This is the authors' final version of the article published in *Physica A*, 563, 125480 (2021). The final publication is available at: https://doi.org/10.1016/j.physa.2020.125480

Opinion Dynamics of Social Learning with a Conflicting Source

Catherine A. Glass[a] and David H. Glass[b]

[a]Economics Research Branch, Agri-Food and Biosciences Institute, 18a Newforge Lane, Belfast, BT9 5PX, UK.[1] Email: catherine.glass@afbini.gov.uk

[b]School of Computing, Shore Road, Newtownabbey, Co. Antrim, BT37 0QB, UK.

Email: dh.glass@ulster.ac.uk

Abstract

The way in which agents are influenced by the truth and/or a conflicting source can have a significant effect on the extent to which social learning is successful. We investigate these influences via several variations of the Hegselmann-Krause model of opinion dynamics. First, we compare two ways of modelling the influence of truth in the absence of a conflicting source and find that in a model where access to the truth is more restricted, increasing the proportion of truth seekers in the society has little effect on convergence to the truth. Second, we investigate the same models of truth in the presence of a conflicting source, which could represent the opinions of a radical group, opinion leader or media source. The results show that a consensus on the truth can be reached in certain cases in both models, but also that in a wide range of cases both models give rise to the same partition of the society into truth seekers and non-truth seekers.

Keywords: social learning; Hegselmann-Krause model; opinion dynamics; truth; opinion leader; stubborn agents.

1 Introduction

The question of how agents learn the truth in a social context has been explored widely in studies of social learning [1–5] and through simulations using various opinion dynamics models [6–14]. In this paper, we employ several extensions of the Hegselmann-Krause (HK) model [15], which is a bounded confidence model of opinion dynamics in which two agents interact only if their opinions lie within a specified distance from each other. An alternative bounded confidence model is the Deffuant-Weisbuch model [16]. The difference between the two models is the updating rule. In the HK model, all agents interact simultaneously with those within their confidence limits whilst in the DW model, all interactions are pairwise and sequential. In this paper the HK model is used to explore social learning in the presence and absence of a conflicting source. In a recent review of opinion dynamics [17], the authors highlight a number of open questions including the importance of developing intervention strategies to manage

[1]Address since November, 2019.

the evolution of opinions with an established purpose. In this study, the purpose is to promote social learning. This is achieved by investigating how a range of factors such as the proportion of truth seekers in a society, the accessibility of truth, the willingness of agents to consider alternative views and the presence of a conflicting source can affect the extent to which social learning occurs.

To start with, we adopt a model introduced by Hegselmann and Krause [6] (henceforth HK06) that extends the standard HK model to incorporate truth in such a way that the truth attracts truth seekers irrespective of whether their opinions are close to the truth. We then compare it directly with an alternative approach in which truth seekers are only drawn to the truth if they lie within its neighbourhood. These models can be thought of as representing scenarios where the truth is readily accessible in the former case and where the truth is only accessible to those with relevant expertise or privileged access to it in the latter case.

We then extend each of these models to include a conflicting source that can also influence agents, again based on the HK model. Various extensions of the HK model have been proposed to explore the influence of radical or stubborn groups, opinion leaders and mass media [18–20], each of which could correspond to a source in conflict with the truth in a given context. Nevertheless, interest in such models is typically to investigate their influence on opinion dynamics in general rather than with specific reference to the truth. The respective influences of competing opinion leaders or groups have been explored in the HK model [21, 22], which is more closely related to the current work since it too involves competing sources. Here, however, the focus is on truth and the influence of truth and the conflicting source are modelled differently.

Section 2 sets out several extensions of the HK model, with truth incorporated in two different ways and then each of these models further modified to allow for the presence of a conflicting source. Section 3 presents the results of computer simulations for the different models of truth, first in the absence, and then in the presence, of a conflicting source. These results enable us to compare the models to see how close the society gets to the truth in each case. Some analytical results are also presented for the long term behaviour of the society in certain cases when both the truth and the conflicting source influence the dynamics. Section 4 presents conclusions.

2 HK models with truth and a conflicting source

We consider four models of social learning, all based on the HK model of opinion dynamics [15]. The first two models include different ways in which the truth can be included in the HK model, while two further models also include a conflicting source of information.

2.1 Model I: HK with truth accessible to all

Consider a group of n agents, m of whom are truth seekers, while the remaining $n - m$ agents will be referred to as normal agents. Let agent i have belief $x_i(t)$ lying in the interval $[0, 1]$ at time t. Agent j will be a neighbour of agent i at t if the belief of agent j lies within a certain bound of confidence of agent i, that is $|x_i(t) - x_j(t)| \leq \varepsilon_i$, where ε_i is the confidence of agent i. Let τ be a value in the interval

$[0,1]$ representing truth. We follow HK06 by specifying update rules for the agents as follows:

$$x_i(t+1) = \begin{cases} \alpha_i \tau + (1-\alpha_i) x_i^{HK}(t+1), & \text{if agent } i \text{ is a truth seeker} \\ x_i^{HK}(t+1), & \text{otherwise} \end{cases} \tag{1}$$

where α_i is the weighting given to the truth and $x_i^{HK}(t+1)$ is the updated opinion for agent i obtained by the standard HK approach:

$$x_i^{HK}(t+1) = \frac{1}{|I(i,t)|} \sum_{j \in I(i,t)} x_j(t), \tag{2}$$

with $I(i,t) = \{j : |x_i(t) - x_j(t)| \leq \varepsilon_i\}$, i.e. the set of agents who are in the neighbourhood of agent i.

In this model, all truth seekers have access to the truth. This could represent scenarios where no specialist knowledge or permissions are required. For example, relevant information might be freely available on the Internet for those who wish to explore it. There is no requirement that a truth seeker's opinion needs to be close to τ before the agent's opinion is influenced by the truth.

2.2 Model II: HK with limited accessibility to truth

In contrast to Model I, it is only truth seekers whose current opinions are sufficiently close to the truth value, τ, who are influenced by the truth in Model II. In this case, we could think of such truth seekers as representing people with specialist knowledge or privileged access to the truth in a given context such as those with relevant medical or legal expertise. Nevertheless, other agents within the society can still be influenced by the truth indirectly.

A truth seeking agent i will then be influenced by truth if τ lies within a certain distance of the agent's opinion. For simplicity, we set this distance to be the agent's bound of confidence, ε_i, though other values could also be considered. One way to implement the influence of truth with limited accessibility would be to follow the approach proposed by Hegselmann and Krause [18], which will be discussed in the context of Model III below. Adopting this approach, truth would be represented by a number of fixed agents within the standard HK updating procedure, but in order to compare with Model I we instead give a weighting to truth as before. This gives the following updating rule:

$$x_i(t+1) = \begin{cases} \alpha_i \tau + (1-\alpha_i) x_i^{HK}(t+1), & \text{if } i \text{ is a truth seeker and } |x_i(t) - \tau| \leq \varepsilon_i \\ x_i^{HK}(t+1), & \text{otherwise} \end{cases} \tag{3}$$

where, as before, $x_i^{HK}(t+1)$ is given by equation (2).

2.3 Model III: HK with truth accessible to all and a conflicting source

In Models I and II, there are two influences on agents' opinions: a social influence that affects all agents, whether they are truth seekers or normal agents, and an influence from the truth that only affects truth seekers. Now we add a further influence arising from a conflicting source, i.e. one that differs from the truth. This could represent an influence due to mass media or an opinion leader, for example.

To achieve this, we adopt the approach of Hegselmann and Krause [18], who represented the influence of a constant signal within the HK model by means of a group of stubborn agents, all of whose opinions are fixed at a single value, but who can influence other agents via the normal HK updating procedure. For computational reasons, this can be implemented by a single stubborn agent that is given an appropriate weighting, so that instead of r stubborn agents, there is one agent that is given a weighting of r in the HK model. This weighting is referred to as the *reputation* of an opinion leader by Chen *et al* [21], who investigated a closely related model that allowed the opinion leader's opinion to evolve over time. The approach in [21] has also been explored in the context of truth in [23] to investigate the influence of characteristics of opinion leaders, but it differs in a number of important respects from the current work. In particular, in [23] truth is modelled in essentially the same way as an opinion leader with a specified reputation. Similarly, while the work in [22] can be interpreted in terms of truth, under such an interpretation truth would again be modelled in the same way as an opinion leader. By contrast, in the current work, different approaches are used to model truth and the influence of an opinion leader as is the case in [24]. Also, in contrast to [22–24], the current work distinguishes between truth seekers and non-truth seekers and between models where the truth is limited and unlimited in terms of accessibility.

Unlike truth, we assume that the conflicting source influences normal agents and truth seekers alike. Of course, it would also be possible to explore alternative models where only the normal agents, or some proportion of them, are influenced by this source, but these options are not pursued here. Following [18], an agent i will take this source into account if it is sufficiently close to the agent's opinion and for simplicity we will assume that it must be within the agent's bound of confidence, ε_i, though other values could also be considered.

Incorporating this conflicting source in Model I gives the following update rule for Model III:

$$x_i(t+1) = \begin{cases} \alpha_i \tau + (1-\alpha_i) x_i^{MHK}(t+1), & \text{if agent } i \text{ is a truth seeker} \\ x_i^{MHK}(t+1), & \text{otherwise} \end{cases} \tag{4}$$

where $x_i^{MHK}(t+1)$ is the updated opinion for agent i obtained by a modified version of the HK model that takes the conflicting source into account as follows:

$$x_i^{MHK}(t+1) = \begin{cases} \frac{\sum_{j \in I(i,t)} x_j(t) + r \cdot \psi}{|I(i,t)| + r}, & \text{if } |x_i(t) - \psi| \leq \varepsilon_i \\ x_i^{HK}(t+1), & \text{otherwise} \end{cases} \tag{5}$$

where ψ represents the opinion of the conflicting source, r represents the weight or reputation of this source, and as before $I(i,t) = \{j : |x_i(t) - x_j(t)| \leq \varepsilon_i\}$, i.e. the set of agents (normal and truth seeking) who are in the neighbourhood of agent i, and $x_i^{HK}(t+1)$ is given by equation (2).

2.4 Model IV: HK with limited accessibility to truth and a conflicting source

This model adopts the same approach to the conflicting source as in Model III, but now incorporates truth via Model II rather than Model I to give:

$$x_i(t+1) = \begin{cases} \alpha_i \tau + (1-\alpha_i) x_i^{MHK}(t+1), & \text{if } i \text{ is a truth seeker and } |x_i(t) - \tau| \leq \varepsilon_i \\ x_i^{MHK}(t+1), & \text{otherwise} \end{cases} \tag{6}$$

where $x_i^{MHK}(t+1)$ is given by equation (5).

2.5 Simulations

In the following sections, we present results to compare Models I and II and then Models III and IV. The results were obtained from simulations that included $n = 100$ agents, with the number of truth seeking agents, m, and hence the normal agents, $n - m$, being varied. Since $n = 100$, we will often refer to m as the percentage of truth seekers. The initial opinions of these agents were generated randomly from a uniform distribution on the interval $[0, 1]$, while the truth value, τ, and, where appropriate, the opinion of the conflicting source, ψ, are specified in the discussion. For simulations involving the conflicting source, its weight or reputation, r, was varied. Since a weighting of r is equivalent to r stubborn agents with a weighting of one and with an opinion ψ, and since $n = 100$, we will often refer to r as the percentage of stubborn agents. Note, however, that these agents should be thought of as *additional* to the n agents, which was fixed at 100.

To constrain the parameter space, we keep the weighting of truth the same for each agent so that $\alpha_i = \alpha$ for all agents i. Similarly, we keep the bound of confidence for each agent the same, $\varepsilon_i = \varepsilon$. However, we then vary both α and ε to explore how they affect the results. We also vary the number of truth seekers, m, the value of truth, τ, and the percentage of stubborn agents, r.

For each simulation, the update rules were iterated until a steady state is reached, which was implemented by requiring that $\sum_i |x_i(t+1) - x_i(t)| < 10^{-4}$. We present results for the root mean square deviation (RMSD) of the final opinions of the normal and truth seeking agents from the truth. This is given by:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n}(x_i^* - \tau)^2}{n}}, \tag{7}$$

where x_i^* is the opinion of agent i at the end of the simulation. We repeated each simulation 200 times and obtained average values for the RMSD. We have carried out simulations with different numbers of repetitions to ensure that the results are converged. We have also obtained results for different numbers of agents, n, and found that this does not change the results significantly.

3 Results

3.1 Incorporating truth: a comparison of Models I and II

Model I was investigated previously in HK06, so our goal in this section is to compare this model with Model II and pave the way for the comparison between Models III and IV in section 3.2. All our findings for Model I are consistent with those in HK06, while Model II is similar for some regions of the parameter space while differing in some important respects in other regions.

As pointed out in HK06, according to Model I, the whole society can end up with a consensus on the truth even if there is only a tiny minority of truth seekers (even just one), provided the values for other parameters are chosen appropriately. Our results show that this is the case for Model II as well, although the region of parameter space where a consensus on the truth obtains is reduced.

3.1.1 Comparison of Models I and II with few truth seekers and different truth values

Figure 1 shows the RMSD from the truth of the final opinions of all normal and truth seeking agents for Models I and II in various scenarios where the percentage of truth seekers is small, $m = 5\%$. Results are presented for three different truth values: $\tau = 0.5$ corresponding to truth being found at the centre of the opinion space, $\tau = 1$ so that truth is at one extreme, and an intermediate value, $\tau = 0.75$. For each value of τ, the confidence of the agents, ε, is varied between 0 and 0.7 while the weighting given to the truth, α, is varied between 0 and 1. The left hand side of the figure provides results for Model I while the right hand side provides results for Model II.

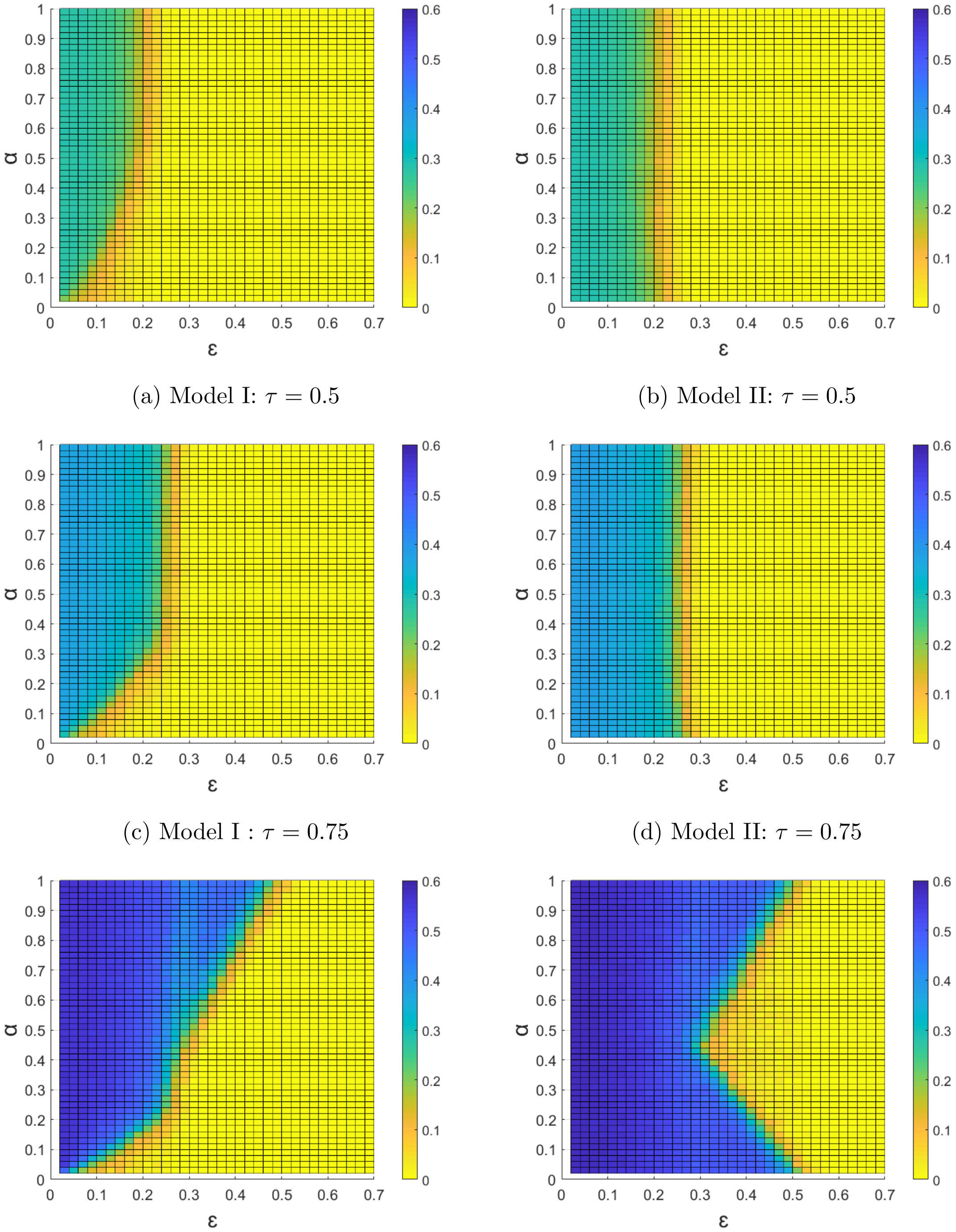
For both models, all agents form a consensus on the truth provided the confidence is large enough. Indeed as pointed out in HK06, "for almost all values of [τ, m, and α] there seems to be a critical value ε^* such that a consensus on the truth is possible [for Model I] only if $\varepsilon > \varepsilon^*$" (HK06, p. 19). This is also borne out in our results for Model II, as illustrated in figure 1 for $m = 5\%$. As also pointed out in HK06, "if the truth is in the centre of the opinion space the critical confidence level is lower than for more extreme truths" (HK06, p. 19) and the results in figure 1 also illustrate this for Model II.

However, HK06 makes the further claim that "the critical confidence level that has to be exceeded *increases* as the strength α of truth attraction *increases*" (HK06, p. 19). Our results in figure 1 bear this out for Model I (see also figure 2), although it should be noted that even for Model I, once α reaches a certain level (around 0.3-0.4) there is very little, if any, further increase in the critical confidence level when τ is non-extreme (0.5 and 0.75 in figures 1(a) and (c) respectively). However, this finding does not apply to Model II. For the non-extreme truth values in figures 1(b) and (d) the results are almost independent of α, so that increasing α has no bearing on the critical confidence level. Furthermore, when $\tau = 1$, the HK06 result applies for values of α above about 0.45, but below this *increasing* α actually *decreases* the critical confidence level.

A related claim of HK06 is that "for almost all [τ, m, and ε] there seems to be a critical value α^* for the strength of the attraction of truth such that for a consensus on the truth it is necessary that $\alpha < \alpha^*$" (HK06, p. 20). Of course, as discussed earlier, there is a consensus on the truth if ε is sufficiently high, but for a range of lower values of ε each of the results for Model I in figures (a), (c) and (d) demonstrate that there is indeed a critical value α^*, which decreases with decreasing ε (see also figure 2 for higher values of m).

However, in general there is no such critical value for α in Model II. As noted above, for the non-extreme truth values the results are almost independent of α. When $\tau = 1$ (figure 1(f)), there is a range of confidence values (ε around 0.3 to 0.5) where there are two critical values of α. For a given confidence in this range, there is an upper critical value α^U such that for a consensus on the truth it is necessary that α is *less than* α^U as in Model I and in line with HK06. There is also a lower critical value α^L such that for a consensus on the truth it is necessary that α is *greater than* α^L and hence exactly the opposite to the dependence proposed in HK06.

What are the mechanisms that result in these critical value of α? First of all, consider Model I. In Model I, as described in equation (1), all truth seekers must get to the truth and a low α can be beneficial because the truth seekers converge to the truth slowly and are able to bring non-truth seekers along with



(a) Model I: $\tau = 0.5$ (b) Model II: $\tau = 0.5$

(c) Model I : $\tau = 0.75$ (d) Model II: $\tau = 0.75$

(e) Model I: $\tau = 1$ (f) Model II: $\tau = 1$

Figure 1: RMSD from the truth, τ, for Models I and II at varying values of τ and m=5%

them (*take-along-effect* in HK06). However, if α is too high, then for a range of values of confidence truth seekers can converge to the truth too quickly leaving normal agents behind (*leaving-behind-effect* in HK06). The critical value α^* marks the transition from the *leaving-behind-effect* (for $\alpha > \alpha^*$) to the *take-along-effect* (for $\alpha < \alpha^*$). A key difference between Models I and II can then be expressed in terms of this transition from the *leaving-behind-effect* to the *take-along-effect* being much less common in Model II than it is in Model I. For example, no such transition is found for low α in any of the results for Model II in figure 1 (see also figure 2). Returning to the case where $\tau = 1$ (figure 1(f)), however, we see that such a transition does occur in Model II at higher values of α for a range of values of confidence; this is the upper critical value of α^U mentioned earlier. For $\alpha > \alpha^U$ truth seekers within the confidence of truth all converge to the truth, but fail to bring their neighbours with them, while for slightly lower values of α, they are able to bring about a consensus to the truth.

There are two reasons for these differences between Models I and II. First, consider the case of low confidence. Recall that only truth seekers close to the truth are influenced by it in Model II. Consequently, while those close to the truth can converge to the truth and bring some of their non-truth seeking neighbours with them, the low value of confidence means they have little or no influence on agents who are further away from the truth. Second, when confidence is higher and $\tau = 1$ (figure 1(f)), a low α can actually be disadvantageous in Model II because the weight of truth may not be sufficiently strong to withstand the influence of agents who are not influenced by the truth. This means that even truth seekers whose opinions are within the confidence of truth can be pulled away from it, which is what occurs at $\tau = 1$; the pull of the crowd is greater than the influence of truth, so that a consensus is formed close to 0.5 and thus far from the truth. Hence, a major difference between Models I and II is not merely that truth seekers are guaranteed to converge to the truth in Model I (if $\alpha > 0$) while this is not the case in Model II, but that even truth seekers within the confidence of truth may fail to converge to the truth in Model II.

Despite the differences noted between Models I and II in figure 1, it is worth returning to some similarities. It is not surprising that for high values of confidence both models give the same results since in the limit of confidence going to one the models are equivalent. In fact, for all values of τ in figure 1, both models yield consensus to the truth for confidence greater than about 0.5 (and for much lower values of confidence when $\tau = 0.5$ and $\tau = 0.75$). More surprisingly, there is also a lot of agreement between the models even for lower confidence for all values of τ, provided α is greater than about 0.4. In both models for non-extreme truth values, full truth convergence is achieved at a confidence value close to 0.3 and the results are mostly independent of α as it increases above about 0.4.

3.1.2 Comparison of Models I and II with varying proportions of truth seekers

So far we have kept the percentage of truth seekers fixed at $m = 5\%$, but in figure 2 we increase m from 5% to 10%, 30%, 50% and 90% in both Models I and II while keeping the value of truth fixed at $\tau = 0.75$ for all the simulations. In Model I, as the proportion of truth seekers is increased, greater truth convergence is achieved for confidence less than 0.2. This, however, can be explained by the fact that in Model I all truth seekers are guaranteed to converge to the truth and so for $m = 90\%$ then the majority of agents will converge to the truth even for extremely low confidence. For higher values of confidence, but still less than 0.2, all the truth seekers will converge to the truth together with normal agents they have managed to bring along with them. It is therefore not surprising that the proportion of agents converging to the truth increases with the percentage of truth seekers.

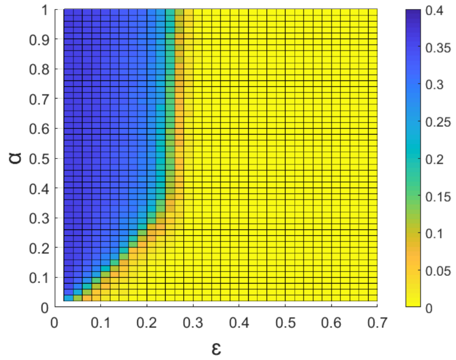
The results for Model II in figure 2 are completely different in this respect, however, since truth convergence is hardly affected at all by an increase in the proportion of truth seekers from 5% to 90%. It may seem very surprising that increasing the percentage of truth seekers in the society so significantly has very little effect on the proportion of agents who end up at the truth. Recall, however, that for a low confidence level, very few truth seekers are within the confidence of truth. Even so, considering the number within the confidence of truth will increase in proportion to the percentage of truth seekers, this might have been expected to have more influence on the society as a whole. However, the problem remains in Model II that agents who start out with opinions far from the the truth can neither be influenced by the truth themselves nor influence others towards the truth if confidence is not high enough. For example, when $\varepsilon = 0.2$, typically agents converge into two groups, one at the truth and the other some distance from it, with both containing a mixture of truth seekers and non-truth seekers.

According to HK06, in a wide range of scenarios "it seems that for a given position of the truth the critical value ε^* slightly decreases if [m decreases]" (HK06, p. 19). This is borne out for results for Model I in the case of $\tau = 0.75$ presented in figure 2 since, for example, at $\alpha = 1$ the critical value of confidence is at about 0.38 when $m = 90\%$ whereas when $m = 10\%$ it is at about 0.3. The reason for this is that when the percentage of truth seekers is larger, this can result in faster convergence to the truth while others are left behind. So it seems that large numbers of truth seekers, not just high α values, can lead to the leaving-behind-effect. In the case of Model II, there is a drop in the critical value ε^* between $m = 90\%$ and 50%, but it changes very little as m decreases to lower values. In general, the results presented in figure 2 for Model II show very little dependence on the percentage of truth seekers or on the value of α.

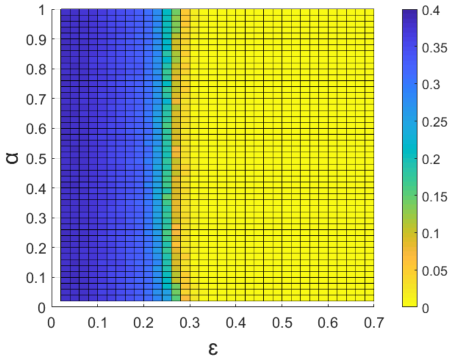
3.1.3 Convergence rates

However, it should be remembered that our focus is on RMSD results. In terms of time taken for the results to converge (i.e. average number of time steps), the results for Model II (and Model I) depend a lot on α and m. For both models, convergence times generally decrease for higher values of both α and m. This is not surprising since, assuming that at least some truth seeking agents converge to the truth, this will occur more quickly if α is high or if m is high since truth seekers will also have an influence on each other. To illustrate this type of dependence, results are presented in figure 3 for both models when the truth is set to $\tau = 0.75$ and $\varepsilon = 0.2$. These results are based on the average convergence time obtained over 1,000 simulations. It should be noted that the results can differ quite a lot in different cases. For example, for high confidence values the difference between the models becomes negligible. In other cases the convergence times can be much greater such as when both α and m are low. Nevertheless, this case provides a good illustration of the dependence on α and m.

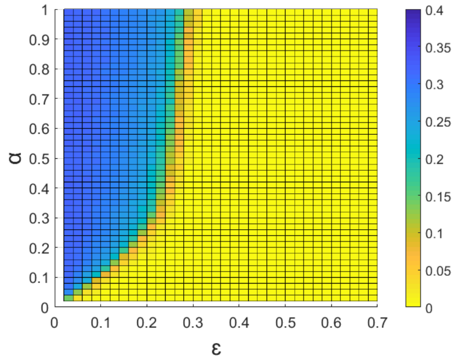
In figure 3(a), α is varied while m is set to 50%. For low α, we see that in this case Model I takes



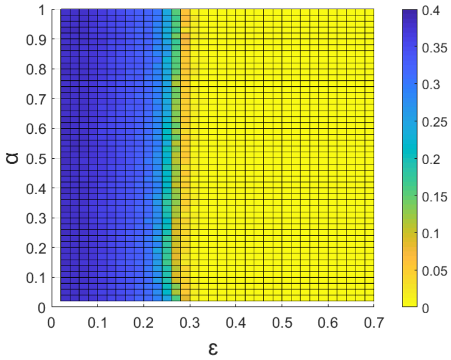
(a) Model I: $m = 10\%$



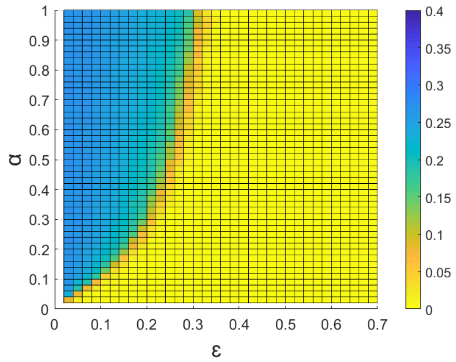
(b) Model II: $m = 10\%$



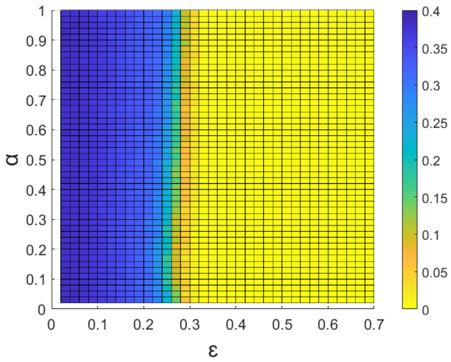
(c) Model I: $m = 30\%$



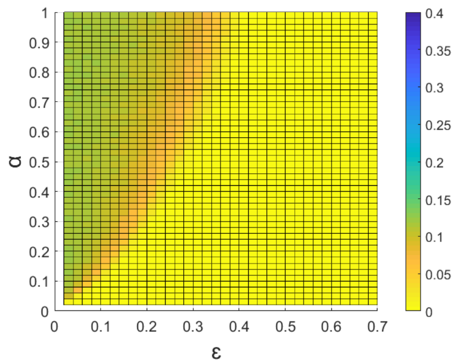
(d) Model II: $m = 30\%$



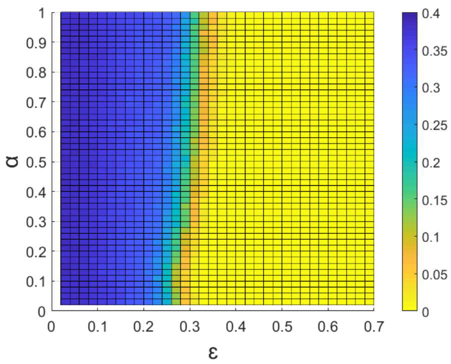
(e) Model I: $m = 50\%$



(f) Model II: $m = 50\%$

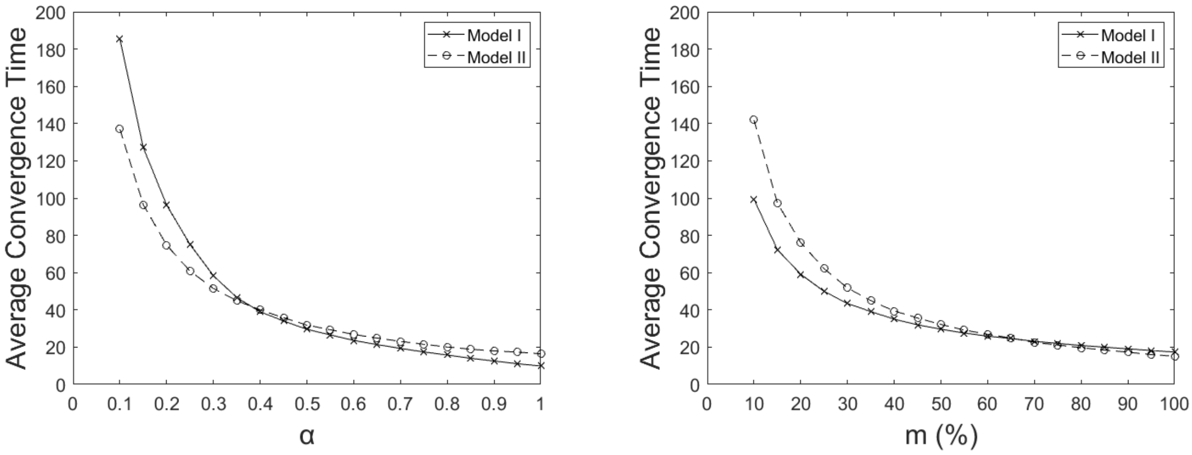


(g) Model I: $m = 90\%$



(h) Model II: $m = 90\%$

Figure 2: RMSD from the truth for Models I and II at varying proportions of truth seekers, m, and $\tau = 0.75$



(a) $m = 50\%$

(b) $\alpha = 0.5$

Figure 3: Convergence rates by α and proportion of truth seekers, m, for Models I and II

longer to converge than Model II. In Model I all truth seekers will eventually converge to the truth, but for low α this can take a long time. In figure 3(b), m is varied while α is set to 0.5. For low m, we see that in this case it is Model II that takes longer to converge. While a significant number of agents converge to the truth in both models in this case, the number of truth seekers within the neighbourhood of the truth is low in Model II and so they have less influence on the normal agents around them than the truth seekers in Model I, hence it takes longer to converge.

We have compared exponential and power law models and found that a power law fits the data better. For example, when varying α in Model I, corresponding to figure 3(a), we obtain the following model for the convergence time, $C(\alpha)$,

$$C(\alpha) = 14.37\,\alpha^{-1.127}, \tag{8}$$

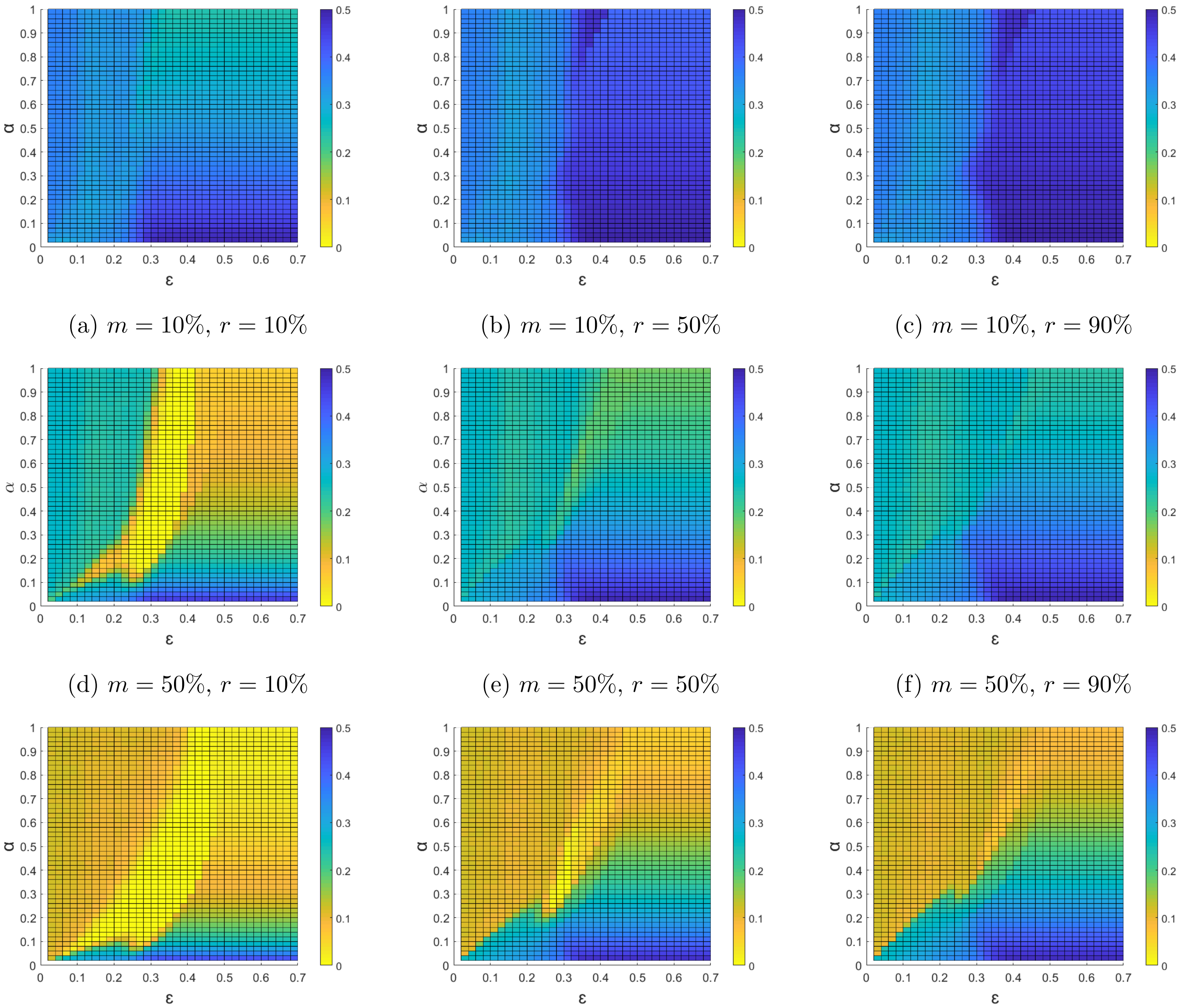
giving $R^2 = 0.9924$. Similarly, when varying m in Model I, corresponding to figure 3(c), we obtain the following model for the convergence time, $C(m)$,

$$C(m) = 17.98\,m^{-0.733}, \tag{9}$$

giving $R^2 = 0.9991$.

3.2 Incorporating truth and a conflicting source

In Models III and IV, we examine truth convergence in the presence of a conflicting source, where the truth is set to 0.75 and the fixed belief of the conflicting source to 0.25. Confidence is varied between 0 and 0.7 and α between 0 and 1. In Model III, we extend Model I to explore how the conflicting source affects truth convergence when truth is accessible to all truth seekers. In Model IV, we extend Model II to look at how truth convergence is affected by the presence of a conflicting source when only those in the neighbourhood of truth are attracted to it.



(a) $m = 10\%$, $r = 10\%$ (b) $m = 10\%$, $r = 50\%$ (c) $m = 10\%$, $r = 90\%$

(d) $m = 50\%$, $r = 10\%$ (e) $m = 50\%$, $r = 50\%$ (f) $m = 50\%$, $r = 90\%$

(g) $m = 90\%$, $r = 10\%$ (h) $m = 90\%$, $r = 50\%$ (i) $m = 90\%$, $r = 90\%$

Figure 4: Impact on truth of varying proportion of truth seekers relative to stubborn group in Model III

3.2.1 Comparison of Models III and IV with varying numbers of truth seekers and stubborn agents

Figures 4 and 5 show the RMSD of the final opinions of the normal and truth seeking agents from the truth for Model III and IV respectively. Recall from section 2 that the conflicting source can be thought of as a group of stubborn agents whose opinion is fixed throughout the simulation. We can therefore represent the influence of this source in terms of the number of stubborn agents, which we represent as a percentage, r, of the number of non-stubborn agents (truth seekers and normal agents), n. Hence, we present results for the following values of both m (percentage of truth seekers) and r (percentage of stubborn agents): $10\%, 50\%$ and 90%. This gives nine sub-figures in each of figures 4 and 5, each of which presents the mean RMSD result as ε is varied between 0 and 0.7 and α between 0 and 1.

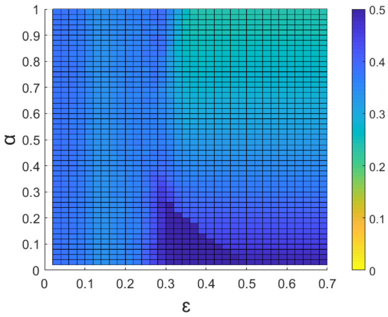
First, compared with the earlier results, it is clear across figures 4 and 5 that truth convergence is greatly reduced in the presence of a conflicting source. With Models I and II, as long as truth was not extreme, full truth convergence was guaranteed when the confidence exceeded around 0.3, and even for an extreme truth value of truth provided confidence exceeded 0.5. With Models III and IV, full convergence is considerably more difficult to achieve. In general, we see that there is no such complete convergence to the truth for high confidence in either figure 4 or 5. Second, as will be explained later, the results are the same for Models III and IV for the different combinations of m and r when the confidence is high, certainly for confidence above 0.5 and in many cases above about 0.4. Third, there is very little difference between the results for the two models when $m = 10\%$ (figures 4(a) - (c) compared to 5(a) - (c)).

Fourth, the results for Model IV vary very little when the confidence is less than about 0.25 for the different combinations of m and r. This is similar to the behaviour of Model II (i.e. in the absence of a conflicting source) as discussed in the context of figure 2. In that case we observed that when $\varepsilon = 0.2$, typically agents converge into two groups, one at the truth and the other some distance from it, with both containing a mixture of truth seekers and non-truth seekers. This is also the case for Model IV except that now the second group converges to the opinion of the conflicting source. Fifth, for Model III truth seekers will converge to the truth for low confidence, but only if α is not too low, otherwise the pull of the crowd and the influence of the conflicting source can prevent some of the truth seekers from getting to the truth. As we shall see below, no truth seekers can converge to the truth if confidence is too high. These findings are in sharp contrast to Model I where truth seekers are guaranteed to converge to the truth.

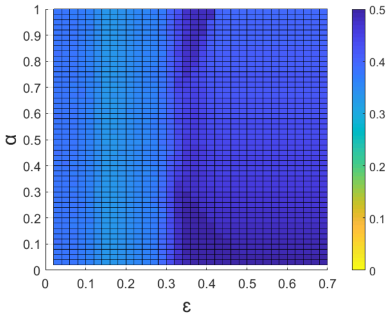
Sixth, a consensus on the truth can occur in both models, but it is only observed in cases where m is greater than r, particularly when $r = 10\%$ and $m = 50\%$ or 90%. It occurs for a limited range of α and ε values, which is greater for Model III than Model IV. As discussed in the context of Models I and II, consensus to the truth depends on the take-along-effect: truth seekers must converge to the truth and bring others along with them. Hence it is not surprising that m needs to be large enough relative to r to ensure truth seekers have enough influence compared to the conflicting source. Also, as already noted, not even all truth seekers will converge to the truth if confidence is too high or α too low. This also occurs for low confidence in Model IV, while for low confidence in Model III if α is too high, the leaving-behind-effect can occur (in figures 4(d), (g) and (h)): truth seekers converge to the truth too rapidly to bring others along with them.

3.2.2 Partitioning of society into truth seekers and non-truth seekers

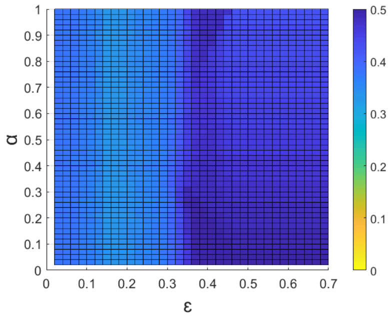
In the standard HK model a consensus is formed when confidence is high and we have seen from the results for Models I and II that a consensus on the truth is formed for high confidence. However, it turns out that no such consensus forms in Models III or IV, either to the truth or elsewhere for $\varepsilon > 0.5$, provided $m > 0$ and $r > 0$. Clearly, such a consensus could not converge to any opinion greater than $\tau = 0.75$ or less than that of the conflicting source, $\psi = 0.25$, since in the former case truth seeking agents would be influenced by the truth and in the latter case all agents would be influenced by the conflicting source. Similarly, no consensus could form at either 0.75 or 0.25 since $\varepsilon > 0.5$ would ensure in the former case that all agents were influenced by the conflicting source and in the latter case that truth



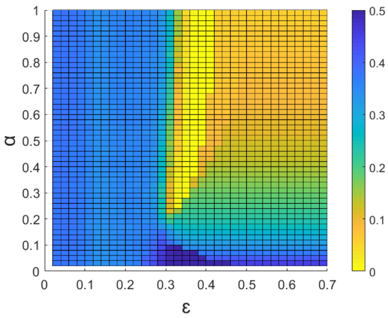
(a) $m = 10\%$, $r = 10\%$



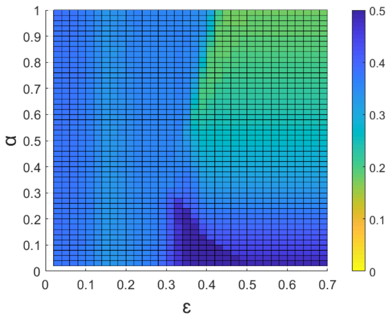
(b) $m = 10\%$, $r = 50\%$



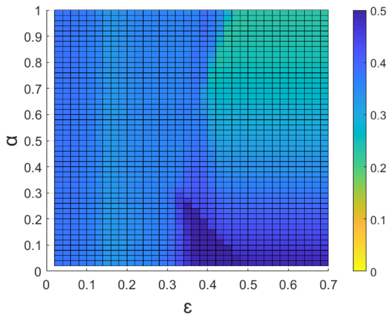
(c) $m = 10\%$, $r = 90\%$



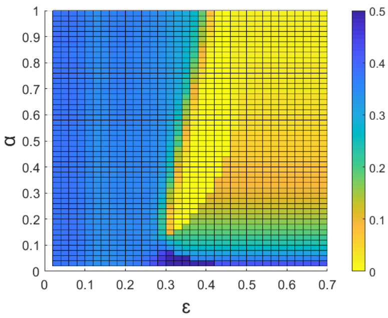
(d) $m = 50\%$, $r = 10\%$



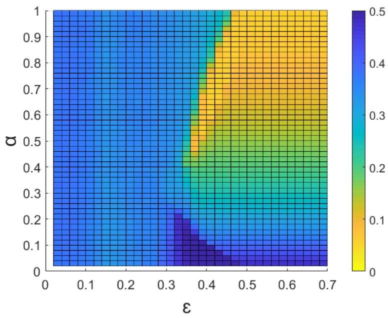
(e) $m = 50\%$, $r = 50\%$



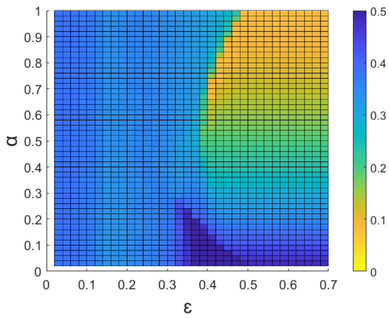
(f) $m = 50\%$, $r = 90\%$



(g) $m = 90\%$, $r = 10\%$



(h) $m = 90\%$, $r = 50\%$



(i) $m = 90\%$, $r = 90\%$

Figure 5: Impact on truth of varying proportion of truth seekers relative to stubborn group for Model IV

seekers would be influenced by the truth. Could a consensus form in between 0.25 and 0.75? No, while truth seekers would be influenced by both truth and the conflicting source, non-truth seekers would only be influenced by the conflicting source, drawing them away from the supposed consensus. So a consensus for $\varepsilon > 0.5$ is impossible.

Since all agents would be expected to converge to an opinion between 0.25 and 0.75, all truth seekers must converge to a single opinion, x^* say, since they will be subject to the same influences and all non-truth seekers must similarly converge to a single opinion, y^* say, since they too will be subject to the same influences. Hence, there will be a partition of truth seekers and non-truth seekers into two separate groups. In such a scenario Models III and IV will be equivalent in terms of their long term behaviour since all truth seekers will converge to an opinion within the confidence of truth and from equation (4) we can see that x^* and y^* will be related as follows

$$x^* = \alpha\tau + (1-\alpha)x^{MHK*} \tag{10}$$

and

$$y^* = \frac{mx^* + (n-m)y^* + r\psi}{n+r}, \tag{11}$$

with

$$x^{MHK*} = y^* \tag{12}$$

where n is the total number of agents (truth seekers and normal agents), m is the number of truth seekers, r represents the reputation or weight of the conflicting source, which can be thought of as a number of stubborn agents, and ψ represents the fixed value of the conflicting source. It is straightforward to show that the solution to these equations is given by

$$x^* = \frac{\alpha(m+r)\tau + (1-\alpha)r\psi}{\alpha m + r} \tag{13}$$

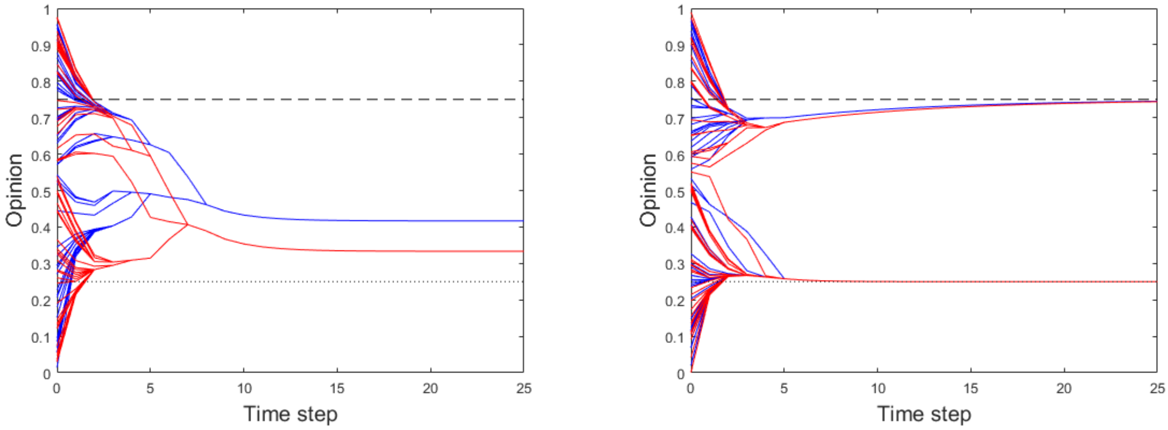
and

$$y^* = \frac{mx^* + r\psi}{m+r}. \tag{14}$$

Note that this result means that Models III and IV have exactly the same long term behaviour for $\varepsilon > 0.5$, forming a partition between truth seekers and non-truth seekers. Also, it shows that even in Model III, truth seekers cannot converge to the truth unless $\alpha = 1$ or else $r = 0$.

From equations (13) and (14), it is easy to see that as either $r \to 0$ or $\alpha \to 1$, then $x^* \to \tau$, while y^* tends to τ in the former case or a weighting of τ and ψ in the latter. Similarly, when either $\alpha \to 0$ or $m \to 0$, then $y^* \to \psi$, while x^* tends to ψ in the former case or a weighting of τ and ψ in the latter. These findings are clearly reflected in figures 4 and 5 at high confidence levels where the average distance from truth decreases with increasing α, increasing m and decreasing r, but is independent of the confidence, ε. Note in particular that in figures 4(g) and 5(g), where $m = 90\%$ and $r = 10\%$, the average distance from the truth is very low when α is close to 1. This is because both x^* and y^* will be close to τ in this case.

In many cases, the same partition of truth seekers and non-truth seekers can occur at lower confidence values as well, although this depends on the various parameters and can differ between Models III and IV.



(a) Model III

(b) Model IV

Figure 6: Examples of individual simulations for Models III and IV when $\varepsilon = 0.3$, $\alpha = 0.2$ and $m = r = 50\%$. Results are presented for truth seekers (—) and non-truth seekers (—). The value of truth (- - -) and the conflicting source (·····) are also shown.

For example, the partition still applies in Model III, as shown in figure 6(a) when $\varepsilon = 0.3$, $m = r = 50\%$ and $\alpha = 0.2$ giving $x^* = 5/12$ and $y^* = 1/3$, whereas this is not the case with Model IV. An example of the behaviour that can occur in Model IV with these parameter values is shown in figure 6(b) where there is a partition in which about half the agents (truth and non-truth seekers) converge to the truth and the rest converge to the opinion of the conflicting source.

Furthermore, while the argument for a partition into truth seekers and non-truth seekers and the related analytical results were based on $\tau = 0.75$ and $\psi = 0.25$, corresponding results can be derived in many other cases. For example, the same result can be derived for any case where $|\tau - \psi| = 0.5$. Similarly, if τ and ψ are closer to the central value of 0.5, a partition will occur for lower values of confidence. For example, if $\tau = 0.7$ and $\psi = 0.3$, then the same result will hold if $\varepsilon > 0.4$. Alternatively, if τ and ψ are further from 0.5, a partition will be guaranteed for higher values of confidence. For example, if $\tau = 0.8$ and $\psi = 0.2$, then the same result will hold if $\varepsilon > 0.6$. Hence, there are a wide range of cases where the society will partition into truth seekers and non-truth seekers as expressed in equations (13) and (14).

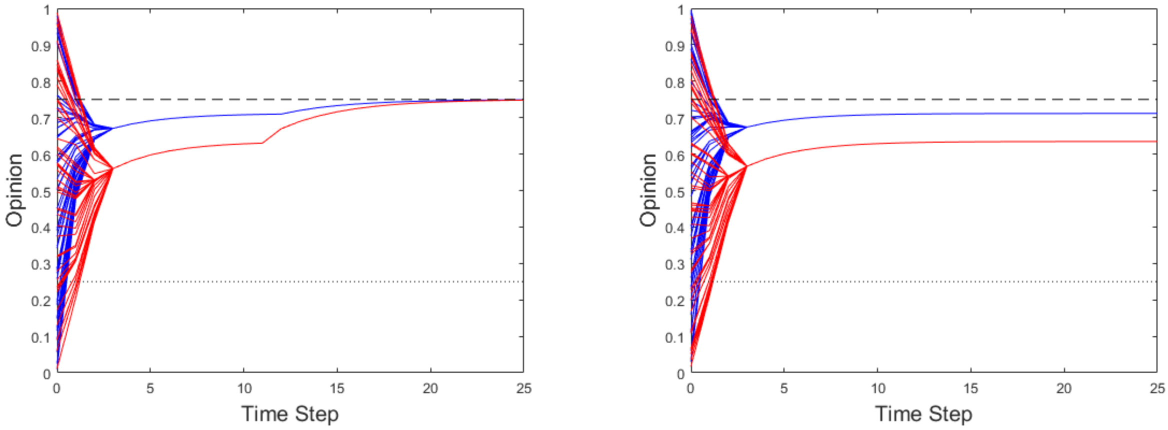
Indeed there are other scenarios where a partition into the two groups can occur in a different way. The derivation of equations (13) and (14) required that x^* was within the confidence of the truth, y^* and the conflicting source, while y^* was within the confidence of x^* and the conflicting source. In some cases, a partition can occur in a different way for range of lower confidence values where x^* is not within the confidence of the conflicting source, though y^* still is. In such a case, equation (11) remains the same but (12) can be replaced by

$$x^{MHK*} = \frac{mx^* + (n-m)y^*}{n} \tag{15}$$

which gives the following solutions

$$x^* = \frac{\alpha n(m+r)\tau + (1-\alpha)(n-m)r\psi}{n(m+r) - (1-\alpha)m(n+r)} \tag{16}$$

and y^* is given by equation (14) as before.



(a) $\varepsilon = 0.38$ (b) $\varepsilon = 0.39$

Figure 7: Examples of individual simulations for Models III where $\alpha = 0.5$, $m = 50\%$ and $r = 10\%$. Results are presented for truth seekers (—) and non-truth seekers (—). The value of truth (- - -) and the conflicting source (·····) are also shown.

For example, consider the case in Model III where $m = 50\%$, $r = 10\%$ and $\alpha = 0.5$ in figure 4(d). For values of confidence from about 0.43 and higher, the earlier partition into truth seekers and non-truth seekers at the values given by equations (13) and (14) occurs. For values of confidence between 0.39 and 0.42, the alternative mechanism for creating a partition occurs and so the truth seekers converge to an x^* value given by equation (16). In this case, $x^* = 0.7115$ and $y^* = 0.6346$, which means that when $\varepsilon = 0.39$, y^* is just within the confidence of the conflicting source. For lower values of confidence, down to about 0.28, the influence of the truth seekers is sufficient to overcome that of the conflicting source on the non-truth seekers and a consensus is formed at the truth. Figure 7 illustrates these differences in behaviour. In figure 7(a), when $\varepsilon = 0.38$, all truth seekers and non-truth seekers converge to the truth, but in figure 7(b), when $\varepsilon = 0.39$, there is a partition into truth seekers and non-truth seekers. For values lower than about 0.28, a partition into two groups becomes more dominant, with one group containing all the truth seekers and some non-truth seekers converging to the truth, while the other group converges to the conflicting source at 0.25.

4 Conclusions

We have investigated social learning in both the presence and absence of a conflicting source and with two different ways of modelling the influence of the truth. To do this, we have used four variations of the HK bounded confidence model of opinion dynamics. In Model I, truth is accessible to all truth seekers, while in Model II, truth only has an influence on truth seekers whose opinions are close to the truth. Models III and IV extend each of these models by including the influence of a conflicting source on the opinion dynamics.

In the absence of a conflicting source, there are some similarities between the results obtained for Models I and II. In particular, a consensus on the truth occurs in both models provided the confidence is sufficiently high and fails to occur in both models if the confidence is low and the influence of the truth is high. There are also some significant differences between the models. In Model I, all truth seekers are guaranteed to converge to the truth, whereas in Model II even truth seekers who are close to the truth can fail to reach the truth as they are pulled away by the influence of the crowd. Furthermore, increasing the proportion of truth seekers results in the society as a whole moving closer to the truth in Model I, but has little effect in Model II.

In the presence of a conflicting source, both approaches to modelling the influence of truth (Models III and IV) yield the same results when the confidence is high, giving the same partition of the society into truth seekers and non-truth seekers. However, the presence of a conflicting source, no matter how weak, ensures that even the truth seekers do not converge to the truth in either model if the confidence is high. Nevertheless, both models can lead to a consensus on the truth in some scenarios. In particular, for both models such a consensus requires the proportion of truth seekers to be sufficiently high relative to the influence of the conflicting source and also requires the confidence and influence of the truth to lie within certain ranges. The biggest difference between the two models is that in Model III the society as a whole ends up closer to the truth when confidence is low than it does in Model IV.

The paper has several implications for various real world contexts. It is clear from comparing Models I and II that the scope for social learning is enhanced by making the truth more accessible to all rather than a privilege of the few. For example, in a public health campaign, it highlights the importance of clear and simple communication of a potentially complex message. Related to this, social learning can be enhanced if there are more truth seekers within the society and so, for example, a health campaign could benefit from stimulating interest and emphasising the relevance of the message to those who might otherwise ignore it. The results also highlight the importance of being aware of how people might be influenced by other views within society. This is particularly relevant in instances where social learning in a society may be hampered by the presence of a conflicting source which stands opposed to the truth as in Models III and IV. This could apply even in cases where the truth has a firm evidence base, for example in medical or scientific research or in well-established government policy. The results demonstrate the importance of taking account of a conflicting source where it exists to minimise the adoption of false beliefs and lead to effective social learning. Given the ability of even a relatively weak conflicting source to prevent social learning in some cases, the results suggest that ignoring such sources and just concentrating on making a positive case may often be inadequate. Instead, false beliefs will often need to be rebutted directly. Further discussion of optimal strategies for maximising social influence in the context of competing opinions is provided in [22], but there the focus was not on truth and hence did not take into account the role of truth seekers. However, stimulating interest in seeking the truth could have an important role to play in many cases, especially given the tendency for society to partition into truth seeking and non-truth seeking groups as the results in section 3.2.2 show. There is a lot of scope for exploring these implications further in the context of scientific, ethical and social issues in a world where social media and fake news have a lot of influence. Towards this end, it would be interesting to explore the possibilities for applying the models considered here to real data available from social media. This is a challenge for research in opinion dynamics more generally as noted in [17], but one possibility might be to draw on work that infers opinions from users' actions [25, 26].

There are also a number of other possible directions for further research. For example, it would be interesting to compare the results here with corresponding models based on the Deffuant-Weisbuch (DW) model of opinion dynamics [16]. Like the HK model, the standard version of the DW model reaches a consensus for sufficiently high confidence levels ($\varepsilon > 1/2$ in the case of DW) and so it might be expected to reach a consensus to the truth in a wide range of cases corresponding to Models I and II. The more interesting case concerns the inclusion of a conflicting source where it would be interesting to see whether the DW model also results in similar partitions of the agents into truth seekers and non-truth seekers. It would also be worthwhile exploring alternative modelling assumptions. For example, the influence of allowing heterogeneous confidence values [19, 27–30] or dropping the assumptions that confidence with respect to the conflicting source in Models III and IV should be the same as the confidence in general and similarly for the confidence of truth seekers with respect to the truth in Models II and IV [21] could be explored. In terms of the respective influences of truth and the conflicting source, we have considered the case where the truth is at $\tau = 0.75$ and the conflicting source at $\psi = 0.25$, so it would be interesting to explore what happens when the conflicting source is closer or further away from the truth. Nevertheless, the simulations and analytical results presented here give a good sense of the kinds of behaviour that could be expected in a range of scenarios.

References

[1] A. V. Banerjee, "A simple model of herd behavior," *The Quarterly Journal of Economics*, vol. 107, no. 3, pp. 797–817, 1992.

[2] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of Political Economy*, vol. 100, no. 5, pp. 992–1026, 1992.

[3] D. Acemoglu and A. Ozdaglar, "Opinion dynamics and learning in social networks," *Dynamic Games and Applications*, vol. 1, no. 1, pp. 3–49, 2011.

[4] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi, "Non-Bayesian social learning," *Games and Economic Behavior*, vol. 76, no. 1, pp. 210 – 225, 2012.

[5] P. Molavi, A. Tahbaz-Salehi, and A. Jadbabaie, "A theory of non-Bayesian social learning," *Econometrica*, vol. 86, no. 2, pp. 445–490, 2018.

[6] R. Hegselmann and U. Krause, "Truth and cognitive division of labour first steps towards a computer aided social epistemology," *Journal of Artificial Societies and Social Simulation*, vol. 9, no. 3, 2006. http://jasss.soc.surrey.ac.uk/9/3/1.html.

[7] K. Malarz, "Truth seekers in opinion dynamics models," *International Journal of Modern Physics C*, vol. 17, pp. 1521–1524, 2006.

[8] I. Douven and C. Kelp, "Truth approximation, social epistemology, and opinion dynamics," *Erkenntnis*, vol. 75, no. 2, p. 271, 2011.

[9] N. E. Friedkin and F. Bullo, "How truth wins in opinion dynamics along issue sequences," *Proceedings of the National Academy of Sciences*, vol. 114, no. 43, pp. 11380–11385, 2017.

[10] Z.-X. Tan and K. Cheong, "Cross-issue solidarity and truth convergence in opinion dynamics," *Journal of Physics A: Mathematical and Theoretical*, vol. 51, 2017.

[11] I. Douven and S. Wenmackers, "Inference to the best explanation versus Bayes's rule in a social setting," *The British Journal for the Philosophy of Science*, vol. 68, no. 2, pp. 535–570, 2017.

[12] Y. Liu and M. Lipo, "Noise-induced truth seeking of heterogeneous Hegselmann-Krause opinion dynamics," *Advances in Mathematical Physics*, vol. 2018, pp. 1–6, 2018.

[13] W. Su and Y. Yu, "Free information flow benefits truth seeking," *Journal of Systems Science and Complexity*, vol. 31, no. 4, pp. 964–974, 2018.

[14] I. Douven, "Optimizing group learning: An evolutionary computing approach," *Artificial Intelligence*, vol. 275, pp. 235 – 251, 2019.

[15] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence: models, analysis and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, 2002. http://jasss.soc.surrey.ac.uk/5/3/2.html.

[16] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, "Mixing beliefs among interacting agents," *Advances in Complex Systems*, vol. 3, pp. 87–98, 2000.

[17] Y. Dong, M. Zhan, G. Kou, Z. Ding, and H. Liang, "A survey on the fusion process in opinion dynamics," *Information Fusion*, vol. 43, pp. 57–65, 2018.

[18] R. Hegselmann and U. Krause, "Opinion dynamics under the influence of radical groups, charismatic leaders, and other constant signals: A simple unifying model," *Networks and Heterogeneous Media*, vol. 10, pp. 477–509, 2015.

[19] M. Pineda and G. Buendía, "Mass media and heterogeneous bounds of confidence in continuous opinion dynamics," *Physica A: Statistical Mechanics and its Applications*, vol. 420, pp. 73 – 84, 2015.

[20] H. Z. Brooks and M. A. Porter, "A model for the influence of media on the ideology of content in online social networks," *Physical Review Research*, vol. 2, no. 2, p. 023041, 2020.

[21] S. Chen, D. H. Glass, and M. McCartney, "Characteristics of successful opinion leaders in a bounded confidence model," *Physica A: Statistical Mechanics and its Applications*, vol. 449, pp. 426–436, 2016.

[22] C. A. Glass and D. H. Glass, "Social influence of competing groups and leaders in opinion dynamics," *Computational Economics*, 2020. http://dx.doi.org/10.1007/s10614-020-10049-7.

[23] S. Chen, D. H. Glass, and M. McCartney, "How opinion leaders affect others on seeking truth in a bounded confidence model," *Symmetry*, vol. 12, no. 8, p. 1362, 2020.

[24] I. Douven and A. Riegler, "Extending the Hegselmann-Krause model I," *The Logic Journal of the IGPL*, vol. 18, no. 2, pp. 323–335, 2010.

[25] C. Monti, G. De Francisci Morales, and F. Bonchi, "Learning opinion dynamics from social traces," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '20, (New York, NY, USA), p. 764–773, Association for Computing Machinery, 2020.

[26] T. Tang and C. G. Chorus, "Learning opinions by observing actions: Simulation of opinion dynamics using an action-opinion inference model," *Journal of Artificial Societies and Social Simulation*, vol. 22, no. 3, p. 2, 2019.

[27] G. Fu, W. Zhang, and Z. Li, "Opinion dynamics of modified Hegselmann–Krause model in a group-based population with heterogeneous bounded confidence," *Physica A: Statistical Mechanics and its Applications*, vol. 419, pp. 558 – 565, 2015.

[28] G. Chen, W. Su, S. Ding, and Y. Hong, "Heterogeneous Hegselmann-Krause dynamics with environment and communication noise," *IEEE Transactions on Automatic Control*, 2019.

[29] W. Su, Y. Gu, S. Wang, and Y. Yu, "Partial convergence of heterogeneous Hegselmann-Krause opinion dynamics," *Science China Technological Sciences*, vol. 60, no. 9, pp. 1433–1438, 2017.

[30] H. Schawe and L. Hernández, "When open mindedness hinders consensus," *Scientific reports*, vol. 10, no. 1, pp. 1–9, 2020.