

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
INSTITUTO DE CIÊNCIAS EXATAS  
PÓS-GRADUAÇÃO EM ESTATÍSTICA

Pedro Henrique Pereira Salgado

*Tópicos sobre o Modelo Bak-Sneppen*

Belo Horizonte

2019

Pedro Henrique Pereira Salgado

*Tópicos sobre o Modelo Bak-Sneppen*

Dissertação apresentada ao Programa de Pós Graduação em Estatística da UFMG, como requisito para a obtenção do grau de MESTRE em Estatística.

**Orientador: Roger William Câmara Silva**

Belo Horizonte

2019

Pereira Salgado, Pedro Henrique

Tópicos sobre o Modelo Bak-Sneppen / Pedro Henrique Pereira Salgado - 2019

125.p

1.Estatística. I.Título.

CDU 519.5

Pedro Henrique Pereira Salgado

*Tópicos sobre o Modelo Bak-Sneppen*

Dissertação apresentada ao Programa de Pós Graduação em Estatística da UFMG, como requisito para a obtenção do grau de MESTRE em Estatística.

**BANCA EXAMINADORA**

---

Roger William Câmara Silva

Doutor em Estatística

---

Bernardo Nunes Borges de Lima

Doutor em Matemática

---

Marcelo Richard Hilário

Doutor em Matemática

## Resumo

O modelo Bak-Sneppen é conhecido por ser um modelo que apresenta uma dinâmica simples e ainda exibe um comportamento crítico auto-organizado. Neste trabalho, vamos abordar, de forma detalhada, alguns tópicos sobre o modelo Bak-Sneppen que, de certa forma, contribuíram para o entendimento da distribuição estacionária do modelo.

Estudaremos os seguintes artigos: *Rigorous self-organised criticality in the modified Bak-Sneppen model* [13], dos autores Meester e Sarkar, *Bounds for avalanche critical values of the Bak-Sneppen model* [3], dos autores Gillett, Meester e Nuyens e *Maximal avalanches in the Bak-Sneppen model* [4], dos autores Gillett, Meester e van der Wal.

No artigo *Rigorous self-organised criticality in the modified Bak-Sneppen model*, os autores verificam a existência de uma lei de potências para o comportamento da duração e do alcance de avalanches em uma versão modificada do modelo Bak-Sneppen, através de um acoplamento com um processo de ramificação.

No artigo *Bounds for avalanche critical values of the Bak-Sneppen model*, os autores obtêm cotas para um dos pontos críticos de avalanches no modelo Bak-Sneppen (para grafos transitivos e localmente finitos) através de dois acoplamentos: para a cota inferior, usa-se o processo de ramificação e, para a cota superior, usa-se o modelo de percolação em sítios.

Por fim, no artigo *Maximal avalanches in the Bak-Sneppen model*, os autores estudam o comportamento da duração de avalanches (em grafos finitos), onde o limitante da avalanche é aleatório, e resultados surpreendentes são obtidos para o valor esperado da duração.

Palavras-chaves: Bak-Sneppen, avalanche, acoplamento, percolação, processo de ramificação.

# Abstract

The Bak-Sneppen model is known to be a simple model, exhibiting self-organized criticality. In this master thesis we'll study in detail some topics about the Bak-Sneppen model that, in a certain sense, contributed to the understanding of stationary distribution of the model.

We'll study the following articles: *Rigorous self-organised criticality in the modified Bak-Sneppen model* [13], by Meester and Sarkar, *Bounds for avalanche critical values of the Bak-Sneppen model* [3], by Gillett, Meester and Nuyens and *Maximal avalanches in the Bak-Sneppen model* [4], by Gillett, Meester and van der Wal.

In the article *Rigorous self-organised criticality in the modified Bak-Sneppen model*, the authors prove that a modified version of the Bak-Sneppen model obeys power law behaviour for avalanche duration and range using a coupling with a branching process.

In the article *Bounds for avalanche critical values of the Bak-Sneppen model*, the authors get bounds for one of avalanche critical values of the Bak-Sneppen model (for transitive and locally finite graphs) through two couplings: to the lower bound, the authors use a branching process and, to the upper bound, the authors use an independent site percolation model.

Finally, in the article *Maximal avalanches in the Bak-Sneppen model*, the authors study the avalanche duration behaviour (in finite graphs) with random threshold, and they get surprising results to the expected avalanche duration.

Keywords: Bak-Sneppen, avalanche, coupling, percolation, branching process.

*Aos meus pais, por tudo!*  
*À CAPES, pelo apoio financeiro.*

# Sumário

<b>Lista de Figuras</b>	<b>6</b>
<b>Introdução</b>	<b>9</b>
<b>1 Preliminares</b>	<b>11</b>
1.1 Teoria de grafos . . . . .	11
1.2 O modelo de percolação em sítios . . . . .	14
1.2.1 Introdução e alguns resultados . . . . .	14
1.2.2 Uma construção dinâmica do modelo de percolação em sítios . . . . .	17
1.3 O processo de ramificação . . . . .	18
1.3.1 Introdução e alguns resultados . . . . .	18
1.3.2 Uma construção alternativa do processo de ramificação . . . . .	21
<b>2 O modelo Bak-Sneppen</b>	<b>25</b>
2.1 Definição do modelo . . . . .	25
2.2 Avalanches . . . . .	26
2.2.1 Pontos críticos das avalanches . . . . .	32
2.3 O modelo Bak-Sneppen sem memória . . . . .	35
<b>3 Uma versão modificada do modelo Bak-Sneppen</b>	<b>41</b>
3.1 O modelo . . . . .	41
3.2 O acoplamento . . . . .	44
3.2.1 O processo MS . . . . .	44



3.2.2	Um exemplo . . . . .	47
3.2.3	A dominação estocástica entre os processos . . . . .	51
3.3	Prova do Teorema 3.1.1 . . . . .	51
<b>4</b>	<b>As cotas para os pontos críticos de avalanches no modelo Bak-Sneppen</b>	<b>58</b>
4.1	O teorema e algumas consequências . . . . .	58
4.2	Prova do Lema 4.1.1 . . . . .	60
4.2.1	O acoplamento . . . . .	60
4.2.2	Um exemplo . . . . .	64
4.3	A construção do acoplamento para a cota superior . . . . .	68
4.3.1	Intuição . . . . .	68
4.3.2	O processo acoplado . . . . .	69
4.3.3	Um exemplo . . . . .	73
4.3.4	A dominação estocástica entre os processos . . . . .	80
4.4	Prova do Lema 4.1.2 . . . . .	83
4.4.1	Um exemplo . . . . .	83
4.4.2	O acoplamento . . . . .	89
<b>5</b>	<b>Avalanches Maximais</b>	<b>93</b>
5.1	Introdução e resultados . . . . .	93
5.2	Prova do Teorema 5.1.1 . . . . .	95
5.3	Prova do Teorema 5.1.2 . . . . .	110
<b>A</b>	<b>Apêndice</b>	<b>121</b>
A.1	Códigos das simulações . . . . .	121

## Lista de Figuras

1	Ecossistema com $N$ espécies. . . . .	9
1.1	Passeio aleatório do processo de ramificação no instante inicial . . . . .	22
1.2	Passeio aleatório do processo de ramificação no instante $n = 1$ . . . . .	22
1.3	Passeio aleatório do processo de ramificação no instante $n = 2$ . . . . .	23
1.4	Passeio aleatório do processo de ramificação no instante $n = 3$ . . . . .	23
1.5	Passeio aleatório do processo de ramificação no instante $n = 7$ . . . . .	24
2.1	Exemplo de atualização no modelo Bak-Sneppen . . . . .	25
2.2	Configuração de $\Lambda_7$ no início de uma $b$ -avalanche . . . . .	28
2.3	Exemplo de uma $b$ -avalanche após a primeira atualização . . . . .	28
2.4	Exemplo de uma $b$ -avalanche após a segunda atualização . . . . .	29
2.5	Configuração de $\Lambda_7$ ao final de uma $b$ -avalanche . . . . .	29
2.6	Simulação do modelo Bak-Sneppen no círculo . . . . .	31
2.7	Atualização do modelo Bak-Sneppen sem memória no instante inicial . . . . .	37
2.8	Atualização do modelo Bak-Sneppen sem memória no instante $n = 1$ . . . . .	38
3.1	Exemplo de atualização no modelo Bak-Sneppen modificado . . . . .	42
3.2	Simulação do modelo Bak-Sneppen modificado . . . . .	44
3.3	Configuração inicial do processo MS . . . . .	47
3.4	Configuração do processo MS após uma atualização . . . . .	47
3.5	Configuração do processo MS após duas atualizações . . . . .	48
3.6	Configuração do processo MS após três atualizações . . . . .	48

3.7	Configuração do processo MS após quatro atualizações . . . . .	49
3.8	Configuração do processo MS após cinco atualizações . . . . .	50
3.9	Configuração do processo MS ao final da avalanche . . . . .	50
4.1	Configuração do processo GMS no instante inicial . . . . .	64
4.2	Configuração do processo GMS após a primeira atualização . . . . .	65
4.3	Configuração do processo GMS após a segunda atualização . . . . .	65
4.4	Configuração do processo GMS após a terceira atualização . . . . .	66
4.5	Configuração do processo GMS após a quarta atualização . . . . .	67
4.6	Configuração do processo GMS após o fim da $p$ -avalanche . . . . .	67
4.7	Configuração do processo acoplado no instante inicial . . . . .	73
4.8	Configuração do processo acoplado após a primeira atualização . . . . .	73
4.9	Configuração do processo acoplado após a segunda atualização . . . . .	74
4.10	Configuração do processo acoplado após a terceira atualização . . . . .	75
4.11	Configuração do processo acoplado após a quarta atualização . . . . .	76
4.12	Configuração do processo acoplado após a quinta atualização . . . . .	78
4.13	Configuração do processo acoplado após a sexta atualização . . . . .	79
4.14	Passo final do processo acoplado . . . . .	79
4.15	Configuração inicial de $\mathbb{Z}$ no processo acoplado . . . . .	82
4.16	Configuração do processo acoplado em $\mathbb{Z}$ ao final do instante $n = 1$ . . . . .	82
4.17	Configuração do processo acoplado em $\mathbb{Z}$ ao final do instante $n = 2$ . . . . .	82
4.18	Amostra de uniformes $(0, 1)$ obtida no modelo Bak-Sneppen . . . . .	90
5.1	Configuração de $\Lambda_7$ no início de uma $b$ -avalanche . . . . .	97
5.2	Exemplo de uma $b$ -avalanche após a primeira atualização . . . . .	98
5.3	Exemplo de uma $b$ -avalanche após 14 atualizações . . . . .	98

5.4	Exemplo de uma $b$ -avalanche após 15 atualizações . . . . .	99
5.5	Exemplo de uma $b$ -avalanche após 16 atualizações . . . . .	100
5.6	Grafo conexo e o circuito obtido a partir do grafo . . . . .	108
5.7	Grafo conexo e o circuito obtido a partir do grafo, após uma atualização. . .	109
5.8	Grafo conexo e o circuito obtido a partir do grafo, após duas atualizações. .	109
5.9	Grafos $K_5$ e $\Lambda_5$ antes do início da avalanche . . . . .	112
5.10	Grafos $K_5$ e $\Lambda_5$ no início da avalanche . . . . .	113
5.11	Grafos $K_5$ e $\Lambda_5$ após a primeira atualização . . . . .	113
5.12	Grafos $K_5$ e $\Lambda_5$ durante o instante de tempo $n = 1$ . . . . .	114
5.13	Grafos $K_5$ e $\Lambda_5$ após a segunda atualização . . . . .	115
5.14	Grafos $K_5$ e $\Lambda_5$ durante o instante de tempo $n = 2$ . . . . .	115
5.15	Grafos $K_5$ e $\Lambda_5$ após a terceira atualização . . . . .	116
5.16	Grafos $K_5$ e $\Lambda_5$ : a avalanche termina em ambos os grafos . . . . .	117
5.17	Grafos $K_5$ e $\Lambda_5$ : a avalanche termina apenas em $\Lambda_5$ . . . . .	118

## Introdução

O modelo Bak-Sneppen foi originalmente introduzido como um modelo simples de evolução por Per Bak e Kim Sneppen em 1993 (*ver [1]*). Considere  $N$  espécies dispostas em um ecossistema, como na Figura 1. A espécies estão representadas pelos pontos (em cinza) e os segmentos (em preto) representam a vizinhança entre as espécies. Inicialmente, sorteamos um valor do intervalo  $(0, 1)$  para cada espécie. Esse valor é denominado *aptidão* e representa o nível de adaptação dessa espécie em relação ao ecossistema. A dinâmica do modelo é dada pelo seguinte algoritmo: a cada instante de tempo  $n$ , executamos o procedimento abaixo, que gera a configuração do ecossistema no instante de tempo  $n + 1$ . Esse procedimento é denominado *atualização*.

1. Localizamos a espécie com valor mínimo dentre as  $N$  espécies;
2. Eliminamos esta espécie;
3. Eliminamos os dois vizinhos desta espécie;
4. Inserimos uma nova espécie no lugar de cada uma das espécies eliminadas;
5. Sorteamos um novo valor, do intervalo  $(0, 1)$ , para cada uma das novas espécies.

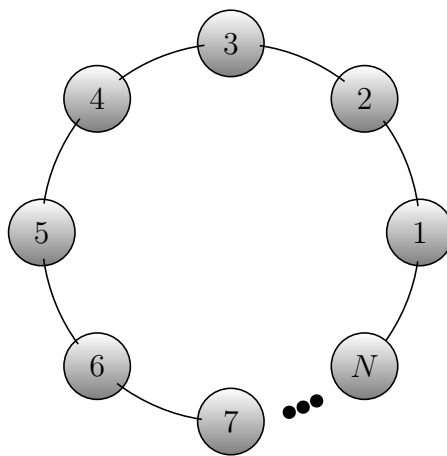


Figura 1: Ecossistema com  $N$  espécies.

---

O modelo pode ser interpretado como um ecossistema em que as espécies mudam devido às relações com outras espécies. A motivação de eliminar a espécie com a menor aptidão vem da lei de sobrevivência (da teoria de evolução) e a motivação em eliminar os vizinhos também vem da lei de sobrevivência e representa a interação entre as espécies no ambiente onde habitam.

Com uma introdução bastante superficial, já é possível abordar um dos principais problemas em aberto relacionados ao modelo. De acordo com [14], se executarmos o mesmo procedimento sem a regra 3, quando o número de atualizações e o número de vértices tendem ao infinito, as distribuições marginais das aptidões são degeneradas em 1, o que é bastante intuitivo. Com a regra 3, o problema ainda está em aberto. Felizmente, a simplicidade do modelo facilita a sua programação a fim de obter simulações numéricas. De acordo com [5], quando o número de atualizações e o número de vértices tendem ao infinito, simulações sugerem que as distribuições marginais das aptidões são independentes e uniformemente distribuídas em  $(b_c, 1)$ , onde  $b_c \approx \frac{2}{3}$ .

O objetivo do modelo não é abranger as mais diversas características de evolução de um ecossistema. Na verdade, o atrativo do modelo é o fato de ser simples, no sentido que cada espécie é associada a um único valor, e que ainda carrega algumas ideias de evolução e, principalmente, apresenta um comportamento crítico auto-organizado (*veja [12]*). De acordo com [2], muitas fontes afirmam que eventos de extinção parecem obedecer leis de potência e, por sua vez, a existência de leis de potência sugerem que a evolução poderia ser um exemplo de criticalidade auto-organizada (*veja mais em [16]*).

Vale ressaltar que o nosso objetivo é abordar, de forma detalhada, alguns tópicos sobre o modelo que contribuíram para o entendimento de sua distribuição estacionária. Para uma abordagem mais ampla, vamos precisar de alguns conceitos que serão abordados a seguir.

# 1 Preliminares

## 1.1 Teoria de grafos

Nesta seção, vamos abordar alguns conceitos básicos da teoria de grafos. Esses conceitos, que são muito importantes no estudo do processo de ramificação e, principalmente, no estudo do modelo de percolação, também serão importantes para o resultado principal do trabalho. Para uma abordagem mais profunda, veja [19].

Um *grafo*  $G$  é um par  $(V, E)$  formado por um conjunto  $V = V(G)$ , o conjunto de vértices (ou sítios) de  $G$  e um conjunto  $E = E(G)$  de elos (ou arestas), onde  $E \subset \{(x, y) : x, y \in V, x \neq y\}$ , isto é,  $E$  é um subconjunto de  $V \times V$ . Os grafos considerados são não direcionados ( $(x, y) = (y, x)$ ), não apresentam *loops* (já que  $x \neq y$ ), e não apresentam elos múltiplos ( $(x, y)$  é único). Para  $x, y \in V$ , dizemos que  $x$  é vizinho de  $y$  em  $G$  se  $(x, y) \in E$ , isto é, se existe um elo  $e$  em  $G$  da forma  $e = (x, y)$ . Dizemos que o elo  $e$  é *incidente* ao vértice  $x$  quando  $x$  é uma extremidade de  $e$ . Para cada  $x \in V$ , definimos  $\Delta_x$  como o *grau* de  $x$ , isto é, o número de elos incidentes a  $x$ , ou, equivalentemente, o número de vértices vizinhos a  $x$ .

Seja  $H = (V_H, E_H)$ . Dizemos que  $H$  é *subgrafo* de um grafo  $G$ , se  $V(H) \subset V(G)$  e  $E(H) \subset E(G)$ . A notação usada é  $H \subset G$ . Dizemos que a sequência alternante  $x_0, e_0, x_1, e_1, \dots, x_{n-1}, e_{n-1}, x_n$ ,  $n \geq 0$ , onde  $e_i = (x_i, y_i)$ ,  $i = 0, \dots, n-1$ , é um *caminho* se  $x_1, \dots, x_n$  são distintos e  $y_i = x_{i+1}$ ,  $i = 0, 1, \dots, n-1$ . O *comprimento de um caminho* é  $n$ , onde  $n$  é o número de elos  $x_i$  que compõem o caminho. Por definição, não há *loops*, já que  $y_n \neq x_1$ . Um *circuito* em  $G$  é uma sequência alternante  $x_0, e_0, x_1, e_1, \dots, x_{n-1}, e_{n-1}, x_n, e_n, x_0$  tal que  $x_0, e_0, \dots, x_{n-1}, e_{n-1}, x_n$  é um caminho e  $e_n = (x_n, x_0)$  (isto é,  $y_n = x_0$ ). O *comprimento de um circuito* é  $n+1$ , onde  $n$  é o número de elos  $x_i$  que compõem o circuito.

Para cada vértice  $x$ , defina  $\mathcal{C}_x$  o conjunto dos vértices de  $G$  que podem ser conectados a  $x$ , isto é, existe um caminho que liga esses vértices a  $x$  e assumamos que  $x \in \mathcal{C}_x$ . Se  $\mathcal{C}_x \cap \mathcal{C}_y \neq \emptyset$ , então  $\mathcal{C}_x = \mathcal{C}_y$ , isto é, diferentes conjuntos  $\mathcal{C}_x$  particionam o conjunto de vértices. Para cada um dos conjuntos  $\mathcal{C}_z$ ,  $z \in V$ , considere o subgrafo de  $G$ ,  $(\mathcal{C}_z, E(\mathcal{C}_z))$ , onde  $E(\mathcal{C}_z) = \{(x, y) \in E : x, y \in \mathcal{C}_z\}$ . Definimos os subgrafos da forma  $(\mathcal{C}_z, E(\mathcal{C}_z))$  como *componentes conexas* de  $G$ . Assim,  $x, y \in V$  estão na mesma componente conexa se, e somente se, existe um caminho em  $G$  que conecta  $x$  a  $y$  (ou  $y$  a  $x$ ).

Em geral, consideramos grafos conexos e localmente finitos. Um grafo  $G$  é *finito*, se os conjuntos  $V(G)$  e  $E(G)$  são finitos. Dizemos que um grafo  $G$  é *conexo* se existe um caminho entre qualquer par de vértices. Podemos também trabalhar com grafos de *grau limitado*, ou grafos *localmente finitos*, que é quando o grau de todos os vértices é uniformemente limitado, isto é,  $\exists M$  tal que  $\Delta_x \leq M$ ,  $\forall x \in V$ . Neste caso, o *grau máximo* de  $G$  é definido como  $\Delta^* = \sup_{x \in V} \Delta_x < \infty$ . Em particular, se  $\Delta_x = \Delta$ ,  $\forall x \in V$ , dizemos que o grafo é  $\Delta$ -regular.

Um outro conceito importante é a definição de isomorfismo em grafos. Sejam  $G_1 = (V_1, E_1)$  e  $G_2 = (V_2, E_2)$  grafos. Definimos um *isomorfismo* de  $G_1$  e  $G_2$  como uma bijeção  $\phi : V_1 \rightarrow V_2$ , com a propriedade de que,  $\forall x, y \in V_1$ ,  $\phi(x)$  e  $\phi(y)$  são vizinhos em  $G_2$  se, e somente se,  $x$  e  $y$  são vizinhos em  $G_1$ . Se existe um isomorfismo de  $G_1$  e  $G_2$ , dizemos que os grafos são *isomorfos*. Intuitivamente, podemos considerar que os grafos são “iguais”. Por exemplo, se  $G_1$  é isomorfo a um subgrafo de  $G_2$ , intuitivamente, podemos dizer que  $G_1 \subset G_2$ .

A *distância*  $\delta(x, y)$  entre dois vértices  $x$  e  $y$  de um grafo  $G$  conexo é definida como o menor comprimento dentre todos os caminhos em  $G$  que conectam  $x$  a  $y$ .

Agora, vamos apresentar quatro exemplos de grafos que serão abordados mais adiante.

**Exemplo 1.1.1** (rede hipercúbica  $d$ -dimensional). O primeiro exemplo é a rede hipercúbica  $d$ -dimensional ( $d \geq 1$ ), que é denotada por  $(\mathbb{Z}^d, \mathbb{E}^d)$ , onde  $\mathbb{Z}^d = \{(x_1, \dots, x_d) : x_i \in \mathbb{Z}, i = 1, \dots, d\}$  e  $\mathbb{E}^d = \{(x, y) : x, y \in \mathbb{Z}^d \text{ e } \|x - y\|_1 = 1\}$  é o conjunto de elos (vizinhos mais próximos). Com um certo abuso de notação, a partir de agora denotaremos a rede



hipercúbica por  $\mathbb{Z}^d$ . Observe que  $\mathbb{Z}^d$  é conexo e localmente finito. Em particular,  $\mathbb{Z}^1$  (ou simplesmente  $\mathbb{Z}$ ) é reta que conecta todos os números inteiros. Por  $\|\cdot\|$ , representamos a norma da soma, isto é,  $\|x\|_1 = \sum_{i=1}^d |x_i|$ , para  $x = (x_1, \dots, x_n) \in \mathbb{R}^d$ . Portanto, a distância entre dois vértices  $x, y \in \mathbb{Z}^d$  é dada por:

$$\delta(x, y) = \|x - y\|_1 = \sum_{i=1}^d |x_i - y_i|,$$

e  $x$  e  $y$  são vizinhos se, e somente se,  $\delta(x, y) = 1$ .

**Exemplo 1.1.2** (árvores). Um outro tipo de grafo são as árvores. Uma *árvore* é um grafo conexo que não possui circuitos. Então, dados dois vértices  $x$  e  $y$ , existe um único caminho que conecta  $x$  a  $y$ . Consequentemente, a distância entre  $x$  e  $y$  é o comprimento deste caminho. Denotaremos por  $T_\Delta$  a *árvore  $\Delta$ -regular* (ou  *$\Delta$ -homogênea*), que é a árvore em que todo vértice tem grau  $\Delta$ , onde  $\Delta \geq 2$ . Observe que  $T_2$  e  $\mathbb{Z}$  são grafos isomorfos. Consideramos também as árvores com raiz. Dizemos que uma árvore tem *raiz* (ou é *enraizada*), se existe um nó (vértice) em que a árvore cresce a partir deste nó. Cada nó tem um número de ramos (elos) e os ramos estão ligados aos *filhos* deste nó. Por  $T_\Delta^*$ , denotaremos a *árvore enraizada  $\Delta$ -regular*, que é a árvore em que a raiz tem grau  $\Delta - 1$  e os demais vértices têm grau  $\Delta$ . Observe que  $T_2^*$  e  $\mathbb{Z}^*$  são grafos isomorfos. Finalmente, por  $T_\Delta^+$ , denotaremos a *árvore enraizada  $\Delta$ -ária*, em que a raiz tem, no máximo, grau  $\Delta - 1$  e os demais vértices têm, no máximo, grau  $\Delta$ . Veja que  $T_\Delta^+ \subset T_\Delta^* \subset T_\Delta$ .

**Exemplo 1.1.3** (grafos circulares). Um *grafo circular* (ou *grafo cíclico*) é um grafo que possui um único ciclo, ou equivalentemente, um número de vértices conectados em um circuito. O grafo circular com  $N$  vértices será denotado por  $\Lambda_N$ . O número de vértices em  $\Lambda_N$  é igual ao número de arestas e todo vértice tem grau 2. Isto é, o grafo circular é um grafo 2-regular. A Figura 1 é um exemplo de grafo circular.

**Exemplo 1.1.4** (grafos completos). Um *grafo completo* é um grafo no qual cada par de vértices distintos é conectado por um único elo. Em outras palavras, um grafo completo é um grafo em que todo vértice é adjacente a todos os outros vértices. No caso do grafo completo, temos que  $E = \{(x, y) : x, y \in V, x \neq y\}$ . Portanto, se  $|V| = N$ , então  $|E| = \frac{N(N-1)}{2}$ . O grafo completo de  $N$  vértices será denotado por  $K_N$ .

Vamos definir o conceito de *transitividade* para grafos. Para isso, precisamos definir automorfismo. Um *automorfismo* de um grafo  $G = (V, E)$  é um isomorfismo de  $G$  em si próprio, isto é, uma bijeção  $\phi : V \rightarrow V$ , com a propriedade de que,  $\forall x \in V$  e  $\forall y \in V$ , temos  $(\phi(x), \phi(y)) \in E$ , se, e somente se,  $(x, y) \in E$ . Por fim, um grafo  $G = (V, E)$  é *transitivo* se,  $\forall x \in V$  e  $\forall y \in V$ , existe um automorfismo  $\phi : V \rightarrow V$  tal que  $\phi(x) = y$ . Os grafos dos exemplos acima são grafos transitivos.

## 1.2 O modelo de percolação em sítios

Nesta seção apresentamos o modelo de percolação em sítios e uma construção dinâmica deste modelo. O modelo será crucial para o estudo de um dos principais resultados da dissertação. Para uma abordagem mais profunda do modelo, veja [6].

### 1.2.1 Introdução e alguns resultados

Considere um grafo  $G = (V, E)$  conexo e localmente finito. A cada sítio de  $V$  será atribuído aleatoriamente o status aberto ou fechado da seguinte maneira: seja  $\chi := \{X(v), v \in V\}$  uma coleção de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com distribuição comum de *Bernoulli* com parâmetro  $p$ , isto é,

$$P_p(X(v) = 1) = p = 1 - P_p(X(v) = 0),$$

para todo  $v \in V$ , onde  $p$  é um número real entre 0 e 1 e  $P_p$  é a probabilidade associada a  $\chi$ . Mais formalmente, o espaço amostral do modelo será

$$\Omega = \{0, 1\}^V = \prod_{v \in V} \{0, 1\}.$$

Assim, uma *configuração* é um ponto  $\omega = (\omega_v : v \in V)$ . Seja  $\mathcal{F}$  a  $\sigma$ -álgebra gerada pelos cilindros em  $\Omega$ . Em outras palavras,  $\mathcal{F}$  é gerada pelos conjuntos que dependem apenas de vértices em subconjuntos finitos de  $V$ . A probabilidade  $P_p$  é a probabilidade produto em

$\Omega$ , atribuindo peso  $p$  a 1's e  $1 - p$  a 0's.  $X_\omega(v)$  é a projeção no vértice  $v$ , isto é, para cada configuração  $\omega \in \Omega$ ,

$$\begin{aligned} X_\omega : V &\longrightarrow \{0, 1\} \\ v &\longmapsto \omega_v, \end{aligned}$$

ou seja,  $X_\omega(v) = 1$  indica que o vértice  $v$  está aberto e  $X_\omega(v) = 0$  indica que o vértice  $v$  está fechado. Dizemos que um caminho está *aberto* se todos os seus vértices estão abertos, assim como um circuito (isto é, se  $X(x_i) = 1$ ,  $i = 0, \dots, n$ ). Definimos  $\{x \leftrightarrow y\}$ , como o evento “existe um caminho aberto que conecta  $x$  a  $y$ ”. Note que  $\{x \leftrightarrow x\}$  ocorre se, e somente se,  $x$  está aberto.

Um *aglomerado* é uma componente conexa do subgrafo obtido a partir de  $G$ , subgrafo no qual são excluídos todos os vértices fechados e todos os elos incidentes a estes vértices. Para todo  $x \in V$ ,  $C_x$  denota o *aglomerado que contém  $\{x\}$*  e  $C$  o *aglomerado que contém a origem*;  $C_x = \emptyset$ , se  $x$  é fechado. O conjunto de vértices  $C_x$  é o conjunto de todos os vértices do grafo que estão conectados a  $x$  por caminhos abertos. Podemos escrever  $C_x = \{y \in V : x \leftrightarrow y\}$ . Denotaremos por  $|C_x|$  o número de vértices em  $C_x$  e por  $\{x \leftrightarrow \infty\}$  o evento em que  $|C_x| = \infty$ . Estamos interessados em  $|C|$ , o volume (ou cardinalidade) do aglomerado da origem, mais precisamente em sua distribuição. Note que  $|C|$  é uma variável aleatória que pode assumir os valores  $1, 2, \dots, \infty$ , isto é,  $|C| : \Omega \longrightarrow \mathbb{N} \cup \{\infty\}$ . Queremos saber se aglomerados infinitos podem ocorrer (com probabilidade positiva).

**Definição 1.2.1.** Seja o evento  $\{\omega \in \Omega : |C| = \infty\} = \{0 \leftrightarrow \infty\}$ , isto é, o evento em que o tamanho do aglomerado é infinito a partir da origem. Denominamos este evento como *percolar*. Assim,

$$\begin{aligned} \theta : [0, 1] &\longrightarrow [0, 1] \\ p &\longmapsto P_p(0 \leftrightarrow \infty), \end{aligned}$$

onde  $\theta(p) := P_p(|C| = \infty) = P_p(0 \leftrightarrow \infty)$ . Em palavras,  $\theta(p)$  é a probabilidade de percolação a partir da origem em um modelo de percolação em sítios independentes, em que a probabilidade de um vértice estar aberto é igual a  $p$ .

*Observação 1.2.1.* Em geral, a probabilidade  $\theta(p)$  depende do vértice  $x$ . Inclusive, alguns autores definem a função como  $\theta(p, x)$ . No caso em que  $G$  é um grafo transitivo, o valor de  $\theta(p, x)$  não depende de  $x$ .

Claramente,  $\theta(0) = 0$  e  $\theta(1) = 1$ . Com efeito, se  $p = 0$ , não há vértices abertos e, conseqüentemente, não há percolação. Se  $p = 1$ , todos os vértices estão abertos, com probabilidade 1, e isto implica que há percolação. Para  $0 < p < 1$ , vamos definir um ponto crítico para o modelo de percolação em sítios.

**Lema 1.2.1.**  $\theta(p)$  é não-decrescente em  $p$ .

O Lema 1.2.1 é bastante intuitivo. Observe que a probabilidade de percolar não pode diminuir quando mais sítios são declarados abertos. A prova usa a técnica de acoplamento descrita na Seção 1.2.2 e pode ser verificada em [6].

**Definição 1.2.2.** Em um grafo localmente finito  $G$ , definimos o ponto crítico de percolação como

$$p_c^{\text{sítios}}(G) = \inf\{p : \theta(p) > 0\}. \quad (1.1)$$

Uma definição análoga, que é usada por alguns autores, é  $p_c^{\text{sítios}}(G) = \sup\{p : \theta(p) = 0\}$ .

Veja que o Lema 1.2.1 é crucial para que o ponto crítico de percolação esteja bem definido. Ao contrário de  $\theta(p)$ , o ponto crítico de percolação não depende da origem. Esta afirmação pode ser verificada em [11]. Agora, vamos enunciar um teorema para  $p_c^{\text{sítios}}(\mathbb{Z}^d)$ , onde  $\mathbb{Z}^d$  é a rede hipercúbica do Exemplo 1.1.1.

**Teorema 1.2.1.** Para  $G = \mathbb{Z}^d$  e  $d \geq 2$ , existe um valor crítico do parâmetro  $p$ , denominado  $p_c^{\text{sítios}}(\mathbb{Z}^d) \in (0, 1)$ , tal que:

- Se  $p < p_c^{\text{sítios}}(\mathbb{Z}^d)$ , então  $\theta(p) = 0$
- Se  $p > p_c^{\text{sítios}}(\mathbb{Z}^d)$ , então  $\theta(p) > 0$ .

A prova deste teorema pode ser vista em [6].

*Observação 1.2.2.* Note que, para  $d = 1$ , o teorema é falso. Com efeito, defina  $C_-$  e  $C_+$  como os vértices à esquerda e à direita da origem, respectivamente. Temos que  $|C_-|$  e

$|C_+|$  são variáveis aleatórias independentes e identicamente distribuídas, com  $P_p(|C_-| \geq k) = p^k$ . Portanto, com probabilidade 1, não há aglomerados infinitos em dimensão 1, se  $p < 1$ , pois  $\lim_{k \rightarrow \infty} P_p(|C_-| \geq k) = 0$ . Logo,  $p_c^{\text{sítios}}(\mathbb{Z}) = 1$ .

Para o caso em que  $G = T_\Delta$ , onde  $T_\Delta$  é uma árvore  $\Delta$ -regular (como no Exemplo 1.1.2), temos outro teorema:

**Teorema 1.2.2.** *Para  $G = T_\Delta$  e  $\Delta \geq 2$ ,*

$$p_c^{\text{sítios}}(T_\Delta) = \frac{1}{\Delta - 1}.$$

A demonstração pode ser vista em [9]. Vimos no Exemplo 1.1.2 que  $T_2$  e  $\mathbb{Z}$  são isomorfos. Logo,  $p_c^{\text{sítios}}(\mathbb{Z}) = p_c^{\text{sítios}}(T_2) = 1$ .

## 1.2.2 Uma construção dinâmica do modelo de percolação em sítios

Agora, vamos introduzir uma construção alternativa do modelo de percolação em sítios. A construção é feita de forma iterativa e a ideia é, começando da origem, avaliar cada um dos vizinhos e decidir se este vizinho está aberto ou não. Se está, adicionamos ao aglomerado. Caso contrário, declaramos esse vizinho como fechado. Podemos repetir esse procedimento, isto é, avaliar a cada passo os vizinhos do aglomerado atual, um a um.

Defina  $H \subset G$  um subconjunto localmente finito, conexo e que contém a origem e fixe uma ordenação para seus elos, isto é, defina  $e_{(1)}, e_{(2)}, \dots$  a ordem fixada dos elos do grafo induzido por  $H$ . Seja  $\mathcal{Y} = \{Y(v), v \in V_H\}$  uma coleção de variáveis aleatórias independentes e uniformemente distribuídas no intervalo  $(0, 1)$ . Para todo  $\rho \in [0, 1]$  e  $x \in V$ , dizemos que  $x$  é  $\rho$ -aberto, se  $Y(x) \leq \rho$  e  $\rho$ -fechado, se  $Y(x) > \rho$ . Assim,  $P(\{x \text{ ser declarado } \rho\text{-aberto}\}) = \rho$  e  $P(\{x \text{ ser declarado } \rho\text{-fechado}\}) = 1 - \rho$ .

Definimos, de forma indutiva, uma sequência  $(S_k)_{k \geq 0}$  de pares ordenados de subconjuntos de  $G$ , onde, para  $\omega \in \Omega$ ,

$$S_0 = (\emptyset, \emptyset) \quad S_1 = \begin{cases} (\{x_1\}, \emptyset), & \text{se } Y_\omega(x_1) = 1 \\ (\emptyset, \{x_1\}), & \text{se } Y_\omega(x_1) = 0 \end{cases}$$

com  $x_1 = 0$  sendo a origem.

Uma vez definidos  $S_1, S_2, \dots, S_n = (A_n, B_n)$ , definimos  $S_{n+1}$  da seguinte maneira:

- 1) Se possível, defina  $e_{(1)} = e_n$ , isto é,  $e_n = (x_n, x_{n+1})$  é um elo em que  $x_n \in A_n$  e  $x_{n+1} \notin A_n \cup B_n$ . Então

$$S_{n+1} = \begin{cases} (A_n \cup \{x_{n+1}\}, B_n), & \text{se } Y_\omega(x_{n+1}) = 1 \\ (A_n, B_n \cup \{x_{n+1}\}) & \text{se } Y_\omega(x_{n+1}) = 0 \end{cases}$$

- 2) Se não for possível (isto é, se  $\nexists x_{n+1}$ ), definimos  $S_{n+1} = S_n$ .

No limite, quando  $n \rightarrow \infty$ , definimos  $S = (A, B)$ , onde  $A = \bigcup_{n=1}^{\infty} A_n$  e  $B = \bigcup_{n=1}^{\infty} B_n$ .

Depois de analisarmos todos os vértices de  $V_H$ , vemos que, se a origem está aberta, eventualmente construiremos um aglomerado aberto  $A$  a partir da origem e que está contido em  $H$ . Como consequência, o conjunto  $B$  é a fronteira de  $A$  em  $H$ .

**Lema 1.2.2.** *Suponha que existe uma constante  $\gamma$  tal que  $\gamma > p_c^{\text{sítios}}(H)$  e*

$$P(Y(x_{n+1}) = 1 \mid S_0, \dots, S_n) \geq \gamma, \forall n. \quad (1.2)$$

Então  $P(|A| = \infty) > 0$ .

Este lema foi proposto e demonstrado em [7].

## 1.3 O processo de ramificação

Nesta seção, introduzimos o processo de ramificação, que será útil posteriormente. Uma abordagem mais completa é feita em [10].

### 1.3.1 Introdução e alguns resultados

O processo de ramificação é uma forma simples de descrever a evolução de uma população durante um intervalo de tempo. Basicamente, o processo é composto por indivíduos e estes indivíduos têm uma quantidade de filhos em um determinado instante

de tempo, que é denominado *geração*. O número de gerações pode ser finito, que é o caso em que a população é extinta; ou infinito, que é o caso em que a população sobrevive “eternamente”. Uma maneira conveniente de descrever o processo de ramificação é usando grafos. Neste caso, cada indivíduo é representado por um vértice e as arestas representam o parentesco entre os indivíduos. Mais especificamente, podemos usar árvores enraizadas (veja o Exemplo 1.1.2). Por exemplo, se dois vértices são vizinhos, isto é, se dois indivíduos estão conectados por uma aresta, dizemos que o vértice mais próximo da raiz é o pai (isto é, o vértice cuja distância para a raiz é a menor) e o outro vértice é o filho. Se definirmos como  $X_{n,i}$  o  $i$ -ésimo filho na  $n$ -ésima geração, podemos determinar a quantidade de filhos em uma determinada geração através da seguinte relação de recursão:

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_{n,i}, \text{ para } n \geq 1, \quad (1.3)$$

onde  $Z_n$  representa a quantidade de filhos na  $n$ -ésima geração e  $Z_0 = 1$ , o que é bastante razoável, pois  $Z_0 = X_{0,1} := X_0$  (ou seja,  $X_0 = 1$ ) e  $X_0$  é exatamente a raiz da árvore.

Vamos associar a cada vértice (indivíduo) uma variável aleatória, isto é, vamos associar uma distribuição, de forma independente, a cada um dos indivíduos no processo de ramificação. É comum atribuímos a mesma distribuição a todos os indivíduos. Nestes casos, podemos tratar os indivíduos como uma matriz de variáveis aleatórias  $(X_{n,k})_{n,k \geq 1}$  i.i.d., onde cada linha  $n$  da matriz corresponde a uma geração, e a distribuição associada a cada  $X_{n,k}$  é discreta, já que estamos analisando a quantidade de filhos de um indivíduo. Dessa forma, a distribuição do número de filhos de um indivíduo pode ser definida como um vetor de probabilidades,  $(p_i)$ , onde a  $i$ -ésima posição do vetor representa a probabilidade de um indivíduo ter  $i$  filhos, ou seja,

$$p_i = P(X = i),$$

onde  $X$  é uma variável aleatória com a mesma distribuição de  $X_{n,k}$ ,  $\forall n, k$ .

Veja que a população é extinta se, e somente se, todos os indivíduos presentes em uma determinada geração não têm filhos. De maneira formal, a população é extinta se, e somente se, ocorre o evento  $\{\exists n \text{ tal que } Z_n = 0\}$ . Com efeito, se o número de filhos em uma geração é igual a 0, claramente não haverá indivíduos nas próximas gerações, fato

que pode ser observado através da Equação (1.3). Como já era de se esperar, estamos interessados na distribuição de  $Z_n$  e se a população pode sobreviver (com probabilidade positiva).

Agora, dado um processo de ramificação, será que é possível determinar se a população será (ou não) extinta? Em outras palavras, é possível determinar sob quais condições esta população sobrevive ou é extinta? Para responder essas perguntas, vamos, finalmente, definir  $\eta$  e enunciar um teorema.

**Definição 1.3.1.** Seja  $\{X_{n,i}\}_{n,i \geq 1}$  uma matriz de variáveis aleatórias i.i.d. de um processo de ramificação e seja  $Z_n$  como na Equação (1.3). Definimos a *probabilidade de extinção* como

$$\eta = P(\exists n \text{ tal que } Z_n = 0). \quad (1.4)$$

**Teorema 1.3.1.** *Seja  $\{X_{n,i}\}_{n,i \geq 1}$  uma matriz de variáveis aleatórias i.i.d de um processo de ramificação e  $X$  uma variável aleatória com a mesma distribuição de  $X_{n,i}$ ,  $\forall n, i$ . Temos que,*

- Se  $E(X) < 1$ , então  $\eta = 1$ ;
- Se  $E(X) > 1$ , então  $\eta < 1$ ;
- Se  $E(X) = 1$  e  $P(X = 1) < 1$ , então  $\eta = 1$ .

Além disso, a probabilidade de extinção,  $\eta$ , é a menor solução no intervalo  $[0, 1]$  de

$$\eta = G_X(\eta), \quad (1.5)$$

onde  $G_X = E(s^X)$ .

O teorema acima fornece uma espécie de ponto crítico para o processo de ramificação, que é o valor esperado do número de filhos de um indivíduo, denotado por  $E(X)$ . A demonstração do teorema pode ser vista em [10].

Seja  $T$  o total de indivíduos em um processo de ramificação. Podemos escrever  $T$  da seguinte forma:

$$T = \sum_{n=0}^{\infty} Z_n.$$



Observe que  $T : \Omega \rightarrow \mathbb{N} \cup \{\infty\}$ .

**Teorema 1.3.2.** *Seja  $\{X_{n,i}\}_{n,i \geq 1}$  uma matriz de variáveis aleatórias i.i.d de um processo de ramificação,  $X$  uma variável aleatória com a mesma distribuição de  $X_{n,i}$ ,  $\forall n, i$ , e  $T$  o total de indivíduos do processo. Para todo  $s \in [0, 1)$ , temos que*

$$G_T(s) = sG_X(G_T(s)).$$

onde  $G_T(s) = E(s^T)$ .

A prova é feita em [10]. O teorema acima oferece uma maneira alternativa de obter  $G_T(s)$ , a partir da distribuição do número de filhos de um indivíduo. Usaremos este teorema na prova do Teorema 3.1.1.

### 1.3.2 Uma construção alternativa do processo de ramificação

Na seção anterior, estudamos o processo de ramificação através das gerações. Isto é, cada instante de tempo era denominado geração, e denotamos por  $Z_n$  o número de filhos na  $n$ -ésima geração, ou equivalentemente, o número de filhos no instante de tempo  $n$ . Agora vamos olhar para cada indivíduo da população, isto é, a cada instante de tempo, vamos olhar para o número de filhos de um único indivíduo. Essa nova construção é denominada *passeio aleatório do processo de ramificação*.

Seja  $\{X'_i\}_{i \geq 1}$  uma sequência de variáveis aleatórias i.i.d. e com a mesma distribuição de  $X$  e defina

$$S'_i = S'_{i-1} + X'_i - 1, \quad i = 1, 2, \dots, \quad (1.6)$$

com  $S'_0 = 1$  e  $T' = \min\{i : S'_i = 0\}$ . Se  $\nexists i$  tal que  $S'_i = 0$ , então  $T' = \infty$ .

O passeio aleatório do processo de ramificação é apenas uma forma diferente de explorar a árvore subjacente do processo. Nessa construção, seria mais difícil de extrair a distribuição de uma geração, por exemplo. Aqui, dizemos que  $S'_i$  é o número de indivíduos ativos após a exploração de  $i$  indivíduos. Já  $T'$  é o total de indivíduos no passeio aleatório do processo de ramificação.

Para explorar a árvore, podemos executar uma busca em largura ou uma busca em profundidade. O tipo de busca não afeta o valor de  $T'$ . No capítulo 3, veremos um

outro tipo de busca, que é feita através de um acoplamento com uma versão modificada do modelo Bak-Sneppen. Aqui, vamos executar a busca em largura, da esquerda para a direita.

A seguir, veremos um exemplo da construção e, então, veremos que a distribuição de  $T'$  coincide com a distribuição de  $T$ , que é o total de indivíduos em um processo de ramificação.

Seja  $A$  o conjunto de vértices ativos, que é o conjunto que contém os vértices não-explorados. Considere a Figura 1.1. Vamos explorar os vértices da árvore, até o instante de tempo em que  $S'_i = 0$ , para algum  $i$ . O processo é descrito da seguinte forma: a cada instante de tempo  $i$ , exploramos o  $i$ -ésimo vértice e adicionamos os filhos deste vértice ao conjunto  $A$ . Então, removemos o  $i$ -ésimo vértice do conjunto  $A$ . A relação descrita aqui fornece a Equação (1.6). No instante inicial, temos que  $A = \{1\}$ . O próximo vértice a ser explorado será o vértice 1.

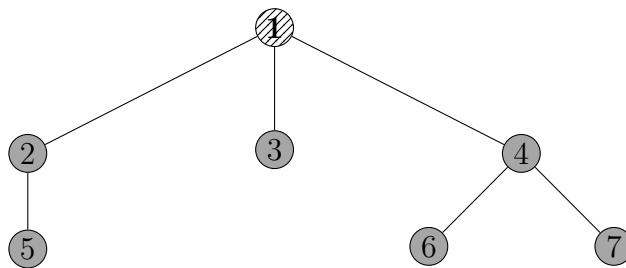


Figura 1.1: Execução do passeio aleatório do processo de ramificação no instante inicial. Os vértices em cinza representam os vértices inativos. O vértice 1 (em linhas diagonais) é o vértice a ser explorado.

Observe que o vértice 1 tem 3 filhos. Portanto,  $X'_1 = 3$  e  $S'_1 = 1 + 3 - 1 = 3$ . O próximo vértice a ser investigado será o vértice 2 (Figura 1.2).

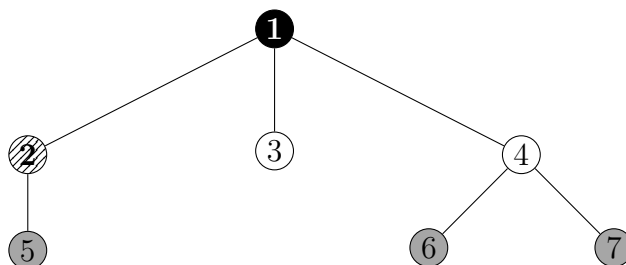


Figura 1.2: Execução do passeio aleatório do processo de ramificação no instante de tempo  $n = 1$ . O vértice 1 (em preto) já foi explorado. Assim,  $A = \{2, 3, 4\}$ .

Observe que o vértice 2 tem apenas um filho. Portanto,  $X'_2 = 1$  e  $S'_2 = 3 + 1 - 1 = 3$ . Como a busca é em largura, investigaremos agora o vértice 3 (Figura 1.3).

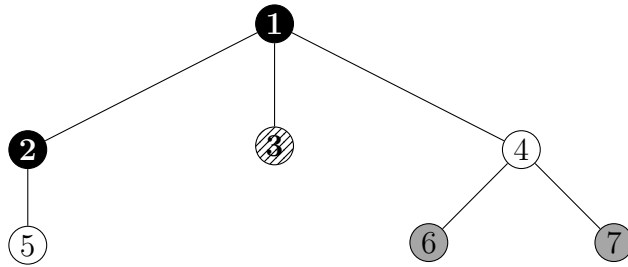


Figura 1.3: Execução do passeio aleatório do processo de ramificação no instante de tempo  $n = 2$ . Temos que  $A = \{3, 4, 5\}$ .

O vértice 3 não tem filhos. Portanto,  $X'_3 = 0$  e  $S'_3 = 3 + 0 - 1 = 2$ . Observe que temos menos vértices ativos. Vamos explorar o vértice 4 (Figura 1.4).

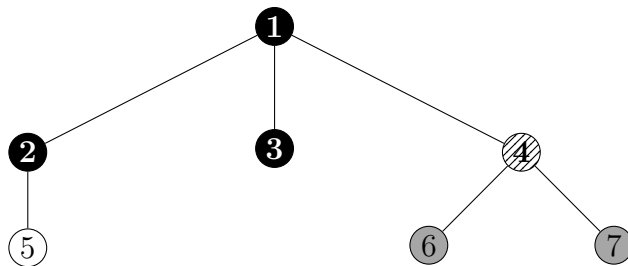


Figura 1.4: Execução do passeio aleatório do processo de ramificação no instante de tempo  $n = 3$ . Temos que  $A = \{4, 5\}$ , já que o vértice 3 não tem filhos.

Veja que o vértice 4 tem 2 filhos. Assim,  $X'_4 = 2$  e  $S'_4 = 2 + 2 - 1 = 3$ . Os próximos vértices a serem explorados são os vértices 5, 6 e 7, respectivamente. Mas veja que eles não têm filhos, ou seja,  $X'_5 = X'_6 = X'_7 = 0$ , e, portanto,  $S'_5 = 2$ ,  $S'_6 = 1$  e  $S'_7 = 0$ . Portanto, o processo é interrompido. A configuração final pode ser observada na Figura 1.5.

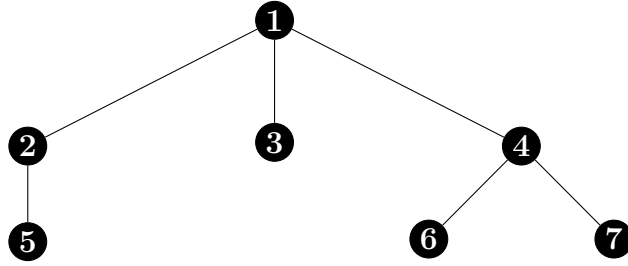


Figura 1.5: Execução do passeio aleatório do processo de ramificação no instante de tempo  $n = 7$ . Note que  $A = \emptyset$ .

Logo,  $T' = 7$ , ou seja, a árvore contém 7 indivíduos. O lema a seguir mostra que  $T'$  e  $T$ , de fato, seguem a mesma distribuição.

**Lema 1.3.1.** *Seja  $S_i$  e  $S'_i$  o número de indivíduos ativos no instante de tempo  $i$  no processo de ramificação e no passeio aleatório do processo de ramificação, respectivamente. O processo estocástico  $\{S'_i\}_{i \in [T']}$  segue a mesma distribuição do processo estocástico  $\{S_i\}_{i \in [T]}$ , onde  $T$  e  $T'$  são o total de indivíduos no processo de ramificação e no passeio aleatório do processo de ramificação, respectivamente.*

A demonstração é feita em [10]. Aqui, vemos que as distribuições dos processos  $\{S_i\}_{i \in [T]}$  e  $\{S'_i\}_{i \in [T']}$ , de fato, coincidem. Este lema será importante para a prova do Teorema 3.1.1 e também pode ser usado para provar o seguinte teorema:

**Teorema 1.3.3.** *Considere um processo de ramificação i.i.d., onde  $X$  uma variável aleatória com a mesma distribuição do número de filhos de um indivíduo e  $T$  o total de indivíduos do processo. Temos que*

$$P(T = n) = \frac{1}{n} P(X_1 + \dots + X_n = n - 1),$$

onde  $\{X_i\}_{i=1}^n$  são cópias de  $X$ .

Este resultado também será importante para a prova do Teorema 3.1.1 e não será provado aqui, mas a prova pode ser encontrada em [10].

## 2 O modelo Bak-Sneppen

Finalmente, veremos o modelo Bak-Sneppen, que é a principal motivação da dissertação.

### 2.1 Definição do modelo

Antes da definição matemática do modelo, vamos adaptar as terminologias propostas na Introdução. Considere um grafo circular  $\Lambda_N$  (veja o Exemplo 1.1.3). A cada vértice atribuiremos um valor, de forma independente, e com distribuição uniforme no intervalo  $(0, 1)$ . Esse valor é o que chamamos de *aptidão* do vértice. Em cada instante de tempo  $n$ , localizamos o vértice com menor valor. Este vértice será chamado de *vértice minimal*. Então, apagamos o valor deste vértice e dos vizinhos, e atribuímos novos valores para cada um, de forma independente, e com distribuição  $U(0, 1)$ . Esse procedimento é o que chamamos de *atualização*.

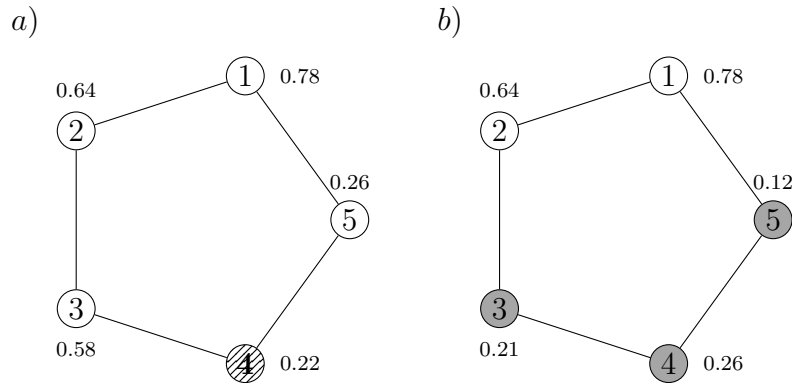


Figura 2.1: Configuração inicial de  $\Lambda_5$  (a) e configuração de  $\Lambda_5$  após uma atualização (b). Em (a), o vértice minimal foi localizado (vértice 4). Em (b), executamos o procedimento de atualização. Em cinza, os vértices que foram atualizados. O próximo vértice minimal seria o vértice 5.

Mais adiante, na Seção 2.3, o termo *aptidão* será usado para definir as variáveis aleatórias associadas aos vértices. Portanto, vamos evitar usar este termo ao longo do

texto e vamos usar a expressão “valor(es) associado(s) ao(s) vértice(s)”, ou simplesmente “valor(es) do(s) vértice(s)”.

A simplicidade do modelo nos permite estendê-lo para diversos tipos de grafos, e não apenas grafos circulares. Para grafos finitos, podemos definir o modelo da seguinte forma: seja  $G_N = (V, E)$  um grafo conexo e com  $N$  vértices. O modelo Bak-Sneppen pode ser definido como uma cadeia de Markov de tempo discreto e com espaço de estados  $\Lambda = [0, 1]^V$ . Para todo  $n \in \mathbb{N}$ , seja

$$\begin{aligned}\Phi_n : V &\longrightarrow [0, 1] \\ v &\longmapsto x_v,\end{aligned}$$

o valor associado ao vértice  $v$  no instante de tempo  $n$ , e considere a sequência de variáveis aleatórias  $\{\Phi_n\}_{n \geq 0}$ , que denotam os estados do modelo Bak-Sneppen, onde cada estado é uma coleção de valores da forma  $\{\Phi_n(v), v \in V\}$ . Seja  $\pi_0$  a distribuição inicial, isto é,  $\Phi_0 \sim \pi_0$ . Defina  $\mathcal{N}(v)$  o conjunto dos vizinhos de um vértice  $v$  e seja  $w_n = \operatorname{argmin}_{v \in V(G)} \Phi_n(v)$  o vértice minimal no instante de tempo  $n$ . A função de transição é dada por

$$\Phi_{n+1}(v) = \begin{cases} U_n^v, & \text{se } v \in \mathcal{N}(w_n) \cup \{w_n\} \\ \Phi_n(v), & \text{caso contrário,} \end{cases}$$

onde  $U_n^v$  é uma variável aleatória uniforme no intervalo  $(0, 1)$ . As variáveis aleatórias que geram os valores para os vértices são independentes entre si.

## 2.2 Avalanches

Como já foi dito, o modelo Bak-Sneppen não tem o que chamamos de parâmetro externo de ajuste. E uma forma bastante conveniente de definir tal parâmetro é analisar o modelo através do que chamamos de avalanches, ou melhor,  $b$ -avalanches.

Uma  $b$ -*avalanche* ocorre no intervalo de tempo  $[s, s + t)$  se

$$w_s \geq b \text{ e } \min\{n \geq s : w_n > b\} = s + t.$$

Em palavras, uma  $b$ -avalanche ocorre entre os instantes de tempo  $n = s$  e  $n = s + t$  se no instante de tempo  $s$ , todos os valores atribuídos aos vértices são maiores ou iguais a  $b$  e no instante de tempo  $s + t$ , todos os valores estão acima de  $b$  pela primeira vez após o instante de tempo  $s$ . O vértice minimal no instante de tempo  $s$  é chamado de *origem* da avalanche. Seja  $m$  o valor do vértice minimal no instante de tempo  $s$ . Note que a definição da ocorrência de uma  $b$ -avalanche no intervalo de tempo  $[s, s + t)$  vale para qualquer  $b \in [0, m]$ . A situação em que escolhemos  $b = m$  é chamada de *escolha máxima*.

Veremos mais adiante que é natural definir o modelo Bak-Sneppen como uma sequência de avalanches. Como diferentes limitantes podem ser escolhidos, podemos realizar uma *decomposição em avalanches*. E chamamos de *decomposição maximal de avalanches* quando, ao final de cada avalanche, sempre fazemos a escolha máxima de  $b$  para a próxima avalanche. No Capítulo 5, voltaremos a falar sobre tal decomposição.

Agora, vamos definir algumas quantidades que envolvem avalanches e vamos dar um exemplo que ilustra o modelo Bak-Sneppen como uma sequência de avalanches. Dizemos que o *conjunto-alcance* de uma  $b$ -avalanche é o conjunto de vértices que passaram por pelo menos uma atualização durante a  $b$ -avalanche. Por sua vez, o *alcance* de uma avalanche é a cardinalidade do conjunto-alcance, isto é, o número de vértices do conjunto-alcance. A *duração* de uma  $b$ -avalanche é dada pela diferença entre os instantes de início e fim da  $b$ -avalanche. Em outras palavras, a duração é o número de atualizações feitas durante a  $b$ -avalanche.

### Um exemplo

Vamos ilustrar o procedimento de uma  $b$ -avalanche em um grafo finito. Considere o grafo circular  $\Lambda_7$ , como nas figuras abaixo, e considere que iniciamos uma  $b$ -avalanche no instante de tempo  $n = s$ . Faremos a escolha máxima para  $b$ , ou seja,  $b = 0.4$ .

As figuras à esquerda mostram o grafo  $\Lambda_7$ , e as figuras à direita mostram uma representação gráfica do processo em  $\Lambda_7$ . No eixo  $x$ , vemos os índices dos vértices. Já os pontos representam valores associados aos vértices.

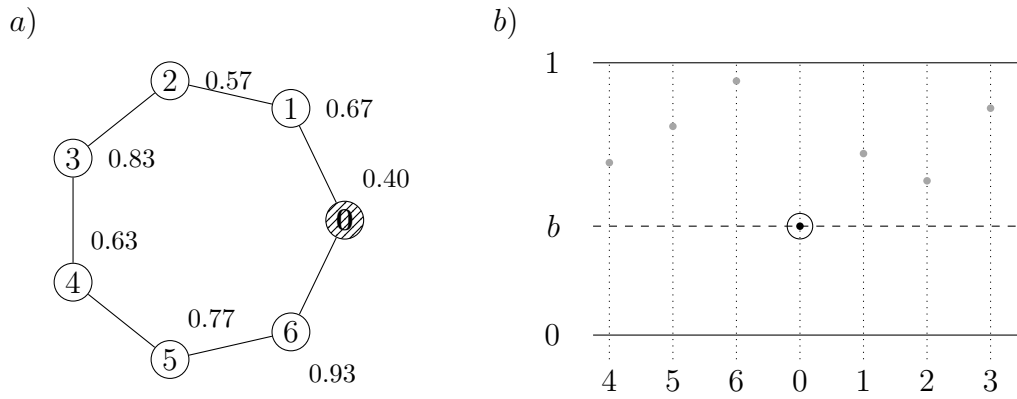


Figura 2.2: Configuração inicial de uma  $b$ -avalanche, onde a origem é o vértice minimal e  $b = 0.4$ .  $n = s$ ; conjunto-alcance =  $\{0\}$ .

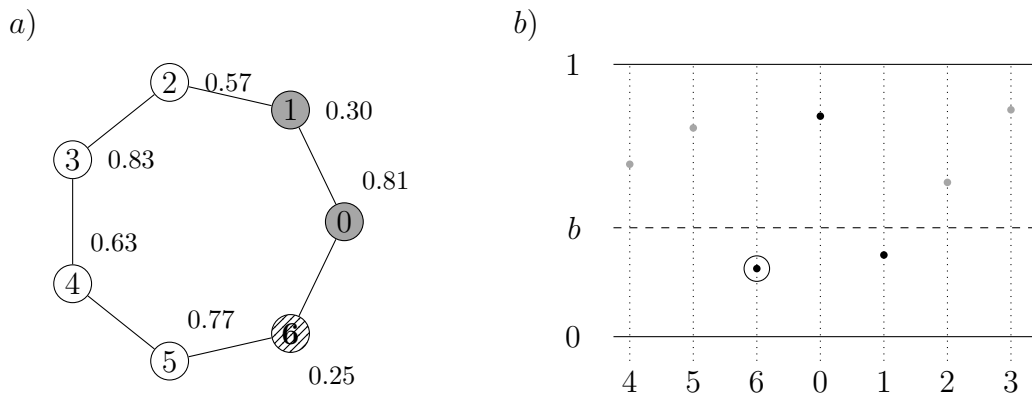


Figura 2.3: Configuração de  $\Lambda_7$  após a primeira atualização. Tanto a origem, quanto os seus dois vizinhos foram atualizados. Agora, o vértice minimal é o vértice 6, com valor igual a 0.25.  $n = s + 1$ ; conjunto-alcance =  $\{0, 1, 6\}$ .



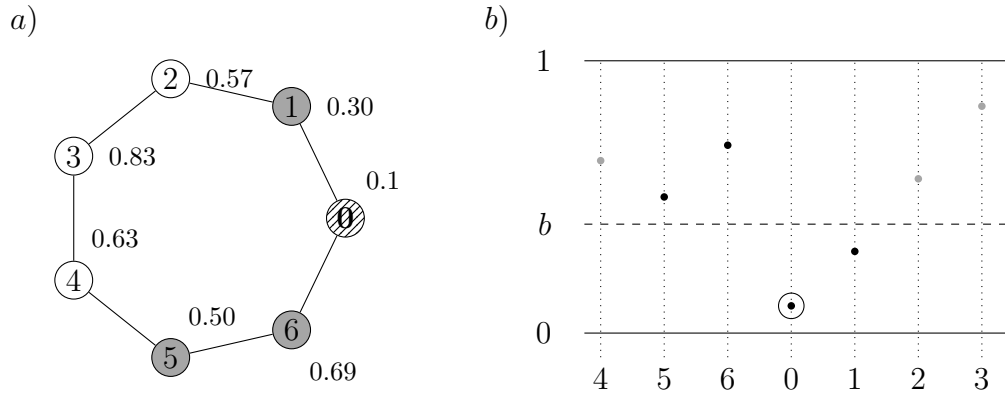


Figura 2.4: Configuração de  $\Lambda_7$  após a segunda atualização. O vértice 5 foi atualizado, por ser vizinho do vértice minimal anterior. Repare que agora o vértice minimal é a origem da avalanche, com valor igual a 0.1.  $n = s + 2$ ; conjunto-alcance =  $\{0, 1, 5, 6\}$ .

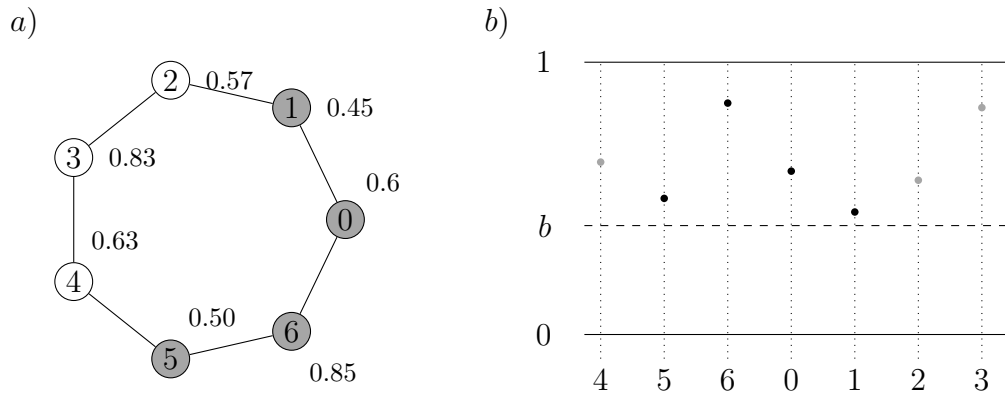


Figura 2.5: Configuração de  $\Lambda_7$  ao final da  $b$ -avalanche. Repare que os valores de todos os vértices estão acima de  $b = 0.4$ .  $n = s + 3$ ; conjunto-alcance =  $\{0, 1, 5, 6\}$ ; duração = 3.

Uma  $b$ -avalanche pode ser considerada como um processo estocástico por si só. A principal característica da origem é que ela é o vértice minimal e, conseqüentemente, será atualizada de imediato. Assim, podemos atribuir qualquer valor para a origem, desde que seja o valor mínimo. Vértices com valores abaixo do limitante são chamados *ativos*, e os demais vértices são chamados *inativos*. Note que pouco importa o valor exato de um vértice inativo, uma vez que estes vértices estão acima do limitante.

O processo segue as regras de atualização do modelo Bak-Sneppen. Seja  $A_n$  o conjunto de vértices ativos em uma  $b$ -avalanche no modelo Bak-Sneppen no instante de tempo  $n$ . Os vértices com valor menor ou igual a  $b$  no instante de tempo  $n$  são incluídos

em  $A_n$ . Qualquer vértice com valor maior que  $b$  não é incluído em  $A_n$ . O processo termina no primeiro instante de tempo em que nenhum vértice está ativo, isto é, quando  $A_n = \emptyset$ , para algum  $n$ .

### Avalanches em grafos localmente infinitos

A abordagem do modelo Bak-Sneppen a partir de uma sequência de avalanches oferece algumas vantagens. Ao contrário do modelo Bak-Sneppen, cuja dinâmica é bem definida apenas para grafos finitos, é possível estudar o comportamento das avalanches em certos grafos infinitos. Para isso, podemos definir a distribuição inicial como  $\pi_0(0) = b$  e  $\pi_0(v) = 1$ , se  $v \neq 0$ , onde  $0$  é a origem da avalanche e  $b \in [0, 1)$ . Assim, a origem sempre será o vértice minimal inicial. Mas ainda seria inviável mensurar  $\mathcal{N}(v)$ , já que não temos um número finito de vizinhos. Por isso, pedimos que o grafo seja localmente finito.

Como já foi dito, note que, ao contrário do próprio modelo Bak-Sneppen, a abordagem do modelo via avalanches nos oferece um parâmetro de ajuste, que chamamos de limitante  $b$ . O resultado a seguir mostra um exemplo de dominação estocástica entre duas avalanches de limitantes distintos.

**Proposição 2.2.1.** *Seja  $G$  um grafo conexo e localmente finito e considere uma  $b'$ -avalanche e uma  $b''$ -avalanche em  $G$ , iniciadas no mesmo instante de tempo. Assuma que  $b'' < b'$ . Então existe um acoplamento entre a  $b'$ -avalanche e a  $b''$ -avalanche no mesmo grafo e com a mesma origem, de forma que se a  $b'$ -avalanche é encerrada, então a  $b''$ -avalanche também é encerrada.*

*Demonstração.* A prova segue por construção. Considere  $G'$  e  $G''$  duas cópias idênticas de  $G$  e seja  $\{U_i\}_{i \geq 1}$  uma sequência de variáveis aleatórias i.i.d. com distribuição comum  $U(0, 1)$ . A cada instante de tempo, atribua os valores aos respectivos vértices de  $G'$  e  $G''$ , usando a mesma sequência de variáveis aleatórias,  $\{U_i\}_{i \geq 1}$ . Como os grafos são idênticos, então  $v' \in V(G')$  e seu vértice correspondente,  $v'' \in V(G'')$ , serão atualizados pela mesma variável aleatória. Assim, os valores associados aos vértices serão idênticos e, portanto, se a  $b'$  avalanche termina, então a  $b''$ -avalanche também é encerrada. Observe que as avalanches são idênticas até o término da  $b''$ -avalanche.

□

De acordo com [4], no modelo original Bak-Sneppen (no círculo), as avalanches ajudam a explicar a criticalidade do modelo. Para  $N$  grande, após consecutivas atualizações, o sistema parece evoluir de maneira auto-organizada para um estado crítico, isto é, para uma  $b_c$ -avalanche, onde  $b_c \approx \frac{2}{3}$ . E o mais interessante é que limitante  $b_c$  não é definido antecipadamente. Além disso, as avalanches parecem seguir uma lei de potência, no sentido de que tanto a duração quanto o alcance podem ser descritos por leis de potência (*veja [12]*). Resultados bastante significativos para o modelo Bak-Sneppen no círculo foram obtidos em [14, 15] através de uma *representação gráfica auto-similar* e da identificação de limitantes não-triviais para o modelo. Ainda nesta seção, tais limitantes, que serão denominados pontos críticos, serão definidos. Vale mencionar também que a distribuição estacionária para o modelo Bak-Sneppen em  $\Lambda_N$  foi calculada de forma explícita para  $N = 4$  e  $N = 5$  em [17, 18], respectivamente.

A Figura 2.6 mostra o modelo Bak-Sneppen em um círculo com  $N = 500$  em um possível estado estacionário, após cerca de  $10^6$  atualizações. No eixo horizontal temos os 500 vértices, e os pontos representam os valores associados aos vértices.

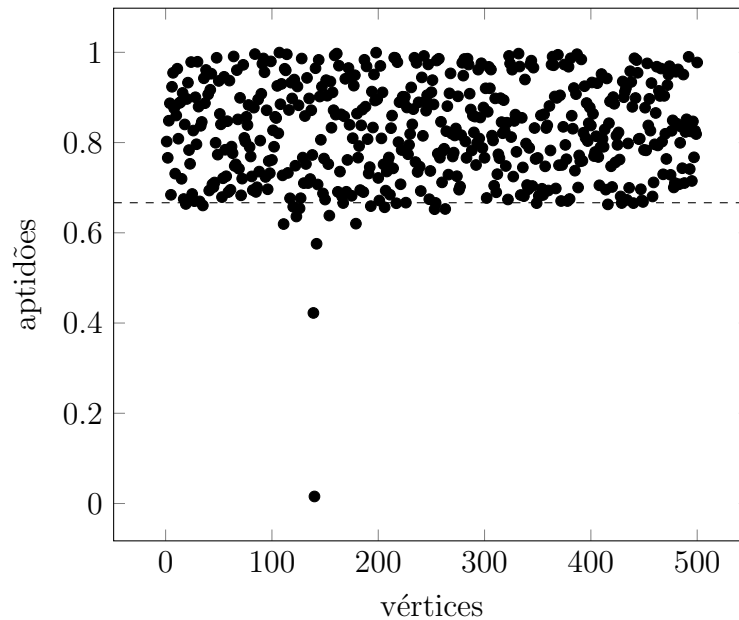


Figura 2.6: Simulação numérica do modelo Bak-Sneppen em  $\Lambda_{500}$ . Foram feitas cerca de  $10^6$  atualizações.

Vamos denotar como  $D_G^b$  e  $R_G^b$  a duração e o alcance de uma  $b$ -avalanche em um grafo  $G$ , respectivamente (lembre-se de que o alcance é a cardinalidade do conjunto-alcance). Temos que

$$D_G^b = \min\{n \geq s : A_n = \emptyset\} - s,$$

onde  $s$  é o instante de tempo de início da  $b$ -avalanche. Se  $\nexists n$  tal que  $A_n = \emptyset$ , então  $D_G^b = \infty$  (neste caso, dizemos que a  $b$ -avalanche é infinita). Já o alcance  $R_G^b$  é o número de vértices atualizados durante a  $b$ -avalanche. Observe que

$$(\Delta^* - 1)D_G^b + \Delta^* \geq R_G^b, \quad (2.1)$$

onde  $\Delta^*$  é o grau máximo de  $G$  e  $\Delta^* \geq 2$ . A relação proposta na Equação (2.1) vem do fato de que, a cada instante de tempo, no máximo,  $\Delta^* - 1$  vértices são adicionados ao conjunto-alcance. Uma  $b$ -avalanche em  $G$  é dita *abrangente* se, em um grafo finito, todos os vértices  $v \in V(G)$  são atualizados durante a avalanche, ou se, em um grafo infinito, o alcance é infinito. Agora, defina

$$|G| = \begin{cases} N, & \text{se } |V| = N < \infty \\ \infty, & \text{caso contrário.} \end{cases}$$

Temos que uma  $b$ -avalanche é abrangente se  $R_G^b = |G|$ . Por último, definimos como  $\mathbf{1}_{\{R_G^b=|G|\}}$  a função indicadora do evento em que a avalanche é abrangente.

### 2.2.1 Pontos críticos das avalanches

Aqui, vamos definir três pontos críticos para avalanches no modelo Bak-Sneppen e vamos enunciar algumas afirmações que relacionam os pontos críticos e as avalanches. Uma vez que  $G$  é um grafo localmente finito, note que as quantidades definidas acima dependem da origem da avalanche. Entretanto, veremos que os pontos críticos não dependem da origem da avalanche e, por isso, a origem foi omitida das definições de duração e alcance.

Sejam  $E(D_G^b)$  e  $E(R_G^b)$  os valores esperados da duração e do alcance, respectivamente, de uma  $b$ -avalanche em  $G$  e seja  $E(\mathbf{1}_{\{R_G^b=|G|\}}) = P(R_G^b = |G|)$  a probabilidade da ocorrência de uma  $b$ -avalanche abrangente em  $G$ . Assim como  $\theta(p)$ , quantidade definida

na Seção 1.2, é usada para definir o ponto crítico do modelo de percolação, os valores esperados serão usados para definir os pontos críticos das avalanches no modelo Bak-Sneppen. Pela Proposição 2.2.1, temos que as quantidades  $E(D_G^b)$ ,  $E(R_G^b)$  e  $P(R_G^b = |G|)$  são não-decrescentes em  $b$ . O seguinte teorema relaciona as avalanches em  $\Lambda_N$  com as avalanches em  $\mathbb{Z}$ :

**Teorema 2.2.1.** *Considere uma  $b$ -avalanche nos grafos  $\Lambda_N$  e  $\mathbb{Z}$  com a seguinte distribuição inicial:*

$$\pi_0(v) = \begin{cases} 0, & \text{se } v = 0; \\ 1, & \text{caso contrário.} \end{cases}$$

Então, para todo  $b > 0$  e  $N \geq 3$ , temos que

$$\begin{aligned} P(R_{\Lambda_N}^b = N) &\geq P(R_{\mathbb{Z}}^b = \infty), & \lim_{N \rightarrow \infty} P(R_{\Lambda_N}^b = N) &= P(R_{\mathbb{Z}}^b = \infty) \\ E(R_{\Lambda_N}^b) &\leq E(R_{\mathbb{Z}}^b), & \lim_{N \rightarrow \infty} E(R_{\Lambda_N}^b) &= E(R_{\mathbb{Z}}^b) \\ E(D_{\Lambda_N}^b) &\leq E(D_{\mathbb{Z}}^b), & \lim_{N \rightarrow \infty} E(D_{\Lambda_N}^b) &= E(D_{\mathbb{Z}}^b). \end{aligned}$$

Este teorema foi proposto e provado em [15]. Agora, vamos definir os pontos críticos das avalanches.

**Definição 2.2.1.** Seja  $G$  um grafo conexo e localmente finito e considere uma  $b$ -avalanche no modelo Bak-Sneppen em  $G$ . Definimos os seguintes pontos críticos:

$$p_c^p(G) = \inf\{b : P(R_G^b = \infty) > 0\}, \quad (2.2)$$

$$p_c^r(G) = \inf\{b : E(R_G^b) = \infty\}, \quad (2.3)$$

$$p_c^d(G) = \inf\{b : E(D_G^b) = \infty\}. \quad (2.4)$$

Vale ressaltar que as quantidades  $E(D_G^b)$ ,  $E(R_G^b)$  e  $P(R_G^b = |G|)$  são não-decrescentes em  $b$ , o que permite que os pontos críticos sejam bem definidos. Assim como o ponto crítico definido na Equação 1.1, os pontos críticos definidos acima também não dependem da origem, embora as quantidades  $E(D_G^b)$ ,  $E(R_G^b)$  e  $P(R_G^b = |G|)$  dependam. De fato,  $\forall v', v'' \in V(G)$ , a probabilidade de uma  $b$ -avalanche com origem  $v'$  terminar com uma configuração idêntica à configuração inicial de uma  $b$ -avalanche em  $v''$  é positiva. Portanto, os pontos críticos não dependem da origem. Os três pontos críticos têm uma ordem

natural. Com efeito, se a probabilidade de uma avalanche abrangente é positiva, então o valor esperado do alcance é infinito. Ou seja, o ponto crítico da Equação (2.2) não pode ser menor que o ponto crítico da Equação (2.3). Por outro lado, pela Equação (2.1), se o valor esperado da duração é finito, então o valor esperado do alcance também é finito, já que  $\Delta^* < \infty$ . Assim, o ponto crítico da Equação (2.3) não pode ser menor que o ponto crítico da Equação (2.4). Portanto,

$$p_c^d(G) \leq p_c^r(G) \leq p_c^p(G). \quad (2.5)$$

No Capítulo 4, mostraremos que  $p_c^d(G) > 0$  (Corolário 4.1.1). Para  $G = \mathbb{Z}$ , sabemos de [14] que  $p_c^p(\mathbb{Z}) \leq 1 - \exp(-68)$ , o que implica que os pontos críticos são não-trivias em  $\mathbb{Z}$ . Vale ressaltar que pontos críticos foram definidos com o objetivo de tentar entender o comportamento do modelo Bak-Sneppen. Até por isso, é razoável o questionamento sobre uma possível equivalência entre os pontos críticos definidos nas Equações (2.2), (2.3) e (2.4). Entretanto, temos apenas o seguinte resultado nessa direção:

**Teorema 2.2.2.** *Em  $G = \mathbb{Z}$ , temos que  $p_c^d(\mathbb{Z}) = p_c^r(\mathbb{Z})$ .*

Como já foi dito, simulações sugerem que as distribuições estacionárias marginais associadas aos vértices em  $\Lambda_N$  tendem para distribuições uniformes em  $(b_c, 1)$ , quando  $N \rightarrow \infty$ , onde  $b_c \approx \frac{2}{3}$  (veja [5]).

**Teorema 2.2.3.** *Em  $G = \mathbb{Z}$ , se vale  $p_c^d(\mathbb{Z}) = p_c^r(\mathbb{Z}) = p_c^p(\mathbb{Z}) = b_c$ , então a distribuição estacionária é o produto de funções distribuições uniformes no intervalo  $(b_c, 1)$ .*

**Teorema 2.2.4.** *Em  $G = \mathbb{Z}$ , temos que  $p_c^r(\mathbb{Z}) \leq \frac{3}{4}$ .*

Os Teoremas 2.2.2, 2.2.3 e 2.2.4 foram propostos e demonstrados em [15]. Observe que, sob a hipótese de igualdade dos pontos críticos definidos nas Equações (2.2), (2.3), (2.4) e  $b_c$ , o Teorema 2.2.3 prova a conjectura principal relacionada ao modelo Bak-Sneppen. A questão é que parte da hipótese do teorema ainda não foi provada. Pelo Teorema 2.2.2, sabemos apenas que  $p_c^d(\mathbb{Z}) = p_c^r(\mathbb{Z})$ .

Agora deve estar mais claro que o conhecimento sobre os pontos críticos é vital para determinar o comportamento do modelo Bak-Sneppen, ainda que o modelo em si não tenha um parâmetro de ajuste.

## 2.3 O modelo Bak-Sneppen sem memória

O objetivo nesta seção é introduzir um maneira alternativa de definir o modelo Bak-Sneppen, que denominaremos por modelo Bak-Sneppen sem memória. A motivação do nome vem do fato de que os valores atribuídos aos vértices são usados apenas para a execução do processo e, ao final de cada instante de tempo, os valores são descartados. Ou seja, ao contrário do modelo Bak-Sneppen, no novo modelo, os valores atribuídos aos vértices não são memorizados.

Antes de definirmos o modelo Bak-Sneppen sem memória, vamos definir o conceito de limiar. Seja  $Y$  uma variável aleatória com distribuição  $U(y, 1)$ , onde  $y \in [0, 1]$ . Dizemos que  $y$  é o *limiar* de  $Y$ . Uma propriedade bastante conveniente dos limiares  $y$  é que eles definem unicamente as distribuições  $U(y, 1)$  das variáveis aleatórias. Uma representação via limiares é abordada em [15].

A diferença entre os modelos Bak-Sneppen e Bak-Sneppen sem memória está no que denominamos por aptidão. Lembre-se de que, no modelo Bak-Sneppen, a aptidão é o valor atribuído ao vértice. No modelo Bak-Sneppen sem memória, a aptidão é uma variável aleatória, isto é, cada vértice é associado a uma distribuição de probabilidade.

Assim como foi dito na Seção 2.1, o termo *aptidão* continuará a ser evitado, com o intuito de prevenir possíveis confusões. Quando nos referirmos as distribuições, vamos usar a expressão “distribuição(ões) associada(s) ao(s) vértice(s)”, ou simplesmente “distribuição(ões) do(s) vértice(s)”.

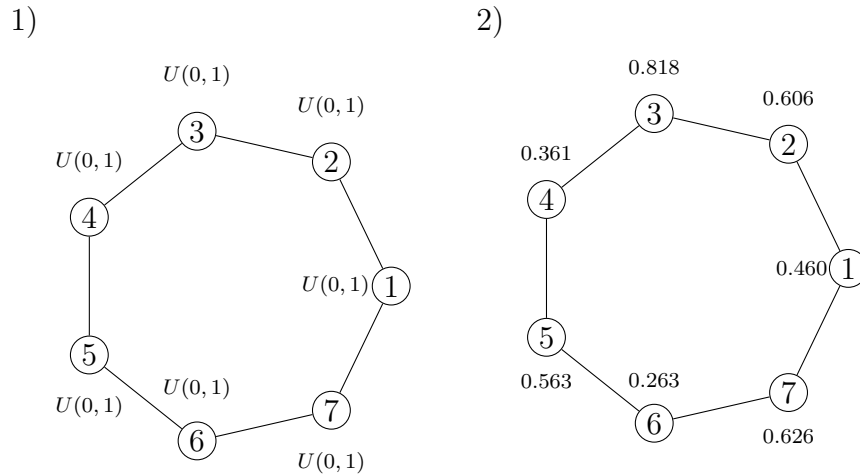
Antes de dar uma definição formal do modelo Bak-Sneppen sem memória, vamos descrever o algoritmo para a construção do modelo (para um grafo finito) e vamos dar um exemplo que ilustra a diferença para o modelo Bak-Sneppen. Seja  $G$  um grafo conexo com  $N$  vértices. Seja  $m$  o valor atribuído ao vértice minimal e  $y_i$  o limiar do vértice  $v_{G,i}$ , para  $i = 1, \dots, N$ .

1. No instante inicial, associamos a cada vértice, de maneira independente, uma distribuição  $U(0, 1)$ . Ou seja, no instante de tempo  $n = 0$ ,  $y_i = 0$ ,  $\forall i$ .
2. Através das respectivas distribuições associadas, geramos um valor para cada vértice.

3. Localizamos o vértice minimal (com valor  $m$ ), fixamos o seu valor no vértice e descartamos os valores dos demais  $N - 1$  vértices.
4. Alteramos as distribuições associadas aos demais  $N - 1$  vértices para  $U(y_i \vee m, 1)$ , onde  $y_i \vee m = \max\{y_i, m\}$ . Em relação ao vértice minimal e aos seus vértices vizinhos, alteramos suas respectivas distribuições para  $U(0, 1)$  e descartamos o valor do vértice minimal.
5. Voltamos ao passo 2.

### Um exemplo

Seja  $G_N = \Lambda_7$  o grafo circular com 7 vértices. Inicialmente, executamos os passos 1 e 2 do modelo Bak-Sneppen sem memória (Figuras 2.7.1 e 2.7.2, respectivamente). Depois de gerados os valores dos vértices, executamos o passo 3 (Figura 2.7.3). No passo 4, uma vez que  $m = 0.063$ , as distribuições dos demais vértices são uniformes acima de  $m$ , ou seja,  $U(0.063, 1)$ . As distribuições dos vizinhos do vértice minimal e do próprio vértice minimal são alteradas para  $U(0, 1)$ . A configuração ao final no instante de tempo  $n = 0$  pode ser visualizada na Figura 2.7.4.





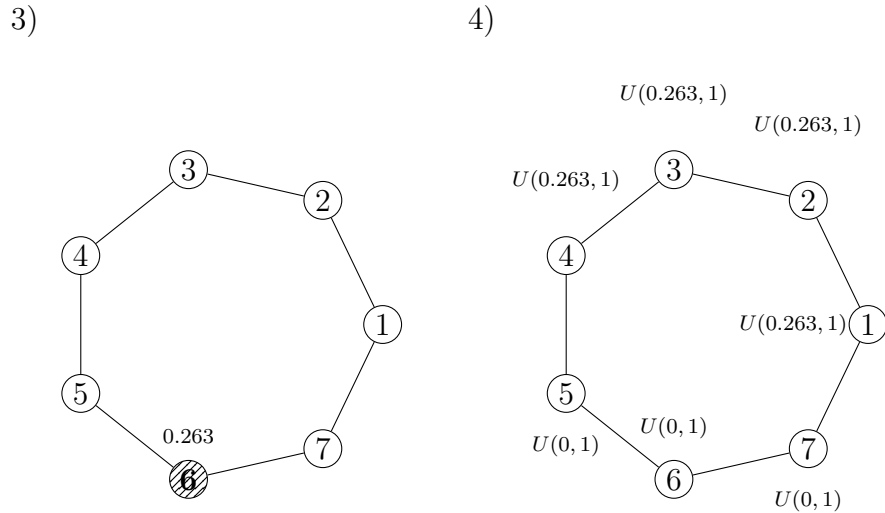
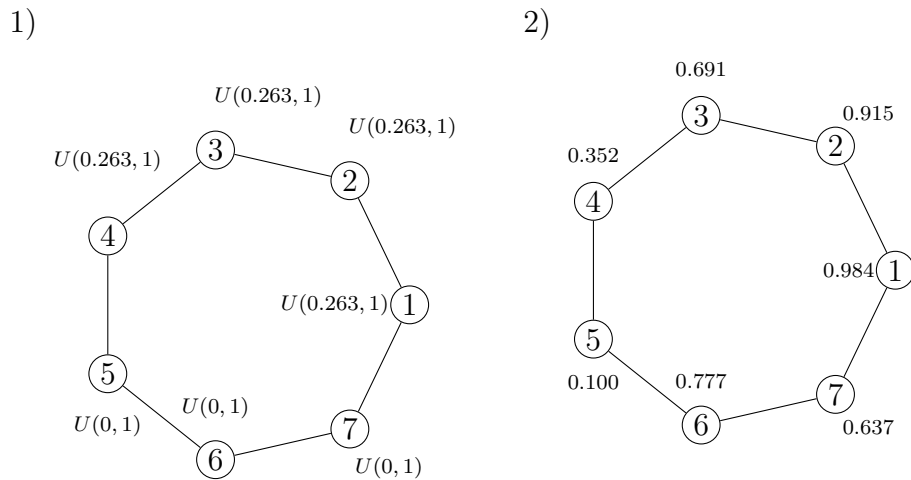


Figura 2.7: Grafo  $\Lambda_7$  após a execução de uma atualização do modelo Bak-Sneppen sem memória, no instante de tempo  $n = 0$ . Nas Figuras 2 e 3, as distribuições foram apenas omitidas, mas ainda estão associadas aos vértices.

Vamos executar o procedimento no instante de tempo  $n = 1$ . Voltando ao passo 2, geramos os valores a partir das distribuições. Veja que todos os valores atribuídos na Figura 2.8.2 são diferentes dos valores atribuídos na Figura 2.7.2, o que não acontecia no modelo Bak-Sneppen. Agora, executamos o passo 3, e o vértice minimal é o vértice 5. No passo 4, alteramos as distribuições do vértice minimal e dos seus vizinhos para  $U(0, 1)$  e alteramos as distribuições dos demais vértices, condicionando à informação do vértice minimal. Observe que as distribuições dos vértices 1, 2 e 3 não foram alteradas, pois  $0.100 < 0.263$ . Já o vértice 7 teve sua distribuição alterada, já que  $0.100 > 0$ .



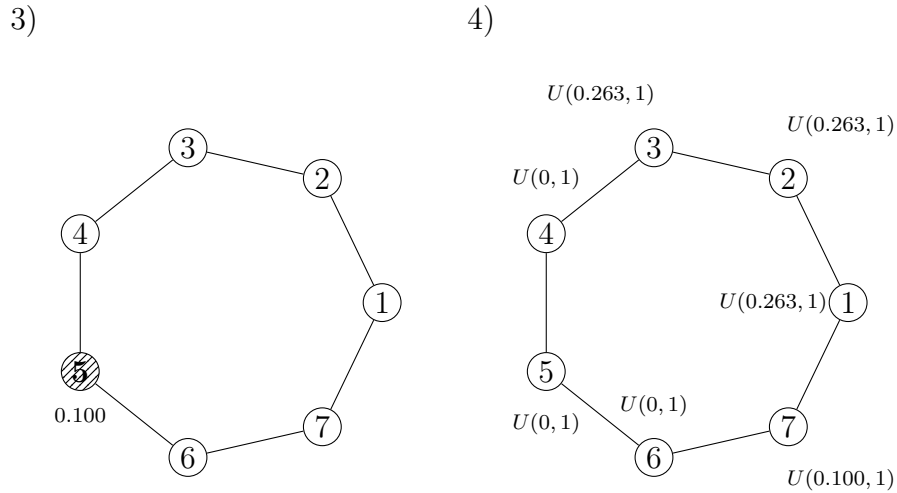


Figura 2.8: Grafo  $\Lambda_7$  após a execução de uma atualização do modelo Bak-Sneppen sem memória, no instante de tempo  $n = 1$ . Observe que os todos valores atribuídos na Figura 2.7.2 são diferentes dos valores na Figura 2, o que evidencia a diferença entre os modelos Bak-Sneppen e Bak-Sneppen sem memória.

**Proposição 2.3.1.** *Seja  $G$  um grafo finito e conexo e seja  $v_G$  um vértice de  $G$ . Em qualquer instante de tempo  $n$ , a distribuição atribuída a  $v_G$  no modelo Bak-Sneppen sem memória coincide com a distribuição deste mesmo vértice no modelo Bak-Sneppen.*

*Demonstração.* A prova da proposição será por indução em  $t$ . Seja  $G$  um grafo conexo e finito. No instante inicial,  $t = 0$ , todos os vértices têm distribuição  $U(0, 1)$ , tanto do modelo Bak-Sneppen, quanto do modelo Bak-Sneppen sem memória. Agora, suponha que a proposição é verdadeira para  $t = n$ . Seja  $Y_i$  uma variável aleatória com distribuição  $U(y_i, 1)$  e  $y_i$  o limiar do vértice  $v_{G,i}$ , para  $i = 1, \dots, N$ , no instante de tempo  $n$  e seja  $m$  o valor do vértice minimal. No instante de tempo  $n + 1$ , temos as seguintes possibilidades:

- $v_{G,i}$  é o vértice minimal ou vizinho do vértice minimal: neste caso, geramos um novo valor para  $v_{G,i}$ , de forma independente e uniformemente distribuído no intervalo  $(0, 1)$ . Portanto, a distribuição de  $v_{G,i}$  é  $U(0, 1)$  no instante de tempo  $n + 1$ .
- $v_{G,i}$  não é o vértice minimal e nem vizinho do vértice minimal: neste caso, sabemos que  $Y_i > m$ . Essa informação pode influenciar na distribuição de  $v_{G,i}$ :

se  $y_i > m$ , então

$$P(Y_i \leq x \mid Y_i > m) = \frac{P(Y_i \leq x, Y_i > m)}{P(Y_i > m)} = \frac{x - y_i}{1 - y_i}.$$

se  $y_i \leq m$ , então

$$\begin{aligned} P(Y_i \leq x \mid Y_i > m) &= \frac{P(Y_i \leq x, Y_i > m)}{P(Y_i > m)} \\ &= \frac{P(m < Y_i \leq x)}{P(Y_i > m)} \\ &= \frac{\frac{x - y_i}{1 - y_i} - \frac{m - y_i}{1 - y_i}}{1 - \frac{m - y_i}{1 - y_i}} \\ &= \frac{x - m}{1 - m}. \end{aligned}$$

Portanto, a distribuição de  $v_{G,i}$  é  $U(y_i \vee m, 1)$  no instante de tempo  $n + 1$ . Por definição, em qualquer um dos casos a distribuição do vértice  $v_{G,i}$  coincide com as distribuições definidas no modelo Bak-Sneppen sem memória. Na prática, o modelo Bak-Sneppen sem memória é simplesmente uma forma diferente de usar as informações oferecidas pelo modelo Bak-Sneppen. É como se olhássemos para as distribuições dos vértices, ao invés dos valores atribuídos aos vértices.

□

Uma consequência direta da Proposição 2.3.1 é que as distribuições dos vértices são sempre uniformes. A única diferença é nos limiares, o que enfatiza o fato dos limiares determinarem unicamente a distribuição associada aos vértices.

Portanto, sem perda de generalidade, podemos usar o modelo Bak-Sneppen sem memória para definir  $b$ -avalanches.

Para grafos localmente finitos, as distribuições associadas dos vértices durante as avalanches no modelo Bak-Sneppen sem memória podem ser definidas como uma cadeia de Markov de tempo discreto e com espaço de estados  $\Lambda = [0, 1]^V$ . Para todo  $n \in \mathbb{N}$ , seja

$$\Gamma_n : V \longrightarrow [0, 1]$$

$$v \longmapsto y_v,$$

o limiar do vértice  $v$ . Considere uma sequência de variáveis aleatórias  $\{\Gamma_n\}_{n \geq 0}$ , que denotam os estados do modelo Bak-Sneppen sem memória, onde cada estado é uma coleção de limiares, isto é,  $\{\Gamma_n(v), v \in V\}$  e  $\Gamma_0(0) = 0$ ;  $\pi_0(v) = 1$ , se  $v \neq 0$ , onde  $0$  é a origem. Seja  $\mathcal{N}(v)$  o conjunto dos vizinhos de um vértice  $v$ ,  $m_n$  o valor associado ao vértice minimal e  $w_n$  o vértice minimal no instante de tempo  $n$ . A função de transição é dada por

$$\Gamma_{n+1}(v) = \begin{cases} 0, & \text{se } v \in \mathcal{N}(w_n) \cup \{w_n\} \\ \max\{m_n, \Gamma_n(v)\}, & \text{caso contrário.} \end{cases}$$

A  $b$ -avalanche termina quando todos os valores atribuídos aos vértices que participam da avalanche estão acima do limitante  $b$ , que é o equivalente a dizer que o vértice minimal tem seu valor acima do limitante. Ou seja, neste modelo, é possível verificar quando uma avalanche terminou. Para isso, basta comparar o valor do vértice minimal com o limitante. Assim, podemos usar o modelo Bak-Sneppen sem memória em grafos localmente finitos para gerar avalanches. Feita a alteração na distribuição inicial dos vértices, a Proposição 2.3.1 é facilmente estendida para grafos localmente finitos.

## 3 Uma versão modificada do modelo Bak-Sneppen

Neste capítulo, mostraremos que o comportamento do alcance e da duração de avanços em uma versão modificada do modelo Bak-Sneppen (denominada *modelo Bak-Sneppen sem memória*) segue uma lei de potências. Para mostrar o resultado, definimos um acoplamento entre o modelo Bak-Sneppen modificado e o processo de ramificação. O material de referência é [13].

### 3.1 O modelo

Na Introdução, foi apresentada a dinâmica do modelo Bak-Sneppen. O procedimento que gera a configuração do modelo Bak-Sneppen modificado não é muito diferente. Basta modificar a regra 3 do procedimento descrito na Introdução, substituindo-a pela seguinte regra:

3'. Sorteamos e eliminamos uma outra espécie;

Em outras palavras, o modelo Bak-Sneppen modificado tem a seguinte dinâmica: ao invés de apagarmos o valor do vértice minimal e dos vizinhos, como era feito no modelo Bak-Sneppen, apagamos o valor do vértice minimal e de um outro vértice, escolhido aleatoriamente. Considere um grafo finito  $G_N$ , onde  $N$  é o número de vértices. A cada vértice atribuiremos um valor, de forma independente, e com distribuição uniforme no intervalo  $(0, 1)$ . Em cada instante de tempo  $n$ , localizamos o vértice com menor valor. Suponha que este vértice seja o vértice  $i$ , para  $i = 1, \dots, N$ . Então, apagamos o valor deste vértice e de um vértice do conjunto  $\{1, \dots, N\} \setminus \{i\}$  escolhido de forma aleatória, e atribuímos novos valores para cada um, de forma independente, e com distribuição  $U(0, 1)$ .

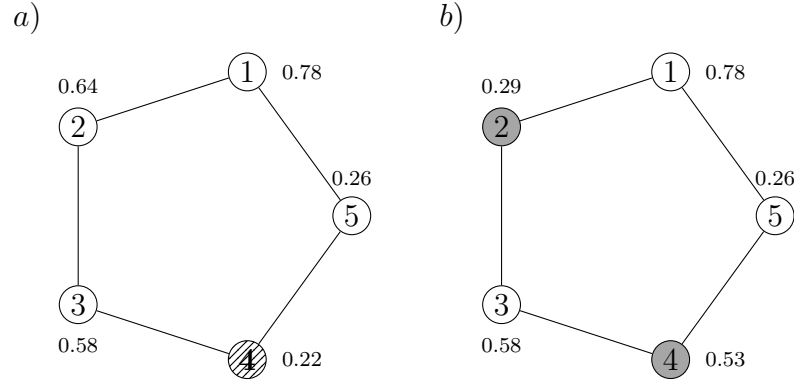


Figura 3.1: Configuração inicial de  $\Lambda_5$  (a) e configuração de  $\Lambda_5$  após uma atualização (b). Em (a), o vértice minimal foi localizado (vértice 4) e o vértice 2 foi sorteado. Em (b), executamos o procedimento de atualização. Em cinza, os vértices que foram atualizados. O próximo vértice minimal seria o vértice 5.

As definições de aptidão, atualização e avalanche no Bak-Sneppen continuam valendo para o modelo Bak-Sneppen modificado, assim como as quantidades que envolvem avalanches: vértice ativo, conjunto-alcance, alcance e duração. Além disso, o modelo Bak-Sneppen modificado ainda carrega as principais características do modelo Bak-Sneppen: é um modelo simples, no sentido de que a cada vértice é associado um único valor; e apresenta o comportamento crítico auto-organizado. Porém, ao contrário do modelo Bak-Sneppen, no modelo Bak-Sneppen modificado conseguimos provar que o comportamento das avalanches segue uma lei de potência. E é o que faremos neste capítulo. Em [14, 15], progressos significativos foram feitos para entender o comportamento das distribuições estacionárias de avalanches no modelo Bak-Sneppen, mas uma possível lei de potência ainda não foi demonstrada.

Seja  $G_N = (V, E)$  um grafo (qualquer) com  $N$  vértices. Assim como o modelo Bak-Sneppen, o modelo Bak-Sneppen modificado também pode ser definido como uma cadeia de Markov de tempo discreto e com espaço de estados  $\Lambda = [0, 1]^V$ . Para todo  $n \in \mathbb{N}$ , seja

$$\begin{aligned} \Phi_n : V &\longrightarrow [0, 1] \\ v &\longmapsto x_v, \end{aligned}$$

o valor associado ao vértice  $v$  no instante de tempo  $n$ , e considere a sequência de variáveis aleatórias  $\{\Phi_n\}_{n \geq 0}$ , que denotam os estados do modelo Bak-Sneppen modificado, onde

cada estado é uma coleção de valores da forma  $\{\Phi_n(v), v \in V\}$ . Seja  $\pi_0$  a distribuição inicial, isto é,  $\Phi_0 \sim \pi_0$ . Defina  $v_n^i = \operatorname{argmin}_{v \in V(G_N)} \Phi_n(v)$  o vértice minimal no instante de tempo  $n$  e  $v_n^j$  um vértice escolhido uniformemente no conjunto  $V \setminus \{v_n^i\}$  também no instante de tempo  $n$ . A função de transição é dada por

$$\Phi_{n+1}(v) = \begin{cases} U_n^v, & \text{se } v \in \{v_n^i, v_n^j\} \\ \Phi_n(v), & \text{caso contrário,} \end{cases}$$

onde  $U_n^v$  é uma variável aleatória uniforme no intervalo  $(0, 1)$ . As variáveis aleatórias que geram os valores para os vértices são independentes entre si.

Aqui, o alcance de uma avalanche de limitante  $b$  em  $G_N$  será denotado por  $R_N^b$  e a duração será denotada por  $D_N^b$ .

**Teorema 3.1.1.** *Seja  $G_N$  um grafo qualquer com  $N$  vértices. No modelo Bak-Sneppen modificado, temos que:*

(a) Para  $b < \frac{1}{2}$  e para  $k$  suficientemente grande, temos que,  $\forall N$ ,

$$P(D_N^b \geq k) \leq e^{-c_1(b)k},$$

onde  $c_1(b)$  é independente de  $N$ .

(b) Para  $b > \frac{1}{2}$  e  $\forall k$ , temos que

$$\lim_{N \rightarrow \infty} P(D_N^b \geq k) \geq c_2(b),$$

onde  $c_2(b) > 0$ .

(c) Para  $b = \frac{1}{2}$ , temos que

$$\lim_{k \rightarrow \infty} \lim_{N \rightarrow \infty} \sqrt{k} P(D_N^b \geq k) = \frac{2}{\sqrt{\pi}}.$$

As afirmações do teorema também são válidas quando substituímos  $D_N^b$  por  $R_N^b$ .

O Teorema 3.1.1 mostra que o comportamento da duração e do alcance de avalanches no modelo Bak-Sneppen modificado segue uma lei de potências em  $b = \frac{1}{2}$ . A Figura 3.1

ilustra o comportamento das avalanches no modelo Bak-Sneppen modificado, após cerca de  $10^6$  atualizações.

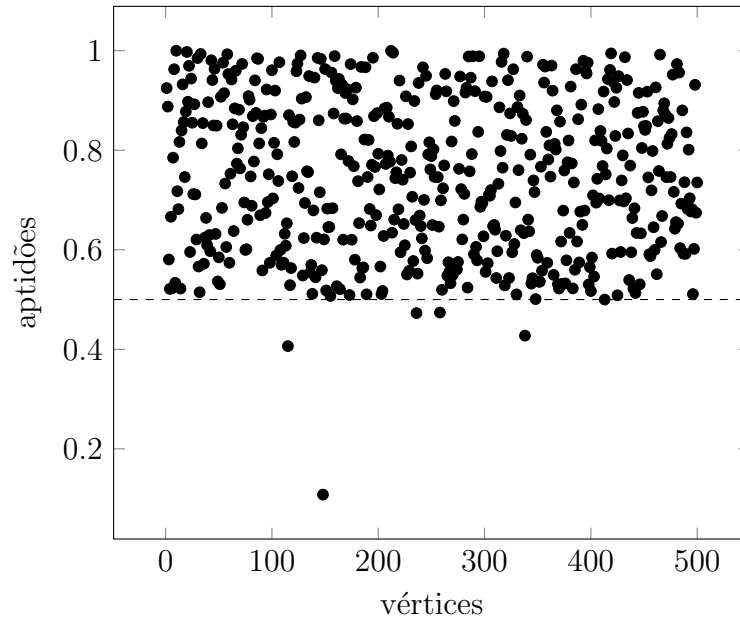


Figura 3.2: Simulação numérica do modelo Bak-Sneppen modificado em um grafo com 500 vértices. Foram feitas cerca de  $10^6$  iterações.

Para demonstrar o Teorema 3.1.1, vamos definir um novo processo, e vamos mostrar que esse processo é estocasticamente dominado por um processo de ramificação.

## 3.2 O acoplamento

Agora vamos descrever a construção de um processo que chamaremos de processo MS. O nome do processo é uma referência aos autores Ronald Meester e Anish Sarkar, responsáveis pelo material de referência deste capítulo.

### 3.2.1 O processo MS

O processo MS consiste em construir uma árvore, de forma iterativa, a partir do modelo Bak-Sneppen modificado, isto é, a cada iteração no modelo Bak-Sneppen modificado, adicionamos (ou não) nós na árvore. Sejam  $G_N$  e  $T_2^+$  os grafos subjacentes do



modelo Bak-Sneppen modificado e do processo MS, respectivamente, onde  $T_2^+$  é uma árvore binária (veja a definição de  $T_\Delta^+$  no Exemplo 1.1.2).

O processo MS é um processo estocástico que toma valores em

$$\{(\mathbb{N} \times [0, 1])^{|A|}, A \subset V(G_N)\},$$

onde o estado  $(i, x)$  significa que o vértice  $i$  tem aptidão igual a  $x$  e  $A$  é um subconjunto de  $V(G_N)$ . O processo é encerrado quando  $A = \emptyset$ . Basicamente, a  $b$ -avalanche no modelo Bak-Sneppen modificado induz o processo MS a construir uma árvore binária, denotada por  $T_2^+$ . Mostraremos que esse processo é dominado estocasticamente por um processo de ramificação, onde a distribuição do número de filhos de um indivíduo é  $\text{bin}(2, b)$ .

No instante inicial ( $n = 0$ ), no modelo Bak-Sneppen modificado, todos os vértices têm valor maior que  $b$ , exceto a origem (que será o vértice 0) que tem valor igual a  $b$ . Portanto,  $A_0 = \{0\}$ . No processo MS, adicionamos a  $T_2^+$  um nó com estado  $(0, b)$ .

Uma vez que os valores associados a todos os vértices são maiores ou iguais a  $b$ , podemos considerar uma  $b$ -avalanche no modelo Bak-Sneppen modificado em um instante de tempo  $n$ .

Seja  $A_n$  o conjunto de vértices ativos no instante de tempo  $n$ . Os nós de  $T_2^+$  correspondem aos vértices de  $G_N$  que se tornaram ativos durante a  $b$ -avalanche. Suponha que, no instante de tempo  $n$ , o vértice minimal no grafo do modelo Bak-Sneppen modificado seja o vértice  $i$ . O nó correspondente na árvore  $T_2^+$  é o nó que foi adicionado mais recentemente e que tem estado  $(i, x)$ , para algum  $x \in [0, 1]$ . Portanto, atualizamos o vértice  $i$  e o vértice  $j$ , que é escolhido aleatoriamente no conjunto  $\{1, \dots, N\} \setminus \{i\}$ . Sejam  $U_i$  e  $U_j$  variáveis aleatórias independentes  $U(0, 1)$ . Usamos  $U_i$  para atualizar o vértice  $i$  e  $U_j$  para atualizar o vértice  $j$ .

No modelo Bak-Sneppen modificado, o conjunto de vértices ativos no instante de tempo  $n + 1$  é dado por

$$A_{n+1} = \begin{cases} A_n \cup \{j\}, & \text{se } \{U_j \leq b\} \cap \{U_i \leq b\} \text{ ocorre} \\ A_n \setminus \{j\}, & \text{se } \{U_j > b\} \cap \{U_i \leq b\} \text{ ocorre} \\ \{A_n \cup \{j\}\} \setminus \{i\}, & \text{se } \{U_j \leq b\} \cap \{U_i > b\} \text{ ocorre} \\ A_n \setminus \{i, j\}, & \text{caso contrário.} \end{cases}$$

Observe que o vértice  $i$  sempre está ativo no instante de tempo  $n$ , já que é o vértice minimal, ou seja,  $i \in A_n$ . Se o evento  $\{U_i > b\}$  ocorre, então  $i \notin A_{n+1}$ . Caso contrário,  $i \in A_{n+1}$ . Já o vértice  $j$  pode ou não estar em  $A_n$  e, portanto, se o evento  $\{U_j \leq b\}$  ocorre, então  $j \in A_{n+1}$ . Caso contrário,  $j \notin A_{n+1}$ .

Agora, sejam

$$X_i = \begin{cases} 1, & \text{se } U_i \leq b \\ 0, & \text{se } U_i > b \end{cases} \quad \text{e} \quad X_j = \begin{cases} 1, & \text{se } U_j \leq b \\ 0, & \text{se } U_j > b \end{cases}$$

e defina  $X = X_i + X_j$ . Temos que  $X \sim \text{bin}(2, b)$ . Portanto, podemos definir o número de vértices em  $A_{n+1}$  da seguinte forma:

$$|A_{n+1}| = |A_n| + X - 1 - \mathbf{1}_{\{j \in A_n\}}, \quad n = 1, 2, \dots, \quad (3.1)$$

onde  $|A_0| = 1$ , por definição. Conseqüentemente, a duração do processo MS é definida por  $\min\{n : |A_n| = 0\}$ , isto é, é o primeiro instante de tempo em que  $A_n = \emptyset$ . Se  $\nexists n$  tal que  $|A_n| = 0$ , então a duração é infinita. Note que a definição da duração do processo MS coincide com a definição da duração da avalanche no modelo Bak-Sneppen modificado. Portanto, sem perda de generalidade, podemos denotar ambas as durações por  $D_N^b$ .

No processo MS, temos as seguintes regras: Se o evento  $\{U_i \leq b\}$  ocorre, adicionamos o vértice  $i$  a  $T_2^+$ , com estado  $(i, u)$ , onde  $u$  é o valor gerado a partir de  $U_i$ , e o conectamos ao nó com estado  $(i, x)$ . Se o evento  $\{U_i \leq b\}$  não ocorre, o nó não é adicionado à  $T_2^+$ . A mesma regra vale para o vértice  $j$ , mas com uma ressalva: se o vértice  $j$  já estava ativo naquele instante de tempo, removemos o nó correspondente ao vértice  $j$  de  $T_2^+$ . Se o evento  $\{U_j \leq b\}$  ocorre, adicionamos o vértice  $j$  a  $T_2^+$ , com estado  $(j, u)$ , onde  $u$  é o valor gerado a partir de  $U_j$ , e o conectamos ao nó com estado  $(i, x)$ . Se o evento  $\{U_j \leq b\}$  não ocorre, o nó não é adicionado à  $T_2^+$ . Observe que apenas vértices ativos em  $G_N$  são adicionados à  $T_2^+$ .

### 3.2.2 Um exemplo

Considere  $G_N = \Lambda_5$  (Figura 3.3.a). Como já foi dito, inicialmente todos os vértices têm valor maior que  $b$ , exceto a origem, que tem valor igual a  $b = 0.401$ . Portanto, iniciamos uma  $b$ -avalanche, com limitante igual a  $0.401$ . No processo MS (Figura 3.3.b), adicionamos a  $T_2^+$  um nó com estado  $(0, 0.401)$ .

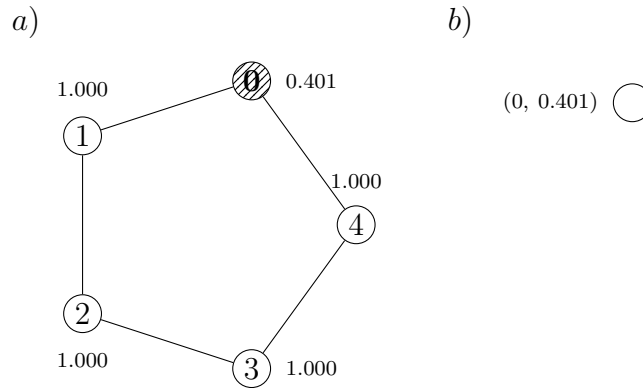


Figura 3.3: Configuração inicial de  $\Lambda_5$  (a) e  $T_2^+$  (b). Em  $\Lambda_5$ , temos: conjunto-alcance =  $\{0\}$ ;  $A_0 = \{0\}$ .

Ainda no instante de tempo  $n = 0$ , executamos o modelo Bak-Sneppen modificado em  $\Lambda_5$  (Figura 3.4.a). Considere que o vértice escolhido aleatoriamente foi o vértice 4 e seja  $(u_0, u_4) = (0.744, 0.380)$  a amostra obtida a partir de  $U_0, U_4$ . Em  $T_2^+$  (Figura 3.4.b), o vértice 5 é adicionado à árvore (pois  $0.380 \leq b$ ) e conectado ao nó com estado  $(0, 0.401)$ , o que não ocorre com o vértice 0 (já que  $0.744 > b$ ).

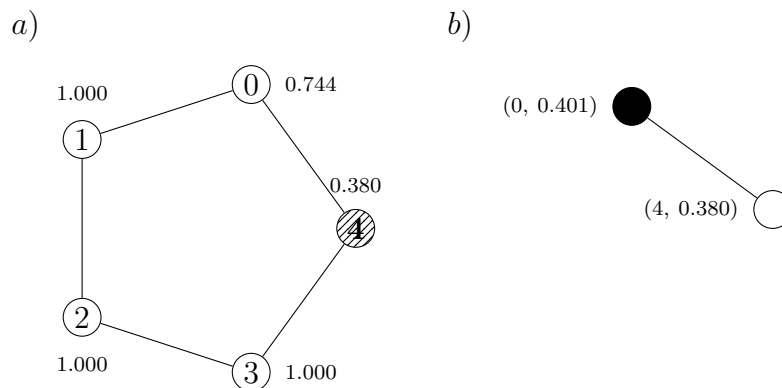


Figura 3.4: Configuração de  $\Lambda_5$  (a) e  $T_2^+$  (b) após uma atualização, ao final do instante de tempo  $n = 0$ . Em  $\Lambda_5$ , temos: conjunto-alcance =  $\{0, 4\}$ ;  $A_1 = \{4\}$ . O vértice 4 será o próximo vértice minimal.

No instante de tempo  $n = 1$ , executamos o modelo Bak-Sneppen modificado em  $\Lambda_5$  (Figura 3.5.a). Considere que o vértice escolhido aleatoriamente foi o vértice 3 e seja  $(u_4, u_3) = (0.009, 0.192)$  a amostra obtida a partir de  $U_4, U_3$ . Em  $T_2^+$  (Figura 3.5.b), os vértices 3 e 4 são adicionados à árvore (pois  $0.009 \leq b$  e  $0.192 \leq b$ ) e conectados ao nó com estado  $(4, 0.380)$ .

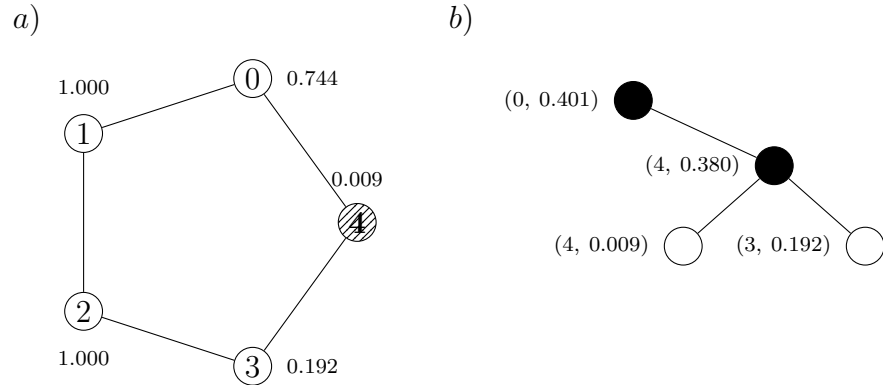


Figura 3.5: Configuração de  $\Lambda_5$  (a) e  $T_2^+$  (b) após duas atualizações, ao final do instante de tempo  $n = 1$ . Em  $\Lambda_5$ , temos: conjunto-alcance =  $\{0, 3, 4\}$ ;  $A_2 = \{3, 4\}$ . O vértice 4 será, novamente, próximo vértice minimal.

No instante de tempo  $n = 2$ , executamos o modelo Bak-Sneppen modificado em  $\Lambda_5$  (Figura 3.6.a). O vértice minimal é o vértice 4. Considere que o vértice escolhido aleatoriamente foi o vértice 1 e seja  $(u_4, u_1) = (0.200, 0.295)$  a amostra obtida a partir de  $U_4, U_1$ . Em  $T_2^+$  (Figura 3.6.b), os vértices 1 e 4 são adicionados à árvore (pois  $0.295 \leq b$  e  $0.200 \leq b$ ) e conectados ao nó com estado  $(4, 0.009)$ .

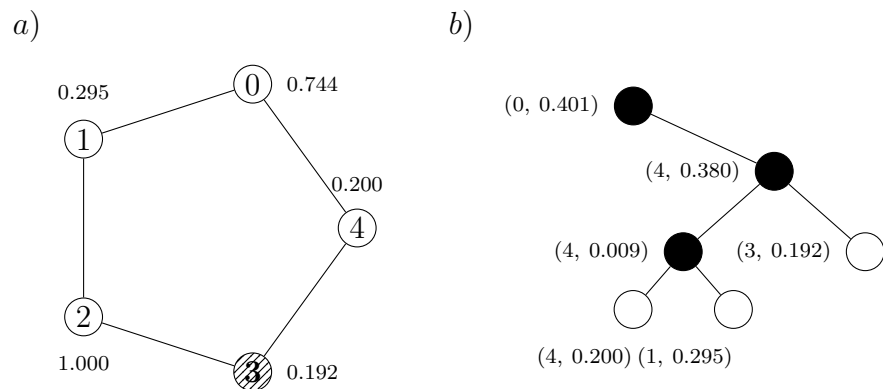


Figura 3.6: Configuração de  $\Lambda_5$  (a) e  $T_2^+$  (b) após três atualizações, ao final do instante de tempo  $n = 2$ . Em  $\Lambda_5$ , temos: conjunto-alcance =  $\{0, 1, 3, 4\}$ ;  $A_3 = \{1, 3, 4\}$ . O vértice 3 será próximo vértice minimal.

No instante de tempo  $n = 3$ , executamos o modelo Bak-Sneppen modificado em  $\Lambda_5$  (Figura 3.7.a). Note que o vértice minimal é o vértice 3, que não aparece na última geração da árvore  $T_2^+$ , o que não é um problema. Considere que o vértice escolhido aleatoriamente foi o vértice 1 e seja  $(u_3, u_1) = (0.493, 0.155)$  a amostra obtida a partir de  $U_3, U_1$ . Em  $T_2^+$  (Figura 3.7.b), observe que já existe um vértice correspondente para o vértice 1. Portanto, removemos este vértice de  $T_2^+$ . Mas veja que o valor associado ao vértice 1 está abaixo do limitante. Portanto, o vértice 1 é adicionado à árvore conectado ao nó com estado  $(3, 0.192)$ .

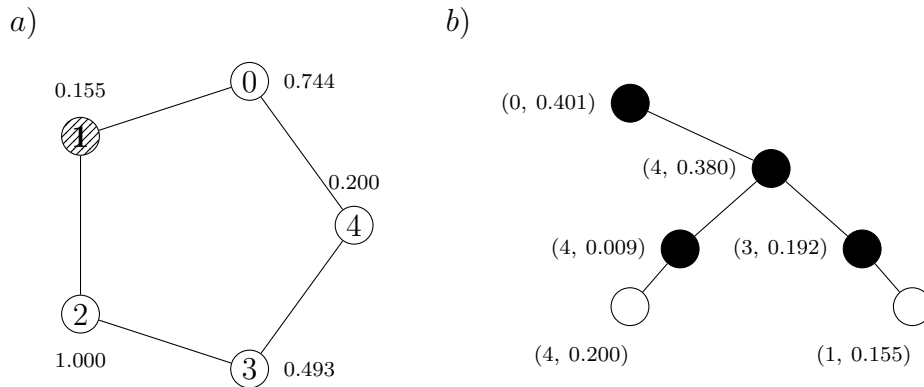


Figura 3.7: Configuração de  $\Lambda_5$  (a) e  $T_2^+$  (b) após quatro atualizações, ao final do instante de tempo  $n = 3$ . Em  $\Lambda_5$ , temos: conjunto-alcance =  $\{0, 1, 3, 4\}$ ;  $A_4 = \{1, 4\}$ . O vértice 4 será próximo vértice minimal.

No instante de tempo  $n = 4$ , executamos o modelo Bak-Sneppen modificado em  $\Lambda_5$  (Figura 3.8.a). Note que o vértice minimal é o vértice 1, e considere que o vértice escolhido aleatoriamente foi o vértice 0 e seja  $(u_1, u_0) = (0.969, 0.386)$  a amostra obtida a partir de  $U_4, U_3$ . Em  $T_2^+$  (Figura 3.8.b), veja que o valor do vértice 0 é menor que o limitante ( $0.394 < b$ ). Portanto, adicionamos o vértice 0 à  $T_2^+$  e o conectamos ao nó com estado  $(1, 0.155)$ . O vértice 1 não é adicionado a  $T_2^+$ , pois  $0.969 > b$ .

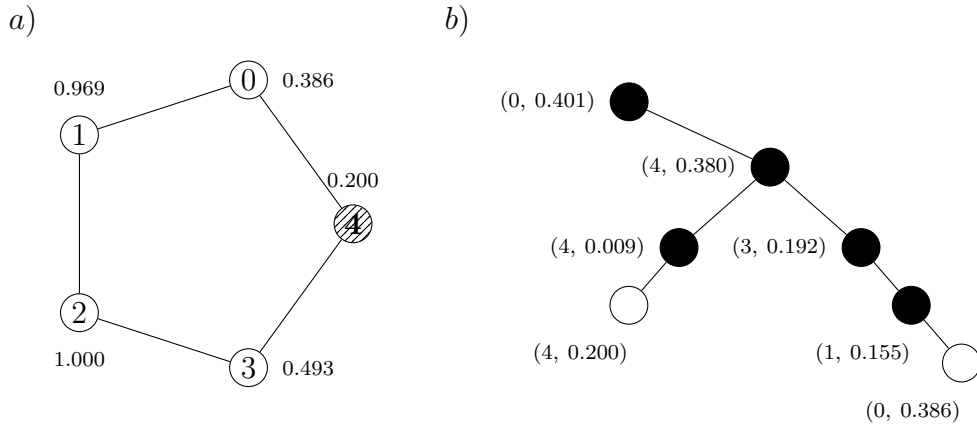


Figura 3.8: Configuração de  $\Lambda_5$  (a) e  $T_2^+$  (b) após cinco atualizações, ao final do instante de tempo  $n = 4$ . Em  $\Lambda_5$ , temos: conjunto-alcance =  $\{0, 1, 3, 4\}$ ;  $A_5 = \{0, 4\}$ . O vértice 4 será próximo vértice minimal.

Por fim, no instante de tempo  $n = 5$ , executamos o modelo Bak-Sneppen modificado em  $\Lambda_5$  (Figura 3.9.a). Note que o vértice minimal é o vértice 4. Considere que o vértice escolhido aleatoriamente foi o vértice 0 e seja  $(u_4, u_0) = (0.420, 0.532)$  a amostra obtida a partir de  $U_3, U_0$ . Veja que os novos valores estão acima do limitante  $b$ , ou seja, nenhum nó é adicionado à árvore  $T_2^+$  (Figura 3.9.b). Em  $\Lambda_5$ , os respectivos valores associados aos vértices estão acima de  $b$ , o que encerra o processo.

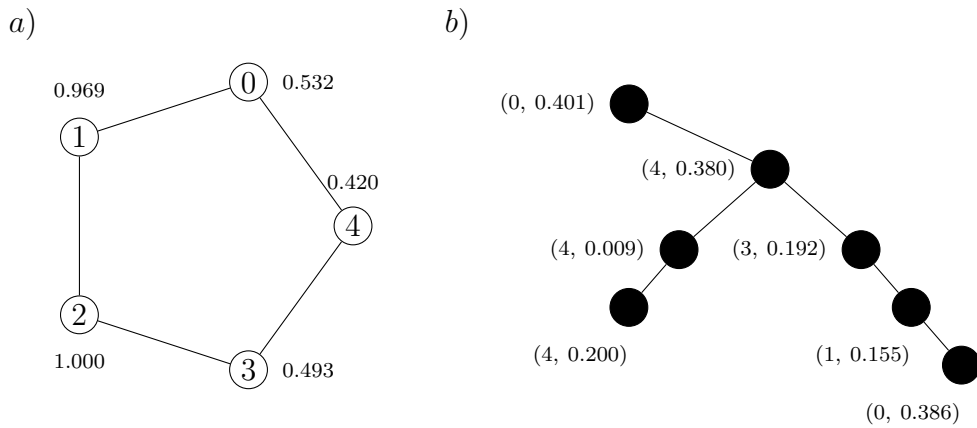


Figura 3.9: Configuração de  $\Lambda_5$  (a) e  $T_2^+$  (b) ao final da avalanche, após seis atualizações. Em  $\Lambda_5$ , temos: conjunto-alcance =  $\{0, 1, 3, 4\}$ ;  $A_6 = \emptyset$ ;  $D_5^b = 6$ .

### 3.2.3 A dominação estocástica entre os processos

Note que processo MS foi construído a partir de uma  $b$ -avalanche no modelo Bak-Sneppen modificado, de forma que suas respectivas durações são equivalentes. Isto é, olhar para a duração de uma  $b$ -avalanche no modelo Bak-Sneppen modificado é o mesmo que olhar para a duração do processo MS. Agora, vamos mostrar que o processo MS é dominado estocasticamente por um processo de ramificação, onde a distribuição do número de filhos de cada indivíduo é  $\text{bin}(2, b)$ .

**Proposição 3.2.1.** *O processo MS é dominado estocasticamente por um processo de ramificação, onde a distribuição do número de filhos de cada indivíduo é  $\text{bin}(2, b)$ . Portanto, se  $T$  é o total de indivíduos do processo de ramificação e  $D_N^b$  é a duração do processo MS, então  $T$  é estocasticamente maior que  $D_N^b$ .*

*Demonstração.* Considere o processo MS'. O processo MS' é definido de forma análoga ao processo MS. A única diferença é que no processo MS' não removemos de  $T_2^+$  os nós correspondentes aos vértices que já estão ativos. Portanto, reescrevendo a Equação (3.1), o número de vértices ativos em  $T_2^+$  no instante de tempo  $n$  é dado por

$$|A'_{n+1}| = |A'_n| + X - 1, n = 1, 2, \dots \quad (3.2)$$

onde  $|A'_0| = 1$  e  $D_N^{b'}$  =  $\min\{n : |A'_n| = 0\}$ . Se  $\nexists n$  tal que  $|A'_n| = 0$ , então  $D_N^{b'} = \infty$ .

Observe que  $D_N^{b'}$  é estocasticamente maior que  $D_N^b$ . Agora, considere  $S'_n$  e  $T'$ , definidos na Equação (1.6) da Seção 1.3. Temos que  $|A'_n| = S'_n, \forall n$ , e  $D_N^{b'} = T'$ . Portanto, pelo Lema 1.3.1, temos que  $D_N^{b'} \stackrel{d}{=} T$ , o que conclui a demonstração da proposição, já que  $D_N^b$  é dominada estocasticamente por  $D_N^{b'}$ .

□

## 3.3 Prova do Teorema 3.1.1

*Demonstração do item (a).* Pela Proposição 3.2.1, temos que

$$P(D_N^b \geq k) \leq P(T \geq k).$$

Portanto, para provar o item *a*, temos que mostrar que, para todo  $b < \frac{1}{2}$  e  $k$  suficientemente grande,

$$P(T \geq k) \leq e^{-c_1(b)k}, \quad (3.3)$$

onde  $T$  é o total de indivíduos de um processo de ramificação com distribuição do número de filhos de um indivíduo  $\text{bin}(2, b)$ .

Para verificar a Equação (3.3), vamos usar o Teorema 1.3.3, com  $X_i \sim \text{bin}(2, b)$ . Temos que

$$\begin{aligned} P(T \geq k) &= \sum_{n=k}^{\infty} \frac{1}{n} P\left(\sum_{i=1}^n X_i = n-1\right) \\ &\leq \sum_{n=k}^{\infty} P\left(\sum_{i=1}^n X_i \geq n-1\right) \\ &\leq \sum_{n=k}^{\infty} P\left(e^{t \sum_{i=1}^n X_i} \geq e^{t(n-1)}\right). \end{aligned} \quad (3.4)$$

Pela desigualdade de Markov, temos que

$$P(T \geq k) \leq \sum_{n=k}^{\infty} e^{-t(n-1)} E\left(e^{t \sum_{i=1}^n X_i}\right).$$

Como a sequência  $\{X_i\}_{i=1}^n$  é i.i.d. e tem a mesma distribuição de  $X$ , temos que

$$P(T \geq k) \leq \sum_{n=k}^{\infty} e^{-t(n-1)} (E(e^{tX}))^n.$$

Mas  $X \sim \text{bin}(2, b)$ , ou seja,  $E(e^{tX}) = (be^t + 1 - b)^2$  e, portanto,

$$\begin{aligned} P(T \geq k) &\leq \sum_{n=k}^{\infty} e^{-t(n-1)} (be^t + 1 - b)^{2n} \\ &\leq \sum_{n=k}^{\infty} e^{-tn} e^{\log(be^t + 1 - b)^{2n}} \\ &\leq \sum_{n=k}^{\infty} e^{-(t - 2 \log(be^t + 1 - b))n}. \end{aligned}$$

Minimizando o lado direito sobre  $t \geq 0$ , obtemos que

$$P(T \geq k) \leq \sum_{n=k}^{\infty} e^{-n \sup_{t \geq 0} \{t - 2 \log(be^t + 1 - b)\}}.$$



Como  $\sup_{t \geq 0} \{t - 2 \log(be^t + 1 - b)\} = \log\left(\frac{1}{4b(1-b)}\right)$ , temos que

$$\begin{aligned} P(T \geq k) &\leq \sum_{n=k}^{\infty} e^{-n \log\left(\frac{1}{4b(1-b)}\right)} \\ &\leq \sum_{n=k}^{\infty} (4b(1-b))^n. \end{aligned}$$

Por hipótese,  $b < \frac{1}{2}$ , o que implica que  $4b(1-b) < 1$ , Portanto, a série acima é convergente.

Assim,

$$P(T \geq k) \leq \frac{(4b(1-b))^k}{1 - 4b(1-b)}.$$

Assim, para  $k$  suficientemente grande, temos que

$$P(T \geq k) \leq e^{-c_1(b)k},$$

onde  $c_1(b) = \log\left(\frac{1}{4b(1-b)}\right)$ .

□

Para verificar os itens  $b - c$ , precisamos do seguinte lema:

**Lema 3.3.1.** *Seja  $T$  o total de indivíduos de um processo de ramificação com a distribuição do número de filhos de um indivíduo  $\text{bin}(2, b)$ . Temos que  $D_N^b \xrightarrow{d} T$ .*

*Demonstração.* Temos que mostrar que, para um  $k$  fixo,  $\lim_{N \rightarrow \infty} P(D_N^b \geq k) = P(T \geq k)$ . Para isso, é suficiente mostrar que  $D_N^b \xrightarrow{P} T$ . Mas veja que  $D_N^b = \min\{k : |A_k| = 0\}$  e  $T \stackrel{d}{=} \min\{k : S'_k = 0\}$ . Portanto, para mostrar que  $D_N^b$  converge em probabilidade para  $T$ , basta mostrar que, para um  $k$  fixo,  $|A_k| \xrightarrow{P} S'_k$  e o resultado segue das respectivas definições de  $D_N^b$  e  $T$ . Pela desigualdade de Markov, temos que,  $\forall \epsilon > 0$ ,

$$P(|A_k| - S'_k \geq \epsilon) \leq \frac{E(|A_k| - S'_k)}{\epsilon}$$

Pelas equações (1.6) e (3.1), temos que  $|A_k| - S'_k = \sum_{n=1}^k \mathbf{1}_{\{j \in A_n\}}$ . Assim,

$$\begin{aligned} P(|A_k| - S'_k \geq \epsilon) &\leq \frac{E\left(\sum_{n=1}^k \mathbf{1}_{\{j \in A_n\}}\right)}{\epsilon} \\ &\leq \frac{\sum_{n=1}^k P(\{j \in A_n\})}{\epsilon} \end{aligned}$$

Agora, seja  $y_n$  o número de vértices ativos no instante de tempo  $n$ . Temos que

$$P(\{j \in A_n\}) = \frac{y_n - 1}{N - 1}.$$

Com efeito, o evento  $\{j \in A_n\}$  é o evento em que o vértice sorteado no instante de tempo  $n$  já está ativo. Uma vez que o vértice minimal não é sorteado, temos que a probabilidade de que um vértice esteja ativo no instante de tempo  $n$  é  $\frac{y_n - 1}{N - 1}$ . Portanto,

$$\begin{aligned} P(|A_k| - S'_k \geq \epsilon) &\leq \sum_{n=1}^k \frac{y_n - 1}{(N - 1)\epsilon} \\ &\leq \frac{(\sum_{n=1}^k y_n) - k}{(N - 1)\epsilon}. \end{aligned}$$

No instante de tempo  $n$ , temos, no máximo,  $n + 1$  vértices ativos. Assim,

$$P(|A_k| - S'_k \geq \epsilon) \leq \frac{(1 + k)k}{2(N - 1)\epsilon}.$$

Logo,  $\forall \epsilon > 0$ ,

$$\lim_{N \rightarrow \infty} P(|A_k| - S'_k \geq \epsilon) = 0,$$

Portanto, o resultado vale para todo  $k$ , ou seja,  $D_N^b \xrightarrow{P} T$ , o que completa a prova do lema. □

*Demonstração do item (b).* Para provar o item  $b$  do Teorema 3.1.1, basta mostrar que, para qualquer  $b > \frac{1}{2}$  e  $\forall k$ ,

$$P(T \geq k) \geq c_2(b),$$

e o resultado segue pelo Lema 3.3.1. Temos que  $T$  é o total de indivíduos em um processo de ramificação com distribuição dos indivíduos  $bin(2, b)$ . Portanto,  $E(X) = 2b$ . Por hipótese, temos que  $b > \frac{1}{2}$ . Portanto, pelo Teorema 1.3.1, temos que o processo de ramificação sobrevive com probabilidade positiva, isto é,  $P(T = \infty) > 0$ . Portanto, pelo Teorema 1.3.3, temos que

$$P(T = \infty) = 1 - P(T < \infty) = 1 - \sum_{n=1}^{\infty} \frac{1}{n} \binom{2n}{n-1} b^{n-1} (1-b)^{n+1}.$$

Por hipótese, temos que  $b > \frac{1}{2}$ , o que implica que

$$P(T = \infty) = \frac{2b - 1}{b^2} = c_2(b).$$

Agora, observe que,  $\forall k$ ,  $\{T \geq k\} \supset \{T \geq k + 1\}$ , ou seja,  $\{T \geq k\} \downarrow \{T = \infty\}$ .

Assim,

$$P(T \geq k) \geq P(T = \infty) = c_2(b).$$

□

*Demonstração do item (c).* Para provar o item *c*, basta mostrar que, para  $b = \frac{1}{2}$ ,

$$\lim_{k \rightarrow \infty} \sqrt{k} P(T \geq k) = \frac{2}{\sqrt{\pi}},$$

e o resultado segue pelo Lema 3.3.1.

Vamos usar a função geradora da sequência  $\{P(T > k)\}_{k=0}^{\infty}$  para obter uma aproximação do termo  $P(T > k)$ , de forma explícita.

$$\begin{aligned} \sum_{k=0}^{\infty} P(T > k) s^k &= \sum_{k=0}^{\infty} \sum_{i=k}^{\infty} P(T = i + 1) s^k \\ &= \sum_{i=0}^{\infty} \sum_{k=0}^i P(T = i + 1) s^k \\ &= \sum_{i=0}^{\infty} P(T = i + 1) \sum_{k=0}^i s^k \\ &= \sum_{i=0}^{\infty} P(T = i + 1) \frac{1 - s^{i+1}}{1 - s} \\ &= \frac{1}{1 - s} \left( \sum_{i=0}^{\infty} P(T = i + 1) - \sum_{i=0}^{\infty} s^{i+1} P(T = i + 1) \right) \\ &= \frac{1 - G_T(s)}{1 - s}, \end{aligned} \tag{3.5}$$

onde  $G_T(s)$  é a função geradora de probabilidades de  $T$ . Agora, vamos usar o Teorema 1.3.2 para determinar  $G_T(s)$ , com  $X \sim \text{bin}(2, \frac{1}{2})$ .

$$\begin{aligned} G_T(s) &= s G_X(G_T(s)) \\ &= s \sum_{i=0}^2 (G_T(s))^i \binom{2}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{2-i} \\ &= \frac{s}{4} + \frac{s G_T(s)}{2} + \frac{s G_T^2(s)}{4}. \end{aligned}$$

Assim, temos que

$$\frac{sG_T^2(s)}{4} + \left(\frac{s}{2} - 1\right)G_T(s) + \frac{s}{4} = 0. \quad (3.6)$$

Resolvendo a equação de segundo grau, obtemos duas raízes para  $G_T(s)$ :

$$G_T'(s) = \frac{2 - s + 2\sqrt{1 - s}}{s} \quad \text{e} \quad G_T''(s) = \frac{2 - s - 2\sqrt{1 - s}}{s}.$$

Para determinar qual das soluções da Equação (3.6) é viável, vamos usar o fato de que  $G_T(0) = 0$  e calcular os limites de  $G_T'(s)$  e  $G_T''(s)$ , quando  $s \rightarrow 0$ . Temos que

$$\lim_{s \rightarrow 0} G_T'(s) = \lim_{s \rightarrow 0} \frac{2 - s + 2\sqrt{1 - s}}{s} = \infty$$

e

$$\lim_{s \rightarrow 0} G_T''(s) = \lim_{s \rightarrow 0} \frac{2 - s - 2\sqrt{1 - s}}{s} \stackrel{(H)}{=} \lim_{s \rightarrow 0} -1 + \frac{1}{\sqrt{1 - s}} = 0.$$

Portanto,  $G_T''(s)$  é a solução viável da Equação (3.6), o que implica que

$$G_T(s) = \frac{2 - s - 2\sqrt{1 - s}}{s}. \quad (3.7)$$

Substituindo a Equação (3.7) na Equação (3.5), obtemos que

$$\begin{aligned} \sum_{k=0}^{\infty} P(T > k)s^k &= \frac{1}{1 - s} - \frac{2 - s - 2\sqrt{1 - s}}{s(1 - s)} \\ &= \frac{-2 + 2s + 2\sqrt{1 - s}}{s(1 - s)} \\ &= -\frac{2}{s} + \frac{2(1 - s)^{-1/2}}{s}. \end{aligned} \quad (3.8)$$

Usando o Teorema Binomial de Newton, temos que

$$\begin{aligned} (1 - s)^{-1/2} &= \sum_{k=0}^{\infty} \binom{-\frac{1}{2}}{k} (-s)^k (1)^{-1/2-k} \\ &= \sum_{k=0}^{\infty} \frac{(-\frac{1}{2})!}{(-\frac{1}{2} - k)!k!} (-1)^k s^k \\ &= \sum_{k=0}^{\infty} \frac{s^k}{2^k k!} \prod_{i=1}^k (2i - 1). \end{aligned}$$

Observe que  $\prod_{i=1}^k (2i-1) = \frac{(2k)!}{2^k k!}$ . Assim,

$$\begin{aligned} (1-s)^{-1/2} &= \sum_{k=0}^{\infty} \frac{s^k (2k)!}{2^k k! 2^k k!} \\ &= 1 + \sum_{k=1}^{\infty} \frac{(2k)!}{2^{2k} (k!)^2} s^k. \end{aligned} \quad (3.9)$$

Pela Fórmula de Stirling, temos que

$$\sqrt{2\pi n} n^n e^{-n} e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi n} n^n e^{-n} e^{\frac{1}{12n}}, \quad n = 1, 2, \dots$$

Reescrevendo a Equação (3.9), temos que

$$1 + \sum_{k=1}^{\infty} \frac{e^{\frac{-18k-1}{144k^2+6k}}}{\sqrt{\pi k}} s^k < (1-s)^{-1/2} < 1 + \sum_{k=1}^{\infty} \frac{e^{\frac{-36k+1}{288k^2+24k}}}{\sqrt{\pi k}} s^k. \quad (3.10)$$

Substituindo a Equação (3.10) na Equação (3.8), obtemos que

$$\sum_{k=0}^{\infty} \frac{2e^{\frac{-18k-1}{144k^2+6k}}}{\sqrt{\pi(k+1)}} s^k < \sum_{k=0}^{\infty} P(T > k) s^k < \sum_{k=0}^{\infty} \frac{2e^{\frac{-36k+1}{288k^2+24k}}}{\sqrt{\pi(k+1)}} s^k,$$

Portanto,

$$\frac{2e^{\frac{-18k+17}{144k^2-282k+138}}}{\sqrt{\pi k}} < P(T \geq k) < \frac{2e^{\frac{-36k+37}{288k^2-552k+264}}}{\sqrt{\pi k}},$$

o que implica que

$$\lim_{k \rightarrow \infty} \sqrt{k} P(T \geq k) = \frac{2}{\sqrt{\pi}},$$

o que completa a prova. □

O teorema também é válido para  $R_N^b$ . Para o item *a*, basta observar que

$$R_N^b \leq D_N^b + 1. \quad (3.11)$$

Para os itens *b* e *c*, basta mostrar que  $R_N^b \xrightarrow{P} D_N^b$ . Com efeito, a probabilidade de um vértice ser sorteado mais de uma vez tende à zero quando  $N \rightarrow \infty$ . Observe que se nenhum vértice for sorteado mais de uma vez, obtemos a igualdade na Equação (3.11). A prova é análoga à prova do Lema 3.3.1.

## 4 As cotas para os pontos críticos de avalanches no modelo Bak-Sneppen

Nosso objetivo é comparar as avalanches no modelo Bak-Sneppen com dois processos mais antigos e mais bem estudados. Esses processos são conhecidos como processos de ramificação e modelo de percolação em sítios. Uma comparação simples com o processo de ramificação fornece uma cota inferior, enquanto que uma comparação mais complexa com o modelo de percolação fornece uma cota superior para os pontos críticos definidos nas Equações (2.2), (2.3) e (2.4).

### 4.1 O teorema e algumas consequências

**Teorema 4.1.1.** *Seja  $G$  um grafo conexo e localmente finito. Temos que*

$$\frac{1}{\Delta^* + 1} \leq p_c^{BS}(G) \leq p_c^{\text{sítios}}(G),$$

onde  $\Delta^*$  é o grau máximo de  $G$ ,  $p_c^{\text{sítios}}(G)$  é definido na Equação (1.1) e  $BS \in \{p, r, d\}$ . Isto é,  $p_c^{BS}(G)$  representa os pontos críticos definidos nas Equações (2.2), (2.3) e (2.4).

A prova do Teorema 4.1.1 será dividida em duas partes: a cota inferior para  $p_c^{BS}(G)$  e a cota superior para  $p_c^{BS}(G)$ . Vamos começar com a cota inferior.

Para provar que vale a cota inferior, basta verificar o lema abaixo, e o resultado da cota inferior do Teorema 4.1.1 segue da Equação (2.5).

**Lema 4.1.1.** *Seja  $G$  um grafo conexo e localmente finito. Temos que*

$$p_c^d(G) \geq \frac{1}{\Delta^* + 1},$$

onde  $p_c^d(G)$  é definido na Equação (2.4) e  $\Delta^*$  é o grau máximo de  $G$ .

Como dito acima, o Lema 4.1.1 combinado com a Equação (2.5) fornece a cota inferior do Teorema 4.1.1, cuja prova será feita na Seção 4.2. O seguinte corolário é uma consequência direta do Lema 4.1.1.

**Corolário 4.1.1.** *Seja  $G$  um grafo conexo e localmente finito. Temos que*

$$p_c^d(G) > 0.$$

*Demonstração.* Com efeito, uma vez que  $\Delta^* < \infty$ , temos que, pelo Lema 4.1.1,

$$p_c^d(G) \geq \frac{1}{\Delta^* + 1} > 0.$$

□

O Corolário 4.1.1 nos diz que os pontos críticos definidos nas Equações (2.2), (2.3) e (2.4) são estritamente positivos. Para verificar a cota superior do Teorema 4.1.1, vamos provar o seguinte lema:

**Lema 4.1.2.** *Seja  $G$  um grafo conexo e localmente finito. Temos que*

$$p_c^p(G) \leq p_c^{\text{sítios}}(G),$$

onde  $p_c^p(G)$  é definido na Equação (2.2) e  $p_c^{\text{sítios}}(G)$  definido na Equação (1.1).

O Lema 4.1.2 combinado com a Equação (2.5) fornece a cota superior do Teorema 4.1.1. O lema implica que, em muitos grafos localmente finitos,  $p_c^p$  é não-trivial. Para a avalanche no modelo Bak-Sneppen em  $\mathbb{Z}$ , o Teorema 4.1.1 dá uma cota superior trivial, já vista na Observação 1.2.2. A prova do Lema 4.1.2 será feita ao longo deste capítulo, a partir da Seção 4.3.

Pelo Teorema 1.2.2, sabemos que o ponto crítico do modelo de percolação em sítios em  $T_\Delta$  é igual a  $1/(\Delta - 1)$ , o que implica no seguinte corolário do Teorema 4.1.1:

**Corolário 4.1.2.** *Seja  $T_\Delta$  uma árvore  $\Delta$ -regular. Temos que*

$$\frac{1}{\Delta + 1} \leq p_c^{BS}(T_\Delta) \leq \frac{1}{\Delta - 1}.$$

Aplicando a expansão proposta por [8] para o ponto crítico do modelo de percolação em  $\mathbb{Z}^d$ , temos outro corolário:

**Corolário 4.1.3.** *Seja  $\mathbb{Z}^d$  uma rede hipercúbica  $d$ -dimensional. Temos que*

$$\frac{1}{2d+1} \leq p_c^{BS}(\mathbb{Z}^d) \leq \frac{1}{2d} + \frac{1}{(2d)^2} + O(d^{-3}), \quad d \rightarrow \infty.$$

*Observação 4.1.1.* Neste capítulo, muitas vezes usaremos a letra  $p$  para nos referir a  $p$ -avalanches. Por isso, é importante observar que este  $p$  não tem nenhuma relação com o  $p$  no ponto crítico definido na Equação (2.2). O  $p$  da Equação (2.2) é apenas para fazer referencia à probabilidade de um alcance infinito.

## 4.2 Prova do Lema 4.1.1

A ideia é acoplar um processo de ramificação i.i.d. com distribuição comum  $\text{bin}(\Delta^* + 1, p)$  a uma avalanche de limitante  $p$  no modelo Bak-Sneppen em  $G$ , de forma que o processo de ramificação domine estocasticamente a  $p$ -avalanche no modelo Bak-Sneppen. Para a prova, vamos definir uma generalização do processo MS (descrito no Capítulo 3), que será denominado processo GMS e será construído a partir de uma  $p$ -avalanche no modelo Bak-Sneppen.

### 4.2.1 O acoplamento

Considere um grafo  $G$  conexo e localmente finito, com o grau máximo de  $G$  igual a  $\Delta^*$ . O processo GMS consiste em construir uma árvore, de forma iterativa, a partir do modelo Bak-Sneppen, de forma que, a cada atualização no grafo do modelo Bak-Sneppen, adicionamos (ou não) nós na árvore. Sejam  $G$  e  $T_{\Delta^*+1}^+$  os grafos subjacentes da avalanche no modelo Bak-Sneppen e do processo GMS, respectivamente, onde  $T_{\Delta^*+1}^+$  é uma árvore enraizada  $\Delta^*$ -ária (veja a definição de  $T_{\Delta}^+$  no Exemplo 1.1.2).

De maneira análoga ao processo MS (definido no Capítulo 3), o processo GMS é um processo estocástico que toma valores em

$$\{(\mathbb{N} \times [0, 1] \times \{a, i\})^{|A|}, \quad A \subset V(G)\},$$



em que o estado  $(v, x, a)$  significa que o vértice  $v$  é associado a um valor igual a  $x$  e está ativo e o estado  $(v, x, i)$  significa que o vértice  $v$  é associado a um valor igual a  $x$  e está inativo. Por fim,  $A$  é um subconjunto de  $V(G)$ . O processo é encerrado quando  $A = \emptyset$ . A  $p$ -avalanche no modelo Bak-Sneppen induz o processo GMS a construir uma árvore  $\Delta^*$ -ária, denotada por  $T_{\Delta^*+1}^+$ . Mostraremos que esse processo é dominado estocasticamente por um processo de ramificação, onde a distribuição do número de filhos de um indivíduo é  $\text{bin}(\Delta^* + 1, p)$ .

No instante inicial ( $n = 0$ ), na avalanche no modelo Bak-Sneppen, todos os vértices têm valor igual a 1, exceto a origem (denotada por 0), que tem valor igual a  $p$ . Portanto,  $A_0 = \{0\}$ . No processo MS, adicionamos a  $T_{\Delta^*+1}^+$  um nó com estado  $(0, p, a)$ .

Agora, considere uma  $p$ -avalanche (com limitante  $p$ ) no modelo Bak-Sneppen em um instante de tempo  $n$ , isto é, os valores de todos os vértices são maiores ou iguais a  $p$ .

Seja  $A_n$  o conjunto de vértices ativos no instante de tempo  $n$ . Os nós de  $T_{\Delta^*+1}^+$  correspondem aos vértices de  $G$  que se tornaram ativos durante a  $p$ -avalanche. Suponha que, no instante de tempo  $n$ , o vértice minimal no grafo do modelo Bak-Sneppen seja o vértice  $w$ . O nó correspondente na árvore  $T_{\Delta^*+1}^+$  é o nó que foi adicionado mais recentemente e que tem estado  $(w, x, a)$ . Portanto, atualizamos o vértice  $w$  e os vizinhos do vértice  $w$ , isto é, atualizamos os vértices pertencentes ao conjunto  $\mathcal{N}(w) \cup \{w\}$ . Seja  $\{U_k\}_{k=1}^{\Delta_w+1}$  uma sequência de variáveis aleatórias independentes  $U(0, 1)$ , onde  $\Delta_w$  é o grau de  $w$ . Usamos essa sequência para atualizar os vértices do conjunto  $\mathcal{N}(w) \cup \{w\}$ . Em cada instante de tempo, os vértices do conjunto serão indexados por  $k = 1, \dots, \Delta_w + 1$ . Assim, se o evento  $\{U_k \leq p\}$  ocorre, o  $k$ -ésimo vértice é associado um valor que está abaixo do limitante  $p$ . Caso contrário,  $k$ -ésimo vértice é associado um valor que está acima de  $p$ . Isso induz o seguinte acoplamento: defina

$$X_k = \begin{cases} 1, & \text{se } U_k \leq p \\ 0, & \text{se } U_k > p \end{cases}$$

e seja  $X_w = \sum_{k=1}^{\Delta_w+1} X_k$ . Temos que  $X_w \sim \text{bin}(\Delta_w + 1, p)$ . Portanto, podemos definir o número de vértices em  $A_{n+1}$  da seguinte forma:

$$|A_{n+1}| = |A_n| + X_w - 1 - \sum_{v \in \mathcal{N}(w)} \mathbf{1}_{\{v \in A_n\}}, \quad n = 1, 2, \dots, \quad (4.1)$$

onde  $|A_0| = 1$ , por definição. Consequentemente, a duração do processo GMS é definida por  $\min\{n : |A_n| = 0\}$ , isto é, é o primeiro instante de tempo em que  $A_n = \emptyset$ . Se  $\nexists n$  tal que  $|A_n| = 0$ , então a duração é infinita. Note que a definição da duração do processo GMS coincide com a definição da duração da avalanche no modelo Bak-Sneppen. Portanto, sem perda de generalidade, podemos denotar ambas as durações por  $D_G^p$ .

No processo GMS, temos as seguintes regras para cada  $v \in \mathcal{N}(w) \cup \{w\}$ :

- 1) Se  $v$  é o vértice minimal (isto é, se  $v = w$ ): alteramos o estado de  $w$  para  $(w, x, i)$ . Se o evento  $\{U_k \leq p\}$  ocorre (onde  $k$  é o índice de  $w$ ), adicionamos o vértice  $w$  a  $T_{\Delta^*+1}^+$ , com estado  $(w, u, a)$ , onde  $u$  é o valor gerado a partir de  $U_k$ , e o conectamos ao nó  $(w, x, i)$ . Se o evento  $\{U_k \leq p\}$  não ocorre, o nó não é adicionado à  $T_{\Delta^*+1}^+$ .
- 2) Se  $v$  não é o vértice minimal: verificamos se o vértice  $v$  já está ativo, isto é, verificamos se  $v \in A_n$ . Se sim, removemos de  $T_{\Delta^*+1}^+$  o nó correspondente a este vértice. Se o evento  $\{U_k \leq p\}$  ocorre (onde  $k$  é o índice de  $v$ ), adicionamos o vértice  $v$  a  $T_{\Delta^*+1}^+$ , com estado  $(v, u, a)$ , onde  $u$  é o valor gerado a partir de  $U_k$ , e o conectamos ao nó  $(w, x, i)$ . Se o evento  $\{U_k \leq p\}$  não ocorre,  $v$  não é adicionado à  $T_{\Delta^*+1}^+$ .

Note que processo GMS foi construído a partir de uma  $p$ -avalanche no modelo Bak-Sneppen, de forma que suas respectivas durações são equivalentes. Isto é, olhar para a duração de uma  $p$ -avalanche no modelo Bak-Sneppen é o mesmo que olhar para a duração do processo GMS. Agora, vamos mostrar que este processo é dominado estocasticamente por um processo de ramificação.

**Proposição 4.2.1.** *O processo GMS é dominado estocasticamente por um processo de ramificação, onde a distribuição do número de filhos de cada indivíduo é  $\text{bin}(\Delta^* + 1, p)$ . Portanto, se  $T$  é o total de indivíduos do processo de ramificação e  $D_G^b$  é a duração do processo GMS, então  $T$  é estocasticamente maior que  $D_G^b$ .*

*Demonstração.* Considere o processo GMS. Primeiro, vamos comparar a Equação (4.1) com a Equação (1.6), definida na Seção 1.3, e vamos mostrar (por indução em  $n$ ) que, em qualquer instante de tempo,  $|A_n|$  é dominada estocasticamente por  $S'_n$ . No instante inicial, temos que,  $|A_0| = 1$  e  $S'_0 = 1$ , por definição. Agora, considere que vale a afirmação para  $n = k$ , isto é,  $|A_k|$  é dominada estocasticamente por  $S'_k$ .

No instante de tempo  $n = k + 1$ , pela Equação (4.1), temos que

$$|A_{k+1}| = |A_k| + X_w - 1 - \sum_{v \in \mathcal{N}(w)} \mathbf{1}_{\{v \in A_k\}}.$$

Observe que  $X_w$  é estocasticamente menor que  $X'$ , onde  $X' \sim \text{bin}(\Delta^* + 1, p)$ , já que  $\Delta_w \leq \Delta^*$ ,  $\forall w$ . Por outro lado, pela hipótese de indução, temos que  $|A_k|$  é dominada estocasticamente por  $S'_k$ . Portanto (com um certo abuso de notação),

$$|A_{k+1}| = |A_k| + X_w - 1 - \sum_{v \in \mathcal{N}(w)} \mathbf{1}_{\{v \in A_k\}} \leq S'_k + X' - 1 = S'_{k+1},$$

o que completa a prova da afirmação. Agora, observe as definições de  $T' = \min\{n : |S'_n| = 0\}$  (Equação (1.6)) e  $D_G^b$ . Veja que  $T'$  domina estocasticamente  $D_G^b$ . Por outro lado, pelo Lema 1.3.1, temos que  $T' \stackrel{d}{=} T$ , onde  $T$  é o total de indivíduos de um processo de ramificação com a distribuição do número de indivíduos sendo a mesma de  $X \stackrel{d}{=} X'$ . Portanto,  $T$  é estocasticamente maior que  $D_G^b$ , o que conclui a demonstração da proposição.

□

A Proposição 4.2.1 nos diz que, se o valor esperado de  $T$  é finito, então o valor esperado de  $D_N^b$  também é finito. Logo, o ponto crítico definido na Equação (2.4) não pode ser menor que o ponto crítico do processo de ramificação, definido no Teorema 1.3.1. Pelo Teorema 1.3.1, temos que  $p = 1/(\Delta^* + 1)$ , já que  $E(X) = (\Delta^* + 1)p$ . E como  $p_c^d(G)$  não pode ser menor que o ponto crítico no processo de ramificação, segue que

$$p_c^d(G) \geq \frac{1}{\Delta^* + 1},$$

o que completa a prova do Lema 4.1.1.

### 4.2.2 Um exemplo

O comportamento dos processos é ilustrado pelo seguinte exemplo, apresentado nas figuras abaixo. Seja  $G = T_4^*$  a árvore enraizada 4-regular (veja a definição no Exemplo 1.1.2). Por fins de ilustração, mostraremos apenas a parte do grafo onde ocorre a dinâmica. Uma vez que o grau máximo de  $G$  é igual a 4 (apenas a raiz tem grau 3), temos que o grafo subjacente do processo GMS é  $T_5^+$ , que será construído a partir das atualizações em  $T_4^*$ . Em  $T_4^*$  (figuras à esquerda), os vértices preenchidos com linhas diagonais são os vértices minimais. Em  $T_5^+$  (figuras à direita), os vértices em branco são os vértices ativos e os vértices em preto são os vértices inativos. Considere uma  $p$ -avalanche no modelo Bak-Sneppen, com  $p = 0.450$ , em  $T_4^*$ . Como já foi dito, no instante inicial ( $n = 0$ ), na avalanche no modelo Bak-Sneppen, todos os vértices têm valor igual a 1, exceto a origem (denotada por 0), que tem valor igual a  $p = 0.450$  (Figura 4.1.a). Portanto,  $A_0 = \{0\}$ . No processo MS, adicionamos a  $T_5^+$  um nó com estado  $(0, 0, a)$  (Figura 4.1.b).



Figura 4.1: Grafos  $T_4^*$  (a) e  $T_5^+$  (b) antes do início da avalanche. Em  $T_4^*$ , o vértice 0 (em linhas diagonais) será o vértice minimal. Em  $T_5^+$ , o estado do vértice 0 (em branco) é  $(0, 0.450, a)$ . Temos que  $A_0 = \{0\}$  e  $|A_0| = 1$ .

Ainda no instante de tempo  $n = 0$ , observe que o vértice 0 é o vértice minimal e tem três vizinhos (vértices 1, 2 e 3), ou seja,  $X_0 \sim bin(4, p)$ . Agora, considere que, após a atualização em  $T_4^*$ , os valores gerados para os vértices 0, 1, 2 e 3 estão abaixo do limitante  $p$ , isto é,  $X_0 = 4$ . A Figura 4.2.a ilustra a configuração de  $T_4^*$ , após a primeira atualização. No grafo  $T_5^+$ , aplicamos as regras do processo GMS. O vértice 0 é declarado inativo, pois o seu vértice correspondente em  $T_4^*$  é o vértice minimal, e adicionamos os vértices 0, 1, 2 e 3 à  $T_5^+$ , pois seus respectivos valores estão abaixo de  $p$ , e os conectamos ao nó  $(0, 0.450, i)$ . A Figura 4.2.b mostra a configuração de  $T_5^+$ , após a primeira atualização.

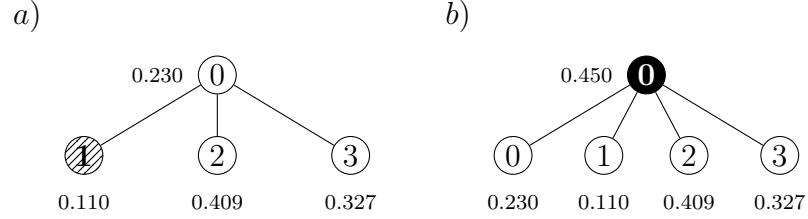


Figura 4.2: Grafos  $T_4^*$  (a) e  $T_5^+$  (b) ao final do instante de tempo  $n = 0$ , após a primeira atualização. Em  $T_4^*$ , o vértice 1 (em linhas diagonais) será o próximo vértice minimal. Em  $T_5^+$ , o vértice 0 (em preto) foi declarado inativo, isto é, seu estado é  $(0, 0.450, i)$ . Temos que  $A_1 = \{0, 1, 2, 3\}$  e  $|A_1| = 1 + 4 - 1 - 0 = 4$ .

No instante de tempo  $n = 1$ , o vértice 1 é o vértice minimal e tem quatro vizinhos (vértices 0, 4, 5 e 6), ou seja,  $X_1 \sim \text{bin}(5, p)$ . Agora, considere que, após a atualização em  $T_4^*$ , apenas o valor gerado para o vértice 0 está abaixo do limitante  $p$ , isto é,  $X_1 = 1$ . A Figura 4.3.a ilustra a configuração de  $T_4^*$ , após a segunda atualização. No grafo  $T_5^+$ , aplicamos as regras do processo GMS. O vértice 1 é declarado inativo, pois o seu vértice correspondente em  $T_4^*$  é o vértice minimal, e adicionamos ao nó  $(1, 0.110, i)$  o vértice 0, pois seu respectivo valor está abaixo de  $p$ . Mas veja que o vértice 0 já estava ativo em  $T_5^+$ . Portanto, removemos o nó  $(0, 0.230, a)$  de  $T_5^+$ , antes do nó  $(0, 0.366, a)$  ser adicionado. A Figura 4.3.b mostra a configuração de  $T_5^+$ , após a segunda atualização.

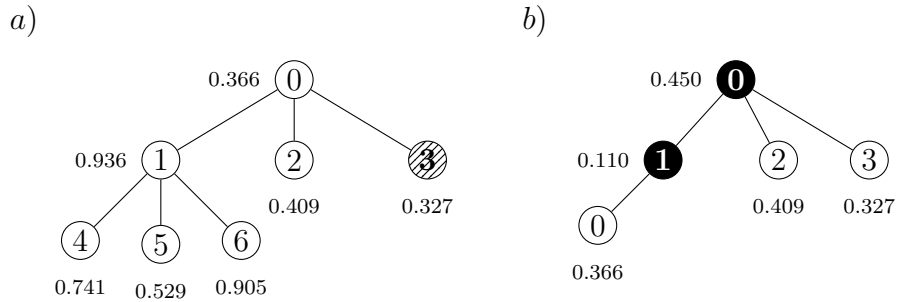


Figura 4.3: Grafos  $T_4^*$  (a) e  $T_5^+$  (b) ao final do instante de tempo  $n = 1$ , após a segunda atualização. Em  $T_4^*$ , o vértice 3 (em linhas diagonais) será o próximo vértice minimal. Em  $T_5^+$ , o vértice 1 (em preto) foi declarado inativo, isto é, seu estado é  $(0, 0.450, i)$ . Temos que  $A_2 = \{0, 2, 3\}$  e  $|A_2| = 4 + 1 - 1 - 1 = 3$ .

No instante de tempo  $n = 2$ , o vértice 3 é o vértice minimal e tem quatro vizinhos (vértices 0, 7, 8 e 9), ou seja,  $X_3 \sim \text{bin}(5, p)$ . Agora, considere que, após a atualização em  $T_4^*$ , os valores gerados para os vértices 0, 3 e 8 estão abaixo do limitante  $p$ , isto é,  $X_3 = 3$ . A Figura 4.4.a ilustra a configuração de  $T_4^*$ , após a terceira atualização. No

grafo  $T_5^+$ , aplicamos as regras do processo GMS. O vértice 3 é declarado inativo, pois o seu vértice correspondente em  $T_4^*$  é o vértice minimal, e adicionamos ao nó  $(3, 0.327, i)$  os vértices 0, 3 e 8, pois seus respectivos valores estão abaixo de  $p$ . Mas veja que o vértice 0 já estava ativo em  $T_5^+$ . Portanto, removemos o nó  $(0, 0.366, a)$  de  $T_5^+$ , antes da adição dos nós  $(0, 0.021, a)$ ,  $(3, 0.077, a)$  e  $(8, 0.191, a)$ . A Figura 4.4.b mostra a configuração de  $T_5^+$ , após a terceira atualização.

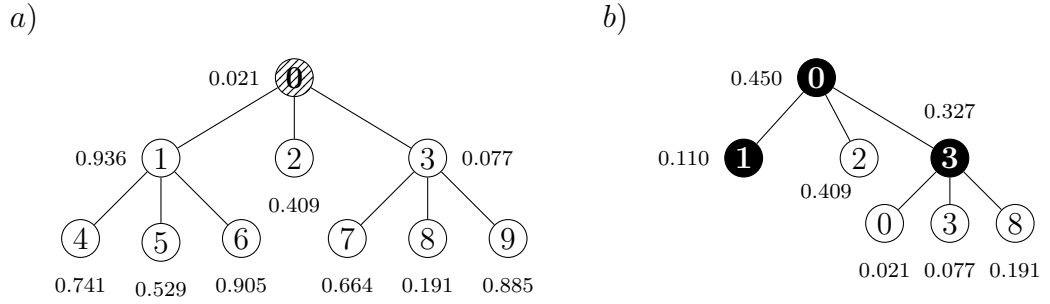


Figura 4.4: Grafos  $T_4^*$  (a) e  $T_5^+$  (b) ao final do instante de tempo  $n = 2$ , após a terceira atualização. Em  $T_4^*$ , o vértice 0 (em linhas diagonais) será o próximo vértice minimal. Em  $T_5^+$ , o vértice 3 (em preto) foi declarado inativo, isto é, seu estado é  $(0, 0.327, i)$ . Temos que  $A_3 = \{0, 2, 3, 8\}$  e  $|A_3| = 3 + 3 - 1 - 1 = 4$ .

No instante de tempo  $n = 3$ , o vértice 0 é o vértice minimal e tem três vizinhos (vértices 1, 2 e 3), ou seja,  $X_0 \sim bin(4, p)$ . Agora, considere que, após a atualização em  $T_4^*$ , apenas o valor gerado para o vértice 3 está abaixo do limitante  $p$ , isto é,  $X_0 = 1$ . A Figura 4.5.a ilustra a configuração de  $T_4^*$ , após a quarta atualização. No grafo  $T_5^+$ , aplicamos as regras do processo GMS. O vértice 0 é declarado inativo, pois o seu vértice correspondente em  $T_4^*$  é o vértice minimal, e adicionamos ao nó  $(0, 0.021, i)$  o vértice 3, pois seu respectivo valor está abaixo de  $p$ . Mas veja que os vértices 2 e 3 já estavam ativos em  $T_5^+$ . Portanto, removemos os nós  $(2, 0.409, a)$  e  $(3, 0.077, a)$  de  $T_5^+$ , antes do nó  $(3, 0.278, a)$  ser adicionado. A Figura 4.5.b mostra a configuração de  $T_5^+$ , após a quarta atualização.

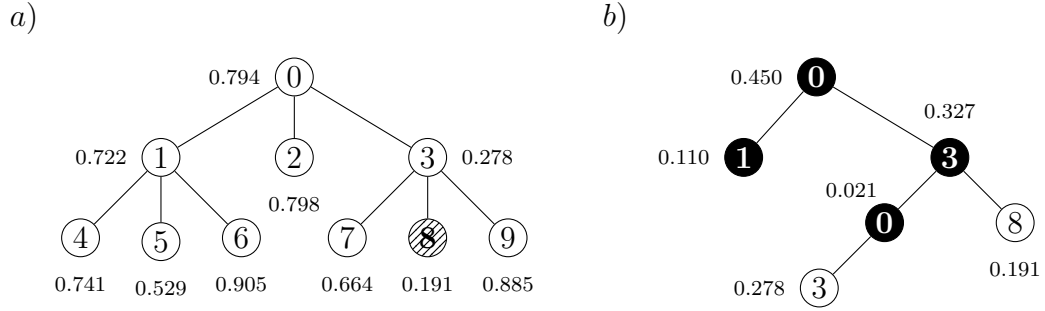


Figura 4.5: Grafos  $T_4^*$  (a) e  $T_5^+$  (b) ao final do instante de tempo  $n = 3$ , após a quarta atualização. Em  $T_4^*$ , o vértice 8 (em linhas diagonais) será o próximo vértice minimal. Em  $T_5^+$ , o vértice 0 (em preto) foi declarado inativo, isto é, seu estado é  $(0, 0.021, i)$ . Temos que  $A_4 = \{3, 8\}$  e  $|A_4| = 4 + 1 - 1 - 2 = 2$ .

Por fim, no instante de tempo  $n = 4$ , o vértice 8 é o vértice minimal e tem quatro vizinhos (vértices 3, 10, 11 e 12), ou seja,  $X_8 \sim bin(5, p)$ . Agora, considere que, após a atualização em  $T_4^*$ , nenhum vértice de  $\mathcal{N}(8) \cup \{8\}$  têm valor abaixo de  $p$ , isto é,  $X_8 = 0$ . A Figura 4.6.a ilustra a configuração de  $T_4^*$ , após a quinta atualização. Note também que todos os valores associados aos vértices estão acima do limitante, o que encerra a  $p$ -avalanche. No grafo  $T_5^+$ , aplicamos as regras do processo GMS. O vértice 8 é declarado inativo, pois o seu vértice correspondente em  $T_4^*$  é o vértice minimal. Como nenhum vértice  $\mathcal{N}(8) \cup \{8\}$  têm valor abaixo de  $p$ , nenhum nó é adicionado à  $T_5^+$ . Por outro lado, o nó  $(3, 0.278, a)$  é removido de  $T_5^+$ , pois é vizinho do vértice minimal. A Figura 4.6.b mostra a configuração de  $T_5^+$ , após a quinta atualização. Note que o número de vértices em  $T_5^+$  corresponde à duração do processo GMS.

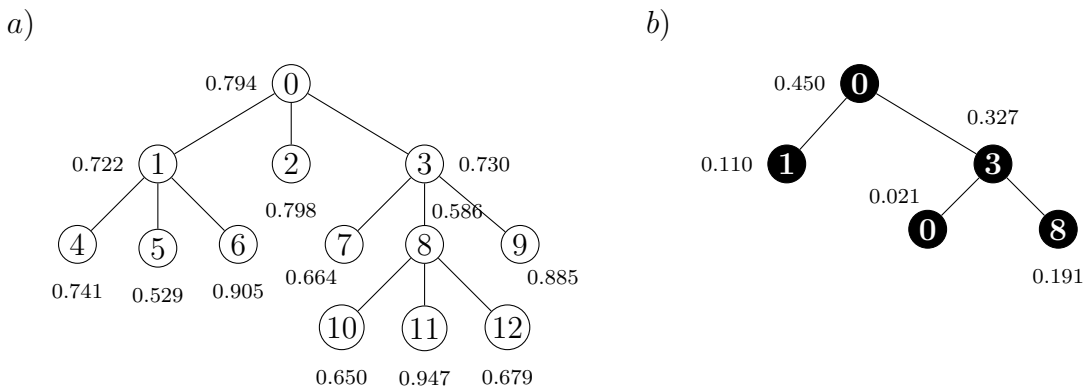


Figura 4.6: Grafos  $T_4^*$  (a) e  $T_5^+$  (b) ao final da  $p$ -avalanche, após a quinta atualização. Em  $T_4^*$ , todos os valores estão acima de  $p$ . Em  $T_5^+$ , o vértice 8 (em preto) foi declarado inativo, isto é, seu estado é  $(0, 0.191, i)$ . Temos que  $A_5 = \emptyset$  e  $|A_5| = 2 + 0 - 1 - 1 = 0$  e, portanto,  $D_{T_4^*}^p = 5$ .

## 4.3 A construção do acoplamento para a cota superior

Essa seção está dividida em quatro partes. Inicialmente, é dada uma intuição por trás do Lema 4.1.2. Então, seguimos com uma descrição detalhada do acoplamento, damos um exemplo e, por fim, verificamos a dominação estocástica obtida a partir do acoplamento.

### 4.3.1 Intuição

Um dos nossos objetivos é comparar a avalanche no modelo Bak-Sneppen com o aglomerado aberto na origem em um modelo de percolação em sítios (independente), com a condição de que a origem esteja aberta com probabilidade 1 (ao invés de  $p$ ), o que não tem efeito no ponto crítico. Normalmente, o modelo de percolação em sítios é estudado como uma estrutura aleatória estática, mas também é possível construir o aglomerado aberto na origem de forma dinâmica. Na Seção 1.2.2, vimos como é feita tal construção.

O crescimento de uma avalanche no modelo Bak-Sneppen e do aglomerado aberto na origem é conduzido pelos vértices *extremos*. Em uma avalanche no modelo Bak-Sneppen, os vértices extremos são aqueles que estão participando da avalanche (isto é, pertencem ao conjunto-alcance) e têm vizinhos fora da avalanche. O alcance da avalanche aumenta somente se um dos vértices extremos se torna o vértice minimal, em algum instante de tempo. No modelo de percolação em sítios, os vértices extremos são aqueles que têm um vizinho no aglomerado (a partir da origem), mas não se sabe se eles próprios estão abertos ou fechados. Estes são exatamente os vértices na borda do aglomerado, e eles irão aumentar o tamanho do aglomerado, se estiverem abertos. Como os vértices extremos são justamente os vértices que contribuem para o crescimento de ambos os processos, no sentido de aumentar a cardinalidade do conjunto-alcance (modelo Bak-Sneppen) e do aglomerado aberto (modelo percolação em sítios), a ideia é relacionar os dois conjuntos de vértices extremos entre si. O problema é que, no modelo Bak-Sneppen, os vértices não-extremos do conjunto-alcance também podem contribuir para aumentar o tamanho



do conjunto, enquanto que, no modelo de percolação em sítios, apenas vértices extremos contribuem para o aumento do aglomerado. Em outras palavras, no modelo Bak-Sneppen, um vértice extremo pode ser atualizado por um vizinho ativo antes de se tornar um vértice minimal, se tornando inativo após a atualização. Este é o caso em que um vértice extremo é atualizado por ser vizinho de um vértice minimal. Por outro lado, um vértice inativo pode ser tornar ativo e, com isso, caso este vértice seja um vértice extremo, um vértice não-extremo será atualizado, dando uma nova oportunidade para tal vértice participar da avalanche. Já no modelo de percolação em sítios, uma vez determinado o estado de um vértice (isto é, aberto ou fechado), o estado não pode mais ser alterado. Isso significa que não é interessante acoplar os dois modelos de forma natural, isto é, não é interessante determinar se um vértice no modelo de percolação é aberto ou fechado apenas pelo valor gerado por uma distribuição no modelo Bak-Sneppen sem memória. No modelo de percolação, o vizinho de um vértice extremo é sempre considerado, enquanto que, no modelo Bak-Sneppen, isso pode não ocorrer, já que é possível que um vizinho de um vértice extremo nunca seja atualizado.

As seguintes intuições tornam o Lema 4.1.2 razoável: Já sabemos que, no modelo Bak-Sneppen sem memória, se um vértice não é minimal, então sua nova distribuição é uniforme e acima do valor mínimo. Por outro lado, se um vértice é atualizado por ser um vizinho do vértice minimal, isso torna sua distribuição estocasticamente menor, o que torna o vértice mais propenso a se tornar ativo e portanto, intuitivamente, é mais provável que a avalanche não seja encerrada. Isso significa que, no modelo Bak-Sneppen, a interferência de vértices não-extremos no estado dos vértices extremos pode ser benéfica para a continuação da avalanche.

### 4.3.2 O processo acoplado

Agora, vamos descrever a construção de um processo que chamaremos de processo acoplado. Como veremos mais tarde, o processo é definido de forma a ser estocasticamente dominado por avalanches no modelo Bak-Sneppen, que é crucial para a demonstração do Lema 4.1.2. Na Seção 5, mostraremos que o processo acoplado, de fato, constrói um aglomerado a partir da origem do modelo de percolação em sítios. Em outras palavras,

mostraremos que as distribuições dos vértices no processo acoplado coincidem com as distribuições dos vértices no modelo de percolação em sítios (com a origem aberta com probabilidade 1). Seja  $V(G)$  o conjunto de vértices do grafo  $G$ . Dizemos que o *processo acoplado* é um processo estocástico que toma valores em

$$\{([0, 1] \times \{d, a, f\})^{|A|}, A \subset V(G)\}.$$

O estado  $(y, d)$  indica que  $y$  é limiar, isto é, o vértice está associado a uma distribuição  $U(y, 1)$ . Já os estados  $(y, a)$  e  $(y, f)$  indicam que o vértice tem um valor fixo  $y$  e é declarado aberto e fechado, respectivamente. O processo é encerrado quando  $A = \emptyset$ .

O nome do processo vem justamente do fato do processo ter sido criado com o intuito de ser acoplado às avalanches no modelo Bak-Sneppen sem memória.

Seja  $G$  um grafo conexo e localmente finito. Faremos duas cópias de  $G$ :  $G_C$  e  $G_B$ . Em  $G_C$ , executaremos o processo acoplado, enquanto que em  $G_B$ , será executado o modelo Bak-Sneppen sem memória. Também vamos definir o *conjunto extremal*  $\mathcal{E}$ , que é o conjunto dos vértices  $v_C \in V(G_C)$  que estão associados a uma distribuição. Em outras palavras, são os vértices de  $G_C$  com estado  $(y, d)$ , para algum  $y \in [0, 1]$ . Os processos serão executados de forma simultânea, de forma que a atualização de um vértice no modelo Bak-Sneppen sem memória implica na atualização do vértice correspondente no processo acoplado. É aqui que definimos o acoplamento entre o processo acoplado e o modelo Bak-Sneppen sem memória. As distribuições dos vértice  $v_C \in \mathcal{E}$  serão determinadas de acordo com as distribuições dos respectivos vértices correspondentes em  $v_B$ .

Considere uma  $p$ -avalanche em um instante de tempo  $n$  e seja  $r_n = R_{G,n}^p$  o número de vértices no conjunto-alcance no tempo  $n$ . Sejam  $U(y_i, 1)$  e  $U(z_i, 1)$  as distribuições dos vértices (no instante de tempo  $n$ ) em  $G_B$  e  $\mathcal{E}$ , respectivamente, para  $i = 0, \dots, r_n - 1$ . O objetivo é definir um acoplamento entre as distribuições dos vértices de  $G_B$  e  $G_C$ . Para isso, seja  $\{U_i\}_{i=0}^{r_n-1}$  uma sequência de variáveis aleatórias independentes  $U(0, 1)$  e defina a função

$$\begin{aligned} F_y : [0, 1] &\longrightarrow [y, 1] \\ u &\longmapsto (1 - y)u + y, \end{aligned}$$

para todo  $y \in [0, 1]$ . Agora, observe que, se  $U \sim U(0, 1)$ , então

$$P(F_y(U) \leq x) = P((1-y)U + y \leq x) = P\left(U \leq \frac{x-y}{1-y}\right) = \frac{x-y}{1-y}, \quad y \leq x < 1.$$

Ou seja,  $F_y(U) \sim U(y, 1)$ .

Portanto, sem perda de generalidade, podemos usar o par de variáveis aleatórias  $(F_{y_i}(U_i), F_{z_i}(U_i))$  para definir o acoplamento. Veja que as distribuições marginais coincidem com as distribuições dos vértices dos processos em questão, e os processos são governados pela mesma sequência de variáveis aleatórias,  $\{U_i\}_{i=0}^{r_n-1}$ .

No instante inicial, temos que  $r_0 = 1$ , pois a origem é o único vértice ativo na avalanche. Portanto, alteramos as distribuições da origem e dos vizinhos para  $U(0, 1)$ . No processo acoplado, o estado da origem é  $(0, a)$  e os vizinhos da origem têm estado  $(0, d)$ . Para determinar o estado dos vértices do grafo  $G_C$  no tempo  $n + 1$ , vamos realizar o seguinte procedimento: para cada  $v_{C,i} \in \mathcal{E}$ , considere o vértice correspondente no modelo Bak-Sneppen sem memória,  $v_{B,i} \in V(G_B)$ . Agora, realizamos uma amostra de  $\{U_i\}_{i=0}^{r_n}$  que vamos denotar por  $(u_0, \dots, u_{r_n-1})$ , e obtemos os valores dos vértices em  $G_B$ . Então, executamos o modelo Bak-Sneppen sem memória em  $G_B$  e localizamos o vértice minimal. Seja  $m$  o valor do vértice minimal.

1. Se  $v_{B,i}$  não é o vértice minimal, então  $F_{y_i}(U_i) > m$ , ou seja,  $U_i > \frac{m-y_i}{1-y_i}$ .
  - Se o estado do vértice correspondente  $v_{C,i}$  é da forma  $(z_i, d)$ , então usamos essa informação para obter a distribuição de  $v_{C,i}$ :

$$\begin{aligned} P(F_{z_i}(U_i) \leq x \mid F_{y_i}(U_i) > m) &= \frac{P\left(\frac{m-y_i}{1-y_i} < U_i \leq \frac{x-z_i}{1-z_i}\right)}{P\left(U_i > \frac{m-y_i}{1-y_i}\right)} \\ &= \frac{\frac{x-z_i}{1-z_i} - \frac{(m-y_i)^+}{1-y_i}}{1 - \frac{(m-y_i)^+}{1-y_i}} \\ &= \frac{x - \left(z_i + (1-z_i)\frac{(m-y_i)^+}{1-y_i}\right)}{1 - \left(z_i + (1-z_i)\frac{(m-y_i)^+}{1-y_i}\right)} \\ &= P(F_{z_i}(U_i) \leq x), \end{aligned}$$

onde  $\hat{z}_i = z_i + (1 - z_i) \frac{(m - y_i)^+}{1 - y_i}$  e  $(m - y_i)^+ = \max(m - y_i, 0)$ . Portanto, o estado do vértice  $v_{C,i}$  será  $(\hat{z}_i, d)$ .

- Se o estado do vértice correspondente  $v_{C,i}$  é da forma  $(z_i, a)$  ou da forma  $(z_i, f)$ , significa que o vértice já está com um valor fixado e, portanto, nada é feito.

2. Se  $v_{B,i}$  é o vértice minimal:

- Se o estado do vértice correspondente  $v_{C,i}$  é da forma  $(z_i, d)$ , então  $F_{y_i}(U_i) = m$ , ou seja,  $U_i = \frac{m - y_i}{1 - y_i}$ . O vértice  $v_{C,i}$  é removido do conjunto  $\mathcal{E}$  e seu valor será  $\hat{z}_i = F_{z_i} \left( \frac{m - y_i}{1 - y_i} \right)$ . Se  $\hat{z}_i \leq p$ , o vértice é declarado aberto, isto é, alteramos o estado de  $v_{C,i}$  para  $(\hat{z}_i, a)$  e adicionamos os vizinhos de  $v_{C,i}$  que ainda não tem o estado definido ao conjunto  $\mathcal{E}$ . O estado destes vértices será  $(0, d)$ . Se  $\hat{z}_i > p$ , alteramos o seu estado para  $(\hat{z}_i, f)$ . Em palavras, o vértice é declarado fechado.
- Se o estado do vértice correspondente  $v_{C,i}$  é da forma  $(z_i, a)$  ou da forma  $(z_i, f)$ , significa que o vértice já está com um valor fixado e, portanto, nada é feito.

Simultaneamente, o modelo Bak-Sneppen sem memória é executado em  $G_B$  através do procedimento descrito na Seção 2.3. O procedimento termina quando o valor do vértice minimal de  $G_B$  é maior que  $p$ , isto é,  $m > p$ . Consequentemente,  $F_{y_i}(U_i) \geq m, \forall i$ . Usamos essa informação para alterar a distribuição de todos os vértices em  $\mathcal{E}$ , como é descrito no caso 1. Em relação ao vértice minimal de  $G_B$ , o procedimento é semelhante ao procedimento descrito no caso 2, e  $v_{C,i}$  será declarado fechado, pois  $\hat{z}_i = z_i + (1 - z_i) \frac{(m - y_i)^+}{1 - y_i} > p$ , já que  $m > p$ . Após a realização dos procedimentos, removemos todos os vértices do conjunto  $\mathcal{E}$  e geramos os valores para estes vértices a partir de suas respectivas distribuições, dadas pelas variáveis aleatórias  $F_{\hat{z}_i}(U_i)$ . Por fim, declaramos os vértices fechados e alteramos seus respectivos estados. Na Seção 4.3.4, veremos que se a  $p$ -avalanche termina, então todos os vértices do processo acoplado são removidos de  $\mathcal{E}$  e declarados fechados.

### 4.3.3 Um exemplo

O comportamento dos processos é ilustrado pelo seguinte exemplo, apresentado nas figuras abaixo. Seja  $G = \mathbb{Z}^2$ . Por fins de ilustração, mostraremos apenas a parte do grafo onde ocorre a dinâmica. Assim como foi descrito na construção do acoplamento, vamos definir duas cópias de  $G$ :  $G_B$  (figuras à esquerda) e  $G_C$  (figuras à direita). Em  $G_B$ , as variáveis aleatórias associadas aos vértices serão denotadas por  $F_y$ , ao invés de  $F_y(U)$ , por fins de praticidade, e os vértices preenchidos com linhas diagonais são os vértices minimais. Em  $G_C$ , os vértices em branco são os vértices declarados abertos, os vértices em cinza são os vértices do conjunto extremal e os vértices em preto são os vértices declarados fechados. Considere uma  $p$ -avalanche, com  $p = 0.450$  no modelo Bak-Sneppen sem memória em  $G_B$ . A Figura 4.7 mostra a configuração dos grafos  $G_B$  e  $G_C$ , respectivamente, no instante inicial, antes do início da  $p$ -avalanche. Por definição, a origem em  $G_C$  é declarada aberta.

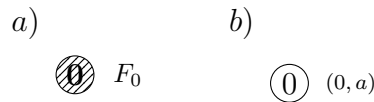


Figura 4.7: Grafos  $G_B$  (a) e  $G_C$  (b) no instante inicial, antes do início da  $p$ -avalanche. Em  $G_B$ ,  $r_0 = 1$ .

Ainda no instante inicial ( $n = 0$ ), a origem é o vértice minimal e, portanto, atualizaremos as distribuições da origem e dos vizinhos para  $U(0, 1)$ , de forma independente (Figura 4.8.a). Enquanto isso, em  $G_C$ , uma vez que o estado da origem é  $(0, a)$  (por definição), seus vizinhos têm estado  $(0, d)$  (Figura 4.8.b).

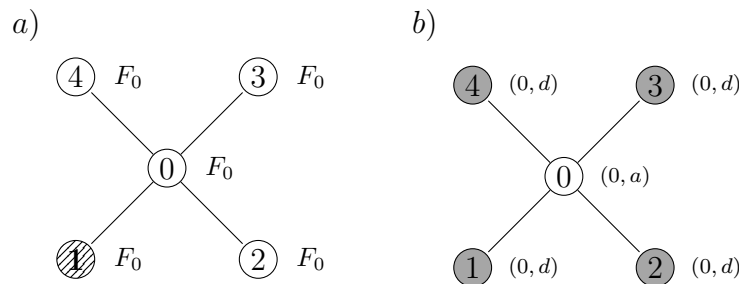


Figura 4.8: Grafos  $G_B$  (a) e  $G_C$  (b) ao final do instante de tempo  $n = 0$  e no início do instante de tempo  $n = 1$ . Em  $G_B$ ,  $r_1 = 5$ . Em  $G_C$ , os vértices cinza são os vértices do conjunto extremal. Portanto,  $\mathcal{E} = \{1, 2, 3, 4\}$ . Após a realização das distribuições, o vértice 1 será o próximo vértice minimal.

No instante de tempo  $n = 1$  (Figura 4.9), realizamos as distribuições em  $G_B$  e localizamos o vértice minimal. Suponha que o vértice minimal é o vértice 1 e seu valor é  $m = 0.110$ . Em  $G_C$ , alteramos os estados dos vértices de  $\mathcal{E}$  de acordo com as distribuições obtidas pelos vértices de  $G_B$  no instante de tempo  $n = 0$ , usando a descrição feita acima. Temos que

$$\hat{z}_i = 0 + (1 - 0) \frac{(0.110 - 0)^+}{1 - 0} = 0.110, \quad i = 1, 2, 3, 4.$$

Como  $0.110 < p$ , o vértice 1 terá estado  $(0.110, a)$  e os vértices 2, 3 e 4 terão estado  $(0.110, d)$ . O vértice 0 já possui valor fixo e, portanto, não terá seu estado alterado. Em  $G_B$ , realizamos o procedimento descrito na Seção 2.3.

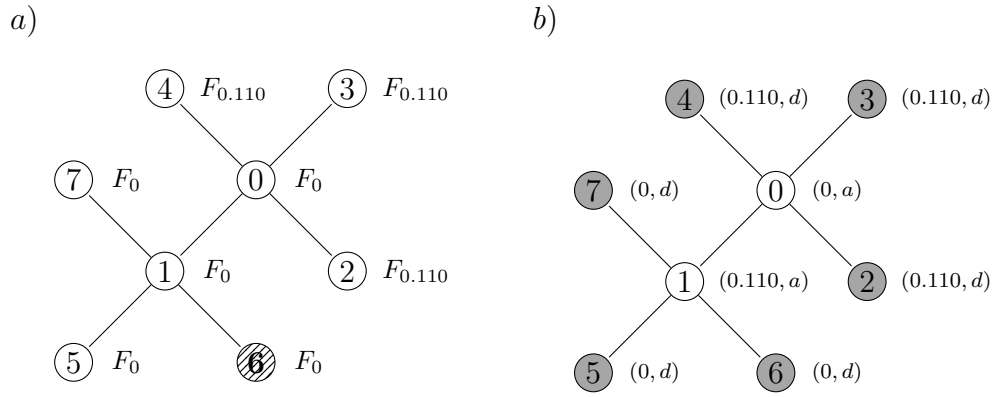


Figura 4.9: Grafos  $G_B$  (a) e  $G_C$  (b) ao final do instante de tempo  $n = 1$  e no início do instante de tempo  $n = 2$ . Em  $G_C$ , observe que o vértice 1 saiu do conjunto extremal, já que agora tem um valor fixo. Como o vértice 1 foi declarado aberto (pois  $0.110 < p$ ) os vértices 5, 6 e 7 entram em  $\mathcal{E}$  com estado  $(0, d)$ . Portanto,  $\mathcal{E} = \{2, 3, 4, 5, 6, 7\}$ . Em  $G_B$ ,  $r_2 = 8$ . Após a realização das distribuições, o vértice 6 será o próximo vértice minimal.

No instante de tempo  $n = 2$  (Figura 4.10), realizamos as distribuições em  $G_B$  e localizamos o vértice minimal. Suponha que o vértice minimal é o vértice 6 e seu valor é  $m = 0.076$ . Em  $G_C$ , alteramos os estados dos vértices de  $\mathcal{E}$  de acordo com as distribuições obtidas pelos vértices de  $G_B$  no instante de tempo  $n = 1$ , usando a descrição do acoplamento. Temos que

$$\hat{z}_i = 0.110 + (1 - 0.110) \frac{(0.076 - 0.110)^+}{1 - 0.110} = 0.110, \quad i = 2, 3, 4;$$

$$\hat{z}_i = 0 + (1 - 0) \frac{(0.076 - 0)^+}{1 - 0} = 0.076, \quad i = 5, 7.$$

Veja que os estados dos vértices 2, 3 e 4 não sofreram alterações. Isso ocorreu porque o mínimo no instante atual é menor que o limiar das suas respectivas distribuições. Como  $0.076 < p$ , o vértice 6 é declarado aberto e terá estado  $(0.076, a)$ . Já os vértices 2, 3 e 4 terão estado  $(0.110, d)$ . Os vértices 0 e 1 já possuem valor fixo e, portanto, não terão os seus estados alterados.

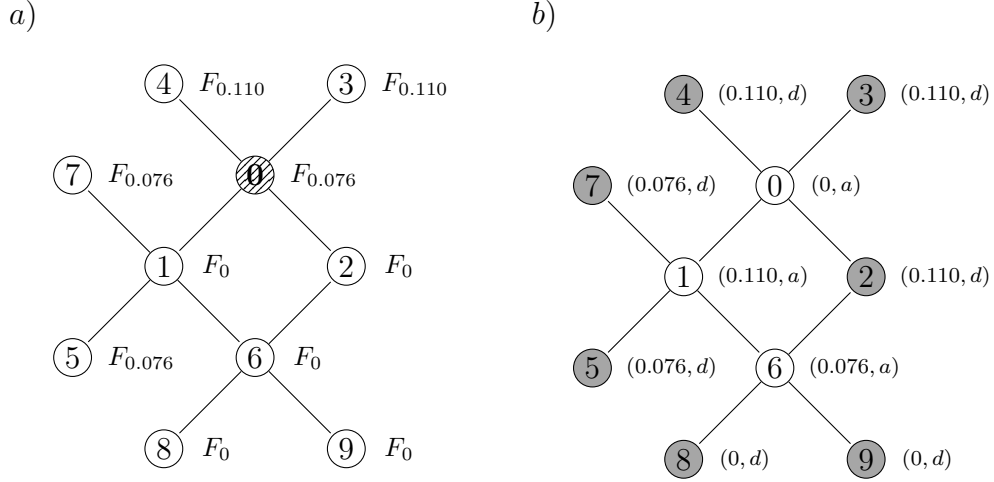


Figura 4.10: Grafos  $G_B$  (a) e  $G_C$  (b) ao final do instante  $n = 2$  e no início do instante de tempo  $n = 3$ . Em  $G_C$ , observe que o vértice 6 saiu do conjunto extremal, já que tem um valor fixo. Como o vértice 6 foi declarado aberto (pois  $0.076 < p$ ) os vértices 8 e 9 entram em  $\mathcal{E}$  com estado  $(0, d)$ . Portanto,  $\mathcal{E} = \{2, 3, 4, 5, 7, 8, 9\}$ . Em  $G_B$ ,  $r_3 = 10$ . Após a realização das distribuições, o vértice 0 será o próximo vértice minimal.

No instante de tempo  $n = 3$  (Figura 4.11), realizamos as distribuições em  $G_B$  e localizamos o vértice minimal. Suponha que o vértice minimal é o vértice 0 e seu valor é  $m = 0.417$ . Em  $G_C$ , alteramos os estados dos vértices de  $\mathcal{E}$  de acordo com as distribuições obtidas pelos vértices de  $G_B$  no instante de tempo  $n = 2$ , usando a descrição do acoplamento. Temos que

$$\hat{z}_2 = 0.110 + (1 - 0.110) \frac{(0.417 - 0)^+}{1 - 0} = 0.481;$$

$$\hat{z}_i = 0.110 + (1 - 0.110) \frac{(0.417 - 0.110)^+}{1 - 0.110} = 0.417, \quad i = 3, 4;$$

$$\hat{z}_i = 0.076 + (1 - 0.076) \frac{(0.417 - 0.076)^+}{1 - 0.076} = 0.417, \quad i = 5, 7;$$

$$\hat{z}_i = 0 + (1 - 0) \frac{(0.417 - 0)^+}{1 - 0} = 0.417, \quad i = 8, 9.$$

Os vértices 0,1 e 6 já possuem valor fixo e, portanto, não terão seus estados alterados. Observe que o vértice 0 de  $G_C$  é um vértice que está fora do conjunto extremal. Como consequência, as distribuições de todos os vértices do conjunto extremal serão estocasticamente maiores em relação ao instante de tempo anterior. E isso aumenta as chances destes vértices serem eventualmente fechados, o que sugere que os vértices que não estão no conjunto extremal não são benéficos para o crescimento do processo acoplado. Por outro lado, em  $G_B$ , o fato do vértice minimal ter sido o vértice 0, que é um vértice do interior do grafo  $G_B$ , pode contribuir para o aumento do conjunto-alcance, já que as distribuições dos vértices 2,3 e 4, que são vértice extremos, foram atualizadas para  $U(0, 1)$ , o que aumenta as chances de um destes vértices se tornarem vértices minimais. Isso mostra que os vértices do interior de  $G_B$  podem contribuir para o crescimento da  $p$ -avalanche.

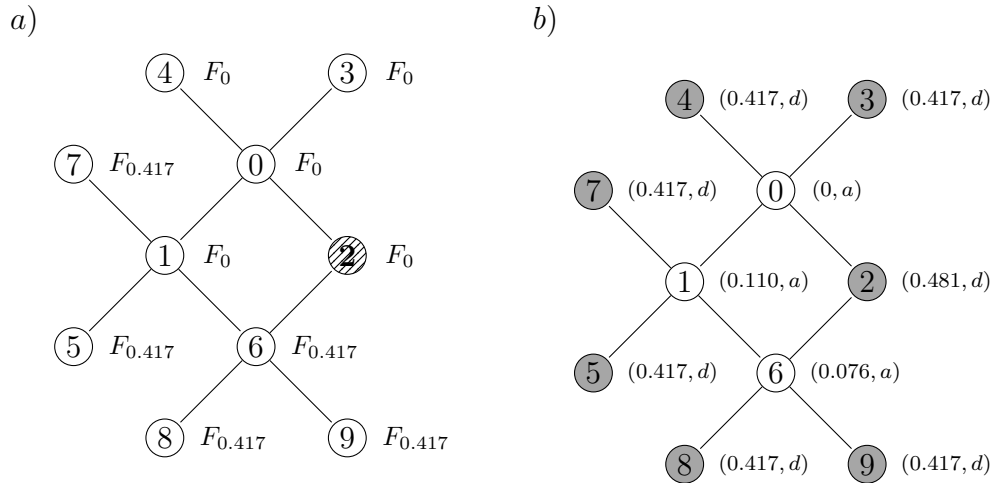


Figura 4.11: Grafos  $G_B$  (a) e  $G_C$  (b) ao final do instante  $n = 3$  e no início do instante de tempo  $n = 4$ . Em  $G_C$ , não temos alterações no conjunto extremal, uma vez que o vértice correspondente ao vértice minimal está fora de  $\mathcal{E}$  (vértice 0). Portanto,  $\mathcal{E} = \{2, 3, 4, 5, 7, 8, 9\}$ . Em  $G_B$ ,  $r_4 = 10$  (também sem alterações). Após a realização das distribuições, o vértice 2 será o próximo vértice minimal.

No instante de tempo  $n = 4$  (Figura 4.12), realizamos as distribuições em  $G_B$  e localizamos o vértice minimal. Suponha que o vértice minimal é o vértice 2 e seu valor é  $m = 0.231$ . Em  $G_C$ , alteramos os estados dos vértices de  $\mathcal{E}$  de acordo com as distri-



buições obtidas pelos vértices de  $G_B$  no instante de tempo  $n = 3$ , usando a descrição do acoplamento. Temos que

$$\hat{z}_2 = 0.481 + (1 - 0.481) \frac{(0.231 - 0)^+}{1 - 0} = 0.601;$$

$$\hat{z}_i = 0.417 + (1 - 0.417) \frac{(0.231 - 0)^+}{1 - 0} = 0.552; \quad i = 3, 4;$$

$$\hat{z}_i = 0.417 + (1 - 0.417) \frac{(0.231 - 0.417)^+}{1 - 0.417} = 0.417; \quad i = 5, 7, 8, 9.$$

Assim como no instante de tempo anterior, os vértices 0, 1 e 6 já possuem valor fixo e, portanto, não terão seus estados alterados pelo procedimento. Agora, observe os vértices do conjunto extremal. Como  $0.601 > p$ , o vértice 2 será declarado fechado e seu estado será  $(0.601, f)$ . Os estados dos vértice 5, 7, 8 e 9 não foram alterados pelo procedimento. Já os vértices 3 e 4 agora têm estado  $(0.552, d)$ . Veja que  $0.552 > p$ , o que mostra que os vértices serão fechados, caso o respectivo correspondente em  $G_B$  se torne o vértice minimal. Na Seção 4.3.4 veremos um exemplo em que todos os vértices do conjunto extremal serão fechados, mas a  $p$ -avalanche em  $G_B$  não necessariamente será encerrada. E isso implica que processo acoplado será finito, mas a avalanche pode ser infinita. Em  $G_B$ , realizamos o procedimento descrito na Seção 2.3. O fato do vértice 0 ter sido o vértice minimal no instante de tempo  $n = 3$  contribuiu para o alcance da  $p$ -avalanche, pois isso tornou a aptidão do vértice 2 estocasticamente menor, o que aumentou as chances do vértice 2 se tornar o vértice minimal no instante de tempo  $n = 4$ .

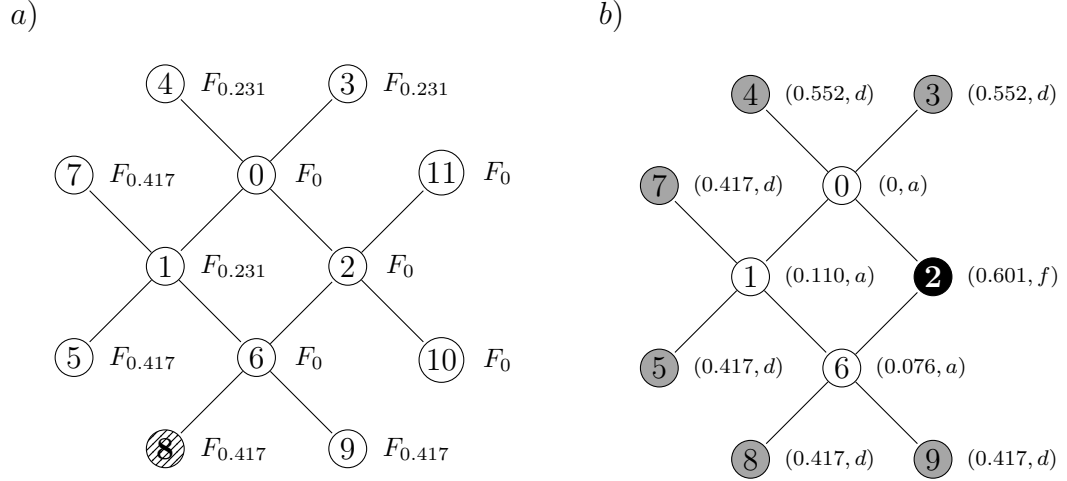


Figura 4.12: Grafos  $G_B$  (a) e  $G_C$  (b) ao final do instante  $n = 4$  e no início do instante de tempo  $n = 5$ . Em  $G_C$ , o conjunto extremal tem um elemento a menos, já que o vértice 2 foi removido. Portanto,  $\mathcal{E} = \{3, 4, 5, 7, 8, 9\}$ . Em  $G_B$ , o conjunto-alcance aumentou e  $r_5 = 12$ . Após a realização das distribuições, o vértice 8 será o próximo vértice minimal.

No instante de tempo  $n = 5$  (Figura 4.13), vamos mostrar o que acontece quando a  $p$ -avalanche é encerrada. Após a realização das distribuições em  $G_B$ , localizamos o vértice minimal. Suponha que o vértice minimal é o vértice 8 e seu valor é  $m = 0.503$ . Veja que  $0.503 > p$ , o que encerra a avalanche. Em  $G_C$ , alteramos os estados dos vértices de  $\mathcal{E}$  de acordo com as distribuições obtidas pelos vértices de  $G_B$  no instante de tempo  $n = 4$ , usando a descrição do acoplamento. Temos que

$$\hat{z}_i = 0.552 + (1 - 0.552) \frac{(0.503 - 0.231)^+}{1 - 0.231} = 0.710; \quad i = 3, 4;$$

$$\hat{z}_i = 0.417 + (1 - 0.417) \frac{(0.503 - 0.417)^+}{1 - 0.417} = 0.503; \quad i = 5, 7, 8, 9.$$

O vértice 8 é declarado fechado e seu estado é  $(0.503, f)$ , pois é o vértice correspondente ao vértice minimal em  $G_B$  e  $0.503 > p$ . Os vértices 0, 1, 2 e 6 já possuem valor fixo e, portanto, não terão seus estados alterados pelo procedimento. Em relação aos demais vértices do conjunto extremal, o estado de cada um é da forma  $(\hat{z}_i, d)$ . Em  $G_B$ , realizamos o passo 3 do procedimento descrito na Seção 2.3 e obtemos as distribuições dos vértices.

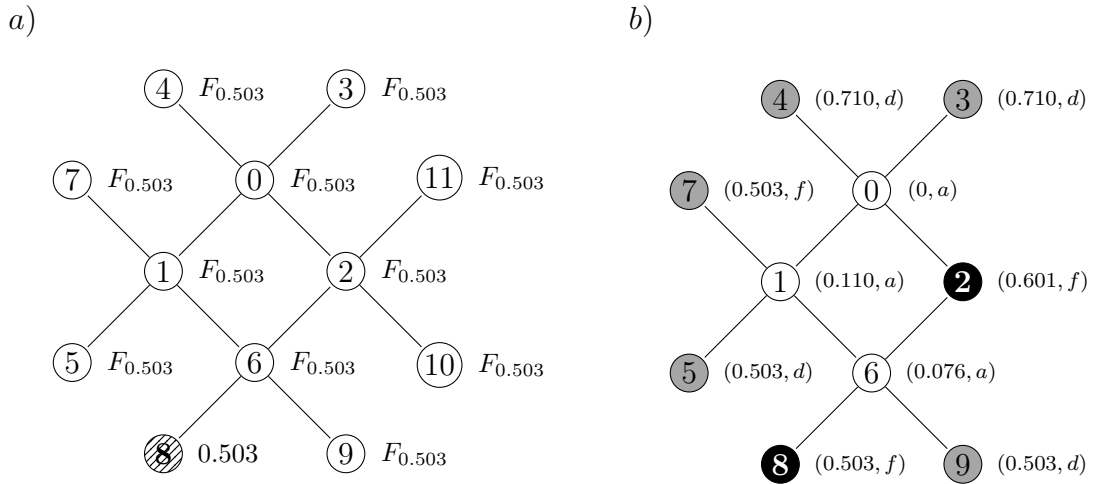


Figura 4.13: Grafos  $G_B$  (a) e  $G_C$  (b) no instante de tempo  $n = 5$ , antes de realizarmos as distribuições e declarar os estados dos vértices. Em  $G_C$ , O vértice 8 (em preto) já foi removido de  $\mathcal{E}$ . Em  $G_B$ ,  $r_5 = 12$ .

Finalmente, removemos os vértices restantes de  $\mathcal{E}$  e, a partir de suas respectivas distribuições, obtemos valores que serão fixados nos vértices que, por sua vez, serão declarados fechados. A configuração final de  $G_C$  pode ser observada na Figura 4.14.b.

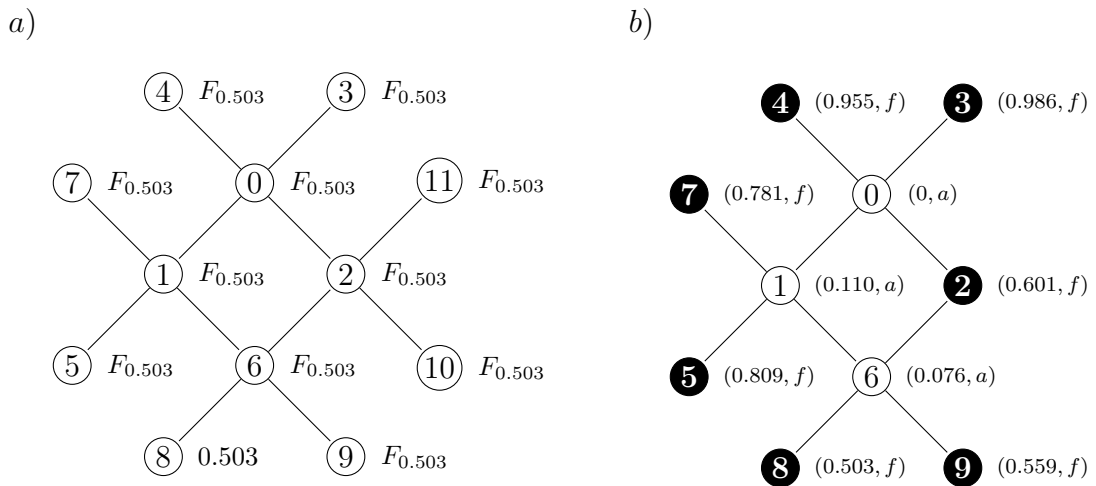


Figura 4.14: Grafos  $G_B$  (a) e  $G_C$  (b) ao final do instante de tempo  $n = 5$ . Em  $G_B$ , a avalanche foi encerrada. Em  $G_C$ , o processo acoplado também foi encerrado, e temos um aglomerado finito a partir da origem, dado por  $C = \{0, 1, 6\}$ .

### 4.3.4 A dominação estocástica entre os processos

Vamos enunciar e demonstrar duas proposições que mostram a dominação estocástica entre as avalanches no modelo Bak-Sneppen e o processo acoplado.

**Proposição 4.3.1.** *Para todo vértice  $v_C \in \mathcal{E}$  e seu vértice correspondente,  $v_B \in G_B$ , em qualquer instante de tempo, a distribuição de  $v_C$  é estocasticamente maior que a distribuição de  $v_B$ .*

*Demonstração.* A prova segue por indução. Seja  $k$  o instante de tempo em que um vértice de  $G_C$  é adicionado ao conjunto  $\mathcal{E}$ . Se o vértice  $v_C$  acabou de entrar no conjunto extremal, então o vértice correspondente  $v_B$  é vizinho do vértice minimal e, portanto, tem distribuição  $U(0, 1)$ . E pela definição do processo acoplado, a distribuição de  $v_C$  também é  $U(0, 1)$ . Portanto, vale a afirmação do teorema.

Agora, suponha que  $v_C \in \mathcal{E}$  no instante de tempo  $k+n$  e que a proposição também é válida em  $k+n$ . Ainda no instante de tempo  $k+n$ , sejam  $U$  uma variável aleatória  $U(0, 1)$ ,  $F_y(U)$  e  $F_z(U)$  as variáveis aleatórias com as distribuições de  $v_B$  e  $v_C$ , respectivamente e  $m$  o valor do vértice minimal. Pela hipótese de indução, vale a afirmação. Isto é,  $\forall x$ ,

$$P(F_z(U) > x) \geq P(F_y(U) > x).$$

Logo,

$$1 - \frac{x - z}{1 - z} \geq 1 - \frac{x - y}{1 - y},$$

ou, equivalentemente,

$$z \geq y.$$

Vamos mostrar que a proposição vale para o instante de tempo  $k+n+1$ . Suponha que  $y < 1$ . Primeira, vamos assumir que  $v_B$  não é o vértice minimal no instante de tempo  $k+n$ . Então a distribuição no instante de tempo  $k+n+1$  de  $v_B$  será  $U(y \vee m, 1)$ . Já a distribuição do vértice correspondente em  $\mathcal{E}$ ,  $v_C$ , será  $U(\hat{z}, 1)$ , onde  $\hat{z} = z + (1-z) \frac{(m-y)^+}{1-y}$ . Temos que mostrar que  $P(F_{y \vee m}(U) > x) \leq P(F_{\hat{z}}(U) > x)$ , ou, equivalentemente, que  $(y \vee m) \leq \hat{z}$ . Neste caso, temos duas possibilidades:

- Se  $(y \vee m) = y$ : Então  $(m-y)^+ = 0$ , o que implica que  $\hat{z} = z$ , ou seja  $\hat{z} \geq y$ .

- Se  $(y \vee m) = m$ : Então  $(m - y)^+ = m - y$ , o que implica que

$$\begin{aligned}
\hat{z} &= z + (1 - z) \frac{(m - y)^+}{1 - y} \\
&= \frac{z - zy + m - y - mz + zy}{1 - y} \\
&= \frac{z - zy + m - y - mz + zy - my + my}{1 - y} \\
&= \frac{m(1 - y) + z - y + my - mz - zy + zy}{1 - y} \\
&= m + \frac{(1 - m)(z - y)}{1 - y} \geq m.
\end{aligned}$$

Portanto,  $(y \vee m) \leq \hat{z}$ . Agora, vamos assumir que  $v_B$  é o vértice minimal no instante de tempo  $k + n$ . Então, no instante de tempo  $k + n + 1$ , o vértice correspondente,  $v_C$ , é removido de  $\mathcal{E}$ , e não há nada para provar, o que conclui a prova da proposição.

□

**Proposição 4.3.2.** *Se o alcance de uma  $p$ -avalanche em  $G_B$  é finito com probabilidade 1, então a probabilidade de haver um aglomerado infinito no processo acoplado é zero.*

*Demonstração.* Com efeito, se a a probabilidade de uma  $p$ -avalanche é zero, então os valores de todos os vértices em  $G_B$  são maiores que  $p$ . Pela Proposição 4.3.1, todos os vértices correspondentes em  $\mathcal{E}$  têm distribuições estocasticamente maiores, o que implica que eles serão declarados fechados e removidos do conjunto extremal. Portanto, a probabilidade de um aglomerado infinito no processo acoplado é zero, o que encerra a prova da proposição.

□

Para encerrar a seção, vamos dar um exemplo que mostra que o processo acoplado é finito, mas a  $p$ -avalanche pode ser infinita. E isso mostra que a dominação estocástica não é uma igualdade estocástica. Seja  $G = \mathbb{Z}$  e seja  $p = 0.7$ . Por fins de ilustração, mostraremos apenas a parte do grafo onde ocorre a dinâmica. A configuração inicial é idêntica à configuração apresentada na Figura 4.7. Ao final instante inicial, temos a seguinte configuração:

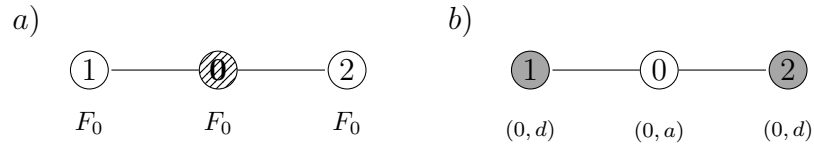


Figura 4.15: Grafos  $G_B$  (a) e  $G_C$  (b) ao final do instante inicial e no início do instante de tempo  $n = 1$ . Após a realização das distribuições, a origem será o próximo vértice minimal.

No instante de tempo  $n = 1$  (Figura 4.16), suponha que o vértice minimal é a origem, com  $m = 0.5$ .

$$\hat{z}_i = 0 + (1 - 0) \frac{(0.5 - 0)^+}{1 - 0} = 0.5; \quad i = 1, 2.$$

Observe que não temos alterações em  $G_B$ . Em  $G_C$  as distribuições dos vértices de  $\mathcal{E}$  (em cinza) são estocasticamente maiores, se comparadas às distribuições dos próprios vértices no instante de tempo anterior.

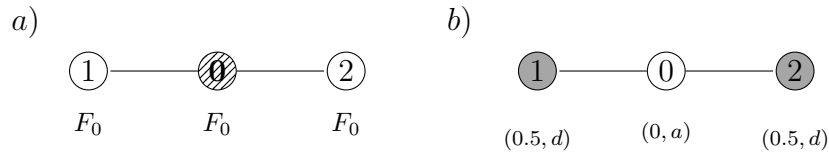


Figura 4.16: Grafos  $G_B$  (a) e  $G_C$  (b) ao final do instante tempo  $n = 1$  e no início do instante de tempo  $n = 2$ . Novamente, a origem será o próximo vértice minimal.

No instante de tempo  $n = 2$  (Figura 4.17), suponha novamente que o vértice minimal é a origem, com  $m = 0.5$ . Veja o que acontece em  $G_C$ :

$$\hat{z}_i = 0.5 + (1 - 0.5) \frac{(0.5 - 0)^+}{1 - 0} = 0.75; \quad i = 1, 2.$$

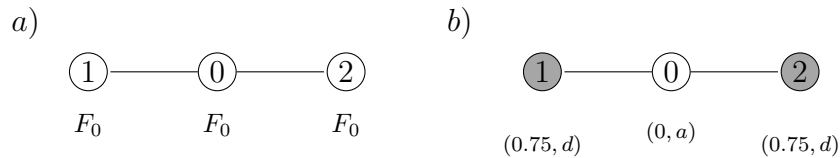


Figura 4.17: Grafos  $G_B$  (a) e  $G_C$  (b) ao final do instante tempo  $n = 2$  e no início do instante de tempo  $n = 3$ .

Como  $0.75 > p$ , os vizinhos da origem eventualmente serão fechados. Para isso, basta que um vértice correspondente de um dos vizinhos da origem seja minimal em  $G_B$ . Portanto, o aglomerado no processo acoplado é finito, mas a avalanche no modelo Bak-Sneppen pode ser infinita, já que ainda não foi encerrada.

## 4.4 Prova do Lema 4.1.2

Para completar a prova do Lema 4.1.2, nos resta mostrar que o processo acoplado, de fato, constrói um aglomerado aberto na origem no modelo de percolação em sítios, com a ressalva que a origem é aberta com probabilidade 1. Para facilitar o entendimento da prova, primeiro damos um exemplo. Ao mesmo tempo, o exemplo mostra a construção do processo acoplado.

### 4.4.1 Um exemplo

Considere a avalanche no modelo Bak-Sneppen, com o limitante igual a  $p$ , e o processo acoplado em  $G = \mathbb{Z}$ , com parâmetro  $p$ . Queremos calcular a probabilidade de que, no processo acoplado, ambos os vizinhos da origem estejam fechados. Note que, para o modelo de percolação em sítios, essa probabilidade é  $(1 - p)^2$  e, portanto, nosso objetivo é mostrar que o mesmo vale para o processo acoplado. Para calcular essa probabilidade, introduzimos a função  $g_p$  e os eventos  $A$  e  $B_x$ :

$$\begin{aligned} g_p : [0, p] &\longrightarrow [0, 1] \\ x &\longmapsto P(A \mid B_x) \end{aligned}$$

onde  $A$  é o evento {Os vizinhos da origem são declarados fechados} e  $B_x$  é o evento {O estado atual dos vizinhos é  $(x, d)$ }. Ou seja,  $g_p(x)$  é a probabilidade dos vizinhos da origem serem declarados fechados, dado que o estado atual dos vizinhos é  $(x, d)$ . Estamos interessados em calcular  $g_p(0)$ .

Começando com o estado  $(x, d)$  para os dois vizinhos, vamos denominar esse passo como o passo inicial. No processo acoplado, os vizinhos da origem são declarados fechados

se seus respectivos valores estão acima de  $p$ . E neste caso, há duas saídas para que os vizinhos da origem sejam declarados fechados (vale lembrar que os dois vizinhos da origem têm distribuição  $U(x, 1)$ , já que o estado de ambos é  $(x, d)$ ):

1. Se, no passo inicial, os valores dos 3 vértices (origem + vizinhos) no modelo Bak-Sneppen são maiores que  $(p-x)/(1-x)$ : se isso acontecer e o vértice minimal não for a origem, então o vértice correspondente ao vértice minimal no modelo Bak-Sneppen terá seu valor fixado em um valor maior que  $x + (1-x) \frac{\frac{p-x}{1-x} - 0}{1-0} = p$ . Portanto, esse vértice será declarado fechado, e os demais vértices serão fechados no passo final. Se o vértice minimal for a origem, então os vizinhos da origem no processo acoplado terão um estado com distribuição acima do mínimo, o que implica que, eventualmente, eles serão declarados fechados.
2. Se o vértice minimal no modelo Bak-Sneppen é a origem e seu valor está abaixo de  $(p-x)/(1-x)$ : neste caso, os vizinhos da origem no processo acoplado não serão declarados fechados imediatamente, já que o estado da origem no processo acoplado é definido com  $(0, a)$  e, portanto, não pode ser alterado. Assim, temos que verificar as atualizações seguintes. As distribuições dos 3 vértices no modelo Bak-Sneppen sem memória serão atualizadas para  $U(0, 1)$ , usando as variáveis aleatórias  $U_0, U_1$  e  $U_2$ . Por outro lado, no processo acoplado, o estado dos vizinhos será  $(x + (1-x)b, d)$ , onde  $b$  é o valor associado a origem no passo inicial. Para o segundo passo, estamos em um situação similar ao passo inicial, exceto que agora o estado dos vizinhos é  $(x + (1-x)b, d)$ , ao invés de  $(x, d)$ .

Em qualquer outro caso, os vizinhos da origem não serão declarados fechados. Com efeito, se, no modelo Bak-Sneppen, o vértice minimal for um dos vizinhos da origem e seu valor for menor que  $(p-x)/(1-x)$ , o vértice correspondente no processo acoplado será declarado aberto, pois seu valor será menor que  $p$ , que é justamente o que não queremos. Segue que

$$g_p(x) := P(A \mid B_x) = P\left(A \cap \left\{M > \frac{p-x}{1-x}\right\} \mid B_x\right) + P\left(A \cap \left\{M \leq \frac{p-x}{1-x}\right\} \mid B_x\right),$$



onde  $M = \min\{U_0, U_1, U_2\}$ .

Assim,

$$g_p(x) = P\left(A \mid \left\{M > \frac{p-x}{1-x}\right\} \cap B_x\right) P\left(M > \frac{p-x}{1-x} \mid B_x\right) + P\left(A \cap \left\{M \leq \frac{p-x}{1-x}\right\} \mid B_x\right).$$

Note que os eventos  $\left\{M > \frac{p-x}{1-x}\right\}$  e  $B_x$  são independentes. Portanto,

$$g_p(x) = P\left(A \mid \left\{M > \frac{p-x}{1-x}\right\} \cap B_x\right) P\left(M > \frac{p-x}{1-x}\right) + P\left(A \cap \left\{M \leq \frac{p-x}{1-x}\right\} \mid B_x\right). \quad (4.2)$$

Na primeira parcela do lado direito da Equação (4.2), temos que  $P(A \mid D \cap B_x) = 1$ , pois essa probabilidade condicional é exatamente o caso 1, já que o valor do vértice minimal é maior que  $(p-x)/(1-x)$ . Assim,

$$g_p(x) = P\left(M > \frac{p-x}{1-x}\right) + P\left(A \cap \left\{M \leq \frac{p-x}{1-x}\right\} \mid B_x\right),$$

que, pela Equação (5.1), pode ser escrita como

$$g_p(x) = \left(1 - \frac{p-x}{1-x}\right)^3 + P\left(A \cap \left\{M \leq \frac{p-x}{1-x}\right\} \mid B_x\right).$$

Podemos reescrever a segunda parcela do lado direito da Equação (4.2) da seguinte forma:

$$g_p(x) = \left(\frac{1-p}{1-x}\right)^3 + \int_0^{(p-x)/(1-x)} P(A \mid \{M = b\} \cap B_x) f_M(b) db, \quad (4.3)$$

onde  $f_M(b)$  é dada pela Equação (5.2).

Também podemos reescrever  $P(A \mid \{M = b\} \cap B_x)$  como

$$\begin{aligned} P(A \mid \{M = b\} \cap B_x) &= \sum_{i=0}^2 P(A, \{K = i\} \mid \{M = b\} \cap B_x) \\ &= \sum_{i=0}^2 P(A \mid \{K = i\} \cap \{M = b\} \cap B_x) P(K = i \mid \{M = b\} \cap B_x). \end{aligned}$$

onde  $\{K = i\}$  é o evento  $\{\text{O } i\text{-ésimo vértice é o vértice minimal}\}$ . Note que  $P(A \mid \{K = i\} \cap \{M = b\} \cap B_x)$  só é diferente de zero quando  $K = 0$ , que é exatamente o caso 2. Vale também que  $\{K = i\}$  é independente dos eventos  $\{M = b\}$  e  $B_x$ . Assim,

$$\begin{aligned} P(A \mid \{M = b\}, B_x) &= P(A \mid \{K = 0\}, \{M = b\}, B_x) P(K = 0) \\ &= \frac{1}{3} P(A \mid \{K = 0\}, \{M = b\}, B_x), \end{aligned}$$

já que  $P(K = i) = 1/3$ , para todo  $i$ .

Veja que o evento  $\{A \mid \{K = 0\} \cap \{M = b\} \cap B_x\}$  é equivalente ao evento  $\{A \mid B_{x+(1-x)b}\}$ . Com efeito, o evento em que os vizinhos da origem são declarados fechados, dado que no modelo Bak-Sneppen a origem é o vértice minimal com valor  $b$  e o estado atual dos vizinhos é  $(x, d)$ , é equivalente ao evento em que os vizinhos da origem são declarados fechados, dado que o estado atual deles é  $(x + (1 - x)b, d)$ . Este é exatamente o caso 2. Portanto,

$$P(A \mid \{M = b\}, B_x) = \frac{1}{3}P(A \mid B_{x+(1-x)b})$$

Pela definição de  $g_p$ , temos que

$$P(A \mid \{M = b\}, B_x) = \frac{1}{3}g_p(x + (1 - x)b). \quad (4.4)$$

Substituindo as Equações (4.4) e (5.2) (para  $N = 3$ ) na Equação (4.3), temos:

$$g_p(x) = \left(\frac{1-p}{1-x}\right)^3 + \frac{1}{3} \int_0^{(p-x)/(1-x)} 3(1-b)^2 g_p(x + (1-x)b) db.$$

Substituindo  $y = x + (1-x)b$  (isto é,  $b = (y-x)/(1-x)$ ), o que implica que  $db = dy/(1-x)$ ), obtemos:

$$\begin{aligned} g_p(x) &= \left(\frac{1-p}{1-x}\right)^3 + \int_x^p \left(\frac{1-y}{1-x}\right)^2 g_p(y) \frac{dy}{(1-x)} \\ &= \frac{1}{(1-x)^3} \left( (1-p)^3 + \int_x^p (1-y)^2 g_p(y) dy \right); \end{aligned} \quad (4.5)$$

e, para  $h$  pequeno,

$$g_p(x+h) = \frac{1}{(1-x-h)^3} \left( (1-p)^3 + \int_{x+h}^p (1-y)^2 g_p(y) dy \right). \quad (4.6)$$

Portanto, a diferença entre as Equações (4.6) e (4.5) nos fornece:

$$g_p(x+h) - g_p(x) = \frac{((1-x)^3 - (1-x-h)^3)}{(1-x-h)^3} g_p(x) - \frac{1}{(1-x-h)^3} \int_x^{x+h} (1-y)^2 g_p(y) dy. \quad (4.7)$$

Podemos usar a Equação (4.7) para mostrar que  $g_p$  é contínua. Como  $g_p \in [0, 1]$ , temos que  $g_p(x+h) \rightarrow g_p(x)$ , quando  $h \rightarrow 0$ . Portanto,  $g_p$  é contínua, e o coeficiente diferenciável é dado por

$$\frac{g_p(x+h) - g_p(x)}{h} = \frac{((1-x)^3 - (1-x-h)^3)}{h(1-x-h)^3} g_p(x) - \frac{1}{h(1-x-h)^3} \int_x^{x+h} (1-y)^2 g_p(y) dy. \quad (4.8)$$

Vamos calcular o coeficiente diferenciável para obter a derivada de  $g(x)$ . Primeiro, vamos calcular o limite à direita do coeficiente diferenciável:

$$\begin{aligned} \lim_{h \rightarrow 0^+} \frac{g_p(x+h) - g_p(x)}{h} &= \lim_{h \rightarrow 0^+} \left( \frac{(1-x)^3 - (1-x-h)^3}{h(1-x-h)^3} g_p(x) - \frac{1}{h(1-x-h)^3} \int_x^{x+h} (1-y)^2 g_p(y) dy \right) \\ &= \lim_{h \rightarrow 0^+} \frac{(1-x)^3 - (1-x-h)^3}{h(1-x-h)^3} g_p(x) - \lim_{h \rightarrow 0^+} \frac{1}{h(1-x-h)^3} \int_x^{x+h} (1-y)^2 g_p(y) dy. \end{aligned} \quad (4.9)$$

Para a primeira parcela do lado direito da Equação (4.9), temos:

$$\begin{aligned} \lim_{h \rightarrow 0^+} \frac{(1-x)^3 - (1-x-h)^3}{h(1-x-h)^3} g_p(x) &= \lim_{h \rightarrow 0^+} \frac{3(1-x)^2 h - 3(1-x)h^2 + h^3}{h(1-x-h)^3} g_p(x) \\ &= \frac{3(1-x)^2 g_p(x)}{(1-x)^3}. \end{aligned} \quad (4.10)$$

Para calcular o limite da segunda parcela da Equação (4.9) vamos usar a regra de L'Hôpital (já que temos uma indeterminação do tipo  $\frac{0}{0}$ ). Para isso, vamos calcular a derivada do numerador, que vamos definir como  $F(h) = \int_x^{x+h} (1-y)^2 g_p(y) dy$ . Defina  $G(y)$  como a primitiva de  $(1-y)^2 g_p(y)$ , isto é,  $G'(y) = (1-y)^2 g_p(y)$ . Pelo teorema fundamental do cálculo,  $G(y)$  é uma função contínua no intervalo  $[x, x+h]$ , pois  $(1-y)^2$  e  $g_p(y)$  são contínuas neste mesmo intervalo. Portanto, novamente pelo teorema fundamental do cálculo,  $F(h) = G(x+h) - G(x)$ , ou seja,  $F'(h) = G'(x+h) = (1-x-h)^2 g_p(x+h)$ . E a derivada do denominador da segunda parcela da Equação (4.9) é igual a  $(1-x-h)^3 - 3h(1-x-h)^2$ . Portanto,

$$\begin{aligned} \lim_{h \rightarrow 0^+} \frac{1}{h(1-x-h)^3} \int_x^{x+h} (1-y)^2 g_p(y) dy &\stackrel{(H)}{=} \lim_{h \rightarrow 0^+} \frac{(1-x-h)^2 g_p(x+h)}{(1-x-h)^3 - 3h(1-x-h)^2} \\ &= \frac{(1-x)^2 g_p(x)}{(1-x)^3}. \end{aligned} \quad (4.11)$$

Substituindo as Equações (4.10) e (4.11) na Equação (4.9), obtemos que

$$\lim_{h \rightarrow 0^+} \frac{g_p(x+h) - g_p(x)}{h} = \frac{3(1-x)^2 g_p(x)}{(1-x)^3} - \frac{(1-x)^2 g_p(x)}{(1-x)^3} = \frac{2g_p(x)}{1-x}.$$

Fazendo exatamente o mesmo procedimento feito para o limite à direita, obtemos o limite à esquerda:

$$\lim_{h \rightarrow 0^-} \frac{g_p(x+h) - g_p(x)}{h} = \frac{2g_p(x)}{1-x}.$$

Ou seja, os limites laterais são iguais, o que implica que o limite existe. Logo,  $g_p(x)$  é diferenciável e  $g'_p(x) = 2g_p(x)/(1-x)$ . Esta é uma equação diferencial ordinária linear de primeira ordem, com condição de contorno dada por  $g_p(p) = 1$ . Esta equação tem uma única solução para cada  $p$  e sua estrutura é dada por:

$$\begin{cases} g'_p(x) = \frac{2g_p(x)}{1-x} \\ g_p(p) = 1 \end{cases}.$$

Para resolver o sistema, assumamos que  $g_p(x) \neq 0$ . Vamos dividir ambos os lados da equação por  $g_p(x)$ :

$$\frac{g'_p(x)}{g_p(x)} - \frac{2}{1-x} = 0.$$

Então, integramos com respeito a  $x$  e obtemos:

$$\ln |g'_p(x)| - \int_x \frac{2}{1-s} ds = c_1(p),$$

onde  $c_1(p)$  é uma constante de integração. Então, tomamos a exponencial em ambos os lados:

$$|g'_p(x)| \exp\left(\int_x -\frac{2}{1-s} ds\right) = e^{c_1(p)}.$$

Ou seja,

$$g_p(x) = \pm e^{c_1(p)} \exp\left(\int_x \frac{2}{1-s} ds\right)$$

$$g_p(x) = \pm e^{c_1(p)} \exp(-2\ln(1-x))$$

$$g_p(x) = \pm e^{c_1(p)} \frac{1}{(1-x)^2}.$$

Definimos uma nova constante  $c_2(p) = \pm e^{c_1(p)}$  e, assim,

$$g_p(x) = \frac{c_2(p)}{(1-x)^2}.$$

Usando a condição de contorno, temos que  $c_2(p) = (1-p)^2$ , o que implica em  $g_p(x) = (1-p)^2/(1-x)^2$ . Em particular, a probabilidade de que, no processo acoplado,

os dois vizinhos da origem sejam declarados fechados é dada por  $g_p(0) = (1 - p)^2$ , como desejado.

Este tipo de cálculo não é generalizado para eventos mais complicados. Portanto, a prova de que o processo acoplado constrói um aglomerado a partir da origem no modelo de percolação em sítios, que veremos agora, tem um estilo diferente.

### 4.4.2 O acoplamento

Nesta seção, vamos olhar para o modelo Bak-Sneppen sem memória sob uma perspectiva diferente. Na Seção 2.3, definimos o modelo Bak-Sneppen sem memória e verificamos que a distribuição dos vértices neste modelo coincide com a distribuição dos vértices no modelo original. Aqui, veremos que, a partir das distribuições (condicionais) dos vértices no modelo Bak-Sneppen sem memória, podemos gerar um vetor aleatório, que vamos chamar de informação, e essa informação será usada para atualizar os estados dos vértices no processo acoplado. Vamos determinar qual é a distribuição dessa informação e as consequências sobre o processo acoplado. Por fim, mostraremos que o processo acoplado, de fato, constrói um aglomerado a partir da origem no modelo de percolação.

#### A informação gerada pelo modelo Bak-Sneppen sem memória

Nosso primeiro objetivo é determinar a distribuição da informação que usamos para gerar o processo acoplado. Por enquanto, vamos esquecer o processo acoplado e focar apenas no modelo Bak-Sneppen. Considere o instante de tempo  $n$  do modelo Bak-Sneppen sem memória, onde o alcance é dado por  $r_n$ . Vamos enumerar estes vértices, de 0 a  $r_n - 1$ , e vamos supor que os respectivos limiares dos  $n$  vértices na avalanche são  $y_0, \dots, y_{r_n - 1}$  (Isto é, cada vértice tem distribuição  $U(y_i, 1)$ ), para  $0 \leq i \leq r - 1$ . Seja  $\{U_i\}_{i=0}^{r_n - 1}$  uma sequência de variáveis aleatórias  $U(0, 1)$  associadas aos respectivos  $r_n$  vértices. Agora, vamos gerar uma amostra desse vetor, ou seja, vamos gerar uma sequência aleatória de  $r_n$  valores distribuídos uniformemente no intervalo  $(0, 1)$ . Usando a função  $F_{y_i}(U_i)$ , obtemos uma nova sequência aleatória. Esse procedimento é equivalente a obter uma amostra diretamente das distribuições marginais  $U(y_i, 1)$ .

Então, localizamos o vértice minimal, que tem valor  $M$  e é o vértice de índice  $K$ , onde  $M$  e  $K$  são variáveis aleatórias. Portanto,

$$U_K = \frac{M - y_K}{1 - y_K}.$$

Condicionando em  $K$  e  $M$ , os valores restantes  $U_i$ ,  $i \neq K$ , são uniformemente distribuídos acima de  $(y_i \vee M)$ , respectivamente. Isto é, sabemos que

$$U_i > \frac{M - y_i}{1 - y_i}, \quad i \neq K.$$

Nós descrevemos uma forma um pouco complexa de amostrar a partir do vetor original  $(U_0, \dots, U_{r_n-1})$ , isto é, geramos uma amostra a partir de  $(F_{y_0}(U_0), \dots, F_{y_{r_n-1}}(U_{r_n-1}))$  e, dada a localização do vértice minimal e do seu valor, é possível extrair informações sobre as coordenadas  $U_i$ , de forma que seja possível gerar uma amostra de  $U(0, 1)$ . Note que não estamos afirmando que  $U_K$  é uniformemente distribuído no intervalo  $(0, 1)$ . Entretanto, como  $K$  é aleatório, isso não contradiz o fato de que o vetor  $(U_0, \dots, U_{r_n-1})$  consiste de variáveis aleatórias independentes e uniformemente distribuídas no intervalo  $(0, 1)$ .

### Um exemplo

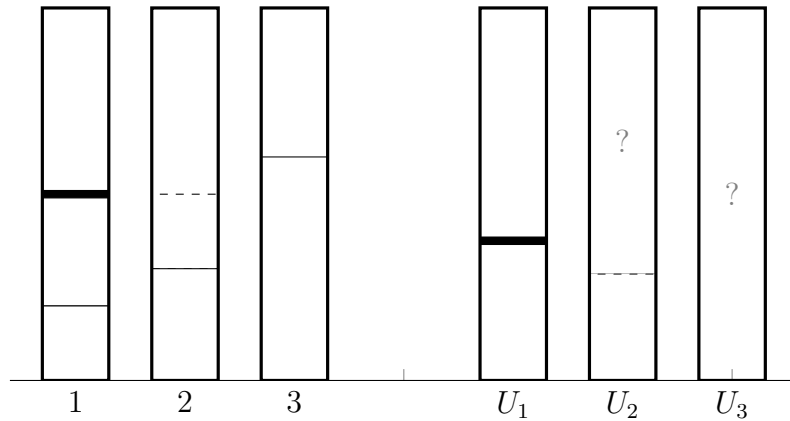


Figura 4.18: Amostra de 3 uniformes no intervalo  $(0, 1)$  recuperada a partir das distribuições dos vértices do modelo Bak-Sneppen sem memória.

Seja  $(U_1, U_2, U_3)$  um vetor aleatório de 3 variáveis i.i.d. com distribuição comum  $U(0, 1)$ . Considere os vértices 1, 2 e 3 no modelo Bak-Sneppen sem memória, como na Figura 4.18, com distribuições  $U(0.2, 1)$ ,  $U(0.3, 1)$  e  $U(0.6, 1)$ , respectivamente. Após os valores serem gerados a partir das distribuições, o vértice 1 é o minimal com va-

lor igual a  $M = 0.5$ . Assim, sabemos a localização do vértice minimal e o seu valor. Consequentemente, é possível obter uma amostra 3 uniformes  $(0, 1)$ . Para isso, basta usar os respectivos limiares das distribuições:  $U_1 = (0.5 - 0.2)^+ / (1 - 0.2) = 3/8$ ,  $U_2 \geq (0.5 - 0.3)^+ / (1 - 0.3) = 2/7$  e  $U_3 \geq (0.5 - 0.6)^+ / (1 - 0.6) = 0$ . Essa amostra, composta por  $U_1 = 3/8$ , uma amostra de uma  $U(2/7, 1)$  e uma amostra de uma  $U(0, 1)$ , é uma amostra de três variáveis aleatórias independentes e uniformemente distribuídas no intervalo  $(0, 1)$ .

### De volta ao processo acoplado

Voltando para a Seção 4.3, deve estar claro que a informação gerada acima, isto é, as variáveis aleatórias independentes e uniformemente distribuídas, são usadas para alterar os estados dos vértices contidos em  $\mathcal{E}$ . Observe que usar  $|\mathcal{E}|$  coordenadas ao invés de  $n$  não afeta suas distribuições marginais, já que os valores das  $U_i$ 's independem dos vértices estarem ou não em  $\mathcal{E}$ .

Agora é possível dar uma descrição direta da construção do processo acoplado. Iniciamos com a origem aberta e observamos os vizinhos da origem, que têm estado  $(0, d)$ , ou seja, distribuição  $F_0$ . Amostramos dessas distribuições usando o vetor  $(U_0, \dots, U_{r_n-1})$ , que foi obtido a partir do modelo Bak-Sneppen. Essa amostragem é feito da seguinte maneira: em cada instante de tempo, pelo menos um vértice é fixado e aos demais são atribuídas distribuições. O caso em que nenhum vértice é fixado é quando o vértice que corresponde ao vértice minimal não pertence a  $\mathcal{E}$ . Em outras palavras, o caso em que há um valor fixado corresponde ao caso em que  $v_{C,K} \in \mathcal{E}$ . Para calcular os novos valores associados aos vértices em  $\mathcal{E} \setminus \{v_{C,K}\}$ , usamos a informação de que as  $U_i$ 's são independentes e uniformemente distribuídas acima de  $(M - y_i)^+ / (1 - y_i)$ , respectivamente. Ou seja, os valores atuais não são fixados, mas as respectivas distribuições dos vértices são alteradas, condicionando nesses valores. Assim que um vértice tem um valor fixado, ele é declarado aberto se, e somente se, esse valor for menor que  $p$ . Sempre que um vértice é declarado aberto, os vizinhos que não têm um valor fixado ou que não pertencem a  $\mathcal{E}$  são adicionados a  $\mathcal{E}$  com estado  $(0, d)$ .

Tendo em vista que os vértices têm distribuições  $U(0, 1)$  independentes quando adiciona-

dos a  $\mathcal{E}$  e a informação que nós usamos para atualizar as distribuições também é de uma amostra de  $U(0, 1)$  independentes, vale o seguinte: se, em qualquer instante de tempo, o procedimento é interrompido e todas as distribuições são amostradas (isto é, são gerados valores a partir de todos os vértices que estão associados a uma variável aleatória), uma amostra de uniformes no intervalo  $(0, 1)$  independentes será recuperada. Portanto, todos os vértices considerados são abertos independentemente e com probabilidade  $p$ . E esse procedimento não é diferente de construir um aglomerado a partir da origem no modelo de percolação em sítios através do método descrito na Seção 1.2.2. Isso completa a prova do Lema 4.1.2.

Veja que, no caso de uma avalanche infinita no modelo Bak-Sneppen, alguns vértices no processo acoplado nunca terão um valor fixo. E isso é bastante razoável, pois é um indicativo de ocorrência de percolação a partir da origem, isto é, um aglomerado a partir da origem é construído dinamicamente e pode ser infinito. De forma equivalente, no modelo de percolação em sítios, nem todos os vértices serão examinados na construção do aglomerado.



## 5 Avalanches Maximais

Até agora, no estudo de avalanches no modelo Bak-Sneppen, nosso foco se restringiu a valores fixos do limitante. A partir de agora, vamos estudar o comportamento de avalanches quando o limitante é uma variável aleatória e investigaremos a duração esperada das avalanches, dentro deste contexto, em grafos finitos e conexos. O material utilizado como referência é o [4].

### 5.1 Introdução e resultados

Na Seção 2.2, vimos que podemos definir o modelo Bak-Sneppen como uma sequência de avalanches. Vimos também que podemos obter diferentes decomposições em avalanches, dentre elas a *decomposição maximal de avalanches*.

Relembrando, na decomposição maximal de avalanches, definimos o limitante inicial como sendo o valor da origem, que é o vértice minimal inicial. Ao final de uma avalanche, o valor do próximo limitante será o valor do novo vértice minimal. Ou seja, sempre fazemos a *escolha máxima* para o limitante, que é o valor mínimo dentre os valores dos vértices naquele instante de tempo.

Uma razão para olhar para a decomposição maximal de avalanches é entender como o modelo Bak-Sneppen tende à criticalidade.

Em grafos transitivos, a duração de uma avalanche depende apenas do limitante, já que todos os vértices têm o mesmo número de vizinhos. Em grafos não-transitivos, a duração depende da origem. Ainda assim, podemos discutir a duração da primeira avalanche com uma origem aleatória em grafos não-transitivos.

Considere o modelo Bak-Sneppen em um grafo conexo e com  $N$  vértices. Se olharmos para a avalanche *inicial* na decomposição maximal, vemos que o limitante inicial é o mínimo de  $N$  variáveis aleatórias independentes e uniformemente distribuídas no in-

tervalo  $(0, 1)$ . Seja  $\{U_i\}_{i=1}^N$  uma sequência de variáveis aleatórias i.i.d.  $U(0, 1)$  e seja  $B = \min\{U_1, \dots, U_N\}$ . Temos que

$$P(B \leq b) = 1 - P(B > b).$$

Veja que podemos reescrever o evento  $\{B > b\}$  como  $\bigcap_{i=1}^N \{U_i > b\}$ . Assim,

$$P(B \leq b) = 1 - P\left(\bigcap_{i=1}^N \{U_i > b\}\right).$$

Como  $\{U_i\}_{i=1}^N$  é uma sequência i.i.d., vale que

$$\begin{aligned} P(B \leq b) &= 1 - \prod_{i=1}^N P(U_i > b) \\ &= 1 - (P(U_1 > b))^N \\ &= 1 - (1 - b)^N, \quad 0 \leq b \leq 1. \end{aligned} \tag{5.1}$$

Isto é, temos uma avalanche com um limitante aleatório  $B$ , cuja função densidade  $f_B(b)$  é dada por

$$f_B(b) = N(1 - b)^{N-1}, \quad 0 \leq b \leq 1. \tag{5.2}$$

Isto é, a distribuição de  $B$  é  $beta(1, N)$ .

Defina a variável aleatória  $D_N^B$  (ou  $D_{G_N}^B$ ) como a duração da avalanche inicial em  $G_N$ , com limitante  $B = \min\{U_1, \dots, U_N\}$ .

**Teorema 5.1.1.** *Seja  $G_N$  um grafo conexo com  $N$  vértices. A duração esperada da primeira avalanche em  $G_N$  é infinita, isto é,  $E(D_N^B) = \infty$ .*

Uma maneira bastante usual de analisar o modelo Bak-Sneppen se dá através de simulações. E este teorema é um prova de que deve-se ter bastante cuidado ao analisar simulações. De acordo com [4], o resultado deste teorema é difícil de ser verificado em simulações, especialmente quando  $N$  é grande. E isso decorre do fato de que é necessário um limitante alto, o que é improvável. Vale ressaltar que aqui o limitante é uma variável aleatória.

**Corolário 5.1.1.** *Se  $G_N$  é transitivo, então qualquer avalanche em  $G_N$  tem duração esperada infinita.*

*Demonstração.* Com efeito, se o grafo é transitivo, o número de vizinhos de qualquer vértice é o mesmo. Portanto, o alcance de cada vértice também é o mesmo, o que faz com que o valor esperado da duração seja independente da origem da avalanche. E como os limitantes das avalanches subsequentes são estocasticamente maiores que o limitante da avalanche inicial, temos que a duração esperada de qualquer avalanche em  $G_N$  é infinita.

□

Seja  $E_j(D_N^B)$  a duração esperada de uma avalanche com limitante  $B$  em um grafo conexo  $G_N$ , onde o valor do vértice minimal é obtido a partir da sequência de variáveis aleatórias  $\{U_i\}_{i=1}^j$ , isto é, a partir de  $j$  variáveis aleatórias  $U(0,1)$ . No Teorema 5.1.1, podemos denotar  $E(D_N^B)$  como  $E_N(D_N^B)$ .

**Teorema 5.1.2.** *Seja  $G_N$  um grafo conexo com  $N$  vértices. Se  $j > N$ , então a duração esperada de uma avalanche em  $G_N$  é finita, isto é,  $E_j(D_N^B) < \infty$ .*

Uma variável aleatória  $U(0,1)$  adicional já é suficiente para obter duração esperada finita.

## 5.2 Prova do Teorema 5.1.1

A prova será feita para o caso em que  $G_N = \Lambda_N$ , onde  $\Lambda_N$  é um círculo com  $N$  vértices. Mais adiante, veremos como é feita a extensão para o caso geral. Para provar o teorema, são enunciados e provados os Lemas 5.2.1, 5.2.2 e 5.2.3. A ideia da prova consiste em dividir a avalanche em blocos (que vamos definir abaixo) e mostrar que, dada uma configuração específica do primeiro bloco, a avalanche só termina após início de um evento, e o tempo para a ocorrência deste evento tem valor esperado infinito.

Usando uma propriedade de esperança condicional, podemos verificar o teorema para  $N = 3$ . Veja que

$$E(D_{\Lambda_3}^B) = E(E(D_{\Lambda_3}^B) | B) = \int_0^1 E(D_{\Lambda_3}^B | B = b) f_B(b) db = \int_0^1 E(D_{\Lambda_3}^b) f_B(b) db.$$

Quando  $N = 3$ , as distribuições dos vértices são i.i.d. em qualquer instante de tempo, já que qualquer vértice de  $\Lambda_3$  é adjacente aos demais vértices. E fixando o limitante, a

probabilidade de sucesso será sempre igual (aqui, o sucesso é o valor estar acima de  $b$ ).

Portanto,

$$E(D_{\Lambda_3}^B) = \int_0^1 \frac{3(1-b)^2}{(1-b)^3} db = 3 \int_0^1 \frac{1}{1-b} db = \infty.$$

Para  $n > 3$ , uma técnica mais sofisticada será necessária, já que as distribuições dos vértices não são independentes.

Para todo  $m \in \mathbb{N}$ , definimos o *bloco*  $m$ , de tamanho  $2N$ , como um conjunto de  $2N$  atualizações no intervalo de tempo  $[2mN, 2(m+1)N - 1]$  onde, em cada atualização (instante de tempo), 3 novas variáveis aleatórias uniformes no intervalo  $(0, 1)$  são geradas e associadas ao vértice minimal e à seus vizinhos (vale ressaltar que o grafo em questão é um círculo e, portanto, cada vértice tem 2 vizinhos). Essa associação é feita da seguinte forma: suponha que no instante de tempo  $n$ , o  $i$ -ésimo vértice é o vértice minimal. Assim, a  $n$ -ésima atualização é feita, isto é, associamos as variáveis aleatórias  $U_{3n}$ ,  $U_{3n+1}$ ,  $U_{3n+2}$  aos vértices  $i-1$ ,  $i$  e  $i+1$ , respectivamente. Na prática, essas variáveis aleatórias geram os valores atribuídas aos vértices. Note que é possível que um vértice esteja associado a mais de uma variável aleatória. Note também que cada bloco contém  $6N$  variáveis aleatórias, uma vez que cada atualização gera 3 variáveis aleatórias.

### Um exemplo

Considere o grafo  $\Lambda_7$ , como na Figura 5.1.

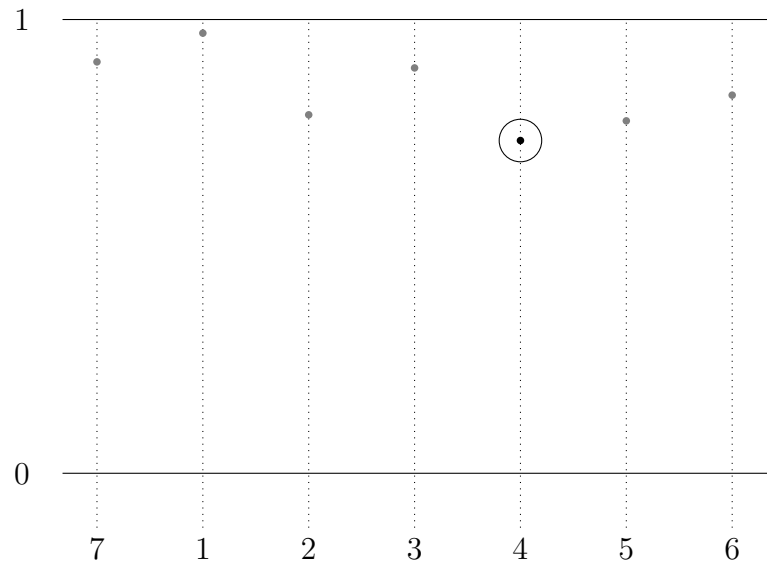


Figura 5.1: Configuração de  $\Lambda_7$  no início da  $b$ -avalanche. Os valores dos vértices, representados pelos pontos no gráfico, foram gerados a partir de uma amostra de 7 variáveis aleatórias independentes e uniformemente distribuídas no intervalo  $(0, 1)$ . O mínimo dessa amostra determina o limitante, que será denotado por  $b$ .

O limitante é definido como o valor mínimo da amostra inicial. Em outras palavras, o limitante é o valor do vértice minimal (que é o vértice 4) antes do início da avalanche. Como o valor é  $b$ , daremos início a uma  $b$ -avalanche (Figura 5.2), e executamos o procedimento de atualização.

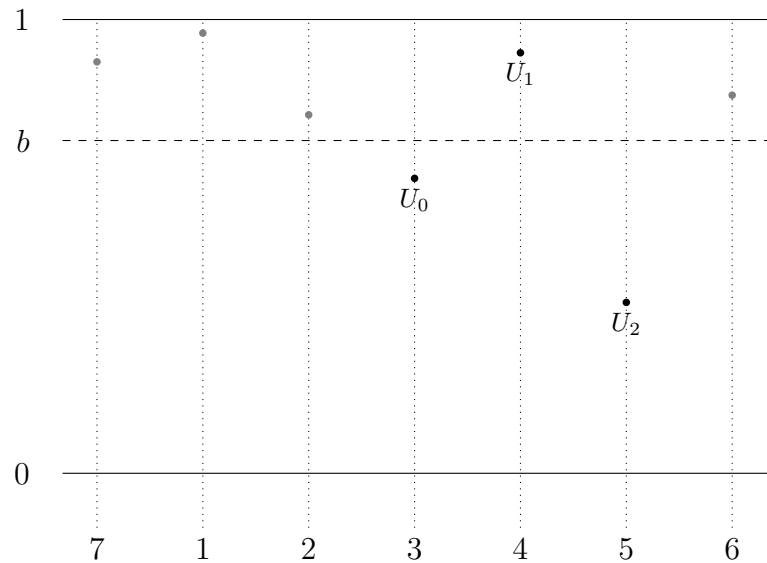


Figura 5.2: Configuração de  $\Lambda_7$  após a primeira atualização, ao final do instante de tempo  $n = 0$ . Os pontos em preto representam os valores dos vértices que foram alterados e, abaixo destes pontos, as variáveis aleatórias  $U_0, U_1$  e  $U_2$  associadas ao vértices 3, 4 e 5, respectivamente.

Agora considere que a  $b$ -avalanche foi executada até o final do bloco 0 (intervalo de tempo  $[0, 13]$ ). Por fins de ilustração, vamos omitir as variáveis aleatórias  $U_0, U_1, \dots, U_{41}$  do bloco 0.

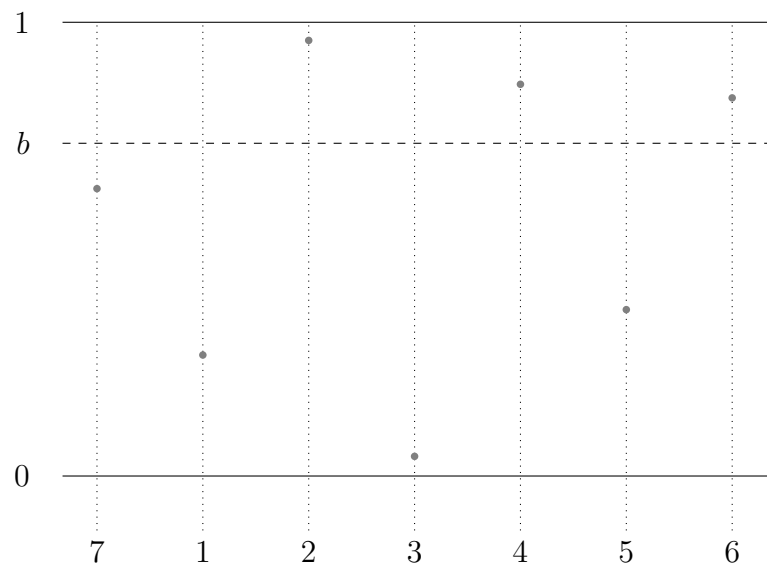


Figura 5.3: Configuração de  $\Lambda_7$  ao final do primeiro bloco (bloco 0), depois de 14 atualizações. Note que a  $b$ -avalanche ainda está em andamento.

No instante de tempo  $n = 2 \cdot 7 = 14$  (Figura 5.4), a avalanche ainda está em andamento, e temos o início do bloco 1. Veja que o vértice 3 é o vértice minimal. Então, associamos as variáveis aleatórias  $U_{42}$ ,  $U_{43}$  e  $U_{44}$  aos vértices 2, 3 e 4, respectivamente.

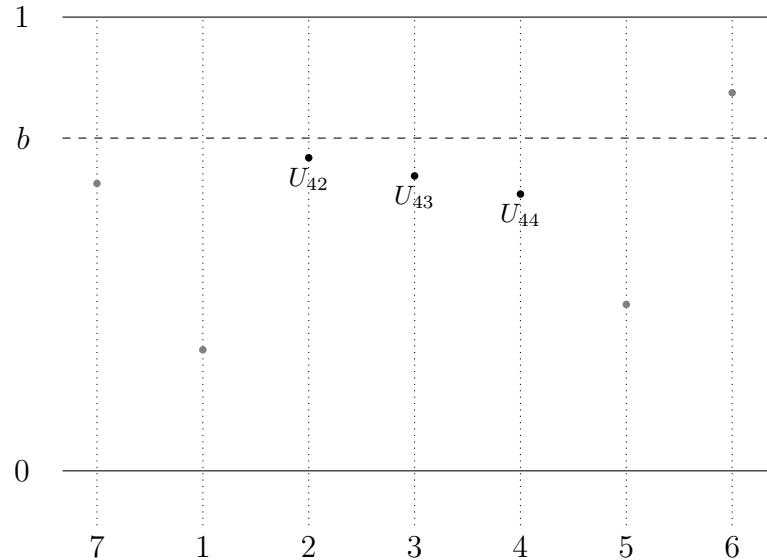


Figura 5.4: Configuração de  $\lambda_7$  ao final no instante de tempo  $n = 14$ , depois de 15 atualizações.

No instante de tempo  $n = 15$  (Figura 5.5), o vértice 1 é o vértice minimal e, consequentemente, associamos variáveis aleatórias  $U_{45}$ ,  $U_{46}$  e  $U_{47}$  aos vértices 7, 1 e 2, respectivamente. Dessa forma, depois de duas atualizações, as variáveis aleatórias  $U_{42}$  e  $U_{47}$  estão associadas a um mesmo vértice (vértice 2).

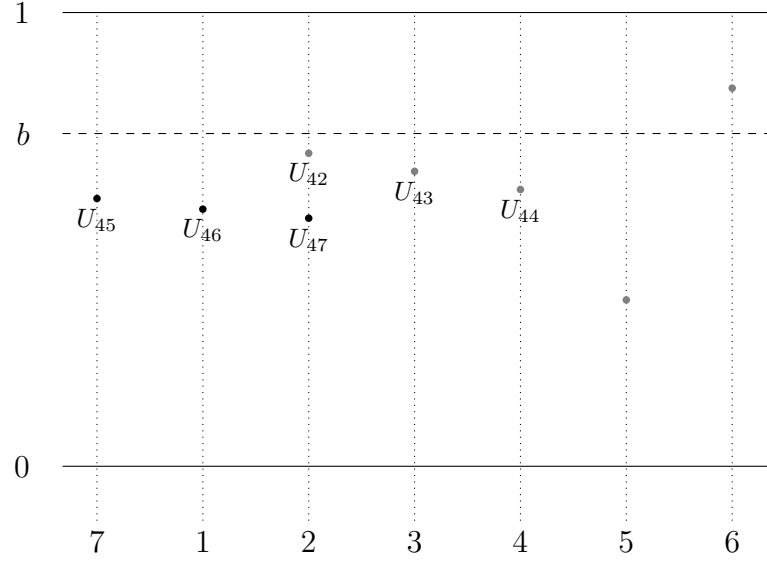


Figura 5.5: Configuração ao final do instante de tempo  $n = 15$ , depois de 16 atualizações.

No exemplo, observe que cada bloco contém  $7 \cdot 6 = 42$  variáveis aleatórias. Dizemos que o bloco  $m$  é *ruim* se o evento

$$\{B > U_{6mN} > U_{6mN+1} > U_{6mN+2} > \dots > U_{6(m+1)N-1}\} \quad (5.3)$$

ocorre. Em palavras, um bloco é ruim se todos os valores do bloco estão abaixo de  $B$  e estão em ordem decrescente. No exemplo anterior, o bloco 0 não pode ser ruim (Figura 5.2), já que os valores gerados pelas variáveis aleatórias no instante de tempo  $n = 0$  já não satisfazem a relação acima. Por outro lado, a Figura 5.5 sugere que o bloco 1 seja ruim. Para que isso de fato aconteça, a relação de monotonicidade deve ser preservada até final do instante de tempo  $n = 27$ .

Seja

$$\begin{aligned} a_i : \mathbb{N} &\longrightarrow [0, 1] \\ n &\longmapsto u \end{aligned}$$

o valor do  $i$ -ésimo vértice no tempo  $n$  e seja  $s_r$  o primeiro instante de tempo logo após o fim do  $r$ -ésimo bloco ruim. Isto é,  $s_r$  é um variável aleatória que toma valores no conjunto  $\{2N, 4N, 6N, \dots\}$ , uma vez que cada bloco tem tamanho  $2N$ .

**Lema 5.2.1.** *Para  $i = 1, \dots, N$ ,  $a_i(s_r) < B$ . Em palavras, no instante de tempo  $s_r$ , todos os vértices estão ativos.*



*Demonstração.* Considere um instante de tempo do início de um bloco ruim. Após a primeira atualização, procuramos o novo vértice minimal. A partir daí, temos duas possibilidades: ou o novo vértice minimal é um vértice que acabou de ser atualizado, ou o novo vértice minimal é um vértice que não foi atualizado agora. Mas veja que a segunda possibilidade não pode acontecer mais do que  $N$  vezes antes da primeira. Agora, assumamos que o vértice minimal é um vértice que acabou de ser atualizado. Portanto, no máximo, em  $N$  atualizações, todos os vértices terão sido atualizados pelo menos uma vez. Assim, depois de  $N + N = 2N$  atualizações, todos os vértices já foram atualizados e seus respectivos valores estão abaixo do limitante  $B$ , já que estamos em um bloco ruim. Uma vez que  $s_r$  é o instante de tempo após as  $2N$  atualizações, o lema está provado. □

Para os lemas a seguir, vamos precisar de mais algumas definições. Dizemos que uma *época* é o intervalo entre dois blocos ruins. Utilizando a notação já definida, a  $m$ -ésima época é o intervalo  $[s_m, s_{m+1} - 1]$ . Note que o tamanho desses intervalos pode ser definido como uma variável aleatória que toma valores no conjunto  $\{0, 2N, 4N, \dots\}$ . A  $m$ -ésima época é definida como *boa* se

$$\sum_{k=3s_m}^{3s_{m+1}-1} \mathbf{1}_{\{U_k \geq B\}} \geq N.$$

Em palavras, a  $m$ -ésima época é boa se pelo menos  $N$  variáveis aleatórias do intervalo  $[3s_m, 3s_{m+1} - 1]$  geram valores acima do limitante  $B$ .

Seja  $A$  o evento em que o bloco 0 é ruim.

**Corolário 5.2.1.** *Se o evento  $A$  ocorre, então uma avalanche só pode terminar em uma época boa.*

*Demonstração.* Se o evento  $A$  ocorre, pelo Lema 5.2.1, todos os vértices devem ter seus respectivos valores abaixo do limitante. Entretanto, para que a avalanche termine, os valores dos  $N$  vértices devem estar acima do limitante  $B$  e, para isso, são necessárias pelo menos  $N$  variáveis aleatórias para gerarem novos valores acima de  $B$ . E esta é justamente a definição de uma época boa. □

Seja  $S$  uma variável aleatória que representa o instante em que começa a primeira época boa. Note que  $S$  toma valores no conjunto  $\{2N, 4N, 6N, \dots\}$ . Antes de enunciarmos o próximo lema, daremos a motivação dos lemas. Lembre-se que o objetivo é mostrar que a duração esperada da primeira avalanche é infinita, isto é,  $E(D_N^B) = \infty$ , para  $G_N = \Lambda_N$ . Observe que

$$E(D_{\Lambda_N}^B) = E(D_{\Lambda_N}^B | A) P(A) + E(D_{\Lambda_N}^B | A^c) P(A^c) \geq E(D_{\Lambda_N}^B | A) P(A).$$

Por outro lado,  $E(D_{\Lambda_N}^B | A) \geq E(S | A)$ . De fato, dada a ocorrência de  $A$ , o instante em que começa uma época boa não pode ser maior que o instante que termina a avalanche. Portanto, para provar o Teorema 5.1.1, basta mostrar que  $E(S | A) = \infty$ , já que  $P(A) > 0$ . E é aqui que entram os próximos lemas.

**Lema 5.2.2.**  $E(S | A) \geq E(S)$ .

*Demonstração.* Seja  $F_{B|A}$  a função distribuição condicional de  $B$  dado o evento  $A$ . Temos que

$$E(S | A) = \int_0^1 E(S | A, B = b) dF_{B|A}(b) db. \quad (5.4)$$

Agora, observe que

$$E(S | A, B = b) = E(S | B = b). \quad (5.5)$$

Note que, quando condicionamos  $S$  a um limitante fixo (isto é,  $B = b$ ), qualquer informação sobre o bloco 0 é irrelevante para a distribuição de  $S$ . E de fato, a distribuição de  $S$  depende apenas de  $B$ , onde  $E(S | B = b)$  é uma função não decrescente em  $b$ . Além disso, veja que,  $\forall b \in [0, 1]$ ,

$$F_{B|A}(b) \leq F_B(b), \quad (5.6)$$

onde  $F_B$  é a função distribuição de  $B$ . Para verificar a equação acima, defina  $U = \max\{U_0, \dots, U_{6N-1}\}$ . Temos que

$$F_{B|A}(b) = P(B \leq b | A).$$

Logo,

$$F_{B|A}(b) = P(B \leq b | B > U_0 > U_1 > \dots > U_{6N-1}),$$

ou, equivalentemente,

$$F_{B|A}(b) = P(B \leq b \mid \{B > U\} \cap \{U_0 > U_1 > \dots > U_{6N-1}\}). \quad (5.7)$$

Observe que o evento  $\{U_0 > U_1 > \dots > U_{6N-1}\}$  é independente do vetor aleatório  $(B, U)$ .

Com efeito,

$$P(U_0 > U_1 > \dots > U_{6N-1} \mid B = b, U \leq u) = P(U_0 > U_1 > \dots > U_{6N-1} \mid U \leq u),$$

pois o limitante é independente das variáveis aleatórias geradas nas atualizações. Agora

nos resta verificar que  $P(U_0 > U_1 > \dots > U_{6N-1} \mid U \leq u) = P(U_0 > U_1 > \dots > U_{6N-1})$ .

Note que o evento  $\{U \leq u\}$  pode ser reescrito como  $\bigcap_{i=0}^{6N-1} \{U_i \leq u\}$ . Logo,

$$P(U_0 > U_1 > \dots > U_{6N-1} \mid U \leq u) = \frac{P\left(\{U_0 > U_1 > \dots > U_{6N-1}\} \cap \left\{\bigcap_{i=0}^{6N-1} \{U_i \leq u\}\right\}\right)}{P\left(\bigcap_{i=0}^{6N-1} \{U_i \leq u\}\right)}.$$

Como as variáveis aleatórias  $U_i$ 's são i.i.d., temos que

$$\begin{aligned} P(U_0 > U_1 > \dots > U_{6N-1} \mid U \leq u) &= \frac{P(u \geq U_0 > U_1 > \dots > U_{6N-1})}{\prod_{i=0}^{6N-1} P(U_i \leq u)} \\ &= \frac{P(u \geq U_0 > U_1 > \dots > U_{6N-1})}{u^{6N}}. \end{aligned}$$

O numerador da fração acima pode ser reescrito da seguinte forma:

$$\begin{aligned} P(u \geq U_0 > U_1 > \dots > U_{6N-1}) &= \int_0^u \underbrace{\int_0^{u_0} \dots \int_0^{u_{6N-2}}}_{6N-1 \text{ integrais}} du_{6N-1} \dots du_1 du_0 \\ &= \frac{u^{6N}}{(6N)!}. \end{aligned} \quad (5.8)$$

Portanto,

$$P(U_0 > U_1 > \dots > U_{6N-1} \mid U \leq u) = \frac{1}{(6N)!}.$$

Por outro lado, substituindo  $u = 1$ , na Equação (5.8), obtemos que

$$P(U_0 > U_1 > \dots > U_{6N-1}) = \frac{1}{(6N)!}.$$

Logo,  $P(U_0 > U_1 > \dots > U_{6N-1} \mid U = u) = P(U_0 > U_1 > \dots > U_{6N-1})$ , o que implica que  $\{U_0 > U_1 > \dots > U_{6N-1}\}$  e  $U$  são independentes. Voltando para a Equação (5.7), temos que

$$F_{B|A}(b) = P(B \leq b \mid B > U).$$

Mas

$$P(B \leq b | B > U) = \frac{P(U < B \leq b)}{P(B > U)}, \quad (5.9)$$

e por definição,

$$P(U < B \leq b) = \int_0^b \int_u^b f_{B,U}(v, u) dv du,$$

onde  $f_{B,U}(v, u)$  é a função densidade conjunta de  $(B, U)$ . E como  $B$  e  $U$  são independentes, temos que

$$P(U < B \leq b) = \int_0^b \int_u^b f_B(v) f_U(u) dv du = \int_0^b f_U(u) [F_B(b) - F_B(u)] du.$$

Usando a desigualdade  $F_B(b)F_B(u) \leq F_B(u)$ , temos que

$$\begin{aligned} P(U < B \leq b) &\leq \int_0^b f_U(u) [F_B(b) - F_B(b)F_B(u)] du \\ &\leq F_B(b) \int_0^b f_U(u) [1 - F_B(u)] du. \\ &\leq F_B(b) \int_0^b f_U(u) \left[ \int_u^1 f_B(v) dv \right] du \\ &\leq F_B(b) \int_0^b \int_u^1 f_U(u) f_B(v) dv du \\ &\leq F_B(b) \int_0^1 \int_u^1 f_U(u) f_B(v) dv du \\ &\leq F_B(b) P(B > U). \end{aligned} \quad (5.10)$$

Substituindo a Equação (5.10) na Equação (5.9), temos que

$$F_{B|A}(b) = P(B \leq b | B > U) \leq \frac{F_B(b)P(B > U)}{P(B > U)} = F_B(b)$$

o que conclui o resultado.

Substituindo as Equações (5.5) e (5.6) na Equação (5.4), obtemos que

$$E(S | A) \geq \int_0^1 E(S | B = b) dF_B(b) = E(S)$$

o que conclui a demonstração do lema. □

**Lema 5.2.3.**  $E(S) = \infty$ .

*Demonstração.* Para provar este lema, vamos usar novamente a propriedade de esperança condicional. Podemos reescrever  $E(S)$  como

$$E(S) = E(E(S | B)) = \int_0^1 E(S | B = b) f_B(b) db.$$

Agora, observe que, fixado um limitante  $B = b$ , a variável aleatória  $S$  tem distribuição geométrica, onde a probabilidade de sucesso é a probabilidade de ocorrer uma época boa. Com efeito, o tamanho das épocas depende da ocorrência de blocos ruins. E como condicionamos  $S$  a um limitante fixo  $B = b$ , as variáveis aleatórias geradas pelas atualizações são i.i.d., o que torna o tamanho dos intervalos das épocas independentes e identicamente distribuídos. Se definirmos  $p(b)$  como a probabilidade de que uma época seja boa, dado que  $B = b$ , temos que  $E(S | B = b) = 1/p(b)$ . Portanto,

$$E(S) = \int_0^1 \frac{N(1-b)^{N-1}}{p(b)} db. \quad (5.11)$$

Agora o nosso objetivo é calcular  $p(b)$ . Como queremos mostrar que  $E(S) = \infty$ , basta obter um limite superior para  $p(b)$ . Para obter tal cota, vamos definir dois eventos. Seja  $F_n$  o evento em que a primeira época não termina antes ou no instante  $s_1 + n$  e  $G_n$  o evento em que

$$\underbrace{\min \left\{ j : \sum_{k=3s_1}^j \mathbf{1}_{\{U_k \geq B\}} = N \right\}}_{X_{N,B}} = n.$$

Em palavras,  $G_n$  é o evento em que o  $n$ -ésimo valor gerado (isto é, o valor gerado por  $U_n$ ) é o  $N$ -ésimo valor acima do limitante  $B$ , após o início da primeira época. Portanto, podemos reescrever  $p(b)$  da seguinte forma:

$$p(b) = \sum_{n=N}^{\infty} P(F_{\lfloor n/3 \rfloor}, G_n | B = b) = \sum_{n=N}^{\infty} P(F_{\lfloor n/3 \rfloor} | G_n, B = b) P(G_n | B = b). \quad (5.12)$$

Observe que evento  $F_n$  só faz sentido para  $n \geq 12N$ , pois a primeira época não acaba antes do instante de tempo  $4N$ . O caso em que a primeira época acaba no instante de tempo  $4N - 1$  é o caso em que os dois primeiros blocos são ruins. Assim, para  $n < 12N$ , poderíamos escrever  $P(G_n | B = b)$  ao invés de  $P(F_{\lfloor n/3 \rfloor}, G_n | B = b)$ . Mas, como  $F_n$  é um evento de probabilidade 1 (se  $n < 12N$ ), escrever  $P(G_n | B = b)$  ou  $P(F_{\lfloor n/3 \rfloor}, G_n | B = b)$  nos casos em que  $n < 12N$  não afeta as parcelas.

Ainda na Equação (5.12), podemos reescrever  $P(G_n|B = b)$  como

$$p(b) = \sum_{n=N}^{\infty} P(F_{\lfloor n/3 \rfloor} | G_n, B = b) \underbrace{P(X_{N,b} = n)}_{P(G_n|B=b)}.$$

Note que para um limitante fixo  $B = b$ , a variável aleatória  $X_{N,b}$  segue distribuição binomial negativa, isto é,  $X_{N,b} \sim \text{binneg}(n; N, 1 - b)$ . Portanto,

$$P(X_{N,b} = n) = \binom{n-1}{N-1} (1-b)^N b^{n-N}, \quad n \geq N.$$

Em palavras, a ocorrência do evento  $G_n$  é a ocorrência de  $N$  sucessos em um número fixo de ensaios,  $n$ , onde o  $N$ -ésimo sucesso ocorre no  $n$ -ésimo ensaio. Conseqüentemente, os demais  $N - 1$  sucessos ocorrem dentre os  $n - 1$  ensaios restantes. Aqui, o sucesso é definido como o evento em que uma variável aleatória gera um valor acima do limitante  $b$ . Como  $b < 1$ , podemos cotar  $P(G_n|B = b)$  por  $\binom{n-1}{N-1} (1-b)^N$ . Assim,

$$p(b) \leq \sum_{n=N}^{\infty} \binom{n-1}{N-1} (1-b)^N P(F_{\lfloor n/3 \rfloor} | G_n, B = b).$$

Observe que, quando  $n > 6N^2$ , temos mais blocos do que sucessos. Com efeito, cada bloco tem  $6N$  variáveis aleatórias, ou seja, se temos  $N$  blocos, temos  $6N^2$  variáveis aleatórias. Portanto, a  $(6N^2 + 1)$ -ésima variável aleatória pertence ao  $(N + 1)$ -ésimo bloco. Neste caso, teríamos pelo menos  $N + 1$  blocos contra  $N$  sucessos. E isto implica que todas as variáveis aleatórias de pelo menos  $\lfloor (n-1)/6N \rfloor - N + 1$  blocos geram valores abaixo do limitante  $b$ . Portanto, a distribuição condicional das variáveis aleatórias de tais blocos são independentes e uniformemente distribuídas no intervalo  $(0, b)$ . Em particular, a probabilidade de qualquer um destes blocos ser ruim é positiva, e vamos limitar esta probabilidade uniformemente por baixo por uma constante  $c_1(N)$ . Isto é equivalente a limitar uniformemente por cima a probabilidade de qualquer um desses blocos não ser ruim. Portanto,  $P(F_{\lfloor n/3 \rfloor} | G_n, B = b) \leq (1 - c_1(N))^{\lfloor (n-1)/6N \rfloor - N + 1}$ , se  $n > 6N^2$ . Para  $n \leq 6N^2$ , vamos cotar  $P(F_{\lfloor n/3 \rfloor} | G_n, B = b)$  por 1. Assim,

$$p(b) \leq \sum_{n=N}^{6N^2} \binom{n-1}{N-1} (1-b)^N + \sum_{n=6N^2+1}^{\infty} \binom{n-1}{N-1} (1-b)^N (1 - c_1(N))^{\lfloor (n-1)/6N \rfloor - N + 1}.$$

Como  $\binom{n-1}{N-1} \leq n^N$ , temos que

$$\begin{aligned} p(b) &\leq \sum_{n=N}^{6N^2} \binom{n-1}{N-1} (1-b)^N + \sum_{n=6N^2+1}^{\infty} n^N (1-b)^N (1-c_1(N))^{\lfloor (n-1)/6N \rfloor - N+1} \\ &\leq (1-b)^N \left( \sum_{n=N}^{6N^2} \binom{n-1}{N-1} + \sum_{n=6N^2+1}^{\infty} n^N (1-c_1(N))^{\lfloor (n-1)/6N \rfloor - N+1} \right). \end{aligned}$$

Agora, defina  $c_2(N) := \sum_{n=N}^{6N^2} \binom{n-1}{N-1}$  e  $c_3(N) := \sum_{n=6N^2+1}^{\infty} n^N (1-c_1(N))^{\lfloor (n-1)/6N \rfloor - N+1}$ . Observe que  $c_2(N) < \infty$ . Então, nos resta mostrar que  $c_3(N)$  também é finito. Podemos reescrever  $c_3(N)$  da seguinte forma:

$$\begin{aligned} c_3(N) &= \sum_{i=0}^{\infty} \sum_{n=1}^{6N} (1-c_1(N))^{i+1} (n+6N^2+6Ni)^N \\ &= \sum_{i=0}^{\infty} (1-c_1(N))^{i+1} \sum_{n=1}^{6N} (n+6N^2+6Ni)^N \\ &\leq \sum_{i=0}^{\infty} (1-c_1(N))^{i+1} 6N(6N+6N^2+6Ni)^N \\ &\leq (6N)^{N+1} \sum_{i=1}^{\infty} (1-c_1(N))^i (N+i)^N. \end{aligned}$$

Aplicando o teste de razão na série acima, obtemos que

$$\begin{aligned} \lim_{i \rightarrow \infty} \left| \frac{a_{i+1}}{a_i} \right| &= \lim_{i \rightarrow \infty} \left| \frac{(1-c_1(N))^{i+1} \cdot (N+i+1)^N}{(1-c_1(N))^i \cdot (N+i)^N} \right| \\ &= (1-c_1(N)) \lim_{i \rightarrow \infty} \left( 1 + \frac{1}{N+i} \right)^N = 1 - c_1(N) < 1. \end{aligned}$$

Portanto, a série converge, o que implica que  $c_3(N) < \infty$ .

Fazendo  $c(N) := c_2(N) + c_3(N)$ , obtemos que

$$p(b) \leq c(N)(1-b)^N < \infty. \quad (5.13)$$

Por fim, substituindo a Equação (5.13) na Equação (5.11), temos que

$$E(S) = \int_0^1 \frac{N(1-b)^{N-1}}{p(b)} db \geq \frac{N}{c(N)} \int_0^1 \frac{(1-b)^{N-1}}{(1-b)^N} db = \frac{N}{c(N)} \int_0^1 \frac{1}{1-b} db = \infty,$$

o que conclui a prova do Lema. □

Provados os Lemas 5.2.1, 5.2.2 e 5.2.3, o Teorema 5.1.1 está provado para  $G_N = \Lambda_N$ . Agora vamos estender a prova para grafos não-circulares. Note que a única parte da prova do teorema que realmente usou a hipótese de que o grafo é circular foi o Lema 5.2.1. Para generalizar a prova para todo  $G_N$  conexo, vamos tentar obter uma noção similar de um bloco ruim, que termina com todos os valores dos vértices abaixo do limitante.

Na Seção 1.1, vimos que um circuito em  $G_N$  é uma sequência alternante  $x_1, e_1, \dots, x_{k-1}, e_{k-1}, x_k, e_k$  tal que  $x_1, e_1, \dots, x_{k-1}, e_{k-1}, x_k$  é um caminho e  $e_k = (x_k, x_1)$ . Para todo grafo  $G_N$  conexo, é possível construir um circuito  $C_{G_N}$  que conecte todos os vértices pelo menos uma vez. Note que é permitido visitar um vértice e atravessar um elo mais de uma vez.

Agora, considere o menor desses circuitos. Vamos caminhar pelo circuito de forma semelhante ao que foi feito no círculo, no Lema 5.2.1. Vamos procurar pelo vértice minimal no circuito. Se existir mais de um vértice minimal, podemos escolher arbitrariamente (por exemplo, na Figura 5.6, isso ocorre se o vértice 1 é o vértice minimal inicial). Então, atualize os valores destes vértices e dos vizinhos no circuito usando as variáveis aleatórias  $U_i$ . A atualização deve ocorrer de forma que a variável aleatória com maior índice fique mais à frente no caminho. Por exemplo, para o grafo da Figura 5.6.a, obtemos o circuito da Figura 5.6.b.

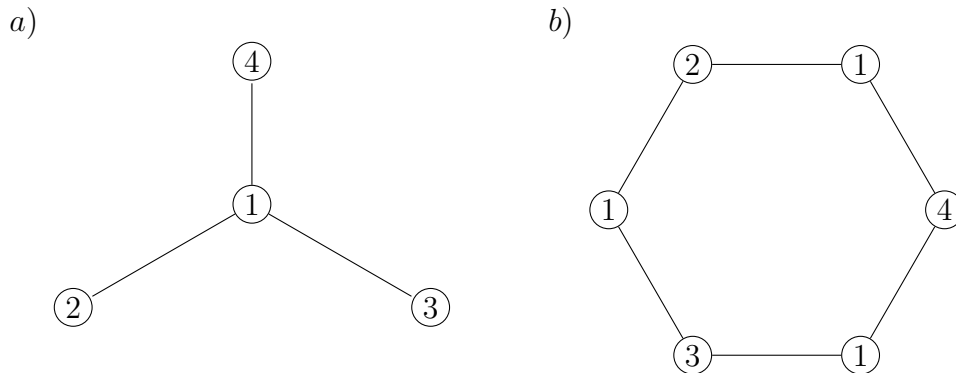


Figura 5.6: Grafo conexo (a) e um circuito (b) obtido a partir deste grafo. Note que o vértice 1 aparece três vezes no circuito e os elos incidentes ao vértice 1 aparecem duas vezes no circuito.

Agora, suponha que o vértice 2 seja o vértice minimal no grafo. No circuito, só temos um vértice correspondente ao vértice 2. Se houvesse mais de um, bastaria escolher um deles. Escolhido o vértice, atualizamos os valores do vértice 2 e dos seus vizinhos. A



atualização no grafo original (Figura 5.7.a) é feita de acordo com os vértices atualizados no circuito (Figura 5.7.b).

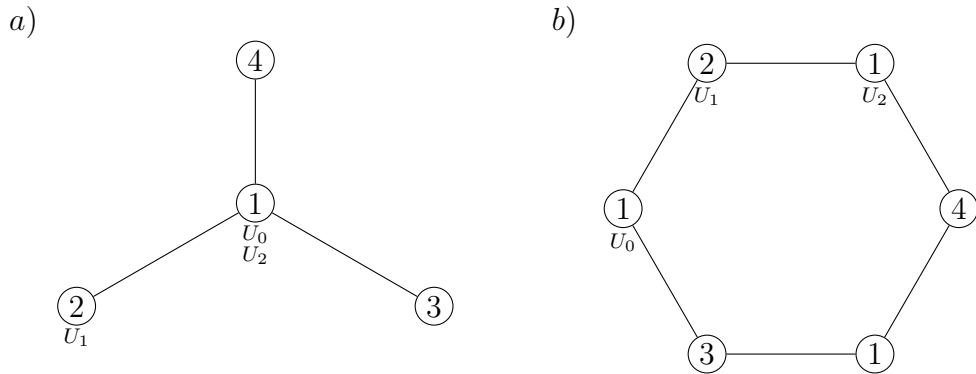


Figura 5.7: Grafo conexo (a) e um circuito (b) obtido a partir deste grafo, depois de uma atualização. Note que em (a), o vértice 1 recebe duas variáveis aleatórias durante a atualização.

Agora, suponha que o vértice 1 é o vértice minimal no grafo. Note que temos três vértices correspondentes no circuito. Mas como a última atualização foi da variável aleatória  $U_2$ , o vértice escolhido no circuito será o vértice associado a  $U_2$ . Agora, atualizamos os valores do vértice 1 e dos seus vizinhos (Figura 5.8.b) A atualização no grafo original é feita de acordo com os vértices atualizados no circuito (Figura 5.8.a). Portanto, o vértice 3, apesar de ser vizinho do vértice 1 no grafo, não é atualizado.

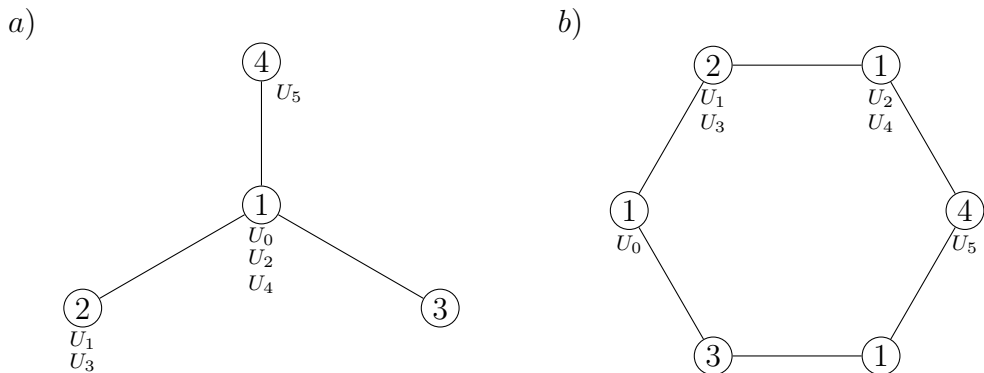


Figura 5.8: Grafo conexo (a) e um circuito (b) obtido a partir deste grafo, depois de duas atualizações.

Devemos ter em mente que a nossa intenção é obter um bloco ruim, a partir do circuito. Se, em qualquer instante de tempo durante um bloco ruim, o vértice minimal for um vértice que acabou de ser atualizado, o vértice minimal sempre será o vértice associado

a  $U_i$  com o maior  $i$ . Para garantir que o vértice minimal seja um vértice que acabou de ser atualizado, vamos cotar por  $l_1$  o número de variáveis aleatórias necessárias para que o valor mínimo venha de um vértice que acabou de ser atualizado. E para que ocorra um bloco ruim, além da ocorrência do evento na Equação (5.3), precisamos garantir que todos os vértices do grafo sejam atualizados pelo procedimento. Vamos cotar por  $l_2$  o número de variáveis aleatórias necessárias para que todos os vértices sejam atualizados. Portanto, um bloco de tamanho  $l = l_1 + l_2$  garante que, ao final do instante de tempo  $l$ , os valores de todos os vértices estarão abaixo do limitante. No caso em que  $G_N = \Lambda_N$ , temos que  $l_1 = 3N$  e  $l_2 = 3N$ . Vale ressaltar que cada bloco de  $\Lambda_N$  tem  $6N$  variáveis aleatórias.

### 5.3 Prova do Teorema 5.1.2

Para provar o Teorema 5.1.2, vamos recorrer a técnica de acoplamento e vamos usar um modelo que já foi definido neste trabalho. Trata-se do modelo Bak-Sneppen sem memória para grafos finitos (veja a Seção 2.3). O objetivo de usar a técnica de acoplamento é obter uma dominação estocástica entre as distribuições dos vértices do grafo  $G_N$  e do grafo  $K_N$ , onde  $K_N$  é um grafo completo (veja o Exemplo 1.1.4) e  $G_N$  é um subgrafo conexo de  $K_N$ . É aqui que entra o modelo Bak-Sneppen sem memória. Ao contrário do modelo Bak-Sneppen, onde estamos interessados nos valores associados aos vértices, no modelo Bak-Sneppen sem memória, o nosso interesse se restringe às distribuições que geram esses valores. E isso é exatamente o que precisamos, uma vez que queremos obter uma dominação estocástica.

Já vimos que para  $j = N$ , a duração esperada em  $G_N$  da primeira avalanche é infinita. Aqui, vamos mostrar que, se  $j > N$ , a duração esperada em  $G_N$  de qualquer avalanche é finita.

Seja  $K_N$  um grafo completo. Então, para  $j \geq N$ :

$$E_j(D_{K_N}^B) = E(E_j(D_{K_N}^B) | B) = \int_0^1 E_j(D_{K_N}^B | B = b) f_B(b) db = \int_0^1 E_j(D_{K_N}^b) f_B(b) db.$$

Observe que a distribuição da variável aleatória  $D_{K_N}^b$  é geométrica, de parâmetro  $(1 - b)^N$ . Com efeito, no grafo  $K_N$ , as distribuições dos vértices são i.i.d. em qualquer

instante de tempo. E como o limitante foi fixado, a probabilidade de sucesso será sempre igual (aqui, o sucesso é o valor estar acima de  $b$ ). Logo,

$$E_j(D_{K_N}^B) = \int_0^1 \frac{j(1-b)^{j-1}}{(1-b)^N} db = j \int_0^1 (1-b)^{j-N-1} db = \frac{j}{j-N} < \infty, \forall j > N.$$

Portanto, para provar o teorema, basta mostrar que  $E_j(D_N^B) \leq E_j(D_{K_N}^B)$ .

Seja  $\{U_i\}_{i=1}^N$  uma sequência de variáveis aleatórias i.i.d. com distribuição comum uniforme no intervalo  $(0, 1)$  e considere  $\{1, \dots, N\}$  os índices dos vértices de  $G_N$ . Seja  $v_{G,i}$  o  $i$ -ésimo vértice do grafo  $G_N$ . Vamos definir dois processos estocásticos: o modelo Bak-Sneppen sem memória em  $G_N$  e o modelo Bak-Sneppen sem memória em  $K_N$ .

O próximo passo é executar os dois processos iterativamente e simultaneamente, de forma que ambos sejam governados pela mesma sequência de variáveis aleatórias,  $\{U_i\}_{i=1}^N$ . Para o modelo Bak-Sneppen sem memória em  $K_N$ , podemos usar a sequência de forma direta, visto que o vértice minimal em  $K_N$  é vizinho de todos os demais vértices do grafo. Portanto, as distribuições de todos os vértices serão sempre uniformes no intervalo  $(0, 1)$ . Já para o modelo Bak-Sneppen sem memória em  $G_N$ , nem sempre as distribuições dos vértices serão uniformes no intervalo  $(0, 1)$ . Assim, será necessário usar o método de transformação de variáveis aleatórias. De forma análoga, vamos usar a mesma função da Seção 4.3, definida  $\forall y \in [0, 1]$  e dada por

$$F_y : [0, 1] \longrightarrow [y, 1]$$

$$u \longmapsto (1-y)u + y.$$

Sabemos que, se  $U \sim U(0, 1)$ , então  $F_y(U) \sim U(y, 1)$ . Portanto, se a distribuição dos vértices de  $G_N$  é  $U(y_i, 1)$ , definimos o acoplamento como o par de variáveis aleatórias  $(U_i, F_{y_i}(U_i))$ . Observe que as distribuições marginais coincidem com as respectivas distribuições dos processos.

### Um exemplo

Antes de dar uma descrição geral, vamos dar um exemplo para um caso particular, onde  $G_N = \Lambda_5$  (Figura 5.9.b). Em cada instante de tempo, uma amostra será gerada

a partir da sequência de variáveis aleatórias  $\{U_i\}_{i=1}^5$  e usada para determinar os valores atribuídos aos vértices nos dois processos.

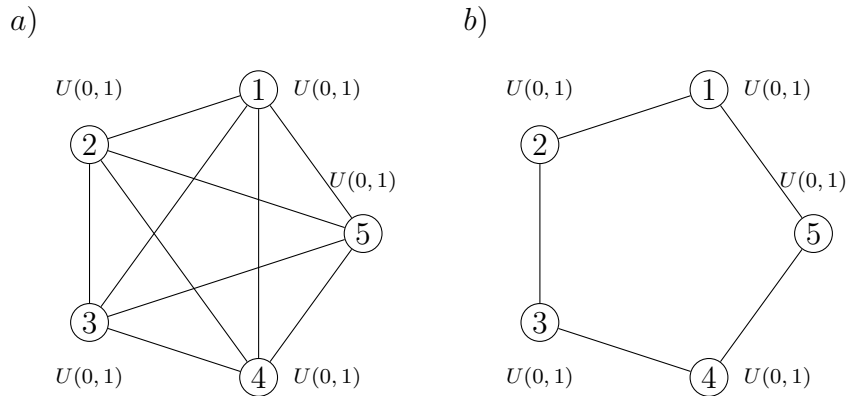


Figura 5.9: Grafos  $K_5$  (a) e  $\Lambda_5$  (b) antes do início das avalanches. Note que associamos uma distribuição a cada vértice.

Seja  $(u_1, u_2, u_3, u_4, u_5) = (0.583, 0.838, 0.972, 0.993, 0.744)$  a amostra no instante de tempo  $n = 0$ . As distribuições iniciais de todos os vértices, tanto em  $K_5$ , quanto  $\Lambda_5$ , são uniformes no intervalo  $(0, 1)$ . No grafo  $K_5$  (Figura 5.10.a), usamos a amostra de forma direta para atribuir os valores aos vértices, ou seja, o valor atribuído ao vértice  $i$  é  $u_i$ ,  $i = 1, 2, 3, 4, 5$ . O mesmo ocorre no grafo  $\Lambda_5$  (Figura 5.10.b), já que  $y_i = 0$  no instante inicial, para  $i = 1, 2, 3, 4, 5$ . Portanto,

$$F_0(u_i) = (1 - 0)u_i + 0 = u_i, i = 1, 2, 3, 4, 5.$$

Assim, obtemos os valores dos vértices em  $\Lambda_5$ :  $(0.583, 0.838, 0.972, 0.993, 0.744)$ . Agora, localizamos o vértice minimal nos processos. Em ambos, o vértice minimal é o vértice 1, e o valor atribuído aos dois vértices é 0.583. Uma vez que estamos no passo inicial, definimos o valor atribuído ao vértice minimal como o limitante da avalanche. Portanto,  $b = 0.583$ .

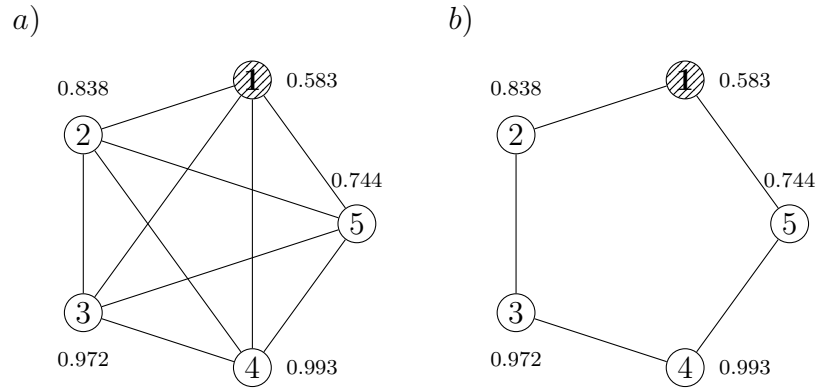


Figura 5.10: Grafos  $K_5$  (a) e  $\Lambda_5$  (b), com os valores atribuídos aos vértices no instante inicial. Em ambos os grafos, o vértice minimal é o vértice 1.

Localizado o vértice minimal, alteramos as distribuições dos vértices minimal e dos seus vizinhos para distribuições uniformes no intervalo  $(0, 1)$ . No grafo  $K_5$  (Figura 5.11.a), todas as distribuições serão  $U(0, 1)$ , pois o vértice minimal é vizinho dos demais vértices. Já no grafo  $\Lambda_5$  (Figura 5.11.b), as distribuições do vértice minimal e dos vértices vizinhos serão  $U(0, 1)$ , enquanto os demais vértices terão distribuição  $U(y_i \vee 0.583, 1)$ . Portanto, os vértices 1, 2, 5 terão distribuição  $U(0, 1)$ , enquanto os vértices 3 e 4 terão distribuição  $U(0.583, 1)$ , já que  $y_3 = y_4 = 0.583$ .

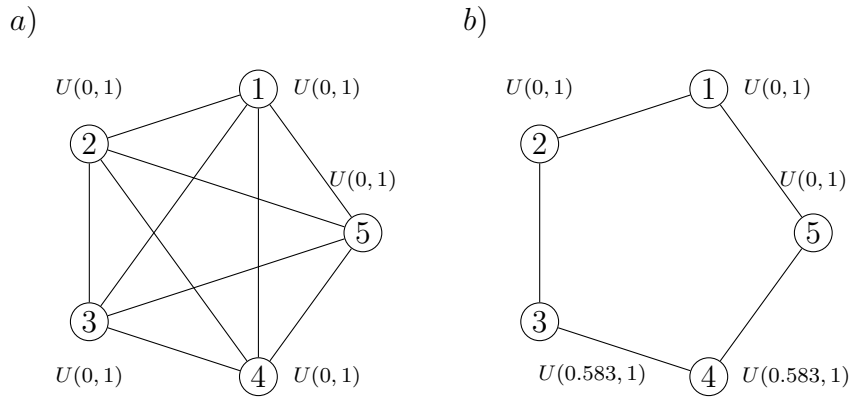


Figura 5.11: Grafos  $K_5$  (a) e  $\Lambda_5$  (b) ao final do instante inicial. Note que as distribuições dos vértices 3 e 4 no grafo  $G_N$  são diferentes da distribuição inicial (Figura 5.9(b)).

Seja  $(u_1, u_2, u_3, u_4, u_5) = (0.192, 0.962, 0.482, 0.464, 0.454)$  a amostra no instante de tempo  $n = 1$ . No grafo  $K_5$  (Figura 5.12.a), usamos a amostra de forma direta para

atribuir os valores aos vértices, ou seja, o valor atribuído ao vértice  $i$  é  $u_i$ ,  $i = 1, 2, 3, 4, 5$ . No grafo  $\Lambda_5$  (Figura 5.12.b), aplicamos a função  $F_{y_i}$  em sua respectiva componente da amostra. Portanto,

$$F_0(u_i) = (1 - 0)u_i + 0 = u_i, i = 1, 2, 5$$

$$F_{0.583}(u_i) = (1 - 0.583)u_i + 0.583 = 0.417u_i + 0.583, i = 3, 4.$$

Assim, obtemos os valores dos vértices em  $\Lambda_5$ : (0.192, 0.962, 0.784, 0.776, 0.454). Então, localizamos o vértice minimal nos processos. Assim como no instante de tempo  $n = 0$ , o vértice minimal de ambos será o vértice 1, e dessa vez o valor atribuído aos dois vértices é 0.192. Observe que, nos dois processos, o valor do vértice minimal é menor que  $b$ , o que mostra que a avalanche de nenhum dos processos terminou.

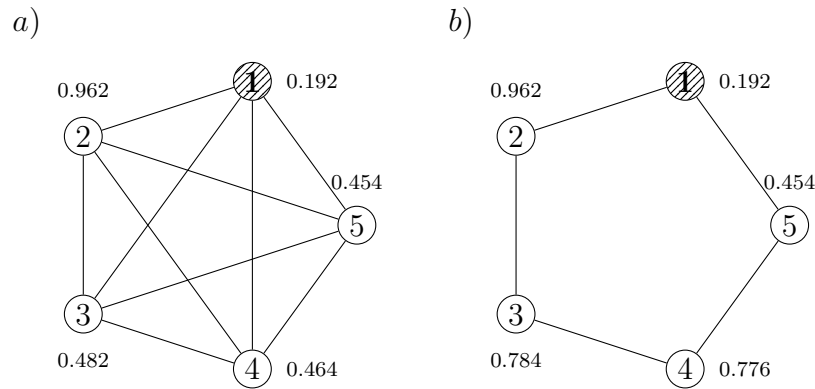


Figura 5.12: Grafos  $K_5$  (a) e  $\Lambda_5$  (b) com os valores atribuídos aos vértices no instante de tempo  $n = 1$ . Note que os valores atribuídos aos vértices 3 e 4 de  $G_N$  são diferentes dos valores atribuídos aos seus correspondentes em  $K_N$ .

Então, alteramos as distribuições do vértice minimal e dos seus vizinhos para  $U(0, 1)$ . No grafo  $\Lambda_5$  (Figura 5.13.b), os vértices que não são vizinhos do vértice minimal terão distribuição  $U(y_i \vee 0.192, 1)$ . Portanto, a distribuição dos vértices 1, 2 e 5 será  $U(0, 1)$ , enquanto os vértices 3 e 4 terão distribuição  $U(0.583, 1)$ , já que  $y_3 = y_4 = 0.583$ .

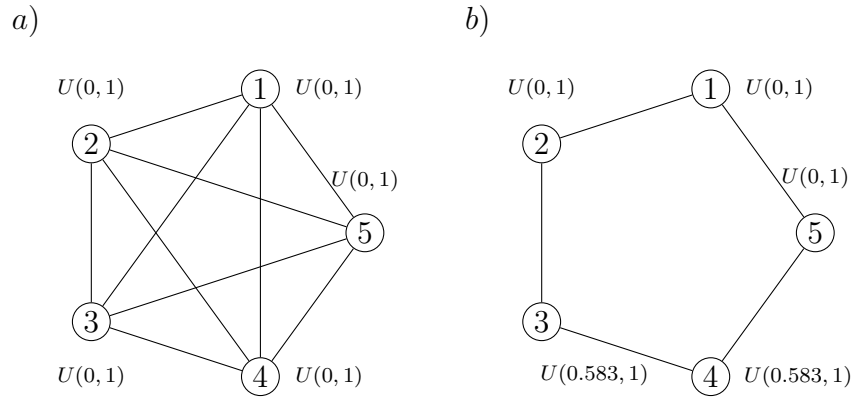


Figura 5.13: Grafos  $K_5$  (a) e  $\Lambda_5$  (b) ao final do instante de tempo  $n = 1$ . Note que os vértices têm exatamente as mesmas distribuições que tinham após a primeira atualização (Figura 5.11).

Seja  $(u_1, u_2, u_3, u_4, u_5) = (0.755, 0.386, 0.387, 0.299, 0.353)$  a amostra no instante de tempo  $n = 2$ . No grafo  $K_5$  (Figura 5.14.a), o valor atribuído ao vértice  $i$  é  $u_i$ ,  $i = 1, 2, 3, 4, 5$ . No grafo  $\Lambda_5$  (Figura 5.14.b), aplicamos a função  $F_{y_i}$  em sua respectiva componente da amostra. Portanto,

$$F_0(u_i) = (1 - 0)u_i + 0 = u_i, i = 1, 2, 5$$

$$F_{0.583}(u_i) = (1 - 0.583)u_i + 0.583 = 0.417u_i + 0.583, i = 3, 4.$$

Assim, obtemos os valores dos vértices em  $\Lambda_5$ :  $(0.755, 0.386, 0.744, 0.708, 0.353)$ .

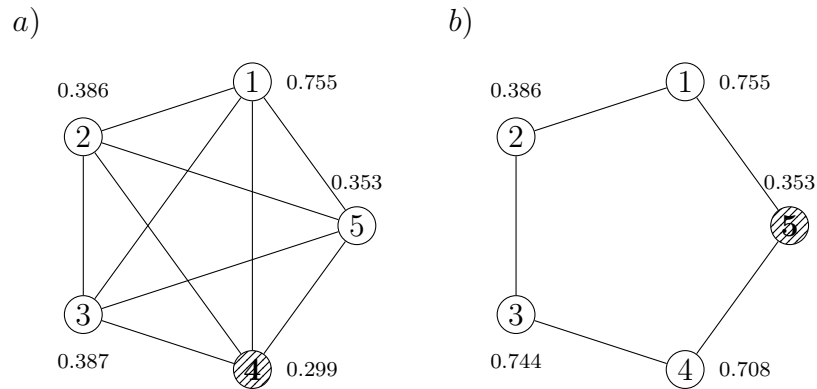


Figura 5.14: Grafos  $K_5$  (a) e  $\Lambda_5$  (b) com os valores atribuídos aos vértices no instante de tempo  $n = 2$ . Observe que os vértices minimais nos grafos são diferentes.

Agora, localizamos o vértice minimal nos processos. Veja que os vértices minimais são diferentes. No grafo  $K_5$  (Figura 5.15.a), o vértice minimal é o vértice 4, com valor igual a 0.299, enquanto que o vértice minimal em  $\Lambda_5$  é o vértice 5, com valor igual a 0.353. A avalanche de nenhum dos dois processos terminou. Veja que em ambos os processos, o valor do vértice minimal é menor que  $b$ . Localizados os vértices minimais, alteramos suas distribuições e as distribuições dos vértices vizinhos para  $U(0, 1)$ . No grafo  $\Lambda_5$  (Figura 5.15.b), os vértices que não são vizinhos do vértice minimal terão distribuição  $U(y_i \vee 0.353, 1)$ . Portanto, a distribuição dos vértices 1, 4 e 5 será  $U(0, 1)$ , a distribuição do vértice 2 será  $U(0.353, 1)$ , pois  $y_2 = 0$ , e a distribuição do vértice 3 será  $U(0.583, 1)$ , pois  $y_3 = 0.583$ .

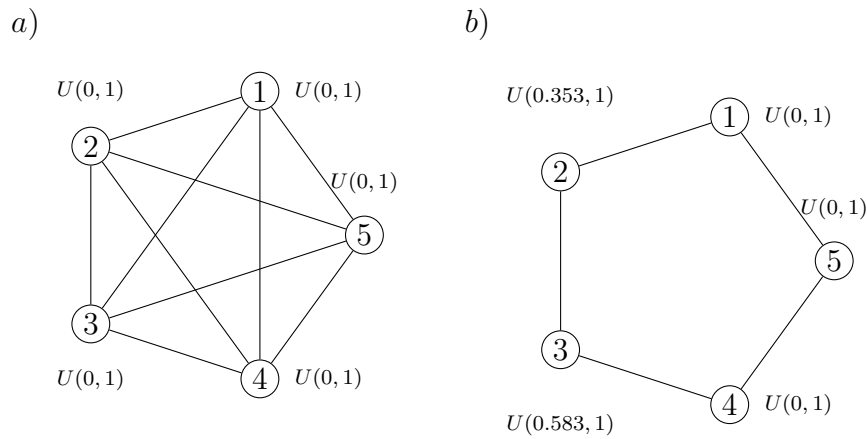


Figura 5.15: Grafos  $K_5$  (a) e  $\Lambda_5$  (b) ao final do instante de tempo  $n = 2$ .

Vamos dividir o próximo instante de tempo em 2 casos: no primeiro caso, mostraremos um exemplo a avalanche termina em  $K_5$  e, conseqüentemente, termina em  $\Lambda_5$ ; no segundo caso, mostraremos que a avalanche pode terminar em  $\Lambda_5$ , mas não necessariamente terminar em  $K_5$ .

• **Caso 1:** Seja  $(u_1, u_2, u_3, u_4, u_5) = (0.682, 0.590, 0.791, 0.838, 0.596)$  a amostra no instante de tempo  $n = 3$ . No grafo  $K_5$  (Figura 5.16.a), o valor atribuído ao vértice  $i$  é  $u_i$ ,  $i = 1, 2, 3, 4, 5$ . No grafo  $\Lambda_5$  (Figura 5.16.b), aplicamos a função  $F_y$  em cada componente da amostra. Portanto,

$$F_0(u_i) = (1 - 0)u_i + 0 = u_i, i = 1, 4, 5$$



$$F_{0.353}(u_2) = (1 - 0.353)u_2 + 0.353 = 0.735$$

$$F_{0.583}(u_3) = (1 - 0.583)u_3 + 0.583 = 0.913.$$

Assim, obtemos os valores dos vértices em  $\Lambda_5$ : (0.682, 0.735, 0.913, 0.838, 0.596). O próximo passo é localizar os vértices minimais nos processos. No grafo  $K_5$ , o vértice minimal é o vértice 2, com valor igual a 0.590 e no grafo  $\Lambda_5$ , o vértice minimal é o vértice 5, com valor igual a 0.596. Em ambos os grafos, o valor do vértice minimal é maior que o limitante (= 0.583), o que implica que as avalanches terminam. Após o término das avalanches, os valores continuam associados aos vértices, isto é, não temos mais uma distribuição associada aos vértices. Este caso é um exemplo de que, se a avalanche termina no grafo  $K_5$ , então a avalanche termina no grafo  $\Lambda_5$ . Esse resultado será provado mais adiante.

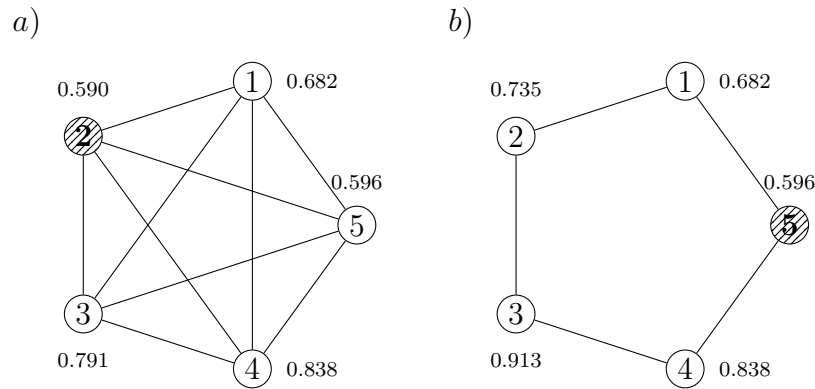


Figura 5.16: Grafos  $K_5$  (a) e  $\Lambda_5$  (b) ao final da avalanche. Os valores atribuídos aos vértices em  $\Lambda_5$  são fixados e a duração é igual a 4.

• **Caso 2:** Seja  $(u_1, u_2, u_3, u_4, u_5) = (0.682, 0.358, 0.791, 0.838, 0.596)$  a amostra no instante de tempo  $n = 3$ . Repare que trocamos apenas o valor de  $u_2$ , em relação ao caso 1. No grafo  $K_5$  (Figura 5.17.a), o valor atribuído ao vértice  $i$  é  $u_i$ ,  $i = 1, 2, 3, 4, 5$ . No grafo  $\Lambda_5$  (Figura 5.17.b), aplicamos a função  $F_y$  em cada componente da amostra. Portanto,

$$F_0(u_i) = (1 - 0)u_i + 0 = u_i, i = 1, 4, 5$$

$$F_{0.353}(u_2) = (1 - 0.353)u_2 + 0.353 = 0.585$$

$$F_{0.583}(u_3) = (1 - 0.583)u_3 + 0.583 = 0.913.$$

Assim, obtemos os valores dos vértices em  $\Lambda_5$ : (0.682, 0.585, 0.913, 0.838, 0.596). O próximo passo é localizar os vértices minimais nos processos. No grafo  $K_5$ , o vértice minimal é o vértice 2, com valor igual a 0.358 e no grafo  $\Lambda_5$ , o vértice minimal também é o vértice 2, mas com valor igual a 0.585. Veja que o valor do vértice minimal em  $K_5$  é menor que o limitante, o que não acontece em  $\Lambda_5$ . Este caso é um exemplo que mostra que avalanche pode terminar em  $\Lambda_5$ , mas não terminar em  $K_5$ . Esse caso sugere qual é a direção da dominação estocástica entre os processos.

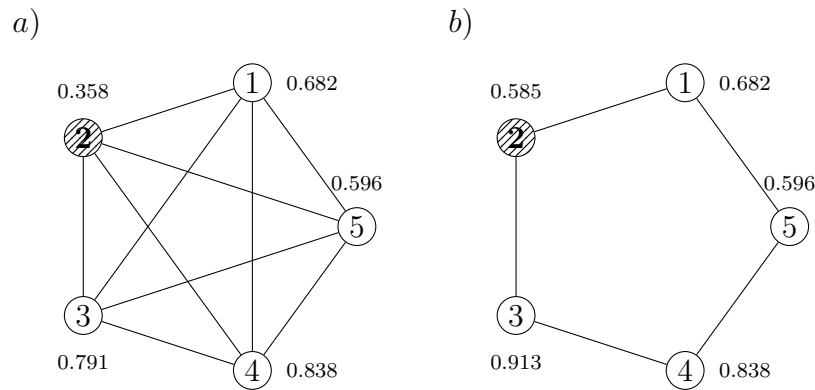


Figura 5.17: Grafos  $K_5$  (a) e  $\Lambda_5$  (b) com os valores atribuídos aos vértices no instante de tempo  $n = 3$ . Em  $\Lambda_5$ , a avalanche será encerrada e os valores fixados, pois todos os valores atribuídos aos vértices estão acima de  $b = 0.583$ . Já em  $K_5$ , a avalanche ainda não acabou e, portanto, o próximo passo seria atribuir distribuições aos vértices e descartar os valores.

Para a descrição geral, considere que estamos em um instante de tempo  $n$ , para  $n > 0$ . No instante de tempo  $n = 0$ , o limitante é o menor valor gerado pela amostra inicial e será denotado por  $b$ . Considere os vértices  $v_{K,i} \in K_N$  e o seu correspondente,  $v_{G,i} \in G_N$ , onde  $1 \leq i \leq N$  e sejam  $U(0, 1)$  e  $U(y_i, 1)$  suas respectivas distribuições. Sem perda de generalidade, podemos usar a  $i$ -ésima variável aleatória da sequência  $\{U_i\}_{i=1}^N$  para determinar a distribuição de  $v_{K,i}$ . Por outro lado, a distribuição de  $v_{G,i}$  nem sempre será  $U(0, 1)$ .

Portanto, usaremos a mesma sequência de variáveis aleatórias, mas aplicaremos cada função  $F_{y_i}$  em sua respectiva  $U_i$ . E como  $U_i \sim U(0, 1)$ , vale que  $F_{y_i}(U_i) \sim U(y_i, 1)$ . Veja que estamos usando a mesma sequência de variáveis aleatórias para governar os dois processos e as distribuições marginais coincidem com as respectivas distribuições dos

processos. Usando as distribuições, geramos os valores que serão atribuídos aos vértices para que os processos sejam executados. Seja  $m_G$  o valor atribuído ao vértice minimal em  $G_N$  e  $m_K$  o valor atribuído ao vértice minimal em  $K_N$ . Como o grafo  $K_N$  é completo, todos os vértices terão suas distribuições atualizadas em todos os instantes de tempo para  $U(0, 1)$ . Já no grafo  $G_N$ , aplicamos as regras do modelo Bak-Sneppen sem memória:

1. Se  $m_K > b$ : as avalanche em ambos os grafos é encerrada e os valores atribuídos aos vértices em  $G_N$  são fixados.
2. Se  $m_K \leq b$ : analisamos os valores atribuídos aos  $N$  vértices de  $G_N$ . Se  $m_G > b$ , apenas o processo no grafo  $G_N$  seria encerrado. Caso contrário,
  - Se  $v_{G,i}$  não é o vértice minimal: alteramos a sua distribuição, condicionando ao fato de que seu valor deve ser maior que  $m_G$ . E já vimos na Seção 2.3 que a distribuição deste vértice será  $U(y_i \vee m_G, 1)$ .
  - Se  $v_{G,i}$  é o vértice minimal: alteramos a sua distribuição e a distribuição dos vértices vizinhos para  $U(0, 1)$ .

Atualizadas as distribuições, geramos os valores que serão atribuídos aos vértices dos processos e voltamos ao passo 1. A descrição do acoplamento está completa.

Agora, vamos enunciar e demonstrar duas proposições que completam a prova do teorema.

**Proposição 5.3.1.** *Em qualquer instante de tempo  $t$  e  $\forall i$ , a distribuição de um vértice  $v_{G,i} \in G_N$  é estocasticamente maior que a distribuição de seu vértice correspondente,  $v_{K,i} \in K_N$ .*

*Demonstração.* Vale ressaltar que as distribuições devem ser condicionadas à mesma sequência de variáveis aleatórias, que neste caso é  $\{U_i\}_{i=1}^N$ . A prova é feita por indução em  $t$ . Vamos assumir que as duas avalanches ainda não terminaram. Quando  $n = 0$ , a afirmação da proposição é verdadeira. Por definição, as distribuições iniciais de todos os vértices em ambos os processos é  $U(0, 1)$ . Suponha que vale a afirmação da proposição para  $t = n$ . Para isso, considere um vértice  $v_{K,i} \in K_N$ , com distribuição  $U(0, 1)$ , e seu

vértice correspondente,  $v_{G,i} \in G_N$ , com distribuição  $U(y_i, 1)$ , onde  $y_i \geq 0$ . No instante de tempo  $t = n + 1$ , geramos os valores que serão associados aos vértices de  $K_N$  e  $G_N$ . Uma vez associados, localizamos o vértice minimal de cada processo. A distribuição do vértice  $v_{K,i}$  será  $U(0, 1)$ , pois  $v_{K,i}$  sempre será o vértice minimal ou vizinho do vértice minimal. Defina  $m_G$  o valor do vértice minimal do grafo  $G_N$ . Se  $v_{G,i}$  é o vértice minimal ou é vizinho do vértice minimal, então sua distribuição será atualizada para  $U(0, 1)$ , e a afirmação da proposição é verdadeira para  $t = n + 1$ . Agora, suponha que  $v_{G,i}$  não seja o vértice minimal e nem vizinho do vértice minimal. Neste caso, sua distribuição seria atualizada para  $U(y_i \vee m_G, 1)$ . Por hipótese,  $y_i \geq 0$  e  $m \geq 0$ , o que completa a prova da proposição.

□

**Proposição 5.3.2.** *A avalanche no modelo Bak-Sneppen no grafo  $G_N$  não pode terminar depois da avalanche no modelo Bak-Sneppen no grafo  $K_N$ . Em outras palavras,  $E_j(D_N^B) \leq E_j(D_{K_N}^B)$ .*

*Demonstração.* Pela proposição anterior, em qualquer instante de tempo, a distribuição de um vértice no grafo  $G_N$  é estocasticamente maior que a distribuição de seu vértice correspondente no grafo  $K_N$ . Portanto, se a avalanche em  $K_N$  é encerrada, isso implica que a avalanche em  $G_N$  também é ou já foi encerrada. Consequentemente, se  $E_j(D_{K_N}^B) < \infty$ , então  $E_j(D_N^B) < \infty$ , o que completa a prova da proposição.

□

Consequentemente, a prova do Teorema 5.1.2 está completa.

# A Apêndice

## A.1 Códigos das simulações

Os códigos abaixo foram produzidos no software *R* e a implementação dos métodos foi baseada no algoritmo descrito na Introdução. Para o modelo Bak-Sneppen modificado, foi feita a alteração descrita no início da Seção 3.1.

### Modelo Bak-Sneppen em um círculo

```

1 #dinamica do modelo Bak-Sneppen em um circulo com N vertices
2
3 #numero de vertices e numero de atualizacoes
4 N <- 500
5 n <- 1000000
6
7 #gera uma amostra de N uniformes (0,1)
8 amostra <- runif(N,0,1)
9
10 #execucao das atualizacoes
11 for (i in 1:n){
12
13     #localiza o vertice minimal
14     indice_m <- which.min(amostra)
15
16     #altera a aptidao do vertice minimal
17     amostra[indice_m] <- runif(1,0,1)
18
19     #altera as aptidoes dos dois vizinhos do vertice minimal
20     if(indice_m > 1) amostra[indice_m-1] <- runif(1,0,1)

```

```
21     else amostra[N] <- runif(1,0,1)
22
23     if(indice_m < N) amostra[indice_m+1] <- runif(1,0,1)
24     else amostra[1] <- runif(1,0,1)
25
26 }
27
28 #desenha o grafico
29 plot(amostra,ylim=c(0,1))
30 abline(h=2/3)
```

### Modelo Bak-Sneppen modificado em um grafo finito

```
1 #dinamica do modelo Bak-Sneppen modificado em grafo com N vertices
2
3 #numero de vertices e numero de atualizacoes
4 N <- 500
5 n <- 1000000
6
7 #indices dos vertices
8 indices <- 1:N
9
10 #gera uma amostra de N uniformes (0,1)
11 amostra <- runif(N,0,1)
12
13 for (i in 1:n) {
14
15     #localiza o vertice minimal
16     indice_m <- which.min(amostra)
17
18     #altera a aptidao do vertice minimal
19     amostra[indice_m] <- runif(1,0,1)
20
```

```
21     #sorteia um outro vertice
22     indice_r = sample(indices[indices!=indice_m],1)
23
24     #altera a aptidao do vertice sorteado
25     amostra[indice_r] <- runif(1,0,1)
26
27 }
28
29 #desenha o grafico
30 plot(amostra,ylim=c(0,1))
31 abline(h=1/2)
```

## Referências

- [1] Bak, P. and Sneppen, K. (1993) Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.*, **71**, 4083–4086.
- [2] Gillett, A. (2007) *Phase transitions in Bak–Sneppen avalanches and in a continuum percolation model*. Ph.D. thesis, Vrije Universiteit Amsterdam.
- [3] Gillett, A., Meester, R. and Nuyens, M. (2005) Bounds for avalanche critical values of the Bak–Sneppen model. *Markov Process. Relat. Fields*, **12**, 679–694.
- [4] Gillett, A., Meester, R. and van der Wal, P. (2006) Maximal avalanches in the Bak–Sneppen model. *Journal of Applied Probability*, **43**, 840–851.
- [5] Grassberger, P. (1995) The Bak–Sneppen model for punctuated evolution. *Physics Letters A*, **200**, 277–282.
- [6] Grimmett, G. (1999) *Percolation*. Springer, 2<sup>nd</sup> edition.
- [7] Grimmett, G. and Marstrand, J. M. (1990) The supercritical phase of percolation is well behaved. *Proceedings: Mathematical and Physical Sciences*, **430**, 439–457.
- [8] Hara, T. and Slade, G. (1995) The self-avoiding-walk and percolation critical points in high dimensions. *Combinatorics, Probability and Computing*, **4**, 197–215.
- [9] Heydenreich, M. and Hofstad, R. (2017) *Progress in High-Dimensional Percolation and Random Graphs*. CRM Short Courses. Springer International Publishing, 1<sup>st</sup> edition.
- [10] Hofstad, R. (2016) *Random Graphs and Complex Networks*, vol. 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press.
- [11] Lebensztayn, E. (2002) *O modelo de percolação em grafos: Um estudo de condições*. Master’s thesis, USP.



- 
- [12] Meester, R. and Quant, C. (2005) Connections between ‘self-organised’ and ‘classical’ criticality. *Markov Processes and Related Fields*, **11**, 355–370.
- [13] Meester, R. and Sarkar, A. (2012) Rigorous self-organised criticality in the modified Bak–Sneppen model. *Journal of Statistical Physics*, **149**, 964–968.
- [14] Meester, R. and Znamenski, D. (2003) Limit behavior of the bak–sneppen evolution model. *The Annals of Probability*, **31**, 1986–2002.
- [15] Meester, R. and Znamenski, D. (2004) Critical thresholds and the limit distribution in the Bak–Sneppen model. *Communications in Mathematical Physics*, **246**, 63–86.
- [16] Newman, M. (1996) Self-organized criticality, evolution and the fossil extinction record. *Proceedings of the Royal Society B: Biological Sciences*, **263**, 1605–1610.
- [17] Schlemm, E. (2012) Asymptotic fitness distribution in the bak–sneppen model of biological evolution with four species. *Journal of Statistical Physics*, **148**, 191–203.
- [18] Schlemm, E. (2015) A differential equation for the asymptotic fitness distribution in the Bak–Sneppen model with five species. *Mathematical Biosciences*, **267**, 53–60.
- [19] Tutte, W. (1984) *Graph Theory*. Section, Combinatorics. Longman Higher Education.