1 *The Cascade Model* The cascade model is a simple toy model to help us think about energy flow in directed acyclic graph such as food webs and some power grids. Imagine N nodes with indices $i = 1, ..., N$ and place an undirected edge between each unique pair with probability p just as in the ER random graph seen in class. Now add a direction to every edge such that each edge goes from node i to node $i' < i$. The index i then informs us of the position of a node in a hierarchy (e.g. trophic level in food webs, or distance from source in power grids) and the model can be used to help us understand the distribution of energy flow through the hierarchy.

a)

b) Since $N - i$ nodes are $>$ node i and $i - 1$ nodes are $<$ node i,

$$\langle k_{in} \rangle + \langle k_{out} \rangle = \langle k \rangle = p(N-1) \simeq pN$$
$$p(N-i) + p(i-1) = p(N-i+i-1)$$
$$\Rightarrow \boxed{\langle k_{in} \rangle = p(N-i)}$$
$$\Rightarrow \boxed{\langle k_{out} \rangle = p(i-1)}$$

c) Since for each node $\leq i$ we only are concerned with the edges coming into it from nodes $> i$, the $\langle k_{in} \rangle$ of interest will be $p(N-i)$ for all nodes $i' < i$, meaning the expected number of edges going from nodes $> i$ to nodes $\leq i$ will be expressed as

$$\sum_{1}^{i}(N-1)p = i(N-i)p = \boxed{p(Ni - i^2)} \tag{1}$$

d) Assuming N is even, to find the **maximum** value of equation 1,

$$\max_{i} \ \left(p(Ni - i^2) \right)$$

We want to take the derivative and set it to 0 to find which i maximizes the equation,

$$\therefore pN - 1pi = 0$$
$$\Rightarrow i = \frac{pN}{2p} = \frac{N}{2}$$
$$\Rightarrow \boxed{\max_{i} \ \left(p(Ni - i^2) \right) = \frac{1}{4}N^2 \quad \text{at } i = \frac{N}{2}}$$

And the **minimum** occurs at $\boxed{i = N}$, at which equation (1) $= 0$

2 *degree-regular networks* While Poisson or ER networks correspond almost exactly to well-mixed systems, degree-regular networks are the simplest generalization of cellular automata to networks. Degree-regular networks are essentially a version of the configuration model seen in class but where every node has the same degree k. Hence, if we randomized the connections between every site in an infinite square grid, we would end up with an infinite degree-regular network of degree 4.

a) For a d-degree-regular network, the degree distribution would be

$$\{p_k\} = \left\{ \begin{array}{ll} 1 & \text{if } k = d \\ 0 & \text{otherwise} \end{array} \right\}$$

And the generating functions $G_0(x)$ & $G_1(x)$ would be,

$$G_0(x) = \sum_{k=0}^{k_{max}} p_k x^k = p_d x^d = \boxed{x^d} \text{ for the degree distribution}$$

$$G_1(x) = \frac{G_0'(x)}{G_0'(1)} = \frac{dx^{d-1}}{d} = \boxed{x^{d-1}} \text{ for the excess degree distribution}$$

b) If we set $u \equiv$ the probability of not reaching the giant component when following an edge[1]. Then it follows that

$$u = G_1(u) = u^{d-1} \tag{2}$$

So, we are interested in the possible solutions for u for different $d \in \{0, \mathbb{Z}_+\}$.   For both $d = 0$ and $d = 1$, the only solution is $u = 1$, meaning we would expect to never arrive at a single, giant component,

$$u = u^{0-1} \qquad\qquad\qquad u = u^{1-1}$$
$$u = u^{-1} \qquad\qquad\qquad u = u^0$$
$$u^2 = 1 \qquad\qquad\qquad u = 1 \tag{3}$$

When $d = 2$,
$$u = u^{2-1} \Rightarrow u = u \tag{4}$$

so, every possible value for $u \in [0, 1]$ is a possible solution, meaning we can just as well expect there to be a giant component as we can expect for there to not be a giant component. At all other values of $d > 2$, the only solutions for u are 0 and 1, but 1 will always be a solution of $u = u^{d-1}$, therefore the unique solution for all of these cases is $u = 0$, and we can expect to always reach a single, giant component.

$$u = u^{3-1} \Rightarrow u = u^2 \tag{5}$$

 Therefore, the **minimal** value for d where we can expect to get a single, connected, giant component is $d = 3$.

---

[1]In other words, the probability of there **not** being one single, giant, component.

c)

As shown above in Eq. (3), for $d = 0$ we can expect to never arrive at a single, giant component, which makes sense and seems obvious, since no node can connect to any other node since they all have 0 edges. So, we would just have a network of singleton, isolated nodes[2].
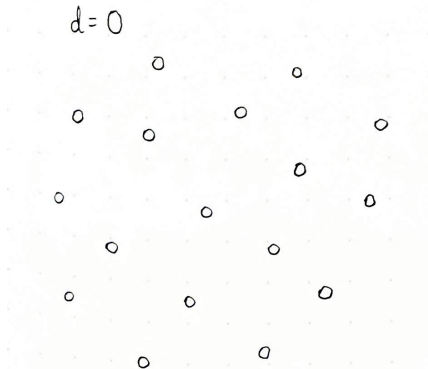
Also shown in Eq. (3), for $d = 1$, we still never expect a single giant component, which is also obvious because every node can only connect to 1 other node, so we would just have many connected pairs, not connected to any other nodes in the network[3].
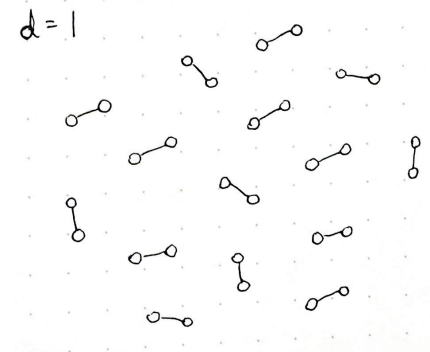
For Eq. (4), we have $d = 2$, our most interesting case, where every node is connected to two other nodes, which allows for either an infinite chain (line) of nodes connected to one another, and/or loops of nodes connected to one another of any size. We would likely get a distribution of these loops of different sizes as we randomly rearrange our network by disconnecting and reconnecting nodes randomly and would have some chance of getting one single infinite chain that all nodes are a part of, but this chance is very small.

---
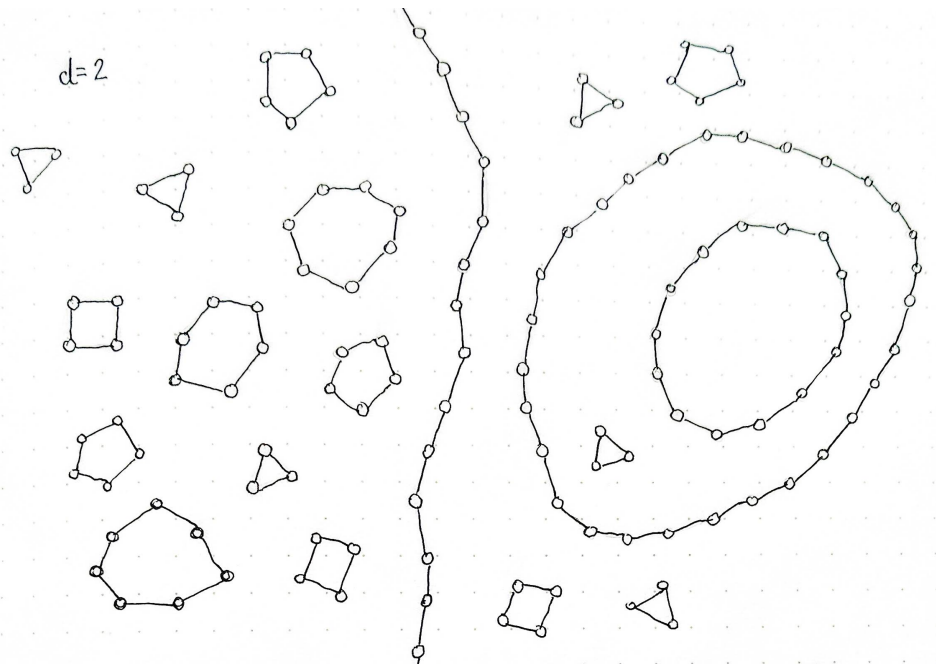
[2]Pretty uninteresting
[3]Also, pretty uninteresting...

d) The pairwise approximation of an SIS model for a degree-regular network of degree 4 would be,

$$\dot{I} = \lambda[SI] - I \qquad \text{where } [SI] \text{ is the number of edges between } S \ \& \ I \text{ nodes.} \qquad (6)$$

$$S = N - I \qquad \text{and } \lambda \text{ is our renormalized parameter } \frac{\beta}{\alpha}.$$

$$[\dot{SI}] = 2[II] - [SI] + \lambda\frac{3}{4}[SI]\left(\frac{2[SS]}{S} - \frac{[SI]}{S}\right) - \lambda[SI]$$

$$\Rightarrow [\dot{SI}] = 2[II] - [SI]\left(1 + \frac{3\lambda}{2S}\left([SS] - \frac{1}{2}[SI]\right) - \lambda\right) \qquad (7)$$

where the moment closure involved, $[SI]\frac{3}{4}\frac{2[SS]}{S}$ is an approximation of the number of triplets, of two S nodes and one I node, using only pairwise (edge) information.

$$[\dot{SS}] = [SI] - \lambda\frac{3}{4}[SI]\left(\frac{2[SS]}{S}\right) = [SI]\left(1 - \frac{3\lambda}{2S}[SS]\right) \qquad (8)$$

$$[II] = \frac{N(4)}{2} - [SI] - [SS] = 2N - [SI] - [SS] \qquad (9)$$

3 *Epidemic spreading*

a) We can think of a version of the SIS model where instead of having compartments S and I, we have k_{max}, I compartments, $\{I_0, I_1, I_2, ..., I_{k_{max}}\}$, where k_{max} is the highest degree in the network. This can take into account the reality that high-degree nodes are more likely to get infected than low-degree. We can then write an equation for the change in the number of infected nodes of each degree k,

$$\dot{I}_k = \lambda k\theta(p_k - I_k) - I_k \qquad (10)$$

where p_k is the fraction of nodes with degree k in the network and $\theta = \left(\sum_k kI_k\right)/\langle k\rangle$

b) If we put in a new state, *vaccinated nodes*, where the transmission rate (to and from these nodes) is reduced by a factor of f, we get a new equation,

$$\dot{I}_k = \lambda k\theta(p_k - I_k - V_k) + f\lambda k\theta(V_k) - (1 + f)I_k \qquad (11)$$
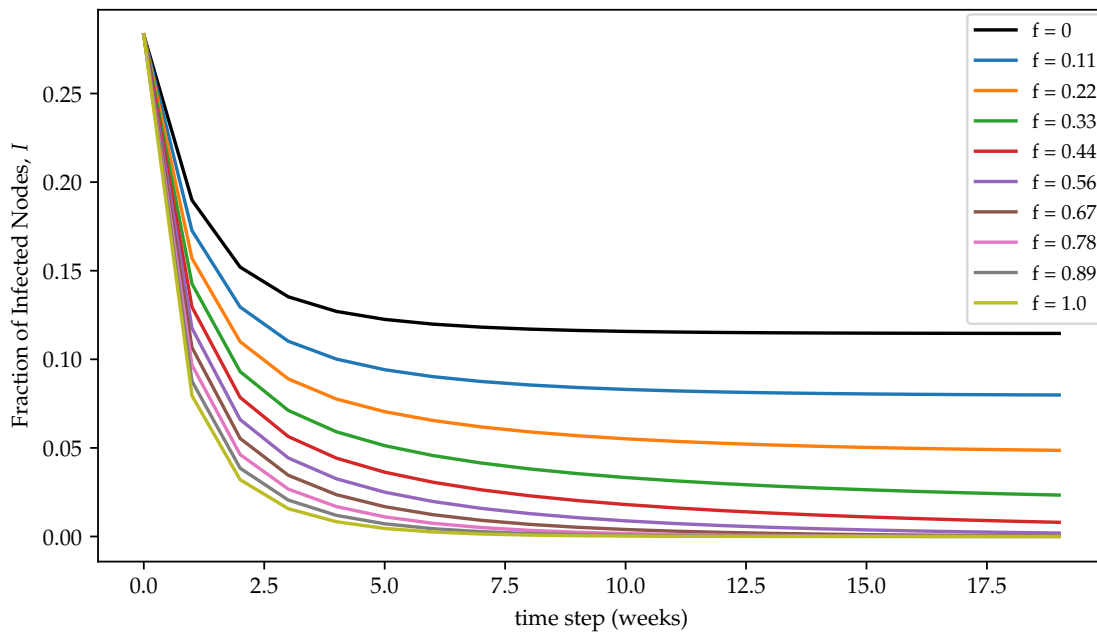
c)
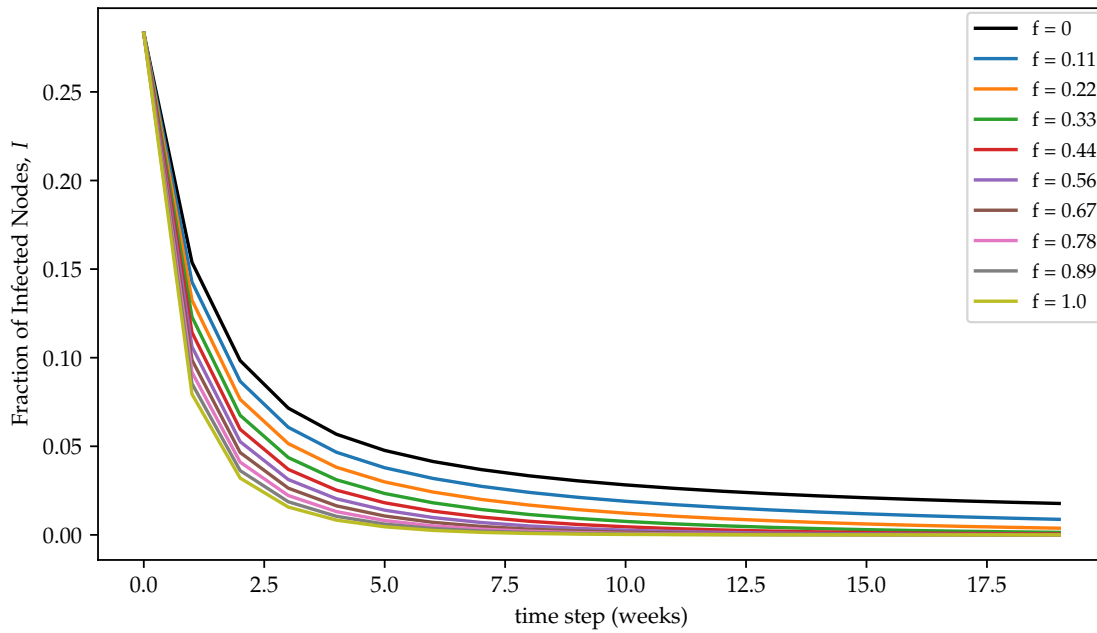
Figure 1

d)

Figure 2

Adding a vaccinated state is also achieved by modeling the interactions between compartmental pairs $I_k V$ and $I_k S$. In the case of the interaction between infected and susceptible nodes, the

equation is very much the same.

$$\frac{dI_kS}{dt} = \beta k(p_k s - I_k S)\theta - I_k S \tag{12}$$

For infected nodes to interact with vaccinated ones however, the equation must incorporate the dampening factro f, giving us the following equation.

$$\frac{dI_kV}{dt} = \beta k f(p_k v - I_k V)\theta - \frac{1}{f}I_k V \tag{13}$$
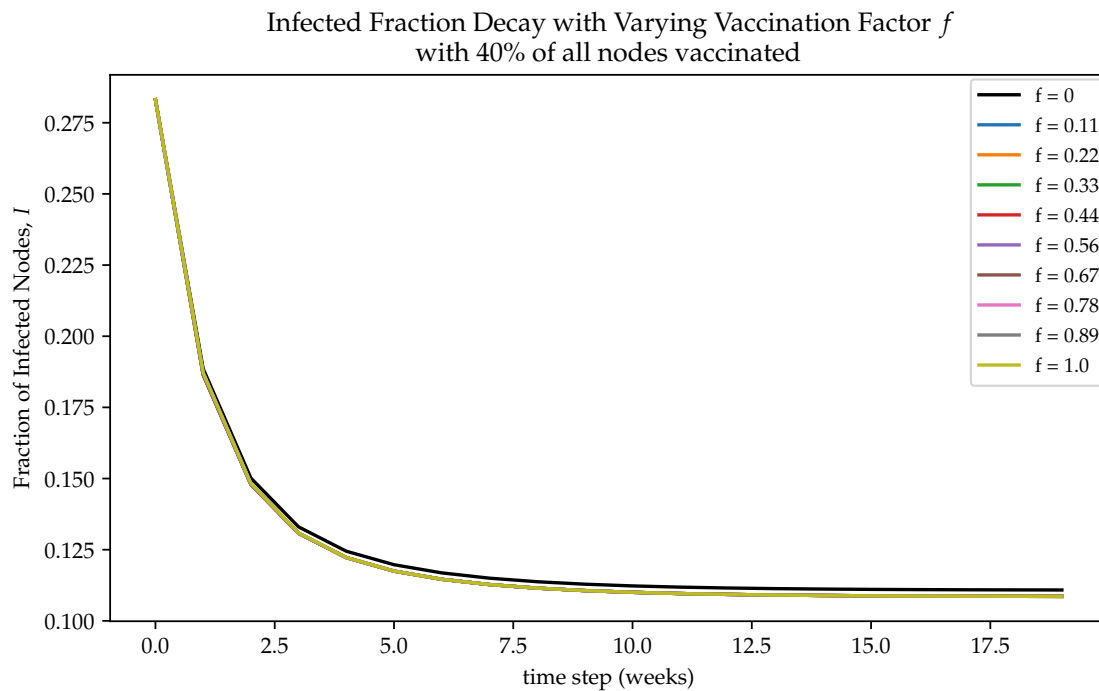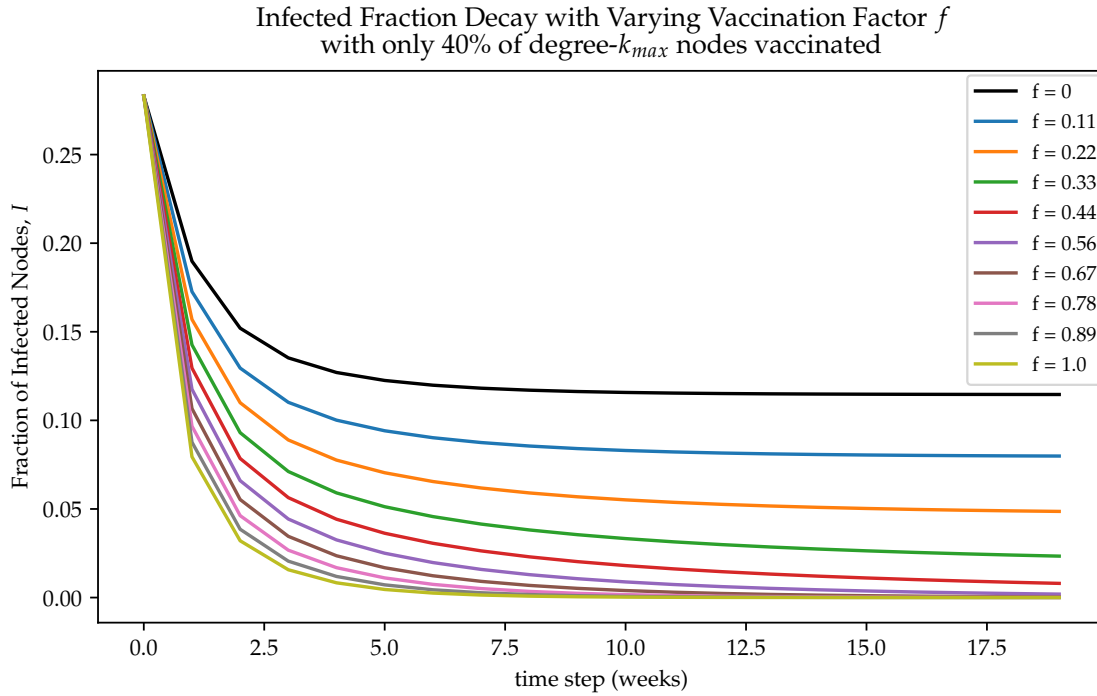
c)

Figure 3

d)

Figure 4

In both cases we believe something in either the equations or our code is incorrect. While $f = 0$ should correspond to a perfect vaccine, in both methods we achieve the exact opposite. Similarly, $f = 1$ counterintuitively is the best case scenario. Additionally, vaccinating the high degree nodes should result in a quicker and more effective campaign, which is not the case for our current solution.

4 *the voter model*

a) The network chosen represents an e-mail network between members of the University Rovira i Virgili, where nodes represent email recipients or senders and edges symbolize that two university affiliates exchanged an email. There are 1133 nodes and 5452 edges in the giant component of this network. The average degree is $\langle k \rangle = 9.62$ and the diameter is 8. Notably, the minimum, maximum, and average closeness centrality across nodes in the graph are (0.18,0.38,0.28) while the min, max, and average betweenness is (0, 0.04, 0.002). This implies that each node in the system not only has many lines of communication open (given the high average degree, low shortest path lengths, and small diameter) but there is also a lack of high throughput nodes; there are very little nodes in the network that act as third parties in order for two nodes to communicate.

b) Below shows a summary of 10 voter model runs of 100 timesteps for different proportions of "red" and "blue" opinions. The plots show the times to achieve consensus, the probability of consensus, and the general trends in opinion ratios for each proportion.

c) Noting that in general it takes less than half of the maximum time for the system to reach consensus, we can postulate that the previously observed high average closeness and low average
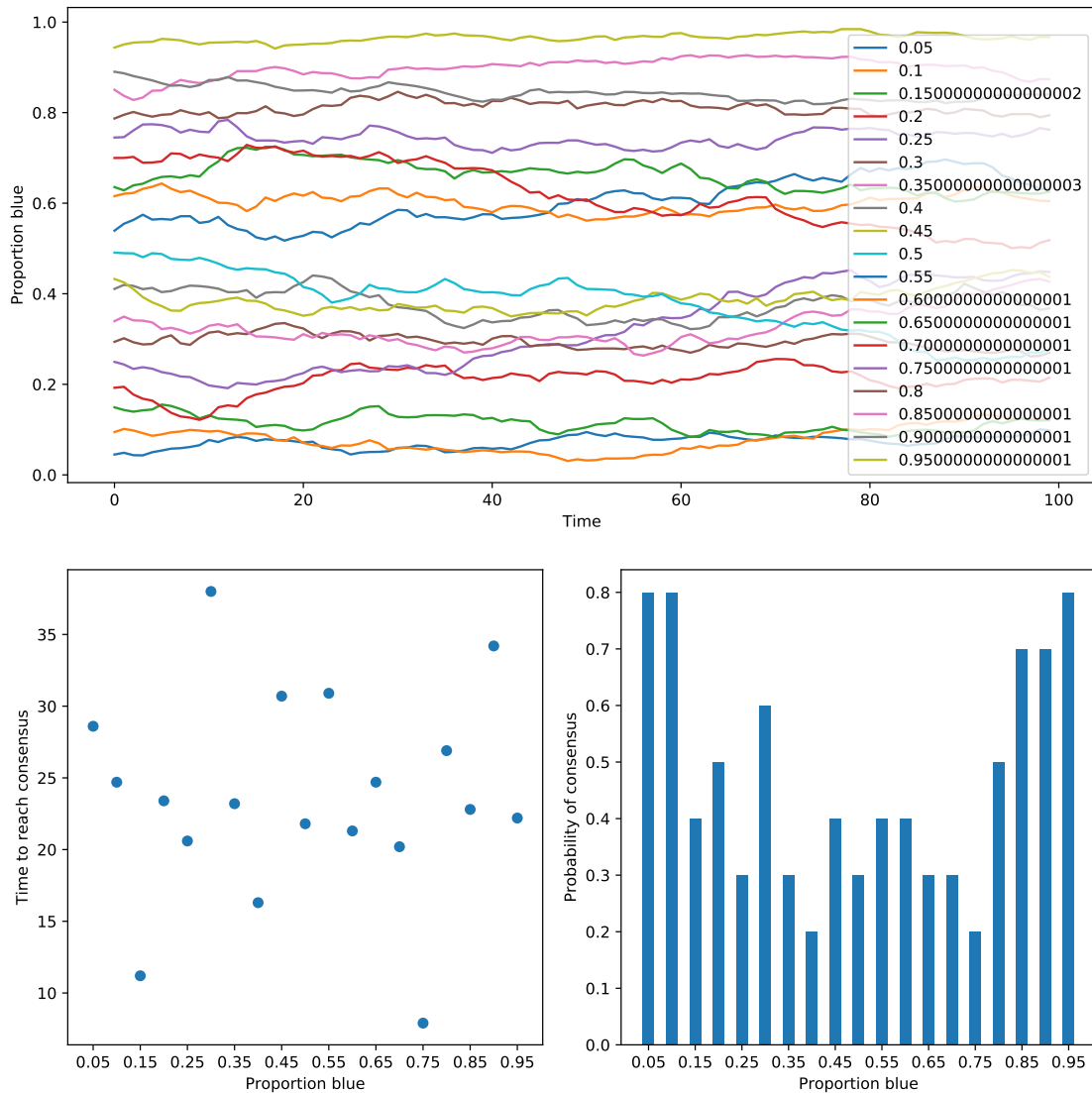
Figure 5: Opinion dynamical trends (Top), time to achieve consensus (bottom left) and probability of consensus (bottom right). For the voter model on the email network. Time to reach consensus for roportions between 0.35 and 0.75 are less exact given their lower probability of consensus.

betweenness opinions to easily propagate through the network. If a node changes their opinion, "news" of that nodes opinion change is likely to move quickly through the network, changing others opinions along the way.