# On two biased graph processes

Gideon Amir * Eyal Lubetzky †

November 21, 2006

**Abstract**

In [1], the authors consider the generalization $\mathcal{G}_K^\vee$ of the Erdős-Rényi random graph process $\mathcal{G}_1$, where instead of adding new edges uniformly, $\mathcal{G}_K^\vee$ gives a weight of size 1 to missing edges between pairs of isolated vertices, and a weight of size $K \in [0,\infty)$ otherwise. This can correspond to the linking of settlements or the spreading of an epidemic. The authors investigate $t_g^\vee(K)$, the critical time for the appearance of a giant component as a function of $K$, and prove that $t_g^\vee = (1+o(1))\frac{4}{\sqrt{3K}}$, using a proper timescale.

In this work, we show that a natural variation of the model $\mathcal{G}_K^\vee$ has interesting properties. Define the process $\mathcal{G}_K^\wedge$, where a weight of size $K$ is assigned to edges between pairs of non-isolated vertices, and a weight of size 1 otherwise. We prove that the asymptotical behavior of the giant component threshold is essentially the same for $\mathcal{G}_K^\wedge$, and namely $t_g^\wedge/t_g^\vee$ tends to $\frac{64\sqrt{6}}{\pi(24+\pi^2)} \approx 1.47$ as $K \to \infty$. However, the corresponding thresholds for connectivity satisfy $t_c^\wedge/t_c^\vee = \max\{\frac{1}{2}, K\}$ for every $K > 0$. Following the methods of [1], $t_g^\wedge$ is characterized as the singularity point to a system of differential equations, and computer simulations of both models agree with the analytical results as well as with the asymptotic analysis. In the process, we answer the following question: when does a giant component emerge in a graph process where edges are chosen uniformly out of all edges incident to isolated vertices, while such exist, and otherwise uniformly? This corresponds to the value of $t_g^\wedge(0)$, which we show to be $\frac{3}{2} + \frac{4}{3e^2-1}$.

## 1 Introduction

### 1.1 The Erdős-Rényi graph process and biased processes

The random graph process, $\mathcal{G}_1(n)$, is a sequence of $\binom{n}{2}+1$ graphs on $n$ vertices, ${\mathcal{G}_1}^0, \ldots, {\mathcal{G}_1}^{\binom{n}{2}}$, where ${\mathcal{G}_1}^0$ is the edgeless graph on $n$ vertices, and ${\mathcal{G}_1}^T$ is obtained by adding an edge to ${\mathcal{G}_1}^{T-1}$, chosen uniformly over all missing edges. This model was introduced by Erdős and Rényi in [9], where it is

*Weizmann Institute, Rehovot, 76100, Israel. Email: gideon.amir@weizmann.ac.il

†School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: lubetzky@tau.ac.il. Research partially supported by a Charles Clore Foundation Fellowship.

shown that for every constant $C < \frac{1}{2}$, the largest component of ${\mathcal{G}_1}^{Cn}$ is typically of size $O(\log n)$, and yet for every constant $C > \frac{1}{2}$ there is typically a single component of linear size in ${\mathcal{G}_1}^{Cn}$ and all other components are of size $O(\log n)$. This single component and its evolution are referred to as the "giant component", and "the double jump phenomenon" respectively. Using a timescale of $n/2$ edges, the threshold for the appearance of the giant component is thus $t_g = 1$. Another classical result of Erdős and Rényi determines that ${\mathcal{G}_1}^{Cn \log n}$ is typically connected for $C > \frac{1}{2}$, and typically disconnected for $C < \frac{1}{2}$. Using a timescale of $\frac{n}{2} \log n$, the connectivity threshold is thus $t_c = 1$. For further information on the evolution of the giant component in the random graph process, as well as on the threshold for connectivity, see, e.g., [3].

There has been extensive study on the thresholds for the appearance of a giant component and for connectivity in different variations of the random graph process. For instance, in a model suggested by Achlioptas, two random edges are chosen uniformly out of the missing edges at each step, out of which some algorithm $\mathcal{A}$ selects one to be added to the graph. For results on upper and lower bounds on the emerging of the giant components for various algorithms in this model, see [4],[5],[6],[10].

In [7], the authors consider a random graph process on multi-graphs, where at each step an edge is added between a random vertex of minimal degree and a random uniformly chosen vertex. The authors analyze the number of vertices of degrees $0, 1, 2$ along the process using the differential equation method for graph processes of Wormald [12], and show that the mentioned graph process becomes connected typically when the minimal degree becomes 3.

The following generalization of the original graph process $\mathcal{G}_1$, which we denote by $\mathcal{G}_K^{\vee}(n)$, was studied in [1]: at each step, $\mathcal{G}_K^{\vee}$ gives a weight of size 1 to missing edges between pairs of isolated vertices, and a weight of size $K \in [0, \infty)$ otherwise (when no isolated vertices are left, the distribution on the missing edges becomes the uniform distribution). This can correspond to the linking process of $n$ initially isolated settlements, or the spreading of an epidemic, where the probability of a new link is affected by whether or not one of its endpoints already has other links. The threshold for the appearance of a giant component in $\mathcal{G}_K^{\vee}$, $t_g^{\vee}$, becomes a continuous function of the parameter $K$, and the authors of [1] use the differential equation method to express $t_g^{\vee}$ as a singularity point to a system of coupled non-linear ordinary differential equations (ODEs). By applying methods of asymptotic analysis of ODEs, it is proved that $t_g^{\vee}(K) = (1 + o(1)) \frac{4}{\sqrt{3K}}$, where the $o(1)$-term tends to 0 as $K \to \infty$.

In this work, we show that a natural variation on the process $\mathcal{G}_K^{\vee}$ has interesting properties. Consider the process $\mathcal{G}_K^{\wedge}$, where instead of placing the weight $K$ when one of the endpoints is non-isolated (as in $\mathcal{G}_K^{\vee}$), it is placed when *both* endpoints are non-isolated. In other words, $\mathcal{G}_K^{\vee}$ gives a weight of size $K \in [0, \infty)$ to missing edges between pairs of non-isolated vertices, and a weight of size 1 otherwise. Notice that for $K = 1$, both processes are equivalent to the original graph process. Furthermore, for any $K$, both processes $\mathcal{G}_K^{\vee}$ and $\mathcal{G}_K^{\wedge}$ apply the rule of the original graph process $\mathcal{G}_1$ once no isolated vertices are left, hence it is interesting to compare the two up to roughly that point.

Applying the methods of [1] on $\mathcal{G}_K^\wedge$, we express its threshold for the appearance of a giant component, $t_g^\wedge(K)$, as a singularity point to a system of coupled non-linear ODEs. For the special case $K = 0$, we obtain a setting where edges are added uniformly at random out of all edges incident to an isolated vertex, until no such vertex is left (from that point on, edges are added uniformly at random). In this special case we get $t_g^\wedge(0) = \frac{3}{2} + \frac{4}{3e^2-1}$.

Using classical methods of asymptotic analysis of ODEs (see, e.g., [2]) we readily calculate the asymptotic behavior of $t_g^\wedge(K)$ for $K \gg 1$, and obtain that $t_g^\wedge(K) = (1+o(1))\frac{\pi}{2\sqrt{2}}\left(1+\frac{\pi^2}{24}\right)\frac{1}{\sqrt{K}}$. It follows that $t_g^\wedge/t_g^\vee \approx 1.47$ for $K \gg 1$, and we note that obtaining this result via combinatorial arguments seems challenging. Numerical approximations of the ODEs validate this asymptotic analysis.

While the behavior of the threshold for the appearance of a giant component is similar for $K \gg 1$, combinatorial arguments yield that for all $K > 0$, $t_c^\vee(K) = 1$, whereas $t_c^\wedge(K) = \max\{\frac{1}{2}, K\}$, hence $t_c^\wedge/t_c^\vee = K$ for every $K \geq \frac{1}{2}$.

It is possible to implement both processes $\mathcal{G}_K^\wedge$ and $\mathcal{G}_K^\vee$ efficiently by choosing an appropriate data-structure for holding the isolated and non-isolated vertices. Our implementation requires $O(n)$ memory and runs in time $O(n \log n)$, and its results validate the above analytical results.

### 1.2 Notations and main results

A property of graphs is a collection of graphs closed under isomorphism. A property is said to be monotone (increasing) if it is closed under the addition of edges. Throughout the paper, we say that a property of graphs on $n$ vertices occurs *with high probability*, or *almost surely*, or that *almost every* process $\mathcal{G}$ satisfies this property, if the probability for the corresponding event tends to 1 as $n \to \infty$. Note that, when proving that certain statements hold with high probability, one may clearly condition on events which hold with high probability.

On several occasions, we examine the processes $\mathcal{G}_K^\vee$ or $\mathcal{G}_K^\wedge$ starting from some arbitrary graph $H$ (instead of the edgeless graph). We denote these processes by $\mathcal{G}_K^\vee|_H$ and $\mathcal{G}_K^\wedge|_H$.

Given a process $\mathcal{G}$, we let $\mathcal{G}^i$ denote $\mathcal{G}$ after $i$ edges. It will be convenient to use two timescales when referring to $\mathcal{G}$, which we denote by:

$$\begin{aligned}\mathcal{G}(t) &:= \mathcal{G}^{tn/2}\ , \\ \mathcal{G}[t] &:= \mathcal{G}^{t\frac{n}{2}\log n}\ .\end{aligned}$$

We say that $t$ is a threshold for the appearance of a giant component in $\mathcal{G}$ if for every $\varepsilon > 0$, with high probability $\mathcal{G}(t-\varepsilon)$ does not contain a giant component and yet $\mathcal{G}(t+\varepsilon)$ does contain one. Similarly, we say that $t$ is a threshold for connectivity if for every $\varepsilon > 0$, with high probability $\mathcal{G}[t-\varepsilon]$ is disconnected whereas $\mathcal{G}[t+\varepsilon]$ is connected. According to these definitions, both thresholds equal 1 for the original random graph process $\mathcal{G}_1$.

Theorem 1.1 determines the threshold for connectivity in both processes, $t_c^\vee$ and $t_c^\wedge$, and shows that $t_c^\vee = 1$ for any $K > 0$ whereas $t_c^\wedge = K$ for any $K > \frac{1}{2}$, hence their ratio is $K$ for any $K > \frac{1}{2}$. Furthermore, as we later state, it follows from [1] that a ratio of $\lceil \max\{\frac{1}{K}, K\} \rceil$ is the maximal possible between the threshold of $\mathcal{G}_K^\wedge$ and the threshold of $\mathcal{G}_1$ for any monotone property, and therefore, $t_c^\vee$ achieves this maximum.

**Theorem 1.1.** *For every $K > 0$, $t_c^\vee(K) = 1$ and yet $t_c^\wedge(K) = \max\{\frac{1}{2}, K\}$. In the special case $K = 0$ we have $t_c^\vee(0) = t_c^\wedge(0) = \frac{1}{2}$.*

For a given graph $G = (V, E)$ on $|V| = n$ vertices, we let $\mathcal{C} = \mathcal{C}(G)$ denote the set of connected components of $G$, and for $i \in \mathbb{N}$, we define $\mathcal{C}_i = \mathcal{C}_i(G)$ to be the set of components of size $i$: $\mathcal{C}_i = \{C \in \mathcal{C}(G) : |C| = i\}$. Whenever it is clear as to which graph process $\mathcal{G}$ we are referring, we use the abbreviation $\mathcal{C}_i^t$ to denote $\mathcal{C}_i(\mathcal{G}^t)$. The fractions of vertices which belong to components of size 1 and 2 are defined as:

$$I(G) = \frac{|\mathcal{C}_1|}{n} \ , \ I_2(G) = \frac{2|\mathcal{C}_2|}{n} \ ,$$

and the *susceptibility* of $G$, $S(G)$, is defined to be the average size of a connected component, averaged over all vertices:

$$S(G) = \frac{1}{n}\sum_{v \in V} |C(v)| = \frac{1}{n}\sum_{C \in \mathcal{C}(G)} |C|^2,$$

where $C(v)$ denotes the connected component of $v$. The relation between $S(G)$ and the existence of a giant component in $G$ is immediate: if $G$ contains a component of size $\alpha n$ for some $\alpha > 0$, then $S(G) \geq \alpha^2 n$, and if $S(G) \geq \alpha n$ then clearly there exists a component of size $\alpha n$. Therefore, $G$ has a giant component iff $S(G) = \Theta(n)$.

The methods used in [1] to analyze $\mathcal{G}_K^\vee$ and describe $t_g^\vee$ as a singularity point to a system of ODEs can in fact be applied to a wider class of graph processes. Namely, these methods can be applied to every process where the weight function $W$ on the missing edges satisfies $\max W / \min W \leq K$ for some constant $K > 0$. Instead of repeating the complete set of arguments of [1] in order to prove analogous results on $\mathcal{G}_K^\wedge$, we summarize the arguments briefly, and proceed to prove the necessary conditions required for them to work.

Following the ideas of [1] and previously of [10], define the following system of coupled ODEs:

$$\begin{cases} y' &= \dfrac{-y}{1 + (K-1)(1-y)^2} \\ y(0) &= 1 \end{cases}, \tag{1}$$

$$\begin{cases} z' &= \dfrac{z^2 + (K-1)(z-y)^2}{1 + (K-1)(1-y)^2} \\ z(0) &= 1 \end{cases}. \tag{2}$$

We further define:

$$\begin{cases} w' &= \dfrac{y^2 - 2wy - 2Kw(1-y)}{1 + (K-1)(1-y)^2} \\ w(0) &= 0 \end{cases}, \tag{3}$$

where $y$ is the solution to (1). Theorem 1.2 states that $I(G)$, $S(G)$ and $I_2(G)$ along the process $\mathcal{G}_K^\wedge$ are approximated by the solutions to the ODEs above, and that $t_g^\wedge$ is equal to the singularity point of the solution to (2) for any $K > 0$:

**Theorem 1.2.** *Let $y(t)$, $z(t)$ and $w(t)$ denote the solutions for* (1),(2),(3)*, and let $x_c$ denote the singularity point of $z(t)$ if such exists, and $\infty$ otherwise. For $0 < \delta < 1$, let $\tau_\delta > 0$ be the minimal point satisfying $y(\tau_\delta) \leq \delta$. The following statements hold almost surely:*

1. *For every $\delta > 0$, $|I(\mathcal{G}_K^\wedge(t)) - y(t)| = o(1)$ and $|I_2(\mathcal{G}_K^\wedge(t)) - w(t)| = o(1)$ for all $0 \leq t \leq \tau_\delta$.*
2. *For every $\varepsilon, \delta > 0$, $|S(\mathcal{G}_K^\wedge(t)) - z(t)| = o(1)$ for all $0 \leq t \leq \min\{\tau_\delta, x_c - \varepsilon\}$.*
3. *For all $K > 0$, $t_g^\wedge = x_c$.*

The value of $t_g^\wedge$ in the special case $K = 0$ and the asymptotic behavior of $t_g^\wedge$ are stated in the following theorem:

**Theorem 1.3.** *The threshold for the appearance of a giant component, $t_g^\wedge(K)$, is a continuous function of $K$ on $(0, \infty)$, and satisfies:*

$$\begin{cases} t_g^\wedge(0) & = & \frac{3}{2} + \frac{4}{3e^2 - 1} \\ t_g^\wedge(K) & = & (1 + o(1))\, \frac{\pi}{2\sqrt{2}} \left(1 + \frac{\pi^2}{24}\right) \frac{1}{\sqrt{K}} \end{cases},$$

*where the $o(1)$-term tends to $0$ as $K \to \infty$.*

The rest of this paper is organized as follows: in Section 2, we prove Theorem 1.1 by examining second moments of processes which are easier to analyze, and provide results of computer simulations of $t_c^\vee$ and $t_c^\wedge$. In Section 3, we prove Theorems 1.2 and 1.3 using analysis of differential equations, and provide results of computer simulations of $t_g^\wedge$.

## 2 Thresholds for connectivity

The proof of Theorem 1.1 uses the following result of [1]:

**Definition.** *Let $M \in \mathbb{N}$. An $M$*-`bounded weighted graph process` *on $n$ vertices, $\mathcal{H} = \mathcal{H}(n)$, is an infinite sequence of graphs on $n$ vertices, $(\mathcal{H}^0, \mathcal{H}^1, \ldots)$, where $\mathcal{H}^0$ is some fixed initial graph, and $\mathcal{H}^t$ is generated from $\mathcal{H}^{t-1}$ by adding one edge at random, according to a distribution of the following type: the probability of adding the edge $e$ to $\mathcal{H}^{t-1}$ is proportional to some weight function $W_t(e)$, satisfying:*

$$\max_{e \notin \mathcal{H}^{t-1}} W_t(e) \leq M \min_{e \notin H^{t-1}} W_t(e) \ .$$

*If for some $\nu \geq 0$ $\mathcal{H}^\nu = K_n$, we define $\mathcal{H}^t = \mathcal{H}^\nu = K_n$ for every $t > \nu$.*

**Theorem 2.1** ([1]). *Let $\mathcal{H}$ denote an $M$-bounded weighted graph process on $n$ vertices, and let $\mathcal{A}$ denote a monotone increasing property of graphs on $n$ vertices. The following statements hold for any $t \in \mathbb{N}$:*

$$\begin{aligned} \Pr[\mathcal{H}^t \in \mathcal{A}] &\leq \Pr[{\mathcal{G}_1}^{Mt}|_{\mathcal{H}^0} \in \mathcal{A}] \ , &(4)\\ \Pr[{\mathcal{G}_1}^{t}|_{\mathcal{H}^0} \in \mathcal{A}] &\leq \Pr[\mathcal{H}^{Mt} \in \mathcal{A}] \ . &(5)\end{aligned}$$

Notice that for every $K > 0$, the processes $\mathcal{G}_K^\vee$ and $\mathcal{G}_K^\wedge$ are both $M$-bounded, where $M = \lceil \max\{\frac{1}{K}, K\} \rceil$. As it is well-known that $I(t)$, the fraction of isolated vertices is $(1 + o(1)) \exp(-C)$ at time $Cn/2$, Theorem 2.1 implies that for any fixed $C$ and $K > 0$, $\mathcal{G}_K^\vee(C)$ and $\mathcal{G}_K^\wedge(C)$ almost surely contain $\Theta(n)$ isolated vertices. On the other hand, it is not difficult to see that $\mathcal{G}_0^\vee(1)$ has at most 1 isolated vertex, and we will later show that $\mathcal{G}_0^\wedge(\frac{3}{2} + o(1))$ almost surely has no isolated vertices. For this reason, we treat the cases $K > 0$ and $K = 0$ separately. Lemmas 2.2 and 2.3 below establish the precise behavior of these parameters when $K > 0$:

**Lemma 2.2.** *Let $K > 0$. For every $0 < \varepsilon < 1$, $\mathcal{G}_K^\vee[1+\varepsilon]$ is almost surely connected, whereas $\mathcal{G}_K^\vee[1-\varepsilon]$ is almost surely disconnected. Altogether, $t_c^\vee = 1$.*

**Lemma 2.3.** *Let $K > 0$. For every $0 < \varepsilon < 1$, $\mathcal{G}_K^\wedge[(1+\varepsilon)\max\{\frac{1}{2}, K\}]$ is almost surely connected, whereas $\mathcal{G}_K^\wedge[K-\varepsilon]$ and $\mathcal{G}_K^\wedge[\frac{1}{2}-\varepsilon]$ are almost surely disconnected. Altogether, $t_c^\wedge = \max\{\frac{1}{2}, K\}$.*

As we stated in the introduction, combining the fact that $t_c = 1$ for $\mathcal{G}_1$ with Theorem 2.1 yields that $t_c^\wedge(K) \leq \lceil \max\{\frac{1}{K}, K\} \rceil$, and indeed Lemma 2.3 shows that for $K \geq 1$, $K \in \mathbb{N}$, this maximum is achieved.

The remaining case $K = 0$ is settled by the next lemma:

**Lemma 2.4.** *The threshold for connectivity for $K = 0$ satisfies $t_c^\vee(0) = t_c^\wedge(0) = \frac{1}{2}$.*

## 2.1 Proof of Lemma 2.2

Let $0 < \varepsilon < 1$; we first show that $t_c^\vee \leq 1 + \varepsilon$. It is easy and well known that for every $\varepsilon' > 0$, there exists some $c = c(\varepsilon')$ such that, with high probability, $\mathcal{G}_1(c)$ has a giant component of size at least $(1-\varepsilon')n$: assume that indeed this holds. Take $0 < \varepsilon' < \min\{\frac{\varepsilon/4}{1+\varepsilon}, K\}$ and $c = c(\varepsilon')$ as above, and notice that:

$$\frac{1-\varepsilon'}{1+\varepsilon'} \geq \frac{1+\frac{\varepsilon}{2}}{1+\varepsilon} \ . \tag{6}$$

By Theorem 2.1, the giant component of ${\mathcal{G}_K^\vee}^{c'n/2}$ typically contains at least $(1-\varepsilon')n$ vertices, where $c' = \lceil \max\{K, \frac{1}{K}\} \rceil c$. Let $H^i$ denote the largest component of ${\mathcal{G}_K^\vee}^i$, and let $A_v^i$, $v \in V$, denote the event: $(v \notin H^i)$. The following holds for every $i \geq c'n/2$:

$$\Pr[\neg A_v^{i+1} | A_v^i] = \frac{K|H^i|}{K(\binom{n}{2} - i) - (K-1)\binom{|\mathcal{C}_1^i|}{2}} \geq \frac{K(1-\varepsilon')n}{K\binom{n}{2} + \binom{|\mathcal{C}_1^i|}{2}} \geq \frac{1-\varepsilon'}{(1+\varepsilon')\frac{n}{2}} \ , \tag{7}$$

where the last inequality is by the fact that $\frac{|\mathcal{C}_1^i|}{n} \leq \varepsilon' < K$. Setting $T_0 = c'n/2$, this gives the following upper bound on $A_v^T$ for $v \notin H^{T_0}$:

$$\Pr[A_v^T | \neg A_v^{T_0}] = \prod_{j=T_0}^{T} \Pr[A_v^j | A_v^{j-1}] \leq \prod_{j=T_0}^{T} \left(1 - \frac{1-\varepsilon'}{(1+\varepsilon')\frac{n}{2}}\right) \leq$$

$$\leq \exp\left(-(T-T_0)\frac{1-\varepsilon'}{(1+\varepsilon')\frac{n}{2}}\right) . \tag{8}$$

Thus, for $T = (1+\varepsilon)\frac{n}{2}\log n$ we get:

$$\Pr[A_v^T] \leq \exp\left(-\frac{1-\varepsilon'}{1+\varepsilon'}\left((1+\varepsilon)\log n - c'\right)\right) \leq \exp\left(-(1+\frac{\varepsilon}{2})\log n + O(1)\right) = o(n^{-1}) ,$$

hence a union bound on the vertices of $V \setminus H^{T_0}$ implies that, with high probability, $\mathcal{G}_K^\vee[1+\varepsilon]$ is connected.

The lower bound on $t_c^\vee$ is slightly more delicate. A simple way to show the lower bound in $\mathcal{G}_1$ is to apply a second moment argument on the number of isolated vertices of $\mathcal{G}_1(1)$. However, in the case of $\mathcal{G}_K^\vee$, using uniform upper and lower bounds on $\mathcal{C}_1^i$ (such as $\varepsilon' n$ and $o(n)$) does not yield useful bounds on the above second moment. The following claim resolves this difficulty:

**Claim 2.5.** *Let $0 < \delta < \frac{1}{2}$, and let $\{H_0\}_n$ denote a family of arbitrary graphs on $n$ vertices with $(1+o(1))n^{1-\delta}$ isolated vertices. If $\mathcal{H}$ is a biased $\mathcal{G}_K^\vee$ process on $n$ vertices which begins with $H_0$, that is: $\mathcal{H} \sim \mathcal{G}_K^\vee|_{H_0}$, then $\mathcal{H}[1-2\delta]$ almost surely contains isolated vertices.*

Indeed, to obtain the lower bound on $t_c^\vee$ from the above claim, fix $\delta = \varepsilon/2$, and let $\tau$ denote the minimal time at which $\mathcal{C}_1^\tau \leq n^{1-\delta}$. Taking $H_0 = {\mathcal{G}_K^\vee}^\tau$, Claim 2.5 implies that $\mathcal{G}_K^\vee|_{H_0}$ almost surely has isolated vertices at time $\tau + (1-2\delta)\frac{n}{2}\log n$, and in particular, we obtain that $t_c^\vee \geq 1-\varepsilon$, as required. It remains to prove Claim 2.5:

*Proof of claim.* Let $H_0$ and $\delta$ be as above, and let $\mathcal{H} \sim \mathcal{G}_K^\vee|_{H_0}$. Let $B_u^i$ denote the event that the vertex $u$ is isolated at time $i$, where $i \leq n\log n$. Clearly:

$$\Pr[\neg B_u^{i+1} | B_u^i] = \frac{K(n - |\mathcal{C}_1^i|) + (|\mathcal{C}_1^i| - 1)}{K(\binom{n}{2} - i) - (K-1)\binom{|\mathcal{C}_1^i|}{2}} .$$

By the assumption on $H_0$, $|\mathcal{C}_1^i| \leq |\mathcal{C}_1^0| = (1+o(1))n^{1-\delta}$, thus:

$$\Pr[\neg B_u^{i+1} | B_u^i] = \frac{2 + O(n^{-\delta})}{n\left(1 + O(\frac{\log n}{n}) + O(n^{-2\delta})\right)} = \frac{2 + O(n^{-\delta})}{n} . \tag{9}$$

Similarly, we can define $B_{u,v}^i = B_u^i \wedge B_v^i$ for $i \leq n\log n$ and get:

$$\Pr[\neg B_{u,v}^{i+1} | B_{u,v}^i] = \frac{2K(n - |\mathcal{C}_1^i|) + 2(|\mathcal{C}_1^i| - 2) + 1}{K(\binom{n}{2} - i) - (K-1)\binom{|\mathcal{C}_1^i|}{2}} = \frac{4 + O(n^{-\delta})}{n} . \tag{10}$$

Let $T = (1-2\delta)\frac{n}{2}\log n$, and define $Y = \sum_u \mathbf{1}_{B_u^T} = |\mathcal{C}_1^T|$ to be the number of isolated vertices of $\mathcal{H}^T$. A straightforward second moment consideration implies that $Y > 0$ almost surely. To see this, first note that (9) along with the well known bound $1 - x \geq \mathrm{e}^{-x/(1-x)}$ for $0 \leq x < 1$ yield:

$$\mathbb{E}Y = |\mathcal{C}_1^0| \prod_{i=1}^{T} \Pr[B_u^i | B_u^{i-1}] = (1+o(1))n^{1-\delta}\left(1 - \frac{2+O(n^{-\delta})}{n}\right)^T \geq$$

$$\geq (1+o(1))n^{1-\delta} \exp\left(-\frac{(2+o(1))\,/n}{1-\frac{2+o(1)}{n}}(1-2\delta)\frac{n}{2}\log n\right) = n^{1-\delta}n^{-1+2\delta+o(1)} = n^{\delta+o(1)}\ . \qquad (11)$$

By (9) and (10), there exist $p = p(n)$ and $q = q(n)$ such that $\Pr[\neg B_{u,v}^{i+1} | B_{u,v}^i] \geq p = \frac{4+O(n^{-\delta})}{n}$ and $\Pr[\neg B_u^{i+1} | B_u^i] \leq q = \frac{2+O(n^{-\delta})}{n}$. The following holds:

$$\mathrm{Cov}(\mathbf{1}_{B_u^T}, \mathbf{1}_{B_v^T}) = \Pr[B_{u,v}^T] - \Pr[B_u^T]^2 \leq (1-p)^T - (1-q)^{2T} =$$

$$= \left(1-p-(1-q)^2\right)\sum_{i=0}^{T-1}(1-p)^i(1-q)^{2(T-1-i)} = \frac{O(n^{-\delta})}{n}T\left(1 - \frac{4+o(1))}{n}\right)^{T-1} \leq$$

$$\leq (1-2\delta)\log n \cdot O(n^{-\delta})\exp\left(-(1+o(1))\frac{4}{n}(1-2\delta)\frac{n}{2}\log n\right) = n^{-2+3\delta+o(1)}\ .$$

Therefore,

$$\sum_{u\in\mathcal{C}_1^0}\sum_{v\in\mathcal{C}_1^0}\mathrm{Cov}(\mathbf{1}_{B_u^T}, \mathbf{1}_{B_v^T}) \leq n^{2(1-\delta)}n^{-2+3\delta+o(1)} = n^{\delta+o(1)} = o\left((\mathbb{E}Y)^2\right)\ .$$

As $\mathrm{E}Y = \omega(1)$, applying Chebyshev's inequality gives:

$$\Pr[Y=0] \leq \frac{\mathrm{Var}Y}{(\mathbb{E}Y)^2} \leq \frac{\mathbb{E}Y + \sum_u\sum_v \mathrm{Cov}(\mathbf{1}_{B_u^T}, \mathbf{1}_{B_v^T})}{(\mathbb{E}Y)^2} = \frac{1}{\mathbb{E}Y} + o(1) = o(1)\ .$$

This completes the proof of the claim and of Lemma 2.2. ■

## 2.2 Proof of Lemma 2.3

Let $0 < \varepsilon < 1$. The bounds $t_c^\wedge \leq (1+\varepsilon)\max\{\frac{1}{2}, K\}$ and $t_c^\wedge \geq K - \varepsilon$ will follow from arguments similar to the ones in the proof of Lemma 2.2, whereas the bound $t_c^\wedge \geq \frac{1}{2} - \varepsilon$ requires more work.

To prove that $t_c^\wedge \leq (1+\varepsilon)\max\{\frac{1}{2}, K\}$, take $\varepsilon' > 0$ which satisfies:

$$\begin{cases} (1-\varepsilon') \geq \dfrac{(1+\varepsilon/4)^2}{1+\varepsilon} \\ 1-(1-\varepsilon')^2 + (1-\varepsilon')^2 K < \left(1+\dfrac{\varepsilon}{4}\right)K \end{cases}\ . \qquad (12)$$

For instance, the following choice is suitable:

$$\varepsilon' < \min\left\{\frac{\varepsilon K}{4}, 1 - \sqrt{1 - \frac{\varepsilon}{2}\min\{K,1\}}\right\}\ .$$

As before, take $c = c(\varepsilon')$ such that, with high probability, $\mathcal{G}_1(c)$ has a giant component of size at least $(1-\varepsilon')n$. Defining $H^i$ to be the largest component of ${\mathcal{G}_K^{\wedge}}^i$, Theorem 2.1 implies that $|H^{c'n/2}| \geq (1-\varepsilon')n$ with high probability, where $c' = \lceil \max\{K, \frac{1}{K}\} \rceil c$, and therefore we may assume that this indeed holds. Let $A_u^i$ denote the event that a vertex $u \in V \setminus H^i$, which belongs to some connected component $C$, joins the giant component at time $i+1$. The following holds:

$$\Pr[A_u^i] = \frac{|H^i| w_C}{\binom{n}{2} + (K-1)\binom{n-|\mathcal{C}_1^i|}{2} - Ki} \text{ , where } w_C = \left\{ \begin{array}{cl} 1 & \text{if } |C| = 1 \\ K|C| & \text{otherwise} \end{array} \right. . \tag{13}$$

By the choice of $c'$, for all $i \geq c'n/2$ we get:

$$\Pr[A_u^i] \geq \frac{w_C(1-\varepsilon')n}{\binom{n}{2} + (K-1)\binom{n-|\mathcal{C}_1^i|}{2}} \geq \frac{2w_C(1-\varepsilon')}{n\left(1-(1-\frac{|\mathcal{C}_1^i|}{n})^2 + K(1-\frac{|\mathcal{C}_1^i|}{n})^2\right)} .$$

For $K \geq 1$, the denominator in the expression above is clearly bounded from above by $nK$, and for $0 < K < 1$ it is bounded from above by $n\left(1-(1-\varepsilon')^2 + (1-\varepsilon')^2 K\right)$, hence (12) implies:

$$\Pr[A_u^i] \geq \frac{2w_C(1-\varepsilon')}{nK(1+\varepsilon/4)} \geq \frac{2w_C(1+\varepsilon/4)}{nK(1+\varepsilon)} . \tag{14}$$

Take $T_0 = c'n/2$ and $T = \max\{\frac{1}{2}, K\}(1+\varepsilon)\frac{n}{2}\log n$, and define the following event for every $u \notin H^{T_0}$:

$$B_u^T = (u \notin H^T) .$$

The definition of $w_C$ in (13) implies that $w_C \geq \min\{1, 2K\}$ and when combined with (14) this gives:

$$\Pr[B_u^T] \leq \prod_{i=T_0}^{T} \left(1 - \frac{2\min\{1,2K\}(1+\varepsilon/4)}{nK(1+\varepsilon)}\right) \leq$$

$$\leq \exp\left(-\frac{1+\varepsilon/4}{1+\varepsilon} \cdot \frac{\min\{\frac{1}{K}, 2\}}{n/2}\left(\max\{\frac{1}{2}, K\}(1+\varepsilon)\frac{n}{2}\log n - c'\frac{n}{2}\right)\right) = n^{-(1+\varepsilon/4)+o(1)} .$$

Therefore, the expected number of vertices of $V \setminus H^{T_0}$ which do not join $H^T$ is clearly $o(1)$, implying that $\mathcal{G}_K^{\wedge}[(1+\varepsilon)\max\{\frac{1}{2}, K\}]$ is almost surely connected.

The lower bound $t_c^{\wedge} \geq K - \varepsilon$ follows from the following claim, analogous to Claim 2.5:

**Claim 2.6.** *Let $0 < \delta < \frac{1}{2}$, and let $\{H_0\}_n$ denote a family of arbitrary graphs on $n$ vertices with $(1+o(1))n^{1-\delta}$ isolated vertices. If $\mathcal{H} \sim \mathcal{G}_K^{\wedge}|_{H_0}$, then $\mathcal{H}[K-2\delta]$ almost surely contains isolated vertices.*

*Proof.* As in the proof of Claim 2.5, let $B_u^i$ denote the event ($u$ is isolated at time $i$) and let $B_{u,v}^i = B_u^i \wedge B_v^i$, where $u, v \in V(H_0)$. Clearly:

$$\Pr[\neg B_u^{i+1} | B_u^i] = \frac{n-1}{\binom{n}{2} + (K-1)\binom{n-|\mathcal{C}_1^i|}{2} - Ki} .$$

The assumption on $H_0$ gives $|\mathcal{C}_1^i| \leq |\mathcal{C}_1^0| = (1+o(1))n^{1-\delta}$, and thus for all $i = O(n\log n)$:

$$\Pr[\neg B_u^{i+1} | B_u^i] = \frac{2}{Kn\left(1 + O(\frac{\log n}{n}) + O(n^{-\delta})\right)} = \frac{2 + O(n^{-\delta})}{Kn} \ , \tag{15}$$

and:

$$\Pr[\neg B_{u,v}^{i+1} | B_{u,v}^i] = \frac{2n-3}{\binom{n}{2} + (K-1)\binom{n-|\mathcal{C}_1^i|}{2} - Ki} = \frac{4 + O(n^{-\delta})}{Kn} \ . \tag{16}$$

Henceforth, a similar calculation to the one in the proof of Claim 2.5 gives the required result. ∎

Choosing $\delta = \varepsilon/2$ and applying the last claim implies that $t_c^\wedge \geq K - \varepsilon$, and it remains to show that $t_c^\wedge \geq \frac{1}{2} - \varepsilon$. This follows from the fact that $\mathcal{G}_K^\wedge[\frac{1}{2} - \varepsilon]$ contains components of size 2 almost surely. Unfortunately, we cannot repeat the last argument in order to show that indeed this is the case, since we have no guarantee that at time $\tau$, the hitting time for the property $|\mathcal{C}_1| \leq n^{1-\delta}$, $\mathcal{G}_K^\wedge$ still satisfies $|\mathcal{C}_2| = n^{1-\alpha}$ for some small $\alpha$. To prove the next claim, which completes the proof of Lemma 2.2, we consider a simplified process, where it is possible to give a lower bound on $|\mathcal{C}_2|$, and show that this process stochastically dominates $\mathcal{G}_K^\wedge$ with regards to $|\mathcal{C}_2|$.

**Claim 2.7.** *For all $K \in (0,1]$ and $\varepsilon > 0$, $\mathcal{G}_K^\wedge[\frac{1}{2} - \varepsilon]$ almost surely contains a connected component of size* 2.

*Proof of claim.* Define the process $\widetilde{\mathcal{G}_K^\wedge}$, which is an *approximated* version of $\mathcal{G}_K^\wedge$, as follows: at each step, assign a weight $K$ to *ordered pairs* of the form $\{(u,v) : u,v \notin \mathcal{C}_1\}$, and a weight 1 otherwise. If the ordered pair $(u,v)$ chosen at some step $i$ is a (self) loop or corresponds to an edge which already exists in $\widetilde{\mathcal{G}_K^\wedge}$, this step is omitted. In other words, $\widetilde{\mathcal{G}_K^\wedge}$ assigns weights to all ordered pairs, and disregards selections of multiple edges or loops. Clearly, if $\widetilde{\mathcal{G}_K^\wedge}^t$ contains $m \leq t$ edges, then $\widetilde{\mathcal{G}_K^\wedge}^t \sim {\mathcal{G}_K^\wedge}^m$. Furthermore, for any $t = o(n^2)$ we have:

$$\mathbb{E}\left(t - |E(\widetilde{\mathcal{G}_K^\wedge}^t)|\right) \leq \sum_{i=1}^{t} \max\{\frac{1}{K}, K\}\frac{t+n}{n^2} = \left(O\left(\frac{t}{n^2}\right) + O\left(\frac{1}{n}\right)\right)t = o(t) \ ,$$

where $E(H)$ denotes the set of edges of the graph $H$. In particular, for any fixed $c$, with high probability $\mathcal{G}_K^\wedge(c) \sim \widetilde{\mathcal{G}_K^\wedge}(c + o(1))$ and $\mathcal{G}_K^\wedge[c] \sim \widetilde{\mathcal{G}_K^\wedge}[c + o(1)]$, and it is sufficient to prove the claim for $\widetilde{\mathcal{G}_K^\wedge}$.

Take $0 < \varepsilon < \frac{1}{2}$. By well known results on the original Erdős-Rényi graph process, for any fixed $t$, with high probability: $|\mathcal{C}_2(\mathcal{G}_1(t))| = \Theta(n)$, with the constant tending to 0 as $t \to \infty$. For a short proof of this fact, one can verify that the following differential equation approximates the graph parameter $I_2$ along $\mathcal{G}_1$ up to an $o(1)$ error term:

$$w'(t) = -2w(t) + y^2(t) \ , \ w(0) = 0 \ ,$$

where $y(t)$ is the approximating function of the fraction of isolated vertices, $I(t)$, and the time-scaling is of $n/2$ edges at each step. Substituting the well known fact (which also follows from Theorem 1.2 when substituting $K = 1$) that $y(t) = \exp(-t)$, it follows that $w(t) = t\exp(-2t)$.

By Theorem 2.1 and the above fact, we deduce that for every sufficiently large $c$ there exist $0 < \alpha_1 < \alpha_2 < \frac{1}{3}$ such that $\widetilde{\mathcal{G}_K^{\wedge}}(c)$ almost surely satisfies $\alpha_1 n \leq |\mathcal{C}_2| \leq \alpha_2 n$. Let:

$$c' = \inf\{t > c : y(t) \leq \frac{\varepsilon K}{2}\} ,$$

where $y$ is the solution to the ODE (1)), and set $\varepsilon' = y(c')$. By Theorem 1.2, with high probability $I\left(\widetilde{\mathcal{G}_K^{\wedge}}(c')\right) = y(c') + o(1)$, and in particular, $\widetilde{\mathcal{G}_K^{\wedge}}(c')$ has $(1 + o(1))\varepsilon' n$ isolated vertices almost surely. Let $H_0$ denote $\widetilde{\mathcal{G}_K^{\wedge}}(c')$, and take $\alpha_1 = \alpha_1(c')$, $\alpha_2 = \alpha_2(c')$ as above. According to these definitions, $S = \mathcal{C}_1(H_0)$ satisfies $s = |S| = (1 + o(1))\varepsilon' n$ almost surely, and $W = \mathcal{C}_2(H_0)$ satisfies $\alpha_1 n \leq |W| \leq \alpha_2 n$ almost surely. Assume that indeed this holds.

We consider the graph process $\mathcal{H}$, which begins with $H_0$, and at each step selects an ordered pair $(u, v) \in V(H_0)^2$ according to the following probabilities:

$$\begin{cases} \frac{1}{Kn^2} & \text{if } (u,v) \text{ is incident to } W \text{ and } S \\ \frac{1}{n^2} & \text{if } (u,v) \text{ is incident to } W \text{ and } V \setminus S \\ \lambda & \text{otherwise} \end{cases} ,$$

where the value of $\lambda > 0$ is chosen such that the probabilities sum up to 1. This is possible, since the probabilities for the first two types of pairs sum up to at most:

$$2\alpha_2 \left(\varepsilon'/K + (1 - \varepsilon')\right) \leq 3\alpha_2 < 1 ,$$

as $\varepsilon' \leq \frac{K}{2}$ and $\alpha_2 < \frac{1}{3}$. We claim that the following two events occur almost surely, and complete the proof of the claim:

$$\mathcal{C}_2(\mathcal{H}[\tfrac{1}{2} - \varepsilon]) \cap W \neq \emptyset , \tag{17}$$

$$\left|\mathcal{C}_2(\mathcal{H}[\tfrac{1}{2} - \varepsilon]) \cap W\right| \leq \left|\mathcal{C}_2(\widetilde{\mathcal{G}_K^{\wedge}}|_{H_0}[\tfrac{1}{2} - \varepsilon]) \cap W\right| . \tag{18}$$

A standard second moment consideration proves that (17) occurs with high probability. To see this, set: $T = \left(\frac{1}{2} - \varepsilon\right) \frac{n}{2} \log n$, and notice that:

$$\frac{2T}{Kn^2}\left(K(n - s - 2) + s\right) = \frac{1}{2}(1 - 2\varepsilon)\left((1 - \varepsilon') + \frac{\varepsilon'}{K} + o(1)\right)\log n = \frac{1}{2}(1 - \varepsilon'' + o(1))\log n \tag{19}$$

for some $\varepsilon \leq \varepsilon'' \leq 2\varepsilon$ (by our choice of $\varepsilon'$). Next, let the random variable $X_e$ ($e \in W$) be the indicator of the event: $\left(e \in \mathcal{C}_2(\mathcal{H}[\frac{1}{2} - \varepsilon])\right)$, and let $X = \sum_{e \in W} X_e$. The following holds:

$$\Pr[X_e = 1] = \left(1 - \frac{2K(n - s - 2) + 2s}{Kn^2}\right)^T ,$$

$$\Pr[X_e = 1 \wedge X_{e'} = 1] = \left(1 - \frac{4K(n - s - 2) - 4K + 4s}{Kn^2}\right)^T ,$$

for every $e, e' \in W$. Thus, a calculation similar to the one in the proof of Claim 2.5 gives:

$$\mathbb{E}X \geq |W| \exp\left(-(2 - o(1))T\frac{K(n - s - 2) + s}{Kn^2}\right) \geq |W| n^{(-1+\varepsilon'')/2 + o(1)} = \omega(\sqrt{n}) ,$$

and:

$$\mathrm{Cov}(X_e, X_{e'}) \le \frac{4K}{Kn^2} \exp\left(-4\frac{K(n-s-4)+K+s}{Kn^2}T\right) \le 4n^{-3+\varepsilon''+o(1)} .$$

Therefore:

$$\sum_{e\in W}\sum_{e'\in W} \mathrm{Cov}(X_e, X_{e'}) \le 4|W|^2 n^{-3+\varepsilon''+o(1)} = o(\mathbb{E}X) ,$$

and in particular, $\mathrm{Var}(X) = (1+o(1))\mathbb{E}X = o\left(\mathbb{E}X)^2\right)$ and by Chebyshev's inequality we deduce that $X > 0$ almost surely.

It remains to prove that (18) occurs almost surely. This is achieved by a coupling argument: we claim that there exists a coupling of the processes $\widetilde{\mathcal{G}_K^{\wedge}}|_{H_0}$ and $\mathcal{H}$ whose support consists of pairs $(G_t, H_t)$ such that: $G_t \sim \widetilde{\mathcal{G}_K^{\wedge}}|_{H_0}^t$, $H_t \sim \mathcal{H}^t$, and $(\mathcal{C}_2(H_t)\cap W) \subset (\mathcal{C}_2(G_t)\cap W)$. This clearly holds for $t = 0$, and by induction, it remains to extend the coupling from $(G_t, H_t)$ to $(G_{t+1}, H_{t+1})$. For this purpose, we apply the following lemma of [1] (Lemma 2.2), which was first proved by Strassen [11] in a slightly different setting:

**Lemma 2.8** ([1])**.** *Let $U, V$ be two finite sets, and let $R \subset U \times V$ denote a relation on $U, V$. Let $\mu$ and $\nu$ denote probability measures on $U$ and $V$ respectively, such that the following inequality holds for every $A \subset U$:*

$$\mu(A) \le \nu(\{y \in V : xRy \textit{ for some } x \in A\}) . \tag{20}$$

*Then there exists a coupling $\varphi$ of $\mu, \nu$ whose support is contained in $R$.*

Let $U, V$ be the set of all $n^2$ ordered pairs selected, representing the next pair selected by $G_t$ and $H_t$ respectively. Let $\mu$ denote the probability measure of each selection in $U$ by $G_t$, and let $\nu$ denote the probability of each selection in $V$ by $H_t$. Define:

$$\begin{cases} X = \mathcal{C}_2(G_t)\cap W & , \quad X_{u,v} = \mathcal{C}_2\left(G_t \cup (u,v)\right)\cap W \\ Y = \mathcal{C}_2(H_t)\cap W & , \quad Y_{x,y} = \mathcal{C}_2\left(H_t \cup (x,y)\right)\cap W \end{cases} .$$

By the induction hypothesis, $Y \subseteq X$, and we define $R$ to be $\{((u,v),(x,y)) : Y_{x,y} \subseteq X_{u,v}\}$.

Clearly, if $(u,v)$ is a loop or an edge which already belongs to $G_t$, it has no effect on $X$, and as $Y_{x,y} \subseteq Y \subseteq X$ we get $(u,v)R(x,y)$ for all $x, y$. Furthermore, every $(u,v)$ which is not incident to any $e \in Y$ also satisfies $(u,v)R(x,y)$ for all $x, y$, as the components $(u,v)$ may remove from $X$ already do not belong to $Y$. Therefore, (20) is satisfied for every $A \subset U$ which contains such pairs $(u,v)$.

It remains to prove that (20) holds for sets $A \subset U$ such that $A \cap E(G_t) = \emptyset$, and $A$ consists entirely of edges incident to edges of $Y$. Let $A = \{e_1, \ldots, e_m\}$, and notice that $e_i \notin E(H_t)$ for all $i$, as $e_i$ is incident to some component $e \in Y$ of size 2 in $H_t$, satisfying $e \ne e_i$ as $Y \subset E(G_t)$. Furthermore, for all $i$ we have $e_i R e_i$, since all components that $e_i$ removes from $X$ in $X_{e_i}$ are also removed from $Y$ in $Y_{e_i}$. Thus, showing that $\mu(e_i) \le \nu(e_i)$ for all $i$ will imply that $A$ satisfies the condition of (20). Indeed, if $e_i = (u,v)$ is such that $u, v \notin S$ ($u, v$ are both non-isolated in $H_0$), then:

$$\mu(e_i) = \frac{K}{n^2 + (K-1)(n - \mathcal{C}_1(G^t))^2} \leq \frac{K}{Kn^2} = \nu(e_i) \ ,$$

and otherwise:

$$\mu(e_i) \leq \frac{1}{n^2 + (K-1)(n - \mathcal{C}_1(G^t))^2} \leq \frac{1}{Kn^2} = \nu(e_i) \ ,$$

by definition of the process $\mathcal{H}$, completing the proof of the claim and the proof of Lemma 2.2. ■

## 2.3 Proof of Lemma 2.4

To prove that $t_c^\vee(0) = \frac{1}{2}$, we recall the following easy facts stated in [1]: by definition, $\mathcal{G}_0^\vee$ adds edges between pairs of isolated vertices until no such pair is left. Hence, after adding $\lfloor n/2 \rfloor$ edges there is at most 1 isolated vertex in $\mathcal{G}_K^\vee$, and $\mathcal{G}_K^\vee$ behaves as $\mathcal{G}_1$ on $\lfloor n/2 \rfloor$ components of size 2 (and possibly 1 additional isolated vertex). Thus, $t_g^\vee(0) = 1 + \frac{1}{2} = \frac{3}{2}$.

Proving a connectivity threshold of $(\frac{1}{2} + o(1))\frac{n}{2}\log n$, we may assume $n$ is even: otherwise, the single isolated vertex from time $\lfloor n/2 \rfloor$ becomes connected almost surely at time $\omega(n)$, and its edge set accounts to at most $n - 1 = o(n \log n)$ edges, not affecting the threshold for connectivity. By the discussion above, for even values of $n$, $\mathcal{G}_0^\vee$ has a linear number of components of size 2 and no isolated vertices at the time of appearance of the giant component. Thus, we deduce from the arguments used for the proofs of Lemmas 2.2 and 2.3 that, with high probability, $\mathcal{G}_0^\vee[\frac{1}{2} - \varepsilon]$ still contains components of size 2, whereas $\mathcal{G}_0^\vee[\frac{1}{2} + \varepsilon]$ is connected. Altogether, $t_c^\vee(0) = \frac{1}{2}$.

It remains to show that $t_c^\wedge(0) = \frac{1}{2}$. Substituting $K = 0$ in equation (1), it reduces to the form:

$$y' = \frac{1}{y-2} \ , \ y(0) = 1 \ ,$$

provided that $y \neq 0$. This yields the solution:

$$y(t) = 2 - \sqrt{1 + 2t} \ . \tag{21}$$

Notice that $y(t)$ is strictly monotone decreasing, and reaches 0 at $t = \frac{3}{2}$. Therefore, if we denote by $\tau_0(\mathcal{G}_0^\wedge)$ the minimal time $t$ at which $I\left(\mathcal{G}_0^\wedge(t)\right) = 0$, Theorem 1.2 implies that for every $0 < \delta < 1$, $|\tau - \frac{3}{2}| < \delta$ almost surely. To see this, let $x_0$ be such that $y(x_0) = \delta$, and let $x_1 = \max\{x_0, \frac{3}{2} - \delta\}$. By Theorem 1.2, with high probability:

$$I(\mathcal{G}_0^\wedge(\frac{3}{2} - \delta)) \geq I\left(\mathcal{G}_0^\wedge(x_1)\right) = y(x_1) + o(1) \geq \frac{1}{2}y(x_1) > 0 \ ,$$

where the second from last inequality holds for sufficiently large values of $n$. On the other hand, $y(x_1) < \delta$, and hence $I\left(\mathcal{G}_0^\wedge(\frac{3}{2} + \delta)\right) \leq I\left(\mathcal{G}_0^\wedge(x_1 + \delta)\right) = 0$ almost surely, since each edge eliminates at least one isolated vertex in $\mathcal{G}_0^\wedge$.

Equation (3) takes the following form after substituting $K = 0$ and the value of $y(t)$ as it is given in (21), provided that $y \neq 0$:

$$w'(t) = \frac{y - 2w}{2 - y} = 2\frac{1 - w}{\sqrt{1 + 2t}} - 1 \ , \ w(0) = 0 \ ,$$

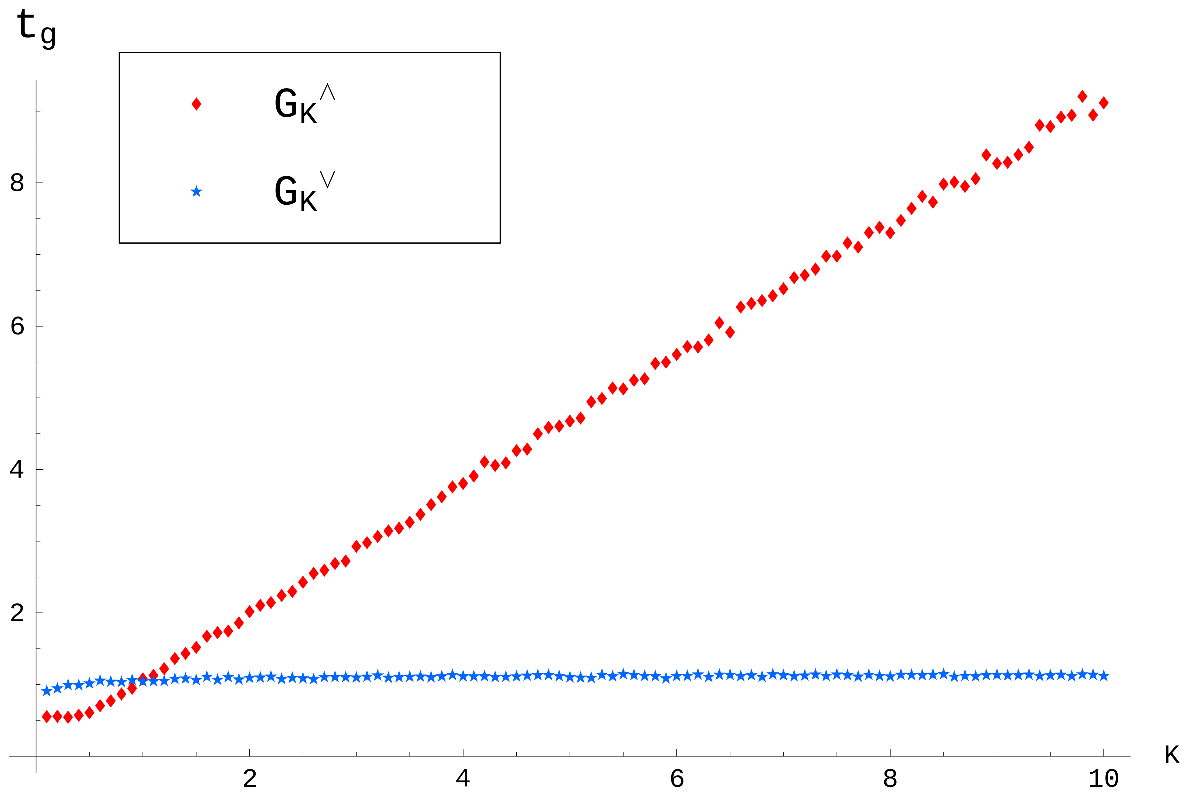


Figure 1: Comparison of the numerical results for $t_c^{\wedge}$ and $t_c^{\vee}$, and estimations of $t_c^{\wedge}$ and $t_c^{\vee}$ according to computer simulations of the model.

which has the solution:

$$w(t) = \frac{5}{4} - \frac{3}{4}\mathrm{e}^{2(1-\sqrt{1+2t})} - \frac{1}{2}\sqrt{1+2t}\ . \tag{22}$$

For $t = \frac{3}{2}$ we have $w(t) = \frac{1}{4} - \frac{3}{4\mathrm{e}^2} \approx 0.1485$, hence from the discussion above we obtain that $H = \mathcal{G}_K^{\wedge}(\tau)$ almost surely satisfies $I_2(H) = \Theta(n)$ (and $I(H) = 0$). From that point on, the process $\mathcal{G}_K^{\wedge}$ is equivalent to $\mathcal{G}_1|_H$, and from the similar arguments to those used in the proofs of Lemmas 2.2 and 2.3 we obtain that $t_c^{\vee}(0) = \frac{1}{2}$. ∎

### 2.4 Computer experiments of $\mathbf{t_c^{\wedge}}$ and $\mathbf{t_c^{\vee}}$

Maintaining the set of isolated vertices and the edges already added to the process provides all the information needed to add the next edge to $\mathcal{G}_K^{\wedge}$ and $\mathcal{G}_K^{\vee}$ at an $O(1)$ cost. In order to recognize the threshold for connectivity, the set of connected components must be efficiently maintained. Our implementation holds the components in linked-lists, according to the Weighted-Union Heuristic (see, e.g., [8] p. 445) which guarantees an average cost of $O(\log n)$ for uniting components.

Figure 1 shows the results of $t_c^{\wedge}$ and $t_c^{\vee}$ according to simulations of both models on $n = 10^4$ vertices. The values of $t_c^{\wedge}$ and $t_c^{\vee}$ were averaged over 100 tests per value of $K$.

# 3 The appearance of a giant component in $\mathcal{G}_{\mathbf{K}}^{\wedge}$

## 3.1 Proof of Theorem 1.2

We begin with a short summary of the methods used in [1] to analyze $t_g^{\vee}(K)$ for $K > 0$. First, the authors prove Theorem 2.1 and deduce that the process $\mathcal{G}_K^{\vee}$ is stochastically dominated by $\mathcal{G}_1$, up to a timescale factor of $\lceil \max\{\frac{1}{K}, K\} \rceil$. By applying the differential equation method of Wormald [12] to the approximated process $\widetilde{\mathcal{G}_K^{\vee}}$ (which selects an ordered pair at each step, similar to the process $\widetilde{\mathcal{G}_K^{\wedge}}$ introduced in the proof of Claim 2.7), the parameters $I(G)$ and $S(G)$ are approximated by $y(t)$ and $z(t)$, solutions to a system of coupled ODEs. Henceforth, a repeated use of Theorem 3.1 of [10], which relates the susceptibility and the appearance of a giant component, implies that the singularity point $z(t)$ is equal to $t_c^{\vee}$.

We note that the methods of [1] can be applied to any family of $M$-bounded processes $\mathcal{P}_M$ (as referred to in Theorem 2.1) provided that the following conditions hold:

1. Attempting to approximate a fixed number of bounded graph parameters $Y_1, \ldots, Y_d$ (such as $I(G)$ and $I_2(G)$) by functions $y_1, \ldots, y_d$ for $1 \le t \le T$, we require that

$$\mathbb{E}\frac{Y_i(\mathcal{P}_M^{t+1}) - Y_i(\mathcal{P}_M^t)}{2/n} = \frac{dy_i}{dt}\left(Y_1(\mathcal{P}_M^t), \ldots, Y_d(\mathcal{P}_M^t), t\right) + \mathrm{err_{y_i}(t)}\ , \tag{23}$$

for all $i$ and $t$, where $\max_{i,t} \mathrm{err}_{y_i}(t) = o(1)$.

2. In addition to the above approximations of $Y_1, \ldots, Y_d$ by $y_1, \ldots, y_d$, when attempting to approximate $S(G)$ by the function $z(t)$ for $1 \le t \le T$, we require that

$$\mathbb{E}\frac{S(\mathcal{P}_M^{t+1}) - S(\mathcal{P}_M^t)}{2/n} = \frac{dz}{dt}\left(Y_1(\mathcal{P}_M^t), \ldots, Y_d(\mathcal{P}_M^t), S(\mathcal{P}_M^t), t\right) + \mathrm{err_z(t)}\ , \tag{24}$$

where $\max_t \mathrm{err}_z(t) = o(1)$.

To show the above, we may use the fact that with high probability, the largest component is of size $O(\log n)$ as long as $t < x_c - \varepsilon$, where $x_c$ is a possible singularity point of $z(t)$. This follows from the proof of Theorem 1.3 of [1], which applies to this generalized setting as well.

3. Finally, if the above function $z(t)$ has a singularity point, we require that it is uniformly bounded (regardless of $M$).

If the above 3 conditions hold, we obtain that $Y_1, \ldots, Y_d, S$ are within $o(1)$ distance from $y_1, \ldots, y_d, z$ along the process $\mathcal{P}_M$ for $1 \le t \le T$. Furthermore, if $z$ has a singularity point at $x_c$ and the above conditions hold with $T = x_c - \varepsilon$ for any $\varepsilon > 0$, it follows that the appearance of a giant component in $\mathcal{P}_M$ is at $t = x_c$.

We first show that the above 3 conditions hold for $K > 0$ under the assumptions of Theorem 1.2. By the well known fact that for any constant $t$, $I(\mathcal{G}_1(t)) = \Theta(n)$ almost surely, Theorem 2.1

implies that $\tau_\delta$ can be taken to be arbitrarily large. In particular, we can take $\tau_\delta > x_c$. In order to verify that (23) holds for $I(G)$ and $I_2(G)$, let $G$ and $G'$ denote $\mathcal{G}_K^{\wedge\, t}$ and $\mathcal{G}_K^{\wedge\, t+1}$ respectively for some fixed $t$, and let $I = I(G)$ and $I_2 = I_2(G)$. We have:

$$\begin{aligned}\mathbb{E}\frac{I(G')-I(G)}{2/n} &= (-1)\frac{I(I-\frac{1}{n})}{1+(K-1)(1-I)^2} + (-1)\frac{2I(1-I)}{1+(K-1)(1-I)^2} = \\ &= \frac{-I}{1+(K-1)(1-I)^2} + O(I/n) \ , \end{aligned} \tag{25}$$

and:

$$\begin{aligned}\mathbb{E}\frac{I_2(G')-I_2(G)}{2/n} &= \frac{I(I-\frac{1}{n}) - 2I_2I - 2KI_2(1-I-I_2) - 2KI_2(I_2-\frac{2}{n})}{1+(K-1)(1-I)^2} = \\ &= \frac{I^2 - 2I_2I - 2KI_2(1-I)}{1+(K-1)(1-I)^2} + O(I/n) + O(I_2/n)) \ , \end{aligned} \tag{26}$$

where in both cases we used the fact that $K > 0$ to obtain an upper bound of $o(1)$ on the error term. To prove (24), set $S = S(G)$ and observe that:

$$\mathbb{E}\frac{S(G')-S(G)}{2/n} =$$

$$\begin{aligned} &= \frac{n/2}{1+(K-1)(1-I)^2}\Bigg(I(I-\frac{1}{n})\frac{2}{n} + 2I\sum_{\substack{C\in\mathcal{C}\\|C|>1}}\frac{|C|}{n}\frac{2|C|}{n} + \sum_{\substack{C_1\in\mathcal{C}\\|C_1|>1}}\sum_{\substack{C_2\in\mathcal{C}\setminus\{C_1\}\\|C_2|>1}}\frac{K|C_1||C_2|}{n^2}\frac{2|C_1||C_2|}{n}\Bigg) = \\ &= \frac{2I(S-I)+I^2+K(S-I)^2}{1+(K-1)(1-I)^2} + O(I/n) + O\Bigg(\sum_{\substack{C\in\mathcal{C}\\|C|>1}}\frac{|C|^4}{n^2}\Bigg) = \\ &= \frac{S^2+(K-1)(S-I)^2}{1+(K-1)(1-I)^2} + O(I/n) + O\Bigg(\sum_{\substack{C\in\mathcal{C}\\|C|>1}}\frac{|C|^4}{n^2}\Bigg) \ , \end{aligned} \tag{27}$$

and the assumption that $|C| = O(\log n)$ gives a bound of $o(1)$ on $\mathrm{err}_z$. The next claim therefore completes the statements of Theorem 1.2 for the case $K > 0$:

**Claim 3.1.** *For every $K > 0$, $t_g^\wedge(K) < 5$.*

*Proof.* Consider the case $K \geq 1$, and let $y(t)$ and $z(t)$ denote the solutions to the ODEs (1) and (2) respectively. Recalling that $y(t) \leq 1$ for every $t \geq 0$, (2) yields that $z'(t) \geq 0$ provided that $z(t) \geq 1$. Thus, the initial condition $z(0) = 1$ implies that $z(t)$ is monotone increasing in $t$, and in particular:

$$z' = \frac{z^2+(K-1)(z-y)^2}{1+(K-1)(1-y)^2} \geq \frac{z^2+(K-1)z^2(1-y)^2}{1+(K-1)(1-y)^2} = z^2 \ , \tag{28}$$

where inequality is by the fact that $y \geq 0$ and $z \geq 1$. By standard considerations in differential analysis , (28) and the initial condition $z(0) = 1$ imply that $z(t) \geq \frac{1}{1-t}$ for every $t \geq 0$ (as $\tilde{z}(t) = \frac{1}{1-t}$ satisfies $\tilde{z}' = \tilde{z}^2$ and $\tilde{z}(0) = 1$), and in particular, $t_g^\wedge \leq 1$ for any $K \geq 1$.

We are left with the case $0 < K < 1$. Clearly, (1) implies that $y' \le 0$ for every $t \ge 0$, and furthermore, $y' < 0$ as long as $y > 0$, hence $y$ is strictly monotone decreasing from 1 to 0. Let $t^*$ be such that

$$y(t^*) = 1 - \sqrt{1-K} \ . \tag{29}$$

Notice that the solutions to the equation $-x^2 + 2x - \frac{K}{1-K} = 0$ are: $x_{1,2} = 1 \pm \sqrt{1 - \frac{K}{1-K}}$ if $0 < K \le \frac{1}{2}$, and no solution exists if $\frac{1}{2} < K < 1$. In both cases, $-x^2 + 2x - \frac{K}{1-K} < 0$ for every $x \le 1 - \sqrt{1-K}$. Therefore, the definition of $t^*$ and the fact that $y$ is monotone decreasing give:

$$-y(t)^2 + 2y(t) - \frac{K}{1-K} \le 0 \ \text{ for every } t \ge t^*,$$

or equivalently:

$$(1-K)y(t)(2-y(t)) \le K \ \text{ for every } t \ge t^* \ . \tag{30}$$

Rewriting equation (2) as:

$$z' = \frac{z^2 - (1-K)(z-y)^2}{1-(1-K)(1-y)^2} = \frac{Kz^2 + (1-K)y(2z-y)}{K + (1-K)y(2-y)} \ , \tag{31}$$

it follows that for every $t \ge t^*$, $z' \ge \frac{1}{2}z^2$. Furthermore, for every $t \ge 0$, $z' > 0$, and hence $z(t^*) > z(0) = 1$. We obtain that the function $w(t) = z(t - t^*)$ satisfies $w' \ge \frac{1}{2}w^2$ for every $t \ge 0$ and $w(0) \ge 1$, and by the same consideration as above, $w(t) \ge \frac{2}{2-t}$ for every $t \ge 0$. Altogether, we deduce that $t_g^{\wedge} \le t^* + 2$, and it remains to provide an upper bound on $t^*$.

For this purpose, define $u(t) = 1 - y(t)$, and consider (1) for $0 \le t \le t^*$:

$$u' = -y' = \frac{1-u}{1-(1-K)u^2} \ge \frac{1-u}{1-u^4} = \frac{1}{1+u+u^2+u^3} \ ,$$

where the inequality is by the fact that $u(t) \le \sqrt{1-K}$ for $0 \le t \le t^*$. Define $w(t)$ to be the solution to the differential equation:

$$w' = \frac{1}{1+w+w^2+w^3} \ , \ w(0) = 0 \ , \tag{32}$$

it follows from the above mentioned argument that $u(t) \ge w(t)$ for $0 \le t \le t^*$. The solution to (32) satisfies: $t = \sum_{j=1}^{4} \frac{w^j}{j}$, hence $w(t_0) = \sqrt{1-K}$ for $t_0 = \sum_{j=1}^{4} \frac{(1-K)^{j/2}}{j} \le \frac{25}{12}$. As $u(t_0) \ge w(t_0)$, it follows that $t^* \le t_0 \le \frac{25}{12}$, completing the proof. ■

In the special case $K = 0$, $\mathcal{G}_0^{\wedge}$ is no longer an $M$-bounded process, however the assertions of statements $1, 2$ of Theorem 1.2 remain valid and follow from (23) and (24), by applying Wormald's differential equation method directly. To see that (23) holds, notice that as long as $I(\mathcal{G}_K^{\wedge}(t)) \ge \delta$ for some fixed $\delta > 0$, the denominators in (25) and (26) are $\Theta(1)$, and the approximation remains valid (note that $z(t)$ has no singularity point for $K = 0$). To show that (24) holds, set $S = S(G)$, and note that:

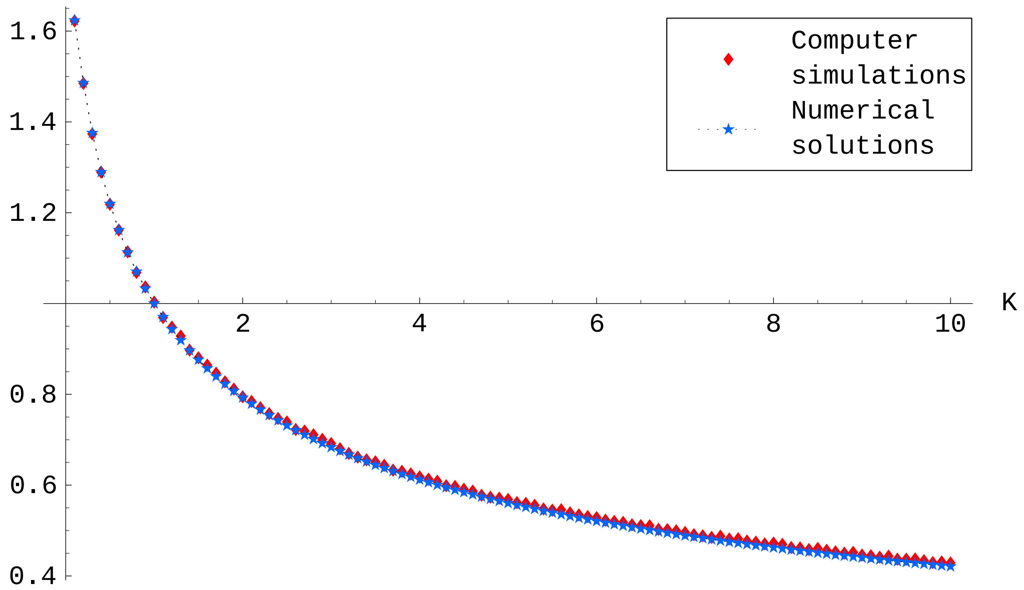


Figure 2: Comparison of the numerical results for $t_g^\wedge$, and estimations of $t_g^\wedge$ according to computer simulations of the model.

$$\begin{aligned}\mathbb{E}\frac{S(G')-S(G)}{2/n} &= \frac{n/2}{1-(1-I)^2}\left(I(I-\frac{1}{n})\frac{2}{n}+2I\sum_{\substack{C\in\mathcal{C}\\|C|>1}}\frac{|C|}{n}\frac{2|C|}{n}\right)= \\ &= \frac{2I(S-I)+I^2}{2I-I^2}+O(I/n)=\frac{2S-I}{2-I}+O(I/n)\ , \qquad (33)\end{aligned}$$

where we used the fact that $I>0$ for $t\le\tau_\delta$. This implies that the error term $\mathrm{err}_z$ is $o(1)$ without making any assumptions on the size of the largest component, and completes the proof of the theorem. ■

### 3.2 Computer experiments of $\mathbf{t_g^\wedge}$

We conducted simulations of $t_g^\wedge$ using the implementation of $\mathcal{G}_K^\wedge$ mentioned in 2.4. In these simulations, the number of vertices was $n=10^6$, and the threshold for the appearance of the giant component was taken to be the minimal time at which $\mathcal{G}_K^\wedge$ contains a component of size $\alpha n$, where $\alpha=0.01$. The value of $t_g^\wedge(K)$ was averaged over 10 tests for each value of $K$.

Figure 2 shows the comparison between the values of $t_g^\wedge$ according to the above computer simulations, and the values obtained by numerically solving the ODEs (1) and (2) by Mathematica.

### 3.3 Proof of Theorem 1.3

First, the fact that $t_c^{\wedge}(K)$ is continuous follows from the general continuous dependence of ODEs on their parameters (in this case,the single parameter $K$).

For the special case $K = 0$, recall that in our treatment of $\mathcal{G}_0^{\wedge}$ for the proof of Lemma 2.4 we showed that for every $\delta > 0$, with high probability $y(t) = 2 - \sqrt{1+2t}$ approximates $I(\mathcal{G}_0^{\wedge}(t))$ for $t \leq \frac{3}{2} - \delta$, and $I(\mathcal{G}_0^{\wedge}(\frac{3}{2}+\delta)) = 0$. Take $\delta > 0$ and consider the interval $[0, \frac{3}{2} - \delta]$; equation (2) takes the following form when substituting $K = 0$ and the solution to $y(t)$:

$$z' = \frac{z^2 - (z-y)^2}{2y - y^2} = \frac{2z - y}{2 - y} = \frac{2(z-1)}{\sqrt{1+2t}} + 1 \ . \tag{34}$$

Taking $w = z - 1$, we obtain the linear equation:

$$w' - \frac{2}{\sqrt{1+2t}} w = 1 \ . \tag{35}$$

Multiplying (35) by its integrating factor and integrating by parts, we get:

$$w(t) = \exp\left\{2\int (1+2t)^{-\frac{1}{2}} dt\right\} \int \exp\left\{-2\int (1+2t)^{-\frac{1}{2}} dt\right\} dx \ =$$

$$= -\frac{1}{2}(\sqrt{1+2t} + \frac{1}{2}) + C\exp(2\sqrt{1+2t}) \ .$$

The initial condition $w(0) = 0$ gives $C = \frac{3}{4}\mathrm{e}^{-2}$, hence:

$$z(t) = \frac{3}{4}\mathrm{e}^{2(\sqrt{1+2t}-1)} - \frac{1}{2}\sqrt{1+2t} + \frac{3}{4} \text{ for all } t \in [0, \frac{3}{2} - \delta] \ .$$

The above solutions for $y(t)$ and $z(t)$ give $y(\frac{3}{2}) = 0$ and

$$z(3/2) = \frac{3}{4}\mathrm{e}^2 - \frac{1}{4} \ . \tag{36}$$

Let $\tau_0 = \tau_0(\mathcal{G}_0^{\wedge})$ denote the first time $t$ at which $\mathcal{G}_0^{\wedge t}$ has no isolated vertices, i.e., the hitting time for the property: $\{G : I(G) = 0\}$. By the above arguments, we have:

$$\tau_0 = \left(\frac{3}{2} + o(1)\right) n/2 \tag{37}$$

almost surely. We claim that proving that with high probability:

$$S(\mathcal{G}_0^{\wedge \tau_0}) = z(3/2) + o(1) \tag{38}$$

implies the required result on $t_g^{\wedge}(0)$. Indeed, once no isolated vertices are left, the process $\mathcal{G}_0^{\wedge}$ adds edges according to the uniform distribution, and hence from that point the susceptibility follows the equation $z' = z^2$ (e.g., see the case $K = 1$ of the analysis of $\mathcal{G}_K^{\wedge}$ or $\mathcal{G}_K^{\vee}$). By (37), we obtain that for some $x_c > \frac{3}{2}$:

$$S(\mathcal{G}_0^{\wedge}(t)) = (1 + o(1))\frac{1}{x_c - t} \text{ for any } \frac{3}{2} \leq t < x_c \ ,$$

and the value of $x_c$ is derived from the initial condition (36):

$$t_g^{\wedge}(0) = x_c = \frac{3}{2} + \frac{4}{3e^2 - 1} \ . \tag{39}$$

It is left to show that (38) indeed holds. The lower bound $S(\mathcal{G}_0^{\wedge\tau_0}) \geq z(\frac{3}{2}) - o(1)$ follows from Theorem 1.2, which states that $z$ approximates $S$ until time $\frac{3}{2} - \delta$ for any $\delta > 0$, and from the continuity and monotonicity of $z$. That is, for any fixed $\xi > 0$, choosing a sufficiently small $\delta > 0$ such that $z(\frac{3}{2} - \delta) > z(\frac{3}{2}) - \xi$ gives:

$$S\left(\mathcal{G}_0^{\wedge}(\frac{3}{2} - \delta)\right) = z(\frac{3}{2} - \delta) + o(1) > z(\frac{3}{2}) - \xi + o(1) \ .$$

For the upper bound $S(\mathcal{G}_0^{\wedge\tau_0}) \leq z(\frac{3}{2}) - o(1)$ we are required to examine the second moment of $S\left(\mathcal{G}_0^{\wedge}(\frac{3}{2} + \delta)\right)$. Assume by contradiction that:

$$\Pr\left[S\left(\mathcal{G}_0^{\wedge\tau_0}\right) > z(\frac{3}{2}) + \xi\right] > \alpha \text{ for some fixed } \alpha, \xi > 0 \ , \tag{40}$$

and choose $\delta > 0$ small enough such that:

$$\left(\frac{z(\frac{3}{2})}{\xi/2}\right)^2 \left(e^{12\delta} - e^{2\delta}\right) < \alpha \quad \text{, and:} \quad \left(e^{4\delta} - 1\right) z(\frac{3}{2}) \leq \xi/2 \ . \tag{41}$$

Set $T_0 = \left(\frac{3}{2} - \delta\right)\frac{n}{2}$ and $T_1 = \tau_0 - 1$, and note that, with high probability, $\Delta := \frac{T_1 - T_0}{n/2}$ satisfies $\delta/2 \leq \Delta \leq 2\delta$, and therefore we may assume that this holds. We consider $\mathcal{G}_0^{\wedge T}$ for $T = T_0, \ldots, T_1$, and let $S_0 = S\left(\mathcal{G}_0^{\wedge T_0}\right)$. As $I(\mathcal{G}_0^{\wedge T}) > 0$, the calculation which yielded (34) gives:

$$\mathbb{E}\left(S(\mathcal{G}_0^{\wedge T+1}) \mid S(\mathcal{G}_0^{\wedge T})\right) = S(\mathcal{G}_0^{\wedge T}) + \frac{2}{n}\left(\frac{2(S(\mathcal{G}_0^{\wedge T}) - 1)}{2 - I(\mathcal{G}_0^{\wedge T})} + 1\right) \ ,$$

and hence:

$$\left(1 + \frac{2}{n}\right) S(\mathcal{G}_0^{\wedge T}) \leq \mathbb{E}\left(S(\mathcal{G}_0^{\wedge T+1}) \mid S(\mathcal{G}_0^{\wedge T})\right) \leq \left(1 + \frac{4}{n}\right) S(\mathcal{G}_0^{\wedge T}) \ .$$

Therefore:

$$e^{\Delta} S_0 \leq \mathbb{E}S(\mathcal{G}_0^{\wedge\tau_0}) \leq e^{2\Delta} S_0 \ .$$

Similarly, we can write the expression for the second moment of $S(\mathcal{G}_0^{\wedge T_1})$:

$$\mathbb{E}\left(S(\mathcal{G}_0^{\wedge T+1})^2 \mid S(\mathcal{G}_0^{\wedge T})^2\right) \leq S(\mathcal{G}_0^{\wedge T})^2 \left(1 + \frac{4}{n} \cdot \frac{3}{2 - I(\mathcal{G}_0^{\wedge T})}\right) \ ,$$

and hence:

$$\mathbb{E}\left(S(\mathcal{G}_0^{\wedge\tau_0})^2\right) \leq e^{6\Delta} S_0^2 \ .$$

We obtain that:

$$\mathrm{Var}S(\mathcal{G}_0^{\wedge\tau_0}) \leq \left(e^{6\Delta} - e^{\Delta}\right) S_0^2 \leq \left(e^{12\delta} - e^{2\delta}\right) S_0^2 \ ,$$

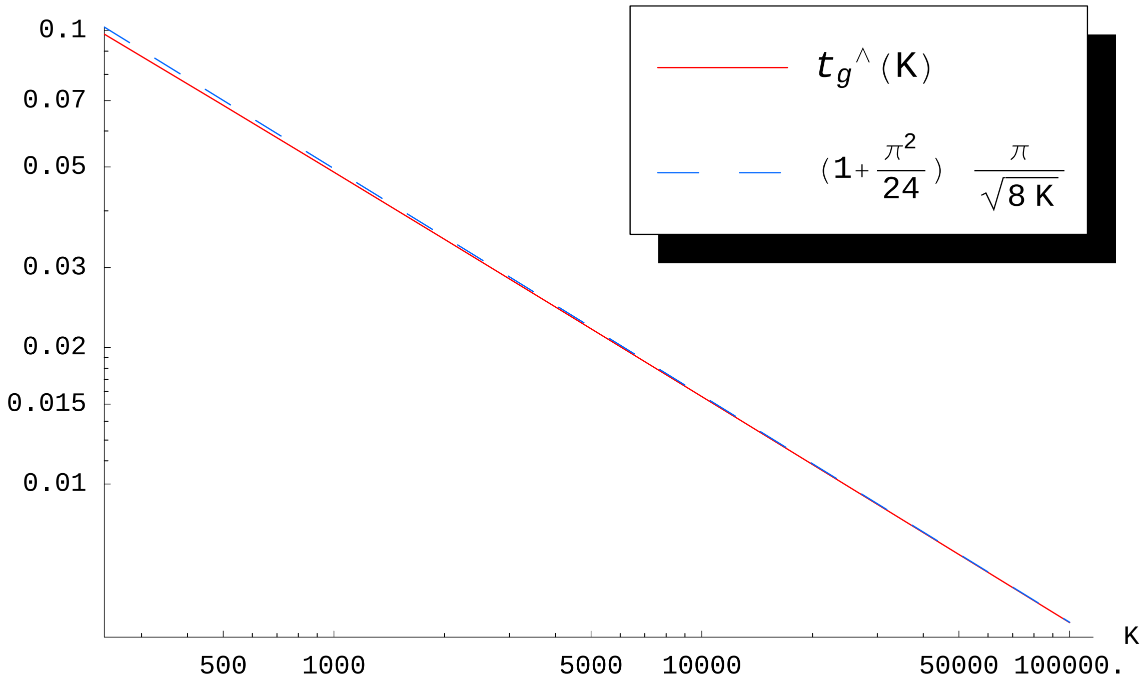


Figure 3: Comparison of the values of $t_g^{\wedge}(K)$, obtained by numerical solutions of the ODEs (1)-(2), and the asymptotic approximation of Theorem 1.3. Logarithmic scale was used in both axes.

where the last inequality is by the fact that $\Delta \leq 2\delta$ almost surely. By (41) and the fact that $S_0 < z(\frac{3}{2})$ almost surely, this implies that:

$$\mathrm{Var} S(\mathcal{G}_0^{\wedge\tau_0}) < \alpha\,(\xi/2)^2 \ . \tag{42}$$

By (41) we have:

$$\Pr\left[S\left(\mathcal{G}_0^{\wedge\tau_0}\right) > z(\frac{3}{2}) + \xi\right] \leq \Pr\left[S\left(\mathcal{G}_0^{\wedge\tau_0}\right) - \mathbb{E}\mathcal{G}_0^{\wedge\tau_0} > z(\frac{3}{2})\left(1 - \mathrm{e}^{4\delta}\right) + \xi\right] \leq$$
$$\leq \Pr\left[\left|S\left(\mathcal{G}_0^{\wedge\tau_0}\right) - \mathbb{E}\mathcal{G}_0^{\wedge\tau_0}\right| > \xi/2\right] \ ,$$

and thus combining Chebyshev's inequality with (42) gives:

$$\Pr\left[S\left(\mathcal{G}_0^{\wedge\tau_0}\right) > z(\frac{3}{2}) + \xi\right] < \alpha \ ,$$

contradicting the assumption (40).

It remains to show that for $K \gg 1$, $t_g^{\wedge}(K) = (1+o(1))\frac{\pi}{2\sqrt{2}}\left(1+\frac{\pi^2}{24}\right)\frac{1}{\sqrt{K}}$. Consider equation (1); the initial condition $y(0) = 1$ suggests that we examine the function $u(t) = 1 - y(t)$, which satisfies $u(t) \ll 1$ near the origin:

$$-u' = \frac{u-1}{u^2(K-1)+1} \ ,$$

and substituting the initial condition of $y$ we have:

$$u' = \frac{1-u}{u^2(K-1)+1} \ , \ u(0) = 0 \ . \tag{43}$$

As $u \ll 1$, the $u^2$-term at the denominator is negligible, and we have:

$$u' \approx \frac{1}{Ku^2+1} \ , \ u(0) = 0 \ , \tag{44}$$

where here and in what the follows, we denote by $\approx$ equality up to leading order terms. Rearranging the last equation to the form $Ku^2u' + u' \approx 1$ and integrating, we obtain that, up to leading order terms, $u$ satisfies the equation:

$$\frac{K}{3}u^3 + u \approx t \ . \tag{45}$$

We note that this immediately gives $u \approx t$ for $u \ll \frac{1}{\sqrt{K}}$, however we are interested in the behavior of $u$ precisely when $t = \Theta(\frac{1}{\sqrt{K}})$. Applying Caradano's solution to the above cubic equation gives the following approximation of $u$ when $t \ll 1$ (and hence $u \ll 1$):

$$u(t) \approx \left(\frac{3t}{2K} + \sqrt{\frac{1}{K^3} + \left(\frac{3t}{2K}\right)^2}\right)^{1/3} + \left(\frac{3t}{2K} - \sqrt{\frac{1}{K^3} + \left(\frac{3t}{2K}\right)^2}\right)^{1/3} \ , \ t \ll 1 \ .$$

Moving on to $z(t)$, we substitute $w = z - 1$ in equation (2) and obtain the following:

$$w' = \frac{(w+1)^2 + (K-1)(w+u)^2}{1 + Ku^2 - u^2} \approx \frac{(w+1)^2 + (K-1)(w+u)^2}{1+Ku^2} \ .$$

Next, we may replace $w+1$ by 1 whenever $w \ll 1$, and furthermore, whenever $w = \Omega(1)$ clearly the dominant term is $K(w+u)^2$. Altogether, we obtain the following uniform approximation for $w$:

$$w' \approx \frac{K(w+u)^2+1}{Ku^2+1} \approx \left(K(w+u)^2+1\right)u' \ .$$

Adding $u'$ to both sides of the equation and rearranging, we obtain:

$$\frac{w'+u'}{K(w+u)^2+2} \approx u' \ ,$$

hence if we define $v = \sqrt{\frac{K}{2}}(w+u)$ we obtain:

$$\frac{v'}{v^2+1} \approx \sqrt{2K}u' \ ,$$

and thus:

$$v \approx \tan(\sqrt{2K}u) \ .$$

Returning to $z$, we get:

$$z \approx 1 + \sqrt{\frac{2}{K}}\tan(\sqrt{2K}u) - u \ .$$

This implies that the singularity point $x_c$ satisfies $\sqrt{2K}u(x_c) = \frac{\pi}{2}$. Recalling equation (45), we have:

$$x_c \approx \frac{K}{3}u(x_c)^3 + u(x_c) = \frac{K}{3}\frac{\pi^3}{16K\sqrt{2K}} + \frac{\pi}{2\sqrt{2K}} = \frac{\pi}{2\sqrt{2}}\left(1 + \frac{\pi^2}{24}\right)\frac{1}{\sqrt{K}} \ .$$

Figure 3 shows the excellent agreement between the above asymptotic approximation of $t_g^{\wedge}(K)$, and its value as obtained by numerically solving the ODEs (1) and (2) by Mathematica. ■

**Acknowledgement** The authors wish to thank Noga Alon for helpful discussions and comments.

## References


[1] G. Amir, O. Gurel-Gurevich, E. Lubetzky, A. Singer, Giant components in biased graph processes, preprint.

[2] C.M. Bender and S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, Springer, New York (1999).

[3] B. Bollobás, *Random graphs*, volume 73 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, second edition (2001).

[4] T. Bohman and A. Frieze, Avoiding a giant component, Random Structures and Algorithms, 19 (2001), 75-85.

[5] T. Bohman and J.H. Kim, A phase transition for avoiding a giant component, Random Structures and Algorithms , to appear.

[6] T. Bohman and D. Kravitz, Creating a giant component, Combinatorics, Probability and Computing , to appear.

[7] M. Kang, Y. Koh, T. Łuczak and S. Ree, The connectivity threshold for the min-degree random graph process, Random Structures and Algorithms, to appear.

[8] T. H. Cormen, C.E. Leiserson, and R.L. Rivest, *Introduction to Algorithms*, The MIT Press/McGraw-Hill (1990).

[9] P. Erdős and A. Rényi, On the evolution of random graphs, Publ. math. Inst. Hungar. Acad. Sci., 5 (1960), 17-61.

[10] J. Spencer and N. Wormald, Birth control for giants, Combinatorica, to appear.

[11] V. Strassen, The existence of probability measures with given marginals, Annals of Mathematical Statistics, 36 (1965), 423-439.

[12] N.C. Wormald, The differential equation method for random graph processes and greedy algorithms, Lectures on Approximation and Randomized Algorithms, M. Karoński and H.J. Prömel (eds), PWN, Warsaw (1999), 73-155.