**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS**

JOÃO LUIZ BUNORO BATISTA

**Dinâmica de opinião de Krause-Hegselmann em
redes complexas**

São Carlos

2012

JOÃO LUIZ BUNORO BATISTA

Dinâmica de opinião de Krause-Hegselmann em redes complexas

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 Ciência.

Área de Concentração: Física Aplicada
Opção: Física Computacional
Orientador: Prof. Dr. 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)

Bunoro Batista, João Luiz
Dinâmica de Krause-Hegselmann em redes complexas /
João Luiz Bunoro Batista; orientador Luciano da
Fontoura Costa -- São Carlos, 2012.
103 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. Dinâmicas de opinião. 3.
Fenômenos coletivos. 4. Krause-Hegselmann. I. da
Fontoura Costa, Luciano, orient. II. Título.

FOLHA DE APROVAÇÃO

João Luiz Bunoro Batista

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: 28/11/2012

Comissão Julgadora

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

Instituição: IFSC/USP

Prof(a). Dr(a). Sérgio Furtado dos Reis

Instituição: UNICAMP/Campinas

Prof(a). Dr(a). Zhao Liang

Instituição: ICMC/USP

AGRADECIMENTOS

- Aos meus pais, e à minha irmã, Mariana, pela convivência, educação e confiança.
- Ao professor Luciano, pela orientação, dedicação, paciência e, principalmente, por toda experiência e conhecimentos transmitidos
- Ao Cesar e ao Thomas, pelas longas discussões sobre redes, pela amizade e pelo incentivo na conclusão dos trabalhos.
- Ao André, pela amizade desde os anos da graduação e pela ajuda com Latex nos momentos finais deste texto.
- Ao Berté, por me aturar por vários anos, pelos debates da madrugada e pela amizade.
- À Maria Luiza, pela confiança e intensa convivência nos últimos anos.
- Aos físicos dos tempos de graduação, Thiago (Yoko), Rafael (Melão), Jaqueline, Willian (Alf) e Tiago (Loro).
- Aos amigos de Grupo, Matheus, Lucas, Renato, Diego, Débora, Gustavo e João Florindo.
- Aos do professores do IFSC, principais responsáveis pela minha formação recente.

RESUMO

BATISTA, J.L.B. **Dinâmica de opinião de Krause-Hegselmann em redes complexas.** 2012. 103p. Dissertação (Mestrado em ciência) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2012.

Fenômenos coletivos em redes sociais como a formação de linguagem ou cultura, crenças, emergência de consenso em relação a algum assunto, aquisição de conhecimento e aprendizagem, dentre outros, tem conduzido a um grande interesse no estudo de comportamentos cooperativos e fenômenos sociais, resultando numa grande variedade de dinâmicas de opinião. Nestes modelos, uma população de agentes interagentes carrega uma variável (ou um conjunto delas) numérica cujo valor representa uma “opinião” sobre um tópico, com interpretações distintas em cada contexto. Inspirados em conceitos de mecânica estatística e mecanismos sociais, estes estados evoluem governados por regras matemáticas que controlam a dinâmica de interação entre os agentes e a influência de fatores externos. Outro ingrediente importante na modelagem de sistemas reais, é que a representação das interações entre agentes difere bastante de reticulados ou misturas homogêneas, sendo mais bem descritas por redes complexas. Neste trabalho, estudamos a dinâmica de opinião de Krause e Hegselmann. Neste modelo, agentes possuem opiniões que assumem valores contínuos e são atualizados de acordo com a vizinhança compatível, definida pelo princípio da *confiança limitada*. Após apresentar uma revisão da literatura, estudamos a dinâmica de opinião no contexto de Redes Complexas, seguido de modificações do modelo que consideram a ação de ruído e campo externo (propaganda). Finalmente, propomos um modelo de consenso cuja interpretação está inserida no contexto de aquisição de conhecimento por agentes interagentes que realizam observações sujeitas a erros. Os resultados mostram como os diferentes tipos de topologia influenciam no comportamento das dinâmicas.

Palavras-chave: Redes complexas. Dinâmicas de opinião. Fenômenos coletivos. Krause-Hegselmann.

ABSTRACT

BATISTA, J.L.B. **Opinion dynamics of Krause-Hegselmann on complex networks** 2012. 103p. Dissertação (Mestrado em ciência) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2012.

Collective phenomena in social networks such as formation of language or culture, beliefs, emergence of consensus on any subject, knowledge acquisition and learning, among others, has led to an increasing interest in the study of cooperative behavior and social phenomena, resulting in great variety of opinion dynamics. In these models, a population of interacting agents holds a variable (or a set of them) whose numerical value is an opinion on a topic, with different interpretations in each context. Inspired by concepts from statistical mechanics and social mechanisms, these states evolve governed by mathematical rules that control the dynamics of interaction between agents and the influence of external factors. Another important ingredient in the modeling of real systems is the representation of the interactions between agents, which strongly differs from lattices or fully mixed states, being better described by complex networks. In the present work, we study the opinion dynamics of Krause and Hegselmann. In this model, agents hold opinions that assume continuous values and are updated according to their compatible neighborhood, defined by the *bounded confidence* principle. After presenting a literature review, we studied the opinion dynamics in the context of complex networks, followed by modifications of the model considering the effect of noise and external field (advertising). Finally, we propose a consensus model interpreted as a process of knowledge acquisition by interacting agents that make observations subject to errors. The results show how the topology influences the dynamic behavior.

Keywords: Complex networks. Opinion dynamics. Collective phenomena. Krause-Hegselmann.

LISTA DE FIGURAS

Figura 1.1 - Divisão conceitual entre dinâmicas de opinião baseada no tipo de variável mantida pelos agentes.	21
Figura 2.1 - Mínimo caminho médio e coeficiente de clusterização para o modelo Watts-Strogatz. Média sobre 500 realizações de redes com tamanho $N = 10000$	26
Figura 2.2 - Dependência do coeficiente de clusterização em termos do parâmetro β para o modelo social de Boguñá. Redes construídas num espaço de dimensão $d = 1$ com tamanho $N = 500$. Média sobre 50 realizações. . .	29
Figura 2.3 - Propriedades, segregadas por grau, verificadas para o modelo social de Boguñá para valores de β iguais à 1.5, 2.5 e 3.5 conforme indicado na Figura. (a) coeficiente de clusterização médio e (b) grau médio dos primeiros vizinhos. Redes construídas num espaço de dimensão $d = 1$ com tamanho $N = 500$. Média sobre 50 realizações.	29
Figura 2.4 - Scatter-plot entre a acessibilidade e o tempo de exploração para 10^5 vetores de tamanho $N = 6$ gerados aleatoriamente. As linhas indicam os casos onde as probabilidades são divididas em dois grupos de mesmo valor, conforme discutido no texto.	32
Figura 3.1 - Evolução temporal das opiniões no modelo de KH com 300 agentes: (a) $\varepsilon = 0.25$, (b) $\varepsilon = 0.15$ e (c) $\varepsilon = 0.05$. Opiniões iniciais são sorteadas no intervalo uniforme $[0,1]$	35

Figura 3.2 - Evolução temporal das opiniões no modelo KH com 300 agentes para valores de ε iguais à (a) 0.1 e (c) 0.3. Figuras (b) e (d) mostram a variação das opiniões dos agentes entre passos de tempo sucessivos. Escala de cores indica a posição inicial dos em $t = 0$	36
Figura 3.3 - Diagrama de bifurcação para a dinâmica de KH. Espaço de opiniões e de ε discretizados em 500 e 100 partes, respectivamente.	40
Figura 4.1 - Comparação da evolução temporal da dinâmica de KH para uma rede ER, de grau médio $\langle k \rangle = 10$, e grafo completo, ambas de tamanho $N = 500$, conforme indicado na Figura.	43
Figura 4.2 - Número de clusters obtidos com realizações da dinâmica de KH em termos de ε para redes do tipo (a) ER, (b) BA e (c) LA com $\langle k \rangle = 10$ e tamanhos conforme indicado na Figura. Média sobre 200 realizações das redes e da dinâmica.	44
Figura 4.3 - Distribuição do tamanho dos clusters obtidos para redes ER e LA de $\langle k \rangle = 10$ e tamanho $N = 1000$ para (a) $\varepsilon = 0.1$ e (b) $\varepsilon = 0.03$. Média sobre 100 realizações das redes e da dinâmica.	45
Figura 4.4 - Tamanhos do primeiro e segundo maior cluster obtidos com realizações da dinâmica de KH em termos de ε para redes do tipo (a) ER, (b) BA e (c) LA com $\langle k \rangle = 10$ e tamanhos conforme indicado na Figura. Setas identificam transições de fase na estrutura da rede de opiniões, conforme discussão no texto. Média sobre 200 realizações das redes e da dinâmica.	47
Figura 4.5 - Tempo para estabilização da dinâmica de KH em termos de ε para redes (a) ER, (b) BA e (c) LA com $\langle k \rangle = 10$ e tamanhos conforme indicado na Figura. Média sobre 200 realizações das redes e da dinâmica.	49
Figura 4.6 - Exemplo de evolução temporal da dinâmica de KH próximo da transição <i>monopolar-multipolar</i> para rede ER de tamanho $N = 500$ e $\langle k \rangle = 10$	50

Figura 4.7 - Tamanho do cluster central em termos do parâmetro ε para rede ER. Média sobre 50 realizações de redes e da dinâmica.	51
Figura 4.8 - Mínimo caminho médio entre os vértices que compõem o maior cluster de opiniões em termos de ε . A seta identifica o ponto de máximo para esta medida e está relacionada com a transição para a fase fragmentada. Simulação considerando redes ER de 750 e 1000 vértices com $\langle k \rangle = 10$. Média sobre 50 realizações da dinâmica de KH.	52
Figura 4.9 - Efeito do grau médio na posição da transição para a fase fragmentada. (a) Mínimo caminho médio entre os vértices que compõem o maior cluster de opiniões considerando redes do tipo ER com diferentes valores de $\langle k \rangle$ (grau médio varia de 5.0 a 15.0, conforme indicado na Figura) em termos do parâmetro de confiança ε . (b) Relação entre os pontos de máximo obtidos em (a) para cada rede e os respectivos valores de $\langle k \rangle$. Média sobre 50 realizações das redes e da dinâmica de KH.	55
Figura 4.10 - Mínimo caminho médio entre os vértices que compõem o maior cluster de opiniões em termos de ε para redes ER e BA com $N = 1000$ e $\langle k \rangle = 10$. Valores de ε próximos da região de transição para a <i>fase fragmentada</i> . Média sobre 50 realizações da dinâmica de KH.	56
Figura 5.1 - Exemplos de realizações da dinâmica de KH na presença de ruído. Séries temporais mostram a evolução das opiniões ao longo do tempo para valores de α iguais a 0.2 (coluna da esquerda) e 0.6 (coluna da direita) e ε iguais a 0.1, 0.3 e 0.5 (de cima para baixo). Níveis de cinza indicam a concentração de agentes ao longo do espaço de opiniões. Resultados obtidos para um grafo completo com $N = 1000$	60

Figura 5.2 - Distribuições estacionárias para o modelo de KH na presença de ruído. Resultados obtidos a partir da solução numérica da equação de campo médio (Eq. 5.1) reescrita em termos da cadeia da Markov interativa (Eq. 5.2). Valores de α variam de 0.1 a 0.6 (esquerda para direita, de cima para baixo). Espaço de opiniões e de ε discretizados em 100 partes. . . . 61

Figura 5.3 - Coeficiente de dispersão, G_M , obtido para a dinâmica de KH em termos da intensidade do ruído, α , e valores do parâmetro de confiança, ε , conforme indicado na Figura. Foi considerado grafo completo com $N = 1000$ agentes. Média sobre 50 realizações da dinâmica e espaço de opiniões discretizado em 100 partes. 62

Figura 5.4 - Coeficiente de dispersão, G_M , em termos da intensidade do ruído, α , para os modelos de rede (a) ER e (b) o modelo social SN de tamanhos N igual 1000. Valores do parâmetro de confiança conforme indicado na Figura. Média sobre 100 realizações das redes e da dinâmica e espaço de opiniões discretizado em 100 partes. 64

Figura 5.5 - Comparação da evolução temporal das opiniões no modelo KH para redes ER e SN, ambas com mesmo tamanho $N = 500$ e grau médio $\langle k \rangle = 10$ sujeitas a $\varepsilon = 0.15$. Parâmetro β do modelo SN igual 3. Níveis de cinza estão relacionadas com a densidade de agentes ao longo do espaço de opiniões. 65

Figura 5.6 - Exemplos de realizações da dinâmica de KH na presença de ruído temporário ($\Delta T = 10$ passos de simulação) para valores de ε iguais a 0,1 e 0,2 conforme indicado na figura. Níveis de cinza indicam a concentração de agentes ao longo do espaço de opiniões. 66

Figura 5.7 - Tempo de convergência da dinâmica de KH em termos do tempo ΔT de aplicação de ruído. Resultados para redes ER e SN considerando valores de ε e intensidade do ruído, α , conforme indicados na Figura. Média sobre 50 realizações das redes e da dinâmica.	66
Figura 5.8 - Relação entre o número de clusters sem aplicação de ruído e após aplicação de ruído durante intervalo de tempo ΔT . Resultados para redes ER e SN com valores de ε e intensidade do ruído, α , conforme indicados na Figura. Média sobre 50 realizações das redes e da dinâmica.	68
Figura 5.9 - Relação entre o tamanho do maior cluster de opinião sem aplicação de ruído e após aplicação de ruído com relação ao parâmetro ΔT . Resultados para redes ER e SN com valores de ε e intensidade do ruído α conforme indicados na Figura. Média sobre 50 realizações das redes e da dinâmica.	68
Figura 5.10 - Tendência de manutenção das opiniões S após aplicação de ruído durante intervalo de tempo ΔT . Resultados para redes ER e SN com valores de ε e intensidade do ruído, α , conforme indicados na Figura. Média sobre 50 realizações das redes e da dinâmica.	69
Figura 5.11 - Diagrama de bifurcação com as distribuições estacionárias para o modelo de KH na presença de campo externo considerando diferentes valores da opinião B . Resultados obtidos a partir da solução da Equação 5.7 reescrita em termos da Cadeia de Markov em 5.8. Valores de B variam de 0.1 a 0.9 (esquerda para direita, de cima para baixo) com $\gamma = 0.5$. Espaço de opiniões e de ε discretizados em 100 partes.	72

Figura 5.12 - Diagrama de bifurcação com as distribuições estacionárias para o modelo de KH na presença de campo externo considerando diferentes valores de γ . Resultados obtidos a partir da solução da Equação 5.7 reescrita em termos da Cadeia de Markov em 5.8. Valores de γ variam de 0.1 a 0.9 (esquerda para direita, de cima para baixo). Espaço de opiniões e de ε discretizados em 100 partes. 73

Figura 5.13 - Exemplos da evolução temporal da dinâmica de KH sob efeito de campo externo com intensidades $\gamma = 0.8$ (esquerda) e $\gamma = 0.1$ (direita). Foi considerada rede ER de tamanho $N = 500$ e $\langle k \rangle = 10$. Utilizado parâmetro de confiança $\varepsilon = 0.2$ e campo na posição $B = 0.25$ (linha vermelha). . . 74

Figura 5.14 - Dependência da efetividade do campo, E_B (parte superior) e do número de clusters, $N_{clusters}$ (parte inferior) para redes ER e SN de tamanho $N = 500$ e $\langle k \rangle = 10$ com campo mantido na opinião $B = 0.25$. Média sobre 50 realizações das redes e da dinâmica. 76

Figura 5.15 - Exemplos da evolução temporal da dinâmica de KH numa rede SN quando campo, na posição $B = 0.5$, é aplicado sobre agente de maior (esquerda) e menor grau (direita). Considerada rede de tamanho $N = 300$ e $\langle k \rangle = 10$ para $\varepsilon = 0.15$. Linha vermelha e verde indicam a posição do campo e do agente afetado, respectivamente. Níveis de cinza indicam a concentração de agentes ao longo do espaço de opiniões. 77

Figura 5.16 - Dependência da efetividade do campo, E_B , em relação à (a) acessibilidade, (b) grau e (c) betweenness do agente sob efeito da opinião externa. Considerado modelo de rede SN de tamanho $N = 300$ e $\langle k \rangle = 10$ 78

Figura 6.1 - Agentes (cientistas) desenvolvem seus modelos através de observações e interações. 81

Figura 6.2 - Erro global em termos do desvio padrão dos graus para rede ER e BA. Probabilidades de erro nas observações perfeitamente correlacionadas com o grau. Cada ponto do gráfico representa a realização de um modelo de rede com $N_a = 200$ e $\langle k \rangle = 5.8$ 87

Figura 6.3 - Erro global em termos do grau dos agentes com erro de observação diferente dos demais para os modelos ER e BA com $N_a = 200$ and $\langle k \rangle = 9.4$: (a) um dos agentes com taxa de erro igual a 0.2 e os outros agentes com $\gamma = 0.4$; e (b) um dos agentes com 0.4 e os demais com $\gamma = 0.2$. A linha pontilhada representa a solução analítica. Resultados obtidos com médias sobre 100 realizações da dinâmica. 89

Figura 6.4 - Erro global em termos do grau dos agentes com maiores erros de observação: um dos agentes com erro proporcional ao seu grau e todos os outros agentes com $\gamma = 0.2$. A linhas pontilhadas e contínuas representam a solução analítica para os modelos ER e BA, respectivamente. Simulações para uma única realização de cada modelo de rede com $N_a = 200$ and $\langle k \rangle = 9.4$. Resultados obtidos com médias sobre 50 realizações da dinâmica. 90

Figura 6.5 - Erros individuais em termos do grau dos agentes no caso em que agente com maior grau possui error de observação igual a 0.4 e todos os outros com $\gamma = 0.2$ a and c): $\square$, o agente com a maior taxa de erro; $\bullet$, a vizinhança deste agente; $\circ$, demais agentes. Figuras (b e c) mostram o valor médio dos erros individuais para o conjunto de agentes com a mesma distância topológica do agente com maior taxa de erro. Simulações para modelos ER e BA com $N_a = 300$ e $\langle k \rangle = 4.7$. Resultados obtidos a partir de médias sobre 50 realizações da dinâmica. 91

Figura 6.6 - Evolução temporal dos erros individuais considerando o agente com o menor (a) e maior (b) grau tendo um erro de observação igual a 0.4 e todos os outros agentes com $\gamma = 0.2$. Acima: todos os agentes são considerados (linha pontilhada representa o a agente com maior erro). Abaixo: somente os agentes com a mesma taxa de erro (linhas pretas) e o valor médio de seus erros individuais. Simulações para modelo BA com $N_a = 200$ and $\langle k \rangle = 9.4$ 92

Figura 6.7 - Dependência entre o número de arestas adicionadas na comunidade com maior erro de observação e o valor médio dos erros individuais dos agentes nas outras comunidades para redes com $N_a = 256$ agentes: (a) todos os agentes com o mesmo grau, igual a 16, agrupados em 4 comunidades de mesmo tamanho e variando o parâmetro de mistura; (b) todos os agentes com o mesmo grau, igual a 16, agrupados em 5 comunidades com diferentes tamanhos para $\mu = 0.2$; (c) agentes com graus distribuídos como lei potência e agrupados em 4 comunidades de mesmo tamanho considerando $\mu = 0.2$. Resultados obtidos a partir de médias sobre 100 realizações da dinâmica. 94

SUMÁRIO

1	INTRODUÇÃO	19
2	Redes Complexas	23
2.1	Primeiros conceitos	23
2.2	Modelos de redes	24
2.2.1	Modelo de Erdős-Rényi	24
2.2.2	Modelo de Watts-Strogatz	26
2.2.3	Modelo de Barabási-Albert	27
2.2.4	Modelo social de Boguñá	28
2.3	Acessibilidade	30
3	O modelo de Confiança Limitada de Krause-Hegselmann	33
3.1	Apresentação do modelo	33
3.2	Descrição da dinâmica de Krause-Hegselmann baseada na densidade de agentes	35
3.2.1	Solução em termos de Cadeias de Markov interativas	38
4	A dinâmica de Krause-Hegselmann em Redes Complexas	41
4.1	Primeiras diferenças	42
4.2	Avaliação de parâmetros de ordem e identificação de transições de fase estruturais	42

4.3	Relação com a Teoria de Percolação	49
5	Modificações da dinâmica de Krause-Hegselmann	57
5.1	Ruído	58
5.1.1	Resultados para o grafo completo	58
5.1.2	Efeito do ruído na dinâmica de KH em Redes Complexas	63
5.2	Aplicação de campo constante	69
5.2.1	Resultados para o Grafo Completo	70
5.2.2	Efeito do campo externo na dinâmica de KH em Redes Complexas . . .	73
6	Aquisição de conhecimento na presença de erros de observação	79
6.1	Definições	80
6.2	Descrição analítica para o Erro Global	84
6.3	Características gerais	86
6.3.1	Efeito produzido por um único agente	88
6.3.2	Variando erros de observação em estruturas com comunidades	91
7	Conclusão	95
	Referências	97
	Apêndice A – TRABALHOS PUBLICADOS DURANTE O MESTRADO	101

1 INTRODUÇÃO

O efeito conhecido como *Wisdom of Crowd* (“Sabedoria das Massas”, numa tradução livre) mostra como os valores médios de estimativas de vários indivíduos podem ser mais precisos do que se fossem realizadas por especialistas individualmente. Este efeito foi primeiramente descrito, de maneira científica, num texto publicado em 1907 na revista *Nature* por Francis Galton (1). No trabalho, intitulado *Vox Populi*, o autor descreve como, num concurso onde apostadores sem grande conhecimento tentam estimar o peso de animais, a mediana das estimativas mostra-se surpreendentemente próxima do valor correto. De fato, diversas outras publicações mais recentes (2–4) estudam as consequências desse efeito em vários outros contextos dentre os quais, na determinação dos preços em mercados de negociação, em eleições, em sistemas de busca ou ferramentas colaborativas na Web. Em contrapartida, em (5, 6) são documentados os resultados para alguns experimentos controlados realizados da seguinte forma: indivíduos são convidados à fornecerem respostas quantitativas à respeito de questões de conhecimento geral; em seguida, os mesmos indivíduos revisam suas respostas tendo conhecimento do valor médio das estimativas do grupo. Após sucessivas rodadas de revisão, o que se verifica é uma piora na qualidade da estimativa média. O motivo para isso, conforme argumentado em (6), é que o efeito “*Wisdom of Crowd*” não é um fenômeno social, mas estatístico, baseado na heterogeneidade das decisões individuais. Assim, a interação indireta entre os indivíduos acrescenta um ingrediente (social) adicional. Como consequência, rodada após rodada, as opiniões dos indivíduos tendem a se aproximar e a dispersão das estimativas reduz drasticamente, mas o valor médio se distancia do valor correto.

Esta tendência à aproximação de opiniões ilustrada no parágrafo anterior, tem sua origem

na influência, ou pressão, social exercida por um indivíduo sobre o outro. Na literatura, este efeito é conhecido como “*Efeito Manada*” (do inglês “Herd Effect”) (7). De maneira mais geral, o processo de imitação entre indivíduos numa sociedade aparece em muitos contextos, dentre os quais, em processos de negociação, no surgimento de partidos políticos, no processo de aquisição de conhecimento, no aparecimento da linguagem, de crenças, de religiões. Desta forma, a sociedade pode ser vista, como um sistema complexo, no qual fenômenos coletivos emergem a partir de dois ingredientes básicos: i) a interação entre os humanos e ii) o comportamento individual.

Na tentativa de melhor compreensão deste sistema, uma crescente quantidade de novos estudos tem sido realizados desde a última década. Muitos destes trabalhos tentam recriar artificialmente as propriedades macroscópicas observadas na sociedade através da simulação de regras locais que indiquem a maneira como indivíduos se comportam. Tipicamente, a modelagem destes problemas envolve a criação de uma população na qual indivíduos possuem uma *opinião*, representada por uma variável numérica e interagem, por exemplo, seguindo uma maioria local ou imitando um vizinho. Obviamente, estas regras microscópicas são simplificações extremas do comportamento humano e a compreensão de descrições mais realistas, com envolvimento de maior quantidade de parâmetros, demonstra uma das (muitas) dificuldades da investigação do problema. Embora não seja exatamente uma área nova de pesquisa (8), o crescente interesse se deve ao fato do aumento da disponibilidade de mais (e maiores) bases de dados, dos recursos computacionais, e, principalmente, do envolvimento de cientistas de diversas áreas do conhecimento. Como referência, uma busca simples pelos termos “*opinion dynamics*” e “*social dynamics*” na ferramenta “*Web of Science*” retorna uma quantidade de trabalhos igual 38, no ano de 2001, e 186, em 2012, ou seja, um aumento de quase 400%.

Os modelos de opinião são normalmente divididos em dois grupos (Figura 1.1). No primeiro, as variáveis carregadas pelos indivíduos assumem um conjunto finito de valores discretos. Exemplos de dinâmicas incluem os modelos Votante (9, 10) e de Sznajd (11), a Regra da Maioria (12) e a Teoria do Impacto Social (13). No segundo grupo, por sua vez, as opiniões variam continuamente e seus valores são representados por números reais. Nesta última categoria,

os modelos mais conhecidos são os modelos de Deffuant (14) e de Krause-Hegselmann (15), objeto de estudo deste trabalho. Nestas duas dinâmicas está presente a hipótese da *Confiança Limitada*, a qual estabelece que indivíduos aos pares podem influenciar um ao outro somente se suas opiniões forem suficientemente próximas, especificado por um valor limiar ε . Uma revisão geral sobre uma grande variedade de dinâmicas de opinião pode ser encontrada em (8).

Figura 1.1 – Divisão conceitual entre dinâmicas de opinião baseada no tipo de variável mantida pelos agentes.

No modelo de Deffuant, agentes, dentro do intervalo de confiança, interagem aos pares. As opiniões inclinam-se uma em direção à outra através da média com peso μ ($0 \leq \mu \leq 1$), sendo que para valores baixos de μ , os agentes são pouco receptivos à aproximação. Conforme pode ser verificado no trabalho original, a inspiração para o modelo de Deffuant vêm de uma dinâmica um pouco mais antiga estudada por Axelrod (16). Ela se baseia em opiniões multidimensionais, onde cada componente representa alguma característica cultural, e a evolução ocorre de maneira acoplada, de modo que indivíduos distantes socialmente deixam de interagir.

Paralelamente, outros autores exploraram o conceito proposto por Deffuant em trabalhos semelhantes. Por exemplo, em (17), descreve-se o chamado *modelo de compromisso*, também no contexto de dinâmicas sociais. As regras de evolução das opiniões são exatamente as mesmas definidas no trabalho de Deffuant. No entanto, ao invés do problema ser estudado através de simulações de Monte Carlo baseada em agentes, as soluções do problema são buscadas através de uma descrição via *campo médio*. De fato, estes mesmos autores, em (18), 4 anos antes, num contexto mais teórico, descrevem um modelo de interação de partículas num gás cuja regra de evolução dos momentos nada mais é do que o modelo de Deffuant na ausência do parâmetro de confiança.

Já na dinâmica de Krause-Hegselmann, cada agente atualiza sua opinião considerando não apenas um indivíduo, mas toda a vizinhança com opiniões compatíveis. Partindo de uma configuração aleatória, em ambos os modelos o que se verifica é a formação de clusters de opinião. Dependendo da quantidade destes clusters, o sistema pode atingir um estado de *consenso* global, ou alternativamente de *polarização* ou *fragmentação*.

Outro aspecto relevante considerado nestas dinâmicas, e fundamental na modelagem de sistemas reais, é o fato de indivíduos interagirem com um número limitado de outros agentes. Mais do que isso, dados empíricos mostram que redes sociais são bastante pouco regulares apresentando distribuições de conectividade heterogêneas, presença de comunidades, efeito *Mundo Pequeno*, ciclos de diversas ordens, dentre outros. O estudo destas características, não limitado às redes sociais, é endereçado numa área bastante produtiva da ciência atual: *Redes Complexas* (19). De fato, a topologia apresenta papel importante na evolução de dinâmicas que ocorrem em redes.

O objetivo deste trabalho, portanto, engloba: i) a revisão de alguns dos principais elementos de Redes Complexas e dos conceitos envolvidos na modelagem de dinâmicas de opinião, com foco exclusivo no modelo de Hegselmann-Krause; ii) compreensão da influência da estrutura das redes nos processos dinâmicos e iii) a incorporação de ingredientes do mundo real na evolução de opiniões.

Os demais capítulos desta dissertação seguem a seguinte estrutura. O Capítulo 2 traz uma revisão sobre alguns dos principais modelos de geração de Redes Complexas, bem como apresenta uma medida para caracterização da diversidade topológica, a *acessibilidade*. No Capítulo 3, detalhamos a dinâmica original de Krause-Hegselmann e apresentamos uma solução baseada em Cadeias de Markov. O Capítulo 4 inicia a etapa de resultados mostrando as características da dinâmica em redes complexas e sua relação com a teoria de percolação. No Capítulo 5, estudamos a ação de ingredientes adicionais no modelo de Krause, o efeito de ruído e de um campo constante, e seus efeitos em redes sociais. Finalmente, no Capítulo 6, propomos uma dinâmica de consenso, na qual agentes interagentes fazem estimativas sujeitas à erros.

2 *REDES COMPLEXAS*

2.1 Primeiros conceitos

Sistemas complexos, sintéticos ou naturais, são formados por um grande número de elementos interagentes os quais podem ser naturalmente representados pelos vértices de um grafo. Embora tradicionalmente a Física estatística tenha lidado com essas redes considerando-as como estruturas homogêneas ou reticulados, na última década grande atenção se voltou para o estudo de estruturas com características heterogêneas (ou Redes Complexas), motivadas pelas propriedades de sistemas reais.

A Teoria de Redes Complexas (20, 21) é fruto da intersecção dos conceitos de Teoria de Grafos e Mecânica Estatística, muitas vezes inspirados em problemas de caráter multidisciplinar. O interesse pela área tem sido demonstrado devido ao fato de redes reais não terem seu comportamento explicado por um conectividade aleatória. Ao contrário, estas redes apresentam leis de potência, hubs, estrutura de comunidades, dentre outras propriedades. Seu sucesso ocorre principalmente por sua flexibilidade e seu caráter ao mesmo tempo geral e reducionista na representação de sistemas reais. Mais do que isso, as redes estudadas são vistas como estruturas onde processos dinâmicos ocorrem influenciados por sua topologia. Associado a esses fatores, uma crescente capacidade computacional facilita a aquisição de grandes bases de dados bem como possibilita o processamento destas.

Uma rede é um grafo definido por um conjunto N de nós (ou vértices), um conjunto M de arestas (ou conexões) e um mapeamento de como estes conjuntos estão interligados. Se

identificarmos cada nó por um valor discreto ($i = 1, 2, 3, \dots, N$), este mapa pode ser simplificado através de uma matriz $\mathbf{A}$, conhecida por *matriz adjacência*, cuja entrada $A[i, j]$, não nula, representa uma conexão entre os nós i e j . Uma rede é dita ser *dirigida* se as conexões possuem direção, caso contrário, a rede é *não dirigida* e a matriz adjacência é uma estrutura simétrica.

Uma imensa variedade de medidas estruturais podem ser definidas de maneira a caracterizar a topologia das redes (22), tanto a nível microscópico (vértices e arestas), como mesoscópico (sub-grafos) e macroscópico (considerando toda a rede). Estas medidas podem considerar: i) grau (distribuição de grau, grau máximo, assortatividade); ii) ciclos (coeficiente de clusterização, comunicabilidade); iii) distâncias (mínimo caminho médio, vulnerabilidade); iv) medidas de centralidade (betweenness, closeness); v) comunidades (modularidade, distribuição dos tamanhos dos clusters); vi) medidas em arestas (reciprocidade, matching index); vii) medidas hierárquicas, dentre outras. Além disso, a utilização destas medidas associadas á métodos de reconhecimento de padrões constitui uma ferramenta poderosa para a classificação de redes reais e artificiais (22).

2.2 Modelos de redes

Associado às medidas citadas na Seção anterior, modelos de construção de redes são propostos na tentativa de reproduzir propriedades de interesse. A seguir, descrevemos alguns dos principais modelos.

2.2.1 Modelo de Erdős-Rényi

Trata-se, provavelmente, do primeiro modelo de rede para geração de grafos aleatórios, em oposição à estrutura fixas, como reticulados. Este modelo foi proposto em 1959 pelos ma-

temáticos húngaros Paul Erdős e Alfréd Rényi (21). Neste modelo, uma rede é construída a partir de um conjunto de N vértices desconectados. Cada possível conexão é realizada com probabilidade p . Assim, em média, cada vértice nesta rede estará conectado à $\langle k \rangle = (N-1)p \simeq Np$. Obviamente, a probabilidade de um vértice possuir k conexões segue a distribuição binomial, independente do valor de N :

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k} \quad (2.1)$$

No limite em que $N \rightarrow \infty$ mantendo-se o grau médio $\langle k \rangle$ fixo, a Equação anterior pode ser aproximada pela distribuição de Poisson:

$$P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}. \quad (2.2)$$

Do ponto de vista topológico, este modelo foi amplamente estudado, em especial no que diz respeito à estrutura dos subgrafos produzidos. É possível mostrar que para a probabilidade crítica $p_c = 1/N$ a estrutura da rede muda abruptamente, saindo de uma configuração esparsa, cujo maior subgrafo é uma árvore, para uma configuração mais complexa, caracterizada pela emergência de um componente gigante.

Tipicamente, estas redes apresentam distâncias pequenas entre os vértices de modo que o mínimo caminho médio escala com o logaritmo do tamanho da rede: $ASPL \sim \log(N)$. As pequenas distâncias verificadas justificam o efeito *Mundo Pequeno* (23) presente neste modelo. Embora esta seja uma propriedade presente em diversas reais, o modelo não é capaz de gerar altos valores para o coeficiente de clusterização, o qual tende à zero quando os tamanhos das redes divergem.

2.2.2 Modelo de Watts-Strogatz

A motivação para este modelo é a captura, simultânea, da propriedade *Mundo Pequeno* presente em redes aleatórias e dos altos valores do coeficiente de clusterização observados em redes reais. De acordo com o modelo original (24), a construção de um grafo se inicia com um reticulado circular com N nós, cada um conectado aos m vizinhos mais próximos em cada direção, num total de $2m$ conexões iniciais para cada vértice. Em seguida, cada aresta é aleatoriamente reconectada com probabilidade p .

A Figura 2.1 mostra os valores para o mínimo caminho médio e para o coeficiente de clusterização observados em redes geradas a partir do modelo. Veja que temos duas situações extremas: para $p = 0$, a rede é um reticulado perfeito com altos valores para o coeficiente de clusterização e para o mínimo caminho médio. Já para $p = 1$, a rede assemelha-se a um grafo aleatório e ambas as medidas apresentam valores baixos. Para valores intermediários ($p \sim 0.05$), no entanto, tanto distâncias típicas pequenas quanto grande quantidade de loops são verificadas. A distribuição é semelhante à encontrada em redes aleatórias, com pico em $\langle k \rangle = 2m$.

Figura 2.1 – Mínimo caminho médio e coeficiente de clusterização para o modelo Watts-Strogatz. Média sobre 500 realizações de redes com tamanho $N = 10000$.

2.2.3 Modelo de Barabási-Albert

Enquanto os dois primeiros modelos apresentados produzem vértices com um grau característico, o modelo proposto por Albert-Lászlo Barabási e Reka Albert em 1999 (25) procura produzir em grafos sintéticos distribuição de grau na forma de leis de potência ($P(k) \sim k^{-\gamma}$) e, conseqüentemente, uma fração de vértices com altíssimo grau. Essa característica de fato se mostra presente numa grande quantidade de redes reais como a rede física da Internet, redes de colaboração científica, redes metabólicas, redes de interação de proteínas, dentre outras.

O modelo de construção é baseado em dois ingredientes: 1) crescimento e 2) ligação preferencial. Inicialmente, parte-se de uma configuração com N_0 vértices. Em seguida, em cada passo, um vértice é adicionado e conectado à rede com m ($m \leq N_0$) arestas. A probabilidade do novo vértice se conectar a qualquer nó i presente na rede é proporcional ao grau de i .

$$\Pi(k_i) \sim k_i / \sum_j k_j. \quad (2.3)$$

Desta forma, nós mais conectados tendem a receber ainda mais conexões (paradigma “ricos ficam mais ricos”), o que tende a ocorrer com os nós mais antigos. Embora este modelo capture alguns dos ingredientes presentes em redes reais ele necessita de modificações para reproduzir apropriadamente os valores do expoente γ (as redes geradas pela regra acima possuem $\gamma = 3$ quando $N \rightarrow \infty$) e os altos valores de coeficiente de clusterização. Além disso, adaptações do mecanismo de ligação preferencial podem ser feitas considerando a idade dos vértices, o custo da adição de novas arestas, entre outros fatores, produzindo desvios da lei de potência. Uma revisão detalhada sobre as diversas modificações do modelo pode ser encontrada em (20).

2.2.4 Modelo social de Boguña

O objetivo deste modelo (26) é a reprodução de duas importantes propriedades presentes em redes sociais: i) o alto valor para o coeficiente de clusterização e ii) a assortatividade. Diferente do que ocorre para os três modelos de rede descritos até agora, redes reais tendem a apresentar correlações de grau, isto é, vértices tendem a se conectar preferencialmente com outros vértices de grau similar. Redes são ditas *assortativas*, se apresentam correlação de grau positiva, caso contrário, são chamadas *dissortativas*. Empiricamente, sabe-se que redes sociais tendem a apresentar comportamento assortativo, deferente de redes tecnológicas, por exemplo, que a apresentam correlação negativa de graus.

O modelo é construído assumindo que pares de vértices são conectados de acordo com sua distância social. Assim, considere N vértices desconectados, os quais são aleatoriamente colocados num espaço de dimensão n . Seja $\vec{h}_i$ o vetor que define a posição do i -ésimo vértice neste espaço. A probabilidade de dois vértices i e j estarem conectados é definida da seguinte forma:

$$r_n(\vec{h}_i, \vec{h}_j) = \frac{1}{1 + \left[b^{-1} d_n(\vec{h}_i, \vec{h}_j) \right]^\beta}, \quad (2.4)$$

onde b é uma escala característica, responsável por ajustar o grau médio e β é um parâmetro que controla a tendência de conexões entre indivíduos semelhantes.

A Figura 2.2 mostra a dependência do coeficiente de clusterização, CC em termos do parâmetro β para redes geradas num espaço unidimensional. Note que, CC vai à zero quando $\beta \rightarrow 1$ e satura em $3/4$ para $\beta \rightarrow \infty$.

A Figura 2.3(a) complementa a avaliação sobre o coeficiente de clusterização, mostrando o valor médio desta medida em função do grau. Note que $\overline{CC}$ é, aproximadamente, independente de k para os três valores de β avaliados. A Figura 2.3(b) mostra a relação entre o grau médio

Figura 2.2 – Dependência do coeficiente de clusterização em termos do parâmetro β para o modelo social de Boguñá. Redes construídas num espaço de dimensão $d = 1$ com tamanho $N = 500$. Média sobre 50 realizações.

dos primeiros vizinhos para vértices de grau k . Note que $\bar{k}_{nn}(k)$ cresce monotonicamente com k apresentando curvas mais acentuadas à medida que o valor de β aumenta.

Figura 2.3 – Propriedades, segregadas por grau, verificadas para o modelo social de Boguñá para valores de β iguais à 1.5, 2.5 e 3.5 conforme indicado na Figura. (a) coeficiente de clusterização médio e (b) grau médio dos primeiros vizinhos. Redes construídas num espaço de dimensão $d = 1$ com tamanho $N = 500$. Média sobre 50 realizações.

2.3 Acessibilidade

Para um dado um nó fonte i , assumamos ser possível alcançar $N_i(h)$ diferentes através de caminhadas aleatórias de tamanho h partindo do i . Então, podemos dizer que i possui N_i vizinhos alcançáveis à distância h . Cada vizinho é alcançado com uma probabilidade diferente, a qual é representada pelo vetor $\mathbf{p}_i^{(h)} = \{p_1^{(h)}, p_2^{(h)}, \dots, p_{N_i(h)}^{(h)}\}$. A heterogeneidade de acessos pode ser definida através de uma medida de entropia, a qual chamaremos acessibilidade. Assim para um nó i , na escala h , temos:

$$\kappa_i(h) = \exp \left(- \sum_j p_j^{(h)} \log p_j^{(h)} \right). \quad (2.5)$$

De acordo com a definição, os valores de acessibilidade ficam contidos no intervalo $[1, N_i(h)]$, sendo o valor máximo obtido para o caso homogêneo, quando todas as probabilidades possuem o mesmo valor $1/N_i(h)$.

Para melhor caracterizar a acessibilidade, podemos relacionar esta medida com o famoso problema do colecionador de cupons. A questão levantada por este problema (27), no contexto de redes, pode ser visto da seguinte forma: em média, quantas caminhadas de tamanho h partindo de um vértice i são necessárias de modo que todos os vizinhos de i à distância h sejam visitados ao menos uma vez? Chamaremos esta quantidade *tempo de exploração* do vértice i , $\tau_i(h)$. A solução deste problema pode ser colocada da seguinte forma:

$$\tau_i(h) = \int_0^\infty \left(1 - \prod_{j=1}^{N_i(h)} \left(1 - e^{-x p_j^{(h)}} \right) \right) dx \quad (2.6)$$

Conjectura-se (28) que τ_i alcança seu valor mínimo para o caso homogêneo, onde todos os vértices são alcançados com a mesma probabilidade, isto é, $p_j^{(h)} = 1/N_i(h)$ para todo j . Neste caso, não é difícil mostrar que:

$$\tau_i^{hom}(h) = N_i(h) \sum_{m=1}^{N_i(h)} \frac{1}{m}. \quad (2.7)$$

Assumindo a conjectura citada acima, podemos dizer que a acessibilidade é máxima sempre que o tempo de exploração é mínimo. A Figura 2.4 mostra esta característica através de um *scatter-plot* entre a acessibilidade, κ , e o tempo de exploração $\tau(h)$, para 10^5 vetores $\mathbf{p}$ de dimensão $N = 6$ aleatoriamente gerados (para cada componente do vetor um valor é escolhido da distribuição uniforme e, em seguida, normalização é imposta). Na mesma Figura uma série de curvas caracterizando as configurações de probabilidades são mostradas. Elas correspondem ao casos onde exatamente $n (\leq \lfloor N/2 \rfloor)$ probabilidades de $\mathbf{p}$ tem um valor μ , enquanto todas as outras $(N - n)$ probabilidades são também iguais entre si, respeitando a condição de que a soma das probabilidades é igual a 1. Assim, a linha contínua está relacionada ao caso onde $n = 1$, de modo que $N - 1$ probabilidades possuem o mesmo valor. Além disso, esta curva corresponde ao valor limite da acessibilidade dado $\tau(h)$, mostrando que todas possíveis configurações de probabilidade $\mathbf{p}$ estão contidas por esta curva. A linha pontilhada corresponde à configuração onde $n = 2$, enquanto a linha central corresponde ao caso onde $n = 3$, ou seja, metade das probabilidades são iguais entre si. Em termos do parâmetro μ , e usando as equações 2.5 e 2.6, podemos obter uma equação geral para caracterizar estas curvas:

$$\tau_C(\mu) = \frac{1}{\mu} \sum_{m=1}^n \frac{1}{m} + \frac{1}{p} \sum_{m=1}^{N-n} \frac{1}{m} - \sum_{m=1}^d \sum_{m'=1}^{N-n} (-1)^{m+m'} \binom{n}{m} \binom{N-n}{m'} \frac{1}{m\mu + m'p} \quad (2.8)$$

e

$$\kappa_C(\mu) = \frac{1}{p} \left(\frac{p}{\mu} \right)^{\mu n}, \quad (2.9)$$

onde $p = (1 - n\mu)/(N - n)$. Observe que μ assume valores no intervalo $[0, 1/n]$. Quando $\mu < 1/N$, a parte superior das curvas é obtida. Neste caso, temos $\kappa_C \rightarrow N - n$ e $\tau_C \rightarrow \infty$ para $\mu \rightarrow 0$. Para $\mu > 1/N$, obtemos a parte inferior das curvas, para as quais $\kappa_C \rightarrow n$ e $\tau_C \rightarrow \infty$ quando $\mu \rightarrow 1/n$. Quando $\mu = 1/N$, acessibilidade é máxima e o tempo de exploração é

mínimo. As setas na Figura indicam a direção na qual μ cresce. Note ainda que, para $n = 3$, não há diferença entre a parte superior e inferior das curvas, já que exatamente metade das probabilidades tem valor igual a μ .

Figura 2.4 – Scatter-plot entre a acessibilidade e o tempo de exploração para 10^5 vetores de tamanho $N = 6$ gerados aleatoriamente. As linhas indicam os casos onde as probabilidades são divididas em dois grupos de mesmo valor, conforme discutido no texto.

3 O MODELO DE CONFIANÇA LIMITADA DE KRAUSE-HEGSELMANN

3.1 Apresentação do modelo

A dinâmica de Kaue-Hegselmann (KH)(15) é definida da seguinte forma: considere um vetor $O \in (\Re_{\geq 0})^n$, cuja entrada $O_i(t)$ indica a opinião do agente $i \in 1, \dots, n$ no tempo $t \in \mathbb{N}$. Ao longo do tempo, um agente i revisa sua opinião recursivamente levando em conta as opiniões dos agentes cuja opinião não estejam separadas por uma distância maior que ε_i , chamado *parâmetro de confiança*. Em notação matricial, podemos escrever como:

$$O(t+1) = \mathbf{H}(O(t), \varepsilon_1, \dots, \varepsilon_n)O(t), \quad (3.1)$$

onde $\mathbf{H}(O(t), \varepsilon_1, \dots, \varepsilon_n)$ é uma matriz $n \times n$, chamada *matriz confiança*, e definida como

$$\mathbf{H}(O(t), \varepsilon_1, \dots, \varepsilon_n)[i, j] := \begin{cases} \frac{1}{\#I_{\varepsilon_i}(i, \mathbf{x})} & \text{if } j \in I_{\varepsilon_i}(i, \mathbf{x}) \\ 0 & \text{caso contrário} \end{cases} \quad (3.2)$$

sendo $I_{\varepsilon_i}(i, \mathbf{x}) := \{j \in \Gamma i; |O_i - O_j| \leq \varepsilon_i\}$. Ou seja, a nova opinião $O_i(t+1)$ é a média aritmética sobre as opiniões dos vizinhos de i que não difiram deste mais do que ε_i .

Se $\varepsilon_1 = \dots = \varepsilon_n = \varepsilon$, o modelo é chamado *homogêneo*, caso contrário, *heterogêneo*. No modelo original de KH (15) é assumida homogeneidade e opiniões estão restritas ao intervalo $[0, 1]$

(opiniões iniciais são sorteadas de maneira uniforme neste mesmo intervalo). Na interpretação de Stauffer (29), agentes no modelo KH são como *oportunistas* uma vez que estes desejam estar no baricentro de todos os agentes de opiniões compatíveis minimizando, assim, suas distâncias em relação aos agentes confiáveis.

Note que no modelo de KH a atualização das opiniões ocorre de maneira simultânea e determinística, ou seja, a partir das condições iniciais, a evolução de $O(t)$ é completamente previsível. É importante destacar ainda que a dinâmica de KH sempre estabiliza e isto ocorre em tempo finito, ou seja, a partir de um tempo t^* o perfil de opiniões O atinge um ponto fixo tal que $O(t^*) = O(t^* + 1) = \dots = O(\infty)$. De fato, conforme demonstrado em (30), três condições suficientes, porém não necessárias, para que um processo genérico, do tipo definido da Equação 3.1 atinja um ponto fixo são:

- 1) A diagonal da *matriz confiança* $\mathbf{H}$ seja positiva, ou seja, para cada agente $i \in 1, \dots, n$, $\mathbf{H}[i, i] > 0$. Isso significa que todo agente deve ter *auto-confiança*;
- 2) Entradas iguais à zero em $\mathbf{H}$ devem ser simétricas e;
- 3) Existe $\delta > 0$ tal que a menor entrada positiva de $\mathbf{H}(t)$ é maior que δ . Ou seja, entradas positivas da matriz $\mathbf{H}$ não convergem para zero.

Veja que para a dinâmica de KH as três condições são satisfeitas.

A Figura 3.1 mostra a evolução temporal de realizações únicas para três valores distintos de ε para $n = 300$ agentes. Veja que o processo converge para uma configuração estacionária com a formação de clusters de mesma opinião (ou seja, $O(t = \infty) = \sum_{k=1}^Q F(Q)\delta(k)$, onde Q é o número de clusters formados e F uma função peso). Intuitivamente, espera-se que se o valor de ε é baixo, pluralidade é alcançada, enquanto que para ε alto consenso é obtido. Valores intermediários de ε levam à configurações polarizadas com dois clusters de tamanhos semelhantes. Essas três situações são ilustradas nas Figuras 3.1(a), 3.1(b) e 3.1(c) para valores

de ε iguais a 0.25, 0.15 e 0.05, respectivamente.

Figura 3.1 – Evolução temporal das opiniões no modelo de KH com 300 agentes: (a) $\varepsilon = 0.25$, (b) $\varepsilon = 0.15$ e (c) $\varepsilon = 0.05$. Opiniões iniciais são sorteadas no intervalo uniforme $[0, 1]$.

Ainda sobre a evolução das opiniões ao longo o tempo, é importante notar que as mudanças no perfil de O iniciam-se nas bordas, com os agentes de opiniões extremas. Isto fica claro com os exemplos mostrados na Figura 3.2. Para valores de ε iguais a 0.1 e 0.25 são mostradas a evolução temporal do perfil de opiniões (Figs. 3.2(a) e 3.2(c)), bem como a diferença $\Delta_i(t)$ entre as opiniões no tempo t e $t - 1$ de cada agente i (Figs. 3.2(b) e 3.2(d)). A escala de cores indica a posição dos agentes em $t = 0$. Como se pode notar, nos primeiros instantes de tempo, os maiores valores absolutos de Δ estão relacionados com os agentes extremos (azul e amarelo). De fato, agentes nas posições centrais sofrem a influência das opiniões de ambas as extremidades resultando em movimentações pequenas, restritas à flutuações aleatórias.

3.2 Descrição da dinâmica de Krause-Hegselmann baseada na densidade de agentes

Na seção anterior definimos o modelo de KH baseado no comportamento microscópico de cada agente. Naquela situação, o objeto de interesse era o valor $O(t)$ de cada uma das opiniões no tempo. A descrição da dinâmica pode ser refeita numa abordagem baseada na densidade de agentes em cada intervalo do espaço de opiniões. Assim, gostaríamos de avaliar a quantidade

Figura 3.2 – Evolução temporal das opiniões no modelo KH com 300 agentes para valores de ϵ iguais à (a) 0.1 e (c) 0.3. Figuras (b) e (d) mostram a variação das opiniões dos agentes entre passos de tempo sucessivos. Escala de cores indica a posição inicial dos em $t = 0$.

$P(o, t)$, a qual representa a densidade de probabilidade para a variável o , bem como sua evolução temporal $\partial/\partial t P(o, t)$, assumindo t uma grandeza contínua. Se $P(o, 0) = 1$ para $0 \leq o \leq 1$, a equação mestra para descrição da dinâmica fica:

$$\frac{\partial P(o, t)}{\partial t} = \int_0^1 dx P(x, t) \left[\delta \left(o - \frac{\int_{x-\epsilon}^{x+\epsilon} dx' x' P(x', t)}{\int_{x-\epsilon}^{x+\epsilon} dx' P(x', t)} \right) - \delta(o - x) \right] \quad (3.3)$$

As duas funções delta-Dirac dentro dos colchetes representam as duas contribuições para a variação de $P(o, t)$ em torno da opinião o no intervalo de tempo dt . O termo de ganho, a razão entre as integrais na primeira delta (a qual denotaremos por $\bar{o}_\epsilon(x)$), nada mais é do que a média das opiniões de todos os agentes compatíveis com x . Assim, sempre que $\bar{o}_\epsilon(x) = o$, novos agentes ocuparão o intervalo de opiniões $[o, o + do]$. Da mesma forma, se $x = o$, haverá uma contribuição negativa para o intervalo $[o, o + do]$.

Como esperado, é fácil notar da equação 3.3 que a dinâmica conserva a população total de

agentes, ou seja, $N(t) = \int_0^1 doP(o,t)$, é uma constante de movimento. De fato, integrando a equação 3.3 sobre a variável o :

$$\int_0^1 do \frac{\partial P(o,t)}{\partial t} = \int_0^1 \int_0^1 do dx P(x,t) [\delta(o - \bar{o}_\varepsilon(x)) - \delta(o - x)], \quad (3.4)$$

vemos que o lado direito é igual a zero, já que a única dependência em o está contida nas duas deltas. E, portanto, $\frac{dN(t)}{dt} = 0$.

Pode-se verificar ainda que o primeiro momento da distribuição também não varia no tempo. Para mostrar esta propriedade devemos assumir que $P(o,t)$ é simétrica com relação à transformação $o \rightarrow 1 - o$ (ou seja, $P(o,t)$ é simétrica com relação à opinião central 0.5). Temos, portanto, que:

$$P(1 - o, t) = P(o, t) \quad (3.5)$$

e

$$\bar{o}_\varepsilon(1 - x, t) = 1 - \bar{o}_\varepsilon(x, t) \quad (3.6)$$

A última igualdade é facilmente obtida aplicando a transformação na definição de $\bar{o}_\varepsilon(x)$. A evolução da média de $P(o,t)$ pode ser calculada multiplicando ambos os lados da eq. 3.3 por o e integrando sobre esta variável no intervalo $[0, 1]$:

$$\int_0^1 doo \frac{\partial P(o,t)}{\partial t} = \int_0^1 \int_0^1 do dx o P(x,t) [\delta(o - \bar{o}_\varepsilon(x)) - \delta(o - x)]. \quad (3.7)$$

Integrando em o , no lado direito:

$$\int_0^1 doo \frac{\partial P(o,t)}{\partial t} = \int_0^1 dx P(x,t) [\bar{o}_\varepsilon(x) - x] \quad (3.8)$$

Aplicando a transformação no lado direito da equação anterior e usando as relações 3.5 e

3.6, obtém-se:

$$\int_0^1 dx P(x, t) [\bar{o}_\varepsilon(x) - x] = \int_1^0 dx P(x, t) [\bar{o}_\varepsilon(x) - x]. \quad (3.9)$$

Desta forma, ambas as integrais são iguais a zero, e, portanto, a opinião média é conservada. Como consequência, se a distribuição inicial de opiniões é uniforme entre 0 e 1, consenso só pode ser obtido quando todos os agentes possuírem mesma opinião igual a 1/2. Além disso, como $P(o, t)$ deve ser simétrica, na existência de mais de clusters, estes devem ser simétricos em relação à $o = 1/2$ e ainda, se o número de clusters for ímpar, então deve haver um cluster em 1/2. Expressões para momentos mais altos podem somente ser obtidos para $\varepsilon = 1$.

3.2.1 Solução em termos de Cadeias de Markov interativas

Embora a evolução temporal, e obviamente a solução estacionária para a Equação 3.3 possam ser obtidas através de métodos de integração tradicionais, reescreveremos o problema numa abordagem em termos de Cadeias de Markov, como descrito em (29). Assim, supondo a discretização do espaço de opiniões em n_d partes, temos a seguinte equação para a evolução temporal da densidade de agentes:

$$P(o, t + 1) = P(o, t) \mathbf{B}(P(o, t), t) \quad (3.10)$$

A cadeia é dita ser *interativa* uma vez que depende da configuração de opiniões ao longo tempo. A matriz transição $\mathbf{B}$ é definida como:

$$\mathbf{B}[i, j] = \begin{cases} 1 & \text{se } j = M_i, \\ \lceil M_i \rceil - M_i & \text{se } j = \lfloor M_i \rfloor, j \neq M_i, \\ \lfloor M_i \rfloor - M_i & \text{se } j = \lceil M_i \rceil, j \neq M_i, \\ 0 & \text{caso contrário} \end{cases} \quad (3.11)$$

A entrada $B[i, j]$ fornece a probabilidade de transição entre os intervalos de opiniões $[\frac{i-1}{n}, \frac{i}{n}]$ e $[\frac{j-1}{n}, \frac{j}{n}]$. Note que cada linha da matriz $\mathbf{B}$ contém apenas 1 ou 2 entradas adjacentes positivas. M_i é a opinião média da massa compatível com a opinião i/n e é obtida como:

$$M_i = \frac{\sum_{m \in n, |i-m| \leq n_d \varepsilon} m P(m)}{\sum_{m \in n, |i-m| \leq n_d \varepsilon} P(m)} \quad (3.12)$$

A Figura 3.3 mostra os resultados obtidos para a solução estacionária da densidade de opiniões, $P(o, t = \infty)$, em termos dos parâmetro de confiança ε , obtida a partir do uso da matriz transição 3.11 considerando $P(o, t = 0) = U([0, 1])$. Esta representação é o chamado *diagrama de bifurcação* e fornece uma visão do comportamento da dinâmica para cada valor de ε (as linhas fornecem a localização dos clusters). Note que uma opinião central sempre mantém-se, ainda que sua densidade varie. Para valores altos de ε consenso sempre é obtido. Conforme o valor do parâmetro de confiança diminui o seguinte padrão repetidamente ocorre, nessa ordem: i) dois clusters simétricos, com distância ε do centro, são criados e a densidade da opinião central diminui; ii) aumento da densidade do cluster central e destruição do par criado e; iii) consolidação da bifurcação com a volta do par de clusters.

Figura 3.3 – Diagrama de bifurcação para a dinâmica de KH. Espaço de opiniões e de ϵ discretizados em 500 e 100 partes, respectivamente.

4 A DINÂMICA DE KRAUSE-HEGSELMANN EM REDES COMPLEXAS

Na seção anterior apresentamos a dinâmica original de *Krause-Hegselmann* (KH) e suas principais características. Esta apresentação esteve restrita à condição de mistura homogênea, na qual cada agente interage com todos os demais, como num grafo completo. Nesta seção, assumiremos uma premissa um pouco mais realista. Iremos verificar o comportamento da dinâmica de KH considerando que os agentes interagem com um número limitado de outros agentes e esta conectividade é representada através de uma rede, onde vértices são os agentes e as arestas definem os possíveis canais de comunicação. De fato, a topologia é um aspecto importante para definição de redes sociais e as particularidades dessas estruturas influenciam diretamente no comportamento de dinâmicas que ocorrem numa rede.

Consideramos redes do tipo Erdős-Rényi (ER), Barabási-Albert (BA) e Lattice circular (LA). Este último obtido do modelo de Watts-Strogatz com probabilidade de troca p igual a zero. Nas redes, a dinâmica de KH deve ser ligeiramente modificada de modo a considerar a presença de links entre os diferentes agentes, ou seja, além do parâmetro de confiança ε , dois agentes i e j não interagem caso a entrada da matriz adjacência $A(i, j)$ seja igual a zero. Assim, para a *matriz confiança* definida no capítulo anterior

$$\mathbf{H}(O(t), \varepsilon_1, \dots, \varepsilon_n)_{ij} := \begin{cases} \frac{1}{\#I_{\varepsilon_i}(i, O)} & \text{if } j \in I_{\varepsilon_i}(i, \mathbf{x}) \\ 0 & \text{caso contrário} \end{cases} \quad (4.1)$$

$I_{\varepsilon_i}(i, \mathbf{x}) := j \in \Gamma_i; |o_i - o_j| \leq \varepsilon_i$. Onde Γ_i representa a vizinhança do vértice i (ou seja, todas as entradas diferentes de zero da i -ésima linha da matriz adjacência).

Importante destacar ainda que todas as redes consideradas nos resultados a seguir são não-dirigidas e fixas nos tamanhos e estruturas ao longo do tempo.

4.1 Primeiras diferenças

A Figura 4.1 compara a evolução temporal da dinâmica de KH para uma rede ER (direita) e um grafo completo (esquerda) considerando valores de ε iguais a 0.25 e 0.1, conforme indicado na Figura. Para os dois casos a dinâmica converge para um estado estacionário no qual os valores de opinião mantêm-se constantes e cuja solução é uma soma de δ 's. No entanto, pelo menos três diferenças importantes podem ser notadas para a evolução da rede ER: i) o tempo para estabilização aumenta substancialmente; ii) o número de clusters (ou δ 's da solução final) também aumenta, em especial nas regiões periféricas do espaço de opiniões e; iii) diferente do que ocorre para grafos completos, a separação entre os clusters (ou δ 's) não é mais limitada ao parâmetro de confiança, ou seja, conforme pode ser notado na Figura, clusters podem apresentar separação menor que ε . Estas diferenças ocorrem por conta da restrição que a estrutura da rede impõe sobre a interação entre os agentes. Como consequência, diferentes topologias podem direcionar para diferentes resultados.

4.2 Avaliação de parâmetros de ordem e identificação de transições de fase estruturais

Apresentamos a seguir alguns estudos sistemáticos do comportamento da dinâmica de KH nas redes em termos do parâmetro de confiança ε . Para tanto, investigamos alguns parâmetros

Figura 4.1 – Comparação da evolução temporal da dinâmica de KH para uma rede ER, de grau médio $\langle k \rangle = 10$, e grafo completo, ambas de tamanho $N = 500$, conforme indicado na Figura.

importantes da dinâmica após a fase transiente, dentre os quais, o número final de clusters ($N_{clusters}$), o tamanho do maior (S_1) e segundo maior cluster (S_2) e o tempo de convergência da dinâmica (T_{CONV}). Nas simulações, o critério utilizado para decisão do fim da fase transiente foi a verificação de mudança não substancial dos valores de opinião: se $\max_i \{|O_i(t^*) - O_i(t^* - 1)|\} < d$, então $t^* := T_{CONV}$, onde d assumimos igual a 10^{-6} .

Conceitualmente, no estado final, clusters são formados por agentes de mesma opinião. No entanto, por construção, estes agentes possuirão opiniões muito próximas, mas não exatamente idênticas, de modo que dois agentes são considerados membros do mesmo cluster se existe um caminho de agentes entre eles na rede tal que a diferença nos valores das opiniões entre agentes consecutivos é menor que o parâmetro de confiança ϵ . Como consequência, pode-se pensar em dois tipos de rede: i) *rede estrutural*, idêntica à rede original formada pelos agentes e suas conexões e ii) *rede de opiniões*, formada pela junção entre a rede original e a diferença de opinião entre os agentes, sendo que dois agentes i e j estão conectados na *rede de opiniões*

se $A(i, j) = 1$ e $|O_i - O_j| < \varepsilon$.

A Figura 4.2 apresenta o número médio de clusters obtidos para as três redes sob estudo considerando diferentes valores de ε . Em todos os casos $N_{clusters}$ apresenta comportamento monotônico e decresce a medida que ε aumenta, convergindo para N quando $\varepsilon \rightarrow 0$ e para 1 quando $\varepsilon \rightarrow 1$. A primeira diferença importante entre as redes ER e BA quando comparadas ao Lattice circular é a dependência funcional com relação a ε . Para os dois primeiros casos pode-se observar comportamento na forma de lei potência, enquanto para a rede LA a dependência é exponencial, ou seja, $N_{clusters}^{ER/BA} \propto \varepsilon^{-\kappa_1}$ e $N_{clusters}^{LA} \propto \kappa_2^{-\varepsilon}$. Desta forma, o número de clusters avança muito mais rapidamente na rede LA do que nas redes ER e BA, permitindo maior proliferação de clusters de tamanho microscópico na primeira rede a medida que o valor de ε diminui.

Figura 4.2 – Número de clusters obtidos com realizações da dinâmica de KH em termos de ε para redes do tipo (a) ER, (b) BA e (c) LA com $\langle k \rangle = 10$ e tamanhos conforme indicado na Figura. Média sobre 200 realizações das redes e da dinâmica.

Para tornar mais claro este último ponto, a Figura 4.3(a) compara a distribuição do tamanho

Figura 4.3 – Distribuição do tamanho dos clusters obtidos para redes ER e LA de $\langle k \rangle = 10$ e tamanho $N = 1000$ para (a) $\epsilon = 0.1$ e (b) $\epsilon = 0.03$. Média sobre 100 realizações das redes e da dinâmica.

dos clusters para as redes ER e LA no ponto $\epsilon = 0.1$. Veja que para a rede LA os clusters são tipicamente muito menores, concentrados em tamanhos que não superam mais do que 10% do tamanho da rede. Já para rede ER observa-se duas partes, uma composta por clusters microscópicos, semelhante ao verificado para rede LA, e outra formada por clusters macroscópicos de tamanhos comparáveis ao tamanho da rede. É preciso lembrar que o Lattice circular possui distribuição de grau igual a uma Delta Dirac ($P^{LA}(k) = \delta(\langle k \rangle)$), enquanto a rede ER apresenta dispersão nesta grandeza. Assim, vértices mais conectados servem como *agregadores* criando clusters maiores e, conseqüentemente, alterando a dependência de $N_{clusters}$ com o parâmetro ϵ .

Ainda com relação à Figura 4.2, a diferença funcional explicitada anteriormente não é válida por toda a extensão de valores de ϵ , conforme pode ser observado numa visualização um pouco mais cuidadosa. Por exemplo, para a rede LA, a dependência exponencial deixa de valer por volta do ponto $p_1 \sim 0.3$. O mesmo ocorre para as redes ER e BA, para as quais $p_1 \sim 0.2$. Este ponto marca o início do processo de consenso nestas redes, ou seja, para valores de ϵ maiores do que aquele indicado pelo ponto p_1 o número de cluster reduz-se significativamente até que exista um único cluster homogêneo. Como consequência, a partir do ponto p_1 , o parâmetro $N_{clusters}$ deixa de ser extensivo com o tamanho N da rede. De fato, deve-se notar que

para valores de ε menores que p_1 , $N_{clusters} \sim N$, independente do tipo de rede.

Na outra extremidade dos valores de ε , para as redes ER e BA, identifica-se ainda um segundo ponto de transição $p_2 \sim 0.04$. Este ponto marca a mudança para um regime no qual não há clusters macroscópicos. Isto fica evidente na Figura 4.3(b) a qual apresenta a distribuição de tamanho dos clusters para $\varepsilon = 0.03$ em redes ER. Note que clusters de tamanho maior que 10% do tamanho da rede tornam-se muito improváveis, semelhante ao que foi verificado para rede LA quando $\varepsilon = 0.1$ (Figura 4.3(a)). Assim, para valores de ε menores que p_2 em redes ER e BA, a dependência em lei de potência com ε deixa de valer e o número de clusters satura com o tamanho da rede. As transições identificadas pelos pontos p_1 e p_2 se tornarão ainda mais evidentes nos resultados a seguir.

A Figura 4.4 mostra a evolução do primeiro (S_1) e segundo (S_2) maiores clusters para as redes (a) ER, (b) BA e (c) LA ao longo do espaço de ε . Para as redes ER e BA, três fases podem ser diretamente observadas com transições identificadas pelos setas 1 e 2. Para o Lattice circular, no entanto, observa-se uma única transição (seta 1) separando apenas duas fases. Note como este comportamento, agora bastante mais evidente, já havia sido verificado quando da avaliação de $N_{clusters}$. Argumentaremos a seguir que estas fases estão relacionadas com mudanças estruturais na *rede de opiniões* a medida que o valor de ε se torna cada vez menor.

Para valores altos do parâmetro de confiança ε o sistema vai a *consenso*. Neste regime, o tamanho do maior cluster S_1 tende à N e S_2 vai a zero. A medida que ε gradativamente diminui, clusters microscópicos passam a coexistir com um cluster gigante reduzindo ligeiramente seu tamanho (como pode ser observado no intervalo de ε [0,2;0,4] para as Figuras 4.4(a) e 4.4(b) e intervalo [0,3;0,4] para a Figura 4.4(c)). Para as redes ER e BA, um salto em S_1/N de 1 para aproximadamente 1/2 ocorre em torno do valor $\varepsilon = 0.18$. Concomitantemente, um segundo cluster S_2 , de tamanho comparável, passa a coexistir com o maior cluster S_1 e o sistema entra numa fase polarizada. Assim, a seta 1 marca a transição entre uma fase monopolar, de domínio absoluto de uma opinião para uma fase bipolar. A partir deste momento, a medida que ε diminui, S_1 e S_2 passam a diminuir gradativamente dando lugar a mais clusters microscópicos.

Figura 4.4 – Tamanhos do primeiro e segundo maior cluster obtidos com realizações da dinâmica de KH em termos de ε para redes do tipo (a) ER, (b) BA e (c) LA com $\langle k \rangle = 10$ e tamanhos conforme indicado na Figura. Setas identificam transições de fase na estrutura da rede de opiniões, conforme discussão no texto. Média sobre 200 realizações das redes e da dinâmica.

Além disso, conforme verificado na Figura 4.2, $N_{clusters}$ torna-se uma grandeza extensiva. Para valores de ε ainda menores, é observada uma segunda transição nestas duas redes, identificada pela seta 2, a qual separa a fase polarizada de uma fase fragmentada na qual não há existência de clusters de tamanho macroscópico. No limite termodinâmico ($N \rightarrow \infty$), as grandezas S_1/N e S_2/N vão a zero e as transições tornam-se cada vez mais abruptas. Note que para rede LA o ponto de transição 1 marca a mudança de uma fase monopolar para uma fase fragmentada sem a existência de uma fase intermediária na qual clusters macroscópicos coexistam. Esta importante diferença fica evidenciada na Figura 4.4(c) e já havia sido notada anteriormente quando

da avaliação de $N_{clusters}$ e $P_{clusters}(S)$ para a rede LA.

Em resumo, os gráficos das Figuras 4.2, 4.3 e 4.4 nos permite categorizar fases na estrutura da *redes de opiniões* da seguinte forma:

- **Fase monopolar:** marcada pela presença de um único cluster de opinião macroscópico e clusters de tamanhos microscópico. No limite de altos valores de ε o sistema entra em *consenso* e um único cluster aglutina toda a rede de agentes. Parâmetro $N_{clusters}$ não é extensivo com o tamanho da rede. Para redes ER e BA apresenta-se, aproximadamente, no intervalo de ε $[0,18;1,00]$; para a rede LA no intervalo $[0,24;1,00]$.

- **Fase polar ou multipolar:** marcada pela presença de mais de um cluster macroscópico em associação à clusters microscópicos. $N_{clusters}$ extensivo com o tamanho da rede. Para redes ER e BA apresenta-se, aproximadamente no intervalo de ε $[0,04;0,18]$; não é verificado na rede LA.

- **Fase fragmentada:** marcada pela presença apenas de clusters microscópicos. Para redes ER e BA apresenta-se aproximadamente no intervalo de ε $[0,00;0,04]$; para a rede LA no intervalo $[0,00;0,24]$.

Ainda no sentido de evidenciar as transições de fase discutidas até o momento, apresentamos na Figura 4.5 o tempo necessário para a estabilização da dinâmica de KH em termos de ε para as três redes. Conforme pode ser verificado, na região de mudança de fase o tempo para convergência aumenta substancialmente. De fato, a divergência no tempo é uma assinatura da transição de fase (31), de modo que a posição e o número de picos observados na Figura 4.5 corroboram com as evidências das Figuras anteriores. Note ainda que os picos se tornam mais evidentes quando o tamanho N do sistema aumenta.

A evidente associação entre a divergência no tempo e uma transição de fase pode ser explicada do ponto de vista microscópico através da redução do número de *agentes mediadores*. A Figura 4.6 mostra a evolução temporal da dinâmica de KH próxima da transição *monopolar-multipolar*. Note que a partir de $T \sim 5$ o sistema entra num estado *metaestável* (32) até que o consenso seja atingido. Neste período três clusters se formam, dois grupos na extremidade,

Figura 4.5 – Tempo para estabilização da dinâmica de KH em termos de ε para redes (a) ER, (b) BA e (c) LA com $\langle k \rangle = 10$ e tamanhos conforme indicado na Figura. Média sobre 200 realizações das redes e da dinâmica.

cuja separação é maior do que ε , e um cluster central. A existência deste grupo próximo da opinião $O \sim 0.5$ é, portanto, fundamental para que no estado estacionário um grande cluster de única opinião se forme. Conforme o valor de ε diminui até o ponto de transição, o tamanho desse cluster central também diminui. Uma vez que dinâmica de KH é baseada num processo de médias entre os agentes, isto implica que o tempo para convergência aumenta. A redução do número de *intermediadores* fica evidenciada na Figura 4.7 a qual apresenta a dependência, no estado transiente, do tamanho médio do cluster cujos agentes possuem opinião em torno do valor central 0.5 em termos do parâmetro de confiança ε para a rede ER. Note que o intervalo de valores de ε está concentrado em torno da transição, e neste ponto obtém-se um mínimo que cresce a medida que N aumenta, conforme esperado.

4.3 Relação com a Teoria de Percolação

Daremos continuidade ao comportamento da dinâmica de KH em redes avaliando com maiores detalhes a estrutura do maior cluster de opinião. Verificamos na Figura 4.4 como seu

Figura 4.6 – Exemplo de evolução temporal da dinâmica de KH próximo da transição monopolar-multipolar para rede ER de tamanho $N = 500$ e $\langle k \rangle = 10$.

tamanho, S_1 , varia em termos de ε . Gostaríamos de avaliar agora a conectividade deste grupo. Para tanto, é preciso lembrar que à medida que ε diminui, a *rede de opiniões* perde cada vez mais conexões de modo que seria esperado que o caminho mínimo médio, $ASPL$, do maior cluster (assim como de toda a rede de opiniões) aumente. Esta hipótese confirma-se através dos resultados mostrados na Figura 4.8 a qual mostra como o mínimo caminho médio do maior componente é afetado com a diminuição de ε . Note que existem dois comportamentos: i) para $\varepsilon > 0.04$, o valor de $ASPL_{S_1}$ é uma função decrescente com ε e ii) para $\varepsilon < 0.04$, $ASPL_{S_1}$ possui comportamento monotonicamente crescente. Devemos lembrar que ponto $\varepsilon^* = 0.04$ marca a transição para a *fase fragmentada* e, portanto, isto significa que enquanto o maior componente da *rede de opiniões* possui tamanho macroscópico sua densidade diminui conforme ε também diminui, até que no momento da transição o mínimo caminho médio atinge o seu valor máximo. Neste ponto, o maior componente é uma estrutura fracamente conectada (assim como as demais outras estruturas macroscópicas menores) e mesmo uma pequena redução do valor de ε leva à desfragmentação do sistema.

Esta é, portanto, uma outra forma de se avaliar o surgimento da fase fragmentada. É preciso destacar que a presença de um ponto de máximo pronunciado na Figura 4.8 torna mais precisa a medição da posição em que ocorre a transição. Além disso, deve-se notar que este é um fenômeno causado exclusivamente pela eliminação de arestas da *rede de opiniões* de acordo com as regras da dinâmica de KH. Trata-se, portanto, de um típico fenômeno de percolação e

Figura 4.7 – Tamanho do cluster central em termos do parâmetro ε para rede ER. Média sobre 50 realizações de redes e da dinâmica.

um paralelo pode ser feito neste sentido, conforme os próximos resultados evidenciarão.

A Teoria de percolação é um tópico amplamente estudado tanto no campo da matemática como da mecânica estatística (33, 34). Originalmente, o problema trata da emergência de percolação num reticulado de dimensão d no qual arestas estão presentes com probabilidade p . Verifica-se a existencia de uma probabilidade crítica $p_c(d)$ para qual se forma um cluster gigante de vértices conectando as extremidades do reticulado. Uma extensão natural é a verificação deste fenômeno em Redes Complexas. A principal motivação é a investigação da resiliencia e rebustez das redes perante a ataques/falhas sobre vértices ou arestas (35). Existem diversos trabalhos na literatura que oferecem detalhadamente aprofundado sobre o tema, destacando-se a descrição de Percolação em *grafos aleatórios generalizados* (36), neste caso considerando o formalismo baseado em *Funções Geratrizes*, e a relação do problema com dinâmicas de propagação de epidemias (37). Uma revisão sobre o tema, bem como sobre outros fenômenos críticos em redes, pode ser encontrado em (38).

Derivamos a seguir alguns dos principais resultados de percolação para grafos aleatórios.

Figura 4.8 – Mínimo caminho médio entre os vértices que compõem o maior cluster de opiniões em termos de ϵ . A seta identifica o ponto de máximo para esta medida e está relacionada com a transição para a fase fragmentada. Simulação considerando redes ER de 750 e 1000 vértices com $\langle k \rangle = 10$. Média sobre 50 realizações da dinâmica de KH.

Para uma rede com distribuição de grau $P(k)$, considere que, com probabilidade g , arestas são eliminadas. O objetivo é encontrar a probabilidade crítica g^* que determina a transição para uma fase fragmentada. Considere a probabilidade, q , de que uma aresta não aponte para um vértice pertencente a um componente macroscópico. Devemos lembrar que para redes sem correlação de grau, a probabilidade de uma aresta estar conectada a um vértice de grau k é $kP(k)/\langle k \rangle$. Assim, negligenciando a existência de ciclos, a quantidade q pode ser calculada como o produto entre a probabilidade de uma aresta apontar para um vértice de grau k e a probabilidade que nenhuma das outras $k - 1$ arestas remanescentes aponte para um componente macroscópico, considerando todos possíveis valores de k . Colocando em termos de uma equação transcendental temos:

$$q = \sum_k \frac{kP(k)}{\langle k \rangle} q^{k-1}. \quad (4.2)$$

O limite em que $q \rightarrow 1$ indica que o sistema não possui um componente macroscópico. Portanto, este componente existe sempre que q possui solução real positiva com $q < 1$. Uma solução da equação 4.2 pode ser encontrada considerando a intersecção entre as curvas $y_1(q) = q$ e $y_2(q) = \sum_k \frac{kP(k)}{\langle k \rangle} q^{k-1}$. Note que y_2 é uma função monotonicamente crescente e côncava com relação à q nos limites $y_2(0) = P(1)/\langle k \rangle < 1$ e $y_2(1) = 1$. Assim, existirá valor de $q^* < 1$ se a derivada de $y_2(q)$ em $q = 1$ for maior ou igual a 1:

$$\left. \frac{d}{dq} \left(\sum_k \frac{kP(k)}{\langle k \rangle} q^{k-1} \right) \right|_{q=1} \equiv \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \geq 1 \quad (4.3)$$

A transição de fase ocorre na condição de igualdade na equação anterior e, então:

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = 2. \quad (4.4)$$

Considerando uma probabilidade de eliminação de arestas igual a g , um vértice de grau originalmente igual a k' terá grau k com probabilidade:

$$\binom{k'}{k} (1-g)^k (g)^{k'-k}. \quad (4.5)$$

E, portanto, a nova distribuição de grau será dada por

$$P_g(k) = \sum_{k'=k}^{\infty} P(k') \binom{k'}{k} (1-g)^k (g)^{k'-k}. \quad (4.6)$$

Consequentemente, para os momentos obtém-se

$$\langle k \rangle_g = \sum_k k P_g(k) = (1-g) \langle k \rangle \quad (4.7)$$

e

$$\langle k^2 \rangle_g = \sum_k k^2 P_g(k) = (1-g)^2 \langle k^2 \rangle + g(1-g) \langle k \rangle \quad (4.8)$$

Utilizando a condição obtida na equação 4.4, obtemos, após algumas manipulações, a probabilidade g^* crítica a qual marca transição para uma fase fragmentada:

$$g^* = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1} \quad (4.9)$$

Fazendo o paralelo com a dinâmica de KH, o processo de percolação na *rede de opiniões* é marcado pela eliminação de arestas sempre que, para dois vértices i e j $|O_i - O_j| < \varepsilon$. Assim, podemos assumir, muito aproximadamente, quando $\varepsilon \ll 1$, que a probabilidade de eliminação g é dada por $1 - 2\varepsilon$. Reescrevendo a equação anterior, obtemos o valor de ε^* que marca a transição para a *fase fragmentada*:

$$\varepsilon^* = \frac{1}{2} \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1} \quad (4.10)$$

Verifica-se, portanto, que a posição de mudança para a *fase fragmentada* depende tanto do grau médio da rede como da heterogeneidade da distribuição de grau. A observação do primeiro ponto está endereçada na Figura 4.9(a), a qual mostra, assim como na Figura 4.8, a relação entre o mínimo caminho médio do maior componente da *rede de opiniões* em termos de ε para redes ER de diferentes valores de $\langle k \rangle$ (letras indicam os máximos para $ASPL_{S_1}$ e, conseqüentemente, o ponto de transição). É preciso lembrar que este modelo de rede segue a distribuição de Poisson para distribuição de grau e, portanto, $\langle k^2 \rangle = 2\langle k \rangle^2$. Logo, utilizando o resultado obtido na equação 4.10, espera-se que $\varepsilon^* \sim \frac{1}{2\langle k \rangle - 1}$. De fato, a Figura 4.9(b) mostra a dependência das posições dos picos encontrados em 4.9(a) com o grau médio das redes, confirmando a regra.

Ao comparar redes com mesmo $\langle k \rangle$, mas diferentes $\langle k^2 \rangle$ espera-se que os valores de ε^* também sejam diferentes, diminuindo à medida que a heterogeneidade da rede aumenta. Devemos lembrar que para a rede BA, o valor de $\langle k^2 \rangle$ diverge no limite em que $N \rightarrow \infty$, já que

Figura 4.9 – Efeito do grau médio na posição da transição para a fase fragmentada. (a) Mínimo caminho médio entre os vértices que compõem o maior cluster de opiniões considerando redes do tipo ER com diferentes valores de $\langle k \rangle$ (grau médio varia de 5.0 a 15.0, conforme indicado na Figura) em termos do parâmetro de confiança ϵ . (b) Relação entre os pontos de máximo obtidos em (a) para cada rede e os respectivos valores de $\langle k \rangle$. Média sobre 50 realizações das redes e da dinâmica de KH.

distribuição é uma lei de potência de expoente igual 3. Isso significa que, por menor que seja o valor do parâmetro de confiança ϵ , sempre existirão componentes de tamanho macroscópico. Para redes de tamanho finito, ainda que $\epsilon_{BA}^* \ll \infty$, espera-se que $\epsilon_{BA}^* > \epsilon_{ER}^*$, de acordo com a Equação 4.10. De fato, a Figura 4.10 apresenta a dependência entre $ASPL_{S_1}$ e ϵ para redes ER e BA de mesmo $\langle k \rangle$ igual 10. Note que a posição de mudança de fase ocorre num valor menor para a rede BA, conforme esperado.

Figura 4.10 – Mínimo caminho médio entre os vértices que compõem o maior cluster de opiniões em termos de ε para redes ER e BA com $N = 1000$ e $\langle k \rangle = 10$. Valores de ε próximos da região de transição para a fase fragmentada. Média sobre 50 realizações da dinâmica de KH.

5 **MODIFICAÇÕES DA DINÂMICA DE KRAUSE-HEGSELMANN**

Mesmo se tratando de um modelo cujo objetivo é a reprodução da evolução das opiniões entre agentes interagentes, a dinâmica de KH segue regras bastante simplificadas. Diversos ingredientes adicionais podem ser incorporados no modelo de modo a reproduzir características do mundo real. Por exemplo, o parâmetro de confiança ε poderia depender dos pares de agentes, i e j , que estejam interagindo, além do valor das suas opiniões, ou seja, $\varepsilon \rightarrow \varepsilon_{i,j}(O_i, O_j)$. De fato, é bastante aceitável que agentes possam ter mais confiança num determinado grupo de indivíduos, ou ainda, que exista uma tendência natural a se posicionar de maneira inclinada sobre um assunto. Outra possível alteração é a mudança do processo de médias fazendo com que certos indivíduos sejam menos susceptíveis à mudanças, em analogia com o parâmetro μ do modelo de Deffuant (14). No que diz respeito à rede de agentes, esta poderia alterar-se ao longo do tempo de modo que arestas seriam criadas ou destruídas conforme pares de agentes tivessem opiniões alinhadas ou contrárias entre si.

Dentre as várias possibilidades, endereçaremos neste capítulo duas importantes modificações da dinâmica de KH. A primeira é de natureza *interna* ao sistema e diz respeito à presença de ruído ou falha de comunicação entre os agentes. A outra modificação, de natureza *externa*, trata da inclusão de um campo (ou propaganda) cuja proposta é de alinhamento das opiniões. Estes efeitos serão apresentados tanto para grafos completos como também para outras topologias, incluindo o modelo de rede social discutido no Capítulo 2.

5.1 Ruído

No modelo de KH a dinâmica temporal evolui de maneira determinística, ou seja, dada uma configuração inicial de opiniões e o grafo no qual os agentes estão inseridos é possível saber exatamente quantos serão o número de clusters finais e os valores das opiniões de cada agente. Nesta seção, iremos inserir um ingrediente estocástico na dinâmica de KH: em cada passo t da simulação, para um agente i , a regra de evolução definida na Equação 4.1 será realizada com probabilidade $(1 - \alpha)$, caso contrário, o agente assume uma nova opinião uniformemente escolhida no intervalo contínuo $[0, 1]$. Na prática, estaremos introduzindo um tipo de ruído na evolução das opiniões cuja *intensidade* depende do parâmetro α , sendo que para $\alpha = 0$ a dinâmica original é recuperada. O efeito deste tipo de ruído é transportar um agente para condição inicial ignorando completamente as opiniões dos seus vizinhos.

Trata-se de uma modificação com a intenção de reproduzir de maneira bastante simplificada eventos do mundo real que causariam alguma alteração imotivada de uma opinião, como por exemplo, problemas de comunicação ou compreensão ou alguma tendência própria de um indivíduo no sentido de produzir uma diferenciação (39). O objetivo desta seção, portanto, é verificar como a presença de ruído altera os perfis de opiniões encontrados anteriormente.

5.1.1 Resultados para o grafo completo

Para a condição de mistura homogênea podemos reescrever a equação mestra 3.3 de modo a incluir o ruído como novo ingrediente. Assim, a evolução temporal deve possuir duas contribuições:

i) a regra de KH original, com peso $(1 - \alpha)$, e a ação do ruído com peso α :

$$\begin{aligned} \frac{\partial P(o, t)}{\partial t} = (1 - \alpha) \int_0^1 dx P(x, t) \left[\delta \left(o - \frac{\int_{x-\varepsilon}^{x+\varepsilon} dx' x' P(x', t)}{\int_{x-\varepsilon}^{x+\varepsilon} dx' P(x', t)} \right) - \delta(o - x) \right] + \\ + \alpha [P_\alpha(o, t) - P(o, t)], \quad (5.1) \end{aligned}$$

onde $P_\alpha(o, t)$ é uma distribuição genérica que representa o ruído. Em nosso caso $P_\alpha(o, 0) = U([0, 1])$. Da mesma forma como fizemos no Capítulo 3, a Equação mestra será resolvida de maneira aproximada através de uma formulação em termos de cadeias de Markov, de modo que $P_{t+1}(o) = P_t(o) \mathbf{B}_\alpha(P_t(o), \varepsilon, t)$, onde $\mathbf{B}_\alpha(P_t(o), \varepsilon, t)$ é uma modificação da Equação 3.11 para inclusão de ruído, como segue:

$$\mathbf{B}_\alpha[i, j] = \begin{cases} \frac{1}{n_d} + (1 - \alpha) & \text{se } j = M_i, \\ \frac{1}{n_d} + (1 - \alpha)(\lceil M_i \rceil - M_i) & \text{se } j = \lfloor M_i \rfloor, j \neq M_i, \\ \frac{1}{n_d} + (1 - \alpha)(\lfloor M_i \rfloor - M_i) & \text{se } j = \lceil M_i \rceil, j \neq M_i, \\ \frac{1}{n_d} & \text{caso contrário} \end{cases} \quad (5.2)$$

onde n_d é o número de discretizações do espaço de opiniões (dimensão de $\mathbf{B}_\alpha$) e M_i é a posição média, no espaço de opiniões, da massa compatível com a opinião i/n_d , conforme definida na equação 5.9. Note que para matriz $\mathbf{B}_\alpha$ é respeitada a condição de normalização da probabilidade de mudança de estado, ou seja, $\sum_j \mathbf{B}_\alpha[i, j] = 1 \forall i$.

A Figura 5.1 apresenta alguns exemplos, realizados através de simulação de Monte Carlo, da evolução temporal das opiniões para o grafo completo considerando alguns valores de ε e α conforme indicado na Figura. Note que existe uma tendência geral de homogeneização da ocupação do espaço de opiniões que se torna mais intensa quando o valor α aumenta. Embora se possa notar uma aglomeração em torno de alguns valores de opinião, que varia de acordo com ε , o estado estacionário não é mais composto por clusters bem definidos, nem as opiniões mantêm-se fixas ao longo do tempo.

Para um visão mais geral, a Figura 5.2 apresenta os diagramas de bifurcação obtidos a partir

Figura 5.1 – Exemplos de realizações da dinâmica de KH na presença de ruído. Séries temporais mostram a evolução das opiniões ao longo do tempo para valores de α iguais a 0.2 (coluna da esquerda) e 0.6 (coluna da direita) e ϵ iguais a 0.1, 0.3 e 0.5 (de cima para baixo). Níveis de cinza indicam a concentração de agentes ao longo do espaço de opiniões. Resultados obtidos para um grafo completo com $N = 1000$.

da solução da Equação 5.1 através da cadeia de Markov definida na Equação 5.2 para valores de α variando de 0.1 a 0.6. Lembremos que estes gráficos representam a solução estacionária do perfil de opiniões em termos de ϵ , ou seja, $P(o, \infty)$. A condição inicial é a mesma das situações anteriores, $P(o, 0) = U([0, 1])$. Os padrões obtidos assemelham-se ao diagrama original da dinâmica na ausência de ruído (Figura 3.3), mas com uma tendência geral à difusão pelo espaço de opiniões e, conseqüentemente, tornando os clusters originais menos definidos. Este efeito se torna cada vez mais intenso a medida que a intensidade do ruído α aumenta.

No capítulo anterior avaliamos como o número de clusters, $N_{clusters}$, era afetado pelo parâmetro ϵ . Na presença de ruído precisaremos definir outro conceito de agrupamento, uma vez que a

Figura 5.2 – Distribuições estacionárias para o modelo de KH na presença de ruído. Resultados obtidos a partir da solução numérica da equação de campo médio (Eq. 5.1) reescrita em termos da cadeia da Markov interativa (Eq. 5.2). Valores de α variam de 0.1 a 0.6 (esquerda para direita, de cima para baixo). Espaço de opiniões e de ϵ discretizados em 100 partes.

definição de cluster de opinião perde sentido. Isto será feito através de uma medida de entropia (40) que capture a heterogeneidade de $P(o, t)$ no regime estacionário. Considere o espaço de opiniões dividido em M partes iguais e para cada i -ésima divisão obtém-se o número de agentes com opinião i , l_i . Assim, para uma quantidade N de agentes, define-se:

$$S_M = - \sum_{i=1}^M (l_i/N) \ln(l_i/N) \quad (5.3)$$

Caso a população esteja concentrada numa única região, $S_M = 0$; caso a ocupação seja

homogênea, $S_M = \ln(M)$. Finalmente, utilizaremos o seguinte coeficiente:

$$G_M = M^{-1} \exp(\bar{S}_M), \quad (5.4)$$

onde a barra superior representa a média temporal de S_M . Note que $1/M \leq G_M \leq 1$, sendo que valores próximos da unidade serão obtidos nas situações de maior desordem, para a qual $P(o, t)$ encontra-se distribuído por todo espaço de opiniões.

A Figura 5.3 mostra a dependência do coeficiente G_M para o grafo completo em termos da intensidade do ruído α para dois valores de ε . Note que G_M apresenta comportamento monotonamente crescente com α saturando em 1, conforme esperado. Além disso, uma vez que a diminuição do parâmetro de confiança ε reduz a convergência entre as opiniões dos agentes, o valor de G_M também aumenta.

Figura 5.3 – Coeficiente de dispersão, G_M , obtido para a dinâmica de KH em termos da intensidade do ruído, α , e valores do parâmetro de confiança, ε , conforme indicado na Figura. Foi considerado grafo completo com $N = 1000$ agentes. Média sobre 50 realizações da dinâmica e espaço de opiniões discretizado em 100 partes.

5.1.2 Efeito do ruído na dinâmica de KH em Redes Complexas

Uma vez verificadas as mudanças que a presença de ruído causa na dinâmica original de KH em grafos completos, mudaremos agora nossa atenção para as situações em que os agentes encontram-se conectados através de redes. Nesta avaliação, consideraremos dois tipos de topologia: i) o modelo de rede de Erdős-Rényi (ER) e ii) o modelo social de Boguñá (SN) (26), cuja descrição foi feita no Capítulo 2.

A Figura 5.4 mostra a dependência do coeficiente G_M com relação à intensidade do ruído α para as redes ER e SN. O comportamento monotônico é mantido, aumentando à medida que α aumenta. É importante destacar que, sistematicamente, os valores de G_M encontrados para a rede SN são maiores, independente de ε . Mesmo na ausência de ruído ($\alpha = 0$), a ocupação do espaço de opiniões é maior para este modelo de rede. De fato, essa evidência é ilustrada na Figura 5.5, a qual apresenta um exemplo da evolução temporal da dinâmica de KH nas redes ER e SN para o mesmo valor de ε e na ausência de ruído. Note que tanto o número de clusters gerados no modelo social quanto o tempo de convergência são substancialmente maiores quando comparados com a rede ER. O motivo para esta diferença pode ser explicado pela correlação de grau positiva existente no modelo SN. A presença da assortatividade introduz o isolamento de pequenos clusters pouco conectados. Como consequência, as opiniões dos agentes contidos nesses clusters acabam se diferenciando. Existe, portanto, uma tendência natural do modelo social de diversificação das opiniões a qual é intensificada na presença de ruído.

Figura 5.4 – Coeficiente de dispersão, G_M , em termos da intensidade do ruído, α , para os modelos de rede (a) ER e (b) o modelo social SN de tamanhos N igual 1000. Valores do parâmetro de confiança conforme indicado na Figura. Média sobre 100 realizações das redes e da dinâmica e espaço de opiniões discretizado em 100 partes.

Efeito de ruído temporário

Até agora verificamos, de maneira bastante geral, o impacto causado pela ação de um ruído atuando constantemente sobre o sistema desde o instante inicial, quando as opiniões ainda são muito diversas. Surge, portanto, uma questão natural: qual seria a influência do ruído sobre um sistema em que clusters de opiniões já estejam formados? Endereçaremos esta pergunta considerando o efeito de uma perturbação que age durante um intervalo reduzido de tempo ΔT . Assim, a evolução temporal da dinâmica possuirá três etapas: 1) uma primeira fase de convergência, sem ruído, guiada pela dinâmica original de KH, até que uma configuração de clusters seja atingida; 2) uma etapa de duração ΔT na qual ruído é aplicado seguindo as regras da seção anterior e 3) mais uma etapa de convergência sem ação de perturbação.

A Figura 5.6 ilustra a situação de ruído temporário no modelo ER, com $\Delta T = 10$, intensidade

Figura 5.5 – Comparação da evolução temporal das opiniões no modelo KH para redes ER e SN, ambas com mesmo tamanho $N = 500$ e grau médio $\langle k \rangle = 10$ sujeitas a $\epsilon = 0.15$. Parâmetro β do modelo SN igual 3. Níveis de cinza estão relacionadas com a densidade de agentes ao longo do espaço de opiniões.

do campo $\alpha = 0.1$ e valores de ϵ iguais a 0.1 e 0.2, conforme indicado na Figura. A distinção das três etapas fica bastante clara na Figura. É visível que o principal efeito causado pela perturbação é a ampliação do número de clusters e a diversificação das opiniões. No entanto, diferente da situação em que o ruído era aplicado constantemente, a ação por tempo reduzido tende a incentivar a ocupação de opiniões bastante separadas dos clusters antes da aplicação do ruído. Assim, o que se verifica é o aumento de minorias de opiniões extremas e, portanto, ainda que o ruído em si não privilegie um opinião, sua ação sobre clusters já formados promove a diversificação nas regiões pouco ocupadas do espaço.

De maneira sistemática, avaliaremos o efeito causado pela ação do ruído temporário nas redes ER e SN em termos do parâmetro ΔT . A Figura 5.7 mostra o tempo de convergência da dinâmica de KH, T_{CONV} , após a aplicação do ruído. Veja que o tempo de convergência cresce a medida que o tempo de aplicação do ruído ΔT aumenta, até que seu valor atinge um patamar após um tempo característico ΔT^* . Note que o valor da intensidade do ruído, α , exerce pouca influência no valor de T_{CONV} , mas altera significativamente o valor crítico ΔT^* . Destacam-se também os valores acentuados encontrados para a rede SN quando comparada com a rede ER.

A Figura 5.8 quantifica o efeito de criação de clusters gerados pela ação do ruído. Ela

Figura 5.6 – Exemplos de realizações da dinâmica de KH na presença de ruído temporário ($\Delta T = 10$ passos de simulação) para valores de ϵ iguais a 0,1 e 0,2 conforme indicado na figura. Níveis de cinza indicam a concentração de agentes ao longo do espaço de opiniões.

Figura 5.7 – Tempo de convergência da dinâmica de KH em termos do tempo ΔT de aplicação de ruído. Resultados para redes ER e SN considerando valores de ϵ e intensidade do ruído, α , conforme indicados na Figura. Média sobre 50 realizações das redes e da dinâmica.

mostra a relação, R_{NC} , entre o número de clusters, no estado estacionário, antes da aplicação do ruído e depois. Como esperado, este valor fica sempre menor do que 1, independente dos valores de ϵ e α . No entanto, a dependência com ΔT apresenta comportamento menos trivial. Aproximadamente no tempo crítico ΔT^* da Figura anterior, o valor de R_{NC} apresenta um ponto de mínimo, ou seja, a diversificação das opiniões é máxima em ΔT^* e diminui caso a aplicação

do ruído seja muito prolongada. Para entender este efeito devemos lembrar como ruído age na dinâmica de KH: em cada passo da dinâmica os agentes, com probabilidade α , ignoram seus vizinhos e escolhem um opinião uniformemente em todo o intervalo de valores possíveis. Assim, caso ΔT seja muito grande, muitos agentes terão suas opiniões alteradas e o sistema se aproximará da configuração inial, $P(o,0)$. Neste momento, quando o ruído é desligado, e como as opiniões estão espalhadas ao longo do espaço de opiniões, muitas das opiniões terão separação menor que ε e, conseqüentemente, muitos agentes ficarão num mesmo cluster. Por outro, se a aplicação do ruído ocorrer durante tempo ΔT^* , grande o suficiente para diversificar as opiniões de alguns agentes, mas pequeno o suficiente para que o espaço de opiniões não seja homogeneamente preechido, então o grupo de agentes com opiniões alteradas tenderão à formar clusters separados. Note que o tamanho do valor mínimo é pouco alterado com a mudança da intensidade α do ruído, embora sua posição dependa deste parâmetro. Outro ponto bastante relevante a se destacar na Figura 5.8 é a diferença entre os resultados para as redes ER e SN. Note que o modelo social sofre muito menos influência da ação ruído em comparação com o modelo aleatório.

A Figura 5.9 mostra a relação, R_G , entre o tamanho do maior cluster antes e depois da aplicação do ruído para as redes ER e SN. O resultado complementa as informações obtidas a até momento e mostra a redução do maior cluster à medida tempo que ΔT aumenta. Assim como mostrado na Figura anterior, o componente gigante da rede SN é menos modificado pela aplicação da perturbação.

Associado à promoção produzida pelo ruído mostramos na Figura 5.10 como a distribuição de opiniões é alterada com a aplicação da perturbação. Considere $P_A(o)$ e $P_D(o)$ as distribuições no estado estacionário das opiniões antes e depois da aplicação do ruído, respectivamente. Para quantificar as mudanças causadas pela perturbação, definimos a seguinte medida:

$$S = \sqrt{\int_0^1 do [P_A(o) - P_D(o)]^2} \quad (5.5)$$

Assim, caso $P_D(o) = P_A(o)$, S é igual a zero. Verifica-se na Figura que o perfil de opiniões

Figura 5.8 – Relação entre o número de clusters sem aplicação de ruído e após aplicação de ruído durante intervalo de tempo ΔT . Resultados para redes ER e SN com valores de ϵ e intensidade do ruído, α , conforme indicados na Figura. Média sobre 50 realizações das redes e da dinâmica.

Figura 5.9 – Relação entre o tamanho do maior cluster de opinião sem aplicação de ruído e após aplicação de ruído com relação ao parâmetro ΔT . Resultados para redes ER e SN com valores de ϵ e intensidade do ruído α conforme indicados na Figura. Média sobre 50 realizações das redes e da dinâmica.

sofre a alterações na presença de ruído temporário e 5.10 apresenta comportamento crescente com o parâmetro ΔT até que um patamar é atingido. No entanto, note que os valores obtidos para a rede SN são maiores do aqueles obtidos para a rede ER. Desta forma, ainda que o número de clusters pouco se altere para o modelo social, as opiniões dos seus agentes são mais modificadas.

Figura 5.10 – Tendência de manutenção das opiniões S após aplicação de ruído durante intervalo de tempo ΔT . Resultados para redes ER e SN com valores de ε e intensidade do ruído, α , conforme indicados na Figura. Média sobre 50 realizações das redes e da dinâmica.

5.2 Aplicação de campo constante

Nesta seção, investigaremos a influência de um campo externo sobre a evolução de opiniões atualizadas pela dinâmica de KH. Enquanto o conceito de ruído explorado na seção anterior tinha um caráter desagregador, com a tendência de diversificação das opiniões, a presença de um campo externo possui a intenção de produzir um alinhamento na opinião para qual ele aponta. Existem alguns trabalhos que exploram o efeito de um ingrediente externo homogeneizador nas dinâmicas baseadas em *confiança limitada*: em (41), um campo constante é visto como efeito de propaganda na dinâmica de *Deffuant* e em (42) a ação do campo no modelo de KH é interpretada

como uma tendência natural ao alinhamento com conceitos estabelecidos. Independente de qual seja a interpretação, a influência do campo a ser adotada aqui é *unidirecional*, ou seja, ele é capaz de direcionar as opiniões dos agentes do sistema, mas as opiniões não são capazes de alterá-lo. Esta é uma diferença importante, já que não se aplica às situações em que os agentes seguem a opinião média ou a opinião da maioria, muito comum em situações reais em que pesquisas de opinião são informações disponíveis aos agentes.

A aplicação do campo no modelo de KH será feita da seguinte forma: em cada passo de tempo, t , um agente, após atualizar sua opinião de acordo com a dinâmica de KH, incorpora o valor da opinião externa, B , com um peso γ ($0 \leq \gamma \leq 1$), desde que a separação da sua opinião com relação à B não seja maior que ε . Assim, para um agente i que respeite essa condição:

$$O_i(t+1) = (1 - \gamma)\bar{O}_{i;\varepsilon}(t) + \gamma B \quad (5.6)$$

onde $\bar{O}_{i;\varepsilon}$ é a opinião média dos agentes compatíveis com o agente i . Mostraremos nas seções a seguir o efeito do campo na dinâmica original de KH, nos modelos de rede, e o papel da acessibilidade na determinação dos agentes mais efetivos para a propagação de uma informação externa na rede.

5.2.1 Resultados para o Grafo Completo

Da mesma forma como fizemos na seção anterior, iniciaremos a investigação do efeito do campo externo na situação em que todos os agentes podem comunicar entre si. Neste caso, a equação mestra para determinação da evolução de $P(o, t)$ será:

$$\frac{\partial P(o, t)}{\partial t} = \int_0^1 dx P(x, t) \left[\delta \left(o - \gamma B - (1 - \gamma) \frac{\int_{x-\varepsilon}^{x+\varepsilon} dx' x' P(x', t)}{\int_{x-\varepsilon}^{x+\varepsilon} dx' P(x', t)} \right) - \delta(o - x) \right] \quad (5.7)$$

Note que, diferente da equação construída para o caso de ruído, a Equação 5.7 possui o mesmos termos da Equação original 3.3, mas é alterada no argumento da primeira δ , ou seja, a massa média em torno de uma opinião é deslocada no sentido do campo B com peso γ . Assim, para a solução em termos de Cadeias de Markov, a matriz de transição $\mathbf{B}(P_i(o))$ é igual à matriz original definida em 3.11, ou seja

$$B[i, j] = \begin{cases} 1 & \text{if } j = M_i, \\ \lceil M_i \rceil - M_i & \text{if } j = \lceil M_i \rceil, j \neq M_i, \\ \lfloor M_i \rfloor - M_i & \text{if } j = \lfloor M_i \rfloor, j \neq M_i, \\ 0 & \text{caso contrário} \end{cases} \quad (5.8)$$

No entanto, M_i deve ser alterada para

$$M_i = (1 - \gamma) \frac{\sum_{m \in n, |i-m| \leq n_d \epsilon} m P(m)}{\sum_{m \in n, |i-m| \leq n_d \epsilon} P(m)} + \gamma B \quad (5.9)$$

A Figura 5.11 mostra o diagrama de bifurcação para a dinâmica de KH na presença de campo externo constante. Os gráficos foram obtidos a partir da cadeia de Markov definida nas Equações 5.8 e 5.9 para a opinião do campo, B , em diferentes posições e γ constante igual à 0.5. Comparando com o diagrama original da dinâmica apresentado no Capítulo 3 (Figura 3.3), duas diferenças podem ser imediatamente observadas: i) à medida que o campo assume valores de opinião extremas consenso é obtido para valores de ϵ mais altos, de modo que um cluster na extremidade contrária ao campo é sustentado e; ii) na situação de consenso, a opinião dominante é deslocada da posição $O = 0.5$ para $O = B$. Mais do que isso, a posição $O = B$ é sempre ocupada, ainda que com diferentes densidades. Portanto, note que a introdução do campo fora da posição central estabelece uma quebra de simetria ao longo do eixo de opiniões. Esta é a primeira vez que esta diferenciação aparece.

Também baseada nas Equações 5.8 e 5.9, a Figura 5.12 mostra o comportamento do diagrama de bifurcação com relação à intensidade do campo γ quando B é mantido na posição 0.25.

Figura 5.11 – Diagrama de bifurcação com as distribuições estacionárias para o modelo de KH na presença de campo externo considerando diferentes valores da opinião B . Resultados obtidos a partir da solução da Equação 5.7 reescrita em termos da Cadeia de Markov em 5.8. Valores de B variam de 0.1 a 0.9 (esquerda para direita, de cima para baixo) com $\gamma = 0.5$. Espaço de opiniões e de ϵ discretizados em 100 partes.

Observe que, para todos os valores de γ , a posição $O = B$ sempre é ocupada, independente de ϵ . O principal destaque, no entanto, diz respeito à posição em que consenso é obtido: à medida que a intensidade do campo, γ , aumenta *consenso* é menos favorecido e um cluster de opinião contrário ao campo permanece para valores altos de ϵ . Trata-se, portanto, de uma diferença importante, mas, a princípio contra-intuitiva. Este comportamento é explicado na Figura a seguir, no contexto de redes.

Figura 5.12 – Diagrama de bifurcação com as distribuições estacionárias para o modelo de KH na presença de campo externo considerando diferentes valores de γ . Resultados obtidos a partir da solução da Equação 5.7 reescrita em termos da Cadeia de Markov em 5.8. Valores de γ variam de 0.1 a 0.9 (esquerda para direita, de cima para baixo). Espaço de opiniões e de ϵ discretizados em 100 partes.

5.2.2 Efeito do campo externo na dinâmica de KH em Redes Complexas

Da mesma forma como fizemos quando consideramos a presença de ruído, investigaremos o efeito da inclusão do campo quando a dinâmica de KH ocorre em redes. Serão considerados

os modelos de Erdős-Rényi (ER) e o modelo social de Boguñá (SN).

A Figura 5.13 exemplifica o efeito do parâmetro α na evolução temporal das opiniões sob efeito de um campo constante na posição $B = 0.25$. O resultado foi obtido para rede ER de tamanho $N = 500$ e $\langle k \rangle = 10$ quando $\epsilon = 0.2$. Assim, como já observado na Figura anterior, note que quando o campo é menos intenso ($\gamma=0.1$) consenso é obtido na posição do campo, enquanto para $\gamma = 0.8$ o sistema encontra-se quase polarizado. É preciso lembrar que em cada passo da simulação, cada agente interage com seus vizinhos e aceita o campo externo, com peso α , desde que sua opinião não seja distante de B mais do que ϵ . Assim, na situação criada na Figura 5.13, caso o campo seja muito intenso, o grupo de agentes próximos de B são rapidamente atraídos para a posição do campo e se distanciam do cluster formado na parte superior do eixo de opiniões. Como consequência, dois clusters principais são formados. Por outro lado, se o campo é intenso o suficiente para atrair os agentes, mas pequeno a ponto de permitir que eles se aproximem, então um único grande cluster é formado e seu efeito acaba sendo mais efetivo.

Figura 5.13 – Exemplos da evolução temporal da dinâmica de KH sob efeito de campo externo com intensidades $\gamma = 0.8$ (esquerda) e $\gamma = 0.1$ (direita). Foi considerada rede ER de tamanho $N = 500$ e $\langle k \rangle = 10$. Utilizado parâmetro de confiança $\epsilon = 0.2$ e campo na posição $B = 0.25$ (linha vermelha).

De modo a quantificar de maneira mais geral essas particularidades, propomos a definição de uma medida que indique a efetividade, E_B , de um campo externo aplicado:

$$E_B = \sum_{i \in N, |i-B| \leq \epsilon} \frac{1}{N}. \quad (5.10)$$

Ou seja, E_B nada mais é do que a fração de agentes com opiniões compatíveis com o campo. A Figura 5.14 apresenta um mapa para a efetividade do campo, E_B , e para o número de clusters obtidos no estado estacionário em termos do valor da intensidade do campo γ e do parâmetro de confiança ϵ considerando campo na posição $B = 0.25$. Os gráficos comparam os modelos ER e SN para tamanhos de rede $N = 500$ e $\langle k \rangle = 10$. De maneira geral, o resultado mostrado no gráfico nos permite chegar às seguintes evidências:

- 1) Dependência com ϵ : conforme o parâmetro de confiança diminui, o número de clusters aumenta e a eficiência do campo diminui. De fato, quanto menor o valor de ϵ menor é a interação dos agentes entre si e com o campo, levando à fragmentação.
- 2) Dependência com γ : para um dado valor de ϵ , E_B diminui a medida que γ cresce e o número de clusters aumenta. Esta observação corrobora com os resultados encontrados nas Figuras anteriores.
- 3) No modelo social, a efetividade do campo, E_B , apresenta valores maiores e o número de clusters é ligeiramente menor.

Sobre a observação **3**, devemos lembrar que a dinâmica original de KH sobre o modelo SN naturalmente produz maior diversidade de opiniões, quando comparado com o modelo ER. Ao adicionarmos mudanças no modelo KH, seja o ruído na Seção anterior, seja o campo externo, obtivemos resultados diferentes: para o ruído o modelo social mostrou-se robusto, no sentido que a quantidade de clusters aumenta pouco; já para o campo, ou propaganda, mais agentes tendem a seguir a opinião externa, e portanto, sua ação é mais efetiva nestas redes do que no modelo aleatório.

Figura 5.14 – Dependência da efetividade do campo, E_B (parte superior) e do número de clusters, $N_{clusters}$ (parte inferior) para redes ER e SN de tamanho $N = 500$ e $\langle k \rangle = 10$ com campo mantido na opinião $B = 0.25$. Média sobre 50 realizações das redes e da dinâmica.

Acessibilidade na determinação da eficiência do campo externo

Até agora verificamos o efeito do campo externo quando aplicado continuamente sobre todos os vértices da rede, conforme as regras definidas nas Seções anteriores. Gostaríamos de explorar agora a situação em que o campo age somente sobre um único vértice da rede. Neste caso, o agente afetado sempre sente o efeito da opinião externa e a considera com peso γ associado à média das opiniões dos vizinhos compatíveis, como na Seção anterior. Esta restrição exemplifica os casos de grupos de agentes isolados em que apenas um indivíduo tem acesso a uma informação vinda do ambiente externo.

Assim, o objetivo aqui é verificar quais agentes numa rede são mais eficazes no processo de propagação de uma opinião externa. Como exemplo, a Figura 5.15 mostra a evolução temporal das opiniões de agentes conectados através de uma rede gerada pelo modelo SN quando o campo

age sobre o agente de maior (direita) e menor grau (esquerda). Conforme esperado, quando o campo influencia o agente mais conectado, menos clusters são formados e a efetividade do campo é maior. Embora este seja um resultado esperado, e intuitivamente sabemos que os hubs devem ser mais eficientes na propagação de uma informação, seria o grau um bom parâmetro para medir a influência de um agente na dinâmica de KH?

Figura 5.15 – Exemplos da evolução temporal da dinâmica de KH numa rede SN quando campo, na posição $B = 0.5$, é aplicado sobre agente de maior (esquerda) e menor grau (direita). Considerada rede de tamanho $N = 300$ e $\langle k \rangle = 10$ para $\varepsilon = 0.15$. Linha vermelha e verde indicam a posição do campo e do agente afetado, respectivamente. Níveis de cinza indicam a concentração de agentes ao longo do espaço de opiniões.

A resposta para essa pergunta é investigada na Figura 5.16. Para uma rede ($N = 300$, $\langle k \rangle = 10$) formada a partir do modelo SN, aplicamos um campo, de intensidade $\gamma = 0.5$ e posição $B = 0.5$ sobre cada agente, individualmente, e verificamos o valor da efetividade do campo na rede no regime estacionário. A Figura mostra a relação entre E_B e três medidas que caracterizam os agentes: i) acessibilidade, A_k ; ii) grau, k ; e iii) a medida de centralidade betweenness, C_B . Fica claro que existe uma relação bastante mais definida para a acessibilidade do que para as outras medidas. De fato, o coeficiente de correlação entre E_B e A_k se aproxima da unidade. Assim, para o processo de propagação de informação definido, mais importante do que a conectividade imediata de um agente, a homogeneidade de acessos num raio dado pelo diâmetro da rede ($h = 5$) é fundamental.

Figura 5.16 – Dependência da efetividade do campo, E_B , em relação à (a) acessibilidade, (b) grau e (c) betweenness do agente sob efeito da opinião externa. Considerado modelo de rede SN de tamanho $N = 300$ e $\langle k \rangle = 10$.

6 AQUISIÇÃO DE CONHECIMENTO NA PRESENÇA DE ERROS DE OBSERVAÇÃO

O modelo de KH pressupõe a existência da propriedade de confiança limitada, representada pelo parâmetro ε . Essa característica é o fator gerador de clusters de diferentes opiniões na situação estacionária. Em muitas situações no mundo real, no entanto, *consenso* entre pequenos grupos de agentes interagentes é uma condição imperativa. Mais do que isso, em processos de *negociação*, os agentes são os maiores interessados em obterem um acordo. Dentre os mais diversos exemplos, podemos citar o processo de valoração na aquisição de ativos, decisões de comitês em ambientes corporativos ou ainda o desenvolvimento de modelos científicos nos grupos de pesquisa.

Nesta seção investigaremos o problema de como um ou mais agentes progridem enquanto modelam um sistema complexo. De fato, por causa da complexidade dos problemas em ciência, a construção de modelos torna-se intrinsecamente dependente da colaboração entre cientistas ou *agentes*. O problema de aquisição de conhecimento e processamento multi-agente é tratado na literatura (por exemplo, (43, 44)), mas, na maioria dos casos, assumindo esquemas simples de interação entre agentes (reticulados, mistura homogênea). Assim, consideramos a situação onde vários agentes interagem através de uma rede de relacionamentos (Figura 6.1). Agentes investigando um sistema realizam observações dos componentes desse sistema a medida que desenvolvem e complementam seus respectivos modelos individuais. Erros, ruídos, incerteza e esquecimento estão tipicamente envolvidos durante a estimação do modelo. As principais

características de interesse incluem a qualidade dos modelos obtidos e o tempo respectivo para adquirí-lo. Embora corresponda a uma versão extremamente simplificada da investigação científica real, essa abordagem captura alguns dos principais elementos caracterizando o envolvimento de agentes interagentes os quais continuamente trocam informação e modificam seus modelos.

Desta forma, construímos um modelo de aquisição de conhecimento também baseado num processo de médias, assim como na dinâmica de KH, mas agora, agentes possuem não somente uma única variável, mas consideram um conjunto de estados representando seu conhecimento sobre os componentes de um sistema sendo modelado. Embora os agentes continuamente compartilhem suas opiniões, é assumido que a estimação dos diferentes componentes do sistema ocorre de maneira independente. Além disso, consideramos que um agente específico interage com todos os seus vizinhos sem levar em conta a propriedade de confiança limitada. A habilidade de poder fazer observações ao longo do tempo é um ingrediente adicional que enriquece a modelagem e continuamente reforça a posição das estimações de cada agente contrabalancando a tendência de tornar-se similar aos demais após cada interação. É importante lembrar que o uso de um vetor de opiniões foi seminalmente estudado no modelo proposto por Axelrod (16) no contexto de dinâmica social. Neste modelo, as opiniões representam características culturais e evoluem de maneira acoplada. Ainda, em (45), os autores investigam condições para consenso na aprendizagem de agentes modelados com perceptrons.

6.1 Definições

A primeira hipótese importante é que o sistema complexo sob estudo seja representado como uma estrutura discreta (grafo), e que as diversas partes deste sistema possam ser observadas, possivelmente com uma dada probabilidade de erro. Por *agente*, entende-se qualquer entidade capaz de realizar observações/medidas do sistema sob análise e armazenar esta informação por um período de tempo. Os erros envolvidos nas observações podem ser de vários

Figura 6.1 – Agentes (cientistas) desenvolvem seus modelos através de observações e interações.

tipos, por exemplo, observar a presença de uma conexão onde não existe. Tais erros são uma consequência direta da acurácia limitada dos dispositivos de medidas bem como das prioridades ou tendências de cada agente.

Desta forma, nossa investigação inclui 4 componentes básicos: (i) o sistema complexo sob análise S ; (ii) um ou mais agentes capazes de realizar observações do sistema S (sujeitas a erro) e de interagirem entre si; (iii) a rede descrevendo o sistema S ; e (iv) a rede de interações entre os agentes.

O caso mais simples é quando um único agente, isoladamente, observa os elementos do sistema sob estudo. Uma vez que este sistema é descrito por uma rede complexa, uma observação consiste em detectar a presença, ou não, de links entre os pares de vértices do sistema. A predição de cada aresta envolve uma probabilidade de erro γ , isto é, observar uma conexão onde não existe e vice-versa. Um possível procedimento adotado pelo agente envolveria levar em conta a média sobre todas as observações individuais até o tempo corrente T , isto é:

$$\langle x \rangle^T = \frac{1}{T} \sum_{t=1}^T x^t, \quad (6.1)$$

onde x é o valor específico da aresta (0 caso não exista e 1 no caso contrário) e x^t é a observação

de x no tempo t . Observe que estamos considerando γ independente do tempo e das partes do sistema sob análise. O erro da estimação de uma aresta, após o tempo T , pode ser definida como:

$$\varepsilon(x)^T = \left| x_S - \langle x \rangle^T \right|, \quad (6.2)$$

onde x_S é o valor original da aresta. E, obviamente

$$\lim_{T \rightarrow \infty} \varepsilon(x)^T = \gamma. \quad (6.3)$$

Assim, a melhor representação da conectividade alcançada por um agente possui erro γ , considerando um intervalo longo de tempo. No caso mais geral, um total de N_a agentes interagem enquanto realizam observações do sistema ao longo do tempo. Da mesma forma, estas observações possuem probabilidade de erro igual a γ_i para um agente específico i . O sistema sob estudo é representado como uma rede complexa cuja matriz adjacência é representada por A_S . A evolução da aprendizagem dos agentes segue a seguinte dinâmica: em cada tempo t , cada agente observa a conectividade do sistema complexo com erro γ_i , levando à matriz adjacência O_i^t , e também recebe as matrizes correntes K_j^t dos seus vizinhos j (conjunto de vértices diretamente conectados a i). O agente i então calcula a matriz média $\langle K \rangle_i^t$ das matrizes K_j^t , isto é:

$$\langle K \rangle_i^t = \frac{1}{k_i} \sum_{j \in \Gamma_i} K_j^t, \quad (6.4)$$

onde Γ_i é o conjunto de vizinhos do agente i e k_i é o seu grau. O agente i então realiza uma média ponderada entre sua matriz corrente K_i^t e a matriz média dos vizinhos:

$$V_i^t = [a \langle K \rangle_i^t + (1 - a) K_i^t], \quad (6.5)$$

onde $0 < a \leq 1$ é o peso relativo. Assumiremo $a = 0.5$, de modo que V_i^t se torna igual à média entre as matrizes corrente e dos vizinhos em cada tempo. Cada agente i , então adiciona este

valor em sua observação corrente, isto é:

$$K_i^{t+1} = V_i^t + O_i^t, \quad (6.6)$$

de modo que depois do tempo T a estimação da matriz adjacência pelo agente i pode ser obtida como

$$A_i^T = K_i^T / T. \quad (6.7)$$

Com esta formulação, a matriz corrente K_i^t é uma variável contínua e ilimitada controlada por um processo de médias com os vizinhos e pelo acúmulo das observações binárias. Entretanto, a matriz adjacência A_i^t aprendida pelo agente i possui entradas com valores no intervalo contínuo $[0, 1]$. Esta dinâmica discreta leva à consequências importantes. Ao fazer a média entre sua observação corrente e a estimação média dos vizinhos, o erro limite pode ser modificado com respeito ao caso de um único agente. É importante notar também que as variáveis foram naturalmente definidas como matrizes, uma vez que os agentes aprendem sobre os componentes de uma rede, a qual é naturalmente representada por uma matriz adjacência. De forma equivalente, esta definição pode ser alterada e a descrição matemática pode ser feita utilizando vetores de observação, ao invés de matrizes. As entradas dos vetores representariam a estimação de uma aresta em uma rede, ou num contexto mais geral, qualquer informação binária a ser observada.

Diversas variações e elaborações deste modelo são possíveis, incluindo a consideração de ruído enquanto observações são transmitidas, degradação no tempo, adoção de outros valores para a , bem como a evolução no tempo da rede de agentes incluindo diferentes níveis de afinidade entre eles. Além disso, o parâmetro γ poderia depender não somente do agente realizando a observação, mas também do componente do sistema sendo modelado, em nosso caso, cada elemento da matriz adjacência A_S . Como consequência, a probabilidade de erro durante a estimação, γ , seria também representada por uma matriz. Outra alternativa é a possibilidade da predição de links omitidos no sistema estudado. Este tipo de problema é investigado por Clauset et al. (31) levando em conta a estrutura hierárquica presente em redes reais.

Para a dinâmica de aprendizagem a performance de cada agente é quantificada em termos do erro entre a rede original e os modelos obtidos pelo agente após um tempo suficientemente longo. Este erro é calculado como:

$$\epsilon_i^T = \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N |A_i^{T,m,n} - A_S^{m,n}|, \quad (6.8)$$

onde A_i^T é a representação da matriz adjacência do modelo do agente i no tempo T . A performance global do conjunto de agentes é expressa em termos da média dos erros individuais pesados pelos respectivos graus. Chamaremos esta quantidade de *erro global* do sistema:

$$E^T = \frac{1}{\sum_{i=1}^{N_a} k_i} \sum_{i=1}^{N_a} k_i \epsilon_i^T. \quad (6.9)$$

6.2 Descrição analítica para o Erro Global

O valor estacionário para o Erro Global (Eq. 6.9) pode ser obtido a partir da equação diferença que caracteriza a dinâmica. Assim, para cada agente i , temos:

$$2K_i^{t+1} = \left[\frac{1}{k_i} \sum_{j \in \Gamma_i} K_j^t \right] + K_i^t + 2O_i^t, \quad (6.10)$$

onde K_i^t é a matriz corrente e O_i^t é a observação do agente i no tempo t . O termo entre colchetes é a média das matrizes recebidas do conjunto Γ_i de vizinhos de i . Considerando a aproximação contínua, a equação anterior pode ser reescrita como:

$$2 \frac{d}{dt} K_i(t) = \left[\frac{1}{k_i} \sum_{j \in \Gamma_i} K_j(t) \right] - K_i(t) + 2O_i(t). \quad (6.11)$$

Dividindo por t e lembrando que $\frac{1}{t} \frac{d}{dt} K_i(t) = \frac{d}{dt} \left(\frac{K_i(t)}{t} \right) + \frac{1}{t^2} K_i(t)$, nós temos

$$2k_i \frac{d}{dt} \left[\frac{K_i(t)}{t} \right] + 2k_i \frac{K_i(t)}{t^2} = \left[\sum_{j \in \Gamma_i} \frac{K_j(t)}{t} \right] - k_i \frac{K_i(t)}{t} + 2k_i \frac{O_i(t)}{t} \quad (6.12)$$

e somando, para cada agente i :

$$\frac{d}{dt} \sum_{i=1}^{N_a} \left[\frac{k_i K_i(t)}{t} \right] = \frac{1}{t} \sum_{i=1}^{N_a} k_i O_i(t) - \frac{1}{t} \sum_{i=1}^{N_a} \left[\frac{k_i K_i(t)}{t} \right], \quad (6.13)$$

já que $\sum_i \sum_{j \in \Gamma_i} \frac{K_j(t)}{t} = \sum_i \frac{k_i K_i(t)}{t}$. Finalmente, define-se $J \equiv \sum_i \frac{k_i K_i(t)}{t}$:

$$t \frac{d}{dt} J(t) = \sum_{i=1}^{N_a} k_i O_i(t) - J(t) \quad (6.14)$$

Então, após o tempo T :

$$J(T) = \sum_{i=1}^{N_a} k_i \left[\frac{1}{T} \int_0^T O_i(t) dt \right] \quad (6.15)$$

Para um agente i a probabilidade de observar uma aresta onde não existe e vice-versa é γ_i . Assim, para uma entrada $A_S^{m,n}$ do sistema sob consideração e para T muito grande, a média da observação será:

$$\lim_{T \rightarrow \infty} \left[\frac{1}{T} \int_0^T O_i^{m,n}(t) dt \right] = \begin{cases} (1 - \gamma_i) & \text{se } A_S^{m,n} = 1 \\ (\gamma_i) & \text{se } A_S^{m,n} = 0 \end{cases} \quad (6.16)$$

Aplicando este resultado na equação 6.15, temos :

$$\lim_{T \rightarrow \infty} J^{m,n}(T) = \begin{cases} \sum_i k_i (1 - \gamma_i) & \text{se } A_S^{m,n} = 1 \\ \sum_i k_i (\gamma_i) & \text{se } A_S^{m,n} = 0 \end{cases} \quad (6.17)$$

Subtraindo $\sum_i k_i A_S^{m,n}$ e somando para todo m and n :

$$\sum_{m=1}^N \sum_{n=1}^N \left| \lim_{T \rightarrow \infty} \left(J^{m,n}(T) - \sum_{i=1}^{N_a} k_i A_S^{m,n} \right) \right| = \sum_{i=1}^{N_a} \left(\sum_{m=1}^N \sum_{n=1}^N \gamma_i k_i \right)$$

$$\sum_{i=1}^{N_a} k_i \left\{ \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N \left| \lim_{T \rightarrow \infty} \left(\frac{K_i^{m,n}(T)}{T} - A_S^{m,n} \right) \right| \right\} = \sum_{i=1}^{N_a} k_i \gamma_i \quad (6.18)$$

para ambos os casos. O termo na somatória é precisamente o erro individual de um agente i pesado pelo seu grau. Então, dividindo pela soma do graus, nós temos a expressão para o error global do sistemas (E) para tempos longos:

$$\begin{aligned} \frac{\sum_{i=1}^{N_a} k_i \epsilon_i}{\sum_i k_i} &= \frac{\sum_{i=1}^{N_a} k_i \gamma_i}{\sum_i k_i} \\ \Rightarrow E &= \frac{1}{\langle k \rangle N_a} \sum_{i=1}^{N_a} k_i \gamma_i \end{aligned} \quad (6.19)$$

6.3 Características gerais

Observando o valor do erro global para tempos longos obtido na Seção anterior (Eq. 6.19), percebe-se imediatamente que no caso de todos os agentes possuírem a mesma probabilidade de erro nas observações, γ , o o erro global também tende a γ .

Podemos, ainda, aplicar a equação obtida nas duas seguintes situações, assumindo uma rede de agentes com grau médio $\langle k \rangle$ e graus máximo e mínimo iguais à k_{max} e k_{min} , respectivamente:

(a) um agente j com grau k_j possuindo probabilidade de erro na observação igual a $\gamma_j = p\gamma$ ($0 \leq p \leq 1/\gamma$) e todos os outros agentes com probabilidade γ :

$$\begin{aligned} E &= \frac{\left(\sum_{i=1}^{N_a} k_i \gamma \right) - \gamma k_j + p \gamma k_j}{\langle k \rangle N_a} \\ E &= \frac{k_j \gamma (p - 1)}{\langle k \rangle N_a} + \gamma \end{aligned} \quad (6.20)$$

a qual é uma relação linear com respeito ao grau do agente com probabilidade de erro diferente.

Note que não há diferença entre quaisquer duas redes de agentes com mesmo grau médio e mesmo número de vértices.

(b) Um agente j com grau k_j possuindo probabilidade de erro de observação proporcional ao seu grau $\gamma_j (\gamma \leq \gamma_j \leq 1)$ e todos os outros agentes com probabilidade γ :

$$E = \frac{k_j(1-\gamma)(k_j - k_{min})}{\langle k \rangle N_a(k_{max} - k_{min})} + \gamma \quad (6.21)$$

a qual é uma relação quadrática com respeito ao grau do agente com probabilidade diferenciada.

No caso mais geral, em que γ é uma variável aleatória, o erro global pode ser reescrito como:

$$E = \frac{\sigma_k \sigma_\gamma \text{CORR}(k, \gamma)}{\langle k \rangle} + \langle \gamma \rangle, \quad (6.22)$$

onde σ_k e σ_γ são desvios padrão do grau e das probabilidades γ , respectivamente. A Figura 6.2

Figura 6.2 – Erro global em termos do desvio padrão dos graus para rede ER e BA. Probabilidades de erro nas observações perfeitamente correlacionadas com o grau. Cada ponto do gráfico representa a realização de um modelo de rede com $N_a = 200$ e $\langle k \rangle = 5.8$.

6.3.1 Efeito produzido por um único agente

Nesta seção verificaremos o efeito no erro global causado por um único agente através de duas situações: i) um dos agentes possui metade ou o dobro da probabilidade de erro em relação ao demais; e ii) um dos agentes possui erro proporcional ao seu grau. Simulações foram feitas considerando γ diferenciado individualmente em cada um dos possíveis agentes em redes de agentes do tipo Erdős-Rényi (ER) and Barabási-Albert (BA).

A Figura 6.3(a) mostra os resultados obtidos considerando probabilidade $\gamma = 0.2$ para um dos agentes e $\gamma = 0.4$ para os demais. É possível identificar uma diferença substancial da qualidade média dos dos modelos, os quais dependem linearmente do grau dos com probabilidade de erro diferente. Mais especificamente, fica claro que os erros globais são muito menores para os agente com graus mais altos. Em outras palavras, os melhores modelos são obtidos quando os hubs possuem probabilidades de erro menores, influenciando fortemente os demais agentes através da difusão das suas estimativas.

A Figura 6.3(b) mostra os resultados obtidos quando o agente diferenciado possui probabilidade $\gamma = 0.4$ enquanto para os demais agentes $\gamma = 0.2$. Neste caso, efeito oposto é verificado, com o erro global crescendo linearmente com o grau do agente diferenciado. Além disso, como esperado, nenhuma diferença é verificada entre os resultados obtidos para o modelo ER e BA.

A Figura 6.4 mostra os resultados obtidos quando um dos agentes possui probabilidade de erro maior que os demais e proporcional ao seu grau. Resultados similares ao caso anterior são obtidos, mas com comportamento não linear. Em adição, de acordo com as previsões analíticas, os resultados mostraram-se diferentes para os modelos ER e BA (ER apresenta valores maiores do que aqueles encontrados para rede BA). Isto significa que quando consideramos dois agentes com o mesmo grau, um em cada rede, o erro global é menor no modelo BA já que seus hubs influenciam mais fortemente no aumento da qualidade dos modelos estimados.

Uma abordagem mais clara da influência do grau na estimação dos modelos feita pelos

Figura 6.3 – Erro global em termos do grau dos agentes com erro de observação diferente dos demais para os modelos ER e BA com $N_a = 200$ and $\langle k \rangle = 9.4$: (a) um dos agentes com taxa de erro igual a 0.2 e os outros agentes com $\gamma = 0.4$; e (b) um dos agentes com 0.4 e os demais com $\gamma = 0.2$. A linha pontilhada representa a solução analítica. Resultados obtidos com médias sobre 100 realizações da dinâmica.

agentes é feita na Figura 6.5. Ela apresenta a situação em que o agente de maior grau, j , possui probabilidade de erro igual ao dobro das probabilidades dos demais agentes ($\gamma_j = 0.4$). As Figuras 6.5(a) and 6.5(c) consideram o ε_i^T de cada agente, individualmente, em termos da seu respectivo grau para os modelos ER e BA, ambos de tamanho $N_a = 300$ e grau médio igual a 4.7. Círculos preenchidos representam a vizinhança imediata do agente diferenciado. Fica claro que a estimação de todos os agentes é afetada, especialmente para aqueles contidos na vizinhança do agente com maior probabilidade de erro. Observe também que, em ambos os casos, o maior erro individual vem do próprio agente diferenciado e os nós com graus mais altos tendem a ser menos influenciados. A Figuras 6.5(b) and 6.5(d) complementam a análise considerando o valor médio dos erros individuais dos agentes com a mesma distância topológica do agente com maior probabilidade de erro. Os resultados mostram que agentes mais próximos tendem a ser mais afetados do que os periféricos. Assim, vemos que a presença de um agente com estimativas piores produz efeito não uniforme entre os agentes. Uma vez que o erro individual de um agente é obtido através de um process de médias do seu modelo com a estimação dos seus vizinhos, o contato direto (e a proximidade) com o agente de maior probabilidade de erro

Figura 6.4 – Erro global em termos do grau dos agentes com maiores erros de observação: um dos agentes com erro proporcional ao seu grau e todos os outros agentes com $\gamma = 0.2$. As linhas pontilhadas e contínuas representam a solução analítica para os modelos ER e BA, respectivamente. Simulações para uma única realização de cada modelo de rede com $N_a = 200$ and $\langle k \rangle = 9.4$. Resultados obtidos com médias sobre 50 realizações da dinâmica.

altera consideravelmente seu modelo final.

Ainda no sentido de verificar a influência da conectividade no processo de aquisição de conhecimento, consideramos a evolução temporal dos erros individuais numa única realização da dinâmica. As Figuras 6.6(a) e 6.6(b) mostram esta avaliação para uma rede BA quando os agentes de maior e menor grau, respectivamente, possuem o dobro da probabilidade de erro dos demais agentes. Vemos que inicialmente existe uma diferença significativa entre os modelos estimados, a qual é reduzida à medida que os agentes interagem. Uma observação mais detalhada dos agentes com mesma probabilidade de erro (parte de baixo) revela que seus modelos são mais influenciados quando o agente diferenciado possui grau maior.

Figura 6.5 – Erros individuais em termos do grau dos agentes no caso em que agente com maior grau possui error de observação igual a 0.4 e todos os outros com $\gamma = 0.2$ a and c): $\square$, o agente com a maior taxa de erro; $\bullet$, a vizinhança deste agente; $\circ$, demais agentes. Figuras (b e c) mostram o valor médio dos erros individuais para o conjunto de agentes com a mesma distância topológica do agente com maior taxa de erro. Simulações para modelos ER e BA com $N_a = 300$ e $\langle k \rangle = 4.7$. Resultados obtidos a partir de médias sobre 50 realizações da dinâmica.

6.3.2 Variando erros de observação em estruturas com comunidades

Após estudar como um único agente pode afetar a performance de toda a rede, nós podemos estender nossa análise para quantificar a influência de um conjunto de vértices quando a dinâmica de aquisição é aplicada em redes de agentes com estrutura de comunidades. Em nossa modelagem, as redes de agentes sintéticas com comunidades foram geradas usando o algoritmo proposto por Lancichinetti et al. (46). Este algoritmo é capaz de gerar redes com distribuições heterogêneas tanto dos graus quando do tamanho das comunidades. Além disso, a fração de arestas que cada agente compartilha com os agentes de outras comunidades pode também ser

Figura 6.6 – *Evolução temporal dos erros individuais considerando o agente com o menor (a) e maior (b) grau tendo um erro de observação igual a 0.4 e todos os outros agentes com $\gamma = 0.2$. Acima: todos os agentes são considerados (linha pontilhada representa o a agente com maior erro). Abaixo: somente os agentes com a mesma taxa de erro (linhas pretas) e o valor médio de seus erros individuais. Simulações para modelo BA com $N_a = 200$ and $\langle k \rangle = 9.4$.*

especificads por um parâmetro de entrada, μ , o *parâmetro de mistura*.

Durante as simulações todos os agentes de uma das comunidades possui o dobro da taxa de erro dos demais (foi mantido igual a 0.4). Nós analisamos como a inclusão de mais links entre os agents com maiores erros de observação influencia os erros individuais dos agentes nas outras comunidades. A dinâmica foi realizada em redes com $N_a = 256$ agentes considerando variações no parâmetro de mistura, na distribuição de grau e no tamanho da comunidade com maior taxa de erro.

Figura 6.7(a) mostra os resultados da dinâmica aplicada para realizações de redes de agentes com 4 comunidades de mesmo tamanho a todos os agentes com mesmo grau (igual a 16), mas variando o parâmetro de mistura. Para cada rede nós consideramos uma das comunidades possuindo uma erro de observação maior. Como mostrado nesta figura, a adição de arestas entre os agentes desta comunidade aumenta os erros individuais dos agentes nas outras comunidades e este efeito cresce com o número de arestas entre as comunidades (o valor de μ). Uma vez que a dinâmica é baseada num processo de médias, no sentido de que os modelos individuais são atualizados considerando os modelos da vizinhança, os agentes com maiores erros de observação

tendem a estimar com pior qualidade quando a comunicação intra-comunidade é intensificada. De fato, as novas arestas implicam em mais agentes com maiores erros influenciando no desenvolvimento das suas estimações. Finalmente, a perda de performance é propagada para outras comunidades através das conexões inter-comunidades

Resultados similares são mostrados nas figuras 6.7(b) e 6.7(c). Para a primeira figura, consideramos uma única realização da rede de agentes na qual todos possuem o mesmo grau, igual a 16, e parâmetro de mistura igual a 0.2, enquanto os agentes foram agrupados em 5 comunidades de diferentes tamanhos. Simulações independentes foram realizadas considerando todos os agentes de cada comunidade possuindo uma maior taxa de erro. Novamente, uma correlação positiva foi encontrada entre o número de arestas adicionadas entre os agentes na comunidade de maior erro de observação e o valor médio dos erros individuais para os outros agentes. Como esperado, este efeito aumenta com o tamanho da comunidade com maior taxa de erro.

Figura 6.7(c) mostra os resultados para agentes agrupados em 4 comunidades de mesmo tamanho com μ igual a 0.2. Os graus dos agentes seguem distribuição na forma de lei de potência (com expoente $\lambda = -2$ e $\langle k \rangle = 16$), de modo que cada comunidade possui grau médio diferente. Como anteriormente, a influência de cada comunidade foi considerada independentemente, enquanto foi medido o valor médio dos erros individuais dos agentes nas comunidades de menor erro. Uma vez que as comunidades possuem o mesmo tamanho e os agentes possuem, muito aproximadamente, a mesma fração de arestas com agentes em outras comunidades, quanto maior o grau médio da comunidade, maior o número de conexões inter-comunidades ligado a esta comunidade. Isto explica o aumento dos erros individuais com o grau médio da comunidade com menor acurácia.

Figura 6.7 – Dependência entre o número de arestas adicionadas na comunidade com maior erro de observação e o valor médio dos erros individuais dos agentes nas outras comunidades para redes com $N_a = 256$ agentes: (a) todos os agentes com o mesmo grau, igual a 16, agrupados em 4 comunidades de mesmo tamanho e variando o parâmetro de mistura; (b) todos os agentes com o mesmo grau, igual a 16, agrupados em 5 comunidades com diferentes tamanhos para $\mu = 0.2$; (c) agentes com graus distribuídos como lei potência e agrupados em 4 comunidades de mesmo tamanho considerando $\mu = 0.2$. Resultados obtidos a partir de médias sobre 100 realizações da dinâmica.

7 CONCLUSÃO

Neste trabalho investigamos a dinâmica de Krause-Hegselmann no contexto de redes complexas e uma série de importantes resultados foram discutidos ao longo dos capítulos 4, 5 e 6. Iniciamos a investigação verificando o efeito do modelo original KH adaptado para redes. Nesta etapa comparamos os resultados entre os modelos de rede de Erdős-Rényi (ER), Barabási-Albert (BA) e Lattice circular (LA). Pudemos verificar que a topologia da rede exerce bastante efeito sobre a dinâmica de KH. Baseado na distribuição de clusters de opinião, verificamos a existência de três fases estruturais: i) fase monopolar, marcada pela existência de único cluster macroscópico; ii) fase multipolar, marcada pela existência de clusters macroscópicos associados à clusters microscópicos e pela extensividade com o tamanho da rede e ; iii) fase fragmentada, caracterizada pela proliferação de clusters de tamanho reduzido. As três fases mostraram-se presentes nos modelos ER e BA. No lattice, verificamos apenas duas destas fases, com transição direta da condição monopolar para a fragmentada. Mostramos ainda a relação dessas propriedades com o processo de percolação em redes, identificando que a transição ocorre para valores menores de ε a medida que a dispersão de grau da rede aumenta.

No Capítulo 5 estudamos duas modificações para a dinâmica de KH: a existência de ruído na atualização das opiniões e a ação de um campo externo. Para o primeiro caso, observamos que o efeito do ruído, de maneira geral, é de homogeneização da ocupação do espaço de opiniões. Para este caso, reescrevemos a Cadeia de Markov definida no Capítulo 3. No contexto de redes, estudamos o ruído agindo por tempo limitado sobre uma configuração de clusters já formados. Após levantar um série de medidas, verificamos que modelo social de Boguñá apresentou menor diversificação das opiniões quando submetido ao ruído. Por outro lado, no caso da aplicação

de campo constante, as redes sociais apresentaram valores maiores para a eficiência do campo produzindo maior homogeneização. Quando o campo foi aplicado sobre um único indivíduo, verificamos que a acessibilidade se mostrou uma boa medida para identificação dos agentes com maior capacidade de propagação da opinião externa.

Por fim, no Capítulo 6, propomos um modelo de consenso bastante simplificado para representação da aquisição de conhecimento, no qual agentes interagem enquanto fazem estimativas na natureza. Estas observações estão sempre sujeitas a um erro. Verificamos, inclusive analiticamente, como o erro global da rede é afetado quando um único vértice possui probabilidade de erro de observação maior ou menor que dos demais agentes.

O trabalho realizado aqui deixa possibilidades para outros no futuro. Por exemplo, a utilização da solução numérica via equação mestra foi explorada apenas no contexto em que agentes interagem numa mistura homogênea. Existiria a possibilidade de reescrever as equações nas situações em que agentes interagem de acordo com um modelo de rede, dada sua distribuição de grau, de modo que, a distribuição das opiniões seria também função do grau dos agentes, ou seja, $P(O) \rightarrow P_k(O)$. Além da possibilidade da solução numérica em si, uma eventual descrição analítica abriria a possibilidade de investigação das propriedades da dinâmica paralelamente ao de simulações. Um outra importante extensão do trabalho seria a consideração de valores heterogêneos para ε , variando de indivíduo para indivíduo. De fato, da mesma forma que a heterogeneidade das propriedades topológicas geram efeitos sofisticados em redes, a variância dos parâmetros dinâmicos também poderia mudar drasticamente o comportamento dinâmica.

REFERÊNCIAS

- 1 GALTON, F. Vox populi. *Nature*, v. 75, n. 7, p. 45, 1907.
- 2 PAGE, S. E. *The difference: how the power of diversity creates better groups, firms, schools, and societies*. Princeton: Princeton University Press, 2007.
- 3 SUROWIECKI, J. *The wisdom of crowds*. New York: Doubleday, 2004.
- 4 MANNES, A. E. Are we wise about the wisdom of crowds? the use of group judgments in belief revision. *Management Science*, v. 55, n. 8, p. 1267–1279, 2009.
- 5 RAUHUT, H.; LORENZ, J. The wisdom of crowds in one mind: how individuals can simulate the knowledge of diverse societies to reach better decisions. *Journal of Mathematical Psychology*, v. 55, n. 2, p. 191–197, 2010.
- 6 LORENZ, J. et al. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences of the United States of America*, v. 108, n. 22, p. 9020–9025, 2010.
- 7 GLAUBER, R. J. Time-dependent statistics of the ising model. *Journal of Mathematical Physics*, v. 4, n. 294, p. 1–14, 1963.
- 8 CASTELLANO, C.; FORTUNATO, S.; LORETO, V. Statistical physics of social dynamics. *Reviews of Modern Physics*, v. 81, n. 2, p. 591–646, 2009.
- 9 CLIFFORD, P.; SUDBURY, A. A model for spatial conflict. *Biometrika*, v. 60, n. 3, p. 581–588, 1973.
- 10 HOLLEY, R. A.; LIGGETT, T. M. Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.*, v. 3, n. 4, p. 643–663, 1975.
- 11 SZNAJD-WERON, K.; SZNAJD, J. Opinion evolution in closed community. *International Journal of Modern Physics C*, v. 11, n. 6, p. 1157–1165, 2000.
- 12 GALAM, S. Minority opinion spreading in random geometry. *The European Physical Journal B*, v. 25, n. 4, p. 403–406, 2002.
- 13 LATANÈ, B. The psychology of social impact. *American Psychologist*, v. 36, n. 4, p. 343–356, 1981.
- 14 DEFFUANT, G. et al. Mixing beliefs among interacting agents. *Advances in Complex Systems*, v. 3, p. 87–98, 2000.
- 15 HEGSELMANN, R.; KRAUSE, U. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, v. 5, n. 3, 2002.

- 16 AXELROD, R. The dissemination of culture: a model with local convergence and global polarization. *Journal of Conflict Resolution*, v. 41, n. 2, p. 203–226, 1997.
- 17 BEN-NAIM, E.; KRAPIVSKY, P. L.; REDNER, S. Bifurcations and patterns diagrams in compromise processes. *Physica D*, v. 183, n. 3, p. 190–204, 2003.
- 18 BEN-NAIM, E.; KRAPIVSKY, P. L. Multiscaling in inelastic collisions. *Physical Review E*, v. 61, n. 1, p. 1–4, 1999.
- 19 BOCCALETTI, S. et al. Complex networks: structure and dynamics. *Physics Reports*, v. 424, n. 4, p. 175–308, 2006.
- 20 ALBERT, R.; BARABÁSI, A. L. Statistical mechanics of complex network. *Reviews of Modern Physics*, v. 74, n. 1, p. 47–97, 2002.
- 21 NEWMAN, M. E. J. *Networks: an introduction*. New York: Oxford University Press, 2010.
- 22 COSTA, L. F. et al. Characterization of complex networks: a survey of measurements. *Advances in Physics*, v. 56, n. 1, p. 167–242, 2007.
- 23 MILGRAM, S. The small world problem. *Psychology Today*, v. 1, n. 1, p. 60, 1967.
- 24 WATTS, D. J.; STROGATZ, S. Collective dynamics of small world networks. *Nature*, v. 393, n. 6684, p. 440, 1998.
- 25 BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *Science*, v. 286, n. 5439, p. 509, 1999.
- 26 BOGUNA, M. et al. Models of social networks based on social distance attachment. *Physical Review E*, v. 70, n. 5, p. 056122, 2004.
- 27 BONEH, A.; HOFRI, M. The coupon-collector problem revisited - a survey of engineering problems and computational methods. *Stochastic Models*, v. 13, n. 1, p. 39–66, 1997.
- 28 FRAJOLET, P.; GARDI, D.; THIMONIER, L. Birthday paradox, coupon collectors, caching algorithms and self-organizing search. *Discrete Applied Mathematics*, v. 39, n. 3, p. 207–229, 1992.
- 29 LORENZ, J. Consensus strikes back in the hegselmann-krause model of continuous opinion dynamics under bounded confidence. *Journal of Artificial Societies and Social Simulation*, v. 9, n. 1, p. 1, 2005.
- 30 LORENZ, J. A stabilization theorem for dynamics of continuous opinion. *Physica A*, v. 355, n. 1, p. 217–223, 2005.
- 31 CLAUSET, C. M. A.; NEWMAN, M. E. J. Hierarchical structure and the prediction of missing links in networks. *Nature*, v. 453, n. 7191, p. 98–101, 2008.
- 32 SUCHECKI, K.; EGUÍLUZ, V. M.; MIGUEL, M. S. Voter model dynamics in complex networks: role of dimensionality, disorder, and degree distribution. *Physical Review E*, v. 72, n. 3, p. 036132, 2005.

- 33 STAUFFER, D.; AHARONY, A. *Introduction to percolation theory*. London: Taylor & Francis, 1992.
- 34 GRIMMETT, G. *Percolation*. Berlin: Springer, 1999.
- 35 ALBERT, R.; JEONG, H.; BARABASI, A. L. Error and attack tolerance in complex networks. *Nature*, v. 409, n. 6819, p. 378, 2000.
- 36 NEWMAN, M. E. J.; STROGATZ, S. H.; WATTS, D. J. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, v. 64, n. 2, p. 026118, 2001.
- 37 NEWMAN, M. E. J. Spread of epidemic disease on networks. *Physical Review E*, v. 66, n. 1, p. 016128, 2002.
- 38 DOROGOVTSSEV, S. N.; GOLTSEV, A. V.; MENDES, J. F. F. Critical phenomena in complex networks. *Reviews of Modern Physics*, v. 80, n. 4, p. 1275–1335, 2008.
- 39 MAS, M.; FLACHE, A.; HELBING, D. Individualization as driving force of clustering phenomena in humans. *PLoS Computational Biology*, v. 6, n. 10, p. 1, 2010.
- 40 LOPEZ, C. Clustering transition in a system of particles self-consistently driven by a shear flow. *Physical Review E*, v. 70, n. 6, p. 066205, 2004.
- 41 CARLETTI, T. et al. How to make an efficient propaganda? *Europhysics Letters*, v. 74, n. 2, p. 222–228, 2006.
- 42 HEGSELMANN, R.; KRAUSE, U. Truth and cognitive division of labour: first steps towards a computer aided social epistemology. *Journal of Artificial Societies and Social Simulations*, v. 9, n. 3, p. 1, 2006.
- 43 BOTTI, V. et al. *Sharing temporal knowledge by multiple agents*. New York: Springer, 1993.
- 44 SAJJA, P. S. Multi-agent system for knowledge-based access to distributed databases. *Interdisciplinary Journal of Information, Knowledge, and Management*, v. 3, n. 1, p. 1–9, 2008.
- 45 VICENTE, A. C. R. M. R.; CATICHA, N. Opinion dynamics of learning agents: does seeking consensus lead to disagreement? *Journal of Statistical Mechanics*, v. 9, n. 3, p. 03015, 2009.
- 46 LANCICHINETTI, A.; FORTUNATO, S.; RADICCHI, F. Benchmark graphs for testing community detection algorithms. *Physical Review E*, v. 78, n. 4, p. 046110, 2008.

APÊNDICE A – TRABALHOS PUBLICADOS DURANTE O MESTRADO

1) *Knowledge acquisition by networks of interacting agents in the presence of observation errors*

João Luiz B. Batista and Luciano da Fontoura Costa

Physical Review E 82, 016103 (2010).

***Abstract:** In this work we investigate knowledge acquisition as performed by multiple agents interacting as they infer, under the presence of observation errors, respective models of a complex system. We focus the specific case in which, at each time step, each agent takes into account its current observation as well as the average of the models of its neighbors. The agents are connected by a network of interaction of Erdős-Rényi or Barabási-Albert type. First, we investigate situations in which one of the agents has a different probability of observation error (higher or lower). It is shown that the influence of this special agent over the quality of the models inferred by the rest of the network can be substantial, varying linearly with the respective degree of the agent with different estimation error. In case the degree of this agent is taken as a respective fitness parameter, the effect of the different estimation error is even more pronounced, becoming superlinear. To complement our analysis, we provide the analytical solution of the overall performance of the system. We also investigate the knowledge acquisition dynamic when the agents are grouped into communities. We verify that the inclusion of edges between agents (within a community) having higher probability of observation error promotes the loss of quality in the estimation of the agents in the other communities.*

2) *Communication Structure of Cortical Networks*

Luciano da Fontoura Costa, João Luiz B. Batista e Giorgio A. Ascoli

Frontiers in Computational Neuroscience 5:6 (2011).

Abstract: *Large-scale cortical networks exhibit characteristic topological properties that shape communication between brain regions and global cortical dynamics. Analysis of complex networks allows the description of connectedness, distance, clustering and centrality that reveal different aspects of how the network nodes communicate. Here, we focus on a novel analysis of complex walks in a series of mammalian cortical networks that model potential dynamics of information flow between individual brain regions. We introduce two new measures called absorption and driftness. Absorption is the average length of random walks between any two nodes, and takes into account all paths that may diffuse activity throughout the network. Driftness is the ratio between absorption and the corresponding shortest path length. For a given node of the network, we also define four related measurements, namely in- and out-absorption as well as in- and out-driftness, as the averages of the corresponding measures from all nodes to that node, and from that node to all nodes, respectively. We find that the cat thalamo-cortical system incorporates features of two classic network topologies, Erdős-Rényi graphs with respect to in-absorption and in-driftness, and configuration models with respect to out-absorption and out-driftness. Moreover, taken together these four measures separate the network nodes based on broad functional roles (visual, auditory, somatomotor, and frontolimbic).*

3) Effective number of accessed nodes in complex networks

Matheus P. Viana, João Luiz B. Batista and Luciano da Fontoura Costa

Physical Review E 85, 036105 (2012).

Abstract: *The measurement called accessibility has been proposed as a means to quantify the efficiency of the communication between nodes in complex networks. This article reports results regarding the properties of accessibility, including its relationship with the average minimal time to visit all nodes reachable after h steps along a random walk starting from a source, as well as the number of nodes that are visited after a finite period of time. We characterize the relationship between accessibility and the average number of walks required in order to visit all reachable nodes (the exploration time), con-*

ture that the maximum accessibility implies the minimal exploration time, and confirm the relationship between the accessibility values and the number of nodes visited after a basic time unit. The latter relationship is investigated with respect to three types of dynamics: traditional random walks, self-avoiding random walks, and preferential random walks.

4) Structure and dynamics: the transition from nonequilibrium to equilibrium in integrate-and-fire dynamics

Cesar H. Comin, João Luiz B. Batista, Matheus P. Viana, Luciano da F. Costa, Bruno A. N. Travençolo and Marcus Kaiser

International Journal of Bifurcation and Chaos 22, 7 (2012).

Abstract: *The transient and equilibrium properties of dynamics unfolding in complex systems can depend critically on specific topological features of the underlying interconnections. In this work, we investigate such a relationship with respect to the integrate-and-fire dynamics emanating from a source node and an extended network model that allows control of the small-world feature as well as the length of the long-range connections. A systematic approach to investigate the local and global correlations between structural and dynamical features of the networks was adopted that involved extensive simulations (one and a half million cases) so as to obtain two-dimensional correlation maps. Smooth, but diverse surfaces of correlation values were obtained in all cases. Regarding the global cases, it has been verified that the onset avalanche time (but not its intensity) can be accurately predicted from the structural features within specific regions of the map (i.e. networks with specific structural properties). The analysis at local level revealed that the dynamical features before the avalanches can also be accurately predicted from structural features. This is not possible for the dynamical features after the avalanches take place. This is so because the overall topology of the network predominates over the local topology around the source at the stationary state.*