Supplemental materials

Sebastian M. Krause and Vinko Zlatić
Theoretical Physics Division, Rudjer Bošković Institute, Zagreb, Croatia

Michael M. Danziger
Department of Physics, Bar Ilan University, Ramat Gan, Israel

List of variables

	Networks
N	Number of nodes
$\bar{k}$	Average degree
k_i	Degree of node i
p_k	Degree distribution
α	Exponent of scale free degree distribution
g_0	Generating function of degree
g_1	Generating function of excess degree
	Colors
C	Number of colors
$c \in 1, 2, \ldots C$	A color
r_c	Color distribution
n_{deg}	Degeneration of the highest color frequency
$\tilde{r}_{c,k}$	degree-dependent color distribution
	Standard percolation ingredients
$\mathcal{L}$	Set of nodes in the largest component (color blind)
u	Prob. of not being connected to giant comp. over a link
S	Size of giant component
$\phi_{\bar{c}}$	Fraction of nodes without color c
$\mathcal{L}_{\bar{c}}$	Set of nodes in the largest component, after nodes of color c deleted
$u_{\bar{c}}$	Prob. of not being connected to giant $\mathcal{L}_{\bar{c}}$ over a link
$S_{\bar{c}}$	Size of giant $\mathcal{L}_{\bar{c}}$
	Percolation over color avoiding paths
$\mathcal{L}_{\text{color}}$	Candidate set of nodes for the largest avoidable colors component
S_{color}	Size of giant $\mathcal{L}_{\text{color}}$
$B_{k,k'}$	Prob. that out of k links k' connect to giant component
$M_{k',\vec{\kappa}}$	Prob. that out of k' links κ_1 connect to color 1 etc.
$P_{\vec{\kappa}}$	Success probability having neighbors of colors acc. to $\vec{\kappa}$
$U_{\bar{c}}$	Prob. that a link fails connecting to $\mathcal{L}_{\text{color}}$ which already connects to $\mathcal{L}$ and a node not having color c
$S_{\text{color},\infty}$	Size of the set of all nodes being connected to giant component over two links or more
β	Critical exponent
$\bar{k}_{\text{crit}}$	Critical value of average degree
k_{step}	Degree above which all nodes have the same color
γ	Fraction of nodes with highest degree

I. SIZE OF GIANT AVOIDABLE COLORS COMPONENT IN THE CONFIGURATION MODEL

We can find analytical results for S_{color} for random graph ensembles with randomly distributed colors in the limit of infinite graphs. These results can be used to estimate the situation in finite quenched networks. We are able to gain a general understanding including phase transitions. This can guide our understanding of real world networks.

We use the generalized configuration model graph ensemble with N nodes, where each degree sequences $\{k_i\}$ occurs with probability $\prod_i p_{k_i}$, with the degree distribution p_k. Additionally we want to assign to every node i a color $c_i \in 1, 2, \ldots, C$. For given degree sequence k_i, the color sequence $\{c_i\}$ has probability $\prod_i \tilde{r}_{c_i,k_i}$ with the degree-dependent color distribution $\tilde{r}_{c,k}$ ($\sum_c \tilde{r}_{c,k} = 1$ for every degree k separately). For a graph G_N out of the graph ensemble, $\mathcal{L}_{\text{color}}$ has a certain size $N_{\text{color}}(G_N)$. For the whole graph ensemble, we have to use the average value. By considering only giant contributions growing with network size, we have

$$S_{\text{color}} = \lim_{N\to\infty} \sum_{G_N} P(G_N) \frac{N_{\text{color}}(G_N)}{N}, \tag{1}$$

where $P(G_N) = \prod_i p_{k_i} \omega \prod_i \tilde{r}_{c_i,k_i}$ is the probability to have the graph G_N of size N, including ω, the probability of the connection scheme of G_N as a matching of half edges.

A. Question and connection to percolation theory

For calculating S_{color} in the random graph ensemble, we will follow ideas of Erdos and Renij [1] and Newman [2]. For calculating the size of the giant component, they used probabilities of connections for a single node in the graph ensemble. As we have to extend the method to a gradual procedure with conditional probabilities, it is useful to introduce the original method in detail with a shifted viewpoint.

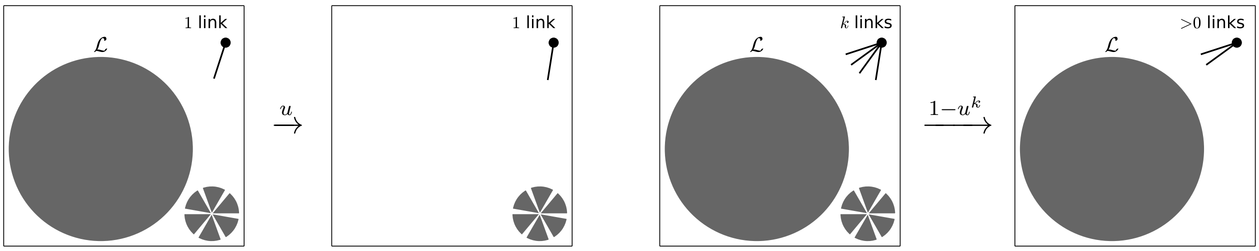

FIG. 1: We base our theory on the method to calculate the size of normal giant components, as illustrated in this figure. Using a self consistency equation, the probability u can be calculated. This is the probability, that a node is not connected to the giant component over a single link (see on the left). On the right, the probability for a node with k links is illustrated to have at least one link connecting to the giant component. u^k is the probability that all links fail.

Lets denote with $\mathcal{L}$ the set of all nodes belonging to the largest component. In figure 1 on the outer left, a possible situation is illustrated. The largest component contains of a large part of the network, and the remaining nodes belong to smaller components. We have to calculate the size S of the giant component, meaning the average relative size of $\mathcal{L}$ in the network ensemble in the limit of infinite network size. For this we can define the average probability u that a node fails to connect to $\mathcal{L}$ over one particular link. This is illustrated in the figure with the left part. Again, the thermodynamic limit $N \to \infty$ is implied. With the definition of u at hand, we can calculate S in two steps: First, using a self consistency equation, u is calculated. The probability u is identical to the probability that the neighbor does not connect to the giant component over any of the remaining links,

$$u = g_1(u), \quad g_1(z) = \sum_k q_k z^k. \tag{2}$$

In this equation, g_1 is the generating function of excess degree $q_k = (k+1)p_{k+1}/\bar{k}$. For important degree distributions as e.g. Poisson or scale-free, the equation for u can only be solved numerically. The second step is an averaging over nodes with various degrees k. The probability to connect to the giant component over any of k links is $(1 - u^k)$, meaning that not all links fail at the same time. This is illustrated in the figure on the right. As a node which connects to the giant component belongs to it,

$$S = \sum_{k=0}^{\infty} p_k(1 - u^k) = 1 - g_0(u), \quad g_0(z) = \sum_k p_k z^k. \tag{3}$$

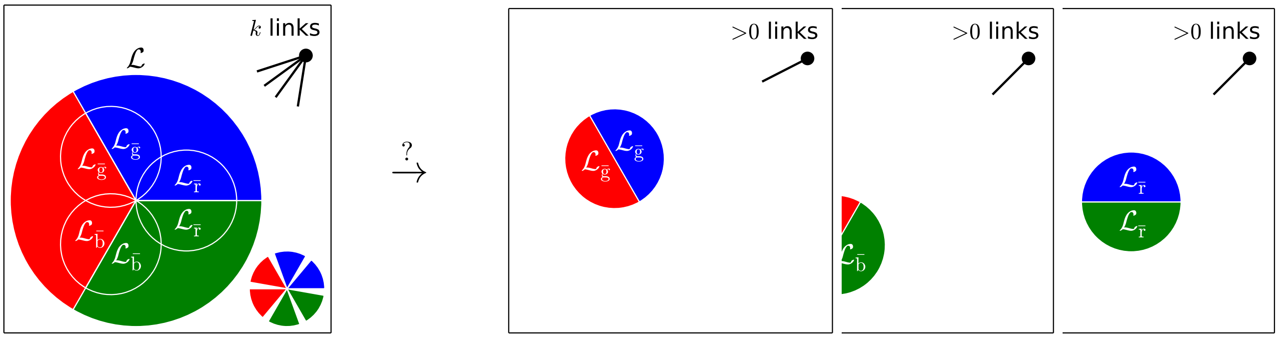

FIG. 2: We have to calculate the probability, if a node with k links is for every color c connected to the giant component $\mathcal{L}_{\bar{c}}$ with deleted color c. All connections over at least one link have to exist at the same time. We illustrate this question with the three colors red ($c=$r), green ($c=$g) and blue ($c=$b). If a link connects to $\mathcal{L}_{\bar{\mathrm{g}}}$, it for sure does not connect to $\mathcal{L}_{\bar{c}}$ for one of the other colors. This kind of dependence forces us to use a stepwise calculation with conditional probabilities.

In analogy to the procedure described above, we will calculate S_{color} as the probability that a randomly chosen node belongs to $\mathcal{L}_{\mathrm{color}}$. This has to be evaluated in the graph ensemble of infinite size. As we will perform an averaging over nodes with various degrees k, the following question has to be answered: What is the probability that a node with k links connects to a giant $\mathcal{L}_{\bar{c}}$ for all colors c at the same time. This is illustrated in figure 2. On the left, the situation for a graph with colors on the nodes is illustrated. Nodes of all colors might be in the largest component. After deleting all nodes of one color c, the remaining largest component $\mathcal{L}_{\bar{c}}$ might still contain a large part of all nodes in $\mathcal{L}$. The condition for the node belonging to $\mathcal{L}_{\mathrm{color}}$ is illustrated on the right of the figure.

We will use the same two steps to attack this problem, as described for calculating the giant component above. First, we provide some single link probabilities which can be used as primitives for the further calculations. Second, we combine the single link probabilities to calculate S_{color}.

B. Single link probabilities

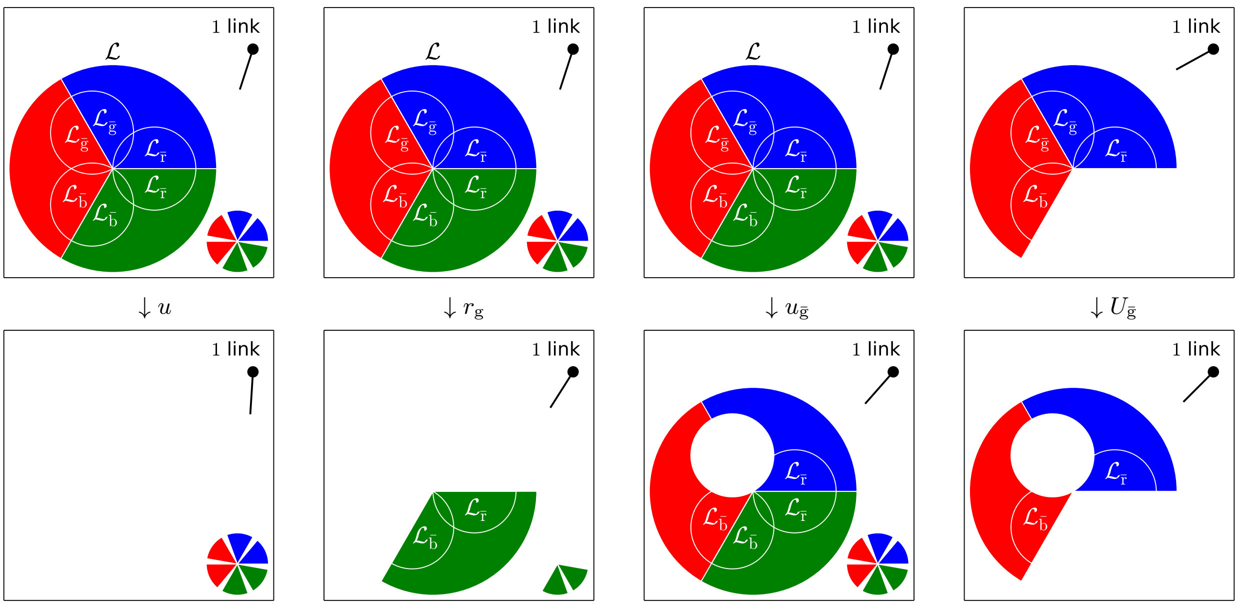

FIG. 3: Probabilities for a single link to connect to different parts of the network. We use these probabilities as primitives to calculate the probability for many links. While u, r_c and $u_{\bar{c}}$ can be calculated with standard methods invented for the configuration model before, the conditional probability $U_{\bar{c}}$ can be calculated as a combination of the others.

We already gave equation 2 for calculating the probability u. In the case of colors on the nodes, as illustrated in figure 3 on the left, the colors can simply be ignored. We further define r_c as the probability to connect to a node of color c. This is illustrated in the second column of the figure. r_c has to include the fact that high degree nodes are reached more probably, therefore the colors on high degree nodes are found with higher probability:

$$r_c = \sum_k k p_k \tilde{r}_{c,k}/\bar{k}. \tag{4}$$

If $\tilde{r}_{c,k}$ does not depend on the degree k, $r_c = \tilde{r}_{c,k}$. We further introduce $u_{\bar{c}}$, the probability that a single link does not connect to a giant $\mathcal{L}_{\bar{c}}$. See the third column of the figure for an illustration. This can be calculated using percolation theory for degree-dependent (targeted) attack by solving

$$u_{\bar{c}} = 1 - f_1(1) + f_1(u_{\bar{c}}), \quad f_1(z) = \sum_k q_k (1 - \tilde{r}_{c,k+1}) z^k. \tag{5}$$

Again, if the color distribution is independent of the degree, we get the simpler expression $u_{\bar{c}} = r_c + (1 - r_c) g_1(u_{\bar{c}})$ (percolation with random attack).

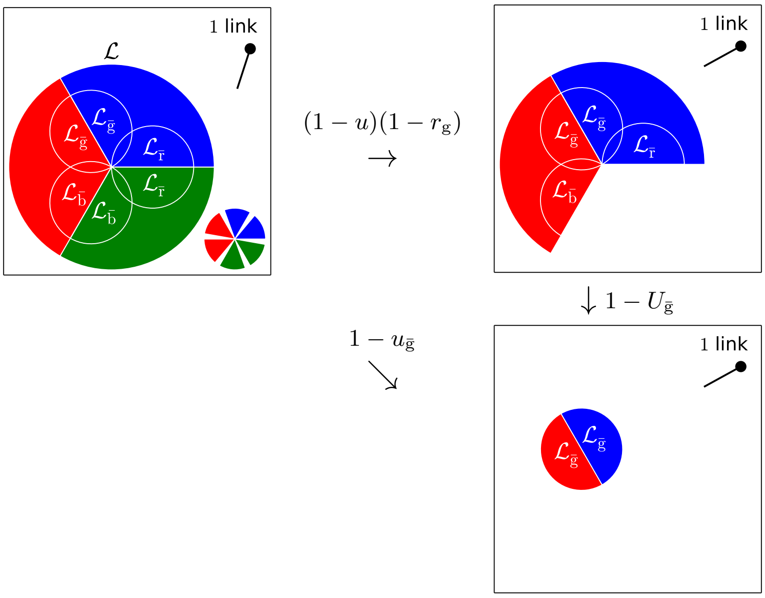

FIG. 4: This figure illustrates the calculation of $U_{\bar{g}}$ using the equality $(1-u)(1-r_{g})(1-U_{\bar{g}}) = 1 - u_{\bar{g}}$. For that, we have assumed independence of the qualities of the link under consideration, especially of the color it connects to and if it connects to the giant component.

Unfortunately, $u_{\bar{c}}$ cannot be used directly for calculating S_{color}. If we look at the same link, the probabilities $u_{\bar{c}}$ are dependent for different colors. The most obvious argument is that always $\Pi_c(1 - u_{\bar{c}}) = 0$, as a link must at least miss one of the $\mathcal{L}_{\bar{c}}$. Instead, we will use the conditional probability $U_{\bar{c}}$, as illustrated with the outer right column of the figure. The precondition is that a link connects to the giant component and the node it connects to has not color c. $U_{\bar{c}}$ is the probability that such a link connects to $\mathcal{L}_{\bar{c}}$. For calculating it, we use the primitives introduced so far, as illustrated in figure 4. Assuming independence of the probabilities $(1-u)$ for connecting to the giant component and $(1-r_c)$ for not connecting to a node of color c, the precondition of $U_{\bar{c}}$ can be constructed. In this way, we can construct $(1 - u_{\bar{c}})$ using the probability we are searching for: $(1 - u_{\bar{c}}) = (1-u)(1-r_c)(1-U_{\bar{c}})$. With this we find

$$U_{\bar{c}} = 1 - \frac{1 - u_{\bar{c}}}{(1-u)(1-r_c)}. \tag{6}$$

If $(1-u)(1-r_c) = 0$, the precondition holds for an empty set of nodes. In this case we define $U_{\bar{c}} = 1$. Notice that the additional information of the explicit color, instead of only stating that the color is not c, does not alter the results, as a further restriction of the colors would meat the numerator and denominator identically and therefore would cancel out.

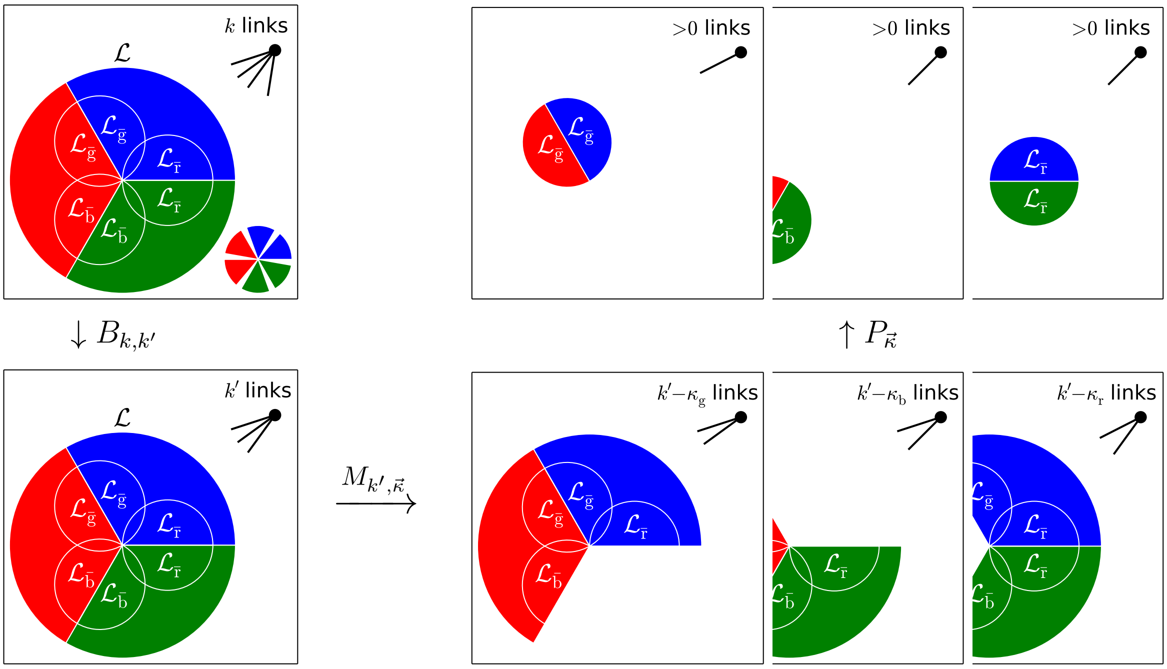

FIG. 5: For calculating the probability of a node with k links to belong to $\mathcal{L}_{\mathrm{color}}$, we have to average over different link constellations which this node might show. First, $B_{k,k'}$ is the probability that out of the k links k' connect to the giant component. It is calculated using u (compare figure 3 on the left). Second, $M_{k',\vec{\kappa}}$ gives the probability for a certain color distribution among the links. It is calculated using r_{g} etc. (compare figure 3, second from left). We assume that this second step is independent of the first step, what is confirmed with the final results. Third, $P_{\vec{\kappa}}$ gives the joint probability that for this color distribution $\mathcal{L}_{\bar{\mathrm{r}}}$, $\mathcal{L}_{\bar{\mathrm{b}}}$ and $\mathcal{L}_{\bar{\mathrm{g}}}$ are connected to at the same time. This is calculated using $U_{\bar{\mathrm{r}}}$ etc. (compare figure 3 on the right).

C. Averaging over link distributions

As in equation 3 for the giant component, we want to get an analytical result for S_{color} by averaging over possible link constellations of a randomly chosen node. Let us give the whole result and then explain it step by step afterwards:

$$S_{\mathrm{color}} = \sum_{k=0}^{\infty} p_k \sum_{k'=0}^{k} B_{k,k'} \sum_{\kappa_1,\ldots,\kappa_C=0}^{k'} M_{k',\vec{\kappa}} P_{\vec{\kappa}}, \tag{7}$$

$$B_{k,k'} = \binom{k}{k'} (1-u)^{k'} u^{k-k'}, \tag{8}$$

$$M_{k',\vec{\kappa}} = \frac{k'!}{\kappa_1! \times \cdots \times \kappa_C!} (r_1)^{\kappa_1} \times \cdots \times (r_C)^{\kappa_C} \, \delta_{k',\kappa_1+\cdots+\kappa_C}, \tag{9}$$

$$P_{\vec{\kappa}} = \prod_{c=1}^{C} [1-(U_{\bar{c}})^{k'-\kappa_c}]. \tag{10}$$

The formulas include the single link probabilities u, r_c and $U_{\bar{C}}$ of equations (2), (4) and (6) (the last depends on (5)). An illustration of the procedure can be seen in figure 5. $B_{k,k'}$ is the binomial probability that out of the k links k' links connect to the giant component. $M_{k',\vec{\kappa}}$ gives the multinomial probability for a certain color distribution among the k' links connecting to the giant component. We assume that this second step is independent of the first step, what is confirmed with the final results. The numbers κ_c count the links which connect to a node of color c in the giant component. Finally, $P_{\vec{\kappa}}$ gives the joint probability that for the color distribution given by $\vec{\kappa}$ all giant $\mathcal{L}_{\bar{c}}$ are connected to at the same time. There is at least one link connecting to $\mathcal{L}_{\bar{c}}$ with probability $1-(U_{\bar{c}})^{k'-\kappa_c}$. The success probabilities for different colors have to be multiplied, as all $\mathcal{L}_{\bar{c}}$ have to be reached at the same time. We tested numerically that e.g. $U_{\bar{1}}$ and $U_{\bar{2}}$ are independent for a link connecting to the giant component and a third color.

II. EXAMINATION OF $S_{\rm color}$

A. Closed form solutions

We now will calculate closed form solutions for $S_{\rm color}$ on graphs with two and three colors. This is done to demonstrate how the extensive summations over k', k and $\vec{\kappa}$ can be performed analytically. In cases where this is not possible, a sampling of values $\vec{\kappa}$ has to be performed. The results can be tested against the analytically tractable situations and by comparing with numerical results. Additionally, our result will help us to gain first insights about the critical behavior.

For evaluating equation 7 with two colors, we first rewrite

$$\sigma_{k'} \equiv \sum_{\kappa_1,\kappa_2=0}^{k'} M_{k',\vec{\kappa}} P_{\vec{\kappa}} \tag{11}$$

$$= \sum_{\kappa_1=0}^{k'} \binom{k'}{\kappa_1} (r_1)^{\kappa_1} (r_2)^{k'-\kappa_1} [1-(U_{\bar{1}})^{k'-\kappa_1}][1-(U_{\bar{2}})^{\kappa_1}] \tag{12}$$

$$= \sum_{\kappa_1=0}^{k'} \binom{k'}{\kappa_1} \left[(r_1)^{\kappa_1}(r_2)^{k'-\kappa_1} - (r_1 U_{\bar{2}})^{\kappa_1}(r_2)^{k'-\kappa_1} - (r_1)^{\kappa_1}(r_2 U_{\bar{1}})^{k'-\kappa_1} + (r_1 U_{\bar{2}})^{\kappa_1}(r_2 U_{\bar{1}})^{k'-\kappa_1}\right] \tag{13}$$

$$= 1 - (r_1 + r_2 U_{\bar{1}})^{k'} - (r_2 + r_1 U_{\bar{2}})^{k'} + (r_1 U_{\bar{2}} + r_2 U_{\bar{1}})^{k'}. \tag{14}$$

In the last step, the binomial formula was used backward. We can use this procedure once more, and with equation 6 and $r_1 + r_2 = 1$ we find

$$S_{\rm color} = \sum_k p_k \sum_{k'=0}^{k} B_{k,k'} \sigma_{k'} \tag{15}$$

$$= \sum_k p_k \sum_{k'=0}^{k} \binom{k}{k'} u^{k-k'} \left[(1-u)^{k'} - ((1-u)(r_1 + r_2 U_{\bar{1}}))^{k'} - \dots\right] \tag{16}$$

$$= \sum_k p_k \left[1 - (u_{\bar{1}})^k - (u_{\bar{2}})^k + (u_{\bar{1}} + u_{\bar{2}} - 1)^k\right] \tag{17}$$

$$= 1 - g_0(u_{\bar{1}}) - g_0(u_{\bar{2}}) + g_0(u_{\bar{1}} + u_{\bar{2}} - 1). \tag{18}$$

This result holds for any degree distribution and color distribution. Notice that $r_c \leq u_{\bar{c}} \leq 1$. The result for two colors does only depend on the probabilities $u_{\bar{c}}$, while conditional probabilities as $U_{\bar{c}}$ were eliminated. This was possible as $\mathcal{L}_{\bar{1}}$ and $\mathcal{L}_{\bar{2}}$ are not overlapping for two colors. For Poisson graphs we find with the according generating function

$$g_0(z) = g_1(z) = e^{\bar{k}(z-1)}. \tag{19}$$

$$S_{\rm color} = [1 - g_0(u_{\bar{1}})][1 - g_0(u_{\bar{2}})]. \tag{20}$$

For more than two colors, $\mathcal{L}_{\bar{c}}$ do overlap. As an example, let us have a look at the result for homogeneous color distributions $\tilde{r}_{c,k} = 1/C$ being

$$S_{\rm color} = \sum_{j=0}^{C} (-1)^j \binom{C}{j} g_0 \left[u + (1-u)\left(\frac{j}{C} U_{\bar{1}}^{j-1} + \frac{C-j}{C} U_{\bar{1}}^{j}\right)\right]. \tag{21}$$

For large numbers of colors we will use an approximation of $\sigma_{k'}$ instead.

B. k'-decomposition and an upper bound

For a further examination of $S_{\rm color}$, it is useful to rearrange equation 7 with focus on k', the number of links connecting a single randomly chosen node to the giant component,

$$S_{\rm color} = \sum_{k'=0}^{\infty} \underbrace{\sum_{k=k'}^{\infty} p_k B_{k,k'}}_{\pi_{k'}} \underbrace{\sum_{\kappa_1,\dots=0}^{1} M_{k',\vec{\kappa}} P_{\vec{\kappa}}}_{\sigma_{k'}} = \sum_{k'=0}^{\infty} \pi_{k'} \sigma_{k'}. \tag{22}$$

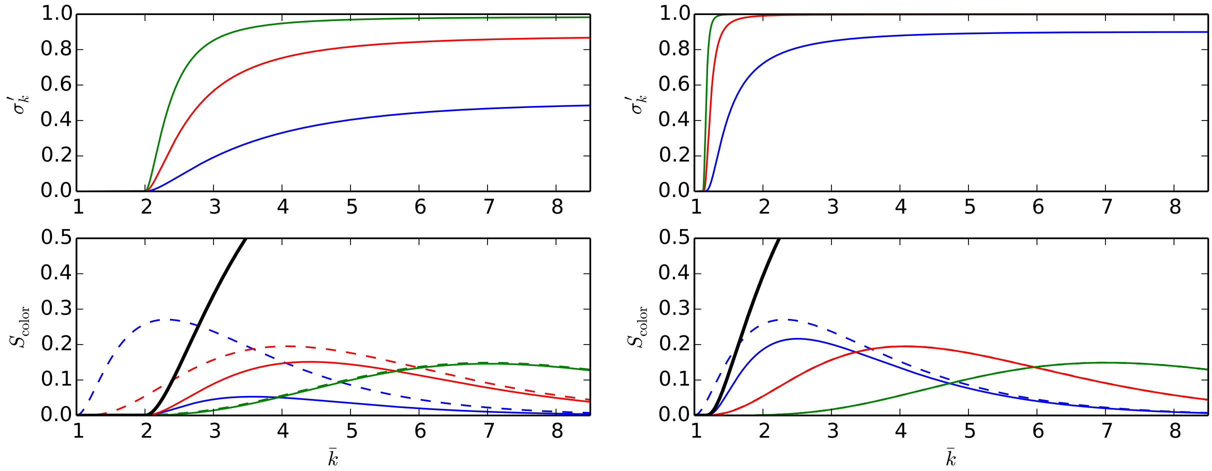

FIG. 6: S_{color} (fat black lines) can be divided into contributions of nodes with different numbers k' of links to the giant component ($k' = 2$ blue, $k' = 4$ red and $k' = 7$ green). Results for Poisson graphs, left $C = 2$, right $C = 10$. On the bottom, $\pi_{k'}$ is shown with dashed lines, and $\pi_{k'}\sigma_{k'}$ is shown with solid lines.

The first term $\pi_{k'}$ gives the probability that a randomly chosen node connects over exactly k' nodes to the giant component. Using equation 8, it can be shown that $\sum_{k'=0}^{\infty} \pi_{k'} = 1$. This quality is meaningful, as every link should have either no or some connections to the giant component. Using the same equation, we can write $\pi_{k'}$ in terms of generating functions,

$$\pi_{k'} = \frac{(1-u)^{k'}}{k'!} \left. \frac{\mathrm{d}^{k'} g_0(z)}{\mathrm{d}z^{k'}} \right|_{z=u} . \tag{23}$$

By construction of the generating function we have $\pi_{k'} = p_k$ for $u = 0$, what should hold of cause. With equation 3 we find $S = 1 - \pi_0 = \sum_{k'=1}^{\infty} \pi_{k'}$. For Poisson graphs, $\pi_{k'}$ can be given in a short and useful form for all k'. Instead of using generating functions, we simply rewrite

$$\pi_{k'} = \sum_{k=k'}^{\infty} \frac{\bar{k}^k e^{-\bar{k}}}{k!} \frac{k!}{(k-k')!k'!} u^{k-k'} (1-u)^{k'} \tag{24}$$

$$= \underbrace{\frac{\bar{k}^{k'} e^{-\bar{k}}}{k'!}}_{p_{k'}} e^{u\bar{k}} (1-u)^{k'} \underbrace{\sum_{k=k'}^{\infty} \frac{(u\bar{k})^{k-k'} e^{-u\bar{k}}}{(k-k')!}}_{1} \tag{25}$$

$$= p_{k'} e^{u\bar{k}} (1-u)^{k'} . \tag{26}$$

The term $\sigma_{k'}$ is the probability that a node connecting over k' links to the giant component belongs to S_{color}. As can be seen with equation 28, $\sigma_0 = \sigma_1 = 0$. In figure 9, $\pi_{k'}$ and $\sigma_{k'}$ are shown for Poisson graphs and homogeneous color distributions for $C = 2$ and $C = 10$ colors. We see that for large numbers of colors and accordingly probabilities $U_{\bar{c}}$ close to zero, $\sigma_{k'} \le 1$ can come close to its upper limit. Therefore the upper limit

$$S_{\mathrm{color}} \le \sum_{k'=2}^{\infty} \pi_{k'} = 1 - \pi_0 - \pi_1 \equiv S_{\mathrm{color},\infty} \tag{27}$$

is useful. Compared to the size of the giant component $S = 1 - \pi_0$, S_{color} misses at least all nodes which are connected to the giant component over $k' = 1$ links. This result is closely connected to k-core percolation with $k = 2$. Note that k-core percolation shows a continuous phase transition for $k = 2$, and only for $k > 2$ has the well known discontinuous behavior.

C. Critical behavior

With the decomposed equation 22, vanishing $\sigma_{k'}$ cause $S_{\text{color}} = 0$. According to

$$\sigma_{k'} = \sum_{\kappa_1,\ldots,\kappa_C=0}^{k'} M_{k',\vec{\kappa}} \prod_{c=1}^{C} [1 - (U_{\bar{c}})^{k'-\kappa_c}]. \tag{28}$$

this is the case if $U_{\bar{c}} = 1$ for any color c. With equation 6 we find that $U_{\bar{c}} = 1$ whenever $u_{\bar{c}} = 1$. Examining equation 5 for $u_{\bar{c}}$ for degree-independent color distributions, we can relate to site percolation (random removal of nodes). Lets use $r_c = \tilde{r}_{c,k}$ to describe the color distribution. With the result of Cohen [3] we find as a condition for the existence of a giant $\mathcal{L}_{\text{color}}$

$$r_c < r_{\text{crit}} = 1 - \frac{\bar{k}}{\langle k^2 \rangle - \bar{k}}, \tag{29}$$

for every color c. For Poisson graphs we have $r_{\text{crit}} = (\bar{k}-1)/\bar{k}$. The color distributions with the lowest value of $\max_c r_c$ is $r_c = 1/C$. Accordingly for connectivity close to $\bar{k} = 1$ many colors are needed. If the color distribution deviates from this optimal case, the largest color frequency defines whether a giant S_{color} exists. An example: If for a connectivity of $\bar{k} = 5$ the largest color frequency is $r_1 = 0.9 > r_{\text{crit}} = 0.8$, we have vanishing S_{color}. For Scale free graphs with exponent in the range $2 \leq \alpha \leq 3$, the largest color frequency can take any value as $r_{\text{crit}} = 1$. However, S_{color} can be very small for large r_c.

If we want to describe the critical behavior of S_{color}, we have to deal with $C-1$ independent variables r_c and one dependent variable due to normalization. We will deal with that only in the case of two colors. To have a clearly defined scenario for $C > 2$, we can shift our focus from critical color frequencies to a critical graph topology. We will analyze the critical behavior for the color distribution $r_c = 1/C$ with the optimal critical value. This is plausible for the question of communication, as the distribution of software versions can be driven to the optimal value on a shorter timescale than the change of topology can be performed. For Poisson graphs with homogeneous color distribution, we have the critical value

$$\bar{k}_{\text{crit}} = C/(C-1). \tag{30}$$

1. *Varying color distribution*

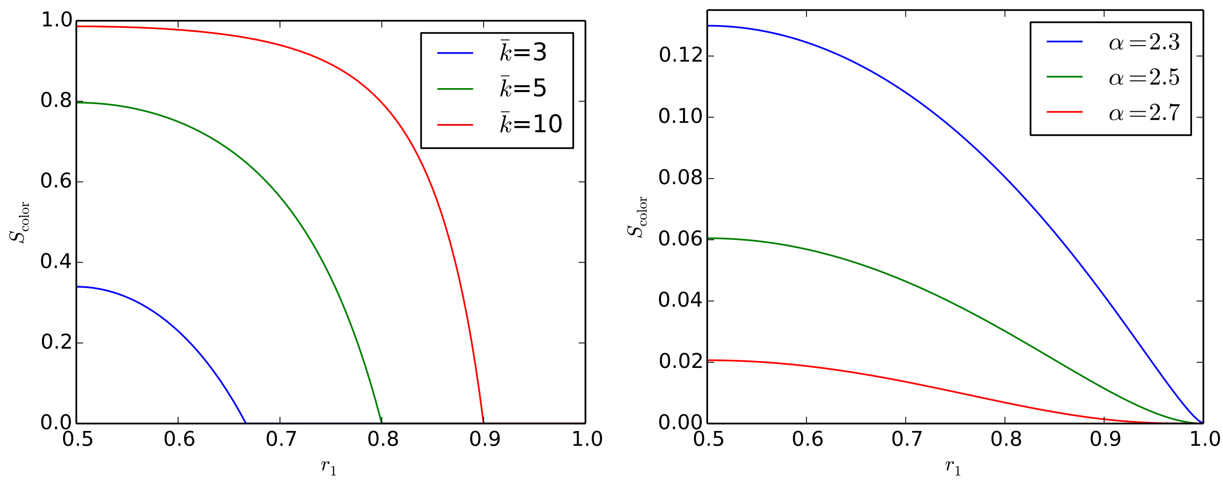

FIG. 7: Results for S_{color} on Poisson graphs (left) and on scale-free graphs (right) for $C = 2$ colors and frequencies $r_1 > r_2$.

For two colors, let us study S_{color} for $r_{\text{crit}} > r_1 > 1/2 > r_2$. For exploring $r_1 \to r_{\text{crit}}$, we define $r_1 = r_{\text{crit}} - \delta$ and $u_{\bar{1}} = 1 - \varepsilon$. Inserting into equation 18, we find using the Taylor expansion

$$S_{\text{color}} = [1 - g_0(1-\varepsilon)] - [g_0(u_{\bar{2}}) - g_0(u_{\bar{2}} - \varepsilon)] \approx \varepsilon \times \left(\left.\frac{\mathrm{d}g_0(z)}{\mathrm{d}z}\right|_1 - \left.\frac{\mathrm{d}g_0(z)}{\mathrm{d}z}\right|_{u_{\bar{2}}(r_2=1-r_{\text{crit}})} \right). \tag{31}$$

Notice that the derivatives are finite for Poisson graphs and for scale-free graphs with exponent above two. Let us compare this with the normal site percolation reading with our notation

$$S_{\bar{c}} = (1 - r_1)[1 - g_0(1 - \varepsilon)] \approx (1 - r_{\mathrm{crit}} + \delta)\varepsilon \times \left.\frac{\mathrm{d}g_0(z)}{\mathrm{d}z}\right|_1 . \tag{32}$$

For Poisson graphs, where $r_{\mathrm{crit}} = (\bar{k} - 1)/\bar{k}$, finite S_{color} is possible for $\bar{k} > 2$. We can find $\varepsilon \propto \delta$ by developing the generating function equation 19 to second order and using equation 5. With that we find

$$S_{\mathrm{color}} \propto \delta^{\beta}, \quad \beta = 1. \tag{33}$$

This corresponds to the behavior known for site percolation.

For scale-free graphs, it is known from the literature [4] that

$$\varepsilon \propto \begin{cases} \delta^1 & \alpha > 4 \\ \delta^{1/(\alpha-3)} & 3 < \alpha < 4 \\ \delta^{1/(3-\alpha)} & 2 < \alpha < 3 \end{cases} . \tag{34}$$

Accordingly are the exponents for $S_{\mathrm{color}} \propto \delta^{\beta}$, with large values close to $\alpha = 3$. This kind of behavior is well known for site percolation. However, notice that for site percolation on scale-free graphs with $2 < \alpha < 3$, the critical exponent differs about one to the result for S_{color}. This is due to $(1 - r_{\mathrm{crit}} + \delta) = \delta$.

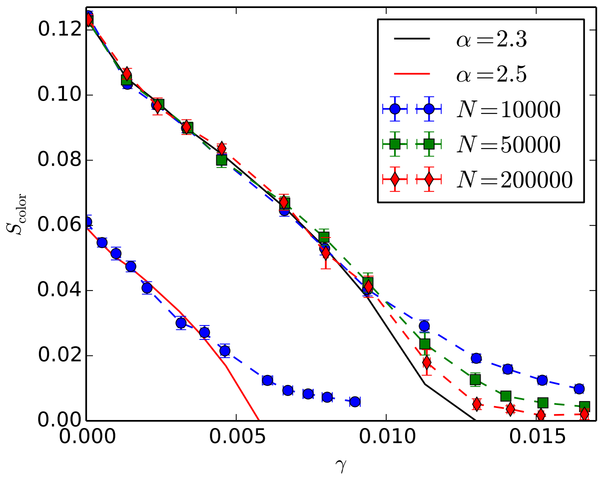

FIG. 8: Results for S_{color} on scale-free graphs for $C = 2$ colors and homogeneity on the largest hubs.

For degree dependent color distributions ...

2. *Varying graph topology*

We have seen that the critical behavior for varying color frequencies far away from the optimal $r_c = 1/C$ implies a behavior of S_{color} similar to results from site percolation. In the following we will see that for $r_c = 1/C$ and varying graph topology, S_{color} shows a surprising critical behavior on Poisson graphs. For Poisson graphs we have

$$\beta = C. \tag{35}$$

This can easily be shown for two colors with equation 18. The case of larger numbers of colors, we have to discuss in more detail below. As S_{color} is converging to $S_{\mathrm{color},\infty}$ as well, we will see a crossover between $\beta = C$ and $\beta = 2$.

3. *Detailed discussion of Poisson graphs*

The normal giant component size S shows a special critical behavior shortly above the transition point, it scales linearly with $\bar{k} - 1$. Here we are interested in the behavior of S_{color} which is a function of $1 - u_{\bar{1}}$ which itself can be

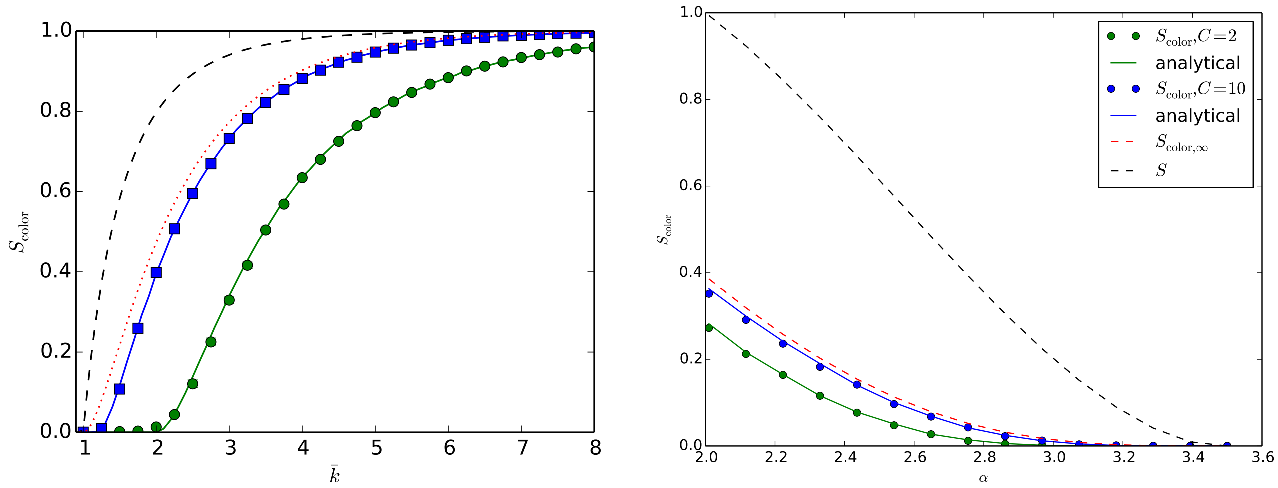

FIG. 9: Results for S_{color} on Poisson graphs (left) and on scale-free graphs (right) for varying graph topology. $C = 2$ colors (green) and $C = 10$ colors (blue).

related to $1 - u = S$. For small arguments $(\bar{k} - \bar{k}_{\mathrm{crit}})$,

$$1 - u_{\bar{1}}(\bar{k} > \bar{k}_{\mathrm{crit}}) \approx (1 - r_1)^2 \left.\frac{\mathrm{d}(1-u)}{\mathrm{d}\bar{k}}\right|_{\bar{k}=1+0} (\bar{k} - \bar{k}_{\mathrm{crit}}) \tag{36}$$

holds due to equation **??**. Inserting into $U_{\bar{1}}$ we find using $1 - u(\bar{k} > 1) \approx \left.\frac{\mathrm{d}(1-u)}{\mathrm{d}\bar{k}}\right|_{\bar{k}=1+0} (\bar{k} - 1)$

$$\varepsilon \equiv 1 - U_{\bar{1}} \approx C(\bar{k} - \bar{k}_{\mathrm{crit}}) \tag{37}$$

if additionally $\bar{k} - \bar{k}_{\mathrm{crit}} \ll \bar{k} - 1$ holds ($1 - u_{\bar{1}}$ small compared to $1 - u$).

For calculating $\sigma_{k'}$, we first need to evaluate $P_{\vec{\kappa}}$ including expressions $1 - (U_{\bar{1}})^{k'-\kappa_c}$. Replacing with ε and applying an approximation we find $1 - (U_{\bar{1}})^{k'-\kappa_c} = 1 - (1-\varepsilon)^{k'-\kappa_c} \approx (k' - \kappa_c)\varepsilon$. This is true at least as long as $k'\varepsilon \ll 1$. With this we find $P_{\vec{\kappa}} \propto (\bar{k} - \bar{k}_{\mathrm{crit}})^C$ independent of $\vec{\kappa}$, and finally

$$\sigma_{k'} \propto (\bar{k} - \bar{k}_{\mathrm{crit}})^C. \tag{38}$$

We will see below that the largest k' giving a notable contribution to S_{color} is smaller then the number of colors C, therefore $k'\varepsilon \ll 1$ can be met and we finally find

$$S_{\mathrm{color}} \propto (\bar{k} - \bar{k}_{\mathrm{crit}})^{\beta}, \tag{39}$$
$$\beta = C. \tag{40}$$

For large numbers of colors C, the functions $\sigma_{k'}$ have a rather extreme scaling behavior $\sigma_{k'} \propto (\bar{k} - \bar{k}_{\mathrm{crit}})^C$. This causes two open questions. First, as the functions $\sigma_{k'}$ are saturating to one for high values of $\bar{k} - \bar{k}_{\mathrm{crit}}$, there could be almost jump-like behavior. Second, contributions with increasing k' become important. Roughly speaking, this is due to the fact that $\sigma_{k'}$ is going to zero very fast with $\bar{k} - \bar{k}_{\mathrm{crit}} \to 0$, and therefore the exponential suppression in $\pi_{k'}$ with k' is compensated for.

For attacking both of these questions, let us apply an approximation. With a homogeneous color distribution $U_{\bar{c}} = U_{\bar{1}}$ and an expected value $\langle \kappa_c \rangle = k'/C \ll k'$, we find

$$P_{\vec{\kappa}} \lessapprox (1 - (U_{\bar{1}})^{k'})^C, \tag{41}$$
$$\sigma_{k'} \lessapprox (1 - (U_{\bar{1}})^{k'})^C. \tag{42}$$

As the approximate $P_{\vec{\kappa}}$ is independent of $\vec{\kappa}$, the summation over $M_{k',\vec{\kappa}}$ was easily executed. Now we are able to compare the contributions $\pi_{k'}\sigma_{k'}$ for $\bar{k} \to \bar{k}_{\mathrm{crit}}$, implying that $k'\varepsilon \ll 1$ holds. We have to find, whether above a certain

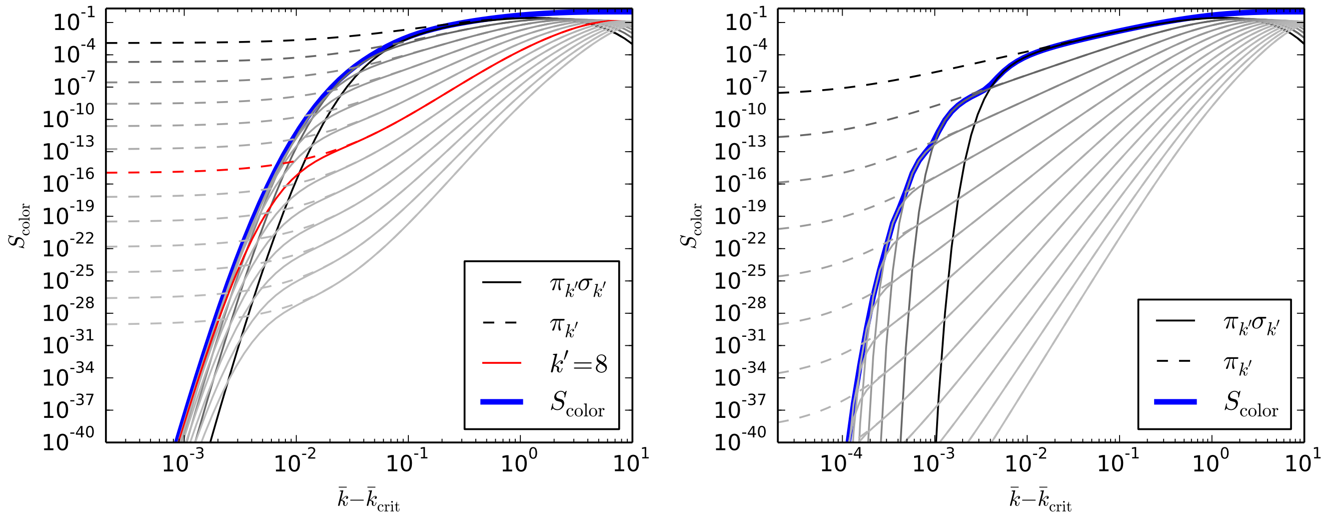

FIG. 10: $S_{\rm color}$ can be divided into contributions of nodes with different numbers k' of links to the giant component. Results on Poisson graphs, left $C = 40$, right $C = 10000$. Results are shown for $k' = 2$ (black) up to $k' = 15$ (gray). For 40 colors $k' = 8$ dominates for small connectivity.

k' the following holds:

$$\log(\frac{\pi_{k'+1}\sigma_{k'+1}}{\pi_{k'}\sigma_{k'}}) < 0. \tag{43}$$

We can find with

$$\log(\sigma_{k'}) \approx C\log[k'(1-u_{\bar{1}})/((1-u)(1-r_1)] \tag{44}$$
$$\approx C\log(k') + \{\text{independent of } k'\}, \tag{45}$$
$$\log(\pi_{k'}) = k'\log(\bar{k}) - \log(k'!) + k'\log(1-u) + \{\text{independent of } k'\}. \tag{46}$$
$$\tag{47}$$

and using $\log(\bar{k}_{\rm crit}) \approx 1/C$ and $\log(1-u(\bar{k}_{\rm crit})) \approx -\log(C)$ and $\log(1=x) \approx x$ for small x:

$$\log(\frac{\pi_{k'+1}\sigma_{k'+1}}{\pi_{k'}\sigma_{k'}}) \approx C/k' - \log(k') + 1/C - \log(C). \tag{48}$$
$$\tag{49}$$

This falls below zero approximately at $k'\log(Ck') = C$, and clearly contributions are exponentially dampened with large k'. Up to the k' with $k'\log(Ck') = C$, contributions $\pi_{k'}\sigma_{k'}$ grow. That is surprising, as nodes with increasing k' should be much less probable. However, if $\sigma_{k'}$ is very small that compensates for the other effect in a certain regime of k'. For estimating the critical region, where $S_{\rm color} \propto (\bar{k}-\bar{k}_{\rm crit})^C$, we use the upper bound for the leading term $k' = C$ and find with $k'\varepsilon = k'C(\bar{k}-\bar{k}_{\rm crit}) \ll 1$ finally

$$(\bar{k}-\bar{k}_{\rm crit}) \ll 1/C^2. \tag{50}$$

This region is not of practical importance for large C, as there are very small values $S_{\rm color}$.

[1] P. Erdos and Renyi, Publicationes mathematicae **6**, 290 (1959).
[2] M. E. Newman, S. H. Strogatz, and D. J. Watts, Physical Review E **64**, 026118 (2001).
[3] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, Physical review letters **85**, 4626 (2000).
[4] R. Cohen, D. Ben-Avraham, and S. Havlin, Physical Review E **66**, 036113 (2002).