

## Research



**Cite this article:** Bhansali R, Schaposnik LP.

2020 A trust model for spreading gossip in social networks: a multi-type bootstrap percolation model. *Proc. R. Soc. A* **476**: 20190826.

<http://dx.doi.org/10.1098/rspa.2019.0826>

Received: 27 November 2019

Accepted: 13 February 2020

**Subject Areas:**

graph theory, combinatorics, mathematical modelling

**Keywords:**

trust model, bootstrap percolation, information propagation, disease propagation

**Author for correspondence:**

Laura P. Schaposnik

e-mail: [schapos@uic.edu](mailto:schapos@uic.edu)

# A trust model for spreading gossip in social networks: a multi-type bootstrap percolation model

Rinni Bhansali<sup>1</sup> and Laura P. Schaposnik<sup>2,3</sup>

<sup>1</sup>Half Hollow Hills High School East, 50 Vanderbilt Pkwy, Dix Hills, NY 11746, USA

<sup>2</sup>University of Illinois at Chicago, Chicago, IL 60607, USA

<sup>3</sup>Simons Center for Geometry and Physics, NY 11794, USA

LPS, 0000-0003-1417-2201

We introduce here a multi-type bootstrap percolation model, which we call  $\mathcal{T}$ -Bootstrap Percolation ( $\mathcal{T}$ -BP), and apply it to study information propagation in social networks. In this model, a social network is represented by a graph  $G$  whose vertices have different labels corresponding to the type of role the person plays in the network (e.g. a student, an educator etc.). Once an initial set of vertices of  $G$  is randomly selected to be carrying a gossip (e.g. to be infected), the gossip propagates to a new vertex provided it is transmitted by a minimum threshold of vertices with different labels. By considering random graphs, which have been shown to closely represent social networks, we study different properties of the  $\mathcal{T}$ -BP model through numerical simulations, and describe its implications when applied to rumour spread, fake news and marketing strategies.

## 1. Introduction

Most people have struggled at some point to find the perfect present for their beloved: we hear from our son's friends that certain 'bacteria growing kit' would be fun for his 10th birthday—but will it actually be safe? Once we hear from our son's friends' parents that the 'bacteria growing kit' is indeed entertaining and safe for that age, we are close to decided on buying it. Is this recurrent phenomenon a consequence of a natural instinct that one has, where having the same information transmitted by different 'types' of people inspires more trust? If so, we naturally wonder:



**Figure 1.** Examples of a social network where different colours—labels—describe the types of people within gossip spread. (Online version in colour.)

*How many different types of people (colleagues, friends, taxi drivers etc. as depicted in figure 1) should we hear a piece of information from, before we start transmitting it as a true fact?*

Or equivalently, and concerning marketing strategies,

*How many different types of people should recommend to us a service or a product before we buy it, and begin recommending it ourselves?*

Having a clear range for sources of information would allow members of the society to disbelieve gossips and differentiate fake news. Moreover, understanding this range would also allow the industry to target wisely a minimum amount of consumers within each type of people, and using the natural propagating process to continue the marketing on its own.

In this paper, we build a new model of information/disease spread which we then use to understand the above questions. Our model builds upon the classical Bootstrap Percolation introduced in [1], but incorporates the concept of different types of members of society. Bootstrap percolation is a particular class of monotone cellular automata describing an activation process which follows certain activation rules, and which has been much used to model interactions within societies. In particular, in the classical  $r$ -neighbour bootstrap process on a graph  $G$ , a set  $A$  of initially ‘infected’ vertices spreads by infecting vertices with at least  $r$  already-infected neighbours (e.g. see [2–7], and also [8–10] for applications to social networks).

In the present manuscript, we shall introduce a multi-type version of bootstrap percolation which we call a *Trusted Bootstrap Percolation* or  *$T$ -Bootstrap Percolation* ( $T$ -BP) to answer an equivalent question to those posed above:

*How does information percolate when a messenger only passes the information if it has been received by a number of different sources of certain types?*

Social media networks and search engines keep track of news a user and his/her friends respond positively to, and then use this information to suggest future articles and advertisements. However, when it comes to news content and discussion of the news, this means one will increasingly only see material that is in line with one’s stated interests. This worsens issues of polarization and group-think. To combat this, we could apply the  $T$ -BP to the algorithms which suggest the news users are shown by adding trust vectors needed to be satisfied before a piece of news is shown—this would prevent the existence of echo chambers by forcing people to see information which might cater to multiple sides of the political spectrum. In what follows, we use  $T$ -BP to numerically understand the spread of gossip in random graphs  $G$  simulating social networks. The simplest form of  $T$ -BP is:

**Definition 1.1.** Consider a finite or infinite graph  $G$ , two natural numbers  $r, m \in \mathbb{N}$ , a vector  $k = (k_1, \dots, k_m) \in \mathbb{N}^m$  of non-negative numbers for which exactly  $r$  of them satisfy  $k_i \neq 0$ , and a set  $A := A_0$  of initially ‘infected’ vertices in  $G$ . After assigning randomly a label in  $\{1, \dots, m\}$  to each vertex, we define  $r$ -bootstrap percolation with trust level  $k$  on  $G$  as the process in which at each

time step all of the vertices which have at least  $k_i$  adjacent vertices infected with label  $i$ , for all  $i$ , become infected.

**Remark 1.2.** One should note that in the above definition, the value of  $r$  indicates the number of different types of people that need to be infected for the disease or gossip to propagate. Moreover, the least amount of vertices of type  $i$  that need to be infected is given by  $k_i$ .

In what follows, we shall introduce and study the most generic form of  $\mathcal{T}$ -BP in §§2–4, and conclude the paper describing the implications of our results within society in §5. The source files for our codes and graphs used can be found in [11].

## 2. Multi-type bootstrap percolation

Given a finite or infinite graph  $G$  with vertex set  $V(G)$ , an integer  $r \in \mathbb{N}$ , and a set  $A$  of initially ‘infected’ vertices, the classical  $r$ -neighbour bootstrap process on  $G$  is defined as the process where, at each time, all of the vertices which have at least  $r$  already-infected neighbours become infected. Hence, one may define the set of infected vertices  $A_t$  at time  $t \in \mathbb{N}$  as

$$A_{t+1} = A_t \cup \{v \in V(G) : |N(v) \cap A_t| \geq r\},$$

where  $N(v)$  denotes the set of adjacent vertices to  $v$  in  $G$ , and  $|S|$  denotes the cardinality of the set  $S$ . The present paper is dedicated to a generalization of this model as follows.

In the simplest form of  $\mathcal{T}$ -BP described in definition 1.1, at each time  $t \in \mathbb{N}$ , the set of infected vertices is given by the set

$$A_{t+1} = A_t \cup \{v \in V(G) : |N_i(v) \cap A_t| \geq k_i \text{ for } i = 1, \dots, m\}, \quad (2.1)$$

where  $N_i(v)$  denotes the set of adjacent vertices to  $v$  in  $G$  with label  $i$ . The most generic form of our multi-type bootstrap percolation model is inspired by the concept of an update family  $\mathcal{U}$  from [12].

**Definition 2.1 ( $\mathcal{T}$ -bootstrap percolation).** A *trust family*  $\mathcal{T} := \{K^1, \dots, K^n\}$  is a tuple composed of trust vectors

$$K^j = (k_1^j, \dots, k_m^j) \in \mathbb{N}^m.$$

Then, we define *Trusted Bootstrap Percolation*, or  *$\mathcal{T}$ -bootstrap percolation*, as the percolation process for which  $A_0 := A$  and which, at time  $t + 1$  has infected vertices

$$A_{t+1} = A_t \cup \{v \in V(G) : \exists K_j \in \mathcal{T} \text{ s.t.} \quad (2.2)$$

and

$$|N_i(v) \cap A_t| \geq k_i^j \quad \text{for } i = 1, \dots, m\}. \quad (2.3)$$

**Remark 2.2.** We have opted to name our vectors  $K^j = (k_1^j, \dots, k_m^j)$  because they indicate the minimum number of people of each type required in order for a disease to be spread, or for a gossip to be transmitted. It gives the minimum number of ‘trusted’ people which have to believe the gossip before it can pass across to a new person.

It is important to note that classical  $r$ -neighbour bootstrap percolation is a particular case of  $\mathcal{T}$ -bootstrap percolation.

**Example 2.3.** An example of  $r$ -neighbour bootstrap is given by  $\mathcal{T}$ -bootstrap percolation for the trust family

$$\mathcal{T} := \left\{ K^j \in \mathbb{N}^m \mid \sum_{i=1}^m k_i^j = r \right\}. \quad (2.4)$$

Specifically, to recover  $r$ -neighbour bootstrap percolation with only one label one may consider  $m = 1$ , and set  $\mathcal{T} = \{K^1 = (r)\}$ .

When studying  $\mathcal{T}$ -BP, it is useful to bear in mind its application to society. For this, in its simplest form, the above set-up of  $\mathcal{T}$ -BP corresponds to considering a society with  $m$  different types of people, and a gossip that spreads only if it is passed by  $k_j$  number of people of type  $j$ . In the most generic set-up, the requirement for a gossip to spread is given by the existence of at least one trust vector  $K^i$  for which the gossip can be passed by  $k_j$  people of type  $j$ , for all types  $j$ .

**Definition 2.4.** We shall refer to *r-neighbour  $\mathcal{T}$ -BP* when considering a  $\mathcal{T}$ -BP model with exactly  $r$  integers  $k^i \in \{0, 1\}$  non-zero.

**Example 2.5.** Consider  $\mathcal{T}$ -BP with  $m = 3$ ,  $r = 2$  and  $k_i = 1$ . In this case, the model can be used to represent the spread of a political rumour among a society of Democrats, Republicans and Independents. Labelling the vertices with 1, 2, 3 to represent each political party, suppose that, in order to limit the spread of biased (and potentially false) information, there exists a rule that an individual will only believe and pass on the rumour, if he/she heard it from two people with different political backgrounds. Then, the trust family in this model would be

$$\mathcal{T} = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\},$$

and this is equivalent 2-neighbour  $\mathcal{T}$ -BP with three labels.

### 3. Immunity to gossip spread

In the following subsections, we shall focus on two forms of  $\mathcal{T}$ -bootstrap percolation that are of particular interest considering their proximity to classical bootstrap percolation, as well as their appearance within the most generic forms of our model: *r*-neighbour  $\mathcal{T}$ -BP, and the simplest  $\mathcal{T}$ -BP, which has single trust vector  $k$ . Within the  $\mathcal{T}$ -BP, a proportion of the vertices is *immune* to the infection: these vertices do not belong to  $A_t$  for any  $t \in \mathbb{N}$ , and we shall formally define the *immune set* by

$$\mathcal{I} := \left\{ v \in G : \forall K^j \in \mathcal{T}, \exists i \in [m] \text{ s.t. } |N_i(v)| < k_i^j \right\},$$

for  $\mathcal{T} = \{K^1, \dots, K^n\}$  the trust family, and  $K^i = \{k_1^i, \dots, k_m^i\} \in \mathbb{N}^m$  trust vectors as before.

In order to study the  $\mathcal{T}$ -BP on  $G$ , the vertices in  $\mathcal{I}$  need to be removed from  $G$ , leading to a modified graph which sometimes will be disconnected, with some components that might never become infected. To understand the likelihood of percolation of a  $\mathcal{T}$ -BP model and how immune vertices disrupt percolation on a network, it is useful to introduce the notion of *diversity*: given a vertex  $v \in G$  in a  $\mathcal{T}$ -BP model, define the *diversity* of  $v$  as the number  $D_v$  of different labels that vertices have:

$$D_v := |\{i \in \{1, \dots, m\} : N_i(v) \neq \emptyset\}|. \quad (3.1)$$

#### (a) Immunity for *r*-neighbour $\mathcal{T}$ -BP

In the case of *r*-neighbour  $\mathcal{T}$ -BP, a vertex is *immune* if it does not have neighbours of at least  $r$  distinct labels. From the definition of *r*-neighbour  $\mathcal{T}$ -BP, one can see that if a vertex is not immune, then  $D_v \geq r$ . Hence, it is of particular interest to understand which vertices  $v$  have  $D_v < r$ , since those will comprise all of the immune vertices.

**Example 3.1.** Returning to example 2.5, if an individual knows only Democrats, then it is impossible for them to be infected with gossip.

Consider  $\mathcal{T}$ -BP on a graph  $G$ , and let each vector component  $k_j \in \{0, 1\}$ . Then, the probability that a vertex  $v$  with  $|N(v)| = d$  has  $D_v \leq r - 1$  is

$$P(D_v \leq r - 1) = \frac{1}{m^d} \sum_{j=1}^{r-1} \left[ \binom{m}{j} (j!) \begin{Bmatrix} d \\ j \end{Bmatrix} \right], \quad (3.2)$$

where  $\left\{ \begin{matrix} d \\ j \end{matrix} \right\}$  is the Sterling number of the second kind (the reader may refer to appendix A for the proof of this statement). In particular, when the number of different labels required is equal to the number of labels that exist, i.e. when  $m = r$ , then

$$P(D_v \leq r-1) = \frac{1}{r^d} \sum_{j=1}^{r-1} \left[ \frac{r!}{(r-j)!} \left\{ \begin{matrix} d \\ j \end{matrix} \right\} \right]. \quad (3.3)$$

### (b) Immunity for the simplest form of $\mathcal{T}$ -BP

The second form of  $\mathcal{T}$ -bootstrap percolation we shall consider is the simplest form of  $\mathcal{T}$ -BP in which there is a single trust vector  $k = (k_1, k_2, \dots, k_m)$ . In this setting, given a vertex  $v$  with degree  $d$  on a graph  $G$ , consider the vector  $x = (x_1, x_2, \dots, x_m)$ , where  $x_i = |N_i(v)|$  is the number of adjacent vertices with label  $i$ . In particular,  $\sum_{i=1}^m x_i = d$ . Moreover, note that a vertex is immune if  $x_i < k_i$  for some integer  $i$ . In this case, the probability of immunity  $p_I^d(k)$  for a vertex  $v$  with  $|N(v)| = d$  is

$$p_I(d, k) = 1 - \sum_{x_m=k_m}^{d-(\sum_{l=1}^{m-1} k_l)} \binom{d}{x_m} \left(\frac{1}{m}\right)^{x_m} \left[ \sum_{x_{m-1}=k_{m-1}}^{d-x_m-(\sum_{l=1}^{m-2} k_l)} \binom{d-x_m}{x_{m-1}} \left(\frac{1}{m}\right)^{x_{m-1}} \left[ \dots \right. \right. \\ \left. \left. \times \left[ \sum_{x_2=k_2}^{d-(\sum_{l=3}^m x_l)-k_1} \binom{d-(\sum_{l=3}^m x_l)}{x_2} \left(\frac{1}{m}\right)^{x_2} \left(\frac{1}{m}\right)^{d-(\sum_{l=2}^m x_l)} \right] \dots \right] \right].$$

The reader should refer to appendix B for the proof of the above equation.

Given a  $\mathcal{T}$ -BP on a graph  $G$ , we shall denote by  $p_I(G, \mathcal{T})$  the expected fraction of immune vertices on  $G$ , and by  $p_I^d(\mathcal{T})$  the probability of immunity for a vertex of degree  $d$  with trust family  $\mathcal{T}$ . Then, one can show that in a  $\mathcal{T}$ -BP on a graph  $G$  with degree distribution  $P(d)$ , one has that the fraction of immune vertices is

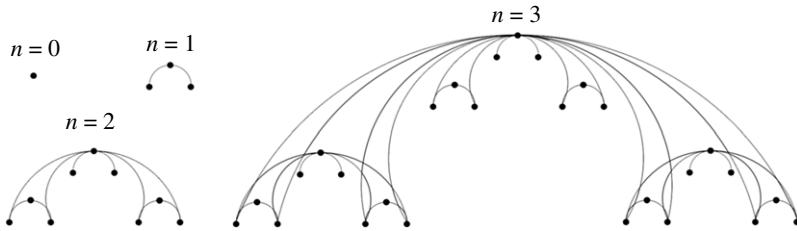
$$p_I(G, \mathcal{T}) = \sum_{j=0}^{m-1} p_I^j(\mathcal{T}) P(j). \quad (3.4)$$

**Example 3.2.** For a  $\mathcal{T}$ -BP with  $m = r = 2$  on  $\mathbb{Z}^2$ , one has that  $p_I^d(\mathcal{T}) = 1/8$ .

### (c) Immunity for $\mathcal{T}$ -BP on social networks

In order to understand the implications of  $\mathcal{T}$ -BP when modelling rumour spread in social networks, one needs to consider graphs  $G$  which accurately represent society. For this, note that many networks have a power-law degree distribution [13], which means that the fraction of vertices in  $G$  having degree  $d$  is approximately  $P(d) = d^{-\gamma}$  for some constant  $\gamma \in \mathbb{R}$ . For social networks particularly,  $\gamma$  is often around between 2 and 3 (see [13,14]). However, although social networks are relatively random, graphs known as deterministic hierarchical networks have power-law degree distributions while also having predetermined configurations, making them simpler to study. Thus, it becomes very interesting to investigate the  $\mathcal{T}$ -BP on a deterministic hierarchical network with  $2 \leq \gamma \leq 3$ . It should be noted that other networks have been shown to closely represent society (see for instance, those in [15]), and take into consideration many further characteristics of a social cluster, such as hubs and authorities, which are key in the spread of information. It would be thus very interesting to expand the present model to such networks, and we hope to come back to this in future publications.

An example of a deterministic hierarchical network with  $2 \leq \gamma \leq 3$  is defined in [16], and shown in figure 2. To construct it, start with one node, a root node. Add two more nodes and connect them to the root. At step  $n$ , add  $3^{n-1}$  nodes each, identical to the figure in the previous iteration (step  $n-1$ ) and connect the  $2^n$  bottom nodes to the root. The degree exponent for these graphs is  $\gamma = 1 + \ln 3 / \ln 2$ .



**Figure 2.** A hierarchical network as constructed in [16]. (Online version in colour.)

At  $n$  iterations there are  $(2/3)3^{n-i}$  vertices with degree  $2^{i+1} - 2$ . Hence, substituting  $k$  with  $2^{i+1} - 2$ , one can see that there are  $2(3^{n-\log_2(k+2)})$  vertices of degree  $d$  at this step. Hence we have that at  $n$  iterations, there are  $2(3^{n-\log_2(d+2)})$  vertices of degree  $d$ . There are also  $3^n$  vertices at  $n$  iterations and thus the probability a vertex has degree  $d$  at  $n$  iterations is  $2(d+2)^{-\log_2 3}$ .

The maximum degree of any vertex on this graph is the degree of the root, which at step  $n$  is given by  $2^{n+1} - 2$  [16]. Hence, this is the upper bound of the summation. To find  $p_I$ , or the expected fraction of initially immune vertices on this model with trust family  $\mathcal{T}$ , we need to find the probability of immunity for a vertex with  $j$  neighbours and then multiply that by the probability that a vertex has  $j$  neighbours. Finally, taking this product across all possible  $j$  one has that the expected fraction of immune vertices  $p_I(G, \mathcal{T})$  for this hierarchical network is given by

$$p_I(G, \mathcal{T}) = \sum_{d=1}^{2^{n+1}-2} 2(d+2)^{-\log_2 3} p_I^d(\mathcal{T}). \quad (3.5)$$

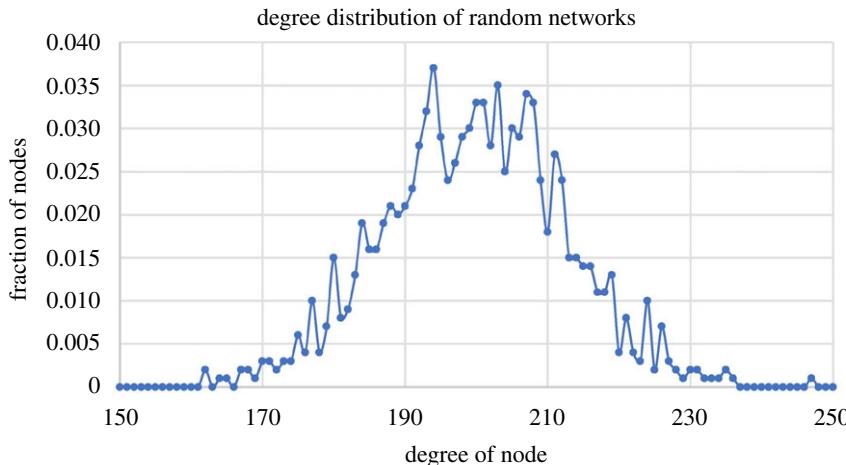
#### 4. $\mathcal{T}$ -BP on random networks

When studying a percolation model, one is particularly interested in the critical probability  $p_c$  describing the initial probability of infection that would make at last half the graph infected by the end of the process [2,3]. Since society can be modelled through certain random graphs (when adding adequate constraints to them), we shall dedicate the next section to the study of  $\mathcal{T}$ -BP on Erdos–Renyi graphs. In particular, we study the following properties, for which the results are primarily analytical:

- (I) The initial probability of infection  $p$  and the critical probability of percolation  $p_c$ ;
- (II) The fraction  $A_t$  of infected individuals at time  $t$ , and the final set of infected vertices  $A_\infty$  once the percolation process has finished.

Through these properties of the model, one can understand how gossip spreads or how marketing models based on  $\mathcal{T}$ -BP behave. Note that the initial probability of infection  $p$  determines the proportion of society that carries a gossip, or that have originally bought the objects for which the marketing campaign is being analysed.

In what follows, we shall consider the properties of  $\mathcal{T}$ -BP on random networks in terms of the main variables of the  $\mathcal{T}$ -BP model: the time  $t$ ; the number  $m$  of labels a vertex may have, the number  $r$  of different labels the set of infected neighbours must include in order for a vertex to be infected, and the probability  $\rho$  of having an edge between two vertices. Finally, throughout this section, for any value requiring multiple trials, we run 100 trials and fix  $n := |V(G)| = 10\,000$  vertices. The graphs used were generated using the code in [11], and have an 10 000 nodes and on average,  $\binom{10\,000}{2} * \text{den} = 4\,999\,500 * \text{den}$  edges. Each node has average degree  $10\,000 * \text{den}$ , where  $\text{den}$  represents the fixed density selected when constructing a random network, or the probability of having an edge between any two given nodes. One can see the computationally derived distribution of the degrees in the following figure 3:



**Figure 3.** Average degree distribution of the random networks considered in the paper, with 1000 nodes and  $\text{den} = 0.2$ . Plot made through 100 000 trials. (Online version in colour.)

Throughout the simulations, the immune vertices are disconnected from the network and considered following the analysis in the previous section, and the model we shall consider, unless otherwise stated, is  $r$ -neighbour  $T$ -BP as defined in definition 2.4.

### (a) Variation of the model's density

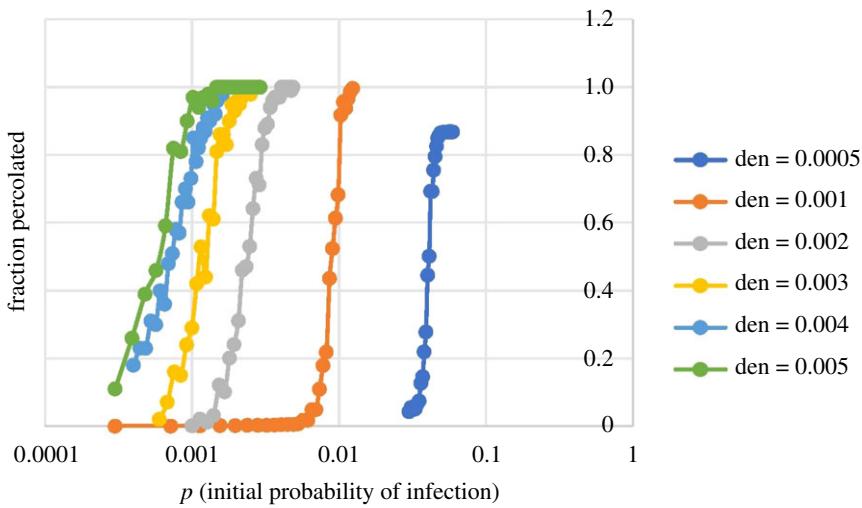
The percolation of the  $T$ -BP model depends heavily on the probability  $\rho$  of having an edge between two vertices. Indeed, one can see that for larger values of  $\rho$ , very low initial probabilities of infection will lead to the whole network being percolated by the end of the process, this is,  $A_\infty = G$ . Moreover, there is a clear wall crossing phenomena for the fraction of the graph that is percolated by the end of the process: a slight shift in  $\rho$  causes a large jump in the fraction percolated and also the probability of percolation.

An example of the above is given in figure 4. Indeed, figure 4 shows the relationship between the fraction of a graph that would end up percolated and the initial probability of infection,  $p$ . We considered this for a  $T$ -BP model with  $m = 3$  and  $r = 2$  on Erdos–Renyi graphs with various values for  $\rho$ . In particular, one can see that for densities  $\rho \geq 0.001$ , the model is very likely to completely percolate for any initial probability of infection  $p > 0.01$ .

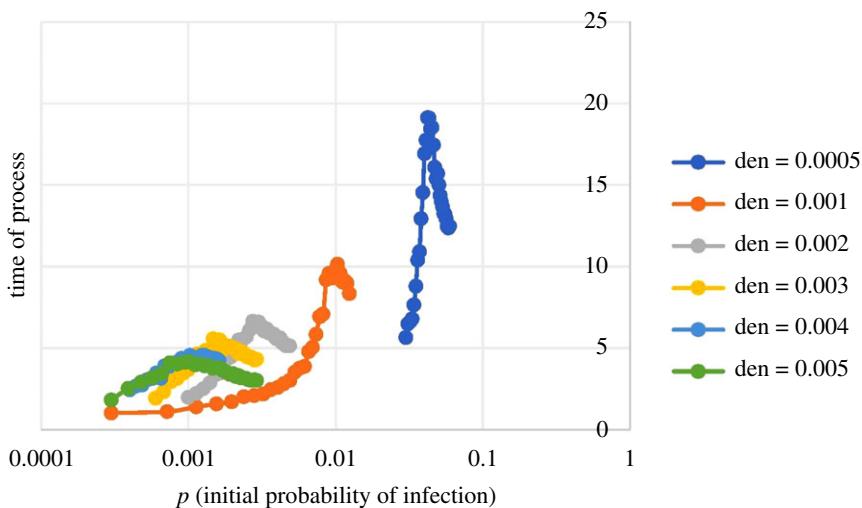
It is interesting to compare the initial probability  $p$  to the time the process takes to complete. Through 100 trials on a model with invariants  $n = 10\,000$ ,  $m = 3$  and  $r = 2$ , we can confirm what one would expect: the value of the initial probability  $p$  at which the time  $t$  is maximized is normally somewhat close to the value of the critical probability  $p_c$ , as depicted in figure 5 below.

**Example 4.1.** Consider the setting of example 2.5, where a society of 10 000 people who are Democrats, Republicans and Independents is modelled through  $T$ -BP with  $m = 3$ ,  $r = 2$  and  $k_i \in \{0, 1\}$ . In this setting, from figure 4 one can see that if people are on average connected to 10 people or more (this is  $\rho > 0.001$ ), and more than 100 people initially believe certain gossip (this is  $p > 0.01$ ), then very quickly the whole society will carry that gossip. On the other hand, if the average person only exchanges gossip with five people or less (considering  $\rho < 0.0005$ ), even if 300 people initially carried the gossip, the whole society would likely not end up carrying the gossip. In fact, we would expect less than 2000 people carrying the gossip by the end of the process (this is, the fraction of the graph percolated would be less than 0.2).

Through our model, one can look at individual curves in graphs as the ones in figure 4, and determine the critical probability  $p_c$  for some fixed value for  $\rho$ .



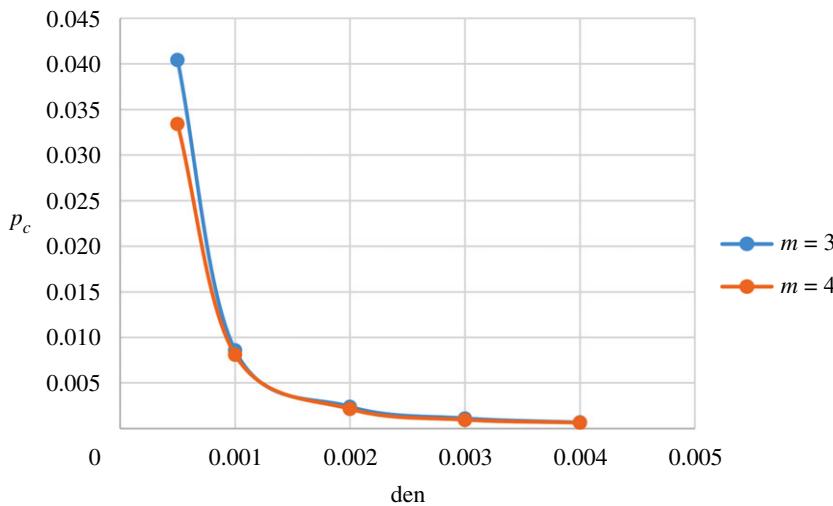
**Figure 4.** Fraction of the graph percolated as the initial probability of infection  $p$  varies, given various values of the density  $\rho$ , taking the integers  $n = 10\,000, m = 3, r = 2$ . Plot made through 100 trials, and with the horizontal axis on a logarithmic scale. (Online version in colour.)



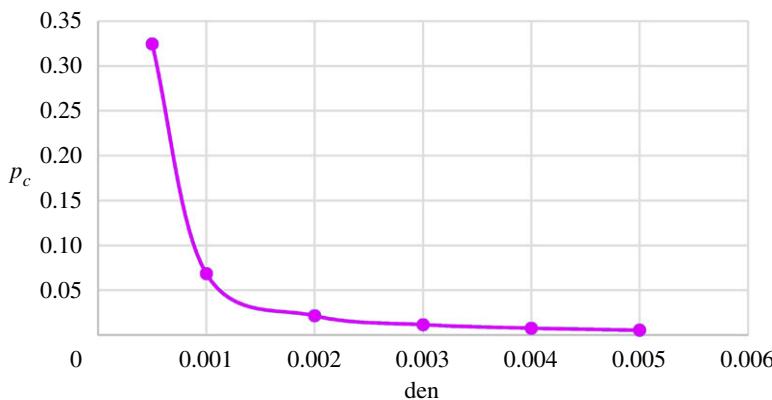
**Figure 5.** Under the same setting of figure 4, a comparison of  $p$  and the time the process takes to finish. (Online version in colour.)

In figure 6, we compare the value of  $p_c$  to the value of  $\rho$  for a 2-neighbour  $T$ -BP model with invariants  $m = 3, 4$ . As expected, one can see that the more labels one has (while requiring the same number of labels to be infected), the higher the critical probability is. In other words, the more different types of people a society has, the higher the initial probability of carrying a gossip needs to be in order for more than half the society to believe the gossip by the end of the process. Moreover, increasing the number of labels required makes the critical probability jump considerably. In particular, one should compare the blue curve in figure 6 and in figure 7.

**Example 4.2.** In order to see the importance of the number of non-zero  $k_i$  a model has, we shall study a variation of example 2.5. Consider a society of 10 000 people who are either Democrats, Liberals, Independents or Politically Agnostic (hence, having  $m = 4$ ). Suppose that the average person exchanges gossips with 20 people (this is, such that  $\rho = 0.002$ ).



**Figure 6.** Variation of  $p_c$  in terms of the density  $\rho$  for  $m = 3, 4$ , where  $r = 2$  and the number of vertices is  $n = 10\,000$ . (Online version in colour.)

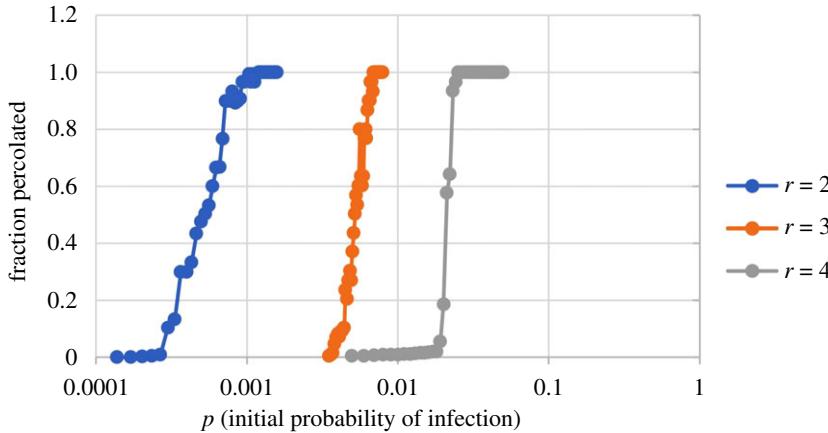


**Figure 7.** Variation of  $p_c$  in terms of the density  $\rho$  for  $m = 4$  and  $r = 3$ , where the number of vertices is  $n = 10\,000$ . (Online version in colour.)

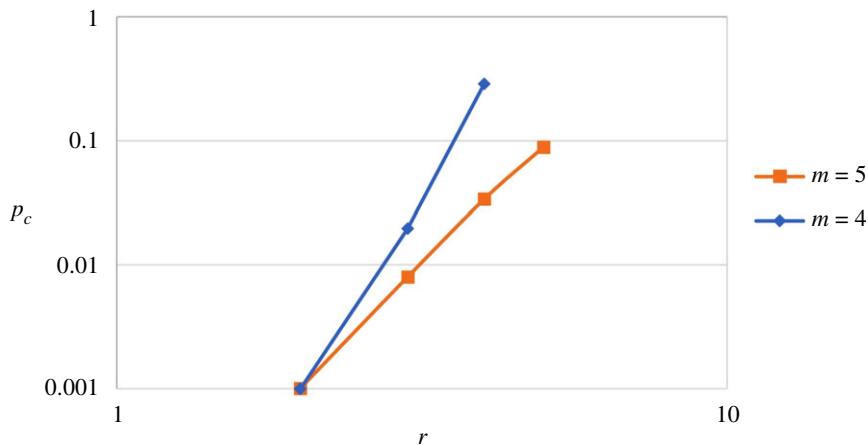
In this setting, if a person only passed a rumour if it had been heard by at least two different types of people, then at least 25 people would need to carry the gossip initially for it to spread to half the society or more (since  $p_c \sim 0.0025$ ). By contrast, if one required a rumour to be heard from at least three different types of people before it could be spread, then the gossip would need to be initially believed by 250 people for it to spread to half the society or more—this is, an order of magnitude more people than if we required only two different types of carriers.

### (b) Variation of the model's type

We shall now consider the relationship between the fraction of a graph that would end up percolated and the initial probability of infection  $p$ , while varying the number  $r$  of required labels. As one would expect, a  $T$ -BP model percolates much faster, and to a much larger proportion of the society, the smaller the number  $r$  of types required is, and this is illustrated in figure 8 where we fixed  $n$ ,  $\rho$  and  $m$ , and vary the number of labels:



**Figure 8.** Fraction of the graph percolated in terms of the variation of  $p$  for various  $r$ , where  $n = 1000$ ,  $m = 4$  and  $\rho = 0.005$ , and considering a log scale for the  $p$ -axis. (Online version in colour.)

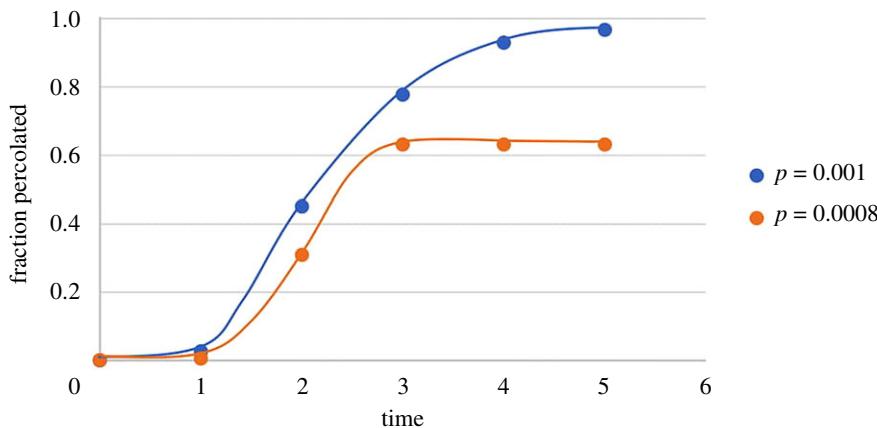


**Figure 9.** Variation of  $p_c$  with respect to  $r$  for  $(m, \rho) = (4, 0.005)$  and  $(m, \rho) = (5, 0.003)$ , for  $|V(G)| = 10\,000$ , where the axes follow log scales. (Online version in colour.)

**Example 4.3.** Similar to example 4.2, consider a town of 1000 people who are either Democrats, Liberals, Independents or Politically Agnostic (hence, having  $m = 4$ ), but suppose now that the average person exchanges gossips with 50 people (this is, such that  $\rho = 0.005$ ). Moreover, suppose that initially only 20 people believe certain gossip ( $p = 0.02$ ). By requiring someone to believe the gossip only if it is heard by every type of people, the gossip would not spread over more than 25% of society. If instead one only required any smaller number of different types of people, the gossip would always spread over the whole society by the end of the process.

In figure 9, we plot  $r$  against  $p_c$ , for various pairs of parameters  $(m, \rho)$ . As in the previous case,  $p_c(r)$  appears to fit a power law curve again, albeit more weakly in some cases. Numerically, one can see that for a fixed  $m$  the probability of percolation  $p_c(r, \rho)$  seems to obey a power law in both variables  $\rho$  and  $r$ , suggesting that

$$p_c(r, \rho) = cf(m)r^{e_1}\rho^{-e_2},$$



**Figure 10.** Average growth curve (comparing the  $t$  to  $\text{frac}_{\text{perc}}$ ) for graphs with  $m = 3$ ,  $r = 2$ ,  $n = 10\,000$  and  $\rho = 0.005$ . (Online version in colour.)

where  $c, e_1, e_2$  are all positive constants. Moreover, varying  $m$  also affects  $p_c$ , which is why the  $f(m)$  component is present.

We shall finally consider the average growth curve for fixed values of total types  $m$ , number of types required by the model  $r$  and density  $\rho$ . In terms of the initial probability of infection, we can see the growth in figure 10 where we considered the same density as in figures 8 and 9 to allow for comparison:

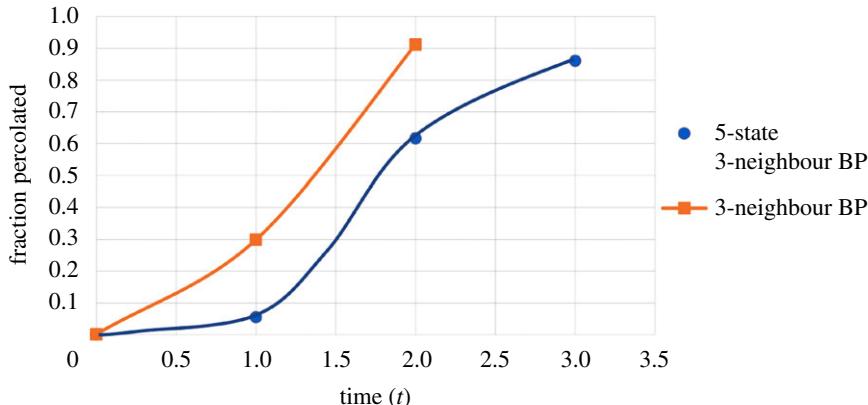
**Example 4.4.** One can see the relevance of figure 10 through the setting of example 2.5. Indeed, consider a society of Democrats, Republicans and Independents where one believes gossip only if at least two people from different parties tells it. Then, if originally only 8 people believe a gossip, by the end of the  $\mathcal{T}$ -BP process more than 30% of the society will not believe the gossip. On the other hand, if two more people believe the gossip originally (hence, having  $p = 0.001$ ), by the end of the percolation process all members of the society will believe it.

### (c) Comparison to standard bootstrap percolation

In order to study the propagation of an infection or a rumour in a random society, one would like to consider graphs with realistic densities. For this, recall that Dunbar found in [17] that the expected number of acquaintances any individual may have is 150. Hence, by consider  $\rho = 0.015$  with our current model, one can obtain representations of society through Erdos–Renyi graphs (although it should be noted that the 150 network is not homogeneous but is highly structured, and thus our model differs from that studied by Dunbar).

In the above setting, one has that a potential application for the  $\mathcal{T}$ -BP is in news-suggesting algorithms on social media: using the  $\mathcal{T}$ -BP as opposed to current algorithms could hinder the spread of fake news. Indeed, one can see the difference of the two models by comparing the growth of infection on the  $\mathcal{T}$ -BP as opposed to on classical  $r$ -neighbour bootstrap percolation, for current news-suggesting algorithms work very similarly to the latter model. As an example, we compare the average growth on 3-neighbour 5-state  $\mathcal{T}$ -BP and on 3-neighbour classical bootstrap percolation in figure 11.

The process for the  $\mathcal{T}$ -BP lasts an entire time step longer, and the growth is more gradual and a lower fraction of the society is infected by the end. In other words, the  $\mathcal{T}$ -BP makes biased news spread more slowly and to a lesser degree, indicating potential applications in replacing current news-suggesting algorithms on these social media platforms. Additionally, note the usage of  $\rho = 0.015$  in these trials, as they are being conducted specifically to determine the potential of the  $\mathcal{T}$ -BP on real social networks.



**Figure 11.** Average growth curve for graphs with  $\rho = 0.015$ ,  $n = 10\,000$ ,  $p = 0.0015$  and  $(m, r) = (5, 3)$  (for the  $\mathcal{T}$ -BP) and then  $r = 3$  (for the classical model). (Online version in colour.)

## 5. Conclusion

Bootstrap percolation has been used for years to model various percolation processes, with applications spanning from epidemiology to rumour spreading. The  $\mathcal{T}$ -BP model has been developed to add an extra layer to the current forms of bootstrap percolation, as the vertices of the graphs on which the infection spreads are now labelled. One should note that while many models have been developed to understand rumour spread (see [18–24]), these models are very different to  $\mathcal{T}$ -BP, and none of them consider multi-type percolations.

The  *$\mathcal{T}$ -Bootstrap percolation* ( $\mathcal{T}$ -BP model) was originally created to represent the spread of information based on the basic human instinct to trust information more if it comes from a variety of sources. This multi-type Bootstrap percolation model can be used to deduce interesting properties of social networks and their behaviour within different contexts (e.g. rumour spread, marketing, infection spread, etc.). While it is interesting to analyse the model on arbitrary random networks, it is of particular interest to consider random networks with plausible sizes and densities that accurately reproduce society (see [16,25–28]).

In this paper, we considered the hierarchical networks introduced in [16] to describe how immunity to rumour or disease spread appears through a  $\mathcal{T}$ -BP model within the network, and also studied the model on random networks with different densities.<sup>1</sup> of density between 2 and 3. In particular, we can see the following remarkable behaviours appear:

- **Delay in the spread.** Infections spread much faster through the classical  $r$ -neighbour bootstrap percolation than via the  $\mathcal{T}$ -BP (figure 11);
- **Containment of the infection.** An infection spreads across a greater percentage of the population in the classical  $r$ -neighbour bootstrap percolation than via the  $\mathcal{T}$ -BP (figure 11);
- **Trust vectors.** By requiring higher values of  $r$ , the percolation of the model is delayed (see for example, figure 8).

In particular, we can see how effective the  $\mathcal{T}$ -BP is in hindering the spread of fake news, and how by requiring higher levels of trust (higher values of  $r$ ), gossip would spread more slowly and to a lesser amount of people. Interestingly within the  $\mathcal{T}$ -BP models, some vertices may be immune, and in this paper, we studied the probability of a vertex being immune. The immunity

<sup>1</sup>It should be noted that this type of network does not guarantee the *small world* phenomenon or other aspects of real OSNs, and thus it would be very interesting to study the model presented in this paper on real OSN graphs which are publicly available online, such as in the SNAP network dataset collection.

of a vertex may be determined before the initial infection of the graph even occurs, as it is based upon a vertex's degree and the update rule for the graph. The probability  $p_I$  of immunity was presented here for the  $\mathcal{T}$ -BP as  $r$ -bootstrap percolation, the simplest form of  $\mathcal{T}$ -BP with only one trust vector, and  $\mathcal{T}$ -BP on the deterministic hierarchical graphs. Moreover, we were able to estimate the expected number of vertices which will be immune on any particular graph with these models. We also offer a rather loose lower bound for  $p_c$ , the critical probability of infection, based on  $p_I$ . Finally, we concluded our investigation by looking to random graphs, as these better represent the irregularities of society, and deriving analytical results on these graphs by running the  $\mathcal{T}$ -BP model on them computationally.

We expect the study of  $\mathcal{T}$ -BP models to be of particular interest from many different perspectives, and we shall conclude this paper mentioning a few of these lines of research which we plan to investigate in the future:

- **Deterministic hierarchical graphs.** With views towards applications of the model to sociology and marketing, and to eventually make the model stochastic by introducing probability of the infection spreading from one vertex to another.
- **Genetic diseases.** Dominant X-linked genetic diseases, which are only passed on if both the male and female parent have the disease, could be modelled by a  $\mathcal{T}$ -BP model with  $r = 2$  (a label for each sex). For recessive X-linked diseases, the model would need to have three labels—an infected female, a female carrier, and an infected male, and would need to be stochastic, as the probability of infection from a female carrier and infected male would only be  $1/2$ .

**Data accessibility.** This article has no additional data.

**Authors' contributions.** Both authors contributed equally to the work.

**Competing interests.** We declare we have no competing interest.

**Funding.** The work of L.P.S. is partially supported by the Humboldt Foundation as well as by the grant nos. NSF DMS 1509693 and NSF CAREER Award DMS 1749013.

**Acknowledgements.** The authors are thankful to PRIMES-MIT for their support, as well as to David Conlon and Dhruv Mubayi for useful conversations, to James Unwin for useful comments on a draft of the paper and to Ivan Smith, Fidel I. Schaposnik and Jan Verschelde for help with the code and for giving us access to a supercomputer. The authors are also very thankful for the numerous improvements suggested by the referees which have made the paper much better. L.P.S. also thank Robin Dunbar for many interesting remarks on the manuscript. L.P.S. is thankful to the Simons Center for Geometry and Physics for the hospitality during part of the preparation of the manuscript. This material is also based upon work supported by the National Science Foundation under grant no. DMS-1440140 while the author was in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Fall 2019 semester.

## Appendix A. The probability of immunity I

**Proposition.** Consider  $\mathcal{T}$ -BP on a graph  $G$ , and let each vector component  $k_j \in \{0, 1\}$ . Then, the probability that a vertex  $v$  with  $|N(v)| = d$  has  $D_v \leq r - 1$  is

$$p_I^d(m, r) := P(D_v \leq r - 1) = \frac{\sum_{j=1}^{r-1} \left[ \binom{m}{j} (j!) \left\{ \begin{smallmatrix} d \\ j \end{smallmatrix} \right\} \right]}{m^d},$$

where  $\left\{ \begin{smallmatrix} d \\ j \end{smallmatrix} \right\}$  is the Stirling number of the second kind.

*Proof.* We shall prove the above statement through the principle of inclusion and exclusion. In order to do this, we shall first calculate the probabilities  $P(D_v = i)$  for  $i = 1, \dots, r$ , and use this to calculate  $p_I^d(m, r)$ .

Given a fixed integer  $n$  such that  $1 \leq n \leq r - 1$ , in order to understand  $P(D_v = n)$  note that there are  $\binom{m}{n}$  ways to choose the  $n$  acceptable labels that the neighbours may have, or in other words,

ways to choose a set  $N \subset \{1, \dots, m\}$  with  $|N| = n$ . Once these labels are chosen, the number of onto functions such that  $f(N(v)) = N$  is given by

$$\sum_{i=0}^n (-1)^{n-i} \binom{k}{i} (n-i)^d. \quad (\text{A } 1)$$

To see this, begin by noting that each of the  $d$  adjacent vertices in  $N(v)$  has  $n$  possible values for its label, so there are  $n^d$  functions from  $N(v)$  to  $N$ . However, not all functions will be different. To visualize this, consider a Venn diagram where each circle  $A_s$  of the Venn diagram corresponds to a label  $s$ , with  $1 \leq s \leq m$ , and is defined as  $A_s := \{f : N(v) \rightarrow N : s \notin N\}$ . For instance, every object in  $A_1$  will be a possible configuration of neighbours of  $v$  such that none of them are of label 1.

All functions which are not surjective to  $N$  must appear in some set  $A_i$  for  $i \in N$ . Therefore, the number of onto functions  $f : N(v) \rightarrow N$  is given by

$$n^d - \left| \bigcup_{i \in N} A_i \right|. \quad (\text{A } 2)$$

In order to calculate (A 2), recall that by the Principle of Inclusion and Exclusion,

$$\left| \bigcup_{i=1}^{\alpha} A_i \right| = \sum_{i=1}^{\alpha} |A_i| - \sum_{1 \leq i < j \leq \alpha} |A_i \cap A_j| \quad (\text{A } 3)$$

$$+ \sum_{1 \leq i < j < k \leq \alpha} |A_i \cap A_j \cap A_k| + \dots + (-1)^{\alpha-1} \left| \bigcap_{i=1}^{\alpha} A_i \right|. \quad (\text{A } 4)$$

Thus, to find  $|\bigcup_{i \in N} A_i|$  one needs to consider  $\sum |A_{x_1} \cap \dots \cap A_{x_j}|$  for all  $j$  such that  $1 \leq j \leq n$ , and  $x_i \in N$ , which is given by

$$\sum |A_{x_1} \cap \dots \cap A_{x_j}| = \binom{n}{j} (n-j)^d, \quad (\text{A } 5)$$

since there are  $\binom{k}{j}$  ways to choose the  $j$  labels out of  $N$  that the neighbours cannot occupy, and so there are  $n-j$  options for every neighbour's label. Then, one has that

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{i=1}^n (-1)^{n-i+1} \binom{n}{i} (n-i)^d, \quad (\text{A } 6)$$

and thus the number in (A 2) of surjective functions  $f : N(v) \rightarrow N$  is

$$\sum_{i=0}^n (-1)^{n-i+1} \binom{n}{i} (n-i)^d. \quad (\text{A } 7)$$

From the above, one can calculate the number of different label assignments that the vertices  $N(v)$  can have, where  $D_v = j$ : there are  $\binom{m}{j}$  possibilities for  $j$  labels, which multiplied by (A 7) leads to

$$\binom{m}{j} \sum_{i=0}^j (-1)^{j-i+1} \binom{j}{i} (j-i)^d. \quad (\text{A } 8)$$

Equivalently, we can write this as  $\binom{m}{j}(j!)\left\{\begin{array}{c} d \\ j \end{array}\right\}$  where  $\left\{\begin{array}{c} d \\ j \end{array}\right\}$  is the Stirling number of the second kind. Hence, the number of different label assignments to the graph  $G$  for which the diversity  $D_v \leq r - 1$  is

$$\sum_{j=1}^{r-1} \left[ \binom{m}{j}(j!)\left\{\begin{array}{c} d \\ j \end{array}\right\} \right]. \quad (\text{A } 9)$$

Recalling that there are  $m^d$  possible functions  $f: N(v) \rightarrow \{1, \dots, m\}$ , the probability of a vertex having diversity  $D_v \leq r - 1$  is given by

$$P(D_v \leq r - 1) = \frac{\sum_{j=1}^{r-1} \left[ \binom{m}{j}(j!)\left\{\begin{array}{c} d \\ j \end{array}\right\} \right]}{m^d}, \quad (\text{A } 10)$$

which concludes the proof.  $\blacksquare$

## Appendix B. The probability of immunity II

**Proposition.** *In a T-BP with  $k = (k_1, \dots, k_m)$ , the probability of immunity  $p_I^d(k)$  for a vertex  $v$  with  $|N(v)| = d$  is*

$$\begin{aligned} p_I^d(K) = 1 - & \sum_{x_m=k_m}^{d-(\sum_{l=1}^{m-1} k_l)} \binom{d}{x_m} \left(\frac{1}{m}\right)^{x_m} \left[ \sum_{x_{m-1}=k_{m-1}}^{d-x_m-(\sum_{l=1}^{m-2} k_l)} \binom{d-x_m}{x_{m-1}} \left(\frac{1}{m}\right)^{x_{m-1}} \left[ \dots \right. \right. \\ & \times \left. \left. \sum_{x_2=k_2}^{d-(\sum_{l=3}^m x_l)-k_1} \binom{d-(\sum_{l=3}^m x_l)}{x_2} \left(\frac{1}{m}\right)^{x_2} \left(\frac{1}{m}\right)^{d-(\sum_{l=2}^m x_l)} \right] \dots \right]. \end{aligned}$$

*Proof.* Let  $f(a, b, c)$  be the function defined as

$$\begin{aligned} f(a, b, c) := & \sum_{x_i=b_i}^{a-(\sum_{l=1}^{i-1} b_l)} \binom{a}{x_i} \left(\frac{1}{c}\right)^{x_i} \left[ \sum_{x_{i-1}=b_{i-1}}^{a-x_i-(\sum_{l=1}^{i-1} k_l)} \binom{a-x_i}{x_{i-1}} \left(\frac{1}{c}\right)^{x_{i-1}} \dots \right. \\ & \times \left. \left[ \sum_{x_2=b_2}^{a-(\sum_{l=3}^i x_l)-b_1} \binom{a-(x_i+\dots+x_3)}{x_2} \left(\frac{1}{c}\right)^{x_2} \left(\frac{1}{c}\right)^{a-(\sum_{l=2}^i x_l)} \right] \dots \right], \end{aligned}$$

where  $b = \{b_1, b_2, \dots, b_i\}$  and  $a, c \in \mathbb{N}$ . Then, one can see that  $p_I = 1 - f(d, k, m)$  for  $k = \{k_1, k_2, \dots, k_m\}$ . Indeed, this can be proven with an inductive argument on the number of available labels  $m$ . For  $m = 2$ , the vector  $k = \{k_1, k_2\}$  satisfies  $k_1 + k_2 \leq d$ . We must find the probability of assigning  $d$  vertices to one of 2 labels such that there are at least  $k_1$  vertices of label 1 and  $k_2$  of label 2—this holds when a vertex is not immune, so we must then subtract this from 1. This is equivalent to saying there may be  $x_1$  vertices of label 1 such that  $k_1 \leq x_1 \leq d - k_2$ , and all other vertices of label 2. Note that the probability that there are  $x_1$  vertices of label 1 is:

$$\binom{d}{x_1} \left(\frac{1}{2}\right)^{x_1} \left(\frac{1}{2}\right)^{d-x_1}. \quad (\text{B } 1)$$

Summing over all possible values for  $x_1$  varying from  $k_1$  to  $d - k_2$ , the overall  $p_d$  for  $m = 2$  and  $k = \{k_1, k_2\}$  is found to be:

$$\sum_{x_1=k_1}^{d-k_2} \binom{d}{x_1} \left(\frac{1}{2}\right)^{x_1} \left(\frac{1}{2}\right)^{d-x_1}. \quad (\text{B } 2)$$

Note that this equals  $f(d, \{k_1, k_2\}, 2)$ . Then, the probability of immunity would be

$$1 - \sum_{x_1=k_1}^{d-k_2} \binom{d}{x_1} \left(\frac{1}{2}\right)^{x_1} \left(\frac{1}{2}\right)^{d-x_1}. \quad (\text{B } 3)$$

Now, move on to  $m = 3$  with  $k = \{k_1, k_2, k_3\}$ . We must find the probability of assigning  $d$  vertices to one of three labels such that there are at least  $k_1$  vertices of label 1,  $k_2$  of label 2, and  $k_3$  vertices of label 3. We can approach this with casework.

First, note that there may be  $x_1$  vertices of label 1 such that  $k_1 \leq x_1 \leq d - (k_2 + k_3)$ . Then, take cases based on the value of  $x_1$ . Note that given  $x_1$ , there are  $d - x_1$  remaining vertices to consider, with 2 possible labels to assign them to. Now this question is almost the same as the one with  $m = 2$ , with the only difference being that the probability a vertex occupies one of these two labels (2 or 3) is not  $1/2$ , but  $1/3$ . So given  $x_1$ , the probability that the number of vertices of label 2 is greater than  $k_2$  and the number of vertices of label 3 is greater than  $k_3$  is just  $f(d - x_1, \{k_2, k_3\}, 3)$ . This means that given  $x_1$ , the probability that the vertex is initially not immune (do not forget that we are using complementary counting) is

$$\binom{d}{x_1} \left(\frac{1}{3}\right)^{x_1} f(d - x_1, \{k_2, k_3\}, 3). \quad (\text{B } 4)$$

Now, as  $x_1$  can go from  $k_1$  to  $d - (k_2 + k_3)$ ,  $p_I^d$  for  $m = 3$  is:

$$1 - \sum_{x_1=k_1}^{d-(k_2+k_3)} \binom{d}{x_1} \left(\frac{1}{3}\right)^{x_1} f(d - x_1, \{k_2, k_3\}, 3). \quad (\text{B } 5)$$

Note that this equals  $f(d, \{k_1, k_2, k_3\}, 3)$ . Therefore,

$$\begin{aligned} p_I^d(\{k_1, k_2, k_3\}) &= 1 - f(d, \{k_1, k_2, k_3\}, 3) \\ &= 1 - \sum_{x_1=k_1}^{d-(k_2+k_3)} \binom{d}{x_1} \left(\frac{1}{3}\right)^{x_1} f(d - x_1, \{k_2, k_3\}, 3), \end{aligned}$$

which provides the intuition for the inductive step: we must prove that

$$p_I^d(\{k_1, \dots, k_m\}) = 1 - f(d, \{k_1, \dots, k_m\}, m),$$

assuming that there exists an  $i$  such that  $p_I^d(\{k_1, \dots, k_i\}) = 1 - f(d, \{k_1, \dots, k_i\}, i)$ .

To prove the statement by induction, consider  $m = i + 1$ , and let  $k = \{k_1, k_2, \dots, k_{i+1}\}$ . Using casework as we did for the example where  $m = 3$ , let the number of vertices of label 1 be  $x_1 \in \{k_1, \dots, d - \sum_{j=2}^{i+1} k_j\}$ . Then, there are  $d - x_1$  vertices which must have labels in  $[i+1] \setminus \{1\}$ . By the inductive assumption, the probability that they have at least  $k_j$  vertices of label  $j$  is  $f(d, \{k_2, \dots, k_{i+1}\}, i)$ . However, note that we must replace the  $1/i$  terms in this formula with  $1/(i+1)$  because the probability any individual vertex has a specific label is now  $1/(i+1)$ , meaning that it is actually  $f(d, \{k_2, \dots, k_{i+1}\}, i+1)$ . Also, the probability that there are  $x_1$  vertices of label 1 is  $\binom{d}{x_1} 1/(i+1)^{x_1}$ , so multiplying these two terms one finds that the probability that a vertex is initially *not* immune given that it has  $x_1$  neighbours of label 1 is

$$\binom{d}{x_1} \left(\frac{1}{i+1}\right)^{x_1} f(d, \{k_2, \dots, k_{i+1}\}, i+1). \quad (\text{B } 6)$$

Since  $x_1$  can be anything from  $k_1$  to  $d - \sum_{j=2}^{i+1} k_j$ , summing over the possible values of  $x_1$  and subtracting this summation from 1 leads to the probability

$$p_I^d(\{k_1, k_2, \dots, k_{i+1}\}) = 1 - \sum_{x_1=k_1}^{d-(\sum_{j=2}^{i+1} k_j)} \binom{d}{x_1} \left(\frac{1}{i+1}\right)^{x_1} f(d, \{k_2, \dots, k_{i+1}\}, i+1). \quad (\text{B } 7)$$

Finally, note that by definition this is simply  $1 - f(d, \{k_1, \dots, k_{i+1}\}, i+1)$ . This concludes the inductive step, and so we have proven that

$$p_I^d(\{k_1, k_2, \dots, k_{i+1}\}) = 1 - f(d, \{k_1, \dots, k_{i+1}\}, i+1),$$

finalizing the proof. ■

## References

- Chalupa J, Leath PL, Reich GR. 1979 Bootstrap percolation on a Bethe lattice. *J. Phys. C.* **12**, L951. (doi:10.1088/0022-3719/12/24/508)
- Balister P, Bollobás B, Przykucki M, Smith P. 2016 Subcritical  $\mathcal{U}$ -bootstrap percolation models have non-trivial phase transitions. *Trans. Am. Math. Soc.* **368**, 7385–7411. (doi:10.1090/tran/6586)
- Balogh J, Bollobás B, Morris R. 2010 Bootstrap percolation in high dimensions. *Combin. Probab. Comput.* **19**, 643–692. (doi:10.1017/S0963548310000271)
- Balogh J, Peres Y, Pete G. 2006 Bootstrap percolation on infinite trees and non-amenable groups. *Combin. Probab. Comput.* **15**, 715–730. (doi:10.1017/S0963548306007619)
- Balogh J, Pittel BG. 2007 Bootstrap percolation on the random regular graph. *Random Struct. Algorithm* **30**, 257–286. (doi:10.1002/rsa.20158)
- Holroyd AE. 2003 Sharp metastability threshold for two-dimensional bootstrap percolation. *Probab. Theory Related Fields* **125**, 195–224. (doi:10.1007/s00440-002-0239-x)
- Janson S. 2009 On percolation in random graphs with given vertex degrees. *Electron. J. Probab.* **14**, 87–118. (doi:10.1214/EJP.v14-603)
- Baxter GJ, Dorogovtsev SN, Goltsev AV, Mendes JFF. 2010 Bootstrap percolation on complex networks. *Phys. Rev. E* **82**, 011103. (doi:10.1103/PhysRevE.82.011103)
- Baxter GJ, Dorogovtsev SN, Goltsev AV, Mendes JFF. 2011 Heterogeneous k-core versus bootstrap percolation on complex networks. *Phys. Rev. E* **83**, 051134. (doi:10.1103/PhysRevE.83.051134)
- Di Muro MA, Valdez LD, Stanley HE, Buldyrev SV, Braunstein LA. 2019 Insights into bootstrap percolation: its equivalence with k-core percolation and the giant component. *Phys. Rev. E* **99**, 022311. (doi:10.1103/PhysRevE.99.022311)
- Bhansali R, Schaposnik LP. source files, Available at this address as well as in <https://github.com/rinbha/A-Trust-Model-in-Bootstrap-Percolation>.
- Bollobás B, Smith PJ, Uzzell AJ. 2015 Monotone cellular automata in a random environment. *Comb. Probab. Comput.* **24**, 687–722. (doi:10.1017/S0963548315000012)
- Onnela J-P, Saramaki J, Hyvonen J, Szabo G, Lazer D, Kaski K, Kertesz J, Barabasi A-L. 2007 Structure and tie strengths in mobile communication networks. *Proc. Natl Acad. Sci. USA* **104**, 7332–7336. (doi:10.1073/pnas.0610245104)
- Choromaski K, Matuszak M, -MieKisz J. 2013 Scale-free graph with preferential attachment and evolving internal vertex structure. *J. Stat. Phys.* **151**, 1175–1183. (doi:10.1007/s10955-013-0749-1)
- Nettleton D. 2013 Data mining of social networks represented as graphs. *Comput. Sci. Rev.* **7**, 1–34. (doi:10.1016/j.cosrev.2012.12.001)
- Barabási AL, Ravasz E, Vicsek T. 2001 Deterministic scale-free networks. *ScienceDirect* **299**, 559–564. (doi:10.1016/S0378-4371(01)00369-7)
- Dunbar RIM. 1992 Neocortex size as a constraint on group size in primates. *J. Human Evol.* **22**, 469–493. (doi:10.1016/0047-2484(92)90081-J)
- Calìò A, Tagarelli A. 2018 Trust-based dynamic linear threshold models for non-competitive and competitive influence propagation. In *17th IEEE Int. Conf. on Trust, Security and Privacy in Computing and Communications, New York, NY, 1–3 August*, pp. 156–162. Piscataway, NJ: IEEE.
- Ferraz de Arruda G, Rodrigues FA, Rodriiguez PM, Cozzo E, Moreno Y. 2016 Unifying Markov chain approach for disease and rumor spreading in complex networks. (<http://arxiv.org/abs/1609.00682>)
- Junior V, Machado F, Zuluaga M. 2011 Rumor processes on N. *J. Appl. Probab.* **48**, 624–636. (doi:10.1239/jap/1316796903)
- Lee MJ, Lee D-S. 2019 Understanding the temporal pattern of spreading in heterogeneous networks: theory of the mean infection time. *Phys. Rev. E* **99**, 032309. (doi:10.1103/physreve.99.032309)
- Lind PG, da Silva LR, Andrade Jr JS, Herrmann HJ. 2007 Spreading gossip in social networks. *Phys. Rev. E* **76**, 036117. (doi:10.1103/physreve.76.036117)
- Pradeep R, VijayKumar VR. 2018 A case study for modeling and analysis of rumour spreading rate in two layer complex network. *IJAICT* **4**, 1296–1300.
- Sajadi FA, Roy R. 2017 On rumour propagation among sceptics. (<http://arxiv.org/abs/1706.02858>)

25. Chakrabarti D, Zhan Y, Faloutsos C. 2004 R-MAT: a recursive model for graph mining. In *Proc. 4th SIAM Int. Conf. on Data Mining* Lake Buena Vista, FL, 22–24 April, pp. 442–446. Philadelphia, PA: SIAM.
26. Fox J, Roughgarden T, Seshadhri C, Wei F. 2018 Finding cliques in social networks: a new distribution free model. In *45th Int. Colloquium on Automata, Languages, and Programming (ICALP 2018)*, Prague, Czech Republic, 9–13 July, article no. 55; pp. 1–15. Wadern, Germany: Schloss Dagstuhl – Leibniz Center for Informatics.
27. Opsahl T, Agneessens F, Skvoretz J. 2010 Node centrality in weighted networks: generalizing degree and shortest paths. *Social Netw.* **32**, 245–251. (doi:10.1016/j.socnet.2010.03.006)
28. Otte E, Rousseau R. 2002 Social network analysis: a powerful strategy, also for the information sciences. *J. Inf. Sci.* **28**, 441–453. (doi:10.1177/016555150202800601)