



# LUND UNIVERSITY

## **Publish Late, Publish Rarely!**

### **Network Density and Group Performance in Scientific Communication**

Angere, Staffan; Olsson, Erik J

*Published in:*

Scientific Collaboration and Collective Knowledge

*DOI:*

[10.1093/oso/9780190680534.003.0001](https://doi.org/10.1093/oso/9780190680534.003.0001)

2017

*Document Version:*

Peer reviewed version (aka post-print)

[Link to publication](#)

*Citation for published version (APA):*

Angere, S., & Olsson, E. J. (2017). Publish Late, Publish Rarely! Network Density and Group Performance in Scientific Communication. In T. Boyer, C. Mayo-Wilson, & M. Weisberg (Eds.), *Scientific Collaboration and Collective Knowledge* (pp. 34-62). Oxford University Press.  
<https://doi.org/10.1093/oso/9780190680534.003.0001>

*Total number of authors:*

2

#### **General rights**

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00



# Publish Late, Publish Rarely! Network Density and Group Performance in Scientific Communication

Staffan Angere  
Erik J. Olsson  
Lund University

Abstract: It is almost a tautology to propose that communication is a good thing and that one should, as a scientist, try to be as “connected” as one can. And yet studies in various disciplines, from social psychology to economics, persistently undermine this notion. These studies have found a lot of communication links in a network of inquirers to be detrimental to group performance in a wide range of cases. We hypothesize that these results are partly due to the effect of network “spamming”: as more links are added so that the density of the network increases, it becomes more sensitive to low quality information. In the present study, we first replicate this negative effect in the Bayesian simulation framework Laputa. We then show how network density can actually be positively correlated with group performance if inquirers agree (i) to communicate only information they judge to be highly certain and (ii) not to recycle information in the absence of new supporting evidence. Finally, we suggest that the results shed doubt on the recommendation “publish early, publish often” frequently given to young scholars. As the scientific community becomes increasingly connected, scholars should, in the interest of science, be increasingly concerned with publishing only material they take to be practically certain and avoid recycling material for which no new supporting evidence has been forthcoming.

## 1. Introduction

Group inquiry is an essential aspect of scientific activity, as we know it. It takes place not only in research groups inquiring into some part of nature or culture, but also in various scientific committees and boards deciding what scientific projects to fund, whom to appoint for an academic position etc. Outside science, group inquiry and deliberation play a prominent role in juries in court, in political committees, in the corporate boardroom, and in many other places. Whereas in the infancy of scientific investigation much activity arguably took place at the individual level, this is not how inquiry is conducted in many, if not most, areas of modern science. And of course, the possibility of instant electronic communication has increased group activity and interaction dramatically.

But is communication actually beneficial in these cases? Does it actually make for better conclusions or decisions? The question may seem unexpected, even imbecile. Of course it is good to communicate, one is inclined to exclaim. And indeed, this gut reaction is perfectly appropriate in many cases in science and elsewhere. Here is one reason for this: Group inquiry typically focuses one or more research questions to be solved by collective means. By a complex question we mean a question that cannot be answered without being further divided into subordinate questions. Thus, the question whether John committed the robbery is more complex than the question whether he had a motive. To be answerable, the former needs to be decomposed into more specific questions, including questions about his motive. Complex questions abound in science. For instance, the question whether there is man-made global warming is a complex question *par excellence* presupposing for instance an answer to the question whether a higher level of CO<sub>2</sub> causes an increase in global temperature. Now, many complex questions are naturally, or even necessarily, treated by means of division of labor where

different research groups work on different sub-problems requiring different competences. At some point, these groups need to pool their results, i.e. communicate their findings among the members, or the overarching question cannot be answered. So, in this case it can hardly be denied that communication is beneficial, indeed essential, just as common sense insists it should be. In the following we will focus instead on (relatively) simple questions that do not obviously call for such division of labor.

Suppose that we wish to maximize group performance in solving a common problem. Suppose, moreover, that the group members are given, and what we can play with is the communication structure of the group, that is, we can decide who is to communicate with whom. If so, how should we “hook people up”? Common sense arguably still dictates that “the more group members can communicate with each other, the more competent the group will become”. If so, then the complete graph, where there is a communication link from everyone to everyone else, is the way to go.

Unfortunately, studies in various disciplines – including social psychology and economics – persistently undermine, or problematize, the commonsense answer. These studies have found a lot of communication links, i.e. high *network density*, to be detrimental to group performance in a wide range of cases. In section 2, we review four such studies. In the remainder of the paper, we explore the issue within the framework of our own Bayesian model of communication, implemented in the simulation environment Laputa. We introduce this model in section 3 and 4. In section 5, we describe the setting within which our simulation experiments are carried out. As we report in section 6, one experiment shows that the “repugnant conclusion” that network density is generally detrimental to group performance can be replicated in our framework. This leads us to asking whether increasing the quality of information in the network can make network density more attractive. We study, in section 7, the effect of imposing two such conditions, which effectively disallow inquirers from *recycling* already posted information under two different interpretations of what recycling means, and we observe that network density can be positively correlated with group performance, construed as average closeness to the truth (in a sense to be defined), if these conditions are imposed. We also observe that there is a tradeoff to be made between group performance and group polarization.

## 2. Background

Economists Bala and Goyal (1998) studied decision making in a group context using mathematical modeling and analysis. The inquirers in the model could choose between two actions without knowing which action is optimal. They could make use of information regarding the pay-off of their own previous action, and the pay-offs of their neighbors in the network. As an example the authors take a case of medical inquirers working on a particular disease. There are two possible states of the world:

$\varphi_1$  = The old method is better.

$\varphi_2$  = The new method is better.

There are, correspondingly, two actions inquirers can take:

$A_1$  = Pursuing the old method.

$A_2$  = Pursuing the new method.

Because the old method is assumed to be well understood, pursuing it will not result in any new information about the state of the world. Unbeknownst to our inquirers, the new method is in fact better. Inquirers can see the results of some of the other inquirers. This gives rise to a graph structure where the nodes represent inquirers and the links communicational connections. Inquirers update their subjective probabilities of  $\varphi_2$  based on their own results and the result of the other inquirers they can see by means of Bayesian conditionalization. A population of inquirers has finished learning just in case one of the following conditions has been met: all inquirers take action  $A_1$  or all inquirers believe that they are in  $\varphi_2$  (with probability greater than 0.9999). A population of inquirers has learnt successfully just in case all inquirers believe that they are in  $\varphi_2$  (with probability greater than 0.9999). Otherwise they have “learnt unsuccessfully”. Bala and Goyal now show that, if inquirers are arranged on a line, all inquirers will learn successfully, i.e. believe truthfully that the new method is better. Surprisingly, however, if there is, in addition, a small “royal family” of inquirers communicating with all the others, inquirers sometimes end up learning unsuccessfully, i.e. pursuing the worse method. Hence the royal family network, which is more “connected”, is actually less reliable, collectively, than the linear network. Bala and Goyal conclude that “more informational links can increase the chances of a society getting locked into a sub-optimal action” (p. 609).

Bala and Goyal’s study raises a number of questions. What is it about the group with a “royal family” that reduces group performance? Is it the particular communicational graph structure? Is it the fact that the graph is more dense, by which we mean that it has a higher number of edges per node? What role does the sheer size of the group play (in cases of finite groups)? Some answers can be found in Zollman (2007), which is a detailed study of Bala and Goyal’s model for finite groups of inquirers using computer simulation. Zollman compared networks with a cycle or wheel structure with the complete graph regarding probability of successful learning and learning speed. He also looked at the effect of group size on these matters. Zollman went on to study the general relationship between communication graph density and group performance for different group sizes. He found that sparser networks are generally more reliable than denser ones. However, learning speed increases with density. In other words, a denser network learns faster but less reliably.

Mason, Jones and Goldstone (2008) studied the propagation of “innovations” in networked groups in laboratory experiments with real participants. Participants had their computers arranged in a virtual network structure and made repeated numerical guesses for which they received scores. The scores were also made available to their neighbors in the network. The task was for each participant to maximize his or her score over 15 rounds. The networks were compared on speed of discovery and convergence to the optimal solution (within a small margin of error) for a number of different score (payoff) functions with different shapes where the score functions used were unknown to the participants. The percentage of participants who guessed the number with the highest global score as a function of the number of “rounds” was studied. Mason et al found that denser networks are generally faster but less reliable than sparser networks, such as a “small world” network: “The advantage of the small-world to the fully connected network is akin to a novel group-based form of the ‘less is more’ effect reported in individual decision making literature” (ibid., p. 430). A suggestive result concerns a case of an easy-to-find local maximum but a difficult-to-find global maximum. The fully connected network converged quickly on the local maximum and stayed there. The lattice structure, composed of “small research groups”, was eventually more reliable

in finding the global maximum. The study of Mason et al indicates that the network structure may influence the learning speed as well as the reliability of a group in ways that depend on the structure of the problem space, i.e. on things like how many locally or globally optimal solutions there are, and how easy it is to find those solutions.

Lazer and Friedman (2007), surveying research in this area, reached similar conclusions: “Remarkably, the highest performing network in the long run was the slowest, most error prone, and had the longest average path length. More generally, our results highlight that given a long time horizon, efficient communication can actually be counterproductive to systemic performance, whereas given a short time horizon, more communication generally contributes to better outcomes. Our results thus suggest a potential dark side to the rush to connectedness in today’s world: it will limit our pathways to developing new and better ways of doing things, even as the demand for new solutions is increasing.”

A reasonable conjecture is that the negative conclusions in the literature, as far as the epistemic value of network density is concerned, are partly due to the effect of network “spamming”: as the density of the network increases, it becomes more sensitive to low quality information. An interesting question which is not dealt with systematically in these studies, or any other study we know of, is whether the conclusions carry over to cases where people are more discriminate as to when to post information in the network, for instance by refraining from posting things whose truth value they are not quite sure of. We will study this issue in the context of the model implemented in the software Laputa which we have developed for the purpose of studying epistemic properties of networked communication, drawing on the influential approach to social epistemology in Goldman (1999).<sup>1</sup>

### 3. A model of scientific inquiry and communication

We define a *research network*  $N$  as a group of people or other agents (including some organizations) engaging in communicative and investigative practices. Examples include university departments, peer-review boards, circles of friends, or even entire societies. As a limiting case, the entirety of humanity during a period can be seen as such a network. The “during a period” clause is necessary here: we are taking the network, i.e. the communicative and investigative practices themselves, to be fairly stable. We also assume that the participants in the network remain fixed. Members of a research network are referred to as *inquirers*.

We assume that there is a single empirical proposition  $p$  whose truth-value the inquirers are trying to ascertain. As the model we are using is Bayesian, we take each inquirer  $\alpha$ , at each time  $t$ , to have a certain *credence*  $C_\alpha^t(p)$  in  $p$ . This value represents each inquirer’s current knowledge about the state of  $p$ . But we also have to model the way that the participants receive new information. There are two fundamentally different inlets for this: inquiry and communication.

Inquiry can here be taken to include any kind of method of altering a credence function which does not base itself on information given by anyone else in the network; in physicists’ terms, it corresponds to *external* forces. Some paradigmatic examples of inquiry are observation, experiment, and perhaps taking advice from persons outside  $N$ . In order for our model to be applicable, the opinions of inquirers in  $N$  must have fairly

---

<sup>1</sup> See for instance Masterton and Olsson (2013), Olsson (2011, 2013), Olsson and Vallinder (2013), Vallinder and Olsson (2003, 2014). Laputa can be downloaded at <http://www.luiq.lu.se/portfolio-item/laputa/>.

little effect on such “external” persons’ opinions, however. An example would be a modern philosopher reading Immanuel Kant, who does not himself need to be taken as a part of our social network, since his opinions cannot be affected by anything we say or do.

Not all inquirers’ approaches to inquiry are the same, and they tend to vary both in their degree of activity and their effectiveness. Let  $S_{i\alpha}^t(p)$  be the proposition “ $\alpha$ ’s inquiry gives the result that  $p$  at time  $t$ ”, let  $S_{i\alpha}^t(\neg p)$  be the proposition “ $\alpha$ ’s inquiry gives the result that not- $p$  at  $t$ ”, and let  $S_{i\alpha}^t \equiv_{df.} S_{i\alpha}^t(p) \vee S_{i\alpha}^t(\neg p)$  be the proposition that  $\alpha$ ’s inquiry gives *some* result at  $t$ . We represent the participants’ properties *qua* inquirers by two probabilities: the chance  $P(S_{i\alpha}^t)$  that, at the moment  $t$ ,  $\alpha$  receives a result from her inquiries, and the chance  $P(S_{i\alpha}^t(p)|S_{i\alpha}^t, p)$  that, when such a result is obtained, it is the right one.<sup>2</sup>  $P(S_{i\alpha}^t)$  will be referred to as  $\alpha$ ’s *inquiry chance*, and  $P(S_{i\alpha}^t(p)|S_{i\alpha}^t, p)$  as her *inquiry accuracy*.

It is important to keep in mind that these, unlike  $C_{\alpha}^t(p)$ , are *objective chances*. They can be interpreted statistically, or perhaps in terms of propensities. In our implementation of the model they will be fed into a pseudorandom number generator to determine what happens during a simulation.

An inquirer without interest in  $p$  would generally have a low value of  $P(S_{i\alpha}^t)$ , while one very interested in  $p$ , but engaging in inquiry using faulty methods would have a high value of  $P(S_{i\alpha}^t)$ , but an inquiry accuracy close to 0.5, or even below that. In the latter case, the results of her inquiry would actually be negatively correlated with the truth. As a simplification, we will assume  $\alpha$ ’s inquiry chance and accuracy to be constant over time, so we will usually write them without the time index  $t$ .

Just as inquiry represents the flow of information *into* the network, communication deals with how this information is disseminated *within* the network. While, as a first approximation, we may be interested only in whether or not  $\alpha$  can receive information from  $\beta$ , we generally need to go deeper. Thus, we take the network  $N$  to include a set of what we will call *links*, each corresponding to a communication channel. Such channels can be as direct as conversation, or mediated like a blog or an instant messaging system. Generally, however, we may want to say that if the “mediator” is able to choose which messages to transmit, it may be more apt to be represented as an inquirer instead of a link, even if it engages in no inquiry itself (this is one case where we may want to set  $P(S_{i\alpha}^t) = 0$ ). A scientific journal could be an example of this, as in the network depicted below.

---

<sup>2</sup> This is only half the story:  $P(S_{i\alpha}^t(p)|S_{i\alpha}^t, p)$  gives the chance of a true report when  $p$  is the case, and does not say anything about reliability when  $\neg p$  holds. To simplify matters, we assume that these probabilities are equal, i.e. that the chance that an agent’s inquiry gives the correct result does not depend on whether  $p$  is true or false.

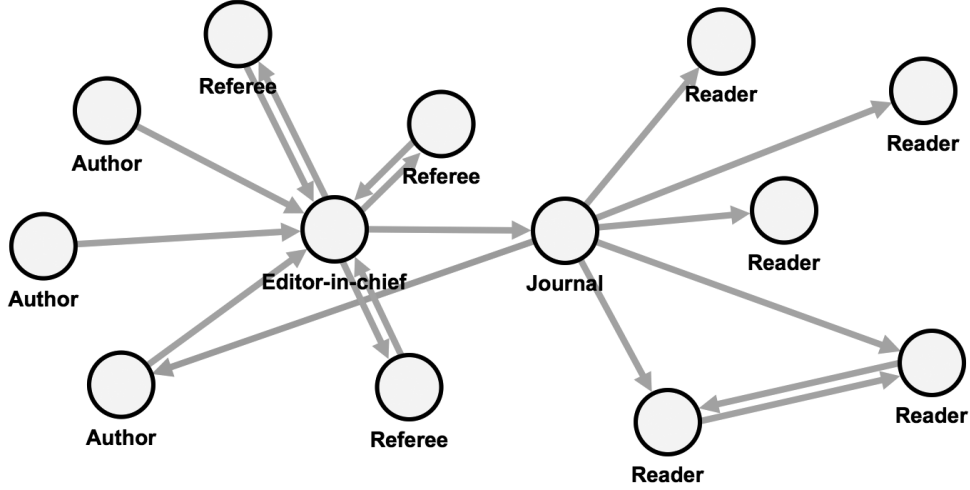


Figure 1: A research network

The links from authors to the editor describe the submission process, the links between the editor and the referees the refereeing process, and the links from the journal to the other inquirers their individual practices of reading that journal.

Like inquirers, links have an internal structure. While links do not have anything corresponding to inquiry accuracy, since the messages they transmit are chosen by the senders and thus not subject to random variation, they have degrees of activity. Two inquirers who seldom talk have weak links between them, while researchers who sit in the same room may be assumed to have stronger links. Analogously to the inquiry notation we define

$$\begin{aligned}
 S_{\beta\alpha}^t(p) &\equiv_{df.} \beta \text{ says that } p \text{ to } \alpha \text{ at } t, \\
 S_{\beta\alpha}^t(\neg p) &\equiv_{df.} \beta \text{ says that not-}p \text{ to } \alpha \text{ at } t, \\
 S_{\beta\alpha}^t &\equiv_{df.} \beta \text{ says either that } p \text{ or that not-}p \text{ to } \alpha \text{ at } t.
 \end{aligned}$$

The strength of a link  $\beta\alpha$ , called its *communication chance*, is representable as a probability  $P(S_{\beta\alpha}^t)$ , being the chance that  $\beta$  communicates that  $p$  or that not- $p$  to  $\alpha$ , at any given moment  $t$ , given that she is certain enough that  $p$  (or not- $p$ ) is indeed the case.

Given that  $\beta$  does communicate with  $\alpha$ , what does she say? Prima facie, it seems that  $\beta$  somehow would indicate how strongly she believes that  $p$ , i.e. she would communicate her credence. But this credence is not in general available to her. As a *subjective degree of belief*, it is a product of betting behavior, and it takes experiments to determine it rather than introspection. We may of course represent  $\beta$ 's beliefs about her own degree of belief as credences as well, but this only pushes the problem further back: somehow, we will still need to determine exactly what to say, given a credence distribution over possible degrees of belief.

Instead, we will adopt a simplified model, which still has enough complexity to be able to represent the important aspects of the situation. We assume that unless she is out to mislead,  $\beta$  will say  $p$  if she believes  $p$ , and not- $p$  if she believes not- $p$ . But belief is not an all-or-nothing matter in the Bayesian tradition, so how sure must she be of  $p$  to say it? This question is answered by a property of the link  $\beta\alpha$  that we will call its *certainty threshold*: a value  $\theta_{\beta\alpha}$  between 0.5 and 1 such that  $p$  can be communicated over  $\beta\alpha$  only if  $C_{\alpha}^t(p) > \theta_{\beta\alpha}$ , and  $\neg p$  can be communicated over  $\beta\alpha$  only if  $C_{\alpha}^t(p) < 1 - \theta_{\beta\alpha}$ .

This combines with the communication chance so that the actual probability of communicating becomes  $P(S_{\beta\alpha}^t)$  if  $\theta_{\beta\alpha}$  is passed, and 0 otherwise.

Are there communicative practices not representable as links of this kind? There is the possibility that  $\beta$  is a systematic liar, but for our current application we will disregard that. We have also already mentioned the scientific journal, which may be more reasonable to model as an inquirer. But what of mass broadcasting systems, like a TV station, or the above-mentioned blog? A link, as we have defined it, has a single source and a single recipient. However, nothing stops us from representing a broadcasting system by a set of links — one for each viewer. Since  $P(S_{\beta\alpha}^t)$  is the probability that  $\alpha$  actually hears what  $\beta$  says, rather than just the probability that  $\beta$  says something, even a broadcasting system ought to have different links for different viewers.

#### 4. Trusting oneself and others

We have described how the participants in a research network  $N$  engage in inquiry and communicate, but we have as yet said nothing about how they react to the results of these practices. It seems that, in general, hearing that  $p$  from someone, or receiving a result indicating  $p$  from inquiry, should influence an inquirer's credence in  $p$ . Fortunately, Bayesianism has a universal answer to the question of how this should be done: belief update proceeds through conditionalization.

But conditionalizing on  $p$  whenever one hears that  $p$  is not reasonable. Straight conditionalization really works only for infallible sources, and in general an inquirer only takes messages proclaiming  $p$  as indications of its truth, and not as conclusive verification. One solution is to use Jeffrey conditionalization instead of regular conditionalization (Jeffrey 1983). This is more of a promissory note than a solution as such, since we still need to decide on a degree of belief to set  $p$  to when one hears that  $p$ . Furthermore, inquirers can receive several messages at the same time, and some of these may even contradict one another, so what we need is a framework that allows us to handle such cases.

One way forward is as follows. To start with, it seems admissible for  $\alpha$  to treat both her own inquiry and the things said to her in roughly the same way: as indications of whether or not  $p$  is the case. We refer to  $\alpha$ 's inquiry  $\iota$  and the other inquirers  $\beta, \gamma, \dots$  who can talk to her as her *sources*. Each of these may have different connections to the truth, and we can represent these by the probability of a source to give the right answer. More specifically, we define *source  $\sigma$ 's reliability for  $\alpha$*  as

$$r_{\sigma\alpha} =_{df}. P(S_{\sigma\alpha}(p)|S_{\sigma\alpha}, p) = P(S_{\sigma\alpha}(\neg p)|S_{\sigma\alpha}, \neg p).$$

This definition presupposes that the probability that any source gives the answer  $p$  if  $p$  is the case is equal to the probability that it gives the answer not- $p$  if not- $p$  is the case. This *source symmetry* simplifies our calculations, although it can be relaxed if we encounter cases where it does not provide a reasonable approximation.<sup>3</sup> As with other

---

<sup>3</sup> There is another way to define reliability, namely, as the probability that  $p$  is the case, given that  $p$  is reported. The two concepts are, however, interdefinable through the equation

$$P_{\alpha}^t(p | S_{\alpha\sigma}(p)) = P_{\alpha}^t(S_{\alpha\sigma}(p) | S_{\alpha\sigma}, p) \frac{P_{\alpha}^t(S_{\alpha\sigma}(p) | S_{\alpha\sigma}, p)}{P_{\alpha}^t(p | S_{\alpha\sigma}, p)}$$

invocations of  $P$ , it is worth remarking that  $r_{\sigma\alpha}$  is concerned with objective chances rather than credences, although inquirers will have credences in propositions such as  $a < r_{\sigma\alpha} < b$  which are *about* such chances. In fact,  $P$  will only appear in such credences, so in this case we do not need to worry about the interpretation of objective chance at all.

It is quickly checked that the reliability of  $\alpha$ 's inquiry is identical to her inquiry accuracy. For other sources, it is an abstraction based on those sources' performances as indications of truth. In general, an inquirer has no direct access to this value, but this does not stop her from forming beliefs about it. Since the number of possible values for the chance  $r_{\sigma\alpha}$  is infinite, we need to represent  $\alpha$ 's credence as a density function instead of a regular probability distribution. Thus, for each inquirer  $\alpha$ , each source  $\sigma$ , and each time  $t$ , we define a function  $\tau_{\sigma\alpha}^t: [0, 1] \rightarrow \mathbb{R}$  which we call  $\alpha$ 's *trust function* for  $\sigma$  at  $t$ , such that

$$C_{\alpha}^t(a < r_{\sigma\alpha} < b) = \int_a^b \tau_{\sigma\alpha}^t(x) dx$$

for  $a, b$  in  $[0, 1]$ . This function is uniquely defined up to a set of measure 0, according to the Radon-Nikodym theorem.  $\tau_{\sigma\alpha}^t(x)$  gives the credence density at  $x$  and we can obtain the actual credence that  $\alpha$  has in propositions about the reliability of her sources by integrating this function. We will also frequently use the expression  $1 - \tau_{\sigma\alpha}^t(x)$ , which represents  $\alpha$ 's credence density for propositions about  $\sigma$  *not* being reliable, and we will refer to this function as  $\overline{\tau_{\sigma\alpha}^t}(x)$ . More generally, we will use the bar  $\bar{x}$  as a shorthand for the expression  $1 - x$ .

Now, it is obvious that an inquirer's credences about chances should influence her credences about the outcomes of these chances. The way this should be done is generally known under a name David Lewis gave to it: the principal principle.<sup>4</sup> This says that if  $\alpha$  knows that the chance that an event  $e$  will happen is  $r$ , then her credence in  $e$  should be exactly  $r$ . Applied to our case, this means that the following dual principle (*PP*) must hold:

$$\begin{aligned} a < C_{\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, p, a < r_{\sigma\alpha} < b) < b \\ a < C_{\alpha}^t(S_{\sigma\alpha}^t(\neg p) | S_{\sigma\alpha}^t, \neg p, a < r_{\sigma\alpha} < b) < b \end{aligned}$$

Hence,  $\alpha$ 's credence in  $\sigma$  giving the report  $p$ , given that she knows that the source gives any report at all, that  $p$  is actually the case, and that  $\sigma$ 's reliability is between  $a$  and  $b$ , should itself be between  $a$  and  $b$ , and similarly for  $\neg p$ . Since  $r_{\sigma\alpha}$  can take a continuum of values, we have framed the principle in terms of intervals rather than specific values. A quantity such as  $C_{\alpha}^t(r_{\sigma\alpha} = x)$ , on the other hand, would be 0 for almost all values of  $x$ .<sup>5</sup>

---

As the definition used here is simpler for our purposes, and as a referee pointed out,  $P_{\alpha}^t(p | S_{\sigma\alpha}(p))$  may be difficult to interpret given that  $p$  is simply true or false, we have adopted this version. The difference corresponds to Goldman's (1986) distinction between *reliability* and *power*.

<sup>4</sup> Lewis (1980). An earlier version is given by Miller (1966). Although there has lately been a large debate on when these principles are admissible, we will only be using two specific instances of the general case. None of these appear to give rise to any problems.

<sup>5</sup> We are indebted to an anonymous referee for pressing us on the exact formulation of our version of the principal principle in terms of intervals rather than point values.

We also have use for an independence postulate. While not strictly necessary, such a postulate will simplify calculations and modeling considerably. The independence assumption we use here will be referred to as *communication independence (CI)*:

$$C_{\alpha}^t(S_{\sigma\alpha}^t, p, a < r_{\sigma\alpha} < b) = C_{\alpha}^t(S_{\sigma\alpha}^t) \times C_{\alpha}^t(p) \times C_{\alpha}^t(a < r_{\sigma\alpha} < b)$$

Communication independence implies that whether  $\sigma$  says anything is independent of whether or not  $p$  actually is true, as well as of what reliability  $\sigma$  has. This is true in the current model, since we have assumed that a source's reliability for reporting that  $p$ , given that  $p$  is the case, is the same as its reliability for reporting that  $\neg p$  is the case, given that  $\neg p$ .

From *PP* and *CI* we can derive the following expression for  $\alpha$ 's credence in  $\sigma$ 's reliability (see the appendix for the actual derivation):

$$C_{\alpha}^t(S_{\sigma\alpha}^t(p)|p) = C_{\alpha}^t(S_{\sigma\alpha}^t) \times E[\tau_{\sigma\alpha}^t] \quad (1)$$

Here,  $E[\tau_{\sigma\alpha}^t]$  is the expectation of the trust function  $\tau_{\sigma\alpha}^t$ . We refer to this as the link  $\sigma\alpha$ 's *expected trust*. Using equation (1), an application of Bayes' theorem together with the law of total probability gives us

$$C_{\alpha}^t(p|S_{\sigma\alpha}^t(p)) = \frac{C_{\alpha}^t(p) \times E[\tau_{\sigma\alpha}^t]}{C_{\alpha}^t(p) \times E[\tau_{\sigma\alpha}^t] + C_{\alpha}^t(\neg p) \times E[\overline{\tau_{\sigma\alpha}^t}]}$$

$$C_{\alpha}^t(p|S_{\sigma\alpha}^t(\neg p)) = \frac{C_{\alpha}^t(p) \times E[\overline{\tau_{\sigma\alpha}^t}]}{C_{\alpha}^t(p) \times E[\overline{\tau_{\sigma\alpha}^t}] + C_{\alpha}^t(\neg p) \times E[\tau_{\sigma\alpha}^t]}$$

Since we, by the requirement of conditionalization, must have that  $C_{\alpha}^{t+1}(p) = C_{\alpha}^t(p|S_{\sigma\alpha}^t(p))$  whenever  $\sigma$  is the only source giving information to  $\alpha$  at  $t$ , and that information consists in the message  $p$ , these formulae completely determine how  $\alpha$  should update her credences in such a case. Some of the consequences of this can be summarized qualitatively. We say that  $\sigma$  is *trusted* when  $E[\tau_{\sigma\alpha}^t] > 0.5$ , *distrusted* when  $E[\tau_{\sigma\alpha}^t] < 0.5$  and neither trusted nor distrusted when  $E[\tau_{\sigma\alpha}^t] = 0.5$ . Furthermore, we say that a message  $m$  is *plausible* if  $C_{\alpha}^t(p) > 0.5$  and  $m \equiv p$  or  $C_{\alpha}^t(p) < 0.5$  and  $m \equiv \neg p$ , *implausible* if  $C_{\alpha}^t(p) < 0.5$  and  $m \equiv p$  or  $C_{\alpha}^t(p) > 0.5$  and  $m \equiv \neg p$ , and neither plausible nor implausible otherwise. Table 1 (see Vallinder and Olsson 2013 for derivations and Collins et al 2015 for some empirical support) gives the qualitative rules for how belief is updated. A "+" means that the message reinforces  $\alpha$ 's current belief (i.e. her credence increases if above 0.5 and decreases if below 0.5), a "-" that the strength of her belief is weakened (i.e. that her credence increases if below 0.5 and decreases if above 0.5), and "0" that her credence is left unchanged.

| Is source trusted? | Is message plausible? |         |    |
|--------------------|-----------------------|---------|----|
|                    | Yes                   | Neither | No |
| Yes                | +                     | +       | -  |
| Neither            | 0                     | 0       | 0  |
| No                 | -                     | -       | +  |

Table 1: Effects of trust on credence

The calculations become slightly more complex when we take into account the possibility of receiving several messages at the same time. Let  $\Sigma_\alpha^t$  be the set of sources from which  $\alpha$  receives information at  $t$ , and let  $m_{\sigma\alpha}^t$  be the message (i.e. either  $p$  or not- $p$ ) that  $\sigma$  gives  $\alpha$  at  $t$ . Conditionalization requires that

$$C_\alpha^{t+1}(p) = C_\alpha^t \left( p \left| \bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) \right. \right)$$

This may be a rather complex expression, but we can simplify greatly it if we take inquirers to treat their sources as independent given the truth or falsity of  $p$ . Formally, this means that we adopt the following axiom, which we call *source independence (SI)*:

$$C_\alpha^t \left( \bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) \left| p \right. \right) = \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | p)$$

Is this a valid assumption to make? In certain cases, it may very well be, at least as an approximation. We generally tend conspiracy or other forms of dependence in the absence of concrete evidence to the contrary. It is this default assumption of independence that makes us attach greater degree of belief to information coming from different sources. Although the assumption relies on a simplification — there may be all kinds of hidden dependencies of which we are unaware — acting on the basis of it may be the best thing to do all things considered.

We should note that communication dependence does not contradict the fact that the credences of inquirers indeed do become probabilistically dependent as a result of their interactions with one another. This is because *SI* applies to credences rather than the probabilities for having these credences. The result is, however, that as the simulation proceeds, the assumption of source independence implies that the inquirers themselves will become less and less informed about the current structure of the network. As noted in the previous paragraph, this is a simplification. However, it is one that is necessary in order to limit the information needed for these to store to manageable levels.

Given source independence, the properties of individual links and inquirers determine how inquirers should update their beliefs when they receive new information. An application of Bayes' theorem together with the theorem of total probability gives that

$$C_\alpha^t \left( \bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) \left| p \right. \right) = \frac{C_\alpha^t(p) \times \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | p)}{C_\alpha^t(p) \times \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | p) + C_\alpha^t(\neg p) \times \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | \neg p)}$$

Since the values  $C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | p)$  and  $C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | \neg p)$  are determined by eq. (1), this equation lets us infer exactly what degree of belief an inquirer  $\alpha$  should give  $p$ , given the information she receives at any time.

However,  $\alpha$ 's credence in  $p$  is not the only part of her epistemic state that needs to be updated in light of new evidence. Equally important is for  $\alpha$  to keep track of how much to trust her sources. A source that generally gives very implausible reports is unlikely to be veridical, and an inquirer should adjust her trust function in light of this. It turns out that our model already determines how to do this: given that  $\alpha$ 's trust function for the

source  $\sigma$  at  $t$  is  $\tau_{\sigma\alpha}^t$ , and that she receives the result that  $p$  or that not- $p$  from  $\sigma$  then, her new trust function  $\tau_{\sigma\alpha}^{t+1}$  must be given either by

$$\tau_{\sigma\alpha}^{t+1}(r) = \tau_{\sigma\alpha}^t(r) \times \frac{r \times C_{\alpha}^t(p) + \bar{r} \times C_{\alpha}^t(\neg p)}{E[\tau_{\sigma\alpha}^t] \times C_{\alpha}^t(p) + E[\bar{\tau}_{\sigma\alpha}^t] \times C_{\alpha}^t(\neg p)}$$

or by

$$\tau_{\sigma\alpha}^{t+1}(r) = \tau_{\sigma\alpha}^t(r) \times \frac{r \times C_{\alpha}^t(\neg p) + \bar{r} \times C_{\alpha}^t(p)}{E[\tau_{\sigma\alpha}^t] \times C_{\alpha}^t(\neg p) + E[\bar{\tau}_{\sigma\alpha}^t] \times C_{\alpha}^t(p)}$$

depending on whether the report received from  $\sigma$  claims that  $p$  or that not- $p$ .<sup>6</sup>

That both  $C_{\alpha}^t(p)$  and  $\tau_{\sigma\alpha}^t$  change as a result of inquiry and communication gives rise to complex interactions. Assume, for example, that  $\alpha$  starts out with a credence in  $p$  slightly below 0.14, and some reasonable degree of trust in her own inquiry, indicated by a value of expected trust at 0.67. Assume that her inquiry keeps indicating that  $p$  actually is the case. How her trust in inquiry changes after  $t$  such inputs then depends crucially on the trust function  $\tau$  she starts out with. But even if the *expected* trust is the same, the shape of the trust function influences the long-term behavior. Figure 2 illustrates three such scenarios during a time lapse of 50 inputs. All the scenarios start out with trust functions having the form of beta distributions with different values of  $\alpha$  and  $\beta$ , although with the same expectation.

---

<sup>6</sup> Again, see the appendix for the derivation.

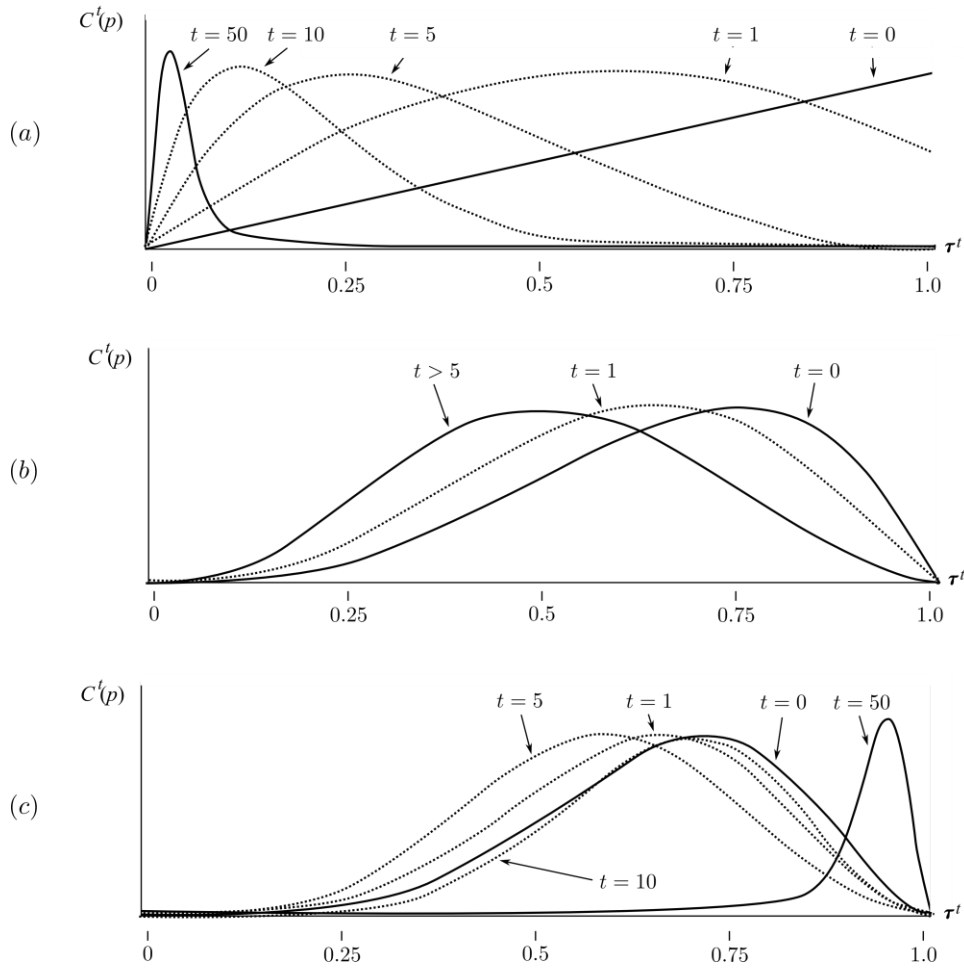


Figure 2: results of iterated trust updates

Here we have three different ways that the inquirer's credence might evolve given the same initial credence, the same obtained evidence, and the same initial expected trust. Although our inquirer happens to be a perfect inquirer insofar as her inquiry always gives the right result, the fairly low stability of her faith in inquiry in (a), together with her prior judgment that  $p$  is unlikely, conspire to make her distrust her own inquiry. This, in turn, gives rise to a vicious circle in which she becomes more and more convinced that  $p$  is false, and that her inquiry is negatively correlated with the truth.

In (b), the inquirer's trust happens to be just enough to counter her prior disbelief in  $p$ , but not enough to get her to believe that  $p$ , with the result that her credence converges to 0.5. In (c), her trust in inquiry is stable enough to overcome her prior disbelief, with the happy consequence that she converges on the truth.

What can we learn from these scenarios? The most important lesson is that trust is a complex issue, impossible to capture in a single number. Although we mentioned 'stability' in the preceding paragraphs, it should be noted that two numbers are still insufficient, so stability should not be seen as a complement to (expected) trust. Although only the expected trust influences an inquirer's posterior credence for a single update, this does not hold for further updates. For arbitrarily iterated updating, all of  $\tau$  is necessary.

## 5. Setting up the experiment

We have given a quick overview of the model we will be using to describe a research network. This model has been implemented in the simulation environment *Laputa*, which allows us to run simulations of such networks and collect their statistics.

For the present investigation, two of the most important properties are the *density* of the network, and the certainty threshold. The density of a graph (and a research network) is the average number of edges (links) per ordered pair of inquirers, excluding pairs with the same first and second element. In a network with  $n$  inquirers and  $m$  links, the density will therefore be

$$\delta = \frac{m}{n \times (n - 1)}$$

*Laputa* can be set to generate networks with a specified density but with different characteristics. The most straightforward method of generating a network of density  $\delta$  is to simply go through each ordered pair of inquirers and add a link between the members of each pair with probability  $\delta$ . Networks obtained this way are called Erdős-Renyi graphs, and have been studied intensely. The edge distribution of an Erdős-Renyi graph is binomial, i.e. plotting a histogram of the number of edges connecting a node gives rise to a binomial distribution.

However, real networks of scientific collaborations, as well as networks of citations, tend to have a different structure, which shows up in the fact that the tails of their edge distributions conform to a power law instead of a binomial distribution.<sup>7</sup> One reason for this is the “rich get richer” mechanism, which makes nodes with more edges proportionally more likely to attract new edges than nodes with less. In terms of citations, papers that are cited more tend to be read more, and thus again be cited more. *Laputa* allows taking this effect into account, and we have done so by assigning inquirers with  $n$  links going out of them with weight  $n + 1$ , rather than 1, which would have been the case in the Erdős-Renyi model.<sup>8</sup> This approximates the model of Barábasi and Albert (1999).

Together with the communication chance of each link, the density  $\delta$  gives a good indication of how much communication is going on in a research network. Since simulations with less links are faster to run than simulations with more links, but lower communication chance, we have opted to vary  $\delta$ , and keep the communication chance random.

When  $\delta = 0$ , no communication at all occurs, and when  $\delta = 1$ , every inquirer can communicate directly with everyone else. However, this does not mean that they actually *do* communicate, and one of the variables that controls whether this happens is the certainty threshold.  $\theta_{\sigma\alpha}$  is part of what may be called the network’s *norms of assertion*: it poses a requirement on when  $p$  can be communicated in terms of the source’s epistemic state (Olsson and Vallinder 2013). Investigating such norms of assertion is one of the objectives of the present study. For simplicity, we assume that all links share a common certainty threshold  $\theta$ .

The other variables of the model are among those we are not currently interested in. Dealing with them properly requires us to randomize them so as to obtain a *randomized experiment*, which will allow us to draw conclusions about the *causal* effect of  $\delta$  and  $\theta$  on

---

<sup>7</sup> See, for example, Newman (2001) and Redner (1998).

<sup>8</sup> This is a novelty in *Laputa* v1.5 of December 2013. Earlier versions lacked this feature, and allowed the generation of Erdős-Renyi networks only.

a network's performance. The specific random distributions that the other variables are picked from should be dependent on the model's intended applications.

We pick the starting degree of belief for each inquirer from a normal distribution with mean 0.5 and standard deviation 0.15. This corresponds to the intended applications being research networks in which the starting degrees of belief are somewhat clustered around 0.5: like good Bayesians, our researchers do not usually start out with strong beliefs about whether or not  $p$  is the case.

Both the inquiry chance and the link communication chance are uniformly distributed in  $[0, 1]$ . This means that we take each degree of activity, of inquirers *qua* researchers of links *qua* communication channels, to be equally likely.

We will study two possible distributions of inquiry accuracy: in the first, it will be normal distributed with mean 0.55 and standard deviation 0.15. This gives inquirers that are, on average, slightly better than chance. The second distribution is also a normal distribution with the same variance, but with mean 0.75, which is intended to describe research networks in which inquirers tend to be highly competent.

Both inquiry trust and the trust functions of each link are set to normal distributions with means randomly picked between 0.5 and 1.0, and standard deviations randomly picked between 1 and 0.15. Hence inquirers will generally start out trusting their inquiry to some degree, and also trusting other inquirers to some degree. This represents a basic form of the principle of charity, which is a reasonable assumption to make when modeling real scientific communities.

To begin with, we are mainly interested in the end result of inquiry, so we let each randomized research group run for 25 steps, which is sufficient for inquirers' opinions to stabilize. Each research group is taken to have 25 members, or, alternatively, we can see it as having from 1 to 25 members, with each number of members weighted by the number of possible graphs of that order, which gives us a uniform distribution over all graphs of order  $\leq 25$ . Since the number of directed graphs of order  $n + 1$  is  $2^{2^n}$  times that of order  $n$ , the difference between considering only groups of 25 inquirers and groups of 1 to 25 inquirers is negligible.<sup>9</sup>

This concludes our specification of the parameters of the model. What we still have to describe is the way we wish to measure a network's performance. In his (1999), Goldman introduces the notion of *veritistic value*, or *V-value*, of a social practice. For the single-proposition case, which is what we are working with, this is calculated as the network's average increase in the credence in  $p$ , where  $p$  is taken to be true. The proposition  $p$  will be assumed true in all the following simulations.

In our case, a *social practice* can be identified with a way of constraining possible evolutions of a research group as it engages in inquiry (Olsson, 2011). But since such evolutions are determined, at least probabilistically, by their initial states, we can take a social practice to be so determined as well. Thus each setup of the model we have (or in Laputa terms, each batch simulation) gives a social practice in this extended sense. We denote the V-value of a network of inquirers  $N$  at a given time  $V^t(N)$ .

V-value gives a straight mean of the influence of a practice, but, just as the GDP per capita of a country says nothing about its distribution of wealth, V-value says nothing about the distribution of knowledge. It is therefore useful to also plot the *polarization* of

---

<sup>9</sup> Sampling uniformly from *isomorphism classes* of graphs is a significantly harder problem. However, we still have that there are about  $6.7 \times 10^6$  more isomorphism classes of graphs with 25 vertices than with 24, and  $2.3 \times 10^{11}$  more than with 23 (see <http://oeis.org/A000088>), so just picking from graphs with 25 vertices still gives a very good approximation. See Masterton (2014) for an extensive investigation into the graph isomorphism issue as it applies to Laputa and related frameworks.

belief that follows from a given practice, by which we mean the root mean square of the deviation of individual credence from the mean, i.e.

$$Pol^t(N) = \sqrt{\frac{1}{|N|} \sum_{\alpha \in N} (C_{\alpha}^t(p) - \mu^t)^2}$$

where  $\mu^t = \frac{1}{|N|} \sum_{\alpha \in N} C_{\alpha}^t(p)$ . A network in which every inquirer has the same credence in  $p$  will thus have minimum polarization 0, while one in which half the inquirers are certain that  $p$  and half are certain that not- $p$  will have maximum polarization 0.5.

A different limitation of V-value is its relative irrelevance for an important type of group decision making: from the V-value of a practice, little can be inferred about the result of a majority vote among the members of a group that have just implemented that practice. We will therefore also measure the *average belief of the majority*, which we define as the average over all simulation runs of the value

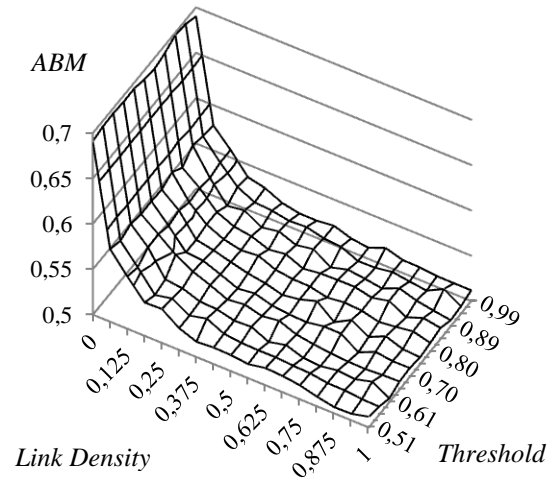
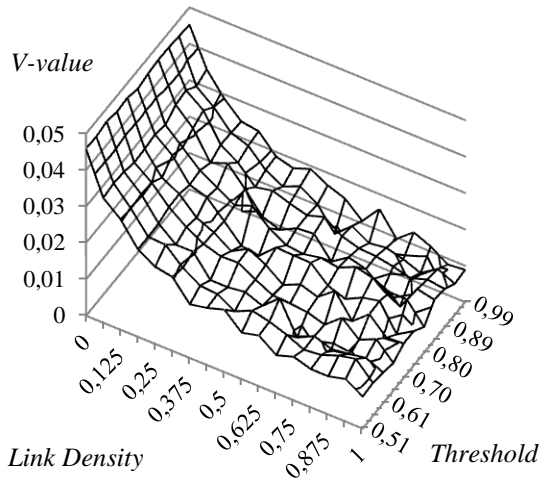
$$BM^t(N) = \begin{cases} 1 & \text{if } |\{\alpha \in N \mid V^t(\alpha) > 0.5\}| > |\{\alpha \in N \mid V^t(\alpha) < 0.5\}| \\ 0 & \text{if } |\{\alpha \in N \mid V^t(\alpha) > 0.5\}| < |\{\alpha \in N \mid V^t(\alpha) < 0.5\}| \end{cases}$$

The average belief of the majority (or ABM) that we are interested in is, like V-value, that which is taken by  $N$  in its *final* states, i.e. after the simulation has run its course of 25 iterations.

## 6. The perils of unrestricted communication

Running the simulation for 25,000 iterations in order to achieve more numerically stable results, using the parameters described in the previous section, produced the following results.

Mean Accuracy 0.55



Mean Accuracy 0.75

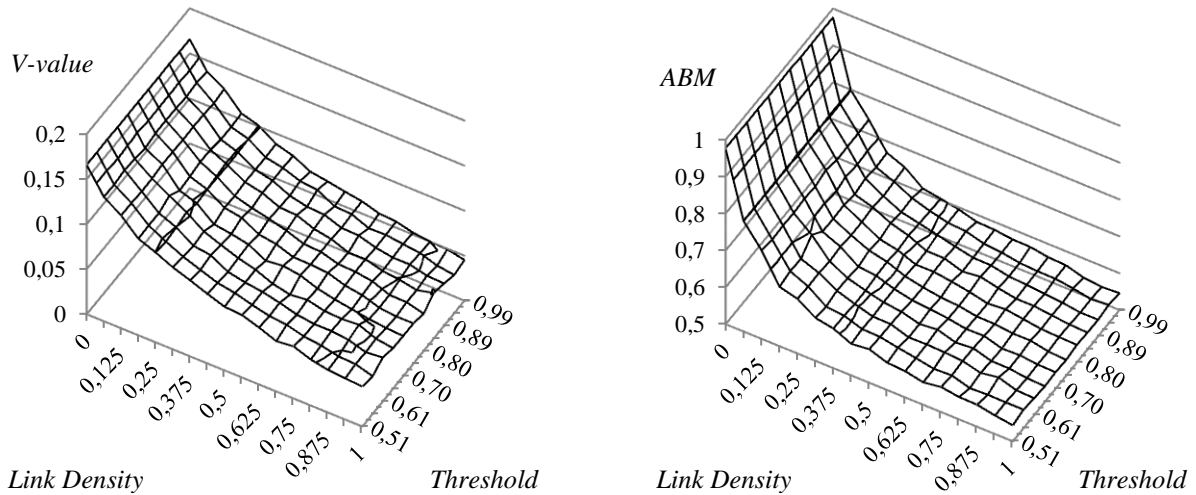


Figure 3: Results of unrestricted communication

We can see that, in all cases, *more communication is bad*, both in terms of V-value and ABM. The threshold does not affect either V-value or ABM, and while the inquirers' accuracy does, it does not make any amount of communication preferable to none at all. Whether inquirer as slightly or highly accurate, no communication at all gives a V-value approximately 5 times higher than full communication.

Note that this does not occur because inquirers in highly dense networks engage in communication *instead of* inquiry: in Laputa, whether a link is used during a time step is independent of whether the source also engages in inquiry during that time. As we mentioned in the introduction, we thus have a puzzle: scientific communication is usually taken for granted to be a good thing, but we have just found an indication that it is not. Why? What is so bad about communication that it makes inquirers less likely to converge on full credence in the truth?

Let us first ask why one might think that groups should perform better than single inquirers in the first place. One reason is the famous Condorcet jury theorem, which states that the majority vote is more reliable than the vote of any given voter under favorable circumstances. This theorem applies to majority belief as a kind of collective "vote", but it does not have any obvious implications for V-value. However, even with regard to majority belief, the theorem has an important presupposition: the voters in question have to be *independent*. If they are able to directly influence one another, the theorem no longer applies. Now scientific communication is a prime example of such influence, and we propose that the downside to communication that makes the results of figure 2 possible is that communication can, and typically does, *reduce independence*. This is particularly problematic when the initial credences are distributed around 0.5 as we have, realistically, assumed them to be. The denser the network, the more sensitive it

will be to misplaced initial credences, i.e. initial credences that lie slightly below 0.5. If the majority starts off on the wrong track, chances are that it will drag down the whole network. For inquirers will then receive a strong social signal that not- $p$  is true when in fact we have assumed  $p$  to be true and hence not- $p$  to be false. This in turn may lead inquirers to downgrade their trust in their own inquiry to the point where they consider it anti-reliable, preventing self-correction through inquiry from taking place. We would expect the probability of this unfortunate development to increase with the network density.

There is an upside of sorts to this, though. The following diagram shows the degree of polarization as a result of the previous practices. Since the numbers, like those of the figure 2, are independent of the threshold, we have plotted the change in polarizations of both values of mean accuracy ( $m$ ) as lines, rather than as surfaces.

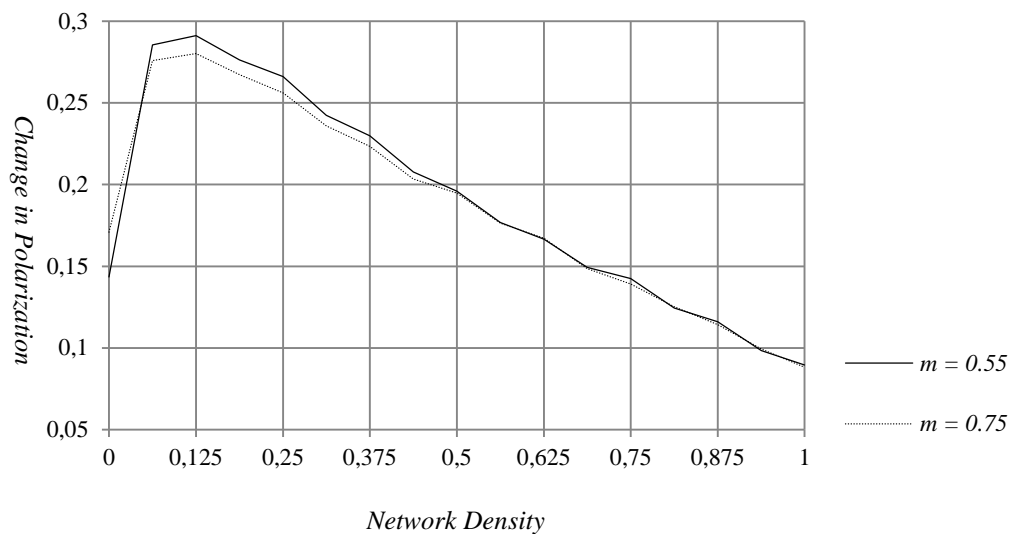


Figure 4: Change in polarization as a result of the practices in figure 2

As seen in the diagram, a degree of communication over zero but below about 0.65 increases the polarization of belief in a research network. On the other hand, high network density decreases it, so while a large amount of communication decreases the V-value, it has an egalitarian aspect: it makes differences of opinion on  $p$  less pronounced. Whether this should count as an *epistemic* advantage may be left unsaid; presumably it can be a practical one.

It is worth pointing out that since the change in polarization is above zero for *all* density values, the process of inquiry generally has the effect of diminishing differences of opinion. The reason for this phenomenon is our handling of trust, which not only models how trust affects belief, but also how belief affects trust. In the typical case, inquirers will start out with beliefs around 0.5, but by chance some get results that not- $p$  from inquiry, either because they have a low accuracy or because of bad luck. Since they typically trust their own inquiry somewhat, this will make them start to believe that not- $p$  is the case. If they happen to receive strong enough confirmation that not- $p$ , messages that  $p$  from others will be taken as indications of untrustworthiness, and will thus only further reinforce the conviction that not- $p$ .

These conclusions may seem somewhat pessimistic, but they follow from a Bayesian approach together with the axioms we have imposed, and the modeling of trust used.

Unlike traditional Bayesians, who rely on the Bernstein-von Mises theorem or some other convergence result in order to show that convergence has to occur, we recognize that prior probabilities influence not only our beliefs about the facts, but also our interpretation of the evidence, i.e. the perceived reliability of our inquiries. This means that we can, and frequently will, get divergences of opinion even in the long run, and that this is not due to irrationality but rather exactly what rationality prescribes.

As we saw, increasing the certainty threshold does not increase either V-value or ABM. But are there other restrictions to communication that could make it more effective? At present, whether a link can be used depends only on chance and the threshold, but this means that inquirers can *spam* the network by having many outgoing links with high communication chance. Such inquirers will transmit their opinion repeatedly. If they start out with initial credences that tend toward not- $p$  rather than  $p$ , which, as we saw, is a serious possibility given our assumptions, there is a fair chance that the false information is repeatedly transmitted in the network. And, as we also noted, the impact of massively false social information on a given inquirer may very well be that she starts distrusting her own inquiry, i.e. starts to regard it as anti-reliable, in which case the probability of self-correction through inquiry is severely reduced.

Stopping this form of recycled information would require asking inquirers not to transmit their opinion across a link unless they have actually received new supporting information since that link was last used. We will consider two conditions along this line:

*New Evidence (NE)*: To use link  $\beta\alpha$  to transmit  $p$  (not- $p$ ) one of the following must be satisfied:

- a)  $\beta\alpha$  has not been used before, so that the information transmitted is based on  $\beta$ 's prior credence in  $p$  (not- $p$ ).
- b) Since the last time  $\beta\alpha$  was used  $\beta$  has received a message that  $p$  (not- $p$ ) from some inquirer other than  $\alpha$ , which has resulted in her increasing her credence in  $p$  (not- $p$ ).
- c) Since the last time  $\beta\alpha$  was used  $\beta$  has received a result that  $p$  (not- $p$ ) from her own inquiry, which has resulted in her increasing her credence in  $p$  (not- $p$ ).

*New Inquiry (NI)*: To use link  $\beta\alpha$  to transmit  $p$  (not- $p$ ), condition c) of *NE* must be satisfied.

Imposing *NE* or *NI* means that an inquirer cannot retransmit information in the network unless she has obtained evidence in support of that information since the last time she transmitted it over the same link. Intuitively, the effect should be to prevent insufficiently supported information from spreading in a massively connected network. This is especially clear in the case of *NI*. To be in a position to retransmit not- $p$  the inquirer would, if *NI* is imposed, have to obtain not- $p$  from inquiry, which, given the assumed reliability of inquiry, is relatively unlikely.

From another perspective, imposing *NE* or *NI* can be seen as a partial fix of one limitation of the model: that the only information transmittable by inquirers in the network is  $p$  or not- $p$ , and no explicit discussion about evidence occurs. With *NE* or *NI* in place, a message that  $p$  can be interpreted as 'there is new evidence that  $p$ ' or 'I made a

new observation of  $p'$ . This gives us a simple way to accommodate evidence in the model.<sup>10</sup>

## 7. Effects of the quality constraints

Imposing  $NE$  or  $NI$  gives the following  $V$ -values, for inquirers with mean accuracy 0.55:<sup>11</sup>

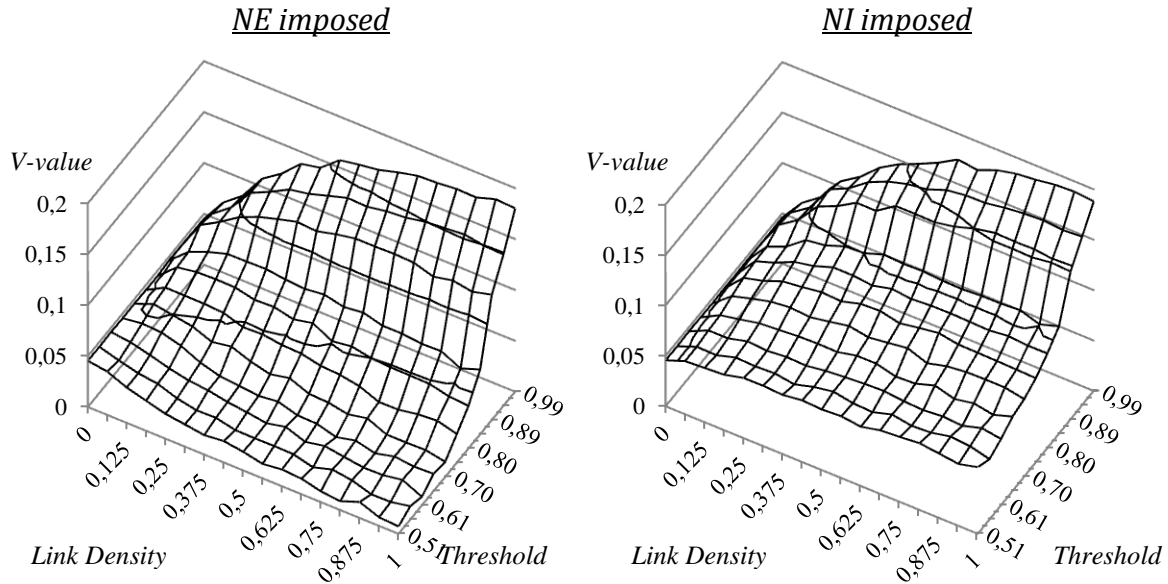


Figure 5: Impact on  $V$ -value of imposing  $NE$  and  $NI$

With the restrictions in place, communication can actually be beneficial in terms of  $V$ -value (figure 5). With a high certainty threshold, the best network density values lie around 0.5 for  $NE$ , and for  $NI$  we get that more communication *never* decreases  $V$ -value. Furthermore, the  $V$ -values for  $NI$  are higher than those for  $NE$ , so among the two, the stricter condition  $NI$  is preferable.

$NE$  and  $NI$  also have interesting consequences for belief polarization. While the threshold had no effect for unrestricted communication, its value is important for the restricted case, as seen in figure 6.

<sup>10</sup> We should, however, recognize that  $NE$  and  $NI$  are only partial fixes, or in programmers' jargon, "hacks". They do not amount to making inquirers actually keep track of particular pieces of evidence. For example,  $NE$  forbids two inquirers from strengthening each other's beliefs by just talking among themselves, but does not do this for three or more inquirers connected in a circle. The problem of extending the model to take evidence into account in a more robust way that excludes such cases as well remains to be solved.

<sup>11</sup> Very similar results were obtained for the case of mean accuracy 0.75, although, as before, with higher  $V$ -values and ABM. To conserve space, we have not reproduced them here.

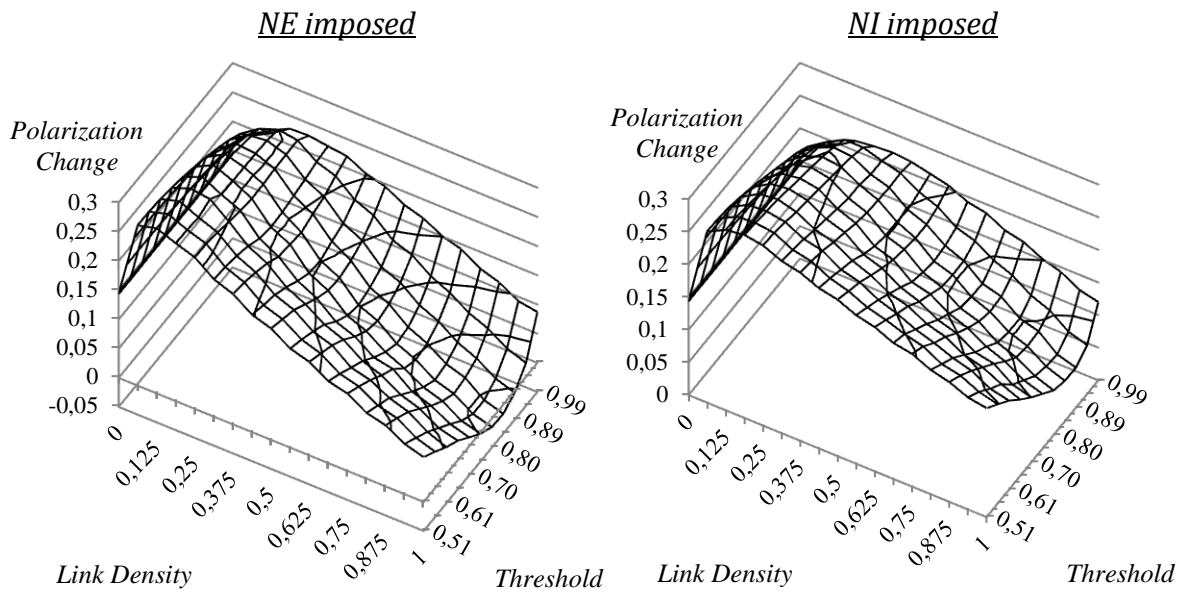


Figure 6: Impact on belief polarization of imposing *NE* and *NI*

We see that, just as for unrestricted communication, a small amount of communication makes the polarization rise, and increasing it further makes it fall again. But the amount it falls depends on the threshold, and using *NE* and setting the threshold at around 0.8 and the density to 1, a minimum of belief polarization is attained.

Taken together, the diagrams for V-value and polarization (figure 5 and 6, respectively) imply a conflict concerning the proper assertion threshold. From the perspective of minimizing polarization, the threshold should be at around 0.80, but V-value is maximized when the threshold is 0.99. There is also a conflict as to what restriction to use: for polarization, *NE* is by far the best, while for V-value, *NI* is somewhat better.

The threshold is also important for the effect of *NE* and *NI* on the average belief of the majority (ABM), as seen in figure 7.

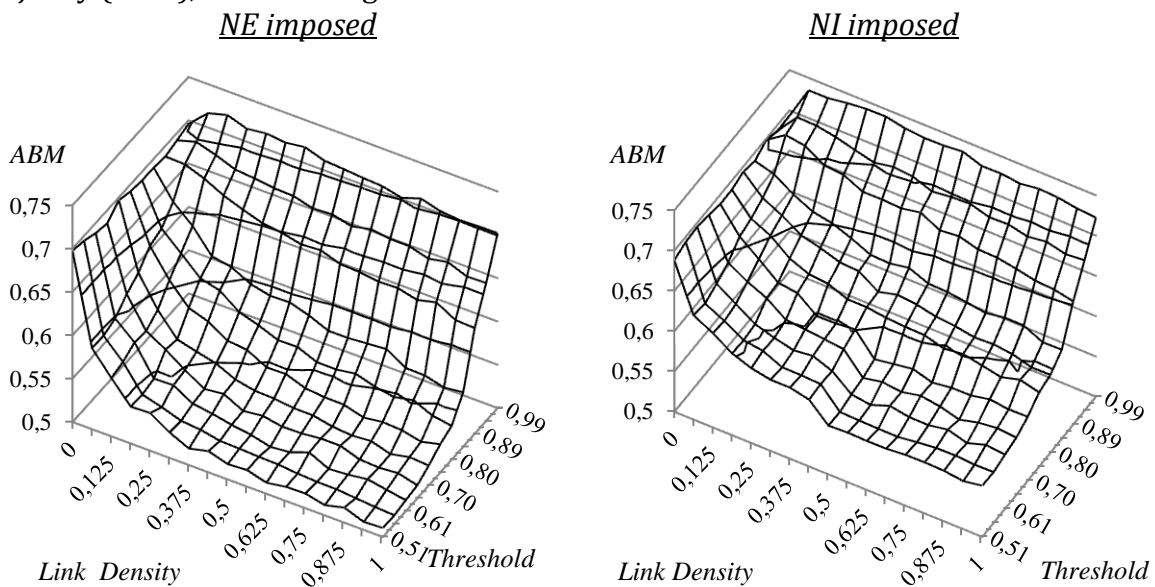


Figure 7: Impact on Average Belief of the Majority of imposing *NE* and *NI*

In terms of ABM, we find that it is still generally the case that the best policy is for researchers not to communicate at all. The only exception is for very high values of the threshold, where a small increase is somewhat beneficial. This is noteworthy because the belief of the majority is arguably more important than V-value in democratically organized groups, where many important decisions are decided through a majority vote. In particular, the results we have obtained seem to count against important assumptions of deliberative democracy, in which group deliberation prior to voting is taken to be beneficial for the result.<sup>12</sup> Our study indicates that, at least epistemically, it would be better if no communication occurred prior to voting.

Olsson and Vallinder (2013) make a related observation about V-value, also using Laputa. As we noted, in the absence of restrictions communication generally decreases the V-value in the long run. But in the short run, Olsson and Vallinder show, communication *can* increase the V-value. While imposing *NE* or *NI* secures a V-value increase even in the long run, Olsson and Vallinder's general observation carries over. The following diagrams (figure 8) show the V-value of network different density values, plotted against the number of simulation steps. The threshold has been set at 0.95.

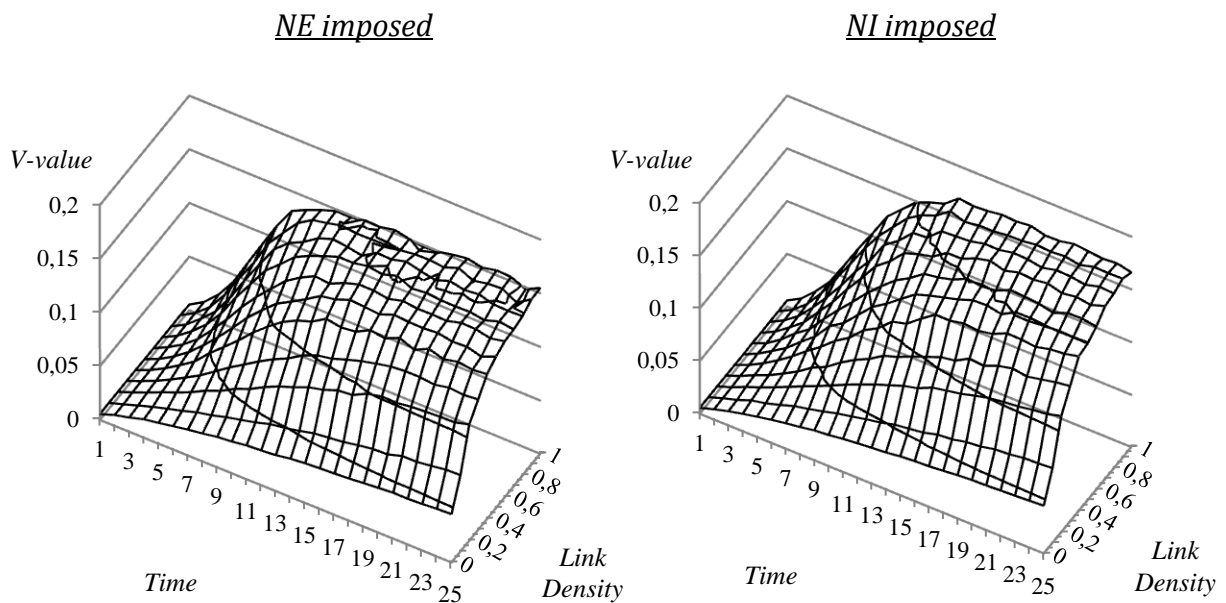


Figure 8: V-values over time

Both with *NE* and *NI*, higher density of the network is clearly beneficial in the V-value sense. In fact, more communication makes the society converge to its final V-value quicker, *as well as* making that V-value higher. *NI* is, on the whole, somewhat better than *NE*, but the differences are fairly small.

Given that shortening the deliberation time makes higher network density lead to higher V-value, it is natural to ask if the same is the case for ABM. Unfortunately, as we can see from the following diagram (figure 9), the effects of lowering the number of simulation steps are rather unimpressive.

<sup>12</sup> Cf. Goodin (2003).

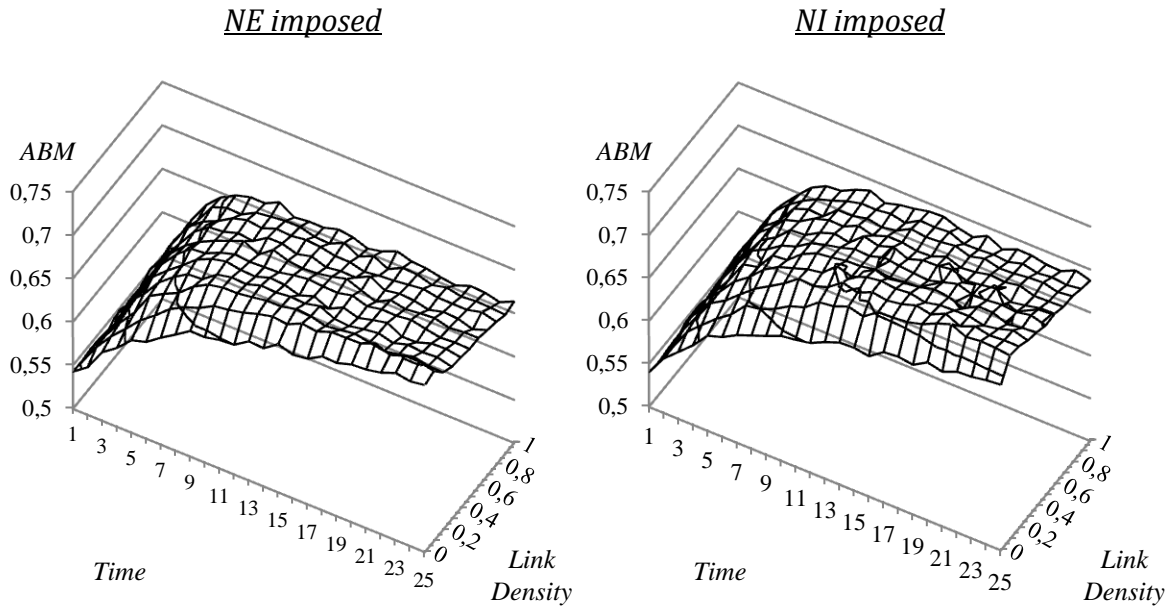


Figure 9: Average belief of the majority over time

As in the long run, there is a small beneficial effect to having some communication rather than none at all, but this effect is actually smaller for short simulations rather than larger. Thus we have as yet no case where a high amount of communication would increase the accuracy of the majority opinion. For a proponent of deliberative democracy, it might be natural to blame Laputa’s simplified model of communication for this effect. Real research networks communicate not only their beliefs, but also the *reasons* for those beliefs. But with the restrictions in place, it is hard to see why this feature should affect the results substantially. Especially if *NI* is adopted, any communication requires new (supporting) results from inquiry, i.e. new reasons (cf. Olsson 2013).

## 8. Discussion and conclusion

As noted in Vallinder and Olsson (2013), the Laputa model can, being Bayesian, be given a normative interpretation. Olsson (2013) is an extended discussion of the normative and descriptive interpretations of Laputa connecting it to the empirically robust Persuasive Argument Theory (PAT) tradition in social psychology and suggests that the former is largely subsumable under the latter. In addition, Olsson shows that inquirers in Laputa survives what he calls the “polarization test” (p. 113): “if initially disposed to judge along the same lines, inquirers in Laputa will adopt a more extreme position in the same direction as the effect of group deliberation, just like members of real argumentative bodies” (ibid.).

The influential simulation model developed in Hegselmann and Krause (2006), the HK-model for short, was one source of inspiration behind Laputa. In Vallinder and Olsson (2013), it was suggested that Laputa and the HK-model are competitors among simulation models that are truth sensitive. The HK-model has been developed in various directions by other researchers and put to interesting use in various philosophical applications (e.g. Douven and Kelp, 2011). However, the HK-model does not as naturally

lend itself to the study of the influence of network density on group competence per se, since the network is not explicitly represented in the model.<sup>13</sup>

A stable result of our investigation is that it is generally beneficial, in terms of V-value, to require inquirers to show some discretion concerning what information to post in the network. Casually asserted or recycled information tends to pull the entire network away from truth, and evidential norms like the ones we have suggested make this less likely to happen. What comes out of our simulations, regarding V-value, is that the importance of adhering to such norms become increasingly important as scientists become ever more connected. Network density can be positively correlated with group performance if inquirers are exercising proper restraint in their communicational behavior.

The present article bears on a matter addressed in Olsson (2013), footnote 3, p. 124: “We have been experimenting with a version of the [Laputa] program in which communication is possible only if new inquiry has taken place. Preliminary simulations suggest that this modification does not have any significant statistical effect on simulation outcome.” The more extensive and systematic simulations carried out in the present article disprove this suggestion. By prohibiting arbitrary repetition of information in the network, the present version of Laputa makes it possible to enforce an “argumentation interpretation” of the model, as defended in Olsson’s 2013 article and further detailed in Masterton and Olsson (2013).

To be more precise, in terms of V-value as well as polarization it is, as long as our restrictions are in effect, best for everyone to be able to communicate with everyone else. But whereas the best effect regarding V-value is obtained for a threshold of 0.99, the optimum level of polarization was obtained for a threshold of 0.80. So how we should set the threshold depends on what we value the most: having accurate credence or avoiding polarization. In this case, an epistemic value (having accurate credences) conflicts with a value of a practical kind (avoiding polarization). We also get different results depending on what kind of *epistemic* value we take to be most central. In terms of average belief of the majority, the best course of action is not to communicate at all, or to do so rarely. This remains true even if our restrictions on communication are imposed. So, again, there is a conflict in what practice should be adopted, depending on what we value the most, but this time the conflicting values are both of an epistemic nature.<sup>14</sup>

Finally, we would like to mention a particularly suggestive aspect of our study that may be of potential practical significance, namely, the fact that it seems to shed doubt on the recommendation to “publish early, publish often” frequently given to young scholars as key to a successful academic career. To see how our study is relevant, recall that our results indicate that unless inquirers agree on certain restrictions on their communications, including scientific publications, an increasingly networked scientific community runs an ever greater risk of betting on the wrong theory, as it were. We went on to show how this effect can be counteracted by scientists agreeing (i) to communicate

---

<sup>13</sup> As Rainer Hegselmann has pointed out to us (personal communication) there are various ways of (re)interpreting the HK-model in terms of a network structure. Unfortunately, a fuller investigation is outside the scope of the present article. A comparison between the two frameworks from another perspective can be found in Vallinder and Olsson (2013), p. 1441.

<sup>14</sup> We are greatly indebted to two anonymous referees for several suggestions for improvements. We have also benefitted from comments by Rainer Hegselmann on interpretational issues and, in particular, on the relation between his model and ours. Our work on this article was supported by a grant from the Swedish Research Council (*Collective Competence in Deliberative Groups: On the Epistemological Foundations of Democracy*).

only information judged to be practically certain and (ii) not to recycle information in the absence of new supporting evidence.

Now it seems to us that the said advice is difficult to reconcile with these two restrictions. In order to publish early and often one would normally have to publish also things that one has some doubts about. A certain amount of recycling of old material also seems necessary in practice if one wants to publish often, as the reader can surely testify based on her own experience. Our study indicates that, as the scientific community becomes increasingly connected, scholars should, in the interest of science as an institution, be increasingly self-critical when deciding whether to publish or not, even if this means publishing late and rarely. Alternatively or complementarily, journals and other scientific gatekeepers should be increasingly unwilling to publish doubtful or recycled articles. Unfortunately, precisely the opposite is arguably happening at the time of writing (July 2016). While the scientific community is becoming increasingly intertwined through electronic communication, there is also a trend of publishing in more or less obscure Open Access journals whose function as scientific gatekeepers could be seriously questioned. Our study gives some reasons to believe that the combined effect of increased networking and less than responsible Open Access publishing may be more problematic than has previously been assumed.

## Appendix: derivation of the credence and trust update functions

We will need the following assumptions for the derivation of the credence function:

*Principal Principle (PP):*

$$\begin{aligned} a &< C_\alpha^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, p, a < r_{\sigma\alpha} < b) < b \\ a &< C_\alpha^t(S_{\sigma\alpha}^t(\neg p) | S_{\sigma\alpha}^t, \neg p, a < r_{\sigma\alpha} < b) < b \end{aligned}$$

*Communication Independence (CI):*

$$\begin{aligned} C_\alpha^t(S_{\sigma\alpha}^t, p, a < r_{\sigma\alpha} < b) &= C_\alpha^t(S_{\sigma\alpha}^t) \times C_\alpha^t(p) \times C_\alpha^t(a < r_{\sigma\alpha} < b) \\ C_\alpha^t(S_{\sigma\alpha}^t, \neg p, a < r_{\sigma\alpha} < b) &= C_\alpha^t(S_{\sigma\alpha}^t) \times C_\alpha^t(\neg p) \times C_\alpha^t(a < r_{\sigma\alpha} < b) \end{aligned}$$

*Source Independence (SI):*

$$\begin{aligned} C_\alpha^t\left(\bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) \middle| p\right) &= \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | p) \\ C_\alpha^t\left(\bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) \middle| \neg p\right) &= \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | \neg p) \end{aligned}$$

where  $0 \leq a < b \leq 1$ ,  $\Sigma_\alpha^t$  is the set of sources that give information to  $\alpha$  at  $t$ , and  $m_{\sigma\alpha}^t$  is the content of the source  $\sigma$ 's message.

Since the trust function, which plays a crucial part in the model, is continuous, the derivation will sometimes need to take a detour through conditional probability densities rather than the conditional probabilities themselves. We will briefly sketch how this can be done here.

We have so far not been specific about the  $\sigma$ -algebra  $Z$  that  $C_\alpha^t$  is defined on. Assume that it is product of several such algebras, the first of which is discrete and generated by atomic events such as  $p, \neg p, S_{\beta\alpha}(p)$  etc., and the others, which are continuous, are generated by events of the form  $a \leq r_{\sigma\alpha} \leq b$ . Call the first algebra  $X$  and the others  $Y_{\sigma_0}, \dots, Y_{\sigma_n}$ . It is clear that, as long as time and the number of inquirers are both finite,  $X$  will have only finitely many elements. On the other hand,  $Y_{\sigma_0}, \dots, Y_{\sigma_n}$  are certainly infinite. As mentioned, we assume that  $Z = X \times Y_{\sigma_0} \times \dots \times Y_{\sigma_n}$ . Given any source  $\sigma_k$  and time  $t$ , we can therefore interpret the part of  $C_\alpha^t$  defined on the subalgebra  $X \times Y_{\sigma_k}$  of  $Z$  as arising from a joint density function  $\kappa_{\sigma\alpha}^t(\varphi; x)$  defined through the equation

$$C_\alpha^t(\varphi, a < r_{\sigma\alpha} < b) = \int_a^b \kappa_{\sigma\alpha}^t(\varphi; x) dx$$

Since we have used the comma to represent conjunction earlier in the paper we use a semicolon here to separate the two variables: the first propositional, and the second real-valued. Like  $\tau$ , this distribution's existence and essential uniqueness are guaranteed by the Radon-Nikodym theorem, and in fact  $\tau_{\sigma\alpha}^t$  is the marginal distribution of  $\kappa_{\sigma\alpha}^t$  with respect to the reliability variable  $r_{\sigma\alpha}$  in question. Since the conditional distribution of a random variable is the joint distribution divided by the marginal distribution of that variable, this means that we have that

$$\kappa_{\sigma\alpha}^t(\varphi|x) = \frac{\kappa_{\sigma\alpha}^t(\varphi; x)}{\tau_{\sigma\alpha}^t(x)}$$

which is what will be used to make sense of what it means to conditionalize on  $r_{\sigma\alpha}$  having a certain value rather than merely being inside an interval. Setting  $r_{\sigma\alpha} = x$ ,  $a = x - \epsilon$  and  $b = x + \epsilon$  in *PP* and *CI* and letting  $\epsilon \rightarrow 0$ , we get the versions

$$\text{Principal Principle}_0 (PP_{\text{lim}}): \quad \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p)|S_{\sigma\alpha}^t, p; x) = \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(\neg p)|S_{\sigma\alpha}^t, \neg p; x) = x$$

$$\text{Communication Independence}_0 (CI_{\text{lim}}): \quad \begin{aligned} \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t, p; x) &= C_\alpha^t(S_{\sigma\alpha}^t) \times C_\alpha^t(p) \times \tau_{\sigma\alpha}^t(x) \\ \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t, \neg p; x) &= C_\alpha^t(S_{\sigma\alpha}^t) \times C_\alpha^t(\neg p) \times \tau_{\sigma\alpha}^t(x) \end{aligned}$$

We can now proceed with the actual derivation. By conditionalization, we must have that  $C_\alpha^{t+1}(p)$  is equal to  $C_\alpha^t(\bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | p)$ . Applying Bayes' theorem and then *SI* to this expression gives

$$\begin{aligned} & C_\alpha^t \left( p \left| \bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) \right. \right) \\ &= \frac{C_\alpha^t(p) \times C_\alpha^t \left( \bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) \mid p \right)}{C_\alpha^t(p) \times C_\alpha^t \left( \bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) \mid p \right) + C_\alpha^t(\neg p) \times C_\alpha^t \left( \bigwedge_{\sigma \in \Sigma_\alpha^t} S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) \mid \neg p \right)} \\ &= \frac{C_\alpha^t(p) \times \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | p)}{C_\alpha^t(p) \times \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | p) + C_\alpha^t(\neg p) \times \prod_{\sigma \in \Sigma_\alpha^t} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t) | \neg p)} \end{aligned}$$

which gives us the posterior credence in terms of the values  $C_\alpha^t(S_{\sigma\alpha}^t(p)|p)$  and  $C_\alpha^t(S_{\sigma\alpha}^t(\neg p)|\neg p)$ . Our next task is thus to derive these expressions. Since  $S_{\sigma\alpha}^t(p)$  is equivalent to  $S_{\sigma\alpha}^t(p) \wedge S_{\sigma\alpha}^t$ , it follows that  $C_\alpha^t(S_{\sigma\alpha}^t(p)|p) = C_\alpha^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t|p)$ . Applying first the definition of conditional probability and then the continuous law of total probability, the definition of conditional probability again, and finally  $CI_{\text{lim}}$ , we get

$$\begin{aligned} C_\alpha^t(S_{\sigma\alpha}^t(p)|p) &= \frac{1}{C_\alpha^t(p)} \times C_\alpha^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, p) \\ &= \frac{1}{C_\alpha^t(p)} \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, p; x) dx \\ &= \frac{1}{C_\alpha^t(p)} \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, p; x) \times \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t, p; x) dx \\ &= \frac{1}{C_\alpha^t(p)} \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, p, x) \times C_\alpha^t(S_{\sigma\alpha}^t) \times C_\alpha^t(p) \times \tau_{\sigma\alpha}^t(x) dx \\ &= C_\alpha^t(S_{\sigma\alpha}^t) \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, p; x) \times \tau_{\sigma\alpha}^t(x) dx \end{aligned}$$

But  $PP_{\text{lim}}$  ensures that  $\kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, p; x) = x$ , so we get

$$C_\alpha^t(S_{\sigma\alpha}^t(p)|p) = C_\alpha^t(S_{\sigma\alpha}^t) \times \int_0^1 x \times \tau_{\sigma\alpha}^t(x) dx = C_\alpha^t(S_{\sigma\alpha}^t) \times E[\tau_{\sigma\alpha}^t]$$

Parallel derivations give that

$$\begin{aligned} C_\alpha^t(S_{\sigma\alpha}^t(\neg p)|p) &= C_\alpha^t(S_{\sigma\alpha}^t) \times E[\overline{\tau_{\sigma\alpha}^t}] \\ C_\alpha^t(S_{\sigma\alpha}^t(p)|\neg p) &= C_\alpha^t(S_{\sigma\alpha}^t) \times E[\tau_{\sigma\alpha}^t] \\ C_\alpha^t(S_{\sigma\alpha}^t(\neg p)|\neg p) &= C_\alpha^t(S_{\sigma\alpha}^t) \times E[\overline{\tau_{\sigma\alpha}^t}] \end{aligned}$$

Now let  $\Sigma_\alpha^t(p) \subseteq \Sigma_\alpha^t$  be the set of sources that give  $\alpha$  the message  $p$  at  $t$ , and let  $\Sigma_\alpha^t(\neg p) = \Sigma_\alpha^t \setminus \Sigma_\alpha^t(p)$ . Plugging the above expressions into our earlier result gives the sought for expression

$$C_\alpha^{t+1}(p) = \frac{C_\alpha^t(p) \times \prod_{\sigma \in \Sigma_\alpha^t(p)} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|p)}{C_\alpha^t(p) \times \prod_{\sigma \in \Sigma_\alpha^t(p)} C_\alpha^t(S_{\sigma\alpha}^t) + C_\alpha^t(\neg p) \times \prod_{\sigma \in \Sigma_\alpha^t(\neg p)} C_\alpha^t(S_{\sigma\alpha}^t(m_{\sigma\alpha}^t)|\neg p)} = \frac{\gamma}{\gamma + \delta}$$

where

$$\begin{aligned} \gamma &= C_\alpha^t(p) \times \prod_{\sigma \in \Sigma_\alpha^t(p)} (C_\alpha^t(S_{\sigma\alpha}^t) \times E[\tau_{\sigma\alpha}^t]) \times \prod_{\sigma \in \Sigma_\alpha^t(\neg p)} (C_\alpha^t(S_{\sigma\alpha}^t) \times E[\overline{\tau_{\sigma\alpha}^t}]) \\ \delta &= C_\alpha^t(\neg p) \times \prod_{\sigma \in \Sigma_\alpha^t(p)} (C_\alpha^t(S_{\sigma\alpha}^t) \times E[\overline{\tau_{\sigma\alpha}^t}]) \times \prod_{\sigma \in \Sigma_\alpha^t(\neg p)} (C_\alpha^t(S_{\sigma\alpha}^t) \times E[\tau_{\sigma\alpha}^t]) \end{aligned}$$

For the derivation of the trust update expression we assume  $PP$  and  $CI$ , but not  $SI$ . The function we wish to derive is

$$\tau_{\sigma\alpha}^{t+1}(x) = \kappa_{\sigma\alpha}^t(x | S_{\sigma\alpha}^t(m_{\sigma\alpha}^t))$$

for a source  $\sigma$  of  $\alpha$ , and a message  $m_{\sigma\alpha}^t$  from that source. Assume that  $m_{\sigma\alpha}^t \equiv p$  (the case  $m_{\sigma\alpha}^t \equiv \neg p$  is completely symmetrical). Applying the definition of conditional probability, the equivalence  $S_{\sigma\alpha}^t \wedge S_{\sigma\alpha}^t(p) \equiv S_{\sigma\alpha}^t(p)$ , and the (discrete) law of total probability, we get

$$\begin{aligned} \kappa_{\sigma\alpha}^t(x | S_{\sigma\alpha}^t(p)) &= \frac{\kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p); x)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p))} = \frac{\kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t; x)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p))} \\ &= \frac{\kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, p; x) + \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, \neg p; x)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p))} \\ &= \frac{\kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, p; x) \times \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t, p; x) + \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, \neg p; x) \times \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t, \neg p; x)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p))} \end{aligned}$$

Now apply  $PP_{\text{lim}}$  and  $CI_{\text{lim}}$  to the factors in both terms of the numerator, and then again the equivalence  $S_{\sigma\alpha}^t \wedge S_{\sigma\alpha}^t(p) \equiv S_{\sigma\alpha}^t(p)$ :

$$\begin{aligned} \kappa_{\sigma\alpha}^t(x | S_{\sigma\alpha}^t(p)) &= \tau_{\sigma\alpha}^t(x) \times C_{\alpha}^t(S_{\sigma\alpha}^t) \times \frac{x \times C_{\alpha}^t(p) + \bar{x} \times C_{\alpha}^t(\neg p)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p))} \\ &= \tau_{\sigma\alpha}^t(x) \times \frac{x \times C_{\alpha}^t(p) + \bar{x} \times C_{\alpha}^t(\neg p)}{C_{\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t)} \end{aligned}$$

We can calculate the denominator in this expression by using the definition of conditional probability and expanding twice using the law of total probability (once using the discrete version, and once using the continuous one):

$$\begin{aligned} C_{\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t) &= \frac{C_{\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t)}{C_{\alpha}^t(S_{\sigma\alpha}^t)} = \frac{C_{\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, p) + C_{\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, \neg p)}{C_{\alpha}^t(S_{\sigma\alpha}^t)} = \\ &= \frac{1}{C_{\alpha}^t(S_{\sigma\alpha}^t)} \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, p; x) dx + \frac{1}{C_{\alpha}^t(S_{\sigma\alpha}^t)} \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p), S_{\sigma\alpha}^t, \neg p; x) dx \\ &= \frac{1}{C_{\alpha}^t(S_{\sigma\alpha}^t)} \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, p; x) \times \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t, p; x) dx \\ &\quad + \frac{1}{C_{\alpha}^t(S_{\sigma\alpha}^t)} \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, \neg p; x) \times \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t, \neg p; x) dx \end{aligned}$$

Let us refer to the last expression as  $\psi$ . Applying  $CI_{\text{lim}}$ , then cancelling, and applying  $PP_{\text{lim}}$ , we get

$$\begin{aligned} \psi &= C_{\alpha}^t(p) \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, p; x) \times \kappa_{\sigma\alpha}^t(x) dx \\ &\quad + C_{\alpha}^t(\neg p) \times \int_0^1 \kappa_{\sigma\alpha}^t(S_{\sigma\alpha}^t(p) | S_{\sigma\alpha}^t, \neg p; x) \times \kappa_{\sigma\alpha}^t(x) dx \\ &= C_{\alpha}^t(p) \times \int_0^1 x \times \tau_{\sigma\alpha}^t(x) dx + C_{\alpha}^t(\neg p) \times \int_0^1 \bar{x} \times \tau_{\sigma\alpha}^t(x) dx \end{aligned}$$

$$= C_{\alpha}^t(p) \times E[\tau_{\sigma\alpha}^t] + C_{\alpha}^t(\neg p) \times E[\overline{\tau_{\sigma\alpha}^t}]$$

Putting it all together, we finally arrive at the result:

$$\tau_{\sigma\alpha}^{t+1}(x) = \tau_{\sigma\alpha}^t(x) \times \frac{x \times C_{\alpha}^t(p) + \bar{x} \times C_{\alpha}^t(\neg p)}{C_{\alpha}^t(p) \times E[\tau_{\sigma\alpha}^t] + C_{\alpha}^t(\neg p) \times E[\overline{\tau_{\sigma\alpha}^t}]}$$

## References

- Bala, V., and Goyal, S. (1998), "Learning from neighbours", *Review of Economic Studies Limited* 65: 595-621.
- Barabási, A.-L. and Albert, R. (1999), "Emergence of scaling in random networks", *Science*, 509-512.
- Collins, P. J., Hahn, U., von Gerber, Y., and Olsson, E. J. (2015), "The Bi-directional Relationship Between Source Characteristics and Message Content," *Proceedings of the annual meeting of the cognitive science society (CogSci 2015)*.
- Cook, J. et al. (2013), "Quantifying the consensus on anthropogenic global warming in the scientific literature", *Environmental Research Letters* 8:1-8.
- Douven, I. and Kelp, C. (2011), "Truth approximation, social epistemology, and opinion dynamics", *Erkenntnis* 75(2): 271-283.
- Goldman, A. I. (1986), *Epistemology and Cognition*, Harvard University Press.
- Goldman, A. I. (1999), *Knowledge in a social world*. Clarendon Press, Oxford.
- Goodin, R. E. (2003), *Reflective democracy*, Oxford University Press.
- Hegselmann, R. and Krause, U. (2006), "Truth and cognitive division of labour: first steps towards a computer-aided social epistemology, *Journal of Artificial Societies and Social Simulation* 9(3) (available at <http://jasss.soc.surrey.ac.uk/9/3/10.html>).
- Jeffrey, R. C. (1983), *The Logic of Decision*, 2nd ed., University of Chicago Press
- Lazer, D., and Friedman, A. (2007), "The network structure of exploration and exploitation", Computer and Information Science Faculty Publications. Paper 1. <http://hdl.handle.net/2047/d20000313>.
- Lewis, D. (1980), "A subjectivist's guide to objective chance", in Richard C. Jeffrey (ed.) *Studies in Inductive Logic and Probability*, vol. 2., University of California Press.
- Mason, W. A., Jones, A., and R. L. Goldstone (2008), "Propagation of innovations in networked groups", *Journal of Experimental Psychology: General* 137(3): 422-433.
- Masterton, G. (2014), "Topological variability of collectives and its import for social epistemology", *Synthese* 191: 2433-2443.
- Masterton, G. and Olsson, E. J. (2013), "Argumentation and belief updating in social networks: a Bayesian approach", in Fermé, E., Gabbay, D. and Simari, G. (Eds.), *Trends in belief revision and argumentation dynamics*, College Publications.
- Miller, D. (1966), "A Paradox of Information", *British Journal for the Philosophy of Science*, 59-61.
- Newman, M. E. J. (2001), "The structure of scientific collaboration networks", *Proc. Natl. Acad. Sci. USA*, 404-409.
- Olsson, E. J. (2011), "A simulation approach to veritistic social epistemology", *Episteme* 8 (2): 127-143.

- Olsson, E. J. (2013), "A Bayesian simulation model of group deliberation and polarization", in Zenker, F. (ed.), *Bayesian Argumentation*, Synthese Library, New York: Springer, 113-134.
- Olsson, E. J., and Vallinder, A. (2013), "Norms of assertion and communication in social networks", *Synthese* 190: 1437-1454.
- Redner (1998), "How popular is your paper? An empirical study of the citation distribution", *The European Physical Journal B*, 131-134.
- Romer, D. (2012), *Advanced Macroeconomics*, 4<sup>th</sup> edition, McGraw-Hill.
- Vallinder, A., and Olsson, E. J. (2013), "Do computer simulations support the argument from disagreement?", *Synthese* 190(8): 1437-1454.
- Vallinder, A., and Olsson, E. J. (2014), "Trust and the value of overconfidence: a Bayesian perspective on social network communication", *Synthese* 191: 1991-2007.
- Zollman, K. J. (2007), "The communication structure of epistemic communities", *Philosophy of Science* 74 (5): 574-587.