

# Chapter 2

## Discontinuous Explosive Percolation with Multiple Giant Components

### 2.1 Introduction

The percolation phase transition models the onset of large-scale connectivity in lattices or networks, in systems ranging from porous media, to resistor networks, to epidemic spreading [1–4]. Percolation was considered a robust second-order transition until a variant with a choice between edges was shown to result in a seemingly discontinuous transition [5]. Subsequent studies have shown similar results for scale-free networks [6, 7], lattices [8, 9], local cluster aggregation models [10], single-edge addition models [11, 12], and models which control only the largest component [13]. It seems a fundamental requirement that in the subcritical regime the evolution mechanism produces many clusters, which are relatively large, though sublinear, in size [10, 11, 14]. Most recently, the notions of “strongly” versus “weakly” discontinuous transitions have been introduced [15], with the model studied in [5] showing weakly discontinuous characteristics, while an idealized deterministic “most explosive” model [14, 16] is strongly discontinuous. Here, we analyze and extend a related model by Bohman, Frieze, and Wormald (BFW) [17], which predates the more recent work, and show that surprisingly [18], multiple stable giant components can coexist and that the percolation transition is strongly discontinuous.

### 2.2 Results

The “most explosive” deterministic process [14–16] begins with  $n$  isolated nodes, with  $n$  set to a power of two for convenience. Edges that connect pairs of isolated nodes, creating components of size  $k = 2$ , are added sequentially until no isolated nodes remain. The cutoff  $k$  is then doubled and edges leading to components of size  $k = 4$  are added sequentially, until all components have size  $k = 4$ .  $k$  is then doubled yet again and the process iterated. By the end of the phase when  $k = n/2$ , only two components remain, each with size  $n/2$ . The addition of the next edge connects those

two components during which the size of the largest component jumps in value by  $n/2$ . Letting  $t$  denote the number of edges added to the graph, we define the critical  $t_c$  as the single edge whose addition produces the largest jump in size of the largest component denoted  $\Delta C_{\max}$ ; here  $t_c = (n - 1)$ , with  $\Delta C_{\max} = n/2$ .

The BFW model begins with a collection of  $n$  isolated nodes (with  $n$  any integer) and also proceeds in phases starting with  $k = 2$ . Edges are sampled one-at-a-time, uniformly at random from the complete graph. If an edge would lead to formation of a component of size less than or equal to  $k$  it is accepted. The edge is otherwise rejected provided that the fraction of accepted edges is greater than or equal to a function  $g(k) = 1/2 + (2k)^{-1/2}$ . Once the next edge rejection event would lead to the fraction of accepted nodes dropping below  $g(k)$ , the phase is augmented to  $k + 1$ , with explicit details given below. Note asymptotically,  $\lim_{k \rightarrow \infty} g(k) = \alpha$  with  $\alpha = 1/2$ .

BFW established rigorous results whereby setting  $g(200) = 1/2$ , all components are no larger than  $k = 200$  nodes (i.e., no giant component exists) when  $m = 0.96689n$  edges out of  $2m$  sequentially sampled random edges have been added to graph. They further establish that a giant component must exist by the time  $m = c^*n$  out of  $2m$  sampled edges have been added, with  $c^* \in [0.9792, 0.9793]$ . Yet, they did not analyze the details of the percolation transition. We show that their model leads to the simultaneous emergence of two giant components (each of size greater than 40% of all the nodes), and show analytically the stability of the two giants throughout the subsequent graph evolution. We then generalize the BFW model by allowing the asymptotic fraction of accepted edges,  $\alpha$ , to be a parameter and show that  $\alpha$  determines the number of stable giant components that emerge and that, in general, smaller values of  $\alpha$  lead to a more delayed and more explosive transition.

Stating the BFW algorithm in detail, let  $k$  denote the stage and  $n$  the number of nodes. Let  $u$  denote the total number of edges sampled,  $A$  the set of accepted edges (initialized to  $A = \emptyset$ ), and  $t = |A|$  the number of accepted edges. At each step  $u$ , an edge  $e_u$  is sampled uniformly at random from the complete graph generated by the  $n$  nodes, and the following algorithm iterated:

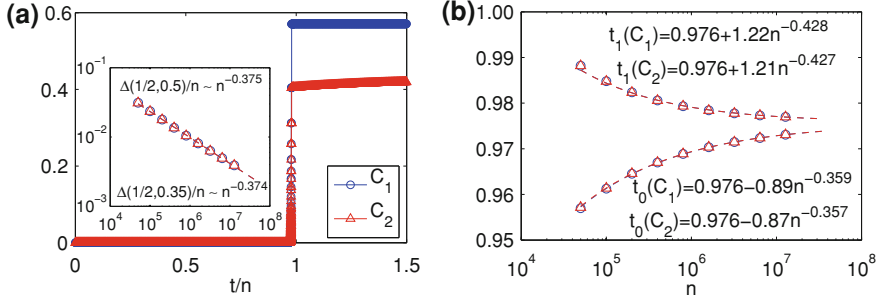
```

Set  $l = \text{maximum size component in } A \cup \{e_u\}$ 
if ( $l \leq k$ ) {
     $A \leftarrow A \cup \{e_u\}$ 
     $u \leftarrow u + 1$  }
else if ( $t/u < g(k)$ ) {  $k \leftarrow k + 1$  }
else {  $u \leftarrow u + 1$  }

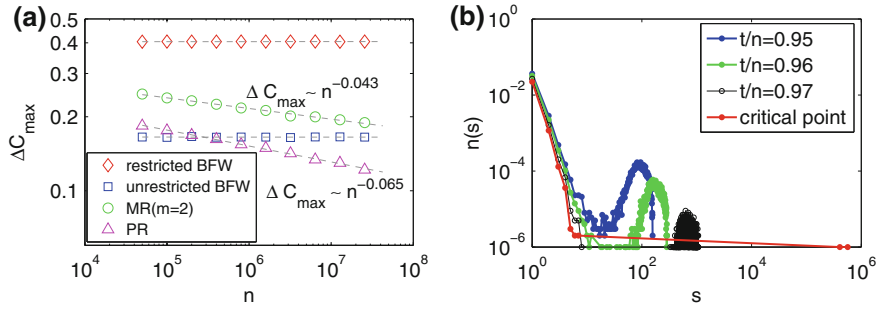
```

Thus, while  $t/u < g(k)$ ,  $k$  is augmented repeatedly until either  $k$  becomes large enough that edge  $e_u$  is accepted or  $g(k)$  decreases sufficiently that edge  $e_u$  can be rejected at which point step  $u$  ends. Note  $g(k) = 1$  requires that all edges be accepted, equivalent to Erdős-Rényi [19].

We numerically implement the BFW model, and measure the fraction of nodes in both the largest and second largest component, denoted  $C_1$  and  $C_2$ , as a function of edge density  $t/n$ . As shown in Fig. 2.1a, two giants appear at the same critical point and remain distinct. To establish that the BFW model shows a seemingly



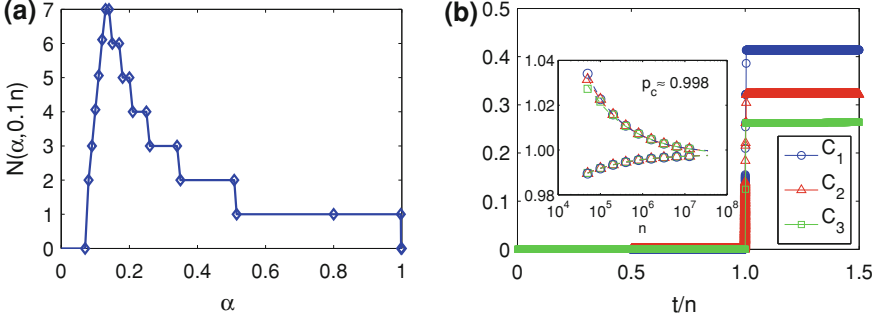
**Fig. 2.1** The BFW model. **a**  $C_1$  and  $C_2$  versus edge density,  $t/n$ , for  $n = 10^6$ , showing the emergence of two stable giant components. Inset:  $\Delta(1/2, 0.5)/n \sim n^{-0.375}$ ,  $\Delta(1/2, 0.35)/n \sim n^{-0.374}$  for  $C_1$  and  $C_2$  versus  $n$ . **b** The lower and upper boundaries of  $\Delta/n$  for both  $C_1$  and  $C_2$ , where data points are averages over 200 to 2,000 realizations and dashed lines are best fits



**Fig. 2.2** **a**  $\Delta C_{\max}$ , the biggest single edge increase in  $C_1$ , is independent of system size  $n$  for BFW and for the restricted BFW process, indicating these transitions are strongly discontinuous. **b** Evolution of the distribution of  $n(s)$ , the fraction of components of size  $s$ , for the BFW model

discontinuous transition we apply the numerical method introduced in [5]. Let  $t_0^{C_i}(n)$  denote the last accepted edge for which  $C_i n \leq n^\gamma$  and  $t_1^{C_i}(n)$  the first accepted edge with  $C_i n \geq A n$ , where  $n$  is system size and  $\gamma$  and  $A$  are parameters.  $\Delta^i(\gamma, A) = t_1^{C_i}(n) - t_0^{C_i}(n)$  denotes the number of accepted edges required between these two points. As shown inset to Fig. 2.3a,  $\Delta^i(\gamma, A)/n \propto n^{-0.375}$ , showing that  $t_0^{C_i}(n)/n$  and  $t_1^{C_i}(n)/n$  converge to same limiting value  $t_c/n = 0.976$  (Fig. 2.1b) for both  $C_1$  and  $C_2$ .

The discontinuous nature is more explicit in Fig. 2.2a showing  $\Delta C_{\max}$ , the largest increase of the largest component due to addition of a single edge, versus  $n$  (blue squares are BFW).  $\Delta C_{\max}$  is independent of  $n$  (strongly discontinuous). Essentially, the same value of  $\Delta C_{\max}$  is always observed and results from the second and third components merging together to overtake what was previously the largest. The model studied in [5] (PR) shows a decrease with  $n$  (weakly discontinuous), with scaling  $n^{-0.065}$  as also recently observed in [12, 15].



**Fig. 2.3** **a** Number of stable giant components of size greater than  $0.1n$  versus  $\alpha$  for  $n = 10^6$ . **b** When  $\alpha = 0.3$ , three giant components emerge simultaneously. Inset shows convergence of the upper and lower boundaries of  $\Delta/n$  for  $C_1, C_2, C_3$

The key to coexisting multiple giant components is the high probability of sampling internal-cluster links in the super-critical region which, by definition, do not increase the component size. We formalize this by first introducing a function  $P(k, t, n)$  defined as the probability of sampling a random link, which leads to a component of size no larger than stage  $k$  at step  $t$  for system size  $n$ :

$$P(k, t, n) = \sum_i C_i^2 + 2 \sum_{C_i + C_j \leq k/n} C_i C_j \quad (2.1)$$

where  $C_i$  denotes the fraction of nodes in component  $i$ . The first term on the right-hand side is the probability of randomly sampling internal-cluster links in all components. The second term is the probability of sampling spanning-cluster links, which lead to a component of size no larger than  $k$ . This is valid for any configuration in phase  $k$ . We also consider  $S(k, n)$ , the probability of sampling random links, which lead to components of size no larger than  $k$  if all possible spanning-cluster links are added in stage  $k$ . Thus,

$$S(k, n) = \sum_i C_i^2. \quad (2.2)$$

(Note the values of the  $C_i$ 's in Eqs.(2.1) and (2.2) can differ from each other.) For any specific stage  $k$ , it is easy to show that  $P(k, t, n) \geq S(k, n)$  since, if  $t$  increases,  $P(k, t, n)$  can only decrease or stay the same. More explicitly, if an internal-cluster link is added then  $P(k, t, n)$  is invariant, while if a spanning-cluster link is added between components  $i$  and  $j$  then the first term increases by  $(C_i + C_j)^2 - (C_i^2 + C_j^2) = 2C_i C_j$  and second term decreases by at least  $2C_i C_j$ . (Additional decreases result if there exist components  $l$  satisfying  $C_i + C_l \leq k/n$ , but with  $(C_i + C_j) + C_l > k/n$ ).

Focusing now on the critical region, let  $k^*$  denote the value of  $k$  at  $t_c$ . Numerical results for a variety of system sizes show that at  $t_c$  when  $n > 10^6$ ,  $k^*/n \sim 0.570$ ,  $C_1 \sim 0.570$  and  $C_2 \sim 0.405$  with error bars of order  $\mathcal{O}(10^{-4})$  obtained over 30

to 300 realizations dependent on  $n$ . Thus, the remaining components have total size density  $\sum_{i \geq 3} C_i = 0.025$ . We can establish a uniform lower bound on  $S(k, n)$  for all  $k \geq k^*$  using the simple intuition that under the normalization condition  $\sum_i C_i = 1$ ,  $S(k, n) = \sum_i C_i^2$  is minimized when the number of components are as numerous as possible and of similar size. Given that  $\sum_{i \geq 3} C_i < C_1 - C_2$ , the lower bound on  $S(k, n)$  is if all small components connect to  $C_2$ . Then  $P(k, t, n) \geq S(k, n) \geq C_1^2 + (C_2 + \sum_{i \geq 3} C_i)^2 = 0.570^2 + (0.405 + 0.025)^2 \sim 0.510$ , so for any stage  $k \geq k^*(n)$ , we have

$$P(k, t, n) > \alpha = 1/2. \quad (2.3)$$

So for  $k \geq k^*$ , the expected fraction of accepted links approaches a positive value strictly larger than  $\alpha$ .

Having established that in expectation  $P(k, t, n) > \alpha$ , for  $k \geq k^*$ , we need to explicitly consider what happens if an edge connecting the two giant components is sampled in this regime. Here  $(C_1 + C_2) > k/n \geq k^*/n$  and, by definition,  $t/u \geq g(k)$ . Consider the case when edge  $e_{u+1}$  connects  $C_1$  and  $C_2$ . If  $t/(u+1) \geq g(k)$  the edge is simply rejected. But if  $t/(u+1) < g(k)$  either  $k$  needs to increase until the edge is accommodated, or (as we show next) a small increase in  $k$  quickly leads to  $t/(u+1) \geq g(k)$ . Setting  $t/u$  to the smallest value possible:

$$\frac{t}{u} = \frac{1}{2} + \sqrt{\frac{1}{2k}} \quad (2.4)$$

Differentiating both sides in Eq. (2.4) by  $k$  we find that

$$\frac{du}{d(k/n)} = \frac{1}{2\sqrt{2}(1/2 + \sqrt{1/2k})^2} \frac{nt}{k^{3/2}}. \quad (2.5)$$

After the critical point we know that  $t \sim \mathcal{O}(n)$  and the stage  $k \sim C_1 n \sim \mathcal{O}(n)$ . Thus, from Eq. (2.5) it follows that  $\frac{du}{d(k/n)} \sim \mathcal{O}(n^{1/2})$  as  $n \rightarrow \infty$  implying that an  $\mathcal{O}(n^{-1/2})$  increase in  $k/n$  results in  $t/(u+1) > g(k)$ , so the link which would lead to merging  $C_1$  and  $C_2$  is rejected and the two giant components are stable throughout the subsequent evolution. We verify this numerically. Letting  $\bar{k}(n)$  denote the largest value of the stage ever attained for system size  $n$ , we find  $(\bar{k}(n) - k^*(n))/n \sim n^{-\gamma}$  with  $\gamma = 0.46 \pm 0.03$ , and as  $n \rightarrow \infty$ ,  $k^*(n)/n$  and  $\bar{k}(n)/n$  converge to the same limiting value of approximately 0.570.

The BFW model samples edges uniformly at random from the complete graph. If we restrict the process to sampling only edges that span distinct clusters, we observe that two components with the same  $C_1 = 0.570$  and  $C_2 = 0.405$  values coexist for several edge additions before merging together. When they do merge the largest jump in  $C_1$ , equal to the size of  $C_2$ , occurs. This is a strongly discontinuous transition as shown in Fig. 2.2a (the red diamonds) with gap size equal to 0.405.

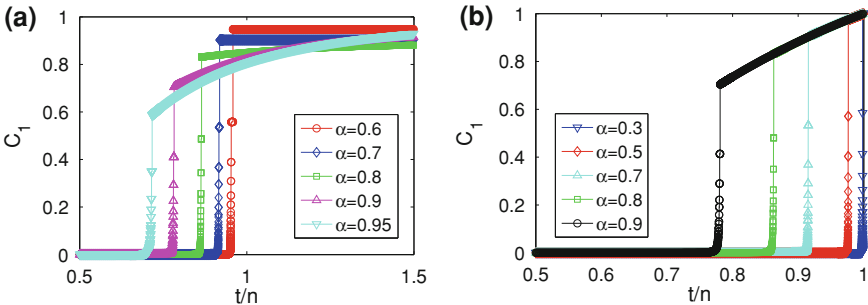
We now generalize the BFW model so that  $g(k) = \alpha + (2k)^{-1/2}$  (i.e., the asymptotic fraction of accepted links is now a parameter  $\alpha$ ). For the unrestricted process

(sampling from the complete graph) we find that  $\alpha$  controls the number of stable giant components. Let  $N(\alpha, m)$  denote the number of stable giant components with size larger than  $m$ , which appear at the critical point and remain throughout the subsequent evolution. Figure 2.3a shows  $N(\alpha, 0.1n)$  versus  $\alpha$  for the unrestricted process, with system size  $10^6$  and each data point averaged over 100 independent realizations (showing no fluctuations). As  $\alpha$  first decreases from  $\alpha = 1$ ,  $N(\alpha, 0.1n)$  increases, going from one giant to two at  $\alpha = 0.511 \pm 0.003$ . Then, once  $\alpha < 0.11$ ,  $N(\alpha, 0.1n)$  decreases. (Using a less stringent criteria that considers all macroscopic components  $C_i n > cn$  where  $c > 0$  a “giant”, then  $N(\alpha, cn)$  actually continues increasing.)

The same reasoning that applied to the original BFW model can be used here to show the stability of the multiple giants. Once  $k \geq k^*$ , in expectation  $P(k, t, n) > \alpha$ . Likewise, once  $k \geq k^*$ ,  $\frac{du}{d(k/n)} \sim \mathcal{O}(n^{1/2})$ , so  $k/n$  increases very slowly and the process frequently samples new links and rejects links that merge any two giants. For example, if  $\alpha = 0.3$ ,  $N(\alpha, 0.1n) = 3$  with  $C_1 = 0.414$ ,  $C_2 = 0.321$ ,  $C_3 = 0.265$ , so  $P(k, t, n) \geq C_1^2 + C_2^2 + C_3^2 \sim 0.345 > \alpha = 0.3$  when  $k \geq k^*(n)$ . See Fig. 2.3b for details.

In Fig. 2.4a we show the typical evolution for the unrestricted BFW process for various values of  $\alpha$  in the regime where only one giant component emerges. We measure the scaling window  $\Delta(\gamma, A)$ , as discussed earlier, and find that smaller  $\alpha$  leads to a more “explosive” transition in that  $A$  is larger and the scaling window shrinks more quickly. Explicitly, for  $\alpha = 0.7, 0.8, 0.9$  (and setting  $\gamma = 1/2$ ), we find respectively that  $A = 0.9, 0.8, 0.7$  and  $t_c \approx 0.915, 0.862, 0.780$ . This delayed and more explosive nature with smaller  $\alpha$  is intuitive in that the more links are rejected at each stage, the longer one stays in that stage, resulting in more components of size  $C_i n \sim k$ .

Figure 2.4b shows the analogous behavior for the restricted BFW process (where only edges that span components are considered). The delayed and more explosive nature of the transition with decreasing  $\alpha$  is also observed here. We also note that the location of  $t_c$  is not affected. For instance, for  $\alpha = 0.3, 0.5, 0.7, 0.8, 0.9$  we find



**Fig. 2.4** The results of varying  $\alpha$  on **a** the unrestricted BFW process in the regime where only one giant component emerges, and **b** the restricted BFW process. Both (a) and (b) show increasing delay and larger  $\Delta C_{\max}$  for smaller  $\alpha$

that  $t_c \approx 0.998, 0.976, 0.915, 0.862, 0.780$  for both the restricted and unrestricted processes.

The behavior of the restricted process can also be explained via Eq. (2.1). Here, because intra-cluster links are not allowed, the first term on the right-hand side vanishes. If the stage stops at some  $k_0 < n$ , then the second term on right side of Eq. (2.1) decreases to 0, which makes  $P(k_0, t, n) < \alpha$  and stage keeps on growing until the two giants merge together. The restriction on sampled links does not change the sublinear nature of  $\Delta(\gamma, A)$ . In fact, we find that  $\Delta(\gamma, A)$  for the general BFW model is sublinear in  $n$  for all  $\alpha \in (0, 0.97]$ , regardless of whether link-sampling is restricted or unrestricted, indicating the transition is discontinuous in all cases.

## 2.3 Summary and Discussion

In summary, we have analyzed the critical behavior of the BFW model and find that two stable giant components emerge at the same critical point in a strongly discontinuous transition (Fig. 2.2a). If we restrict the sampled links to only spanning-cluster edges, multiple giants coexist for a few moments until they merge together in a larger discontinuous jump (Fig. 2.2a) and ultimately only one giant component emerges. We further generalize BFW by making the asymptotic fraction of accepted links a parameter  $\alpha$ , and find that number of stable giant components increases while  $\alpha$  decreases.

The existence of multiple stable macroscopic components is surprising and unanticipated [18], and has not been previously observed in stochastic percolation. A model of cluster aggregation where largest clusters are occasionally “frozen” and prevented from growing, does lead to multiple giant components [20], however the freezing is imposed on the system. Simple algorithms that generate multiple giant components may create a new range of applications. In addition to providing insight and a potential mechanism for controlling gel sizes during polymerization [20], they may be useful for creating communication networks consisting of multiple large components operating on different frequencies or for analyzing epidemic infections simultaneously arising in distinct, independent groups. In the unrestricted process (the original BFW process) multiple links between two nodes and self-loops are allowed. It would be important to understand what happens when only links not yet added to the graph are sampled for finite networks. In the asymptotic size limit, there should be no difference as there are  $O(n^2)$  available edges and  $O(n)$  added edges.

The nature of the transition observed in [5] was recently analyzed using cluster aggregation models with choice, where a set of candidate edges are simultaneously inspected at each step [21]. The mechanism here, in contrast, inspects one edge individually at each time. As shown in Fig. 2.2a the models in [21] (labeled MR ( $m=2$ )) and in [5] (labeled PR) show weakly discontinuous transitions, where  $\Delta C_{\max}$  decreases with system size. In contrast both the restricted and unrestricted BFW models are strongly discontinuous, with a jump independent of system size. Finally, we show the evolution of the component size distribution for the original BFW model

in Fig. 2.2b ( $n(s)$  is the fraction of components of size  $s$ ). This bimodal distribution has a large right-hand tail, which deviates from a power law. Whether it would show a power law at  $t_c$  as  $n \rightarrow \infty$  is not obvious.

## References

1. Stauffer, D., Aharony, A.: Introduction to Percolation Theory. Taylor & Francis, London (1994)
2. Sahimi, M.: Applications of Percolation Theory. Taylor & Francis, London (1994)
3. Pastor-Satorras, R., Vespignani, A.: Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200–3203 (2001)
4. Moore, C., Newman, M.E.J.: Epidemics and percolation in small-world networks. *Phys. Rev. E* **61**, 5678–5682 (2000)
5. Achlioptas, D., D’Souza, R.M., Spencer, J.: Explosive percolation in random networks. *Science* **323**, 1453–1455 (2009)
6. Cho, Y.S., Kim, J.S., Park, J.: Kahng, B., Kim, D.: Percolation transitions in scale-free networks under the Achlioptas process. *Phys. Rev. Lett.* **103**, 135702 (2009)
7. Radicchi, F., Fortunato, S.: Explosive percolation in scale-free networks. *Phys. Rev. Lett.* **103**, 168701 (2009)
8. Ziff, R.M.: Explosive growth in biased dynamic percolation on two-dimensional regular lattice networks. *Phys. Rev. Lett.* **103**, 045701 (2009)
9. Ziff, R.M.: Scaling behavior of explosive percolation on the square lattice. *Phys. Rev. E* **82**, 051105 (2010)
10. D’Souza, R.M., Mitzenmacher, M.: Local cluster aggregation models of explosive percolation. *Phys. Rev. Lett.* **104**, 195702 (2010)
11. Cho, Y.S., Kahng, B., Kim, D.: Cluster aggregation model for discontinuous percolation transitions. *Phys. Rev. E* **81**, 030103(R) (2010)
12. Manna, S.S., Chatterjee, A.: A new route to explosive percolation. *Phys. A* **390**, 177–182 (2011)
13. Araújo, N.A.M., Herrmann, H.J.: Explosive percolation via control of the largest cluster. *Phys. Rev. Lett.* **105**, 035701 (2010)
14. Friedman, E.J., Landsberg, A.S.: Construction and analysis of random networks with explosive percolation. *Phys. Rev. Lett.* **103**, 255701 (2009)
15. Nagler, J., Levina, A., Timme, M.: Impact of single links in competitive percolation. *Nature Phys.* **7**, 265–270 (2011)
16. Rozenfeld, H.D., Gallos, L.K., Makse, H.A.: Explosive percolation in the human protein homology network. *Eur. Phys. J. B* **75**, 305–310 (2010)
17. Bohman, T., Frieze, A., Wormald, N.C.: Avoidance of a giant component in half the edge set of a random graph. *Random Struct. Algorithms* **25**, 432–449 (2004)
18. Spencer, J.: The giant component: The golden anniversary. *Not. AMS* **57**, 720–724 (2010)
19. Erdős, P., Rényi, A.: On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.* **5**, 17 (1960)
20. Ben-Naim, E., Krapivsky, P.L.: Percolation with multiple giant clusters. *J. Phys. A* **38**, L417–L423 (2005)
21. da Costa, R.A., Dorogovtsev, S.N., Goltsev, A.V., Mendes, J.F.F.: Explosive percolation transition is actually continuous. *Phys. Rev. Lett.* **105**, 255701 (2010)



Explosive Percolation in Random Networks

Chen, W.

2014, XV, 63 p. 22 illus., 9 illus. in color., Hardcover

ISBN: 978-3-662-43738-4