Supplementary Materials - Efficient network immunization under limited knowledge

Yangyang Liu,^{1,*} Hillel Sanhedrai,^{2,*} GaoGao Dong,^{3,†} Louis M.

Shekhtman,^{2,4,†} Fan Wang,² Sergey V. Buldyrev,⁵ and Shlomo Havlin²

¹Department of Systems Science, College of Liberal Arts and Sciences,

National University of Defense Technology, Changsha, Hunan 410073, China

²Department of Physics, Bar-Ilan University, Ramat Gan 5290002, Israel

³Faculty of Science, Jiangsu University, Zhenjiang, Jiangsu 212013, China

⁴Networks Science Institute, Northeastern University, Boston, MA 02115

⁵Department of Physics, Yeshiva University, New York, New York 10033, USA

(Dated: April 2, 2020)

I. SOLUTION OF THE ODE FOR F(k,t)

Consider the ODE

$$-F + (N-t)\dot{F} = -F^n, \qquad (S1)$$

where F = F(k, t), with the initial condition

$$F(k, t = 0) = F(k).$$
 (S2)

Using separation of variables we get

$$\frac{\mathrm{d}F}{F-F^n} = \frac{\mathrm{d}t}{N-t}$$

Thus,

$$\frac{1}{n-1} \log \left| \frac{F^n}{F^n - F} \right| = -\log(N-t) + C_1$$
$$\left| \frac{F^n}{F^n - F} \right| = e^{(n-1)(-\log(N-t) + C_1)}$$
$$F^{1-n} = 1 + C_2 e^{(n-1)\log(N-t)}$$
$$F = \left(1 + C_2 e^{(n-1)\log(N-t)} \right)^{-\frac{1}{n-1}}$$

Recalling that F(k, t = 0) = F(k), and substituting t = 0 gives

$$F(k)^{1-n} = 1 + C_2 e^{(n-1)\log(N)}$$

$$C_2 = \left(F(k)^{1-n} - 1\right) e^{-(n-1)\log(N)}.$$

As a result,

$$F(k,t) = \left(1 + \left(F(k)^{1-n} - 1\right)e^{(n-1)\log[(N-t)/N]}\right)^{-\frac{1}{n-1}},$$
(S3)

or

$$F_p(k) = \left(1 + \left(F(k)^{1-n} - 1\right)p^{n-1}\right)^{-\frac{1}{n-1}},\tag{S4}$$

where $F_p(k)$ is the cumulative distribution of degree of the occupied nodes after removing 1-p fraction of nodes. For n = 1, the solution of the ODE is F = Const., namely $F_p(k) = F(k)$ as expected. Also Eq. (S4) converges to it if we take the limit $n \to 1$.

II. DERIVATION OF THE EXPONENTIAL CONVERGENCE OF p_c FOR LARGE n

First, we find $F_p(k)$ and $P_p(k)$ for large n. Examining Eq. (S4) where $n \to \infty$, as elaborated below, we recognize two behaviors depending on F(k) < p or F(k) > p. For the leading term

$$F_p^{\infty}(k) = \begin{cases} \frac{F(k)}{p}, & F(k) < p\\ 1, & F(k) > p \end{cases}$$
(S5)

See Fig. S1 for illustration.

Let us denote the next term as



FIG. S1. Illustration of theory for $F_p(k)$ for large *n*. For small $k, F(k) < p, F_p(k) \rightarrow F(k)/p$, and for large $k, F(k) > p, F_p(k) \rightarrow 1$. The dashed horizontal line is at the value of $p. k_{<}$ and $k_{>}$ are the consecutive degrees for which $F(k_{<}) < p < F(k_{>})$ respectively. For determining the decay rate α in Eq. (S17), one should take $p = p_c$, and find the minimum of $|\log(F(k)/p_c^{\infty})|$ among all k, which is in fact among $k_{<}$ and $k_{>}$.

$$F_p(k) = F_p^{\infty}(k) + \epsilon(k, p, n)$$
 (S6)

Then, we approach finding the next term, ϵ , in each range of degree k.

^{*} Y. Liu and H. Sanhedrai contributed equally to this work

[†] To whom correspondence may be addressed. Email: dfocus.gao@gmail.com or lsheks@gmail.com

For small k, such that F(k) < p,

$$F_p(k) \approx \left(1 + \left(\frac{p}{F}\right)^n - p^n\right)^{-\frac{1}{n}} \approx \left(1 + \left(\frac{p}{F}\right)^n\right)^{-\frac{1}{n}}$$
$$= \frac{F}{p} \left(1 + \left(\frac{F}{p}\right)^n\right)^{-\frac{1}{n}}$$
$$= \frac{F}{p} \exp\left[-\frac{1}{n} \log\left(1 + \left(\frac{F}{p}\right)^n\right)\right]$$
$$\approx \frac{F}{p} \exp\left[-\frac{1}{n} \left(\frac{F}{p}\right)^n\right] \approx \frac{F(k)}{p} \left[1 - \frac{1}{n} \left(\frac{F(k)}{p}\right)^n\right]$$
$$\approx \frac{F(k)}{p} - \frac{1}{n} e^{-(n+1)[\log p - \log F(k)]}$$
(S7)

For the particular case that there exists a degree satisfying F(k) = p,

$$F_p(k) \approx \left(1 + \left(\frac{p}{F}\right)^n - p^n\right)^{-\frac{1}{n}}$$

$$\approx 2^{-\frac{1}{n}} = e^{-\frac{1}{n}\log 2} \approx 1 - \frac{\log 2}{n}$$
(S8)

For the limit $F(k) \to p$ such that $(p/F(k))^n \to 1$, let us define $\zeta = 1 - p/F$ which is small. Then,

$$(p/F)^n = (1-\zeta)^n = \exp(n\ln(1-\zeta))$$

$$\approx 1 + n\ln(1-\zeta) \approx 1 - \zeta n$$

$$= 1 - n(1-p/F).$$

Thus,

$$F_{p}(k) \approx \left(1 + \left(\frac{p}{F}\right)^{n} - p^{n}\right)^{-\frac{1}{n}} \\\approx \left(2 - n(1 - p/F) - p^{n}\right)^{-\frac{1}{n}} \\\approx \exp\left[-\frac{1}{n}\ln\left(2 - n(1 - p/F) - p^{n}\right)\right] \\\approx \exp\left[-\frac{1}{n}\ln 2 - \frac{1}{n}\ln\left(1 - n(1 - p/F)/2 - p^{n}/2\right)\right] \\\approx \exp\left[-\frac{\ln 2}{n} + \frac{1 - p/F}{2} + \frac{p^{n}}{2n}\right] \\\approx 1 - \frac{\ln 2}{n} + \frac{1 - p/F}{2} + \frac{p^{n}}{2n}$$
(S9)

For large k, namely p < F(k) < 1,

$$F_p(k) \approx \left(1 + \left(\frac{p}{F}\right)^n - p^n\right)^{-\frac{1}{n}}$$

= $\exp\left[-\frac{1}{n}\log\left(1 + \left(\frac{p}{F}\right)^n - p^n\right)\right]$
 $\approx \exp\left[-\frac{1}{n}\left(\frac{p}{F}\right)^n + \frac{p^n}{n}\right]$ (S10)
 $\approx 1 - \frac{1}{n}\left(\frac{p}{F(k)}\right)^n + \frac{p^n}{n}$
 $\approx 1 - \frac{1}{n}\exp\left(-n\left[\log F(k) - \log p\right]\right).$

In the limit $F(k) \to 1$, we should keep the last term above, p^n/n , in order to get $F_p(k) \to 1$. In more detail, if $F(k)^n \to 1$ we can denote a small term $\xi(k) = 1 - F(k)$, which satisfies

$$F(k)^{-n} = \exp(-n\ln F(k)) \approx 1 - n\ln F(k) = 1 - n\ln(1 - \xi(k)) \approx 1 + n\xi(k).$$

Substituting this into Eq. (S10) gives

$$F_p(k) \approx 1 - \frac{1}{n} \left(\frac{p}{F(k)}\right)^n + \frac{p^n}{n}$$

$$\approx 1 - \xi(k)p^n = 1 - p^n(1 - F(k)).$$
(S11)

Finally, for F(k) = 1, it easy to see that also $F_p(k) = 1$.

To summarize, the next term of $F_p(k)$ is

$$\epsilon(k, p, n) = \begin{cases} -\frac{1}{n}e^{-\alpha_k n}, & F(k) < 1\\ 0, & F(k) = 1 \end{cases},$$
 (S12)

where $\alpha_k = |\log [p/F(k)]|$.

If there are degrees which obey the limits $F(k)^n \to 1$, or $[F(k)/p]^n \to 1$, then they satisfy

$$\epsilon(k,p,n) = \begin{cases} -\frac{1}{n} \left(\ln 2 - n \left[1 - \frac{p}{F(k)} \right] \right), & \left(\frac{p}{F(k)} \right)^n \to 1\\ -p^n (1 - F(k)), & F(k)^n \to 1 \end{cases}$$
(S13)

Having an approximation for $F_p(k)$, we approach now to find p_c using the equation for criticality,

$$1 = \frac{p}{\langle k \rangle} \sum_{k=0}^{\infty} k(k-1) P_p(k).$$
 (S14)

Before, we denote several special degrees: the degrees

where F(k) is close to p, and the maximal degree, as following

$$k_{<} = \max_{F(k) < p} k,$$

$$k_{>} = \min_{F(k) \ge p} k,$$

$$K = \min_{F(k) = 1} k.$$

Plugging Eqs. (S5),(S12),(S6) into Eq. (S14), and using change of summation indexes, one can obtain

$$\begin{split} \frac{\langle k \rangle}{p_c} &= \sum_{k=2}^{\infty} k(k-1) \Delta F_p(k) \\ &= \sum_{k=2}^{\infty} k(k-1) \Delta F_p^{\infty}(k) + \sum_{k=2}^{K} k(k-1) \Delta \epsilon(k, p_c, n) \\ &= \sum_{k=2}^{k_{<}} k(k-1) \frac{P(k)}{p_c} + \sum_{k=k_{>}+1}^{\infty} k(k-1) \cdot 0 \\ &+ k_{>}k_{<} \left[1 - \frac{F(k_{<})}{p_c} \right] - 2 \sum_{k=1}^{K} k \epsilon(k, p_c, n). \end{split}$$

Hence,

$$\frac{\langle k \rangle}{p_c} = \sum_{k=2}^{k_{<}} k(k-1) \frac{P(k)}{p_c} + k_{>}k_{<} \left[1 - \frac{F(k_{<})}{p_c} \right] - 2 \sum_{k=1}^{K} k\epsilon(k, p_c, n)$$
(S15)

and for $n \to \infty$, for the leading term, p_c^{∞} ,

$$\frac{\langle k \rangle}{p_c^{\infty}} = \sum_{k=2}^{k_<} k(k-1) \frac{P(k)}{p_c^{\infty}} + k_> k_< \left[1 - \frac{F(k_<)}{p_c^{\infty}} \right]$$
$$p_c^{\infty} = F(k_<) + \frac{1}{k_> k_<} \left(\langle k \rangle - \sum_{k=2}^{k_<} k(k-1)P(k) \right),$$

where $k_{<}$ and $k_{>}$ are determined by p_c^{∞} , but are independent of n.

For the next term, returning to the equation (S15) and obtain

$$p_{c} = p_{c}^{\infty} + \frac{2p_{c}^{\infty}}{k_{>}k_{<}} \sum_{k=1}^{K} k\epsilon(k, p_{c}^{\infty}, n)$$
(S16)

The terms $\epsilon(k, p_c^{\infty}, n)$ in the summation decay exponentially with n by rate $\alpha_k = |\ln[p_c^{\infty}/F(k)]|$, and because $n \to \infty$, the dominant term is the one with the lowest rate $\alpha = \min_k \{\alpha_k\}$, and the corresponding k will be denoted by k_{slow} , meaning $\alpha = \alpha_{k_{\text{slow}}}$. Therefore,

$$p_c \sim p_c^{\infty} - A \frac{1}{n} e^{-\alpha n} \tag{S17}$$

where $\alpha = \min_k |\log[p_c^{\infty}/F(k)]|$ and $A = (2p_c^{\infty}k_{slow})/(k_{>}k_{<})$. It is obvious that k_{slow} is $k_{<}$ or $k_{>}$ because F(k) is monotonic.

We know that if F(k) = 1 then $\epsilon = 0$, so if $F(k_{\text{slow}}) = F(k_{>}) = 1$ then $k_{<}$ should be taken as k_{slow} . It should also be noted that if $F(k_{\text{slow}}) = p_c^{\infty}$ or very close, then $\epsilon(k_{\text{slow}}, p_c^{\infty}, n) = -n^{-1} \log 2$. Another special case is where k_{slow} is not unique, then we take all of them, and k_{slow} is replaced by $\sum_{k_{\text{slow}}} k_{\text{slow}}$.



FIG. S2. **Examples of** F(k). The decay rate α determines the regimes of power law or exponential convergence of $p_c(n \to \infty)$. α depends on the minimal ratio between p_c^{∞} and F(k). Hence α is influenced by the density of F(k) in the neighborhood of p_c^{∞} . Here we show the patterns of F(k) for ER network with $N = 10^4$ and k = 2 (a), and for SF network with $N = 10^4$ and $\gamma = 2.5$ (b). One can see that for larger p_c^{∞} , F(k) is much more dense for SF network, which makes α become small and extend the power-law regime.

The behavior of Eq. (S17) depends on the value of n, and splits into two regimes. Where $n \ll 1/\alpha$, convergence will be according to a power-law, and where $n \gg 1/\alpha$ the convergence is exponential. It means that the width of these two regimes depends on α . The value of α depends on the value of p_c^{∞} and the details of the degree distribution. Therefore, taking conditions providing $p_c^{\infty} \to 1$ might lead to a broader power law regime. That is because the tail of the degree distribution, P(k), has very small values, which can be $\sim 1/N$. That implies that there are values of $F(k) \sim 1 - 1/N$, which yields $\alpha = \min_k |\log (p_c^{\infty}/F(k))| \approx \log(1) - \log(1-1/N) \approx 1/N$, and makes the power law regime being any $n \ll N$ in this case. In contrast, if p_c^{∞} is not very close to 1, then α is expected to be finite because F(k) in not very dense in this regime. Therefore, the exponential behavior will rule even for small n. See in Fig. S2 examples of ER and SF networks.

III. DERIVATION OF p_c VS n FOR SCALE-FREE NETWORKS WITH SMALL n

Here we consider SF network with $2 < \gamma < 3$. If n = 1, then $p_c \to 0$. In order to find out for which n, p_c becomes non-zero, we analyze Eqs. (S4) and (S14) for large k (high degrees govern the behavior in SF network), small n and p as follows.

It can be shown that

$$F(k) \approx 1 - (k/m)^{1-\gamma}$$
, (S18)

Then,

$$F(k)^n \approx \left(1 - (k/m)^{1-\gamma}\right)^n \approx \exp\left[n\log\left(1 - (k/m)^{1-\gamma}\right)\right]$$
$$\approx \exp\left[-n\left(k/m\right)^{1-\gamma}\right] \approx 1 - n\left(k/m\right)^{1-\gamma},$$
$$F(k)^{1-n} \approx 1 + (n-1)\left(k/m\right)^{1-\gamma},$$

where we assume $(k/m)^{\gamma-1} \gg n$ for large degrees. Plugging this into Eq. (S4),

$$F_{p}(k) = \left(1 + \left(F(k)^{1-n} - 1\right)p^{n-1}\right)^{-\frac{1}{n-1}} \\\approx \left(1 + (n-1)\left(k/m\right)^{1-\gamma}p^{n-1}\right)^{-\frac{1}{n-1}} \\\approx \exp\left[-\frac{1}{n-1}\log\left(1 + (n-1)\left(k/m\right)^{1-\gamma}p^{n-1}\right)\right] \\\approx \exp\left[-\frac{1}{n-1}(n-1)\left(k/m\right)^{1-\gamma}p^{n-1}\right] \\\approx 1 - p^{n-1}\left(k/m\right)^{1-\gamma}.$$

That leads to

$$P_p(k) \approx \frac{\partial F_p(k)}{\partial k} \approx \frac{\gamma - 1}{m} p^{n-1} \left(\frac{k}{m}\right)^{-\gamma} = p^{n-1} P(k).$$
(S19)

Now we should notice that $P_p(k)$ has a new natural cutoff, K_p , which depends on p as following

$$\int_{K_p}^{\infty} p^{n-1} A k^{-\gamma} = \frac{1}{Np},$$
(S20)
$$K_p \sim p^{n/(\gamma-1)} N^{1/(\gamma-1)}.$$

Taking the result of Eq. (S19) into Eq. (S14) and keeping the leading terms, where $\langle k_{p_c} \rangle$ (average degree of the remaining nodes) is finite and $\langle k_{p_c}^2 \rangle$ increases with N, we obtain

$$1 = \frac{p_c}{\langle k \rangle} \sum_{k=0}^{\infty} k(k-1) P_{p_c}(k) \sim \frac{p_c}{\langle k \rangle} \langle k_{p_c}^2 \rangle$$

But

$$\langle k_{p_c}^2 \rangle \sim \int_{m}^{K_{p_c}} k^2 p_c^{n-1} A k^{-\gamma} \sim p_c^{n-1} K_{p_c}^{3-\gamma} \sim p_c^{n-1+n\beta} N^{\beta},$$
(S21)

where $\beta = (3 - \gamma)/(\gamma - 1)$. Substituting $\langle k_{p_c}^2 \rangle$ into the Eq. above provides

$$1 \sim \frac{p_c}{\langle k \rangle} \langle k_{p_c}^2 \rangle \sim \frac{p_c^{n+n\beta}}{\langle k \rangle} N^{\beta}.$$

Therefore

$$p_c \sim N^{-\delta/n} \sim \exp\left[-\delta \log(N)/n\right],$$
 (S22)

where

$$\delta = \frac{\beta}{1+\beta} = \frac{3-\gamma}{2}.$$
 (S23)

This scaling is with N. The pre-factor depends on n but not in N. The scaling is valid for large N.

From Eq. (S22), it is easy to see that if $n \ll \log N$, $p_c \rightarrow 0$, while if $n \sim \log N$, p_c becomes non-zero.

After we completed the calculation, for consistency, we should return to check our initial assumption saying $(k/m)^{\gamma-1} \gg n$ for the large degrees. We should apply this condition on K_{p_c} that we found. $n \ll (K_{p_c}/m)^{\gamma-1} \sim p_c^n N/m^{\gamma-1}$, and substituting p_c that we found, we get $n \ll N^{1-\delta} = N^{(\gamma-1)/2}$. To conclude, we can say that Eq. (S22) is valid for

$$n \ll N^{(\gamma-1)/2}.\tag{S24}$$

IV. EXAMPLE OF A SIMPLE NETWORK WITH ONLY TWO DEGREES

Here we consider an example of a simple degree distribution which can be analyzed simply as following

$$P(k) = \begin{cases} r, & k = 1\\ 1 - r, & k = 3\\ 0, & \text{else} \end{cases}$$
(S25)

Then F(1) = r, F(2) = r, F(3) = 1.

Assuming that $r < p_c < 1$, one can observe that this assumption is consistent with the final result. Hence, from Eq. (S7)

$$F_{p}(k) = \begin{cases} 0, & k = 0\\ \frac{r}{p_{c}} - \frac{1}{n} \left(\frac{r}{p_{c}}\right)^{n+1}, & k = 1\\ \frac{r}{p_{c}} - \frac{1}{n} \left(\frac{r}{p_{c}}\right)^{n+1}, & k = 2\\ 1, & k = 3 \end{cases}$$
(S26)

Thus

$$\frac{\langle k \rangle}{p_c} = \sum_{k=2}^{3} k(k-1)\Delta F_p(k) = 2 \cdot 1 \cdot 0$$
$$+ 3 \cdot 2 \cdot \left[1 - \frac{r}{p_c} + \frac{1}{n} \left(\frac{r}{p_c}\right)^{n+1}\right],$$
$$\frac{3 - 2r}{p_c} = 6 \left[1 - \frac{r}{p_c} + \frac{1}{n} \left(\frac{r}{p_c}\right)^{n+1}\right],$$
$$p_c = \frac{3 + 4r}{6} - \frac{r}{n} \left(\frac{r}{p_c}\right)^n.$$

For the limit of $n \to \infty$

$$p_c^{\infty} = \frac{1}{2} + \frac{2r}{3}.$$
 (S27)

Note that r < 3/4, otherwise, there is no giant component even without attack.

Then, we conclude for large n that $p_c = p_c^\infty + o(1).$ As a result, for the leading term

$$p_c = p_c^{\infty} - \frac{r}{n} \left(\frac{r}{p_c^{\infty}}\right)^n = p_c^{\infty} - \frac{r}{n} \exp(-\alpha n), \quad (S28)$$

where $\alpha = \log(2/3 + 1/(2r))$.

The other limit of n = 1 is the random attack, and p_c is obtained by the simple known formula for configuration model $p_c = 1/(\kappa - 1)$. Is is easy to find that $\langle k \rangle = 3 - 2r$ and $\langle k^2 \rangle = 9 - 8r$, resulting

$$p_c(n=1) = \frac{1}{2} + \frac{1}{6} \frac{r}{1-r}.$$
 (S29)

Now we analyze a similar but a little bit different case where

$$P(k) = \begin{cases} r, & k = 2\\ 1 - r, & k = 3\\ 0, & \text{else} \end{cases}$$
(S30)

Here F(1) = 0, F(2) = r, F(3) = 1. Thus,

$$F_p(k) = \begin{cases} 0, & k = 1\\ \frac{r}{p_c} - \frac{1}{n} \left(\frac{r}{p_c}\right)^{n+1}, & k = 2\\ 1, & k = 3 \end{cases}$$
(S31)

Therefore,

$$\frac{\langle k \rangle}{p_c} = \sum_{k=2}^{3} k(k-1) \Delta F_p(k) = 2 \cdot 1 \cdot \left[\frac{r}{p_c} - \frac{1}{n} \left(\frac{r}{p_c} \right)^{n+1} \right] + 3 \cdot 2 \cdot \left[1 - \frac{r}{p_c} + \frac{1}{n} \left(\frac{r}{p_c} \right)^{n+1} \right], \frac{3-r}{p_c} = 6 - 4 \left[\frac{r}{p_c} - \frac{1}{n} \left(\frac{r}{p_c} \right)^{n+1} \right], p_c = \frac{1+r}{2} - \frac{2r}{3n} \left(\frac{r}{p_c} \right)^n.$$

For the limit of $n \to \infty$

$$p_c^{\infty} = \frac{1}{2} + \frac{r}{2}.$$
 (S32)

Therefore, for the next leading term

$$p_c = \frac{1+r}{2} - \frac{2r}{3n} \left(\frac{2r}{1+r}\right)^n.$$
 (S33)

Now, if we will take $r \sim 1 - b/N$, where b is finite, then

$$p_c \approx 1 - \frac{2}{3n} \left(1 - \frac{b}{N} \right)^n \approx 1 - \frac{2}{3n} e^{-b\frac{n}{N}}.$$
 (S34)

Hence, if $n \ll N$ then

$$1 - p_c \sim \frac{1}{n}.\tag{S35}$$

This happens because $p_c \to 1$ where $r \to 1$, but F(2) = r, hence $F(2) \to p_c$, which yields a power law behavior of $p_c(n)$.