CS 261 Final Project : Hegselmann-Krause Confidence Bound

Kyle D'Souza, Daniel Guillen, and Claire Mai

Abstract

In this paper, we explore the Hegselmann-Krause (HK) system, which is used to model opinion dynamics (8). We outline the dynamical system and present a literature survey of related research. We later turn to our general focus on measuring the relationship of confidence bounds on the number of converged clusters in the HK 1-D model. We present an algorithm[1], which given a random initialization of agents, solves for the minimum confidence bound ϵ such that agents deterministically converge to a target number of clusters, for which we simulate for a target number 1, 2, and 3. We also map a relationship between the algorithm's determined minimum confidence bound and target converged cluster size. We hope this work can be a strong start to further research and exploration, one that will only gain momentum as researchers from sociology to computer science try to better understand consensus in HK models and ultimately aim to better map the effect of agent influence on community opinions.

1 Introduction

Network-based dynamical systems have received increasing attention as mathematical models have improved in explanatory power (19). In particular, opinion dynamics mesh together sociophysics and social dynamics, employing computational and mathematical models to investigate opinion spread across communities (21). This work applies to any field that aims to track the dynamics of influence. Computer simulations tend to provide approximations for observed opinion spread phenomenon; however, rigorous results remain rare. Nevertheless, over the past few decades, the Hegselmann-Krause confidence bound model (abbreviated as HK-model from now on) has become an exception to this rule. The

[1] Code accessible at `https://github.com/kvdsouza/hkconfidencebounds`

paper has been cited over 2700 times across various disciplines due to its simplicity and accessibility (8).

The HK-model, as described in section 2, has agents update their locations over time until convergence. The model has a linear update rule, preserves order of agent positions, and has agents act synchronously and in a completely deterministic manner (19). However, the length of time for HK-stabilization has been the subject of much debate, with researchers originally positing a convergence time bounded by $n^{O(n)}$ and conjectured to be polynomial (4). Over the next decade, researchers proceeded to prove the upper-bound convergence time to be $O(n^5)$ for dimension $= 1$ and then proved the upper bound to be $O(n^3)$, before most recently proving a quadratic lower bound of $O(n^2)$. We want to take this analysis in a slightly different direction by exploring the impact of confidence bounds on the number of converged clusters.

In this paper we will first provide some preliminaries and some definitions to better understand the HK system. Next, we outline past research, current research, and future research directions for this problem in the form of a Literature Review. Finally, we will find an algorithm that calculates the minimum confidence bound needed for a random distribution of agents to converge to a target number of clusters.

2 Preliminaries

The HK-model works as follows: Given a finite number of n agents, for every time step $t = \{1, 2, 3, ...\}$, and for each $i \in [n]$, the position of agent i is measured by $x_t(i) \in \mathbb{R}$. There is a fixed parameter ϵ that represents the confidence bound — this confidence bound represents the maximum distance at any time step that an agent will consider as their neighborhood. We represent ϵ as a positive, non-zero value, such that $\mathcal{N}_t(i) = \{j : |x_t(j) - x_t(i)| \leq \epsilon\}$, and

$$x_{t+1}(i) = \frac{1}{|\mathcal{N}_t(i)|} \sum_{j \in \mathcal{N}_t(i)} x_t(j).$$

Thus, at the each time step, the agent updates its position according to the average opinions of its neighbors (i.e. those that are within its confidence bound ϵ). The algorithm converges at time t if $x_t(i) = x_{t+1}(i)$ for all $i \in [n]$ (1).

Below, we define terms that will be used throughout the paper.

2.1 Clusters

A *cluster* is defined as a maximal set of agents that agree. Two agents are defined as in agreement if $x(i) = x(j)$ at the point of convergence. The size of a cluster is equivalent to the number of agents in that cluster at convergence. We define a target cluster as the maximal number of possible clusters at convergence. In our algorithm, we set a target cluster size to derive our minimum confidence bound. Based on our algorithmic procedure, we consistently find that the eventual number of clusters at convergence is equivalent to the target cluster size, as desired.

2.2 Consensus & Convergence

Consensus and more generally, the final configuration of the HK model, empirically depends on the initial positions as well as the confidence bounds used. A configuration is defined as converged (or frozen) if $|x(i) - x(j)| > 1 \leftrightarrow x(i) \neq x(j)$. As with any convergence, in the long run $x_t(i) = x_{t+1}(i)$. As we have touched on above, the configuration has been shown to freeze after time $O(n^2)$ in the lower bound in the 1-D case (19).

Notably, with regard to configuration, in the 1-D case, the final configuration will always be ordered in the same format as the initialized configuration (i.e. for all $i, j \in [n]$, if $x_0(i) \leq x_0(j)$, then $x_t(i) \leq x_t(j)$ for all $t > 0$).

3 Literature Survey

Below we break down the relevant related research around HK systems into three sections:

1. Past research, tracking the first 15 years of research on Hegselmann-Krause (1997-2012);

2. Current research, tracking the past decade of research;

3. Future research, cataloging a list of related open questions.

3.1 *Past Research*

HK systems are a type of discrete-time, linear dynamical model that was initially introduced by Ulrich Krause in 1997 to study opinion dynamics and opinion formation (9). These models later came to be known under the Hegselmann-Krause (HK) moniker after paper (8), where the authors explored the model under the context of opinion formation and established some initial analytical convergence results (under certain conditions).

HK systems have since attracted significant attention and emerged as a research platform to study social dynamics and collective agent behavior in a variety of fields including physics, mathematics, computer science, among others. In HK models, agents hold a set of continuous opinions which they can gradually adjust if they hear the opinions of others. Examples of possible opinion dimensions include political issues, prices, tax rates, etc.; however, these examples are not necessarily limited to purely social science topics.

HK systems incorporate the idea of confidence bounds, specifically the idea that agents will only influence, or be influenced, by other agents if they are close in opinion to each other. In a real life corollary, one's opinions are often more successfully influenced by those who share opinions that share a relevant perspective or background, due to both social proximity as well as similar perspective. The central premise of HK models is that agents hold continuous opinions that change over time given their interaction with others, and the confidence bound principle of the HK model incorporates the premise that agents only interact with others who have relatively close or similar opinions. Moreover, each agent in the system can either have the same confidence bound (called a homogeneous system) or can have varying confidence bounds (in a heterogenous system) (12). Both variations are subjects of study with most of the current convergence and running time results achieved around homogeneous systems.

Several open problems remained unanswered in the early years, from understanding convergence guarantees and running time guarantees, to the role of initial conditions. These topics remained unexplained until around 2007 (12).

Some of the early results to tackle these problems around HK systems included showing that HK systems converge to some final configuration when the confidence bound is homogeneous (10) (7). Similarly, additional results included proving the type of final opinion profile that a homogenous HK system converges to — i.e. for any pair of two opinions x_i, x_j they will either belong to the same cluster (i.e. be equal) or be separated by a distance larger than the bounded confidence parameter (12).

In 2007, the same HK dynamical system was explored under the context of robotic coordination algorithms, where the authors proved an upper bound of $O(n^5)$ on the time complexity of the "move-towards-average law" for the 1-dimensional agent case (14). In 2012, the convergence time was improved to $O(n^4)$ for the 1-dimensional agent case (i.e. one singular continuous opinion) (17).

Still, multiple questions remained open at the end of the 2000's (11). Firstly, the task of understanding the impact of initial conditions and distribution on the final outcomes, whether that be the number of equilibrium clusters, cluster locations, and convergence times, remained unanswered. Moreover, the task of predicting equilibrium clusters given initial configuration profiles remained and still remains an exciting area of research. Questions around multidimensional opinion results were raised, from final opinion profiles to running time guarantees for $d > 1$ opinion dimensions, as well as analysis into the role of dimensionality on convergence. Researchers also began exploring heterogenous confidence bound results, namely understanding drifting clusters as well as better grappling with consensus and final opinion profiles (18).

Most of the work done on HK systems have focused on the homogeneous case, with even more focused on the 1-dimensional case. One of the more commonly observed by unexplained phenomenon revolves around the $2R$ conjecture (2). The conjecture states that given enough agents and time, an HK system will evolve into a series of clusters such that each cluster is separated by distances roughly equal to 2R where R is the homogeneous bound. We know by paper (12) that clusters need to be separated by at least the confidence bound but the empirically observed phenomenon of converged clusters being separated by about double the confidence bound remains unexplained. Blondel, Hendricx and Tsitsiklis presented another proof of convergence as well as a discussion of some empirical observations around the 2R conjecture in 2009 (3). This paper proposes one analytical attempt to try to develop a tighter lower bound on the cluster distances.

Finally, from early on, the research community spent time developing understanding observed HK system phenomenon. In 2010, Lorenz found that systems with both "closed" and "open" minded agents (i.e. agents with high and low confidence bounds) could converge to consensus even when both bounds of confidence are relatively low (13). The paper also observed the "drift" phenomenon, where indicating that cluster drift may be a commonly occurring component of heterogeneous systems (13). Open questions that Lorenz left unanswered included better understanding what conditions in a heterogeneous system lead to consensus along the central opinion, and how extreme splits or drifting clusters toward one extreme came into existence.

3.2 *Current Research*

After discovering HK-systems and their use on opinion dynamics, researchers began to generalize the idea of opinion dynamical systems with authors introducing ideas around influence systems (5). Influence systems are an attempt to unify the different varieties and modifications of HK systems (and more generally, time-dependent, agent dynamics systems) that are used in social dynamics and collective agent behavior research under a single framework and calculus. More specifically, one paper took a look at diffusive influence systems (5). In paper (5), diffusive is defined as an agent only capable of moving within the convex hull of its neighbors for each time step. Diffusive influence systems are crucial to social dynamics to a certain extent because they extend the fundamental concept of diffusion to autonomous agents operating within dynamic, heterogeneous environments (5). This has led to a new finding that heterogeneous bounded-confidence models are periodic.

As part of this new path of generalizing the idea of opinion dynamical systems, one paper studied the total s-energy of multi-agent systems with time dependencies proving a general theorem (1.4) on communication count that is used to evaluate / prove an upper bound on the convergence time for a bounded-confidence HK model over any dimension (4). This established an upper bound running time (over any dimension, not just d=1) of HK systems to $O(n^{(O(n))})$.

Some of the ideas around s-energy were used when developing the influence system literature (5) — authors noted that by "restricting ourselves to influence systems, the bidirectional kind are known to be attracted to a fixed point while expending a total s-energy at most exponential in the number of agents and polynomial in the reversible case" — still convergence times for influence systems are still only known in the smaller cases like that of HK systems.

However, lately, researchers have developed more results around convergence and running times for both heterogeneous and homogeneous HK systems. In 2012, Chazelle showed that a high generality of their results on diffusive influence systems can be applied to heterogenous HK. In 2013, Bhattacharyya et al. improved the upper bound for HK system convergence time from $O(n^{O(n)})$ to polynomial in both n and d while also improving the lower bound for the 1-dimensional HK case to $O(n^3)$ (1). Bhattacharyya et al. were able to achieve these results by injecting a geometric perspective to the current technology for HK systems. This is done by having n agents initially located at vertices of a regular n-gon with side lengths of 1. They geometrically used polygons to create a lower bound algorithm requiring at least $\frac{n^2}{28}$ steps to converge (1). In 2015, authors proved a lower-bound on the convergence running time for the 1-dimensional, homogeneous case (20). In 2016, Chazelle & Wang proved that the heterogeneous HK system with closed-minded agents (where agents that have a confidence bound of either 1 or 0) converges by evaluating the s-energy, where s=2, for a more general class of dynamical system called an inertial HK system (6). This new method sheds light on the convergence properties of inertial HK systems where instead of being forced to move to the mass center of its neighbors for each step, each agent can move towards the mass center by any fraction of length.

As of now, most of the research around HK systems has been geared towards proving convergence for various models, and bounding the converge time. In 2013, Yilun Shang derived a critical threshold for the Deffuant model where the opinions converge toward the average value of the initial opinion distribution with probability one (16). However, Shang also tested numerous simulations on different using probability distributions such as uniform, beta, power-law and normal distributions as the initial agent opinions. The lack of research done on the impact of initial conditions on HK models indicates a field of potential research, along with the impact of various confidence bounds, and the bounds of the number of converged clusters.

3.3 *Future Research*

Future research can explore possible bounds for convergence on heterogeneous HK system. In 2016, Chazelle & Wang introduced an anchored variant of HK systems and proved that the variable was equivalent to the symmetric heterogeneous model and that they both converge asymptotically to a fixed configuration (20). However, they never produce a lower or upper bound to the convergence of symmetric heterogeneous HK models. Furthermore, many papers look at HK models, like symmetric heterogenous HK models, that have undirected graphs or networks. Chazelle's earlier 2012 paper exposed a huge gap in the expressivity of directed and undirected dynamic networks (5). While undirected networks have already been proven to lead to a stable agreement, directed graphs are a much more complex environment.

Similar to what current research has been exploring, future research could also investigate tighter bounds of convergence for HK systems in higher dimensions and in the 1-D case. However, while much of the research focuses on convergence times and upper and lower bounds for various HK models, little research has conducted on figuring out the upper or lower bounds on the number of converged clusters. Smaller questions can be asked regarding the confidence bound, where given a point set, what set of characteristics are needed for all agents to converge to exactly one cluster.

4 Simulation

4.1 *Problem statement*

We wish to explore the relationship between confidence bounds and the number of converged clusters. Specifically, we present and explore an algorithmic approach to determine the minimum confidence bound necessary so that a given set of agents will converge to a target number of clusters.

4.2 *Algorithm*

Our algorithm takes a binary-search approach to identifying the minimum confidence bound required. But, before the core component runs, we first normalize the set of given agents into the range 0-to-1 in order to have a starting lower and upper bound on the confidence bound — specifically 0 is a strict lower bound on the minimum confidence bound (since with a confidence bound of 0, the agents never change) and 1 is an inclusive upper bound on the minimum confidence bound (since the entire normalized agent / opinion space is a distance of 1). From then, we run a binary-search procedure where at each iteration, we determine the midpoint between the strict lower bound and inclusive upper bound and then simulate the given set of agents to convergence using the midpoint confidence bound value — after converging, we determine the number of clusters. If the number of clusters is less than or equal to the target number, we decrease our inclusive upper bound to the midpoint, and if the the number of clusters is strictly greater than the target number, we increase the strict lower-bound. This procedure repeats until the lower and upper-bounds converge and we identify the minimum confidence bound value.

Listing 1: Algorithm Pseudocode

```
\\ Finds the minimum epsilon / confidence bound that leads to a consensus convergence.
\\ Assumes agents are within the 0-to-1 range.
def findMinConfidenceBound(agents, targetClusters):
    high = 1.0
    low = 0.0
    while low is not equal to high (or high - low >= some tolerance):
        mid = (high + low) / 2
        numClusters = NumberOfClusters(SimulateToConvergence(agents, eps=mid)
        if numClusters <= targetClusters:
            high = mid
        else:
            low = mid
    return low, high

def findMinConfidenceBound(agents, targetNumClusters):
    \\ First, we normalize the agents to the 0-to-1 range.
    agentsNorm = [elem / max(agents) for elem in agents]

    \\ Find the solution confidence bound for the normalized agents.
    lowNorm, highNorm = findMinConfidenceBoundNormalized(agentsNorm, targetNumClusters)

    \\ Un-normalize the determined confidence bound
    low, high = max(agents) * lowNorm, max(agents) * highNorm

    return low, high
```

In our actual implementation, we determine convergence for either the HK system of agents or the minimum confidence bound's bounds if the difference in changes is below some tolerance and the number of clusters is evaluated by counting the number of agents that have the same value (difference is below some tolerance) — see below for code implementation.[2]

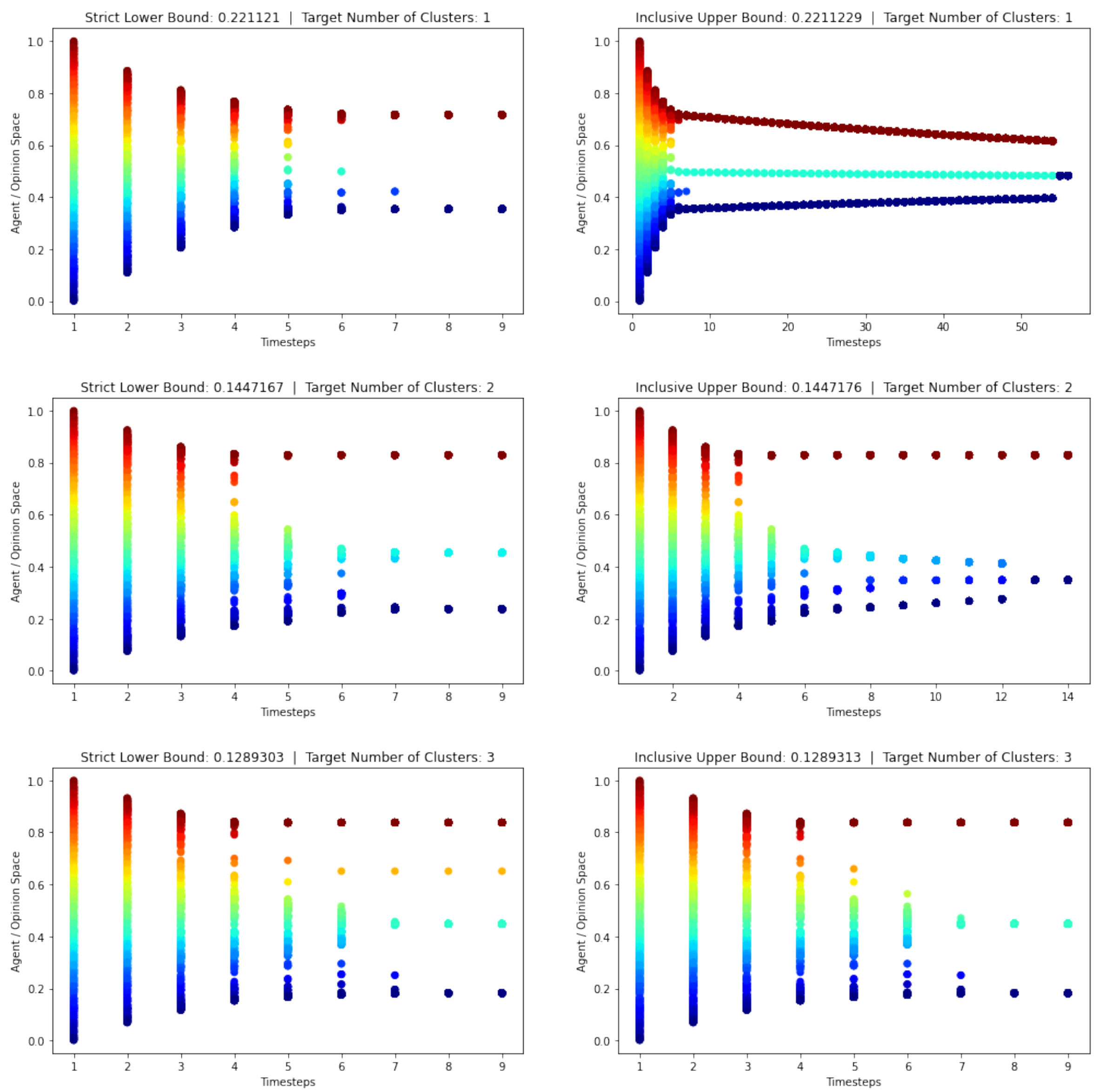
[2] https://github.com/kvdsouza/hkconfidencebounds

In the next section, we present various simulations of our algorithm.

4.3 *Simulations & Results*

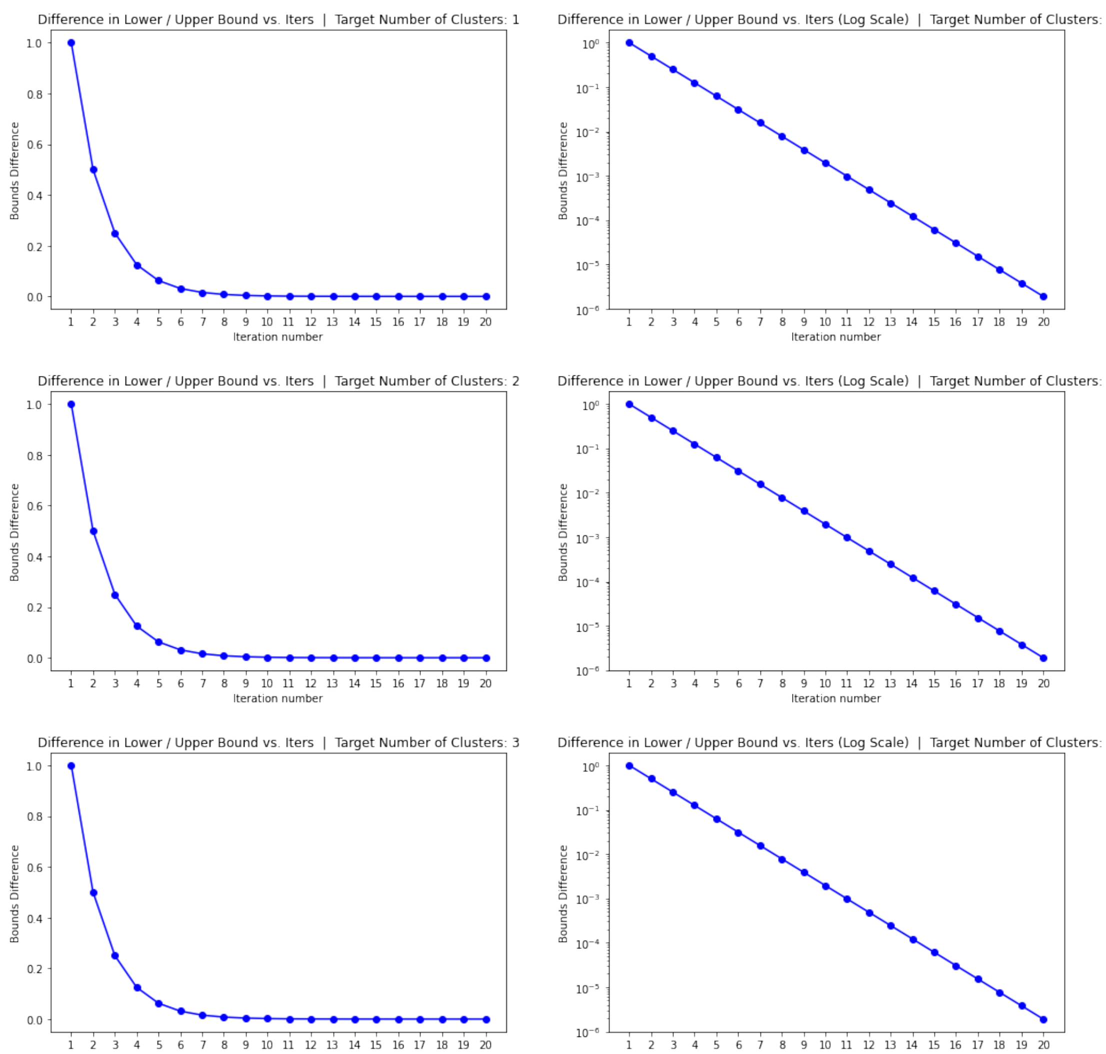
In the first set of simulations (Figure 1), we use 626 agents (generated uniformly between 0 and 1) and run the algorithm to find the lower / upper bounds on the minimum confidence bounds for 1, 2, and 3 clusters — we then test these identified confidence bounds on the original set of agents, simulating the agents to convergence.

Figure 1: First set of simulations



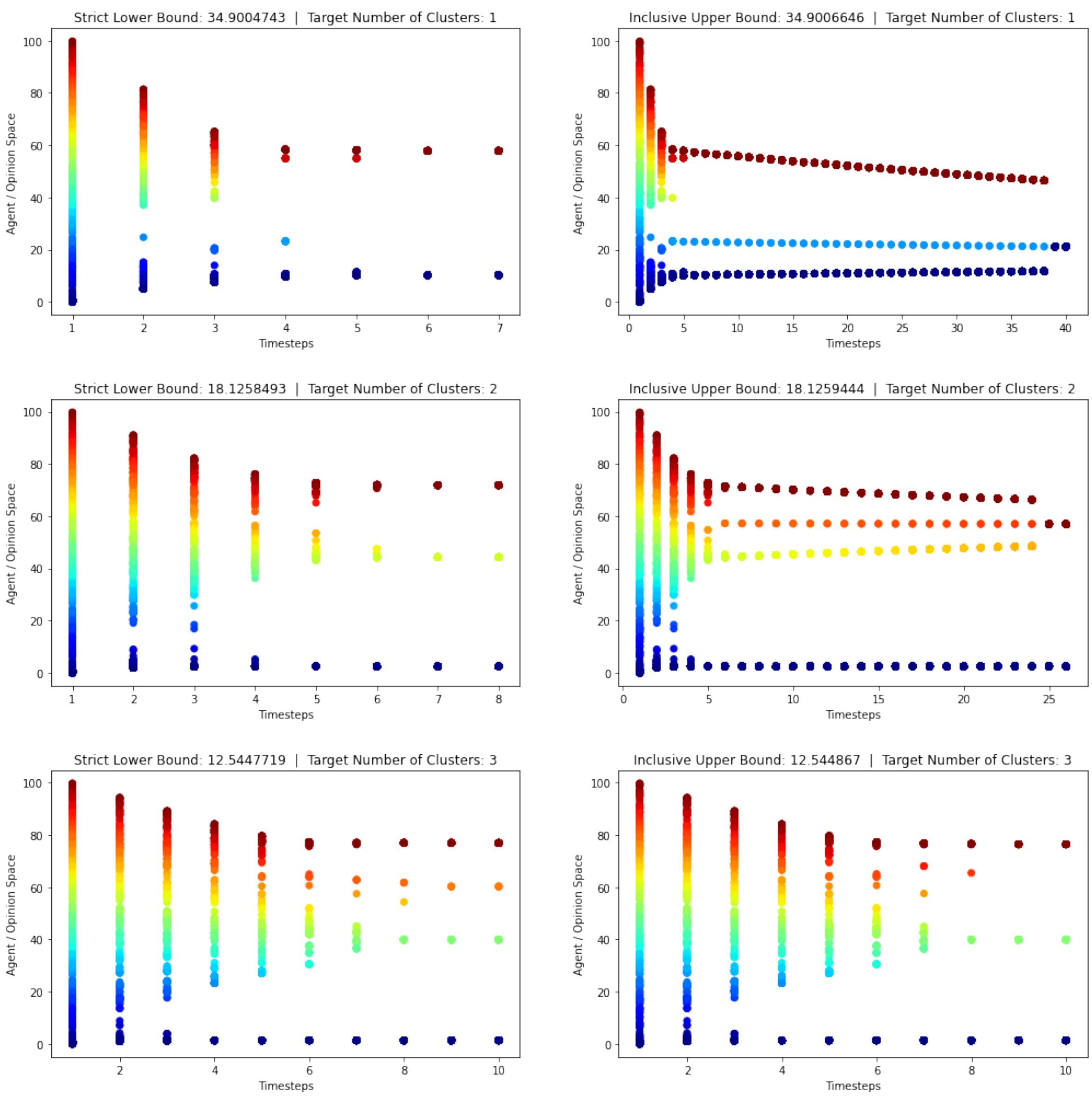
The left column of Figure 1 represents the simulation results of using the strict lower bound of the minimum confidence bound — the agents don't converge to the target number of clusters, as the confidence bound is not large enough. The right column represents the simulation results of using the inclusive upper bound of the minimum confidence bound — we see that the agents converge to the target number of clusters, as required. In each case, you can see how tight the bound is on the minimum confidence bound required, indicating that the algorithm converges quite close to the minimum confidence bound required.

Figure 2: Convergence plot for each of the first set of simluations



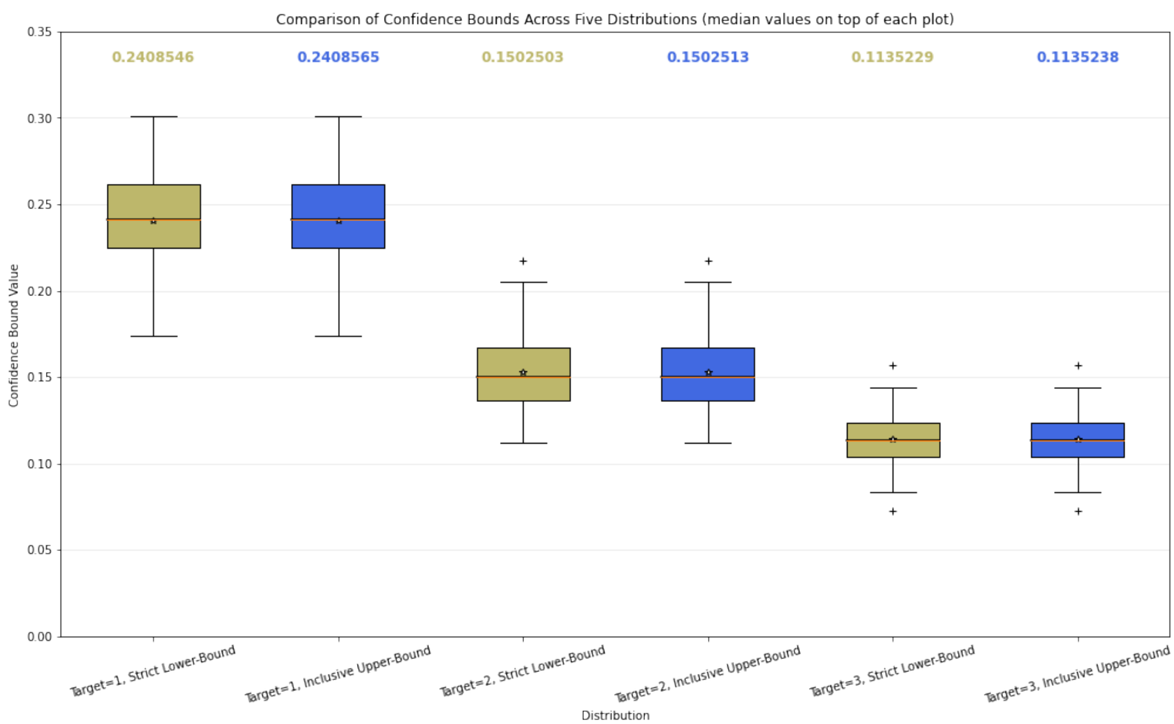
In addition, for the first set of simulations, we also graph the difference between the strict lower bound and inclusive upper bound (which is measured on each iteration to evaluate whether the answer has converged) and the corresponding iteration in order to get a sense for the convergence rate of the number of binary-search-step iterations required by the algorithm (see Figure 2). We can see by looking at the figures how the convergence rate is effectively linear on the log scale in each case — of course this doesn't take into account the iterations used in the simulation components of the algorithm (which are bounded in the 1-dimensional case) but at least this algorithm presents an upper-bound on the convergence of this algorithm.

Figure 3: Second set of simulations



In the next set of simulations, we again use 626 agents, half of which are generated uniformly between 0 and 1 and the other half of which are also generated uniformly between 0 and 1 but are then multiplied by 100 to dramatically distance the two starting agent distributions. We do this to show how the algorithm is both robust to agent magnitudes that are larger than the 0 to 1 range and how it is also still able to determine a tight strict lower bound and tight inclusive upper bound on the minimum confidence bound required to converge to the target number of clusters. We see again in Figure 3 how the identified strict lower bounds on the confidence bounds are not able to converge to the target number of clusters while the identified inclusive upper bounds on the confidence bounds are able to.

Figure 4: Third set of simulations



Finally, we plot a distribution of the minimum confidence bound (Figure 4) needed to converge to the target number of clusters by randomly generating 100 sets of 100 agents and running the algorithm for each set, for 1, 2, and 3 target number of clusters — we plot the inclusive upper bound. In the sample distribution, we see how the minimum confidence bound largely becomes smaller the more lenient (i.e. larger) the target number of clusters is. This makes sense considering that less distance is required to converge since the farthest points do not need to converge to each other.

5 Conclusion

In this paper, we have explored the related literature surrounding HK clustering and confidence bounds. We provide a binary-search based algorithm to identify the minimum confidence bound in 1-D to reach 1, 2, or 3 clusters at convergence. This approach leverages our understanding of lower and upper bounds of our confidence bound value by reducing the difference between the bounds to identify the minimum value. Of note, we find that over repeated samplings, we find an inverse relationship between the target number of clusters and the size of the confidence bound, as seen in Figure 3. This relationship is monotonically decreasing with diminishing rate of change as cluster size increases. Moreover, we show that the convergence rate in lower / upper bound difference of our algorithm is effectively linear on the log scale over iterative steps.

There are several exciting next steps. We, or other researchers, could examine more rigorously through simulations the relationship between min cluster size and min confidence bound. Is there an equation that can relate the two factors perfectly based on randomized initial distributions? Alternatively, given different initial distributions, how will these distributions affect our predictive relationship between cluster size and confidence bound? And more generally, how do other non-random distributions (say Poisson, multimodal) impact this relationship?

Secondly, work remains to be done to try to improve runtime of our algorithm or explore related applications. Can one empirically prove the lower bound to this problem beyond binary search? Additionally, can one use this relationship between bounds and clusters to provide predictive power that can obfuscate the longer polynomial runtime of the classic HK algorithm? We see several exciting future vectors from this work.

To conclude, opinion dynamics are becoming increasingly important to study, especially in online communities and forums. Whether it be measuring the impact of opinion influence in social media networks to preventing echo chambers and testing deliberative online democracy, mathematical models for these questions have never been more sorely needed.

It is our hope that in an age of increasing algorithmic personalization (15), models like HK systems and future papers like ours can shed light on the mathematical intricacies of the dynamical systems underlying modern social interactions.

References

[1] A. Bhattacharyya, M. Braverman, B. Chazelle, and H. L. Nguyen. On the convergence of the hegselmann-krause system. *arXiv*, 2012.

[2] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. On the 2r conjecture for multi-agent systems. In *2007 European Control Conference (ECC)*, pages 874–881. IEEE, 2007.

[3] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. On krause's multi-agent consensus model with state-dependent connectivity. *IEEE transactions on Automatic Control*, 54(11):2586–2597, 2009.

[4] B. Chazelle. The total s-energy of a multiagent system. *SIAM J. Control and Optimization*, 49(4):1680–1706, 2011.

[5] B. Chazelle. The dynamics of influence systems. *arXiv:1204.3946v2*, (nlin.AO), 2012.

[6] B. Chazelle and C. Wang. Inertial hegselmann-krause systems. *arXiv:1502.03332v3*, (cs.SY):1926–1941, 2016.

[7] J. C. Dittmer. Consensus formation under bounded confidence. *Nonlinear Analysis: Theory, Methods & Applications*, 47(7):4615–4621, 2001.

[8] R. Hegselmann, U. Krause, et al. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of artificial societies and social simulation*, 5(3), 2002.

[9] U. Krause. Soziale dynamiken mit vielen interakteuren. eine problemskizze. *Modellierung und Simulation von Dynamiken mit vielen interagierenden Akteuren*, 3751:2, 1997.

[10] J. Lorenz. A stabilization theorem for dynamics of continuous opinions. *Physica A: Statistical Mechanics and its Applications*, 355(1):217–223, 2005.

[11] J. LORENZ. Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics C*, 18(12):1819–1838, Dec 2007.

[12] J. Lorenz. Repeated averaging and bounded confidence modeling, analysis and simulation of continuous opinion dynamics. 2007.

[13] J. Lorenz. Heterogeneous bounds of confidence: meet, discuss and find consensus! *Complexity*, 15(4):43–52, 2010.

[14] S. Martínez, F. Bullo, J. Cortés, and E. Frazzoli. On synchronous robotic networks—part ii: Time complexity of rendezvous and deployment algorithms. *IEEE Transactions on Automatic Control*, 52(12):2214–2226, 2007.

[15] N. Perra and L. E. Rocha. Modelling opinion dynamics in the age of algorithmic personalisation. *Scientific reports*, 9(1):1–11, 2019.

[16] Y. Shang. Deffuant model with general opinion distributions: First impression and critical confidence bound. *Complexity*, 19:38–49, 2013.

[17] B. Touri and A. Nedic. Hegselmann-krause dynamics: An upper bound on termination time. In *IEEE Conference on Decision and Control*, 2012.

[18] C. Wang, Q. Li, E. Weinan, and B. Chazelle. Noisy hegselmann-krause systems: Phase transition and the 2r-conjecture. *Journal of Statistical Physics*, 166(5):1209–

1225, 2017.

[19] E. Wedin and P. Hegarty. A quadratic lower bound for the convergence rate in the one-dimensional hegselmann-krause bounded confidence dynamics. *CoRR*, abs/1406.0769, 2014.

[20] E. Wedin and P. Hegarty. A quadratic lower bound for the convergence rate in the one-dimensional hegselmann–krause bounded confidence dynamics. *Discrete Compute Geometry*, 53:478–486, 2015.

[21] H. Xia, H. Wang, and Z. Xuan. Opinion dynamics: A multidisciplinary review and perspective on future research. *IJKSS*, 2:72–91, 10 2011.