscillators. *Journal of Theoretical Biology*, v. 16, n. 1, p. 15–42, 1967.
- 93 ACEBRÓN, J. A. et al. The kuramoto model: a simple paradigm for synchronization phenomena. *Review of Modern Physics*, v. 77, n. 1, p. 137–185, 2005.
- 94 KURAMOTO, Y. Self-entrainment of a population of coupled nonlinear oscillators. *International Symposium on Mathematical Problems in Theoretical Physics*, v. 39, p. 420–422, 1975.
- 95 GÓMEZ-GARDENES, J. et al. Explosive synchronization transitions in scale-free networks. *Physical Review Letters*, American Physical Society, v. 106, p. 128701, Mar 2011.
- 96 PERKEL, D. H.; GERSTEIN, G. L.; MOORE, G. P. Neuronal spike trains and stochastic point processes: I. the single spike train. *Biophysical Journal*, v. 7, n. 4, p. 391–418, 2008.
- 97 COVER, T. M.; THOMAS, J. A. *Elements of information theory*. 2nd ed. New York: Wiley-Interscience, 2006.

- 98 STRONG, S. P. et al. Entropy and information in neural spike trains. *Physical Review Letters*, v. 80, p. 197–200, Jan 1998.
- 99 SAKAGUCHI, H.; SHINOMOTO, S.; KURAMOTO, Y. Local and global self-entrainments in oscillator lattices. *Progress in Theoretical Physics*, v. 77, n. 5, p. 1005–1010, 1987.
- 100 STUDENT. The probable error of a mean. *Biometrika*, v. 6, n. 1, p. 1–25, 1908.
- 101 JOAN, F. B. Guinness, gosset, fisher, and small samples. *Statistical Science*, v. 2, n. 1, p. 45–52, 1987.
- 102 HOTELLING, H. The generalization of student's ratio. *Annals of Mathematical Statistics*, v. 2, n. 3, p. 360–378, 1931.
- 103 FUKUNAGA, K. *Introduction to statistical pattern recognition*. San Diego: Academic Press, 1990.
- 104 DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification*. New York: Wiley-Interscience, 2000.
- 105 MIRKIN, B. *Core concepts in data analysis: summarization, correlation and visualization*. London: Springer, 2011.
- 106 WARD, J. H. J. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, v. 58, n. 301, p. 236–244, 1963.
- 107 PEARSON, K. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, v. 2, n. 6, p. 559–572, 1901.
- 108 HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, v. 24, n. 7, p. 417–441, 1933.
- 109 LOVÁSZ, L. *Random walks on graphs: a survey*. Hungary: Janos Bolyai Mathematical Society, 1993. (Combinatorics, Paul Erdos is Eighty, v. 1).
- 110 NOH, J. D.; RIEGER, H. Random walks on complex networks. *Physical Review Letters*, v. 92, p. 118701, Mar 2004.
- 111 MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. *Multivariate analysis*. London: Academic Press, 1980.
- 112 NEWMAN, M. E. J.; BARKEMA, G. T. *Monte carlo methods in statistical physics*. Oxford: Oxford University Press, 1999.

- 113 EGUILUZ, V. M. et al. Scale-free brain functional networks. *Physical Review Letters*, v. 94, p. 018102, Jan 2005.
- 114 KAISER, M. Brain architecture: a design for natural computation. *Philosophical Transactions of the Royal Society A*, v. 365, n. 1861, p. 3033–3045, 2007.
- 115 WHITE, J. G. et al. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London. Series B*, v. 314, n. 1165, p. pp. 1–340, 1986.
- 116 WARE, R. W. et al. The nerve ring of the nematode *Caenorhabditis elegans*: sensory input and motor output. *The Journal of Comparative Neurology*, v. 162, n. 1, p. 71–110, 1975.
- 117 WOOD, W. B. *The nematode Caenorhabditis elegans*. New York: Cold Spring Harbor Laboratory Press, 1988. (The Cold Spring Harbor Monograph Series).
- 118 LIU, Y.-Y.; SLOTINE, J.-J.; BARABÁSI, A. L. Controllability of complex networks. *Nature*, v. 473, n. 7346, p. 167–173, 2011.
- 119 REICHARDT, J.; BORNHOLDT, S. Detecting fuzzy community structures in complex networks with a potts model. *Physical Review Letters*, v. 93, p. 218701, Nov 2004.
- 120 WU, F. Y. The potts model. *Review of Modern Physics*, v. 54, p. 235–268, Jan 1982.
- 121 NEWMAN, M. E. J.; GIRVAN, M. Finding and evaluating community structure in networks. *Physical Review E*, v. 69, p. 026113, Feb 2004.
- 122 LEICHT, E. A.; NEWMAN, M. E. J. Community structure in directed networks. *Physical Review Letters*, v. 100, p. 118703, Mar 2008.
- 123 KIM, Y.; SON, S.-W.; JEONG, H. Finding communities in directed networks. *Physical Review E*, v. 81, p. 016103, Jan 2010.
- 124 FRUCHTERMAN, T. M. J.; REINGOLD, E. M. Graph drawing by force-directed placement". *Software: practice and experience*, v. 21, n. 11, p. 1129–1164, 1991.
- 125 SON, S.-W.; JEONG, H.; NOH, J. D. Random field ising model and community structure in complex networks. *European Physical Journal B*, v. 50, n. 3, p. 431–437, 2006.
- 126 ZHOU, H. Distance, dissimilarity index, and network community structure. *Physical Review E*, v. 67, p. 061901, Jun 2003.
- 127 LATAPY, M.; PONS, P. Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications*, v. 10, n. 2, p. 191–218, 2006.

128 ARENAS, A.; DIAZ-GUILERA, A.; PÉREZ-VICENTE, C. J. Synchronization reveals topological scales in complex networks. *Physical Review Letters*, v. 96, p. 114102, Mar 2006.

129 ARENAS, A.; DIAZ-GUILERA, A. Synchronization and modularity in complex networks. *The European Physical Journal*, v. 143, n. 1, p. 19–25, 2007.

130 HENDRICKS, W. A.; ROBEY, K. W. The sampling distribution of the coefficient of variation. *Annals of Mathematical Statistics*, v. 7, n. 3, p. 129–132, 1936.

APÊNDICE A – Tempo para a estabilização da dinâmica

Durante esse trabalho utilizamos três medidas sobre a dinâmica integra-e-dispara, definidas na seção 2.2.3, que devem ser tomadas em um tempo longo o suficiente para garantir que a dinâmica já tenha passado pela fase de transiente. Aqui mostraremos o nível de variação que essas medidas apresentaram para o intervalo de tempo que consideramos. Na Figura A.1 mostramos a ativação da rede, R , definida na equação 2.14 para o modelo geográfico de Waxman e os mesmos parâmetros dinâmicos utilizados durante o trabalho. Exercitamos 100 realizações da dinâmica para 100 redes distintas e calculamos o valor médio em cada iteração. Vemos que, ao menos globalmente, a dinâmica parece convergir rapidamente para um patamar com pequenas flutuações de baixa significância. A janela que utilizamos para calcular os valores médios começa na iteração 600 e se estende até a iteração 1100, e portanto ela é definida em um longo tempo após a transição de R .

Uma forma mais precisa de verificar se não há muita variação nos valores medidos é, basicamente, obter as mesmas medidas para outras janelas definidas em iterações anteriores as de fato utilizadas. O que fazemos é definir uma janela de largura 500 (a mesma largura que utilizamos), e deslizarmos essa janela ao longo do tempo, começando em $t_w = 0$ e terminando em $t_w = 600$. Por exemplo, em $t_w = 0$ nossa janela começa em $t = 0$ e termina em $t = 500$, em $t_w = 150$ a janela vai de $t = 150$ a $t = 650$ e assim por diante. Para cada t_w obtemos o valor das medidas dinâmicas consideradas, e verificamos se elas não possuem uma variação significativa entre janelas subsequentes.

Na Figura A.2(a) mostramos a média da taxa, r , calculada para 100 redes geográficas distintas, onde vemos que esse valor praticamente não varia ao longo das diversas janelas. Entretanto

Figura A.1 – Ativação dos nós na dinâmica integra-e-dispara ao longo do tempo. A rede utilizada segue o modelo de Waxman e os parâmetros dinâmicos são os mesmos utilizados no texto principal.

Figura A.2 – Média e variação das três medidas consideradas para diferentes janelas utilizadas no cálculo. Novamente os parâmetros são os mesmos utilizados no texto principal para a rede geográfica.

o valor médio não é uma medida precisa do que está acontecendo a nível de nó na dinâmica, pois ele é basicamente uma medida global, que inclusive se assemelha muito ao parâmetro R . Uma outra forma (ainda não ideal) de medirmos a variação é tirarmos o módulo da diferença (ou derivada) da medida x obtida para $t_w = t$ e $t_w = t + \delta t$, e normalizarmos o valor obtido pelo valor da medida, isto é

$$\delta x = \frac{x(t_w + \delta t) - x(t_w)}{x(t_w)} \quad (\text{A.1})$$

a quantidade obtida indica o quanto a medida está variando em relação ao seu valor. No contexto estatístico esse valor é conhecido como coeficiente de variação (130). Na Figura A.2(b) mostramos o valor médio e desvio padrão de δr ao considerarmos todos os nós de uma única rede. O fato de utilizarmos apenas uma rede é devido à suavização que ocorre em δr ao realizarmos a média em um conjunto. Os mesmos valores típicos foram observados em outras redes seguindo o mesmo modelo. Vemos que δr típico é da ordem de 0.005 em relação ao valor da medida, e o desvio padrão possui picos da ordem de 0.01, o que consideramos uma variação insignificante, e portanto os valores medidos de r praticamente não variariam se escolhêssemos outra janela. Nas Figuras A.2(c) e (d) mostramos os valores obtidos para a entropia de intervalos (ISE), onde vemos que o valor médio da medida novamente não varia, mas o desvio de δISE possui valores maiores do que δr . Entretanto consideramos que esse desvio ainda é insignificante. As Figuras A.2(e) e (f) mostram o mesmo processo feito sobre o maior intervalo (I^0), onde vemos que essa medida possui a menor variação entre as três medidas.

Nas Figuras A.3 e A.4 mostramos todo o processo aplicado em redes geradas pelo modelo ER, onde vemos que todas as variações foram tipicamente menores ou iguais ao modelo geográfico, e portanto consideramos que as medidas foram corretamente obtidas. Finalmente observamos que no modelo BA as variações foram ainda menores do que no modelo ER.

Figura A.3 – Ativação dos nós na dinâmica integra-e-dispara ao longo do tempo para o modelo ER e os mesmos parâmetros dinâmicos utilizados no texto principal.

Figura A.4 – Média e variação das três medidas consideradas para diferentes janelas utilizadas no cálculo. As redes seguem o modelo ER.