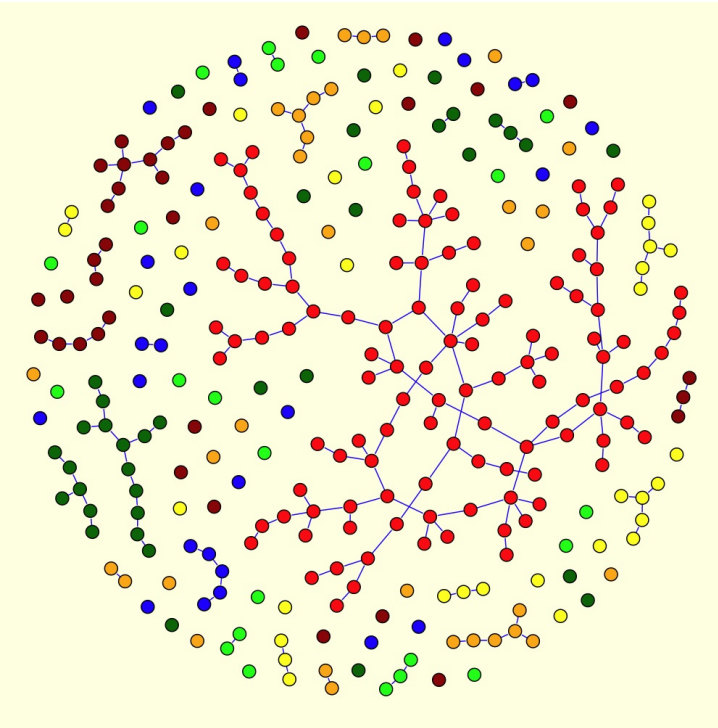
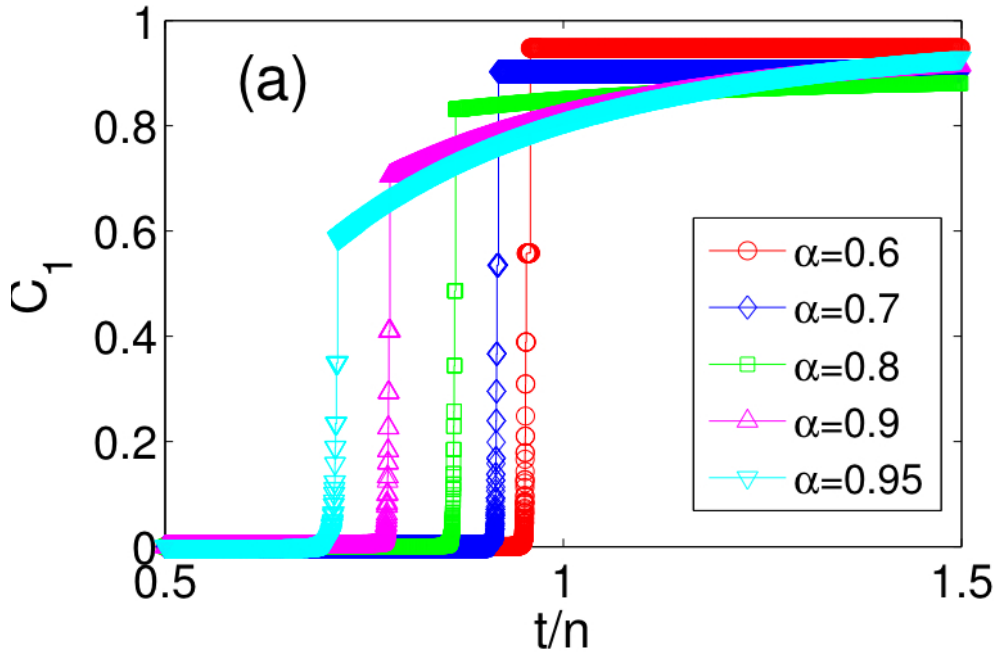
Explosive percolation:
What we've learned in the past two years

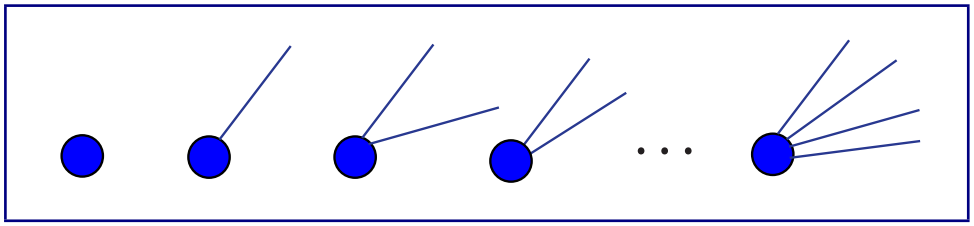




Raissa D'Souza
University of California, Davis
External Professor, **Santa Fe Institute**

Modeling networks as random graphs

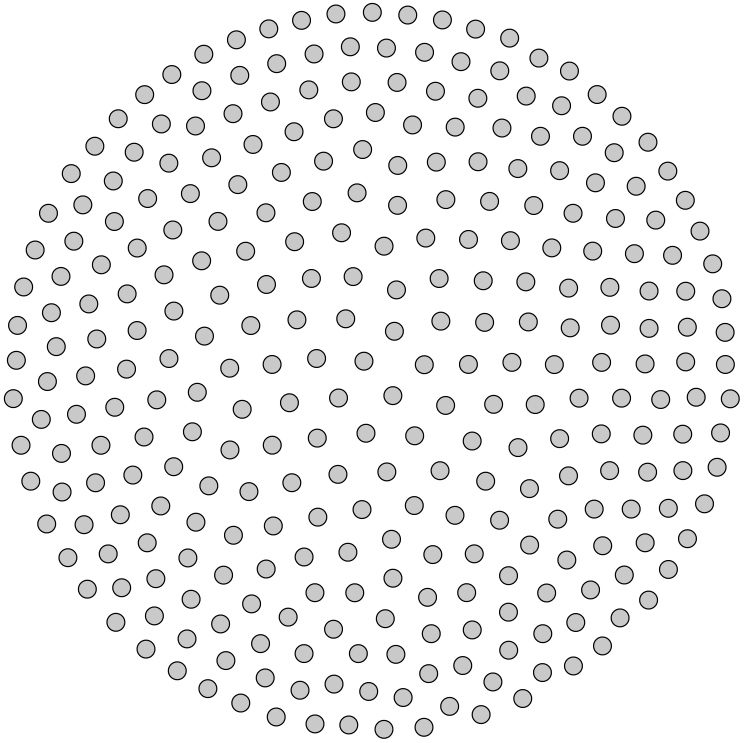
- Erdős and Rényi random graphs (1959, 1960). Phase transition.
- Configuration models (Bollobás 1980, Molloy and Reed *RSA* 1995). Enumerating over all networks with specified $\{p_i\}$.



- Preferential attachment (Barbási-Albert 1999, etc.)
- Growth by copying (Kumar, Raghavan, Rajagopalan, Sivakumar, Tomkins, Upfal *FOCS* 2000), including duplication/mutation (Vazquez, Flammini, Maritan, Vespignani, *ComPlexUs* 2003)
- Many more . . .

Building a random instance of a network, $G(N, p)$

- P. Erdős and A. Rényi, "On random graphs", *Publ. Math. Debrecen.* 1959.
- P. Erdős and A. Rényi, "On the evolution of random graphs", *Publ. Math. Inst. Hungar. Acad. Sci.* 1960.
- E. N. Gilbert, "Random graphs", *Annals of Mathematical Statistics*, 1959.

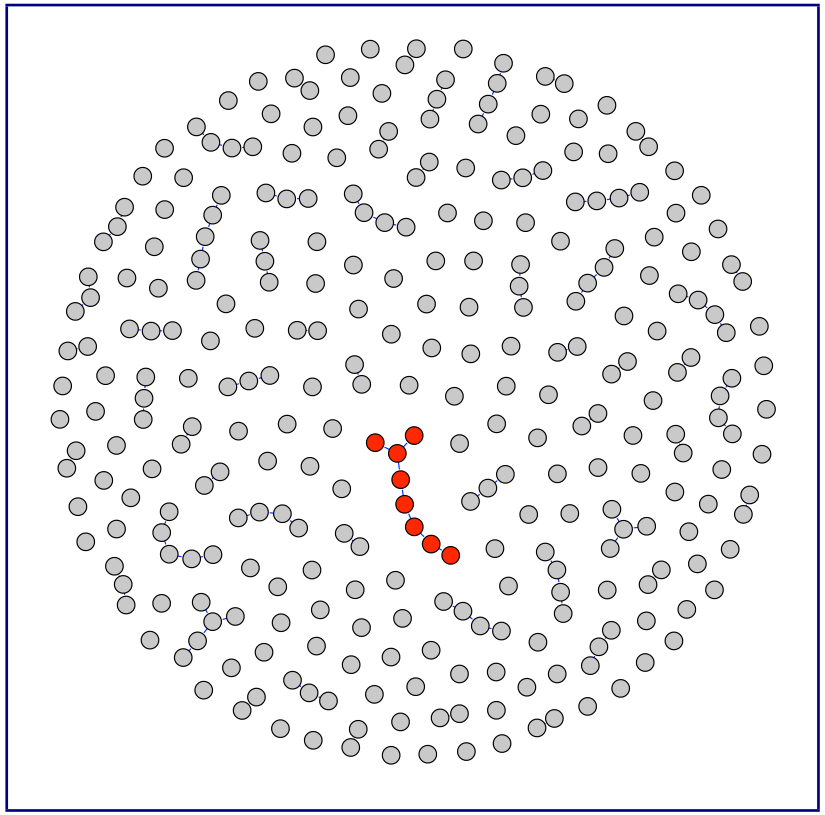


- Start with N isolated vertices.
- Consider each possible edge, and add it with probability p.

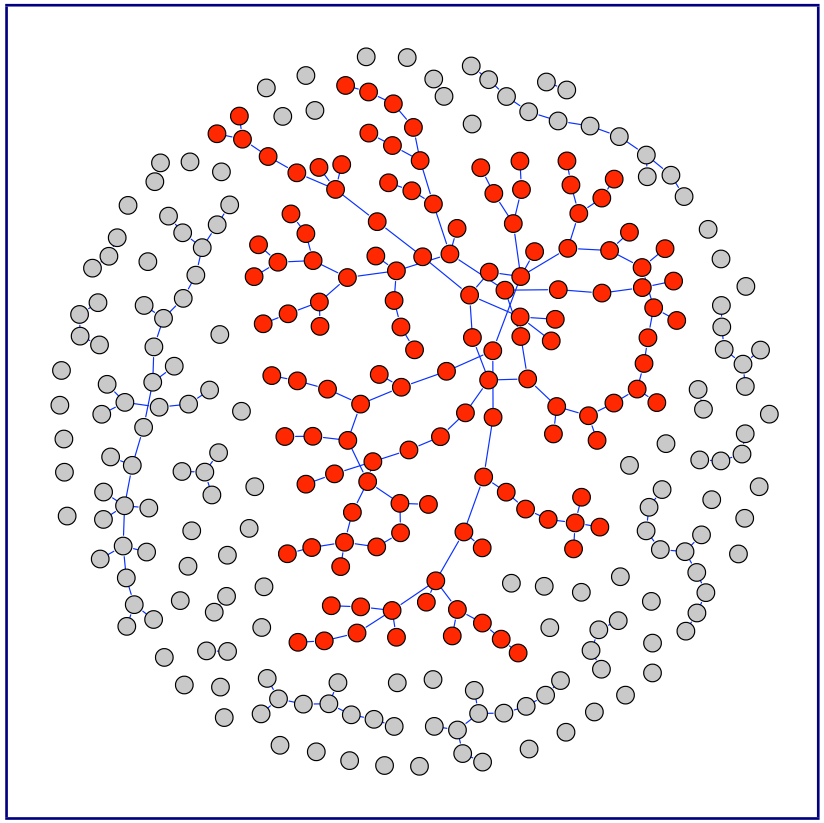
What does the resulting graph look like?

(Typical member of the ensemble)

G(N=300,p)

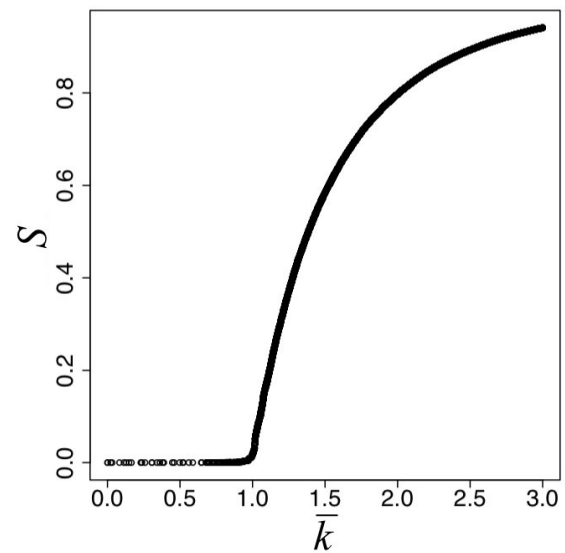


$p = 1/400 = 0.0025$



$p = 1/200 = 0.005$

Erdős-Rényi, Emergence of unique "giant component"



- $t < 1, \quad C_{\max} \sim \mathrm{O}(\ln n)$
- $t = 1, \quad C_{\max} = n^{2/3}$
- $t > 1, \quad C_{\max} \sim An,$ with $A > 1$
- The critical window

  Bollobás, *Trans. Amer. Math. Soc.*, **286** (1984).

  Luczak, *Random Structures and Algorithms*, **1** (1990).

$$t = 1 + \lambda n^{-1/3} \qquad \text{(where } t = 2e/n\text{)}$$

- Mean field critical exponents

  e.g., Grimmett, *Percolation*. 2nd Edition. Springer-Verlag. 1999.

$$\chi \sim (t_c - t)^{-\gamma}, \quad \text{with } \gamma = 1.$$

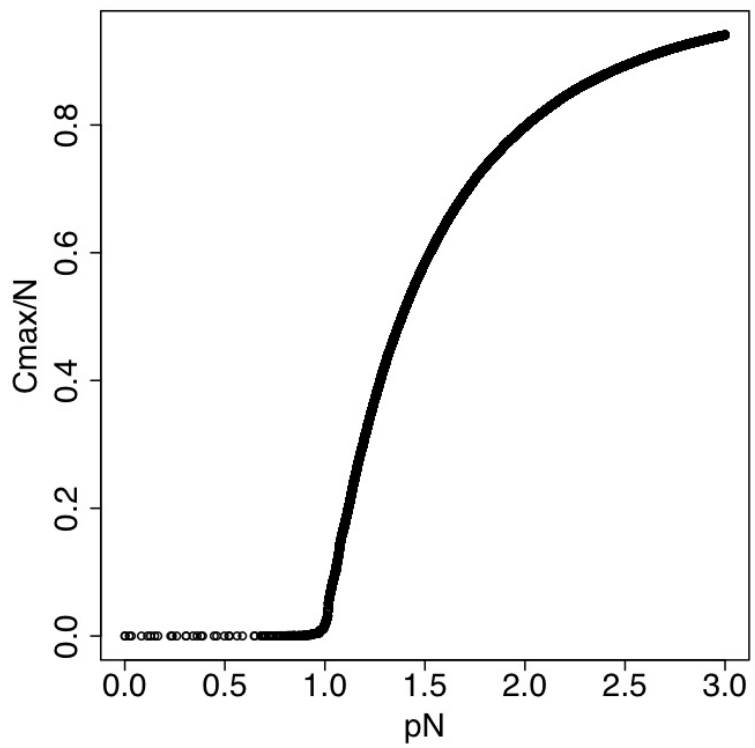
where χ is the expected size of the component to which an arbitrarily chosen vertex belongs.

Connectivity – good or bad?

- Communications, Transportation, Synchronization, ...

  versus

- Spread of human or computer viruses



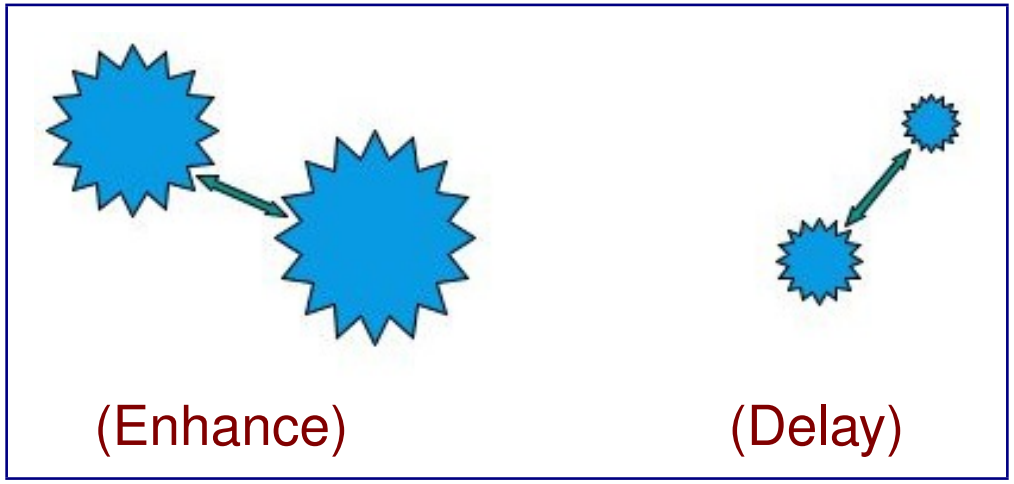
Percolation, onset of: large scale connectivity, epidemic threshold, global cascades...

Can any **limited perturbation** change the phase transition?

[Bohman, Frieze, *RSA* **19**, 2001]

[Achlioptas, D'Souza, Spencer, *Science* **323**, 2009]

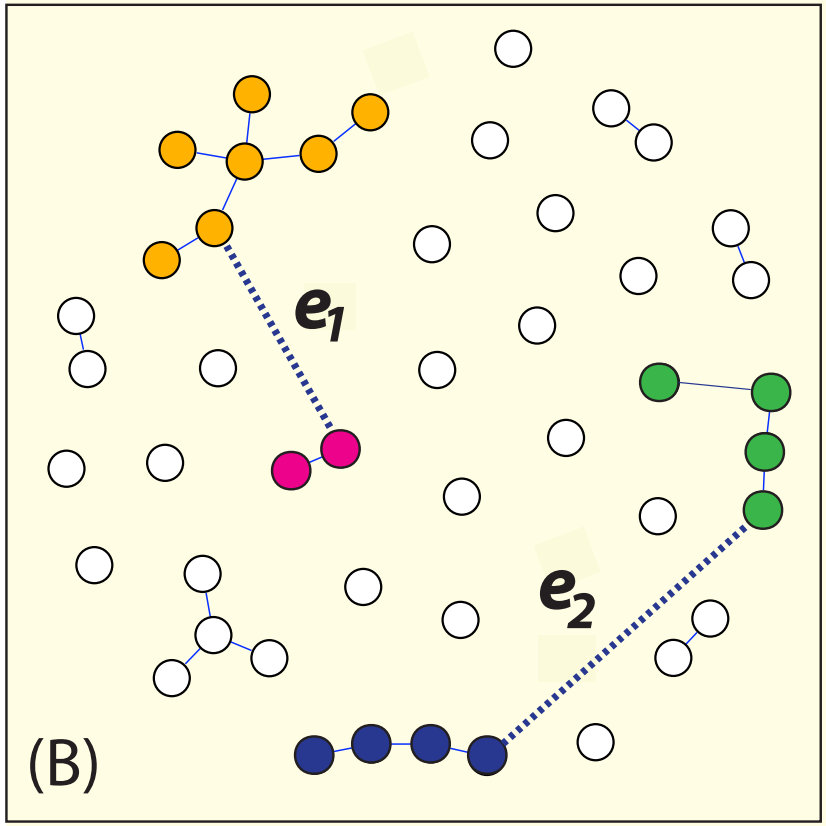
- Possible to **Enhance** or **Delay** the onset?
- The **"Product Rule"**
  - Choose ***two*** edges at random each step.
  - Add only the desirable edge and discard the other.



- The Power of Two Choices in randomized algorithms.

  Azar; Broder; Mitzenmacher; Upfal; Karlin;

ProdRule: Explicit example

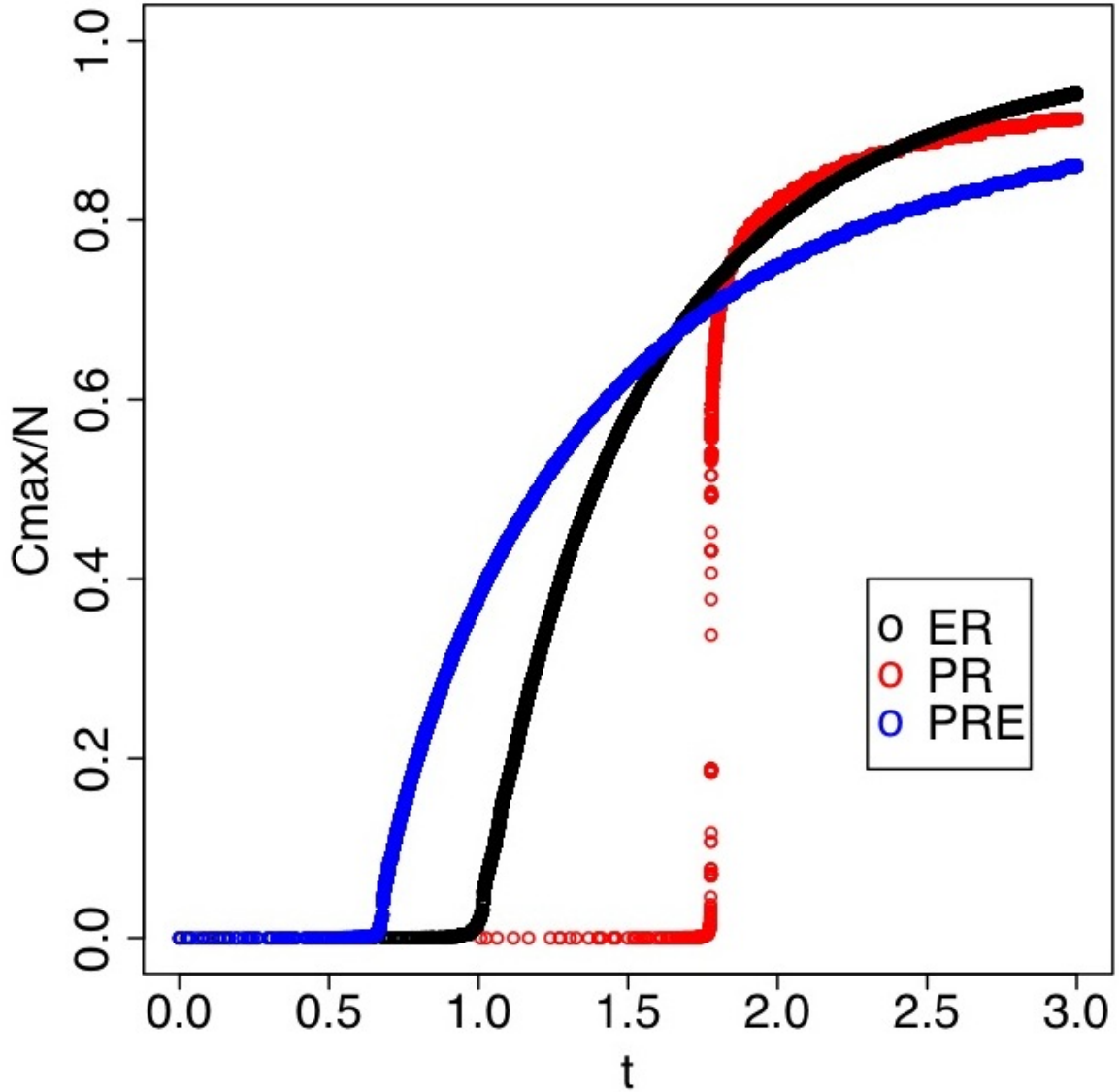


- Prod $e_1 = (7) \times (2) = 14$
- Prod $e_2 = (4) \times (4) = 16$
- To *enhance* choose e_2. To *delay* choose e_1.

Product Rule

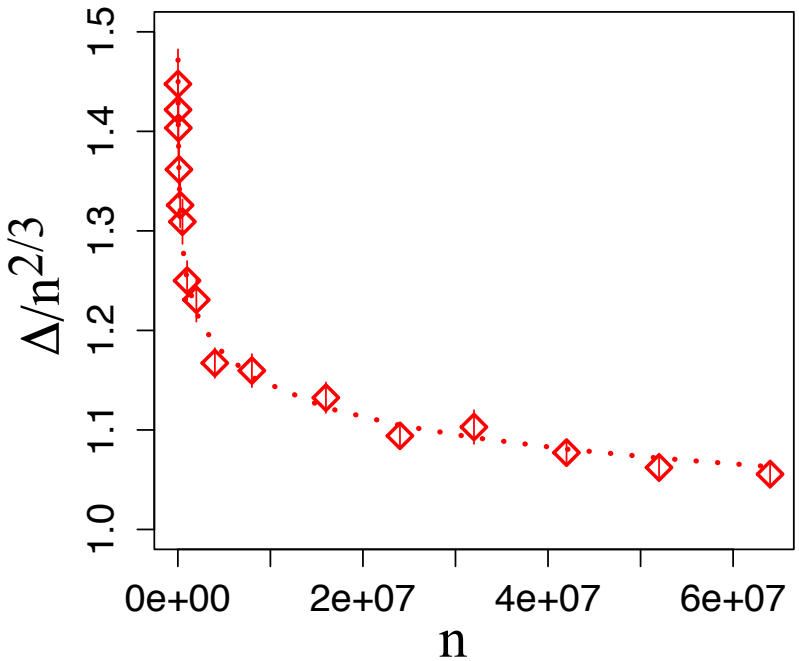
- **Enhance** – similar to **ER** but with earlier onset.

- **Delay** – Extremely abrupt

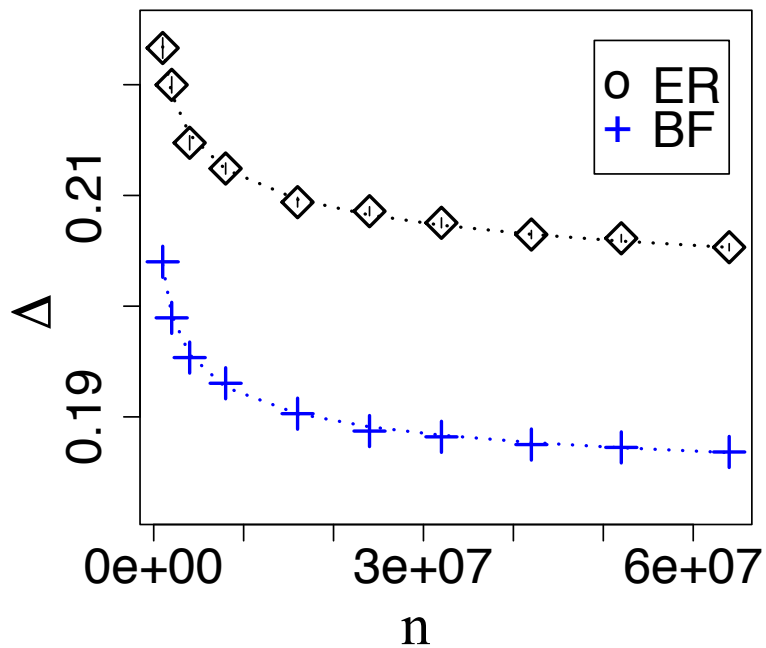


The scaling window, Δ from $n^{1/2}$ to $0.5n$

- Let e_0 denote the last edge added for which $C_{max} < n^{1/2}$. (Recall ER has $n^{2/3}$ at p_c.)
- Let e_1 denote the first edge added for which $C_{max} > 0.5n$.
- Let $\Delta = e_1 - e_0$.



PR $\Delta \sim n^{2/3}$



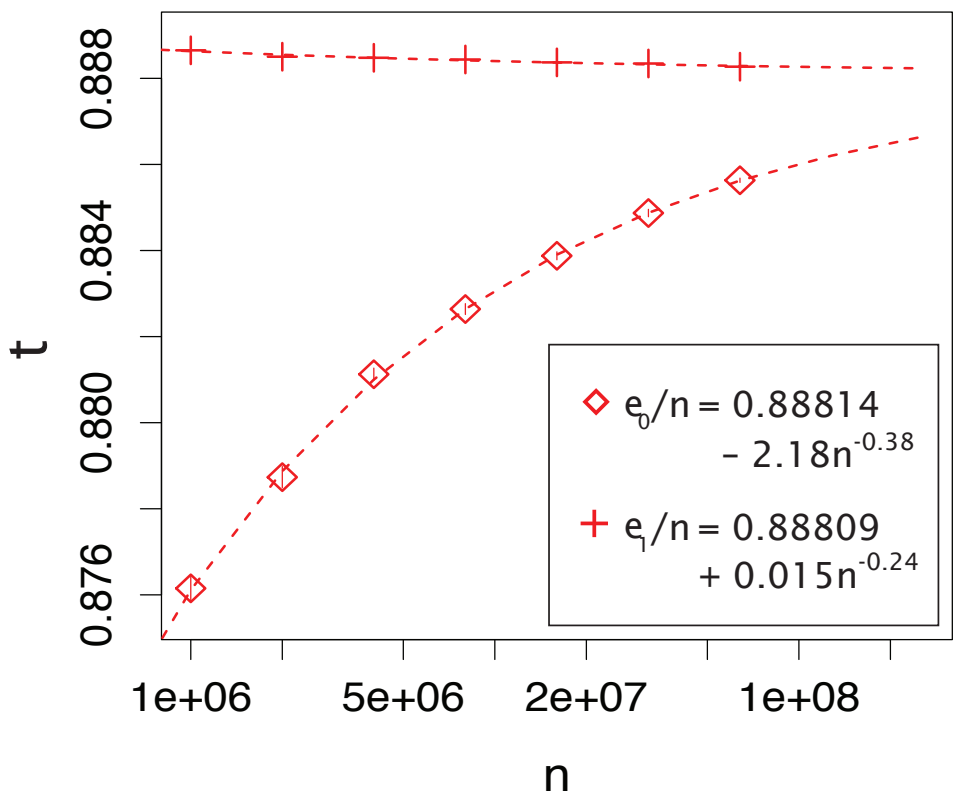
ER (and BF) $\Delta \sim n$.

PR From $n^{1/2}$ to $0.5n$ in number of edges that is sublinear in n.

In terms of edge density or "time", t_c, where $t = e/n$

(Note, for ER, $t_c = 1/2$)

- For $t < t_c$, $C_{\max} < n^{1/2}$.
- For $t > t_c$, $C_{\max} > 0.5n$.

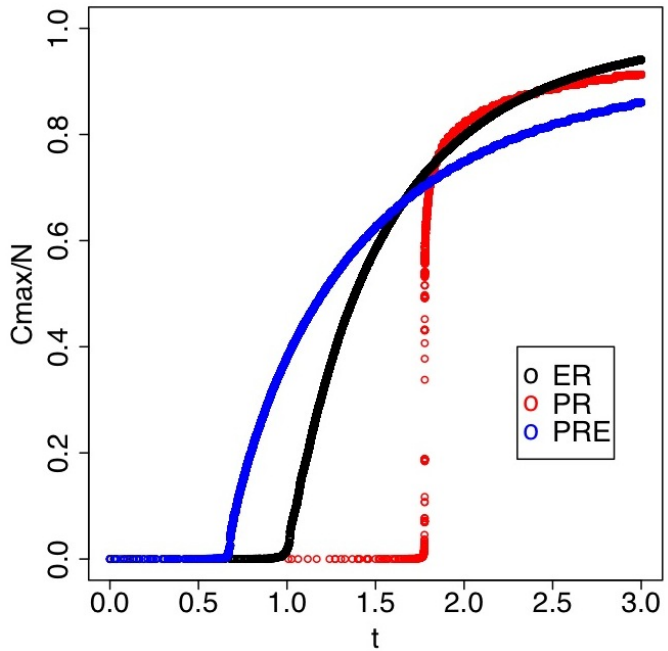


Jumps "instantaneously" from $C_{\max}$ = $n^{1/2}$ to $0.5n$.

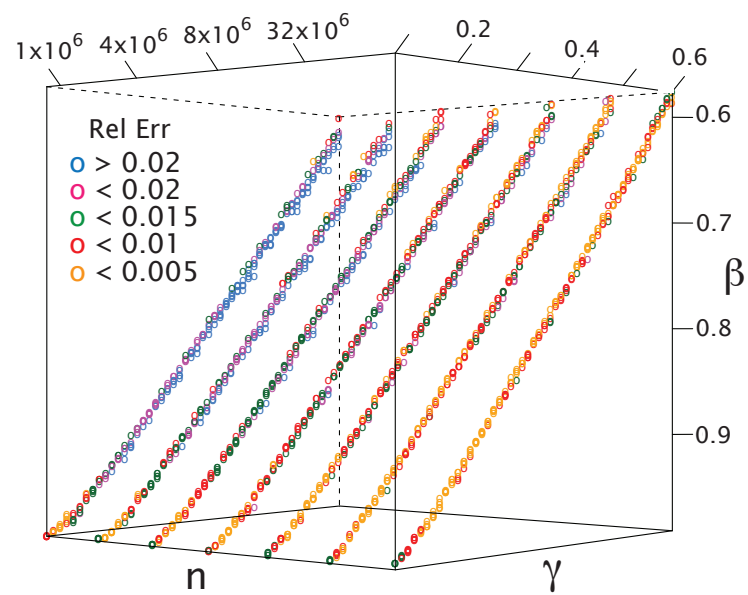
"Explosive Percolation in Random Networks"

From n^γ to greater than $0.6n$ "instantaneously"
(Compelling evidence that the transition is discontinuous)

$C_{\max}$ jumps from sublinear n^γ
to $\geq 0.5n$ in n^β edges, with $\beta, \gamma < 1$.



Nontrivial Scaling behaviors
$\gamma + 1.2\beta = 1.3$ for $A \in [0.1, 0.6]$



Achlioptas, D'Souza, Spencer, *Science*, **323** (5920), 2009

Similar "Explosive percolation" ...

- R. Ziff, *Phys. Rev. Lett.* 103, 045701 (2009).
  "Explosive Growth in Biased Dynamic Percolation on 2-D Regular Lattice Networks"
- Y. S. Cho, J. S. Kim, J. Park, B. Kahng, D. Kim, *Phys. Rev. Lett.* 103, 135702 (2009).
  "Percolation Transitions in Scale-Free Networks under the Achlioptas Process"
  (Chung-Lu weighted node power law growth model)
  - $p_c > 0$ for $\gamma > 2.3$ or 2.4 and discontinuous.
- F. Radicchi, S. Fortunato *Phys. Rev. Lett.* 103, 168701 (2009).
  "Explosive percolation in scale-free networks"
  (Configuration model power law)
  - $p_c > 0$ for $\gamma > 2.2$, discontinuous for $\gamma > 3$.
- E. J. Friedman, A. S. Landsberg *Phys. Rev. Let.* 103, 255701 (2009).
  "Construction and Analysis of Random Networks with Explosive Percolation"
- Y.S. Cho, B. Kahng, D. Kim; *Phys. Rev. E* 030103(R), (2010).
  "Cluster aggregation model for discontinuous percolation transition"
- Rozenfeld, Gallos, Makse; *Eur. Phys. J. B*, 75, 305-310, (2010).
  "Explosive Percolation in the Human Protein Homology Network"
- Araújo, Herrmann, Phys. Rev. Lett. 105, 035701 (2010).
  "Explosive percolation via control of the largest cluster"

More “Explosive percolation” ...

- D’Souza, Mitzenmacher, Phys. Rev. Lett. 104, 195702 (2010).
  “Local cluster aggregation models of explosive percolation”.
- Kumar Pan, Kivelä, Jari Saramäki, Kaski, Kertész, Phys. Rev. E 83, 046112 (2011).
  “Using explosive percolation in analysis of real-world networks”
- Araújo, Andrade Jr, Ziff, Herrmann, Phys. Rev. Lett. 106, 095703 (2011).
  “Tricritical point in explosive percolation”
- Chen, R.D., Phys. Rev. Lett. 106, 115701 (2011)
  “Explosive percolation with multiple giant components”
- Hooyberghs, Van Schaeybroeck, Phys. Rev. E 83, 032101 (2011)
  “Criterion for explosive percolation transitions on complex networks”
- Gomez-Gardenes, Gomez, Arenas, Moreno, Phys. Rev. Lett. *in press*. *
  “Explosive Synchronization Transitions in Scale-free Networks”
- Cho, Khang
  “Explosive percolation transitions in diffusion-limited cluster aggregation model”
- Bastas, Kosmidis, Argyrakis
  “Explosive site percolation and finite size hysteresis”
- . . .

(*Percolations talk session)

Beyond "Product Rule": Models with fixed choice

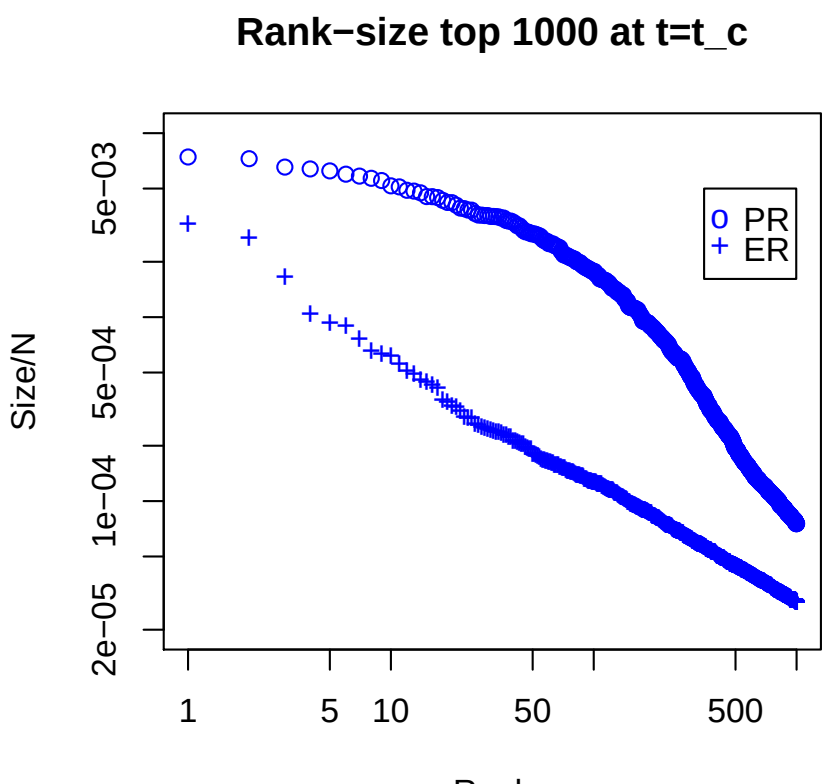
- Nothing special about two edges; need a fixed number greater than one.

  (An "**Achlioptas process**": examine fixed number of edges, add the one that optimizes a pre-set criteria.)

- "Sum rule", Adjacent edge, Triangle rule, k-clique rule, etc., all also work.

- These rules keep largest components similar in size in subcritical regime:

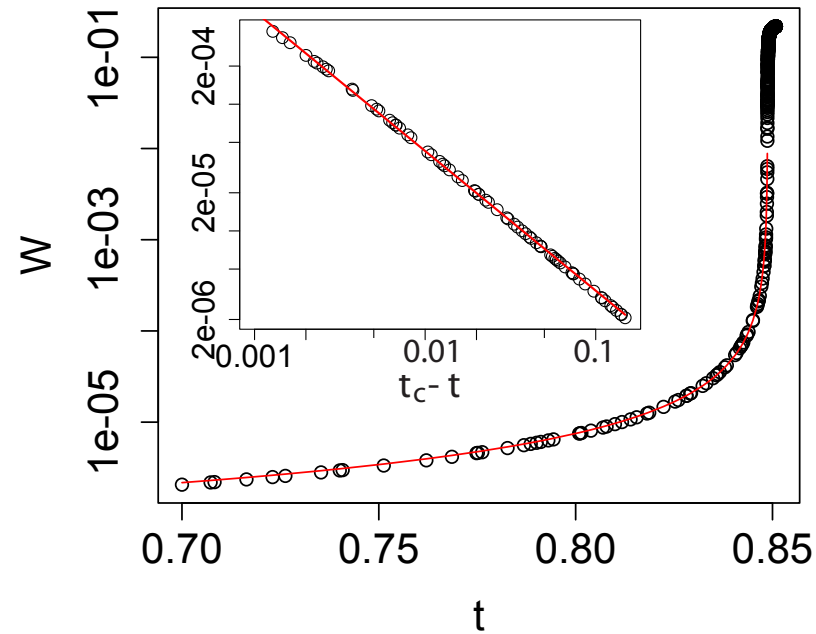
  – "Powder Keg" of Friedman and Landsberg *PRL* (2009).
  – Starting ER from proper initial state; Cho, Khang, Kim *PRE* (2010).



Puzzle 1: Hybrid Transitions!

(Discontinuous change, but scaling behavior)

- Component density $n_i \sim i^{-\tau}$
- "Susceptibility", $W \sim |t - t_c|^{-\alpha}$
- Second largest, $C_2 \sim |t - t_c|^{-\mu}$



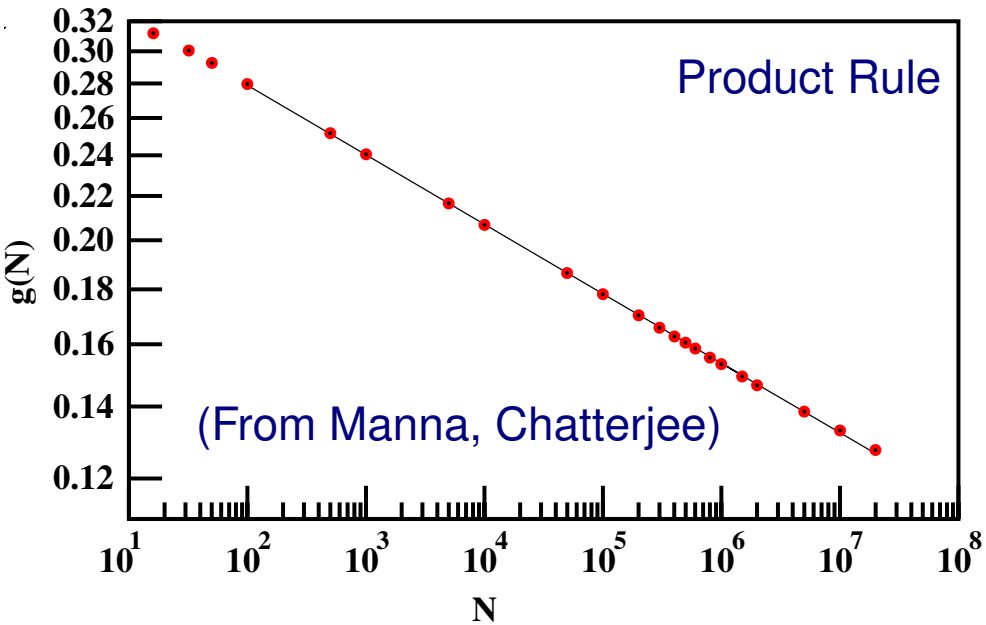
	PR	AE	TR
t_c	0.888	0.796	0.848
τ	2.1*	2.1	2.1
α	1.17	1.13	1.13
μ	1.17	1.13	1.13

* Also seen in Radicchi, Fortunato, *PRE* (2010); Cho, Kahng, Kim, *PRE*, (2010), Manna, Chatterjee *Physica A* (2011).

Puzzle 2: "Weakly" discontinuous

Rather than Δ (the number of edges in the scaling window), measure: the maximum impact from adding *one single edge.*

- C_1 := the fraction of nodes in the largest component.
- ΔC_1 := largest change in C_1 due to addition of a single edge.



$$\Delta C_1 \sim n^{-0.065}$$

- **Nagler**, Levina, Timme, *Nature Phys.* 7, 265 (2011). (Nagler poster)
- Manna, Chatterjee *Physica A* 390 (2011).

The truth: In limit $n \to \infty$, the Product Rule is continuous!

- da Costa, Dorogovtsev, Goltsev, Mendes, *Phys. Rev. Lett.* 105, (2010).*

  Assume component density scales at critical point, $n_i \sim i^{-\delta}$, and then can show ultimately continuous but with extremely slow decay.

  For $n = 10^{18}$ still see $\Delta C_1 = 0.1$ from a single edge (10% of system size!).

  (PR at t_c, $\Delta C = 0.1$ for 10^9, 0.07 for 10^{11}.)

- Riordan, Warnke, arXiv:1102.5306:

  Rigorous proof: Any *fixed choice* process ultimately continuous!
  Δ will ultimately crossover to linear in n, but no estimate of crossover.

- Grassberger, Christensen, Bizhani, Son, Paczuski, arXiv:1103.3728*

  Unusual finite size behavior

- Lee, Kim, Park, arXiv:1103.4439

  Finite size scaling data collapse.

(*See Percolations talks)

- **Are any real social or technologial networks of size $n \sim 10^{18}$?**

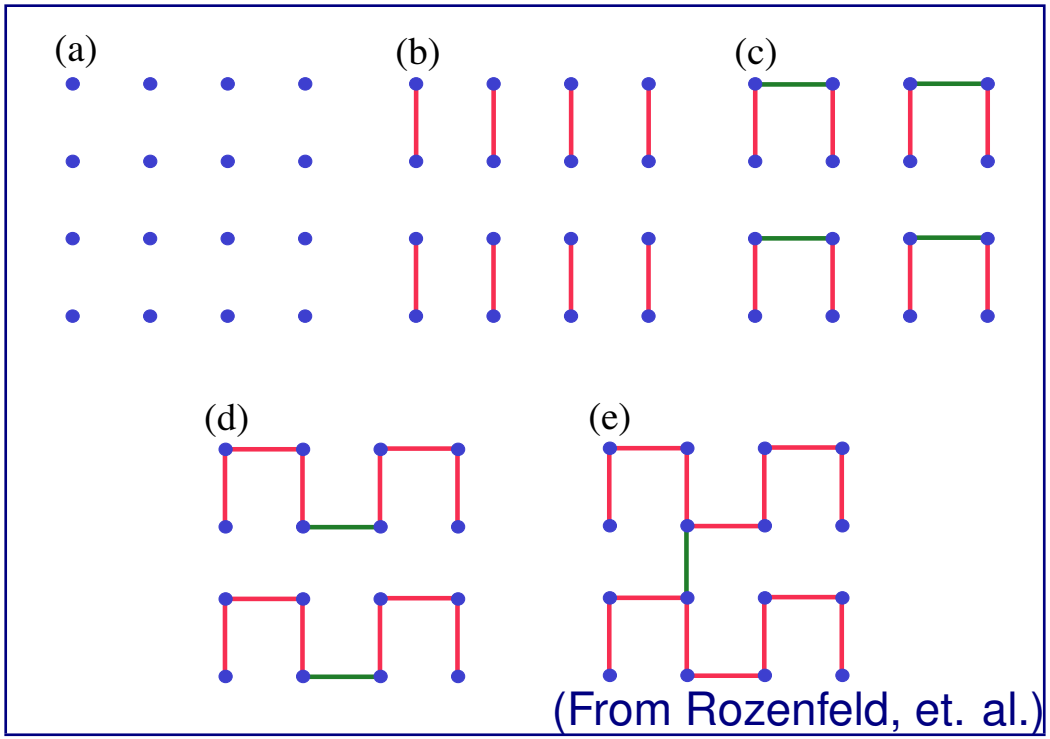
  (100 billion = 10^{11})

The search for a truly discontinuous percolation transition

Friedman, Landsberg *PRL* (2009); Rozenfeld, et. al. *EPJB* (2010);
Nagler, Levina, Timme, *Nature Phys.* (2011)

A deterministic model

- (a) Phase $k = 2$, merge all isolated nodes into pairs.
- (b) Phase $k = 4$, merge pairs into size 4 components.
- (c) Phase $k = 8$, merge pairs of 4's into 8's.
- etc.



(From Rozenfeld, et. al.)

- At edge $e = n$ (time $t = 1$) one giant of size n emerges

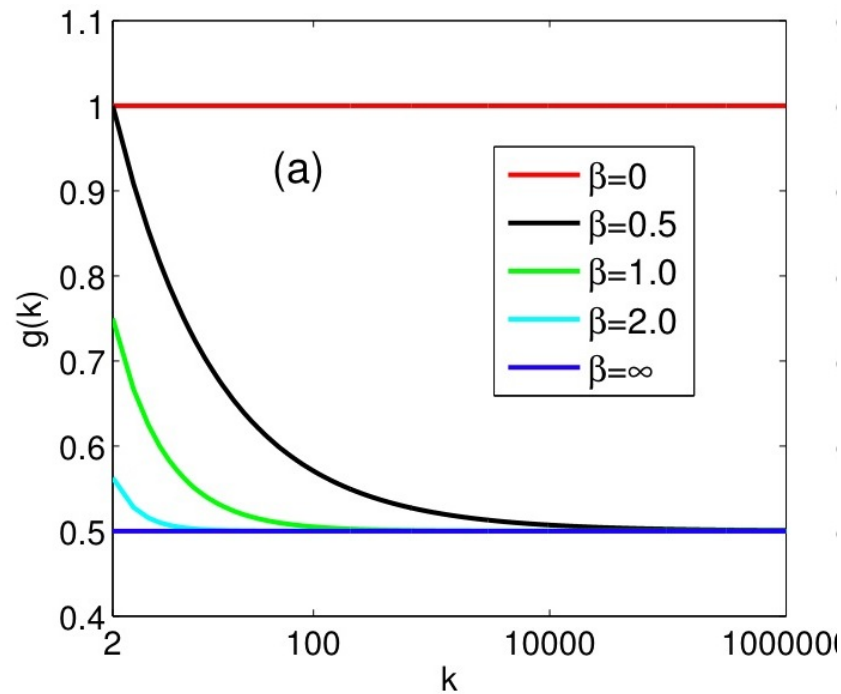
(Giant emerges when only one component remains)

Re-visiting the Bohman Frieze Wormald model (BFW)

(*Random Structures & Algorithms*, 25(4):432-449, (2004))

- A **stochastic model**, which exams a **single-edge at a time**. (Not a model with choice or edge competition).
- Like deterministic, start with n isolated vertices, and stage $k = 2$.
- Sample edges uniformly at random from the complete graph on n nodes.
- Can simply *reject* edges but in phase k must accept at least $g(k) = 1/2 + (2k)^{-1/2}$ fraction of sampled edges.

$$g(k) \to 1/2$$



- **Rigorous proof**: (bounded-size differential eqns)
  – No component of size greater than 200 when $e = 0.96689n$ edges added.
  – Giant component must exist once $e = cn$ edges, with $c \in [0.9792, 0.9793]$.
  – ($t_c < 1$: an infinite number of components exist at transition.)

The BFW model in words

- Start with n isolated vertices, and cap on maximum component set to $k = 2$.
- Sample an edge uniformly at random from the complete graph on n nodes, and examine the result of adding the edge:

1. If the resulting component size $\leq k$, accept the edge.
2. Otherwise reject that edge if possible (meaning the fraction of accepted edges remains $\geq g(k)$).
3. Else augment $k \to k+1$, and repeat steps (1) and (2), with (3) if necessary. (Recall augmenting k decreases $g(k)$).

The BFW model stated formally

- Initially n isolated nodes with cap on maximum size set to $k = 2$.
- Let u denote the total number of edges sampled
- A the set of accepted edges (initially $A = \emptyset$)
- $t = |A|$ the number of accepted edges.

At each step u, select edge e_u uniformly at random from complete graph, and apply the following loop:

```
Set l = maximum size component in A ∪ {e_u}
    if (l ≤ k) {
        A ← A ∪ {e_u}
        u ← u + 1 }
    else if (t/u < g(k)) { k ← k + 1 }
    else { u ← u + 1 }
```

- If the edge e_u is troubling and $t/u < g(k)$, augment k repeatedly until either:

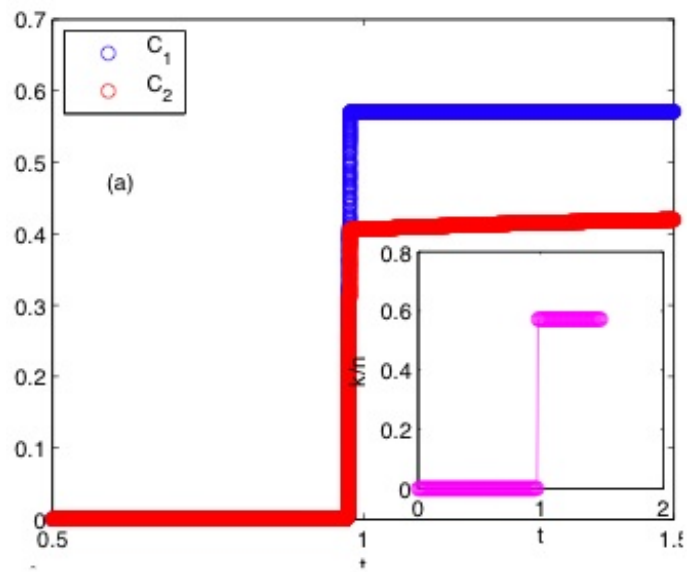
(i) k increases sufficiently that e_u is accepted or

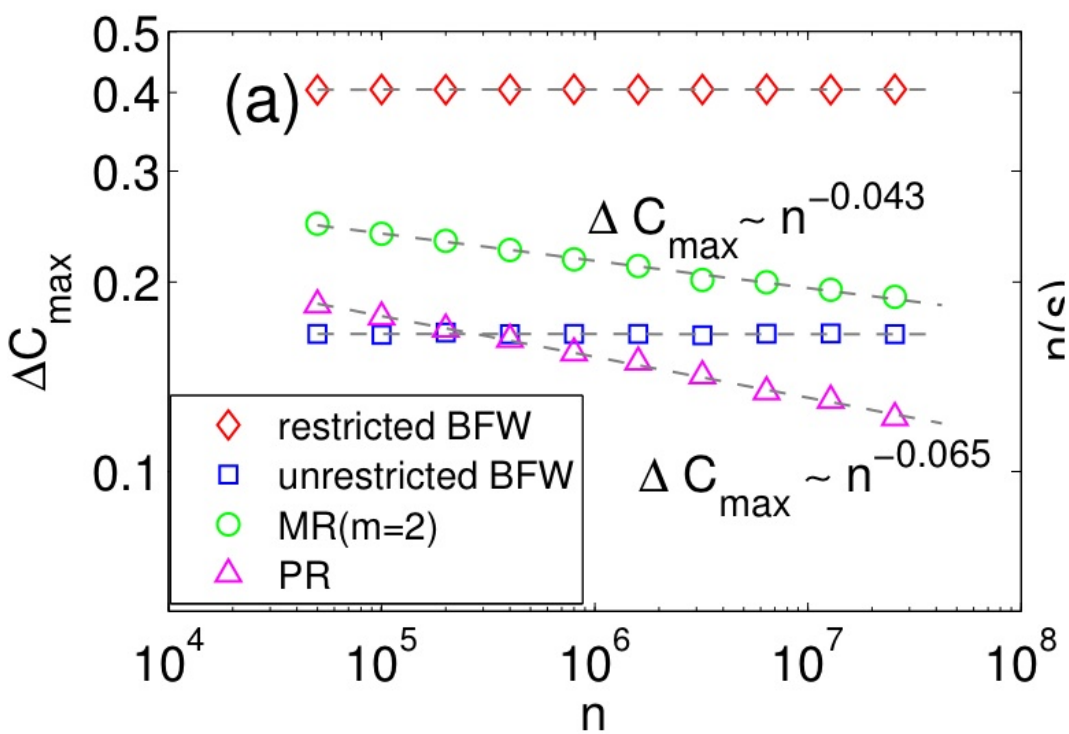
(ii) $g(k)$ decreases sufficiently that e_u is rejected.

Simultaneous emergence of multiple stable giants in a strongly discontinuous transition

(Wei Chen and R.D. *Phys. Rev. Lett.* 83 (2011).)

- Two stable giants!
  $(C_1 = 0.570, C_2 = 0.405.)$
  - Fraction of internal cluster edges $> 1/2$.
  - (If restrict to sampling only edges that span clusters, only one giant ultimately.)





"Strongly" discontinuous (gap independent of n)

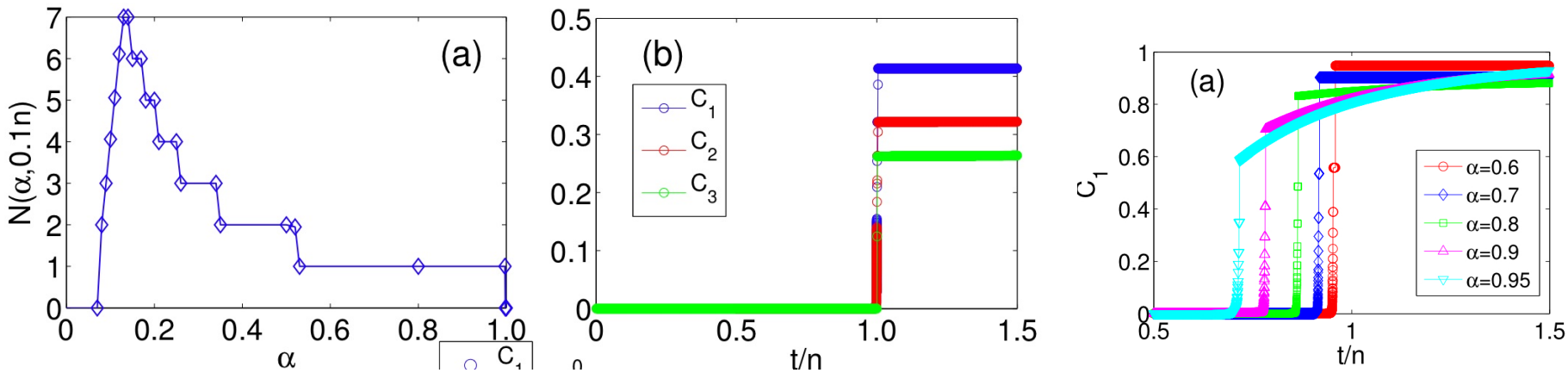
$\Delta C_{\max} \approx 0.165$

Tuning the number of stable giants

(Wei Chen and R.D. *Phys. Rev. Lett.* 83 (2011).)

- Now let $g(k) = \alpha + (2k)^{-1/2}$. Smaller α more edges can be rejected.

  α determines number of stable giants!



- Multiple stable giants, not anticipated.
  ("uniqueness of the giant component" / gravitational coalescence)

- Applications for multiple giants? (Communications, epidemiology, building blocks for modular networks, polymerization (Krapivsky, Ben-Naim)...)

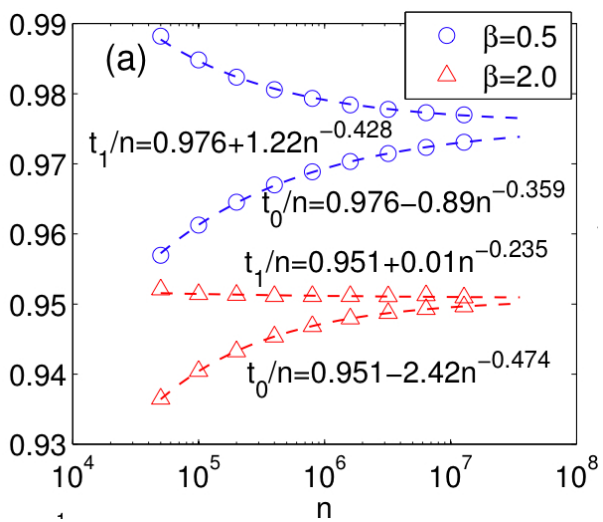
More generally, Discrete jump: multiple giants coexist in critical window

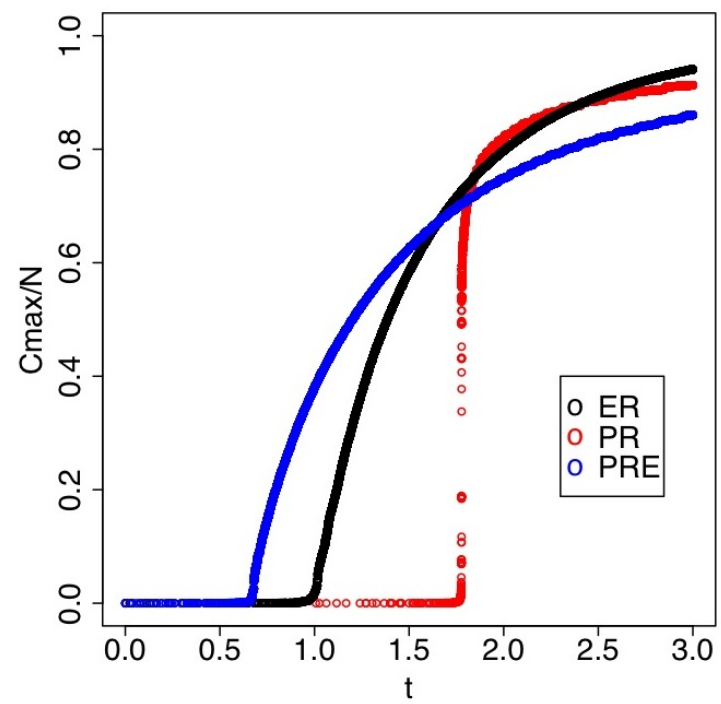
- Note, like Nagler, Manna, da Costa, etc, we define as the critical point t_c, the single edge who's addition causes the biggest change, $\Delta C_{\max}$.

  (Recall C_1 is the *fraction* of nodes in the largest component.)

- If $\Delta C_1 > 0$ (i.e. if we see a discrete jump) then there necessarily existed another macroscopic component. e.g. If $\Delta C_{\max} = 0.1$ (as in da Costa model with $n = 10^{18}$), that means C_1 merged with a component of size $|C_j| = 0.1n$.

- So, t_c is just beyond the "post-critical" regime for Erdős-Rényi.



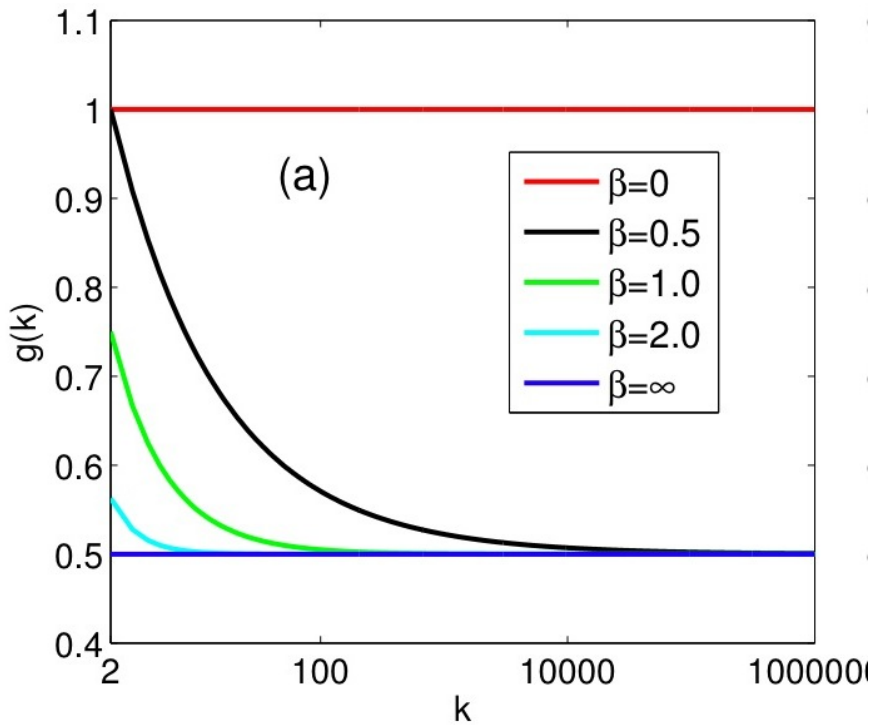


Deriving the underlying mechanism:
Slow decay of $g(k)$ leads to growth by overtaking

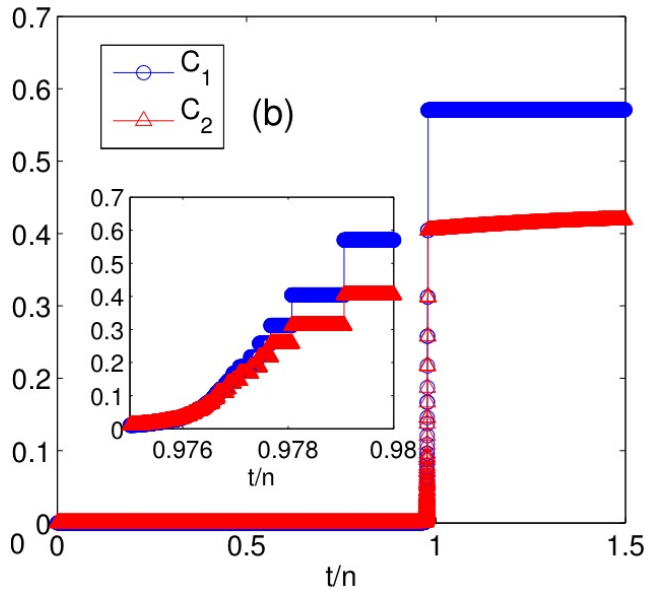
(Wei Chen and R.D, arxiv Fri/Mon,
http://mae.ucdavis.edu/dsouza/Pubs/bfw.pdf)

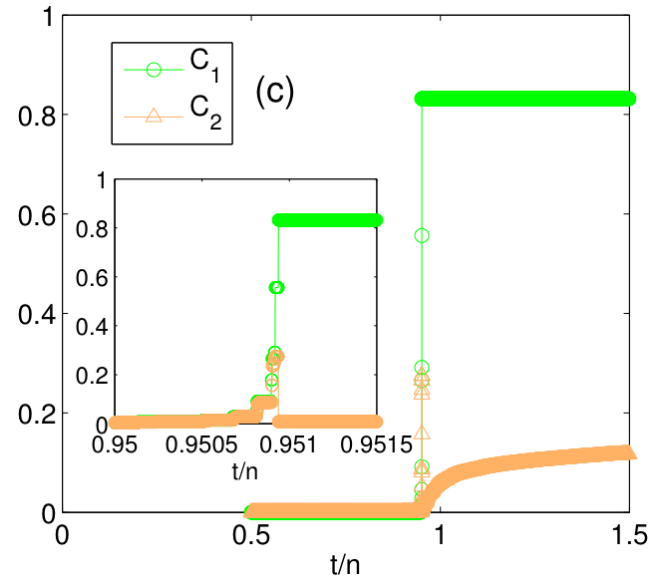
- Instead of $g(k) = 1/2 + (2k)^{-1/2}$ now let $g(k) = 1/2 + (2k)^{-\beta}$

- Proceedure: analyze by how much k must grow before $g(k)$ would decrease sufficiently to reject troubling edge.



- For $\beta \in (0.5, 1]$, an increase in $k \sim n^{\beta}$ is always sufficient to reject a troubling edge. Slow increase in k means:
  - Growth by overtaking*: two smaller components merge becoming new C_1.
  - Multiple components of size $O(n)$ before the largest jump.

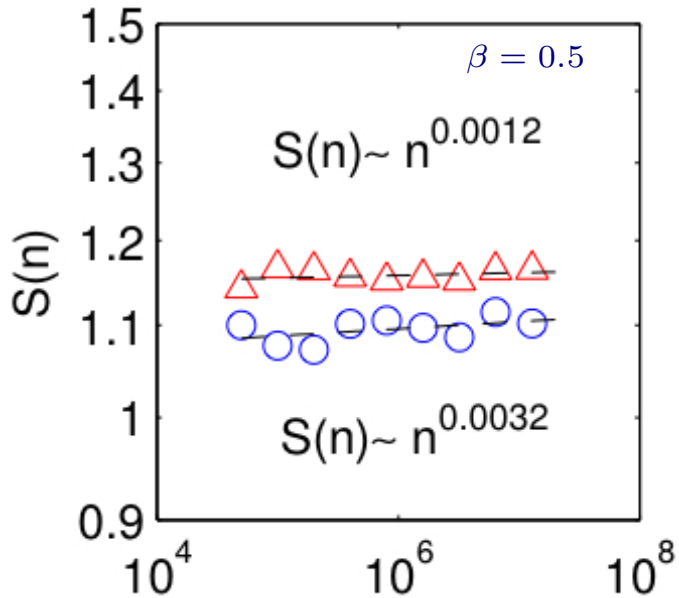


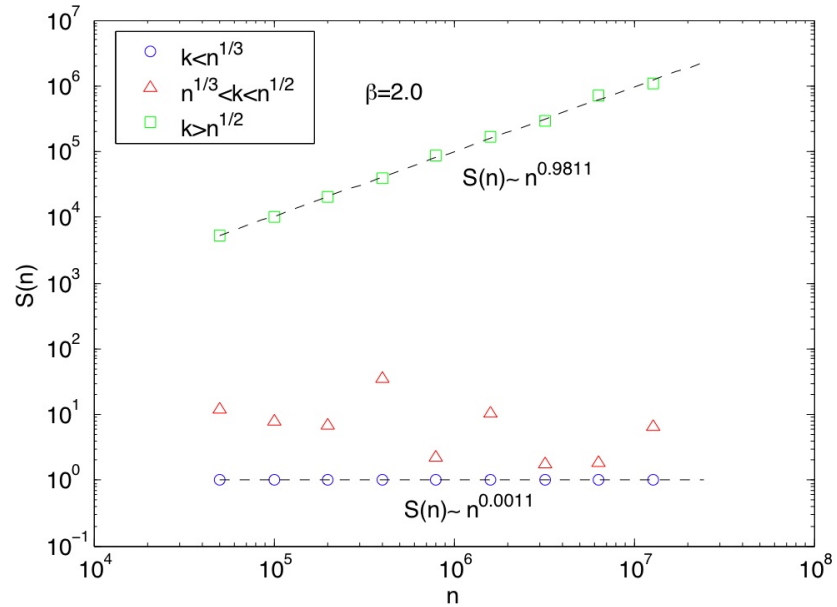


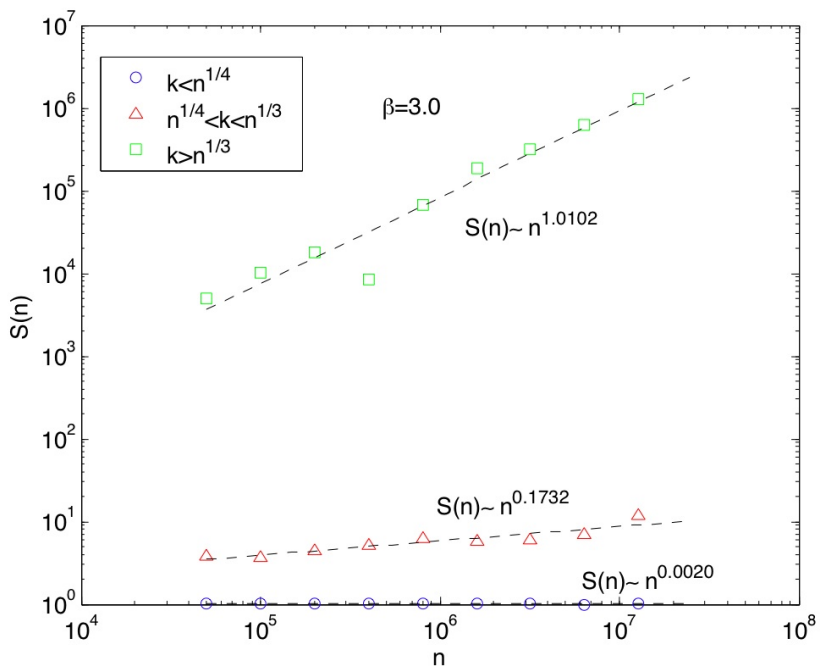
- For $\beta > 1$, once stage $k = n^{1/\beta}$, an increase in $k = C_1$ would be needed, allowing C_1 to even double (the max increase possible). Thus once $k = n^{1/\beta}$ troubling edges **must** be accepted at times, leading to large direct growth of C_1, and a weakly discontinuous transition.

* Consistent with Nagler, et. al., *Nature Phys* (2011), for direct growth forbidden.

$S(n)$, Average size of direct growth in C_1



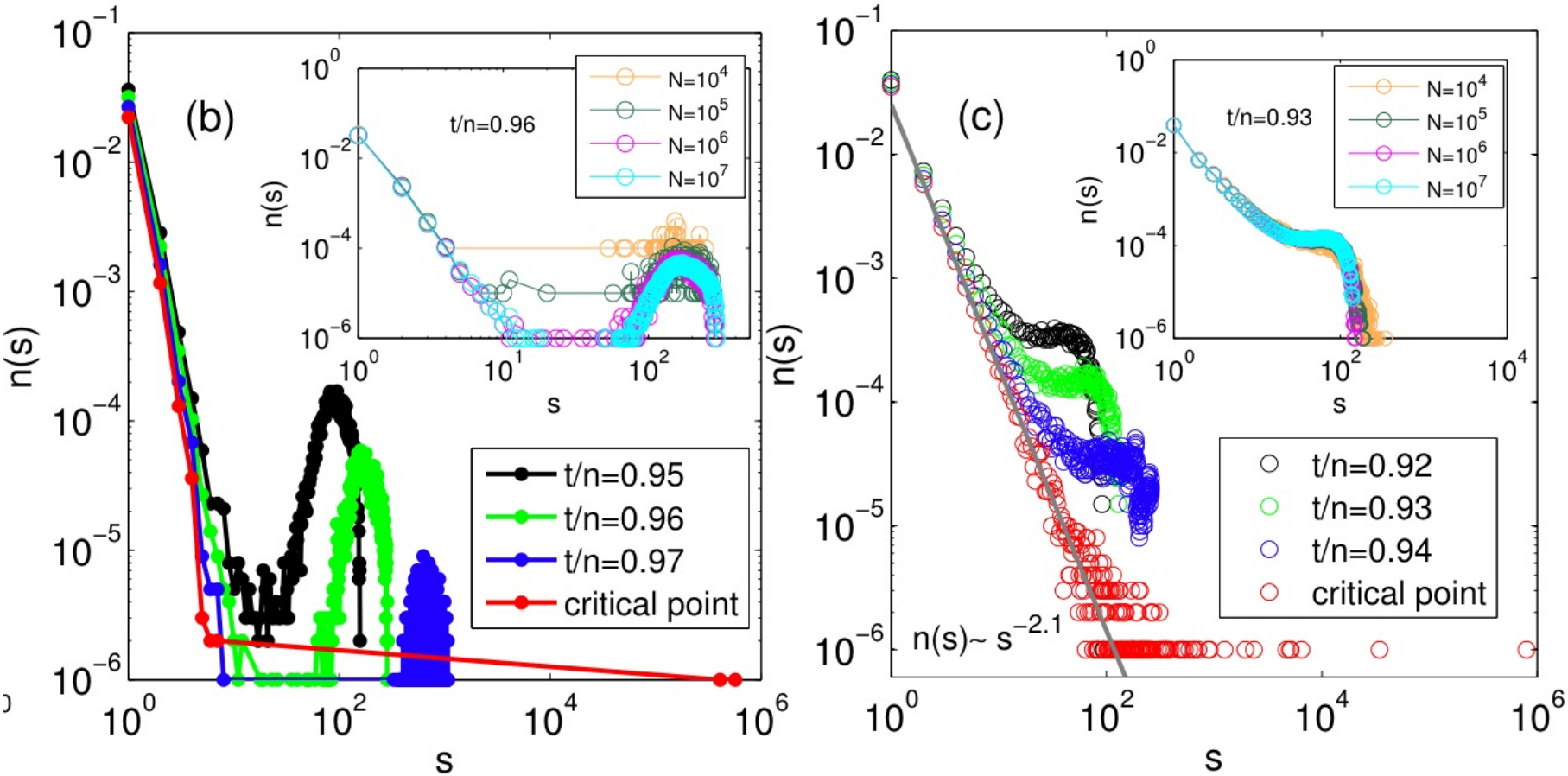




For $\beta > 1$, once $k = n^{1/\beta}$ direct growth merging with size $O(n)$.

Evolution of component density for BFW

$\beta = 0.5$ $\beta = 2.0$



- For $\beta = 0.5$ no scaling. Separates into components of size $O(n)$ and $< \log(n)$.
- For $\beta = 0.5$ and $\beta = 2.0$ no finite size effects in the location of the "hump" (inset), unlike for PR where location depends on n. (c.f. Lee, Kim, Park: data collapse)
- No scaling, no "early warning signs" (Scheffer, et. al. *Nature* (2009).

A rigorous proof of a discontinuous transition for a variant on Erdős-Rényi

Panagiotou, Spöhel, Steger, Thomas, arxiv:1104.1309

- An Erdős-Rényi-like evolution, but at every step connect two vertices, one chosen randomly from all vertices, and one chosen randomly from a restricted set of vertices.

Does EP apply to any real physical process or provide useful applications?

- Growth of wikipedia in 5 different languages:
  disconnected topical clusters merging together over a few days.
  (G. Bounova)

- Evolution of protein homology network: evolution of spanning cluster
  Rozenfeld, Gallos, Makse; *Eur. Phys. J. B*, 75, 305-310, 2010.

- Mapping sub-critical clusters to community structure in real social networks
  Kumar Pan, Kivelä, Jari Saramäki, Kaski, Kertész,
  Phys. Rev. E 83, 046112 (2011).

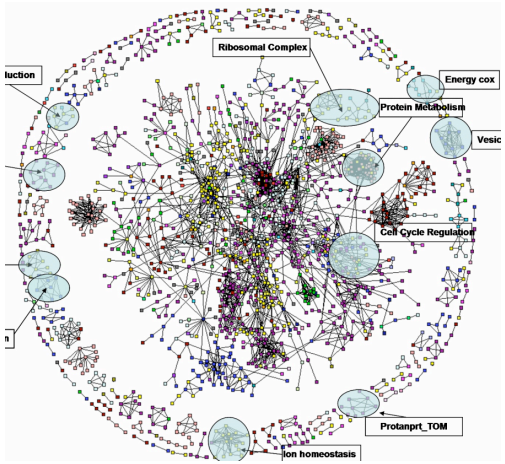
- Applications to network discovery (with Z. Toroczkai)

- Multiple giants...

From multiple giants to interacting networks

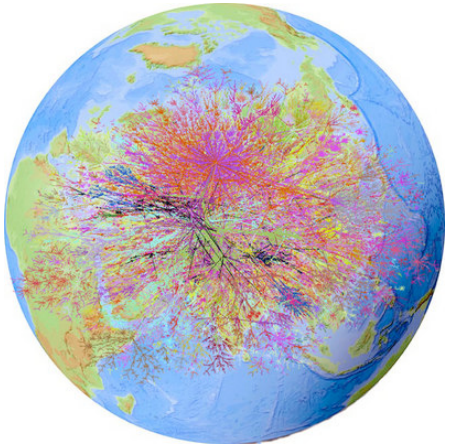
Networks:

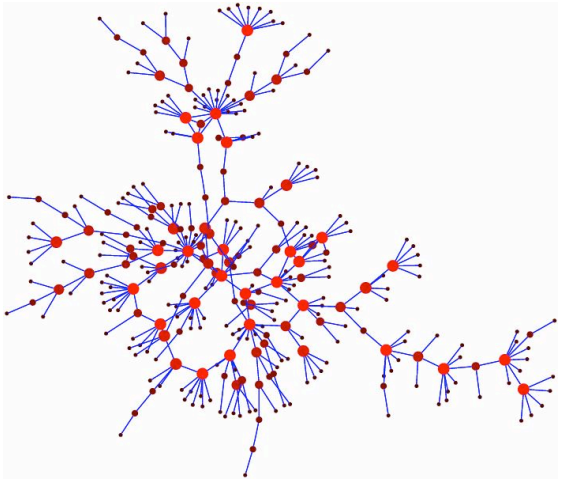
Transportation Networks/ Power grid
(distribution/ collection networks)



Biological networks
- protein interaction
- genetic regulation
- drug design

Computer networks

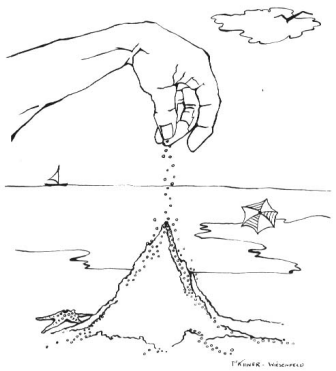


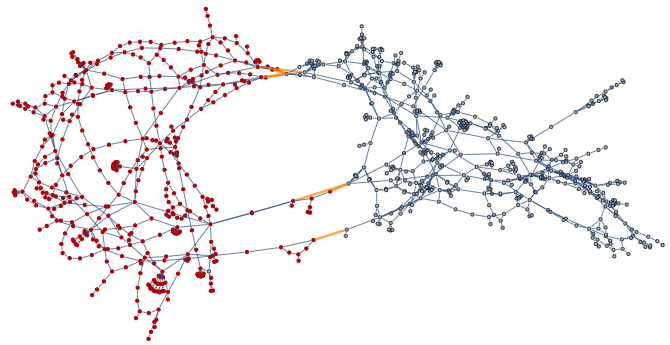


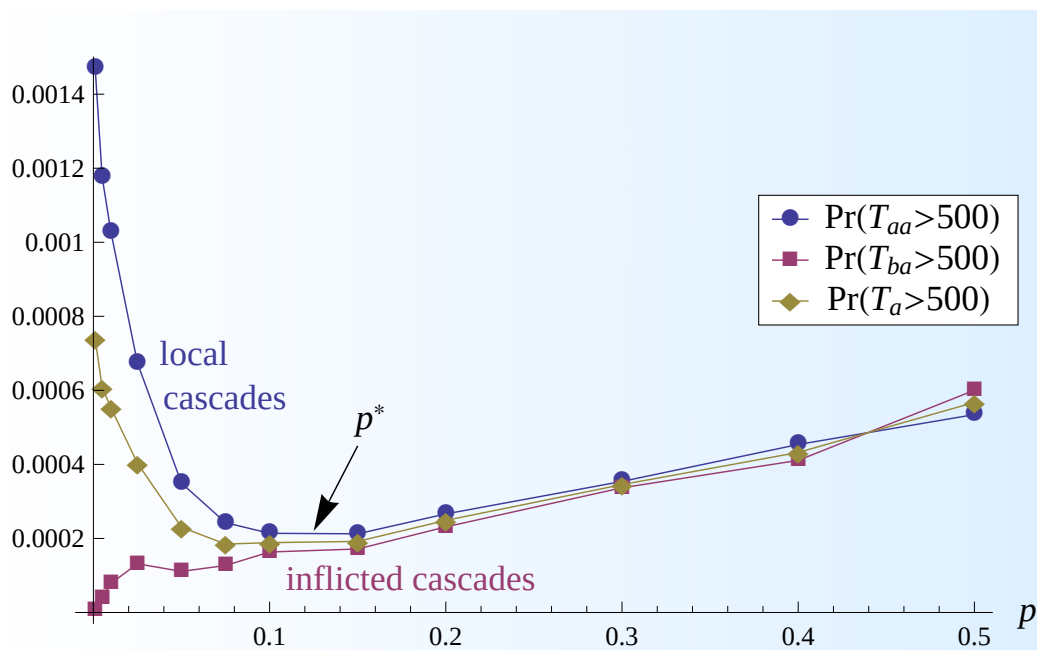
Social networks
- Immunology
- Information
- Commerce

Sandpile cascades on interacting networks

C. Brummitt, R. D'Souza, E. A. Leicht (arxiv on Friday)







Balancing the benefits and detriments of interconnectivity.

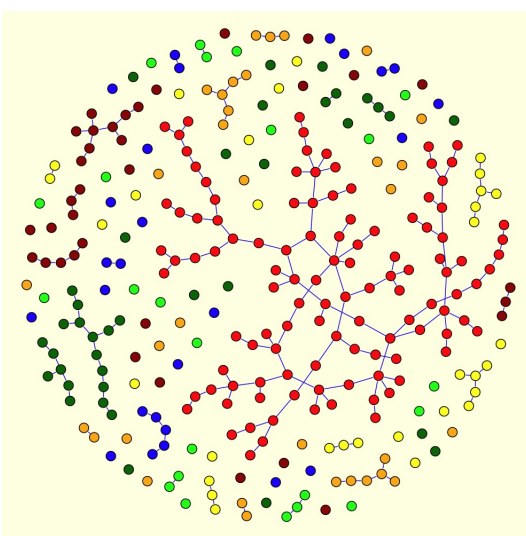
Explosive percolation in random graphs – Conclusions

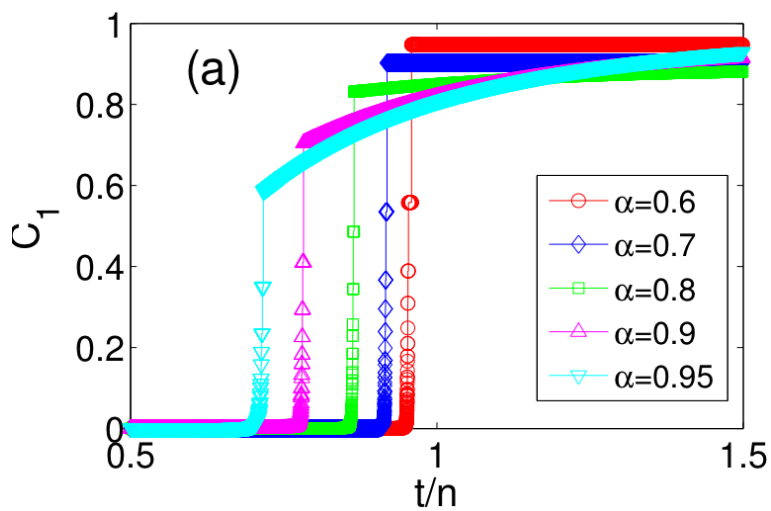
- Models with fixed choice are "weakly" discontinuous (gap decays *slowly* with system size)... Is there a better name?

- In the true $n \to \infty$ limit, they are in fact *continuous*, but the crossover may be greater than 10^{18}. **In what regime to real-world networks exist?**

- **Truly discontinuous percolation transitions do exist** (BFW, restricted Erdős-Rényi).

- **Multiple giant components** necessarily exist *before* the discrete jump (and, c.f. BFW, can co-exist in supercritical region).





Thanks to

- Collaborators: D. Achlioptas, J. Spencer, M. Mitzenmacher, W. Chen

- Funding Agencies:
  – Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053.
  – Defense Threat Reduction Agency, Basic Research Award No. HDTRA1-10-1-0088.
  – National Academies Keck Futures Initiative.

- To you for your attention.

Potential postdoctoral opportunities: raissa@cse.ucdavis.edu