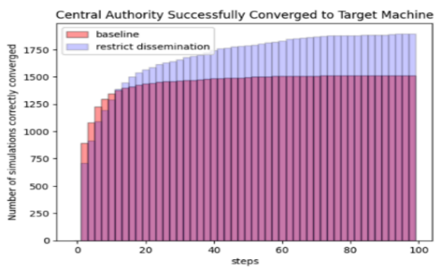
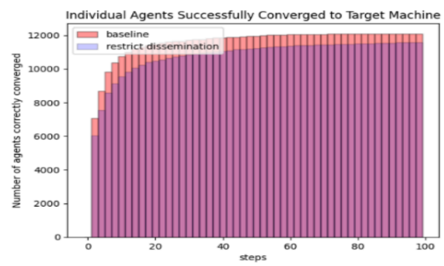
Summary of Results on Network Simulations - Spring 2024

<u>Central Authority vs Individual Agents</u>

These simulations are the same as in Wagner and Herington's setup, but with keeping track of individual beliefs as well. In virtually all cases, the individual agents not only do not learn the correct arm as well as the central authority does when dissemination is restricted, but they even perform worse than the baseline.

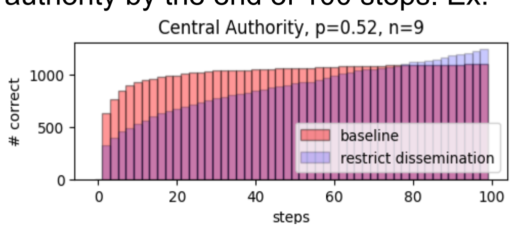


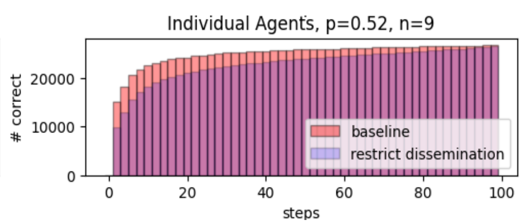


This graph shows p=0.6 and n=9, but this trend holds for most other parameters. (including when inertia and epsilon features are added)

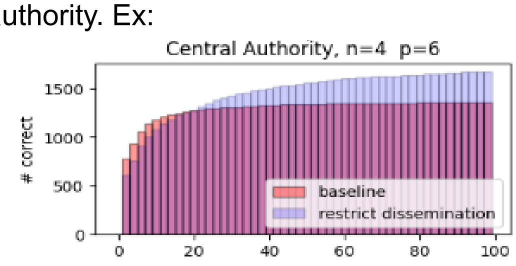
The two slight exceptions for this are:

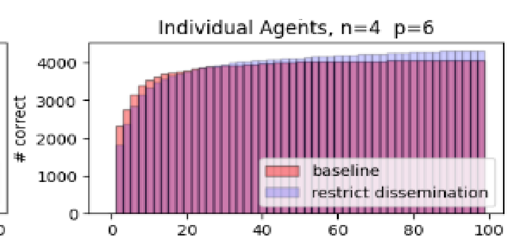
1. For low target p values (.51-.54 ish), the individual agents initially seem to learn better than the central authority relative to the baseline case, but are surpassed by the central authority by the end of 100 steps. Ex:





2. For networks with few agents (3-5 ish) the individual agents do often benefit from restricting dissemination, however they still perform relatively worse than the central authority. Ex:





Notes:

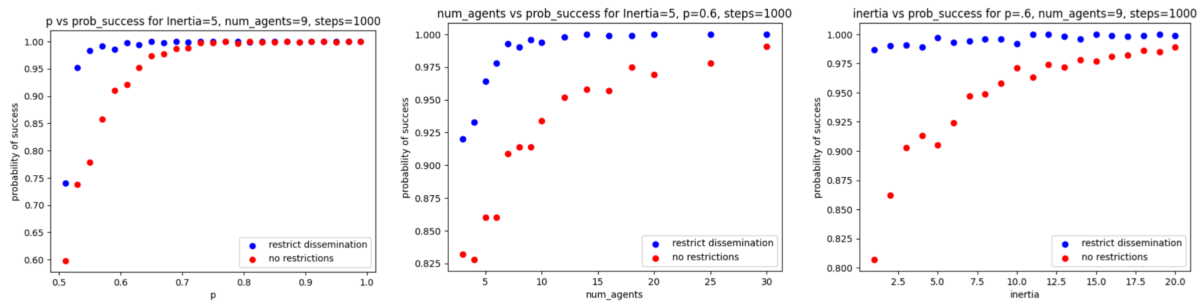
- Y-axis' are not standardized between graphs, so findings are only based on relative performance without direct comparison between parameter values.
- Many of the graphs seem like they continue to evolve past 100 steps, but in most practical cases 100 steps is probably enough to evaluate a network's success.

In the next 3 sections, we add a new feature in an attempt to make the simulations more realistic. Each feature is isolated from the rest to focus only on its own effect.

Inertia

In these simulations, agents don’t switch their action until their belief has been changed for x consecutive number of steps. This feature helps the network tremendously in all cases (both baseline and restricting dissemination)

So far, results suggest that restricting dissemination is always better than the baseline for any parameters. The only exception is when the target p value is very high, where both the restricted dissemination and baseline case are practically guaranteed to succeed anyway.

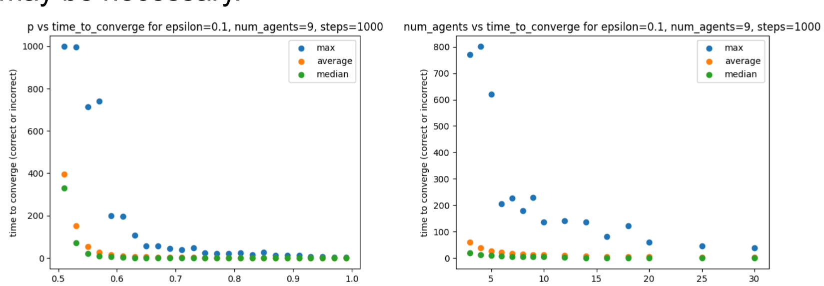


Effect of various parameters on central authorities' probability of successful learning after 1000 steps.

Epsilon-greedy

In these simulations, agents explore what they believe to be the worse arm with probability epsilon. Results suggest that for any values of p, number of agents, or epsilon (minimum is 0.01), the network is virtually guaranteed to successfully converge. This would seem to suggest that restricting dissemination does not have a substantial effect on the network, although all these results are after 1000 steps, so the effect on time to converge remains to be completely seen.

Some analysis on this shows that for small values of p and num_agents, it is possible for the network to take a *very* long time to converge, but otherwise it is usually quite fast. Though further analysis may be necessary.

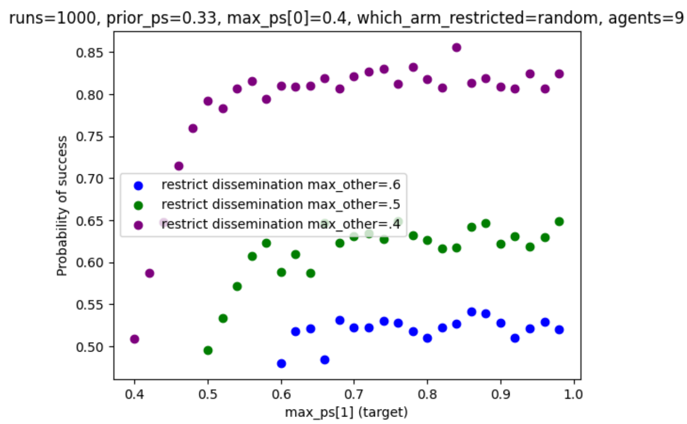


Time to converge statistics for restricted network

Dynamic p Values

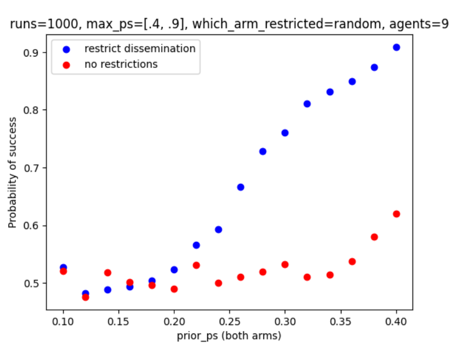
In these simulations, an arm's p-value gets a bonus for each successful pull that has been seen by the agent pulling it, capped at the machine's max p value. These diverge the most from the original setup, since now we are concerned not only with both arms' initial/current p-value, but also their max p-value, which serves as the basis for determining the objectively better target arm. For this reason, all the simulations run so far have the initial p value for both the target arm and the other arm start at the same value. Here are some of the trends:

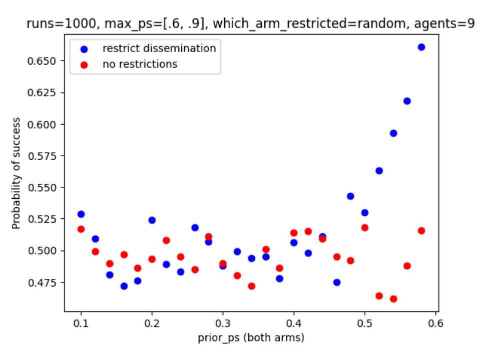
1. In general, restricting dissemination always causes the probability of success to improve or stay the same, but never hurts it.
2. Increasing the max p value of the target arm (starting at approximately equal to the max p value of the other arm) initially provides a huge benefit to the probability of success, but quickly converges to some capped value, and this value is inversely proportional to the max p value of the other arm.



(here the no restrictions case is always around .50 for all max_ps[1] (target) values, with maybe a barely perceptible positive slope)

3. Increasing the initial p values of both arms initially does nothing, with both restricted and non restricted cases hovering at about .50 probability of success until a certain threshold is hit, at which point the probability of success rapidly increases with the initial p values. The threshold that this occurs seems to be proportional to the max p value of the other arm.





4. The baseline case only sees a boost in success probability whenever the initial p value and the max p value of the other arm are the same or close together.