Resilient Distributed Hypothesis Testing

Bo Wu, Steven Carr, Suda Bharadwaj, Zhe Xu, and Ufuk Topcu

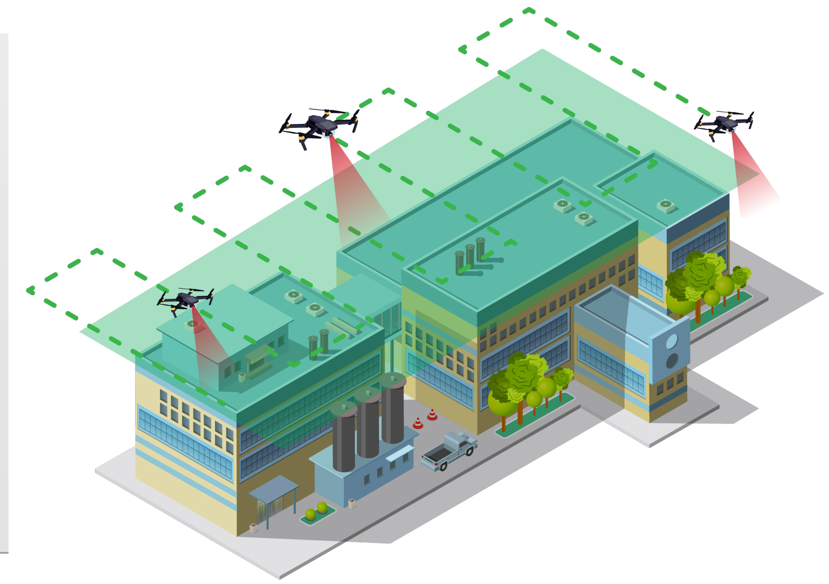
ASSURED AUTONOMY

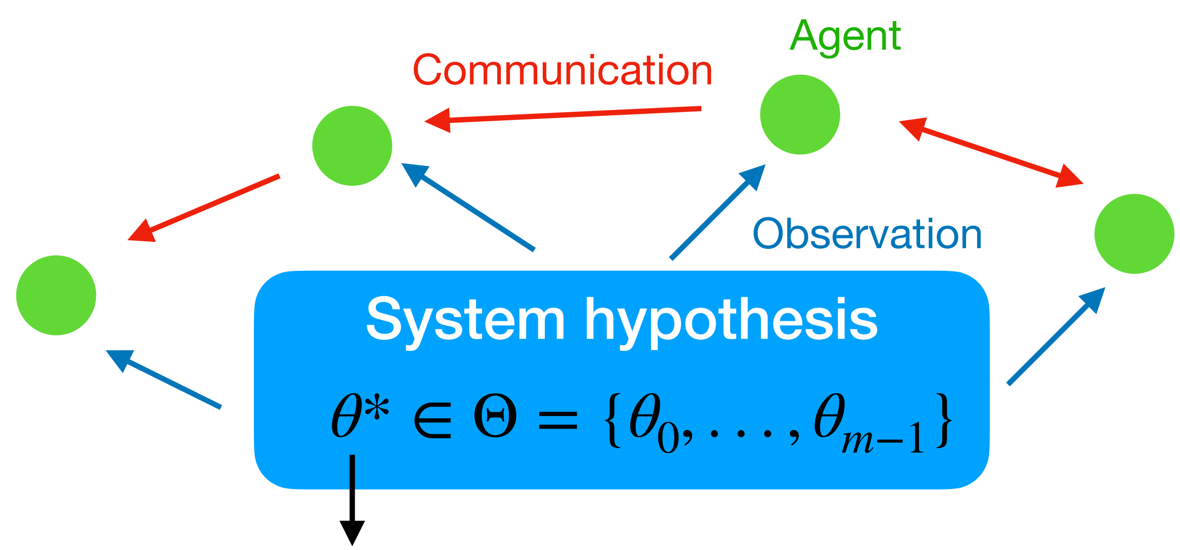
aUTonomous SYSTEMS GROUP

Motivation

A network of mobile agents collaboratively monitors a large-scale environment

- Heterogenous sensing capabilities
- Noisy sensors
- Limited communication range
- Time-varying network topology

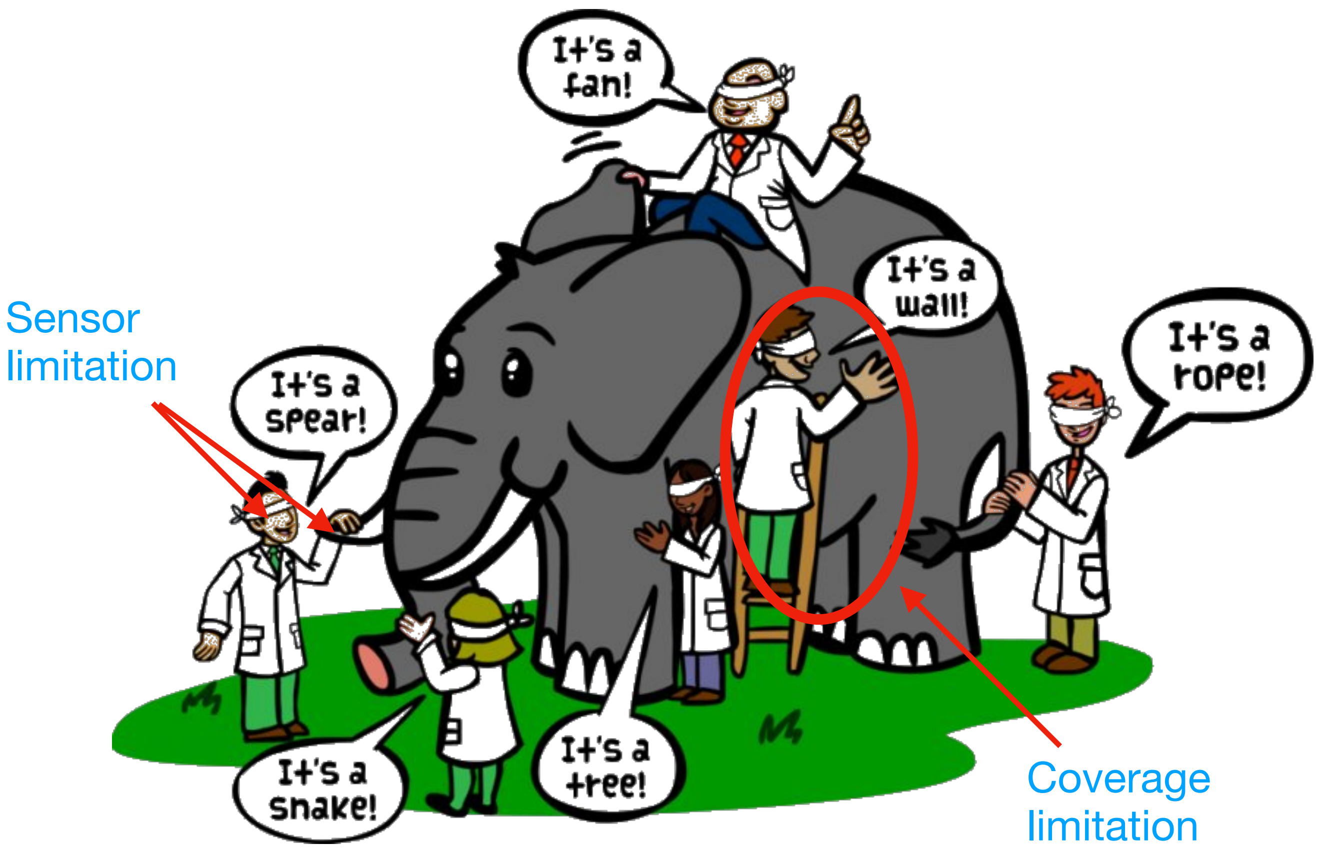




Unknown true hypothesis of the system
(e.g., true position of the object of interest)

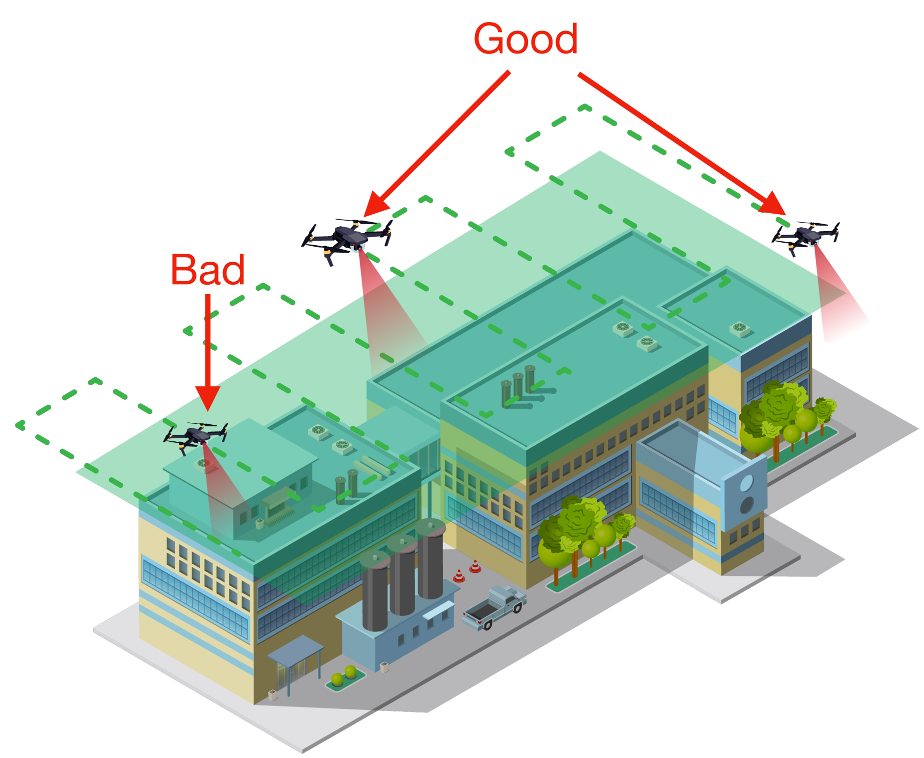
Collaboration in the Presence of Adversaries

Each agent only receives partial information about the state of the system



- Requires information sharing

May contain (Byzantine) adversaries



- Requires resilience against adversaries

How to process local and shared information so that all the good (non-adversarial) agents can collaboratively detect the true hypothesis of the system?

The Modeling Framework

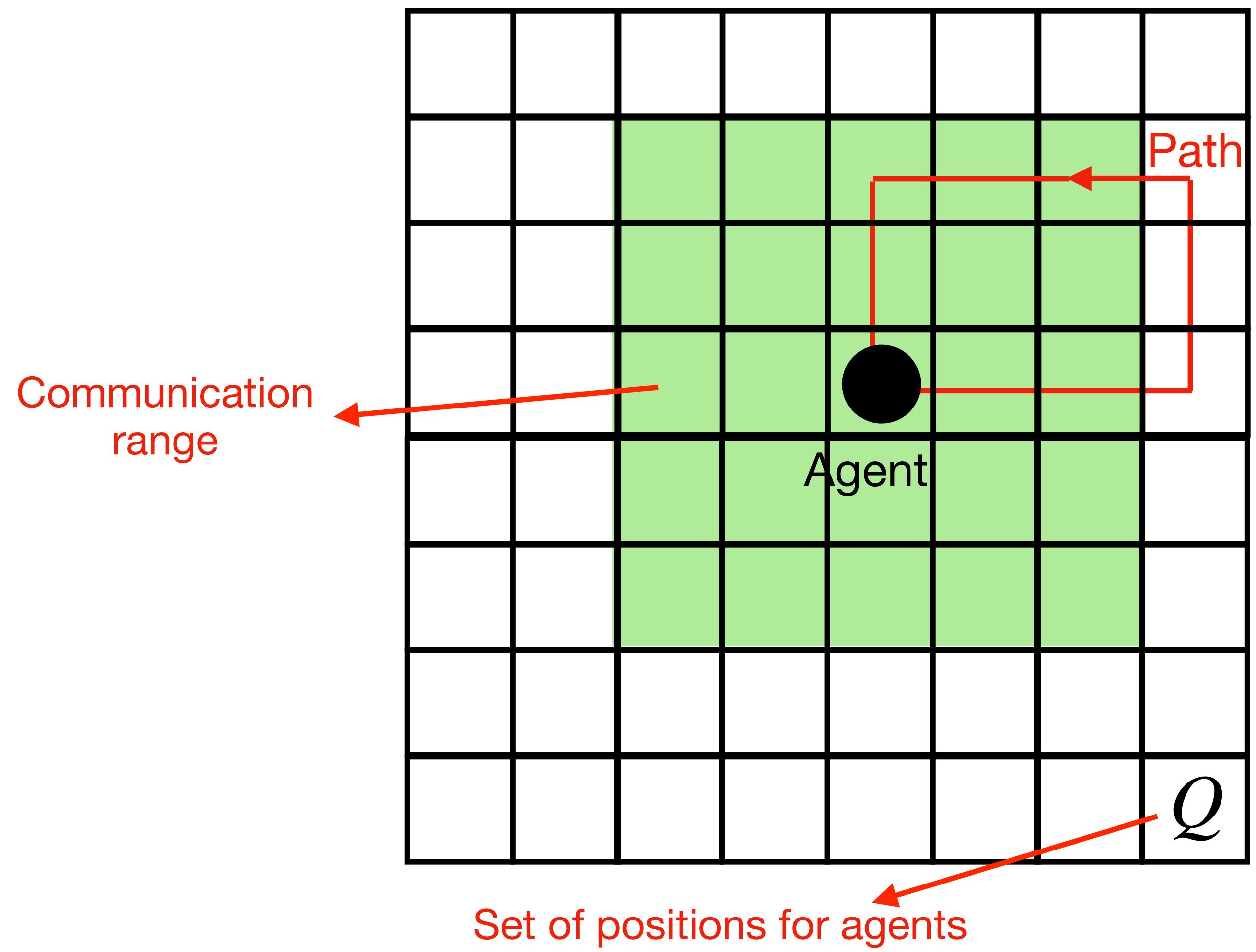
N agents

At most f adversaries $(f < N - f)$

m possible hypotheses, $\Theta = \{\theta_0, \ldots, \theta_{m-1}\}$

One true hypothesis of the environment $\theta^* \in \Theta$

Each agent follows a given path



At time t, agent i gets a local observation $s_{i,t}$ with a probability given by a local likelihood function

$$l_i(s_{i,t} \mid q_{i,t}, \theta^*)$$

Position of agent i at time t

Unknown true hypothesis of the system

How will things work?

1

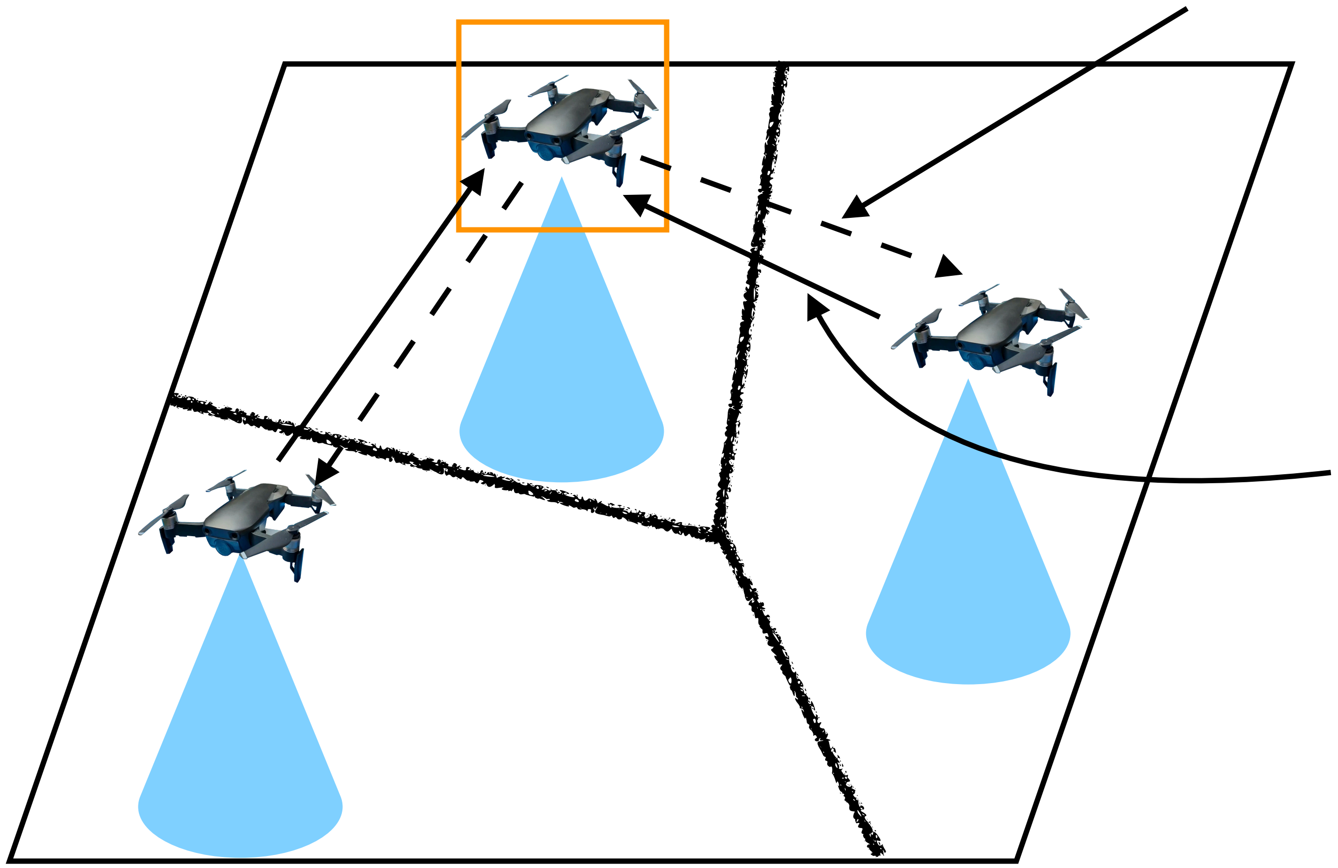
- Makes frequent local observations
- Updates its local belief

2

Shares its actual belief with neighbors

3

- Receives neighbors' actual beliefs
- Updates its own actual belief



Distributed Hypothesis Testing Problem

Each agent maintains two beliefs (probability distributions over Θ)

Likelihood function

Prior local belief

Local belief

$$b^l_{i,t+1}(\theta) = \frac{l_i(s_{i,t+1} \mid \theta, q_{i,t+1}) b^l_{i,t}(\theta)}{\sum_{p=1}^{m} l_i(s_{i,t+1} \mid \theta_p, q_{i,t+1}) b^l_{i,t}(\theta_p)}$$

Based on purely local information

Current local belief

Actual belief

(shared with neighbors)

$$b^a_{i,t+1}(\theta) = g(b^a_{i,t}(\theta), b^l_{i,t+1}(\theta), \{b^a_{j,t}(\theta)\})$$

Update rule

Prior actual belief

Actual beliefs from current neighbors

Merges shared beliefs from neighbors

Adversarial agents can share false actual beliefs!

Problem:
Design resilient actual belief update rule g such that $b^a_{i,t}(\theta^*) \to 1.$

Related Work

Consensus-based belief update algorithms are established in [1-3]

- Asymptotic convergence is proved
- No adversaries
- Assumes (periodic) connectivity of the network topology

Byzantine-resilient belief update algorithms are established in [4-6].

- Asymptotic convergence is proved in the presence of adversarial agents
- Static network topology
- Assumes global connectivity of the network topology

[1] R. Olfati-Saber, et al, “Belief consensus and distributed hypothesis testing in sensor networks,” in Networked Embedded Sensing and Control.Springer, 2006, pp. 169–182
[2] A. Lalitha, et al, “Social learning and distributed hypothesis testing,” IEEE Transactions on Information Theory, vol. 64, no. 9, pp. 6161–6179, 2018.
[3] A. Nedic, et al, “Fast convergence rates for distributed non-bayesian learning,” IEEE Transactions on Automatic Control, vol. 62, no. 11, pp. 5538–5553, 2017.
[4] L. Su and N. H. Vaidya, “Defending non-bayesian learning against adversarial attacks,” Distributed Computing, vol. 32, no. 4, pp. 277–289, 2019.
[5] A . Mitra, et al, “A new approach for distributed hypothesis testing with extensions to byzantine-resilience,” in 2019 American Control Conference (ACC), July 2019, pp. 261–266.
[6] . Mitra, et al, “A new approach to distributed hypothesis testing and non-bayesian learning: improved learning rate and byzantine-resilience,” arXiv 1907.03588, 2019.

Contributions

- Resilient algorithms for distributed hypothesis testing with time-varying network topology

- No requirement on the global connectivity of the network topology (the underlying communication graph does not have to be connected)

- Accommodating different sensor noise levels

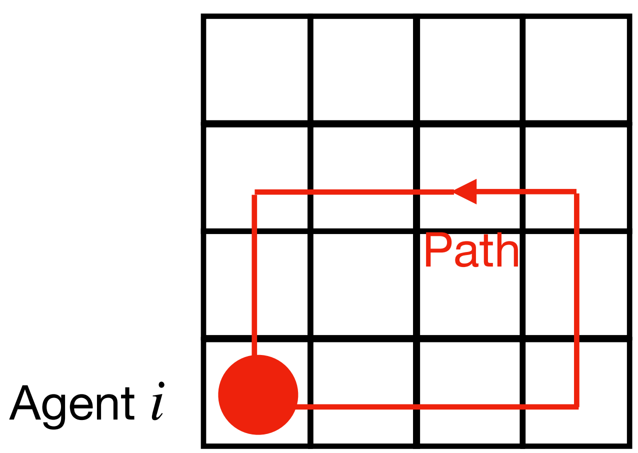
[1] B. Wu et al. "Resilient distributed hypothesis testing with time-varying network topology", American Control Conference, 2020, to appear.
[2] B. Wu et al. "Byzantine-resilient distributed hypothesis testing with time-varying network topology", journal version work in progress.

What can local belief evolution tell?

Given a pair of hypotheses θ and θ', agent i is a source agent, denoted as $i \in S(\theta, \theta')$, if it visits at least one position q infinitely often, where q satisfies

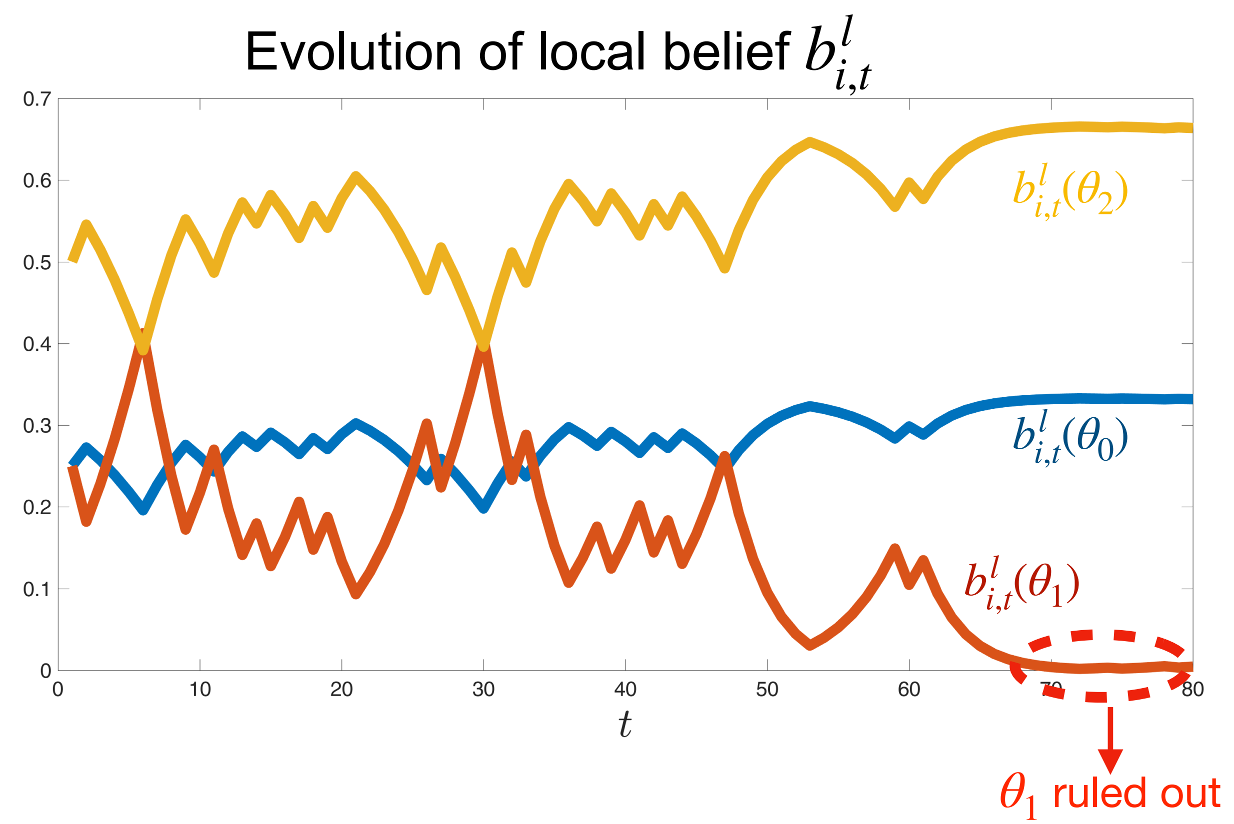
$$l_i(\cdot \mid q, \theta) \neq l_i(\cdot \mid q, \theta')$$

$i \in S(\theta, \theta') \rightarrow$ Agent i can locally distinguish between θ and θ'



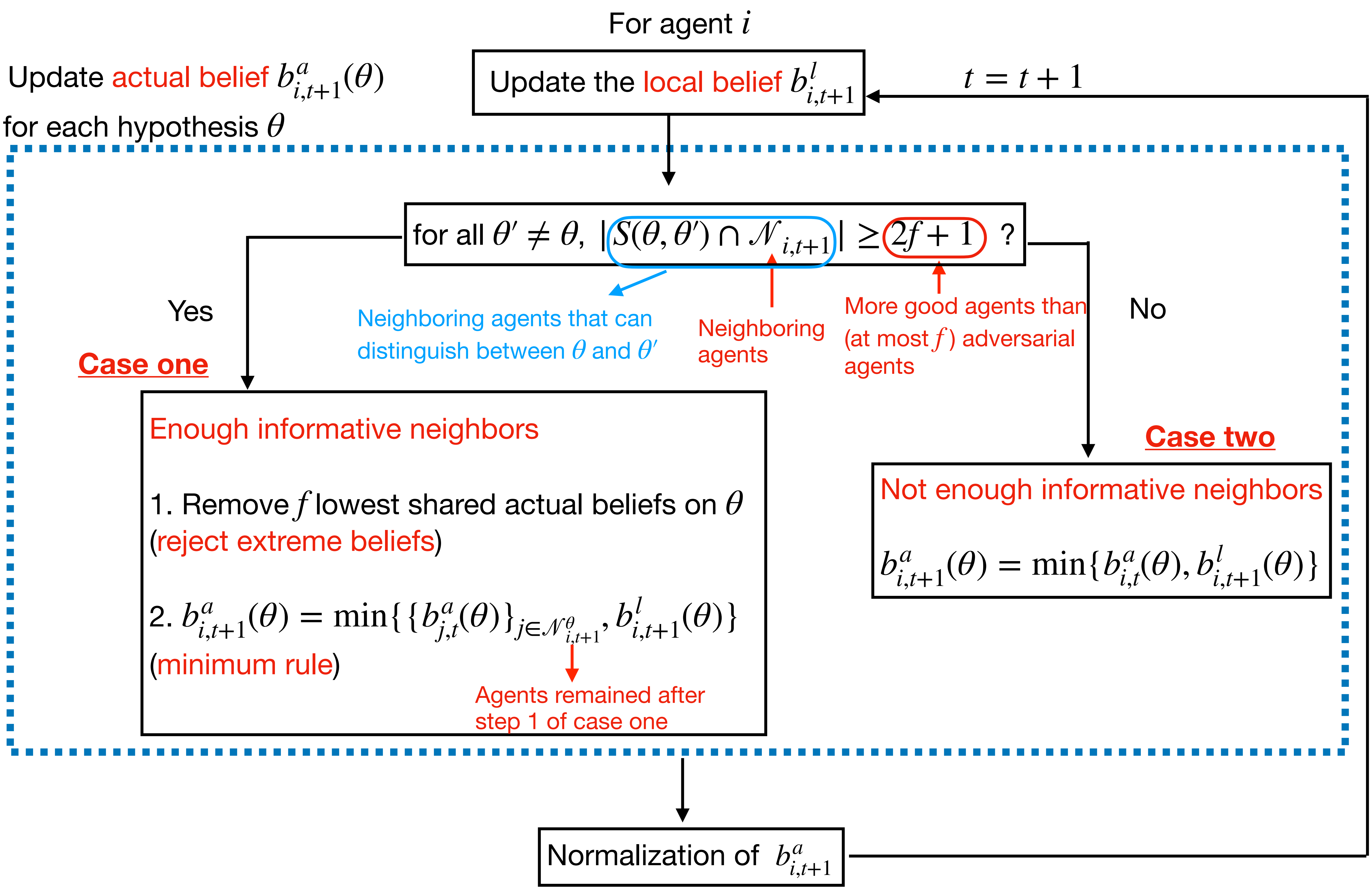
For every q along the path

$$l_i(s \mid q, \theta) = \begin{array}{cc} \begin{array}{cc} s_0 & s_1 \end{array} & \\ \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} & \begin{array}{l} \theta_0 \\ \theta_1 \\ \theta_2 = \theta^* \end{array} \end{array}$$



Local belief itself may rule out of some of the hypotheses

Synchronized Distributed Hypothesis Testing (SDHT)



Convergence of SDHT

For any non-adversarial agent i, suppose the following conditions hold:

1. Non-zero initial local and initial actual beliefs for every $\theta \in \Theta$

2. Every agent that does not interact "often enough" can distinguish by itself.

That is, if case one in SDHT happens only finitely often for a hypothesis $\theta \in \Theta$, then $i \in S(\theta, \theta')$ for any $\theta' \neq \theta$.

Then SDHT guarantees that $b^a_{i,t}(\theta*) \to 1$ almost surely as $t \to \infty$.

How to Make Better Use of the Shared Information?

Recall in SDHT

Case one: If for all $\theta' \neq \theta$, $\boxed{|S(\theta, \theta') \cap \mathcal{N}_{i,t+1}| \geq 2f + 1}$

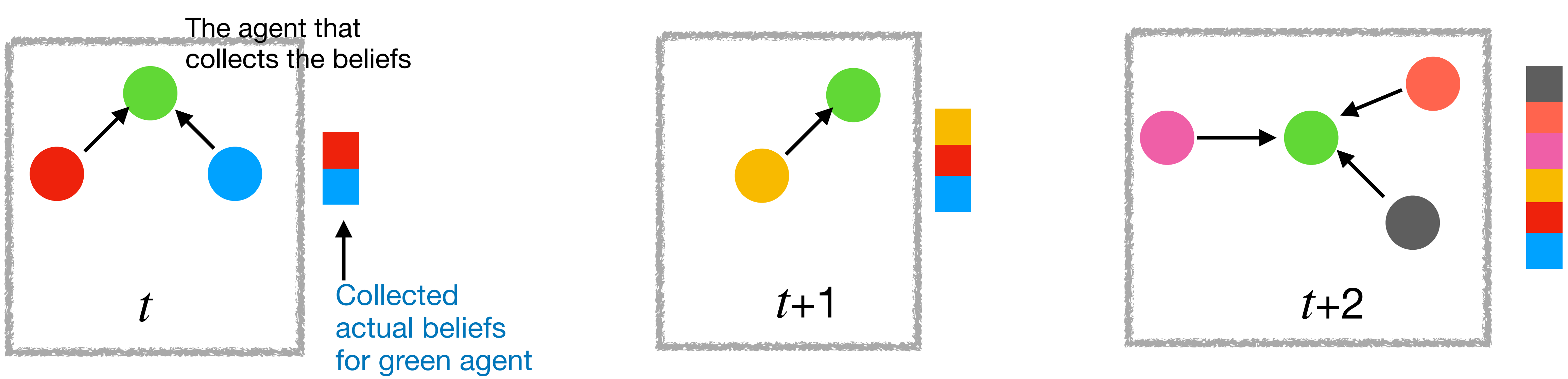
An agent must have enough number of informative neighboring agents to make use of the shared information.

Enough number of neighbors at the same time instant!

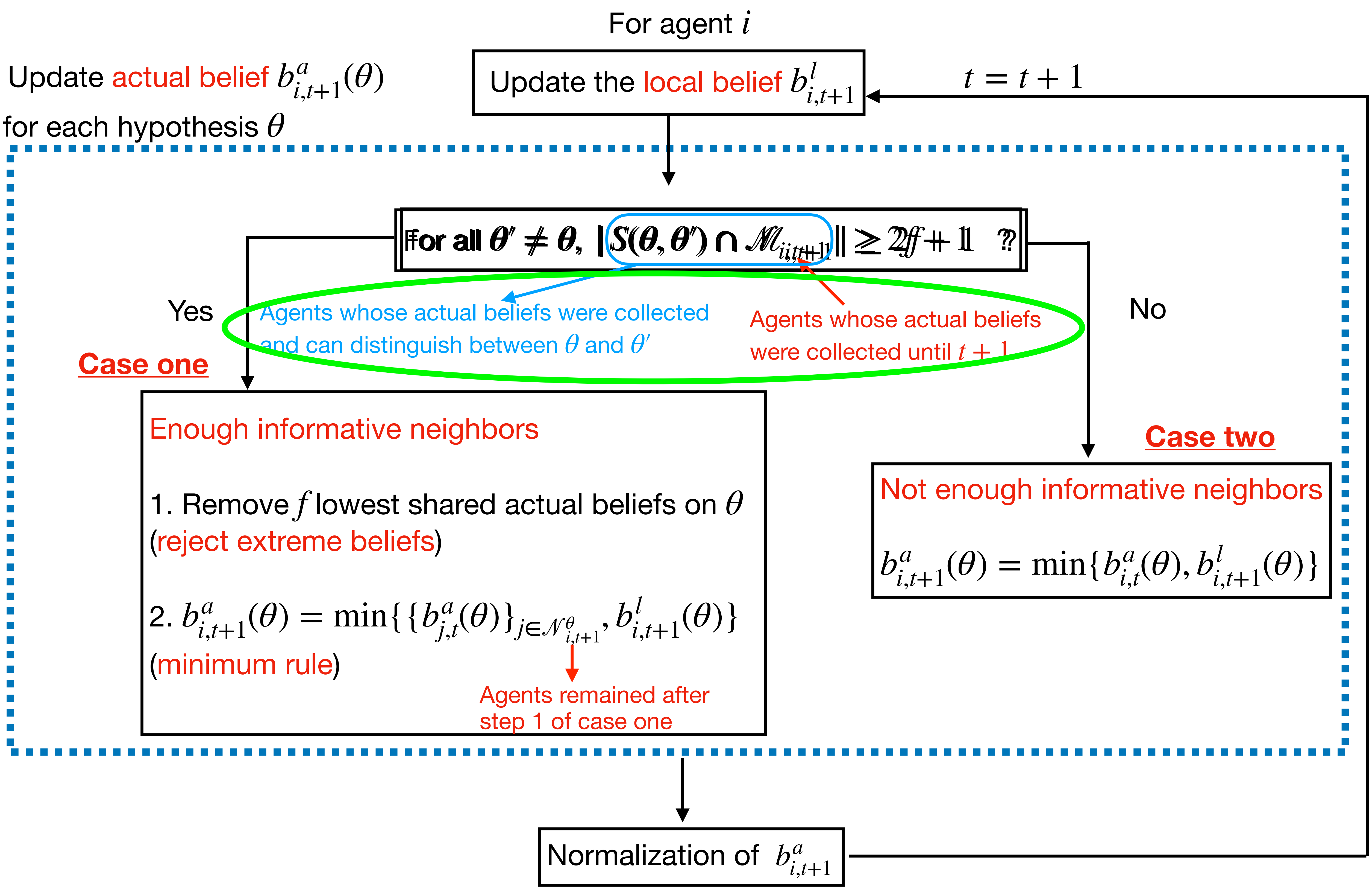
Can we make more frequent use of the shared information?

Key idea:

- Keep collecting shared actual beliefs over time
- Until enough actual beliefs collected for actual belief update



Asynchronous Distributed Hypothesis Testing (ADHT)



"Case 1": Minimum Rule versus Averaging Rule in Actual Belief Update

Minimum rule

1. Remove f lowest shared actual beliefs on θ
2. $b^a_{i,t+1}(\theta) = \min\{b^a_j(\theta)\}_{j \in \mathcal{N}^\theta_{i,t+1}}, b^l_{i,t+1}(\theta)\}$

Agents remained after step 1

Averaging rule

1. Remove f lowest and f highest shared actual beliefs on θ
2. $b^a_{i,t+1}(\theta) = \min\{avg\{b^a_j(\theta)\}_{j \in \mathcal{N}^\theta_{i,t+1}}, b^l_{i,t+1}(\theta)\}$

Agents remained after step 1

Pro: Quickly rules out unlikely hypothesis

Con: May have oscillations with noisy sensors

Pro: Averages out the effect of noisy sensors

Con: May take more time to suppress unlikely hypotheses

Case Study: Compromised UAV Classification

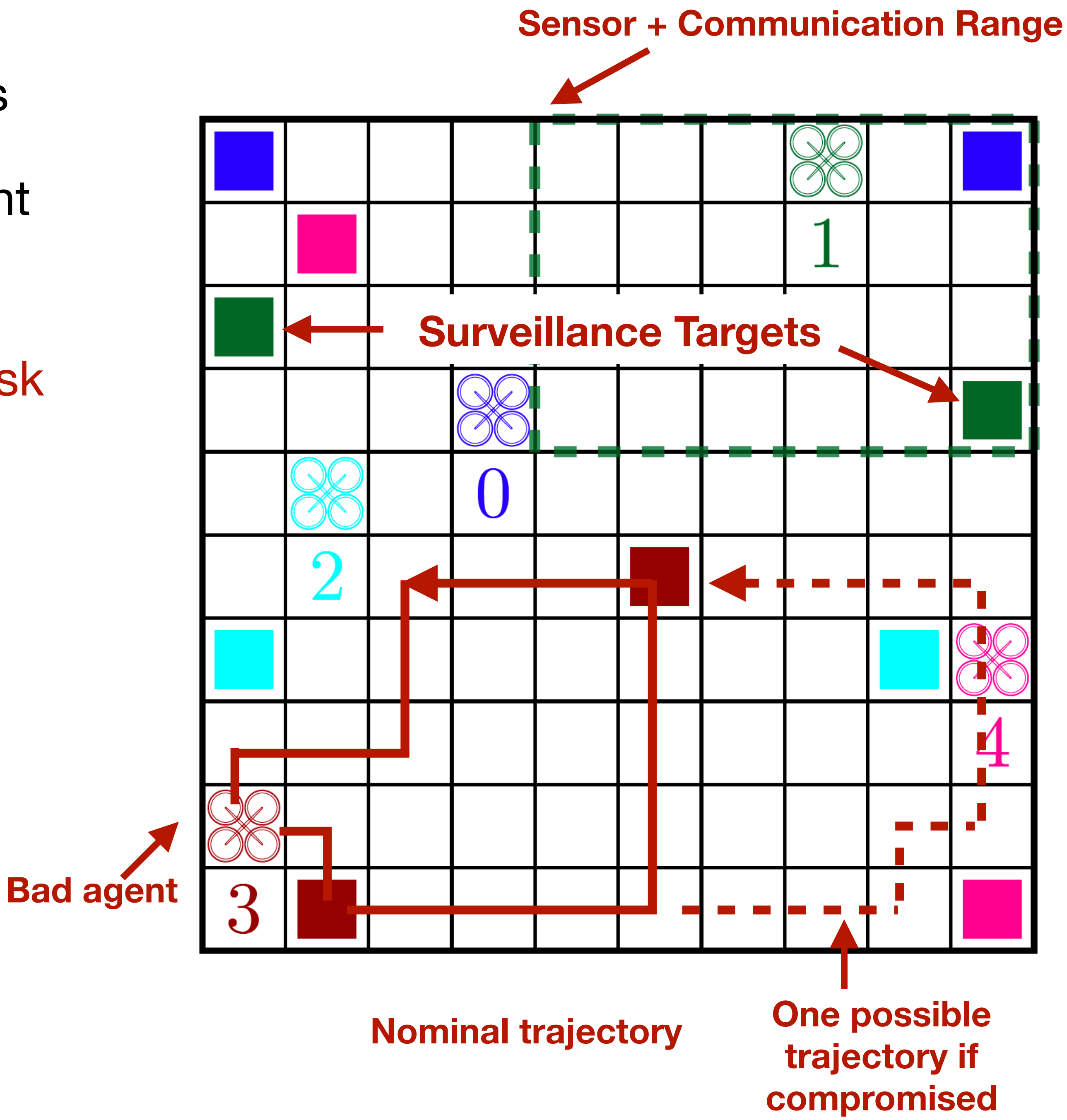
A set of agents with persistent surveillance tasks

Five agents — four good agents, one bad agent

Each agent
- Has a trajectory for prescribed surveillance task
- May follow a set of possible alternative trajectories if compromised
- Bad agent shares randomly generated beliefs

Noisy sensors
- When an agent is in range, its position is sensed from a distribution across the viewable positions

Objective: identify the bad agent



Hypothesis $\theta = \langle \theta(0), \theta(1), \theta(2), \theta(3), \theta(4) \rangle$

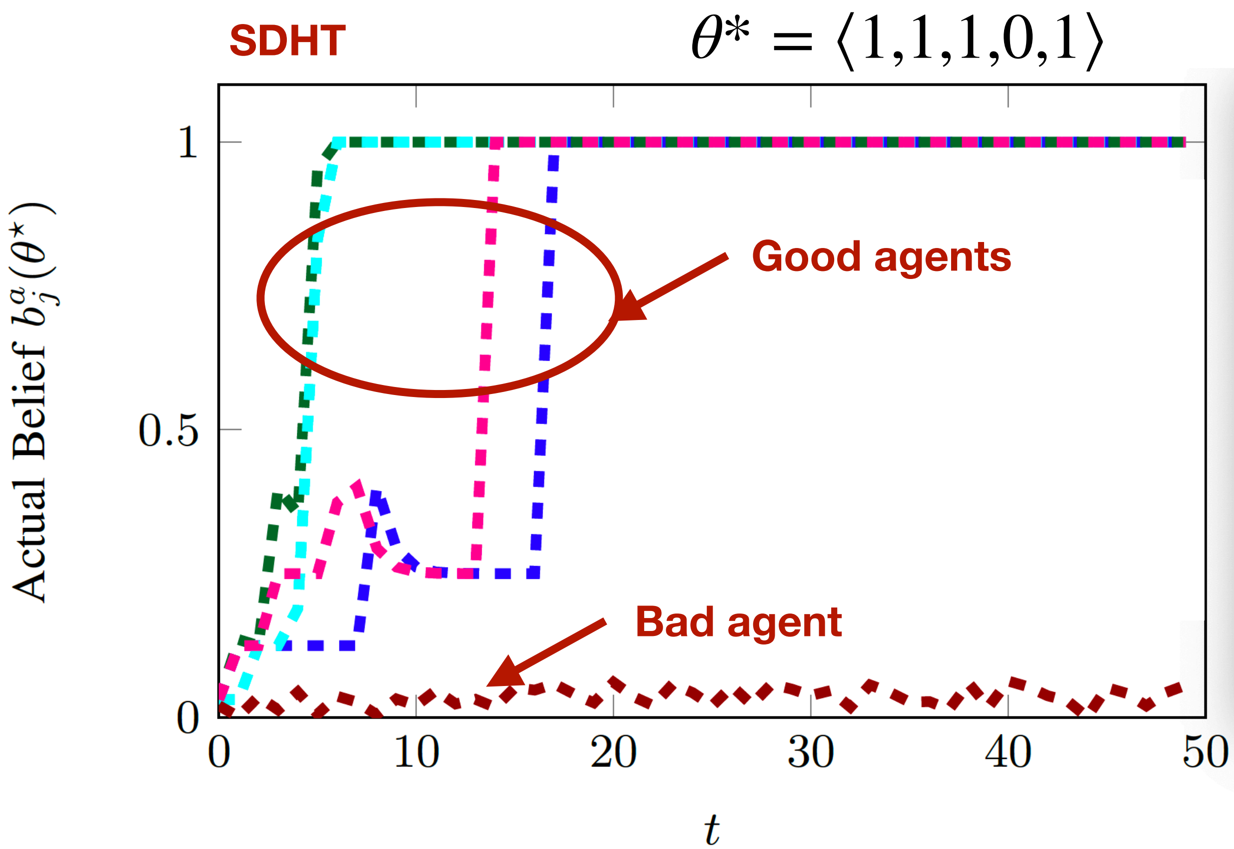
$\theta(i) \in \{0$ (**bad**), 1 (**good**)$\}$

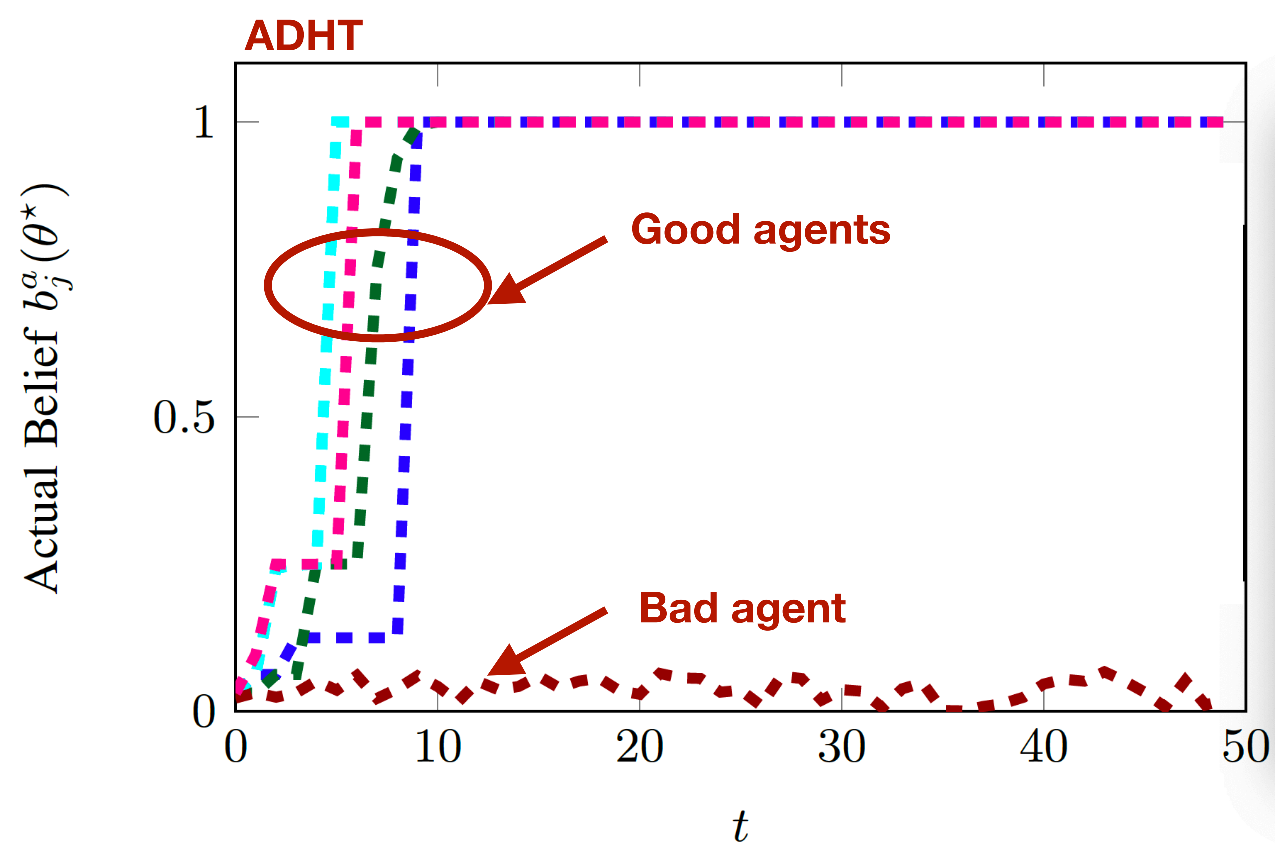
True hypothesis $\theta^* = \langle 1,1,1,0,1 \rangle$

Agent 3 is bad, ($\theta(3) = 0$).

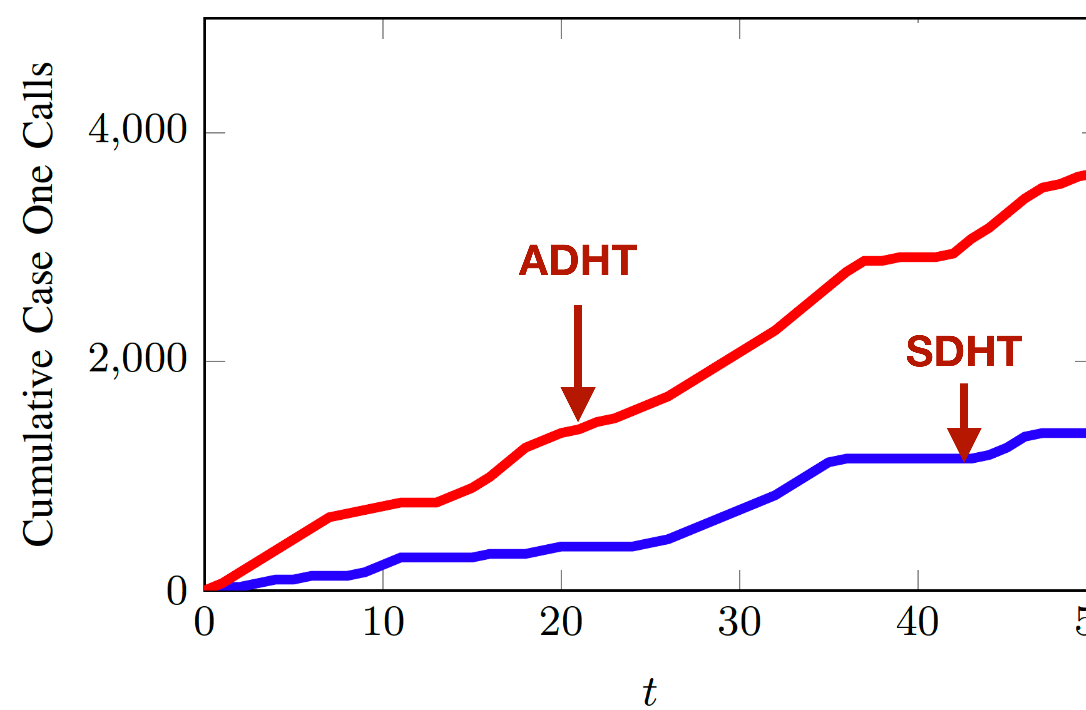
Case Study: Results

- Both SDHT and ADHT converge to the true hypothesis
- ADHT converges faster

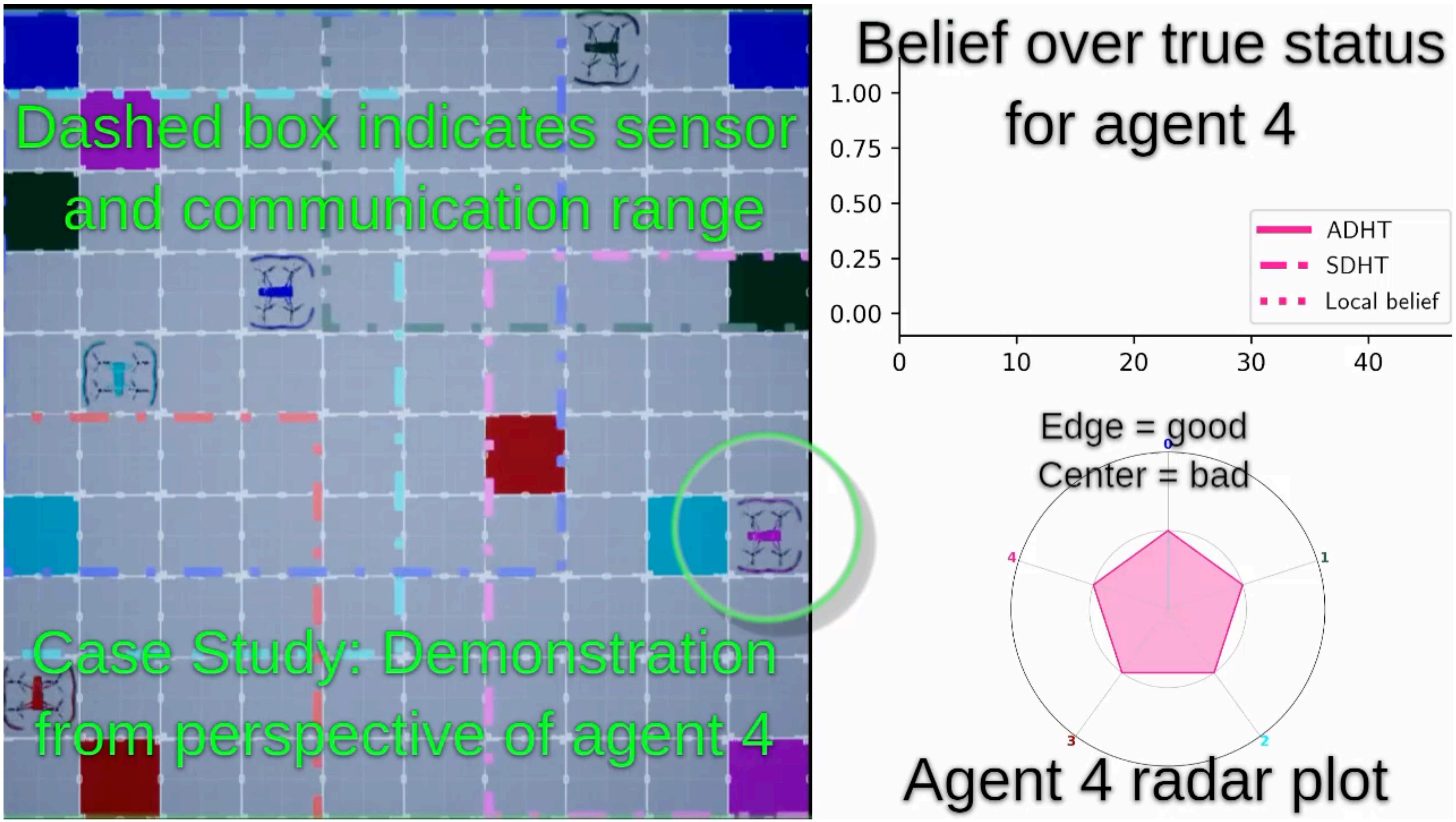




- ADHT uses shared beliefs more frequently than SDHT

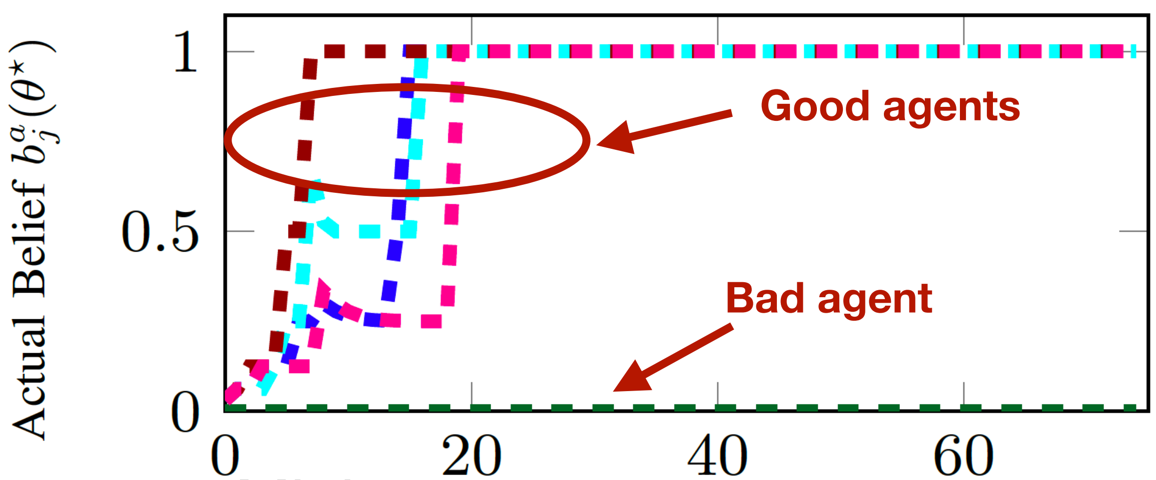


Case Study: Simulation

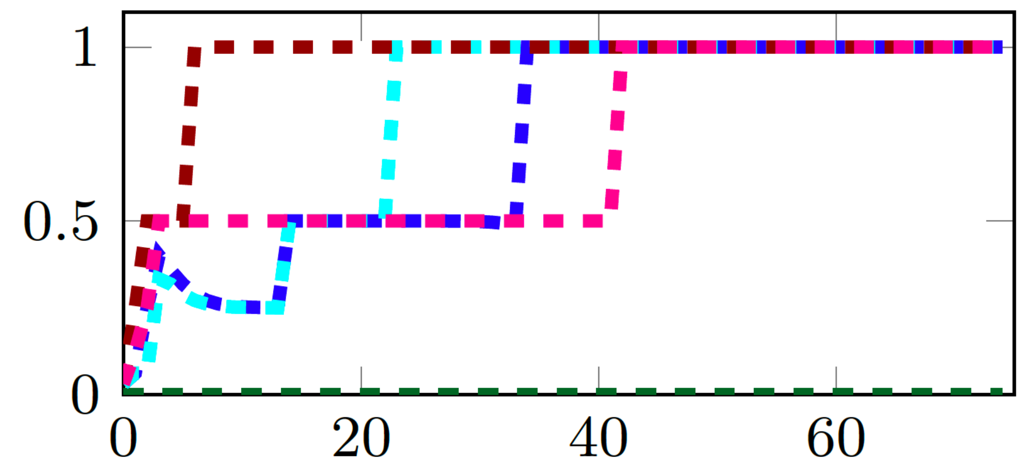


Minimum Rule vs Averaging Rule

Minimum rule converges faster with a low noise sensor

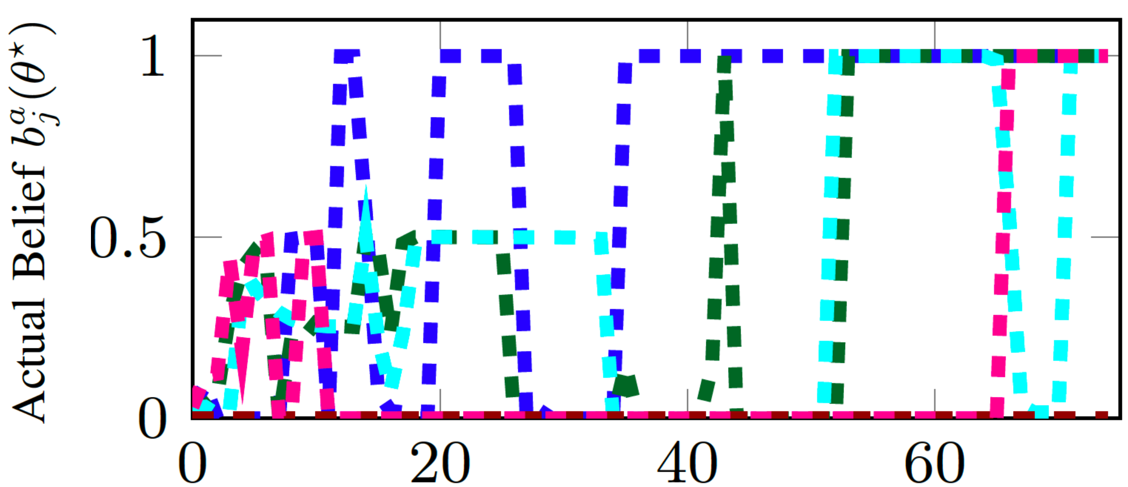


Minimum

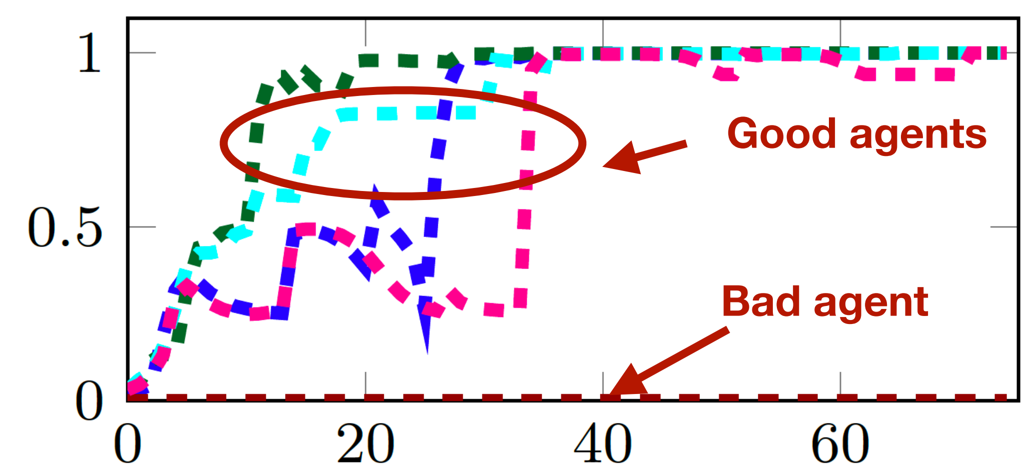


Averaging

Averaging rule converges faster with a high noise sensor



Minimum

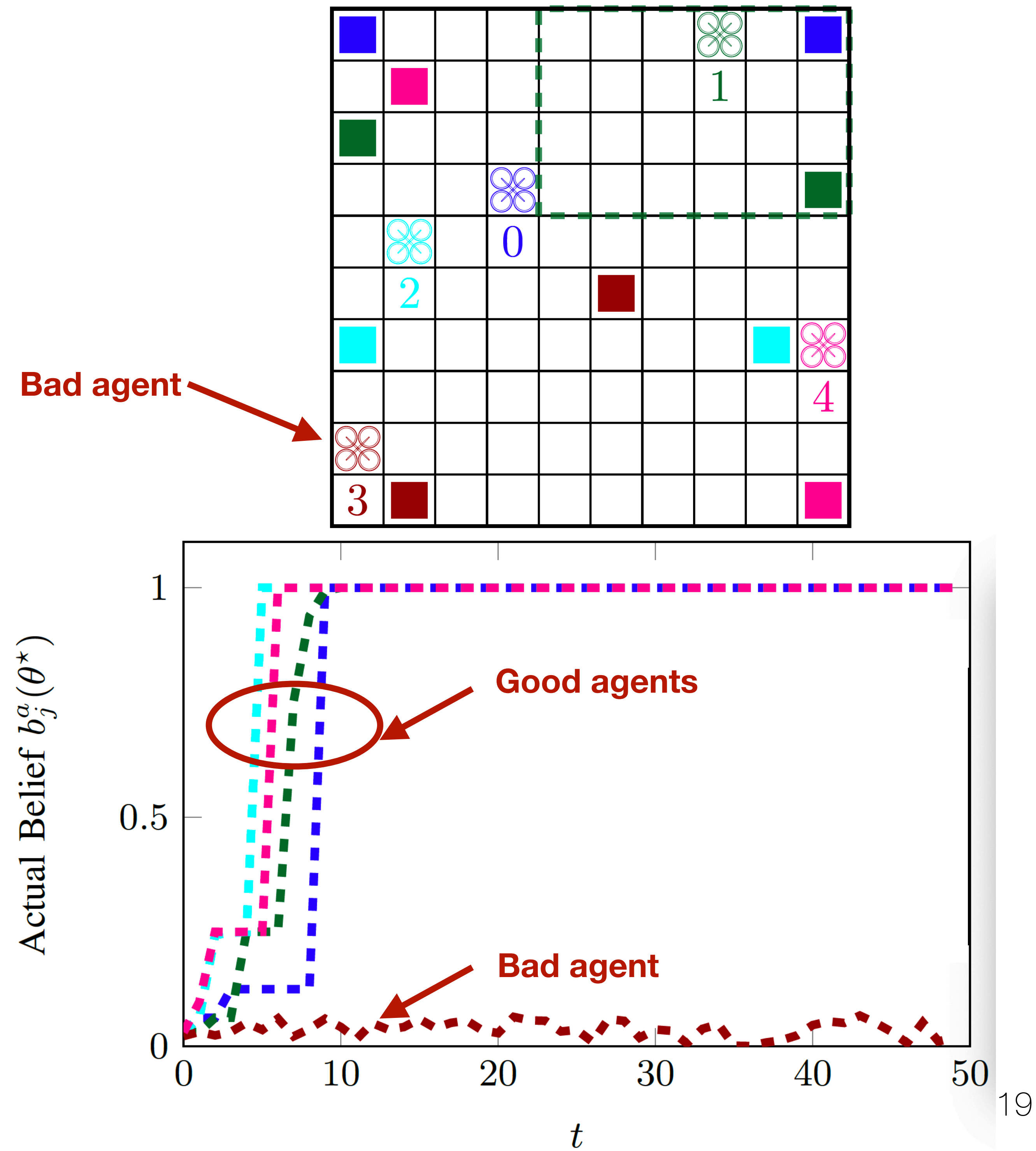


Averaging

Summary

- New distributed learning algorithms that are resilient to Byzantine adversaries
- Convergence is guaranteed without global connectivity constraints
- Minimum and average rules for different levels of sensing noises

$$\boxed{b_{i,t}^{a}(\theta^{*}) \rightarrow 1 \text{ almost surely as } t \rightarrow \infty}$$



What is next?

- How to plan the motion for each agent to guarantee convergence?
- Convergence rate analysis