Robust Multi-Agent Bandits Over Undirected Graphs

DANIEL VIAL, University of Texas at Austin, USA
SANJAY SHAKKOTTAI, University of Texas at Austin, USA
R. SRIKANT, University of Illinois Urbana-Champaign, USA

We consider a multi-agent multi-armed bandit setting in which n honest agents collaborate over a network to minimize regret but m malicious agents can disrupt learning arbitrarily. Assuming the network is the complete graph, existing algorithms incur $O((m + K/n) \log(T)/\Delta)$ regret in this setting, where K is the number of arms and Δ is the arm gap. For $m \ll K$, this improves over the single-agent baseline regret of $O(K \log(T)/\Delta)$.

In this work, we show the situation is murkier beyond the case of a complete graph. In particular, we prove that if the state-of-the-art algorithm is used on the undirected line graph, honest agents can suffer (nearly) linear regret until time is doubly exponential in K and n. In light of this negative result, we propose a new algorithm for which the i-th agent has regret $O((d_{\mathrm{mal}}(i) + K/n) \log(T)/\Delta)$ on any connected and undirected graph, where $d_{\mathrm{mal}}(i)$ is the number of i's neighbors who are malicious. Thus, we generalize existing regret bounds beyond the complete graph (where $d_{\mathrm{mal}}(i) = m$), and show the effect of malicious agents is entirely local (in the sense that only the $d_{\mathrm{mal}}(i)$ malicious agents directly connected to i affect its long-term regret).

CCS Concepts: • **Theory of computation** → **Online learning theory**; *Sequential decision making*; *Regret bounds*; **Multi-agent learning**; *Distributed algorithms*; *Adversarial learning*.

Additional Key Words and Phrases: multi-armed bandits; malicious agents

1 INTRODUCTION

Motivated by applications including distributed computing, social recommendation systems, and federated learning, a number of recent papers have studied multi-agent variants of the classical multi-armed bandit problem. Typically, these variants involve a large number of agents playing a bandit while communicating over a network. The goal is to devise communication protocols that allow the agents to efficiently amalgamate information, thereby learning the bandit's parameters more quickly than they could by running single-agent algorithms in isolation.

Among the many multi-agent variants considered in the literature (see Section 1.5), we focus on a particular setting studied in the recent line of work [18, 49, 52, 56]. In these papers, n agents play separate instances of the same K-armed bandit and are restricted to $o(T)$ pairwise and bit-limited communications per T arm pulls. We recount two motivating applications from this prior work.

EXAMPLE 1. *For an e-commerce site (e.g., Amazon), the agents model n servers choosing one of K products to show visitors to the site. The product selection problem can be viewed as a bandit – products*

Authors' addresses: Daniel Vial, dvial@utexas.edu, University of Texas at Austin, Austin, Texas, USA; Sanjay Shakkottai, sanjay.shakkottai@utexas.edu, University of Texas at Austin, Austin, Texas, USA; R. Srikant, rsrikant@illinois.edu, University of Illinois Urbana-Champaign, Urbana-Champaign, Illinois, USA.

are arms, while purchases yield rewards – and communication among the agents/servers is restricted by bandwidth.

EXAMPLE 2. *For a social recommendation site (e.g., Yelp), the agents represent n users choosing among K items, such as restaurants. This is analogously modeled as a bandit, and communication is limited because agents/users are exposed to a small portion of all reviews.*

To contextualize our contributions, we next discuss this line of work in more detail.

1.1 Fully cooperative multi-agent bandits

The goal of [18, 49, 52] is to devise fully cooperative algorithms for which the cumulative regret $R_T^{(i)}$ of each agent i is small (see (5) for the formal definition of regret). All of these papers follow a similar approach, which roughly proceeds as follows (see Section 3 for details):

- The arms are partitioned into n subsets of size $O(K/n)$, and each agent is assigned a distinct subset called a *sticky set*, which they are responsible for exploring.
- Occasionally ($o(T)$ times per T arm pulls), each agent i asks a random neighbor i' for an arm recommendation; i' responds with the arm they believe is best, which i begins playing.

For these algorithms, the regret analysis essentially contains two steps:

- First, the authors show that the agent (say, $i^\star$) with the true best arm in its sticky set eventually identifies it as such. Thereafter, a *gossip* process unfolds. Namely, $i^\star$ recommends the best arm to its neighbors, who recommend it to their neighbors, etc., spreading the best arm to all agents. The spreading time (and thus the regret before this time) is shown to be polynomial in K, n, and $1/\Delta$, where Δ is the gap in mean reward between the two best arms.
- Once the best arm spreads, agents play only it and their sticky sets, so long-term, they effectively face $O(K/n)$-armed bandits instead of the full K-armed bandit. By classical bandit results (see, e.g., [3]), this implies $O((K/n) \log(T)/\Delta)$ regret over horizon T.

Hence, summing up the two terms, [18, 49, 52] provide regret bounds of the form[1]

$$R_T^{(i)} = O\left(\frac{K}{n}\frac{\log T}{\Delta} + \text{poly}\left(K, n, \frac{1}{\Delta}\right)\right),\tag{1}$$

as compared to $O(K \log(T)/\Delta)$ regret for running a single-agent algorithm in isolation.

1.2 Robust multi-agent bandits on the complete graph

Despite these improved bounds, [18, 49, 52] require all agents to execute the prescribed algorithm, and in particular, to recommend best arm estimates to their neighbors. As pointed out in [56], this may be unrealistic: in Example 2, review spam can be modeled as bad arm recommendations, while in Example 1, servers may fail entirely. Hence, [56] considers a more realistic setting where n *honest agents* recommend best arm estimates but m *malicious agents* recommend arbitrarily. For this setting, the authors propose a robust version of the algorithm from [18] where honest agents *block* suspected malicious agents. More specifically, [56] considers the following blocking rule:

- If agent i' recommends arm k to honest agent i, but arm k subsequently performs poorly for i – in the sense that the upper confidence bound (UCB) algorithm does not select it sufficiently often – then i temporarily suspends communication with i'.

As shown in [56], this blocking scheme prevents each malicious agent from recommending more than $O(1)$ bad arms long-term, which (effectively) results in an $O(m + K/n)$-armed bandit ($O(m)$ malicious recommendations, plus the $O(K/n)$-sized sticky set). Under the assumption that honest

---

[1]More precisely, [18, 49] prove (1), while the K/n term balloons to $(K/n) + \log n$ in [52].

and malicious agents are connected by the *complete graph*, this allows [56] to prove

$$R_T^{(i)} = O\left(\left(\frac{K}{n} + m\right)\frac{\log T}{\Delta} + \text{poly}\left(K, n, m, \frac{1}{\Delta}\right)\right). \tag{2}$$

In [56], it is also shown that blocking is necessary: for any $n \in \mathbb{N}$, if even $m = 1$ malicious agent is present, the algorithm from [18] (which does not use blocking) incurs $\Omega(K \log(T)/\Delta)$ regret. Thus, one malicious agent negates the improvement over the single-agent baseline.

1.3 Objective and challenges

Our main goal is to generalize the results of [56] from the complete graph to the case where the honest agent subgraph is only connected and undirected. This is nontrivial because [56] relies *heavily* on the complete graph assumption. In particular, the analysis in [56] requires that $i^\star$ (the agent with the best arm in its sticky set) *itself* recommends the best arm to each of the other honest agents. In other words, each honest agent $i \neq i^\star$ relies on $i^\star$ to inform them of the best arm, which means $i^\star$ must be a neighbor of i. Thus, to extend (2) beyond complete graphs, we need to show a gossip process unfolds (like in the fully cooperative case): $i^\star$ recommends the best arm to its neighbors, who recommend it to their neighbors, etc., spreading it through the network.

The challenge is that, while blocking is necessary to prevent $\Omega(K \log(T)/\Delta)$ regret, it also causes honest agents to *accidentally block each other*. Indeed, due to the aforementioned blocking rule and the noisy rewards, they will block each other until they collect enough samples to reliably identify good arms. From a network perspective, accidental blocking means that edges in the subgraph of honest agents temporarily fail. Consequently, it is not clear if the best arm spreads to all honest agents, or if (for example) this subgraph eventually becomes disconnected, preventing the spread and causing the agents who do not receive the best arm to suffer $\Theta(T)$ regret.

Analytically, accidental blocking means we must deal with a gossip process over a *dynamic graph*. This process is extremely complicated, because the graph dynamics are driven by the bandit algorithms, which in turn affect the future evolution of the graph. Put differently, blocking causes the randomness of the communication protocol and that of the bandit algorithms to become *interdependent*. We note this does not occur for the original non-blocking algorithm, where the two sources of randomness can be cleverly decoupled and separately analyzed – see [18, Proposition 4]. Thus, in contrast to existing work, we need to analyze the interdependent processes directly.

1.4 Our contributions

Failure of the existing blocking rule: In Section 4, we show that the algorithm from [56] fails to achieve a regret bound of the form (2) for connected and undirected graphs in general. Toward this end, we define a natural "bad instance" in which $n = K$, the honest agent subgraph is an undirected line (thus connected), and all honest agents share a malicious neighbor. For this instance, we propose a malicious strategy that causes honest agents to repeatedly block one another, which results in the best arm spreading extremely slowly. More specifically, we show that if honest agents run the algorithm from [56], then the best arm does not reach honest agent n (the one at the end of the line) until time is *doubly exponential* in $n = K$. Note [56] shows the best arm spreads polynomially fast for the complete graph, so we demonstrate a doubly exponential slowdown for complete versus line graphs. This is rather surprising, because for classical rumor processes that do not involve bandits or blocking (see, e.g., [50]), the slowdown is only exponential (i.e., $\Theta(\log n)$ rumor spreading time on the complete graph versus $\Theta(n)$ on the line graph). As a consequence of

the slow spread, we show the algorithm from [56] suffers regret

$$R_T^{(n)} = \Omega \left(\min \left\{ \log(T) + \exp \left(\exp \left(\frac{n}{3} \right) \right), \frac{T}{\log^7 T} \right\} \right), \tag{3}$$

i.e., it incurs (nearly) linear regret until time $\Omega(\exp(\exp(n/3)))$ and thereafter incurs logarithmic regret but with a huge additive term (see Theorem 1).

Refined blocking rule: In light of this negative result, we propose a refined blocking rule in Section 5. Roughly, our rule is as follows: agent i blocks i' for recommending arm k if

- arm k performs poorly, i.e., it is not chosen sufficiently often by UCB,
- *and* agent i has not changed its own best arm estimate recently.

The second criterion is the key distinction from [56]. Intuitively, it says that agents should *not* block for seemingly-poor recommendations until they become confident that their *own* best arm estimates have settled on truly good arms. This idea is the main new algorithmic insight of the paper. It is directly motivated by the negative result of Section 4; see Remark 5.

Gossip despite blocking: Analytically, our main contribution is to show that, with our refined blocking rule, the best arm quickly spreads to all honest agents. The proof is quite involved; we provide an outline in Section 7. At a very high level, the idea is to show that honest agents using our blocking rule eventually stop blocking each other. Thereafter, we can couple the arm spreading process with a much more tractable noisy rumor process that involves neither bandits nor blocking (see Definition 1), and that is guaranteed to spread the best arm in polynomial time.

Regret upper bound: Combining our novel gossip analysis with some existing regret minimization techniques, we show in Section 5 that our refined algorithm enjoys the regret bound

$$R_T^{(i)} = O \left(\left(\frac{K}{n} + d_{\mathrm{mal}}(i) \right) \frac{\log T}{\Delta} + \mathrm{poly} \left(K, n, m, \frac{1}{\Delta} \right) \right), \tag{4}$$

where $d_{\mathrm{mal}}(i)$ is the number of malicious neighbors of i (see Theorem 2). Thus, our result generalizes (2) from the complete graph (where $d_{\mathrm{mal}}(i) = m$) to connected and undirected graphs. Moreover, note the leading $\log T$ term in (4) is *entirely local* – only the $d_{\mathrm{mal}}(i)$ malicious agents directly connected to i affect its long-term regret. For example, in the sparse regime $d_{\mathrm{mal}}(i) = O(1)$, our $\log T$ term matches the one in (1) up to constants, which (we recall) [18, 49] proved in the case where there are no malicious agents anywhere in the network. In fact, for honest agents i with $d_{\mathrm{mal}}(i) = 0$, we can prove that the $\log T$ term in our regret bound matches the corresponding term from [18], *including* constants (see Corollary 2). In other words, we show that for large horizons T, the effects of malicious agents do not propagate beyond one-step neighbors. Furthermore, we note that the additive term in (4) is polynomial in all parameters, whereas for the existing algorithm it can be doubly exponential in K and n, as shown in (3) and discussed above.

Numerical results: In Section 6, we replicate the experiments from [56] and extend them from the complete graph to $G(n + m, p)$ random graphs. Among other findings, we show that for $p = 1/2$ and $p = 1/4$, respectively, the algorithm from [56] can perform *worse* than the non-blocking algorithm from [18] and the single-agent baseline, respectively. In other words, the existing blocking rule becomes a *liability* as p decreases from the extreme case $p = 1$ considered in [56]. In contrast, we show that our refined rule has lower regret than [18] across the range of p tested. Additionally, it outperforms [56] on average for all but the largest p and has much lower variance for smaller p.

Summary: Ultimately, the high-level messages of this paper are twofold:

- In multi-agent bandits with malicious agents, we can devise algorithms that simultaneously (1) learn useful information and spread it through the network via gossip, and (2) learn who

is malicious and block them to mitigate the harm they cause. Moreover, this harm is local in the sense that it only affects one-hop neighbors.

- However, blocking must be done carefully; algorithms designed for the complete graph may spread information extremely slowly on general graphs. In particular, the slowdown can be *doubly* exponential, much worse than the exponential slowdown of simple rumor processes.

1.5 Other related work

In addition to the paper [56] discussed above, several others have considered multi-agent bandits where some of the agents are uncooperative. In [6], the honest agents face a non-stochastic (i.e., adversarial) bandit [4] and communicate at every time step, in contrast to the stochastic bandit and limited communication of our work. The authors of [48] consider the objective of best arm identification [2] instead of cumulative regret. Most of their paper involves a different communication model where the agents/clients collaborate via a central server; Section 6 studies a "peer-to-peer" model which is closer to ours but requires additional assumptions on the number of malicious neighbors. A different line of work considers the case where an adversary can corrupt the observed rewards (see, e.g., [11, 12, 25, 26, 29, 33, 40, 41, 44], and the references therein), which is distinct from the role that malicious agents play in our setting.

For the fully cooperative case, there are several papers with communication models that differ from the aforementioned [18, 49, 52]. For example, agents in [15, 17] broadcast information instead of exchanging pairwise arm recommendations, communication in [34, 36, 47] is more frequent, the number of transmissions in [45] depends on Δ^{-1} so could be large, and agents in [37] exchange arm mean estimates instead of (bit-limited) arm indices.

More broadly, other papers have studied fully cooperative variants of different bandit problems. These include minimizing simple instead of cumulative regret (e.g., [28, 54]), minimizing the total regret across agents rather than ensuring all have low regret (e.g., [22, 57]), contextual instead of multi-armed bandits (e.g., [19, 23, 24, 35, 55]), adversarial rather than stochastic bandits (e.g., [7, 16, 31]), and bandits that vary across agents (e.g., [10, 53, 58]). Another long line of work features collision models where rewards are lower if multiple agents simultaneously pull the same arm (e.g., [1, 5, 13, 21, 30, 42, 43, 46, 51]), unlike our model. Along these lines, other reward structures have been studied, such as reward being a function of the agents' joint action (e.g., [8, 9, 32]).

1.6 Organization

The rest of the paper is structured as follows. We begin in Section 2 with definitions. In Section 3, we introduce the algorithm from [56]. Sections 4 and 5 discuss the existing and proposed blocking rules. Section 6 contains experiments. We discuss our analysis in Section 7 and close in Section 8.

2 PRELIMINARIES

Communication network: Let $G = ([n + m], E)$ be an undirected graph with vertices $[n + m] = \{1, \ldots, n + m\}$. We call $[n]$ the *honest agents* and assume they execute the forthcoming algorithm. The remaining agents are termed *malicious*; their behavior will be specified shortly. For instance, honest and malicious agents represent functioning and failed servers in Example 1. The edge set E encodes which agents are allowed to communicate, e.g., if $(i, i') \in E$, the i-th and i'-th servers can communicate in the forthcoming algorithm.

Denote by $E_{\text{hon}} = \{(i, i') \in E : i, i' \in [n]\}$ the edges between honest agents and $G_{\text{hon}} = ([n], E_{\text{hon}})$ the honest agent subgraph. For each $i \in [n]$, we let $N(i) = \{i' \in [n + m] : (i, i') \in E\}$ denote its neighbors, $N_{\text{hon}}(i) = N(i) \cap [n]$ its honest neighbors, and $N_{\text{mal}}(i) = N(i) \setminus [n]$ its malicious neighbors. We write $d(i) = |N(i)|$, $d_{\text{hon}}(i) = |N_{\text{hon}}(i)|$, and $d_{\text{mal}}(i) = |N_{\text{mal}}(i)|$ for the associated
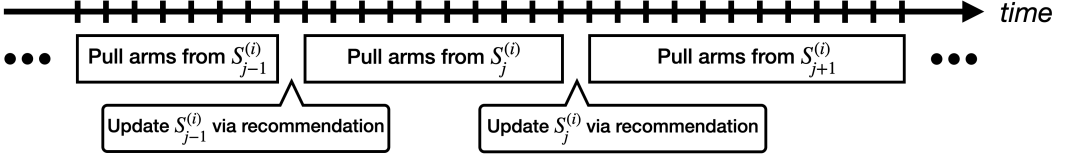
Fig. 1. Illustration of the phases in Algorithm 1; see beginning of Section 3.1 for details.

degrees, and $\bar{d} = \max_{i \in [n]} d(i)$, $\bar{d}_{\text{hon}} = \max_{i \in [n]} d_{\text{hon}}(i)$, and $\bar{d}_{\text{mal}} = \max_{i \in [n]} d_{\text{mal}}(i)$ for the maximal degrees. We make the following assumption, which generalizes the complete graph case of [56].

ASSUMPTION 1. *The honest agent subgraph G_{hon} is connected, i.e., for any $i, i' \in [n]$, there exists $l \in \mathbb{N}$ and $i_0, i_1, \ldots, i_l \in [n]$ such that $i_0 = i$, $(i_{j-1}, i_j) \in E_{hon} \ \forall \ j \in [l]$, and $i_l = i'$.*

Multi-armed bandit: We consider the standard stochastic multi-armed bandit [38, Chapter 4]. Denote by $K \in \mathbb{N}$ the number of arms and $[K]$ the set of arms. For each $k \in [K]$, we let ν_k be a probability distribution over $\mathbb{R}$ and $\{X_{i,t}(k)\}_{i \in [n], t \in \mathbb{N}}$ an i.i.d. sequence of *rewards* sampled from ν_k. The interpretation is that, if the i-th honest agent chooses the k-th arm at time t, it earns reward $X_{i,t}(k)$. The objective (to be formalized shortly) is reward maximization. In Example 2, for instance, $[K]$ represents the set of restaurants in a city, and the reward $X_{i,t}(k)$ quantifies how much person i enjoys restaurant k if they dine there on day t.

For each arm $k \in [K]$, we let $\mu_k = \mathbb{E}[X_{i,t}(k)]$ denote the corresponding expected reward. Without loss of generality, we assume the arms are labeled such that $\mu_1 \geq \cdots \geq \mu_K$. We additionally assume the following, which generalizes the $\nu_k = \text{Bernoulli}(\mu_k)$ and $\mu_1 > \mu_2$ setting of [56]. Notice that under this assumption, the *arm gap* $\Delta_k \triangleq \mu_1 - \mu_k$ is strictly positive.

ASSUMPTION 2. *Rewards are $[0, 1]$-valued, i.e., for each $k \in [K]$, ν_k is a distribution over $[0, 1]$. Furthermore, the best arm is unique, i.e., $\mu_1 > \mu_2$.*

Objective: For each $i \in [n]$ and $t \in \mathbb{N}$, let $I_t^{(i)} \in [K]$ denote the arm chosen by honest agent i at time t. Our goal is to minimize the *regret* $R_T^{(i)}$, which is the expected additive loss in cumulative reward for agent i's sequence of arm pulls $\{I_t^{(i)}\}_{t=1}^T$ compared to the optimal policy that always chooses the best arm 1. More precisely, we define regret as follows:

$$R_T^{(i)} \triangleq \sum_{t=1}^T \mathbb{E}\left[X_{i,t}(1) - X_{i,t}(I_t^{(i)})\right] = \sum_{t=1}^T \mathbb{E}\left[\mu_1 - \mu_{I_t^{(i)}}\right] = \sum_{t=1}^T \mathbb{E}\left[\Delta_{I_t^{(i)}}\right]. \tag{5}$$

3 ALGORITHM

We next discuss the algorithm from [56] (Algorithm 1 below), which modifies the one from [18] to include blocking. For ease of exposition, we begin by discussing the key algorithmic design principles from [18] in Section 3.1. We then define Algorithm 1 formally in Section 3.2. Finally, we introduce and discuss one additional assumption in Section 3.3.

3.1 Key ideas of the non-blocking algorithm

We assume $m = 0$ this subsection and describe the non-blocking algorithm from [18].
  • **Phases:** In [18], the time steps $1, \ldots, T$ are grouped into *phases*, whose role is twofold. First, within the j-th phase, the i-th honest agent only pulls arms belonging to a particular subset $S_j^{(i)} \subset [K]$. We call these *active sets* and detail their construction next. Second, at the end of the j-th phase, the agents construct new active sets by exchanging *arm recommendations*
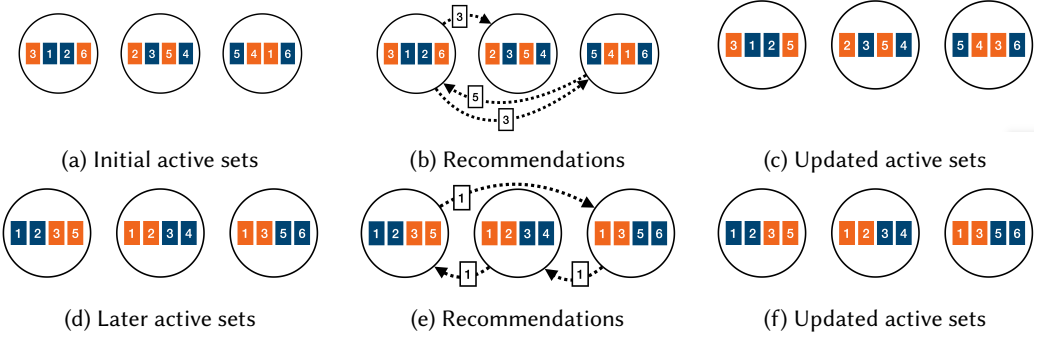
Fig. 2. Illustration of the active sets in Algorithm 1; see Example 3 for details.

with neighbors, in a manner to be described shortly. See Figure 1 for a pictorial description. Notice that the phase durations are increasing, which will be discussed in Section 3.2.

- **Active sets:** The active set $S_j^{(i)}$ will always contain a subset of arms $\hat{S}^{(i)} \subset [K]$ that does not vary with the phase j. Following [18, 56], we call $\hat{S}^{(i)}$ the *sticky set* and its elements *sticky arms*. The sticky sets ensure that each arm is explored by some agent, as will be seen in the forthcoming example. In addition, $S_j^{(i)}$ will contain two non-sticky arms that are dynamically updated across phases j based on arm recommendations from neighbors.

- **Arm recommendations:** After the j-th phase, each agent i contacts a random neighbor, who responds with whichever of their active arms performed "best" in the current phase. Upon receiving this recommendation, i adds it to its active set and discards whichever currently-active non-sticky arm (i.e., whichever element of $S_j^{(i)} \setminus \hat{S}^{(i)}$) performed "worse". (We quantify "best" and "worse" in the formal discussion of Section 3.2.)

EXAMPLE 3. *Each subfigure of Figure 2 depicts $n = 3$ honest agents as circles and their active sets as rectangles. The blue rectangles are sticky sets, the orange rectangles are non-sticky arms, and the arms are sorted by performance. For example, the left agent in Figure 2a has sticky set $\{1, 2\}$ and active set $\{1, 2, 3, 6\}$ and believes arm 3 to be the best of these. Note the blue sticky sets partition $[K] = [6]$, so at each phase, each arm is active for some agent. This ensures the best arm is never permanently discarded during the arm recommendations discussed above. Figure 2b shows agents recommending the active arms they believe are best, and Figure 2c depicts the updated active sets. For instance, the left agent replaces its worse non-sticky arm 6 with the recommendation 5. Figure 2d shows a later phase where the best arm 1 has spread to all agents, who have all identified it as such. Thereafter, all agents recommend 1, so the active set remains fixed (Figures 2e and 2f). Hence, all agents eventually exploit the best arm while only exploring a subset of the suboptimal arms (three instead of five here).*

3.2 Formal definition of the blocking algorithm

The algorithm in [56] supplements the one from Section 3.1 with a blocking procedure. Specifically, honest agents run the algorithm from [18] while maintaining *blocklists* of neighbors they are unwilling to communicate with. This approach is defined in Algorithm 1 and detailed next.

Inputs (Line 1): The first input is a standard UCB exploration parameter $\alpha > 0$, which will be discussed shortly. The input $\beta > 1$ controls the lengths of the phases; namely, the j-th phase encompasses times $1 + A_{j-1}, \ldots, A_j$, where $A_j \triangleq \lceil j^\beta \rceil$. Note the phase duration $A_j - A_{j-1} = O(j^{\beta-1})$

Algorithm 1: Multi-agent bandits with blocking (executed by $i \in [n]$)

---

1 **Input:** UCB parameter $\alpha > 0$, phase parameter $\beta > 1$, sticky arms $\hat{S}^{(i)}$ with $|\hat{S}^{(i)}| = S \leq K - 2$

2 Initialize $A_j = \lceil j^\beta \rceil$, $P_j^{(i)} = \emptyset \; \forall j \in \mathbb{N}$ (communication times and blocklists)

3 Set $j = 1$ (current phase), let $\{U_j^{(i)}, L_j^{(i)}\} \subset [K] \setminus \hat{S}^{(i)}$ be two distinct non-sticky arms and
   $S_j^{(i)} = \hat{S}^{(i)} \cup \{U_j^{(i)}, L_j^{(i)}\}$ the initial active set

4 **for** $t \in \mathbb{N}$ **do**

5 $\quad$ Pull $I_t^{(i)} = \arg\max_{k \in S_j^{(i)}} \left(\hat{\mu}_k^{(i)}(t-1) + \sqrt{\alpha \log(t)/T_k^{(i)}(t-1)} \right)$ (UCB over active set)

6 $\quad$ **if** $t = A_j$ *(if communication occurs at this time)* **then**

7 $\quad\quad$ $B_j^{(i)} = \arg\max_{k \in S_j^{(i)}} \left(T_k^{(i)}(A_j) - T_k^{(i)}(A_{j-1}) \right)$ (most played active arm in this phase)

8 $\quad\quad$ $\{P_{j'}^{(i)}\}_{j'=j}^{\infty} \leftarrow$ Update-Blocklist$(\{P_{j'}^{(i)}\}_{j'=j}^{\infty})$ ([56] uses Alg. 3; we propose Alg. 4)

9 $\quad\quad$ $(H_j^{(i)}, R_j^{(i)}) =$ Get-Recommendation$(i, j, P_j^{(i)})$ (see Alg. 2)

10 $\quad\quad$ **if** $R_j^{(i)} \notin S_j^{(i)}$ *(if recommendation not already active)* **then**

11 $\quad\quad\quad$ $U_{j+1}^{(i)} = \arg\max_{k \in \{U_j^{(i)}, L_j^{(i)}\}} \left(T_k^{(i)}(A_j) - T_k^{(i)}(A_{j-1}) \right)$ (best non-sticky active arm)

12 $\quad\quad\quad$ $L_{j+1}^{(i)} = R_j^{(i)}$ (replace worst non-sticky active arm with recommendation)

13 $\quad\quad\quad$ $S_{j+1}^{(i)} = \hat{S}^{(i)} \cup \{U_{j+1}^{(i)}, L_{j+1}^{(i)}\}$ (new active set is sticky set and two non-sticky arms)

14 $\quad\quad$ **else**

15 $\quad\quad\quad$ $S_{j+1}^{(i)} = S_j^{(i)}$ (keep the same active set, since recommendation is already active)

16 $\quad\quad$ $j \leftarrow j + 1$ (increment phase)

Algorithm 2: $(H_j^{(i)}, R_j^{(i)}) =$ Get-Recommendation$(i, j, P_j^{(i)})$ (black box to $i \in [n]$)

---

1 **Input:** Agent $i \in \{1, \ldots, n\}$, phase $j \in \mathbb{N}$, blocklist $P_j^{(i)}$

2 Sample $H_j^{(i)}$ from $N(i) \setminus P_j^{(i)}$ (non-blocked neighbors) uniformly at random

3 **if** $H_j^{(i)} \leq n$ *(if the sampled agent is honest)* **then**

4 $\quad$ Set $R_j^{(i)} = B_j^{(H_j^{(i)})}$ (honest agents recommend most played arm from this phase)

5 **else**

6 $\quad$ Choose $R_j^{(i)} \in [K]$ arbitrarily (malicious agents recommend arbitrary arms)

7 **Output:** $(H_j^{(i)}, R_j^{(i)})$

---

grows with j, as shown in Figure 1. The final input is an S-sized sticky set $\hat{S}^{(i)}$ ($S = 2$ in Example 3), which, as in [56], we assume are provided to the agents (see Section 3.3 for more details).

Initialization (Lines 2-3): To start, i initializes the times A_j at which the j-th phase ends, along with the blocklist $P_j^{(i)}$. Additionally, i chooses two distinct (but otherwise arbitrary) non-sticky arms $U_1^{(i)}$ and $L_1^{(i)}$ and constructs the active set $S_1^{(i)} = \hat{S}^{(i)} \cup \{U_1^{(i)}, L_1^{(i)}\}$. Notice that the active set contains the sticky set and two arms that depend on the phase, as described in Section 3.1.

UCB over the active set (Line 5): As was also mentioned in Section 3.1, i only pulls arms from its current active set $S_j^{(i)}$. More specifically, at each time t during phase j, i chooses the active arm

$I_t^{(i)} \in S_j^{(i)}$ that maximizes the UCB in Line 5 (see [38, Chapters 7-10] for background). Here $T_k^{(i)}(t-1)$ and $\hat{\mu}_k^{(i)}(t-1)$ are the number of pulls of k and the empirical mean of those pulls, i.e.,

$$T_k^{(i)}(t-1) = \sum_{s \in [t-1]} \mathbb{1}(I_s^{(i)} = k), \quad \hat{\mu}_k^{(i)}(t-1) = \frac{1}{T_k^{(i)}(t-1)} \sum_{s \in [t-1]: I_s^{(i)} = k} X_{i,s}(k),$$

where $X_{i,s}(k) \sim \nu_k$ as in Section 2 and $\mathbb{1}$ is the indicator function.

Best arm estimate (Line 7): At the end of phase j (i.e., when $t = A_j$), i defines its best arm estimate $B_j^{(i)}$ as the active arm it played the most in phase j. The intuition is that, for large horizons, the arm chosen most by UCB is a good estimate of the true best arm [14]. Thus, because phase lengths are increasing (see Figure 1), $B_j^{(i)}$ will be a good estimate of the best active arm for large j.

Blocklist update (Line 8): Next, i calls the Update-Blocklist subroutine to update its blocklist $P_j^{(i)}$. The implementation of this subroutine is the key distinction between [56] and our work. We discuss the respective implementations in Sections 4 and 5, respectively.

Arm recommendations (Line 9): Having updated $P_j^{(i)}$, i requests an arm recommendation $R_j^{(i)}$ via Algorithm 2. Algorithm 2 is a black box (i.e., i provides the input and observes the output), which samples a random non-blocked neighbor $H_j^{(i)} \in N(i) \setminus P_j^{(i)}$. If $H_j^{(i)}$ is honest, it recommends its best arm estimate, while if malicious, it recommends an arbitrary arm.[2]

Updating the active set (Lines 10-15): Finally, i updates its active set as in Section 3.1. In particular, if the recommendation $R_j^{(i)}$ is not currently active[3], i defines $U_{j+1}^{(i)}$ to be the non-sticky arm that performed better in phase j, in the sense that UCB chose it more often (following the above intuition from [14]). The other non-sticky arm $L_{j+1}^{(i)}$ becomes the recommendation $R_j^{(i)}$, and the new active set becomes $S_{j+1}^{(i)} = \hat{S}^{(i)} \cup \{U_{j+1}^{(i)}, L_{j+1}^{(i)}\}$ (the sticky set and two other arms, as above).

3.3 Additional assumption

Observe that Algorithm 1 does not preclude the case where the best arm is not in any honest agent's sticky set, i.e., $1 \notin \cup_{i=1}^{n} \hat{S}^{(i)}$. In this case, the best arm may be permanently discarded, which causes linear regret even in the absence of malicious agents. For example, this would occur if 1 was not a sticky arm for the left agent in Figure 2 (since the right agent discards 1 in Figure 2c). To prevent this situation, we will follow [18, 49, 52, 56] in assuming the following.

ASSUMPTION 3. *There exists $i^\star \in [n]$ with the best arm in its sticky set, i.e., $1 \in \hat{S}^{(i^\star)}$.*

REMARK 1. *As discussed in [18, Appendix N], Assumption 3 holds with high probability if S (the size of the sticky set input to Algorithm 1) is set to $\tilde{\Theta}(K/n)$ and each sticky set $\hat{S}^{(i)}$ is sampled uniformly at random from the S-sized subsets of $[K]$.*

REMARK 2. *The choice $S = \tilde{\Theta}(K/n)$ from Remark 1 requires the honest agents to know an order-accurate estimate of n, i.e., they need to know some $n' = \tilde{\Theta}(n)$ in order to set $S = \tilde{\Theta}(K/n')$ and ensure that $S = \tilde{\Theta}(K/n)$. As discussed in [56, Remark 7], this amounts to knowing order-accurate estimates of $n + m$ and $n/(n + m)$. The former quantity is the total number of agents, knowledge of which is rather benign and is also assumed in the fully-cooperative setting [18, 49, 52]. The latter requires the agents to know that, e.g., half of the others are honest, which is similar in spirit to the assumptions in related problems regarding social learning in the presence of adversarial agents (e.g., [39]).*

---

[2]Technically, malicious recommendations need to be measurable; see [56, Section 3] for details.

[3]If the recommendation is currently active, the active set remains unchanged (see Line 15 of Algorithm 1).

REMARK 3. *Alternatively, we can avoid Assumption 3 entirely by defining the set of the arms to be those initially known by the honest agents (i.e., their sticky sets), rather than sampling the sticky sets from a larger "base set" as in Remark 1. In this alternative model, the honest agents aim to identify and spread through the network whichever of the initially-known arms is best, similar to what happens on platforms like Yelp (see Example 2). In contrast, the Section 2 model allows for the pathological case where the base set contains a better arm than any initially known to honest agents (e.g., where no honest Yelp user has ever dined at the best restaurant). Coping with these pathological cases either requires Assumption 3, or another mode of exploration (i.e., exploration of base arms) that obfuscates the key point of our work (collaborative bandit exploration amidst adversaries). For these reasons, we prefer the alternative model, but to enable a cleaner comparison with prior work [56], we restrict attention to the Section 2 model (which generalizes that of [56]).*

4 EXISTING BLOCKING RULE

We can now define the blocking approach from [56], which is provided in Algorithm 3. In words, the rule is as follows: if the recommendation $R_{j-1}^{(i)}$ from phase $j-1$ is not i's most played arm in the subsequent phase j, then the agent $H_{j-1}^{(i)}$ who recommended it is added to the blocklists $P_j^{(i)}, \ldots, P_{j^\eta}^{(i)}$, where $\eta > 1$ is a tuning parameter. By Algorithm 2, this means i blocks (i.e., does not communicate with) $H_{j-1}^{(i)}$ until phase $j^\eta + 1$ (at the earliest). Thus, agents block others whose recommendations perform poorly – in the sense that UCB does not play them often – and the blocking becomes more severe as the phase counter j grows. See [56, Remark 4] for further intuition.

Algorithm 3: $\{P_{j'}^{(i)}\}_{j'=j}^\infty$ = Update-Blocklist (executed by $i \in [n]$, existing rule from [56])

---

1 **if** $j > 1$ and $B_j^{(i)} \neq R_{j-1}^{(i)}$ (*if previous recommendation not most played*) **then**
2 $\quad$ $P_{j'}^{(i)} \leftarrow P_{j'}^{(i)} \cup \{H_{j-1}^{(i)}\} \ \forall j' \in \{j, \ldots, \lceil j^\eta \rceil\}$ (*block the recommender until phase j^η*)

In the remainder of this section, we define a bad instance (Section 4.1) on which this blocking rule provably fails (Section 4.2). Our goal here is to demonstrate a single such instance in order to show this blocking rule must be refined. Therefore, we have opted for a concrete example, which includes some numerical constants (e.g., 13/15 in (6), the 7 in the $\log^7 T$ term in Theorem 1, etc.) that have no particular meaning. Nevertheless, the instance can be generalized; see Remark 4.

4.1 Bad instance

The network and bandit for the bad instance are as follows:

- There are an even number of honest agents (at least four) arranged in a line, increasing in index from left to right, and there is a malicious agent connected to each of the honest ones. Mathematically, we have $n \in \{4, 6, 8, \ldots\}$, $m = 1$, and $E = \{(i, i+1)\}_{i=1}^{n-1} \cup \{(i, n+1)\}_{i=1}^{n}$.
- There are $K = n$ arms that generate deterministic rewards (i.e., $\nu_k = \delta_{\mu_k}$) with

$$\mu_1 = 1, \quad \mu_k = \frac{13}{15} + \sum_{h=1}^{(n/2)-k} 2^{-2^{h+1}} \ \forall k \in \{2, \ldots, n/2\}, \quad \mu_k = 0 \ \forall k > n/2. \qquad (6)$$

Intuitively, there are three sets of arms: the best arm, $(n/2) - 1$ mediocre arms, and $n/2$ bad arms. We provide further intuition in the forthcoming proof sketch. For now, we highlight three key properties. First, the gap from mediocre to bad arms is constant, i.e., $\mu_{k_1} - \mu_{k_2} \geq 13/15$ when $k_1 \leq n/2 < k_2$. Second, the gaps between mediocre arms are doubly exponentially

small, i.e., $\mu_k - \mu_{k+1} = 2^{-2^{(n/2)-k+1}}$ for $k \in \{2, \dots, (n/2) - 1\}$. Third, the gap Δ_2 from the best to the mediocre arms is at least $1/15$, as shown in Appendix C.

OBSERVATION 1. *Since rewards are deterministic, the most played arm $B_{j+1}^{(i)}$ in phase $j + 1$ is a deterministic function of the number of plays of the active arms at the beginning of the phase, i.e., of the set $\{T_k^{(i)}(A_j)\}_{k \in S_{j+1}^{(i)}}$. Hence, when the j-th recommendation is already active (i.e., when $R_j^{(i)} \in S_j^{(i)}$, which implies $S_{j+1}^{(i)} = S_j^{(i)}$ in Algorithm 1), $B_{j+1}^{(i)}$ is a function of $\{T_k^{(i)}(A_j)\}_{k \in S_j^{(i)}}$, which is information available to the malicious agent at the j-th communication time A_j. Consequently, the malicious agent can always recommend some $R_j^{(i)} \in S_j^{(i)}$ such that $B_{j+1}^{(i)} = R_j^{(i)}$ to avoid being blocked by i.*

We make the following assumptions on Algorithms 1 and 2:
- The parameters in Algorithm 1 are $\alpha = 4$ and $\beta = 2$, while $\eta = 2$ in Algorithm 3.
- Sticky sets have size $S = 1$ and for any $i \in \{1 + n/2, \dots, n\}$, i's initial active set satisfies $\min S_1^{(i)} > n/2$. Thus, active sets contain three arms, and the right half of the honest agents are initially only aware of the bad arms, i.e., of those that provide no reward.

REMARK 4. *Note that Assumptions 1-3 all hold for this instance, and the choices $\alpha = 4$ and $\beta = \eta = 2$ are used for the complete graph experiments in [56]. Additionally, the instance can be significantly generalized – the key properties are that K and n have the same scaling, the gaps from mediocre arms to others are constant, the gaps among mediocre arms are doubly exponentially small, and a constant fraction of agents on the right initially only have bad arms active.*

Finally, we define a particular malicious agent strategy. Let $J_1 = 2^8$ and inductively define $J_{l+1} = (J_l + 2)^2$ for each $l \in \mathbb{N}$. Then the malicious recommendations are as follows:
- If $j = J_l$ and $i \in \{l + 1 + n/2, l + 2 + n/2\}$ for some $l \in [(n/2) - 1]$, set $R_j^{(i)} = 1 - l + n/2$.
- Otherwise, let $R_j^{(i)} \in S_j^{(i)}$ be such that $B_{j+1}^{(i)} = R_j^{(i)}$ (see Observation 1).

Similar to the arm means, we will wait for the proof sketch to explain this strategy in more detail. For now, we only mention that the phases J_l grow doubly exponentially, i.e.,

$$J_{l+1} = (J_l + 2)^2 > J_l^2 > \cdots > J_1^{2^l} \ \forall l \in \mathbb{N}. \tag{7}$$

4.2 Negative result

We can now state the main result of this section. It shows that if the existing blocking rule from [56] is used on the above instance, then the honest agent n at the end of the line suffers nearly linear regret $\tilde{\Omega}(T)$ until time T exceeds a doubly exponential function of $n = K$.

THEOREM 1. *If we run Algorithm 1 and use Algorithm 3 as the `Update-Blocklist` subroutine with the parameters and problem instance described in Section 4, then*

$$R_T^{(n)} = \Omega \left(\min \left\{ \log(T) + \exp \left(\exp \left(n/3 \right) \right), T/\log^7 T \right\} \right).$$

PROOF SKETCH. We provide a complete proof in Appendix C but discuss the intuition here.
- First, suppose honest agent $1 + n/2$ contacts the malicious agent $n + 1$ at all phases $j \in [J_1 - 1]$ (this occurs with constant probability since J_1 is constant). Then the right half of honest agents (i.e., agents $1 + n/2, \dots, n$) only have bad arms (i.e., arms $1 + n/2, \dots, n$) in their active sets at phase J_1. This is because their initial active sets only contain such arms (by assumption), $n + 1$ only recommends currently-active arms before J_1, and no arm recommendations flow from the left half of the graph to the right half (they need to first be sent from $n/2$ to $1 + n/2$, but we are assuming the latter only contacts $n + 1$ before J_1).

- Now consider phase J_1. With constant probability, $1 + n/2$ and $2 + n/2$ both contact $n + 1$, who recommends a currently active (thus bad) arm and the mediocre arm $n/2$, respectively. Then, again with constant probability, $2 + n/2$ contacts $1 + n/2$ at the next phase $J_1 + 1$; $1 + n/2$ only has bad arms active and thus recommends a bad arm. Therefore, during phase $J_1 + 2$, agent $2 + n/2$ has the mediocre arm $n/2$ and some bad recommendation from $1 + n/2$ in its active set. The inverse gap squared between these arms is constant, thus less than the length of phase $J_1 + 2$ (for appropriate J_1), so by standard bandit arguments (basically, noiseless versions of best arm identification results from [14]), $n/2$ will be most played. Consequently, by the blocking rule in Algorithm 3, $2 + n/2$ blocks $1 + n/2$ until phase $(J_1 + 2)^2 = J_2$.
- We then use induction. For each $l \in [(n/2) - 1]$ ($l = 1$ in the previous bullet), suppose $l + 1 + n/2$ blocks $l + n/2$ between phases $J_l + 2$ and J_{l+1}. Then during these phases, no arm recommendations flow past $l + n/2$, so agents $\geq l + 1 + n/2$ only play arms $\geq 1 - l + n/2$. At phase J_{l+1}, the malicious agent recommends $k \geq 1 - l + n/2$ and $-l + n/2$ to agents $l + 1 + n/2$ and $l + 2 + n/2$, respectively, and at the subsequent phase $J_{l+1} + 1$, $l + 1 + n/2$ recommends $k' \geq l + 1 + n/2$ to $l + 2 + n/2$. Similar to the previous bullet, we then show $l + 2 + n/2$ plays arm $-l + n/2$ more than k' during phase $J_{l+1} + 2$ and thus blocks $l + 1 + n/2$ until $(J_{l+1} + 2)^2 = J_{l+2}$, completing the inductive step. The proof that $-l + n/2$ is played more than k' during phase $J_{l+1} + 2$ again follows from noiseless best arm identification, although unlike the previous bullet, the relevant arm gap is no longer constant (both could be mediocre arms). However, we chose the mediocre arm means such that their inverse gap squared is at most doubly exponential in l, so by (7), the length of phase J_{l+1} dominates it.

In summary, we show that due to blocking amongst honest agents, $l + 1 + n/2$ does not receive arm $1 - l + n/2$ until phase J_l, given that some constant probability events occur at each of the times $J_1, \ldots, J_l$. This allows us to show that, with probability at least $\exp(-\Omega(n))$, agent n does not receive the best arm until phase $J_{n/2} = \exp(\exp(\Omega(n)))$, and thus does not play the best arm until time $\exp(\exp(\Omega(n)))$ in expectation. Since Δ_2 is constant, we can lower bound regret similarly. □

5 PROPOSED BLOCKING RULE

To summarize the previous section, we showed that the existing blocking rule (Algorithm 3) may result in honest agents blocking too aggressively, which causes the best arm to spread very slowly. In light of this, we propose a relaxed blocking criteria (see Algorithm 4): at phase j, agent i will block the agent $H_{j-1}^{(i)}$ who recommended arm $R_{j-1}^{(i)}$ at the previous phase $j - 1$ if

$$T_{R_{j-1}^{(i)}}^{(i)}(A_j) \leq \kappa_j \quad \text{and} \quad B_j^{(i)} = B_{j-1}^{(i)} = \cdots = B_{\lfloor \theta_j \rfloor}^{(i)}, \tag{8}$$

where $\kappa_j \leq A_j$ and $\theta_j \leq j$ are tuning parameters. Thus, i blocks if *both* of the following occur:

- The recommended arm $R_{j-1}^{(i)}$ performs poorly, in the sense that UCB has not chosen it sufficiently often (i.e., at least κ_j times) by the end of phase j.
- Agent i has not changed its own best arm estimate since phase θ_j. Intuitively, this can be viewed as a confidence criterion: if instead i has recently changed its estimate, then i is currently unsure which arm is best, so should not block for recommendations that appear suboptimal at first glance (i.e., those for which the first criterion in (8) may hold).

Algorithm 4: $\{P_{j'}^{(i)}\}_{j'=j}^{\infty}$ = Update-Blocklist (executed by $i \in [n]$, proposed rule)

---

1 **if** $j > 1$ *and* (8) *holds* **then** $P_{j'}^{(i)} \leftarrow P_{j'}^{(i)} \cup \{H_{j-1}^{(i)}\} \ \forall j \in \{j, \ldots, \lceil j^\eta \rceil\}$ (block recommender);

REMARK 5. *The first criterion in* (8) *is a natural relaxation of demanding the recommended arm be most played. The second is directly motivated by the negative result from Section 4. In particular, recall from the Theorem 1 proof sketch that $l + 1 + n/2$ blocked $l + n/2$ shortly after receiving a new mediocre arm from the malicious agent. Thus, blocking amongst honest agents was always precipitated by the blocking agent changing its best arm estimate. The second criterion in* (8) *aims to avoid this.*

REMARK 6. *Our proposed rule has two additional parameters compared to the existing one: κ_j and θ_j. For our theoretical results, these will be specified in Theorem 2; for experiments, they are discussed in Section 6. For now, we only mention that they should satisfy two properties. First, κ_j should be $o(A_j)$, so that the first criterion in* (8) *dictates a sublinear number of plays. Second, $j - \theta_j$ should grow with j, since (as discussed above) the second criterion represents the confidence in the best arm estimate, which grows as the number of reward observations increases.*

In the remainder of this section, we introduce a further definition (Section 5.1), provide a general regret bound under our blocking rule (Section 5.2), and discuss some special cases (Section 5.3).

5.1 Noisy rumor process

As discussed in Section 1.4, we will show that under our proposed rule (1) honest agents eventually stop blocking each other, and (2) honest agents with the best arm active will eventually recommend it to others. Thereafter, we essentially reduce the arm spreading process to a much simpler rumor process in which each honest agent i contacts a uniformly random neighbor i' and, if i' is an honest agent who knows the rumor (i.e., if the best arm is active for i'), then i' informs i of the rumor (i.e., i' recommends the best arm to i). The only caveat is that we make *no assumption* on the malicious agent arm recommendations, so we have *no control* over whether or not they are blocked. In other words, the rumor process unfolds over a dynamic graph, where edges between honest and malicious agents may or may not be present, and we have no control over these dynamics.

In light of this, we take a worst-case view and lower bound the arm spreading process with a noisy rumor process that unfolds on the (static) honest agent subgraph. More specifically, we consider the process $\{\bar{I}_j\}_{j=0}^{\infty}$ that tracks the honest agents informed of the rumor. Initially, only $i^\star$ (the agent from Assumption 3) is informed (i.e., $\bar{I}_0 = \{i^\star\}$). Then at each phase $j \in \mathbb{N}$, each honest agent i contacts a random honest neighbor i'. If i' is informed (i.e., if $i' \in \bar{I}_{j-1}$), then i becomes informed as well (i.e., $i \in \bar{I}_j$), subject to some Bernoulli(Υ) noise, where $\Upsilon \leq d_{\text{hon}}(i)/d(i)$. Hence, i becomes informed with probability $|\bar{I}_{j-1} \cap N_{\text{hon}}(i)|\Upsilon/d_{\text{hon}}(i) \leq |\bar{I}_{j-1} \cap N_{\text{hon}}(i)|/d(i)$. Note the right side of this inequality is in turn upper bounded by the probability with which they receive the best arm in the process of the previous paragraph.

More formally, we define the noisy rumor process as follows. The key quantity in Definition 1 is $\bar{\tau}_{\text{spr}}$, the first phase all are informed. Analogous to [18], our most general result will be in terms of the expected time that this phase occurs, i.e., $\mathbb{E}[A_{\bar{\tau}_{\text{spr}}}]$. Under Assumption 1, the latter quantity is $\tilde{O}((n\bar{d}_{\text{hon}}/\Upsilon)^\beta)$, which cannot be improved in general (see Appendix D.4). However, Section 5.3 provides sharper bounds for $\mathbb{E}[A_{\bar{\tau}_{\text{spr}}}]$ in some special cases.

DEFINITION 1. *Let $\Upsilon = \min_{i \in [n]} d_{hon}(i)/d(i)$. For each honest agent $i \in [n]$, let $\{\bar{Y}_j^{(i)}\}_{j=1}^{\infty}$ be i.i.d. Bernoulli(Υ) random variables and $\{\bar{H}_j^{(i)}\}_{j=1}^{\infty}$ i.i.d. random variables chosen uniformly at random from $N_{hon}(i)$. Inductively define $\{\bar{I}_j\}_{j=0}^{\infty}$ as follows: $\bar{I}_0 = \{i^\star\}$ (the agent from Assumption 3) and*

$$\bar{I}_j = \bar{I}_{j-1} \cup \{i \in [n] \setminus \bar{I}_{j-1} : \bar{Y}_j^{(i)} = 1, \bar{H}_j^{(i)} \in \bar{I}_{j-1}\} \; \forall \, j \in \mathbb{N}.$$

Finally, let $\bar{\tau}_{spr} = \inf\{j \in \mathbb{N} : \bar{I}_j = [n]\}$.

5.2 Positive result

We can now present the main result of this section: a regret upper bound for the proposed blocking rule. We state it first and then unpack the statement in some ensuing remarks. The proof of this result (and all others in this section) is deferred to Appendix D.

THEOREM 2. *Let Assumptions 1-3 hold. Suppose we run Algorithm 1 and use Algorithm 4 as the* Update-Blocklist *subroutine with $\theta_j = (j/3)^{\rho_1}$ and $\kappa_j = j^{\rho_2}/(K^2 S)$ in (8). Also assume*

$$\beta > 1, \quad \eta > 1, \quad 0 < \rho_1 \le \frac{1}{\eta}, \quad \alpha > \frac{3}{2} + \frac{1}{2\beta} + \frac{1}{2\rho_1^2}, \quad \frac{1}{2\alpha - 3} < \rho_2 < \rho_1(\beta - 1). \tag{9}$$

Then for any honest agent $i \in [n]$ and horizon $T \in \mathbb{N}$, we have

$$R_T^{(i)} \le 4\alpha \log(T) \min \left\{ \frac{2\eta - 1}{\eta - 1} \sum_{k=2}^{d_{mal}(i)+3} \frac{1}{\Delta_k} + \sum_{k=d_{mal}(i)+4}^{S+d_{mal}(i)+4} \frac{1}{\Delta_k}, \sum_{k=2}^{K} \frac{1}{\Delta_k} \right\} + 2\mathbb{E}[A_{2\bar{\tau}_{spr}}] + C_\star, \tag{10}$$

where $\Delta_k = 1$ by convention if $k > K$. Here $C_\star$ is a term independent of T satisfying

$$C_\star = \tilde{O} \left(\max \left\{ d_{mal}(i)/\Delta_2, (K/\Delta_2)^2, S^{\beta/(\rho_1^2(\beta-1))}, (S/\Delta_2^2)^{\beta/(\beta-1)}, \bar{d}^{\beta/\rho_1}, nK^2 S \right\} \right),$$

where $\tilde{O}(\cdot)$ hides dependencies on $\alpha, \beta, \eta, \rho_1,$ and ρ_2 and log dependencies on $K, n, m,$ and Δ_2^{-1}.

REMARK 7. *The theorem shows that our algorithm's regret scales as $(d_{mal}(i) + S) \log(T)/\Delta$, plus an additive term $2\mathbb{E}[A_{2\bar{\tau}_{spr}}] + C_\star$ that is independent of T and polynomial in all other parameters. When $S = O(K/n)$ (see Remark 1), the first term is $O((d_{mal}(i) + K/n) \log(T)/\Delta)$, as stated in Section 1.4. Also, when $d_{mal}(i)$ is large, we recover the $O(K \log(T)/\Delta)$ single-agent bound (including the constant 4α), i.e., if there are many malicious agents, honest ones fare no worse than the single-agent case.*

REMARK 8. *In addition to Assumptions 1-3, the theorem requires the algorithmic parameters to satisfy (9). For example, we can choose $\beta = \eta = 2, \rho_1 = 1/2, \alpha = 4,$ and $\rho_2 = 1/3$. More generally, we view these five parameters as small numerical constants and hide them in the $\tilde{O}(\cdot)$ notation.*

REMARK 9. *The bound in Theorem 2 can be simplified under additional assumptions. For instance, in Example 2, it is reasonable to assume $K = \Theta(n)$ (i.e., the number of restaurants is proportional to the population) and $\bar{d} = O(1)$ (i.e., the degrees are constant, as in sparse social networks). Under these assumptions, the choice $S = O(K/n) = O(1)$ from Remark 1, and the parameters from Remark 6, the theorem's regret bound can be further upper bounded by*

$$R_T^{(i)} \le \sum_{k=2}^{O(1)} \frac{48 \log T}{\Delta_k} + 2\mathbb{E}[A_{2\bar{\tau}_{spr}}] + \tilde{O}(\max\{(K/\Delta_2)^2, \Delta_2^{-4}, nK^2\}).$$

REMARK 10. *Note the parameters from Remark 8 were also used for the bad instance of Section 4. There, we had $\Delta_k > 1/15, S = d_{mal}(i) = 1,$ and $\mathbb{E}[A_{2\bar{\tau}_{spr}}] = \tilde{O}(n^\beta)$, so our regret is $O(\log T)$ plus a polynomial additive term that is much smaller than the doubly exponential term in Section 4.*

PROOF SKETCH. Let $\tau_{spr} = \inf\{j \in \mathbb{N} : B_{j'}^{(i')} = 1 \ \forall \ i' \in [n], j' \ge j\}$ denote the first phase where the best arm is most played for all honest agents at all phases thereafter. Before this phase (i.e., before time $A_{\tau_{spr}}$) we simply upper bound regret by $\mathbb{E}[A_{\tau_{spr}}]$. The main novelty of our analysis is bounding $\mathbb{E}[A_{\tau_{spr}}]$ in terms of $C_\star$ and $\mathbb{E}[A_{\bar{\tau}_{spr}}]$. We devote Section 7 to discussing this proof.

After phase τ_{spr}, the best arm is active by definition, so i incurs logarithmic in T regret. We let

$$\tau_{blk}^{(i)} = \inf\{j \in \mathbb{N} : H_{j'-1}^{(i)} \in P_{j'}^{(i)} \setminus P_{j'-1}^{(i)} \ \forall \ j' \ge j \ s.t. \ R_{j'-1}^{(i)} \ne 1\}$$

be the earliest phase such that i blocks for all suboptimal recommendations thereafter. We then split the phases after τ_{spr} into two groups: those before $\tilde{\tau}^{(i)} \triangleq \tau_{\text{spr}} \vee \tau_{\text{blk}}^{(i)} \vee T^{1/(\beta K)}$ and those after.

For the phases after τ_{spr} but before $\tilde{\tau}^{(i)}$, we consider three cases:

- $\tilde{\tau}^{(i)} = \tau_{\text{spr}}$: In this case, there are no such phases, so there is nothing to prove.
- $\tilde{\tau}^{(i)} = T^{1/(\beta K)}$: Here we have an effective horizon $A_{\tilde{\tau}^{(i)}} = (\tilde{\tau}^{(i)})^\beta = T^{1/K}$, so similar to [56], we exploit the fact that the best arm is active and modify existing UCB analysis to bound regret by $O(K \log(T^{1/K})/\Delta) = O(\log(T)/\Delta)$, which is dominated by (10) (in an order sense).
- $\tilde{\tau}^{(i)} = \tau_{\text{blk}}^{(i)}$: Here we are considering phases j where the best arm is most played by i (since $j \geq \tau_{\text{spr}}$) but i does not block suboptimal recommendations (since $j \leq \tau_{\text{blk}}^{(i)}$). Note that no such phases arise for the existing blocking rule, so here the proof diverges from [56], and most of Appendix D.1 is dedicated to this case. Roughly speaking, the analogous argument of the previous case yields the regret bound $O((K/\Delta)\mathbb{E}[\log \tilde{\tau}^{(i)}])$, and we prove this term is also $O(\log(T)/\Delta)$ by deriving a tail bound for $\tilde{\tau}^{(i)}$. The tail amounts to showing that, once the best arm is active, i can identify suboptimal arms as such, within the phase. This in turn follows from best arm identification results and the growing phase lengths.

After phase $\tilde{\tau}^{(i)}$, the best arm is most played for all honest agents (since $\tilde{\tau}^{(i)} \geq \tau_{\text{spr}}$), so they only recommend this arm. Thus, i only plays the best arm, its S sticky arms, and any malicious recommendations. Consequently, to bound regret by $O((S + d_{\text{mal}}(i)) \log(T)/\Delta)$ as in (10), we need to show each malicious neighbor i' only recommends $O(1)$ suboptimal arms. It is easy to see that i' can only recommend $O(\log K)$ such arms: if i' recommends a bad arm at phase $\tilde{\tau}^{(i)}$, they will be blocked until phase $T^{\eta/(\beta K)}$ (since $\tilde{\tau}^{(i)} \geq \tau_{\text{blk}}^{(i)} \vee T^{1/(\beta K)}$), then until phase $(T^{\eta/(\beta K)})^\eta = T^{\eta^2/(\beta K)}$, etc. Thus, the $(\log_\eta K)$-th bad recommendation occurs at phase $T^{\eta^{\log_\eta K}/(\beta K)} = T^{1/\beta}$, which is time T by definition $A_j = j^\beta$. Finally, an argument from [56] sharpens this $O(\log K)$ term to $O(1)$. □

5.3 Special cases

We next discuss some special cases of our regret bound. First, as in [18], we can prove an explicit bound assuming the honest agent subgraph G_{hon} is d-regular, i.e., $d_{\text{hon}}(i) = d \ \forall \ i \in [n]$.

COROLLARY 1. *Let the assumptions of Theorem 2 hold and further assume G_{hon} is d-regular with $d \geq 2$. Let ϕ denote the conductance of G_{hon}. Then for any honest agent $i \in [n]$ and horizon $T \in \mathbb{N}$,*

$$R_T^{(i)} \leq 4\alpha \log(T) \min \left\{ \frac{2\eta-1}{\eta-1} \sum_{k=2}^{d_{mal}(i)+3} \frac{1}{\Delta_k} + \sum_{k=d_{mal}(i)+4}^{S+d_{mal}(i)+4} \frac{1}{\Delta_k}, \sum_{k=2}^{K} \frac{1}{\Delta_k} \right\} \tag{11}$$

$$+ \tilde{O} \left(\max \left\{ d_{mal}(i)/\Delta_2, (K/\Delta_2)^2, S^{\beta/(\rho_1^2(\beta-1))}, (S/\Delta_2^2)^{\beta/(\beta-1)}, \bar{d}^{\beta/\rho_1}, nK^2S, (\phi\Upsilon)^{-\beta} \right\} \right).$$

REMARK 11. *This corollary includes the complete graph case studied in [56], where $d_{mal}(i) = m$, $\phi = \Theta(1)$, and $\Upsilon = \Theta(n/(n+m))$. In this case, the term (11) matches the corresponding term from [56] exactly, i.e., for large T, Corollary 1 is a strict generalization. Our additive term scales as*

$$\max \left\{ (m/\Delta_2), (K/\Delta_2)^2, S^{\beta/(\rho_1^2(\beta-1))}, (S/\Delta_2^2)^{\beta/(\beta-1)}, (n+m)^{\beta/\rho_1}, nK^2S, ((n+m)/n)^\beta \right\}$$

whereas the additive term from [56] scales as $\max\{(m/\Delta_2), (K/\Delta_2), (S/\Delta_2^2)^{2\beta\eta/(\beta-1)}, (n+m)^\beta, nK^2S\}$. Notice our dependence on the arm gap is $\Delta_2^{-2\beta/(\beta-1)}$, which matches the fully cooperative case [18], whereas the dependence is $\Delta_2^{-2\beta\eta/(\beta-1)}$ in [56], which is potentially much larger.

REMARK 12. *In the setting of Remark 9, the corollary's regret bound becomes*

$$R_T^{(i)} \leq \sum_{k=2}^{O(1)} \frac{48 \log T}{\Delta_k} + \tilde{O}(\max\{(K/\Delta_2)^2, \Delta_2^{-4}, nK^2, (\phi \Upsilon)^{-\beta}\}).$$

The key difference is the dependence on conductance $\tilde{O}(\phi^{-\beta})$, which matches the result from [19].

PROOF SKETCH. In light of Theorem 2, we only need to show $\mathbb{E}[A_{\bar{\tau}_{\mathrm{spr}}}] = \tilde{O}((\phi \Upsilon)^{-\beta})$. To do so, we let $\underline{\mathcal{I}}_j$ denote the noiseless version of $\bar{\mathcal{I}}_j$ (defined in the same way but with $\Upsilon = 1$) and $\underline{\tau}_{\mathrm{spr}} = \inf\{j : \underline{\mathcal{I}}_j = [n]\}$. We then construct a coupling between $\bar{\mathcal{I}}_j$ and $\underline{\mathcal{I}}_j$, which ensures that with high probability, $\bar{\tau}_{\mathrm{spr}} \leq j \log(j)/\Upsilon$ whenever $\underline{\tau}_{\mathrm{spr}} \leq j$. Finally, using this coupling and a tail bound for $\underline{\tau}_{\mathrm{spr}}$ from [18] (which draws upon the analysis of [20]), we derive a tail bound for $\bar{\tau}_{\mathrm{spr}}$. This allows us to show $\mathbb{E}[A_{\bar{\tau}_{\mathrm{spr}}}] = O(((\log n)^2 \log(\log(n)/\phi)/(\phi \Upsilon))^{\beta}) = \tilde{O}((\phi \Upsilon)^{-\beta})$, as desired.[4]　　□

Finally, we can sharpen the above results for honest agents without malicious neighbors.

COROLLARY 2. *For $i \in [n]$ with $d_{mal}(i) = 0$, the terms (10) and (11) from Theorem 2 and Corollary 1, respectively, can (under their respective assumptions) be improved to $4\alpha \log(T) \sum_{k=2}^{S+2} \Delta_k^{-1}$.*

REMARK 13. *The improved term in Corollary 2 matches the $\log T$ term from [18], including constants. Thus, the corollary shows that for large T, agents who are not directly connected to malicious agents are unaffected by their presence elsewhere in the graph.*

PROOF SKETCH. Recall from the Theorem 2 proof sketch that the $\log T$ term arises from regret after phase τ_{spr}. At any such phase, the best arm is most played for all honest agents (by definition), so when $d_{\mathrm{mal}}(i) = 0$, i's neighbors only recommend this arm. Therefore, i's active sets after τ_{spr} are *fixed*; they contain the best arm and $S + 1$ suboptimal ones. Thus, i only plays $S + 1$ suboptimal arms long-term, so in the worst case incurs the standard UCB regret $4\alpha \log(T) \sum_{k=2}^{S+2} \Delta_k^{-1}$.　　□

6 NUMERICAL RESULTS

Thus far, we have shown the proposed blocking rule adapts to general graphs more gracefully than the existing one, at least in theory. We now illustrate this finding empirically.

Experimental setup: We follow [56, Section 6] except we extend those experiments to $G(n + m, p)$ graphs, i.e., each edge is present with probability p. For each $p \in \{1, 1/2, 1/4\}$ and each of two malicious strategies (to be defined shortly), we conduct 100 trials of the following:

- Set $n = 25$ and $m = 10$ and generate G as a $G(n + m, p)$ random graph, resampling if necessary until the honest agent subgraph G_{hon} is connected (see Assumption 1).
- Set $K = 100$, $\mu_1 = 0.95$, and $\mu_2 = 0.85$, then sample the remaining arm means $\{\mu_k\}_{k=3}^K$ uniformly from $[0, 0.85]$ (so $\Delta_2 = 0.1$). For each $k \in [K]$, set $\nu_k = \mathrm{Bernoulli}(\mu_k)$.
- Set $S = K/n$ and sample the sticky sets $\{\hat{S}^{(i)}\}_{i=1}^n$ uniformly from the S-sized subsets of $[K]$, resampling if necessary until $1 \in \cup_{i=1}^n \hat{S}^{(i)}$ (see Assumption 3).
- Run Algorithm 1 with the existing (Algorithm 3) and proposed (Algorithm 4) blocking rules, along with two baselines: a no communication scheme, where agents ignore the network and run UCB in isolation, and the algorithm from [18], where they do not block.

Algorithmic parameters: We set $\alpha = 4$ and $\beta = \eta = 2$ as in Remarks 4 and 8. For the parameters in the proposed blocking rule, we choose $\kappa_j = j^{1.5}$, and $\theta_j = j - \log j$. While these are different

---

[4]When $\Upsilon = 1$, [18] shows $\mathbb{E}[A_{\underline{\tau}_{\mathrm{spr}}}] = \mathbb{E}[A_{\bar{\tau}_{\mathrm{spr}}}] = O((\log(n)/\phi)^{\beta})$, so our bound generalizes theirs up to log terms.
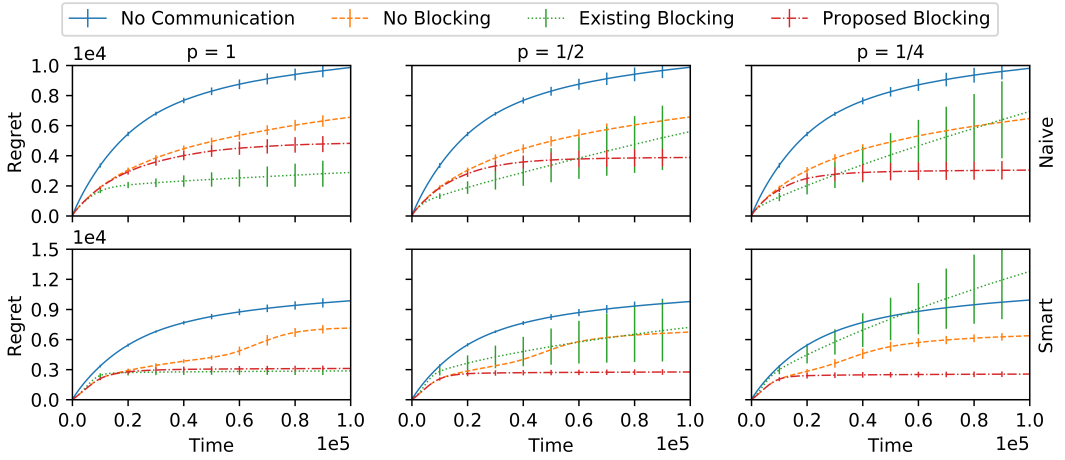
Fig. 3. Empirical results for synthetic data. Rows of subfigures correspond to the malicious strategy, while columns correspond to the edge probability p for the $G(n + m, p)$ random graph.

from the parameters specified in our theoretical results (which we found are too conservative in practice), they do satisfy the key properties discussed in Remark 6.

Malicious strategies: Like [56], we use strategies we call the *naive* and *smart* strategies (they are called *uniform* and *omniscient* in [56]). The naive strategy simply recommends a uniformly random suboptimal arm. The smart strategy recommends $R_j^{(i)} = \arg\min_{k \in \{2,...,K\} \setminus S_j^{(i)}} T_k^{(i)}(A_j)$, i.e., the least played, inactive, suboptimal arm. Intuitively, this is a more devious strategy which forces i to play $R_j^{(i)}$ often in the next phase (to drive down its upper confidence bound). Consequently, i may play it most and discard a better arm in favor of it (see Lines 11-13 of Algorithm 1).

Results: In Figure 3, we plot the average and standard deviation (across trials) of the per-agent regret $\sum_{i=1}^{n} R_T^{(i)}/n$. For the naive strategy, the existing blocking rule eventually becomes worse than the no blocking baseline as p decreases. More strikingly, it even becomes worse than the no communication baseline for the smart strategy. In other words, honest agents would have been better off *ignoring the network* and simply running UCB on their own. As in Section 4, this is because accidental blocking causes the best arm to spread very slowly. Additionally, the standard deviation becomes much higher than all other algorithms, suggesting that regret is significantly worse in some trials. In contrast, the proposed blocking rule *improves* as p decreases, because it is mild enough to spread the best arm at all values of p, and for smaller p, honest agents have fewer malicious neighbors (on average). We also observe that the proposed rule outperforms both baselines uniformly across p. Additionally, it improves over the existing rule more dramatically for the smart strategy, i.e., when the honest agents face a more sophisticated adversary. Finally, it is worth acknowledging the existing rule is better when $p = 1$ – although not in a statistically significant sense for the smart strategy – because it does spread the best arm quickly on the complete graph (as shown in [56]), and thereafter more aggressively blocks malicious agents.

Other results: As in [56], we reran the simulations using arm means derived from the MovieLens dataset [27]. We also experimented with new variants of the smart and naive strategies, where the malicious agents follow these strategies if the best arm is active (in hopes of forcing honest agents to discard it) and recommend the *second* best arm otherwise. Intuitively, these variants differ in that malicious agents recommend good arms (i.e., the second best) more frequently, while still never

revealing the best arm (the only one that leads to logarithmic regret). For all experiments, the key message – that the proposed blocking rule adapts to varying graph structures more gracefully than the existing one – is consistent. See Appendix B for details.

7 GOSSIP DESPITE BLOCKING

As discussed above, the main analytical contribution of this work is proving that the best arm spreads in a gossip fashion, despite accidental blocking. In this (technical) section, we provide a detailed sketch of this proof. We begin with a high-level outline. The key is to show that honest agents eventually stop blocking each other. This argument (roughly) proceeds as follows:

- **Step 1:** First, we show that honest agents learn the arm statistics in a certain sense. More specifically, we provide a tail bound for a random phase τ_{arm} such that for all phases $j \geq \tau_{\text{arm}}$ (1) each honest agent's most played arm in phase j is close to its true best active arm and (2) any active arm close to the true best one is played at least κ_j times by the end of phase j.
- **Step 2:** Next, we show that honest agents communicate with their neighbors frequently. In particular, we establish a tail bound for another random phase τ_{com} such that for any $j \geq \tau_{\text{com}}$, each honest agent contacts all of its honest neighbors at least once between θ_j and j.
- **Step 3:** Finally, we use the above to show that eventually, no blocking occurs amongst honest agents. The basic idea is as follows. Consider a phase j, an honest agent i, and a neighbor i' of i. Then if i has had the same best arm estimate k since phase θ_j – i.e., if the second blocking criterion in (8) holds – i' would have contacted i at some phase $j' \in \{\theta_j, \ldots, j\}$ (by step 2) and received arm k. Between phases j' and j, the most played arm for i' cannot get significantly worse (by step 1). Thus, if i asks i' for a recommendation at j, i' will respond with an arm whose mean is close to μ_k, which i will play at least κ_j times (by step 1). Hence, the first criterion in (8) fails, i.e., the two cannot simultaneously hold.

In the next three sub-sections, we discuss these three steps. Then in Section 7.4, we describe how, once accidental blocking stops, the arm spreading process can be coupled to the noisy rumor process from Definition 1. Finally, in Section 7.5, we discuss how to combine all of these steps to bound the term $\mathbb{E}[A_{\tau_{\text{spr}}}]$ from the Theorem 2 proof sketch.

7.1 Learning the arm statistics

Recall we assume $\mu_1 \geq \cdots \geq \mu_K$, so for any $W \subset [K]$, $\min W$ is the best arm in W, i.e., $\mu_{\min W} = \max_{w \in W} \mu_w$. Therefore, for any $\delta \in (0, 1)$, $G_\delta(W) \triangleq \{w \in W : \mu_w \geq \mu_{\min W} - \delta\}$ is the subset of arms at least δ-close to the best one. For each honest agent $i \in [n]$ and phase $j \in \mathbb{N}$, define

$$\Xi_{j,1}^{(i)} = \left\{ B_j^{(i)} \notin G_{\delta_{j,1}}(S_j^{(i)}) \right\}, \quad \Xi_{j,2}^{(i)} = \left\{ \min_{w \in G_{\delta_{j,2}}(S_j^{(i)})} T_w^{(i)}(A_j) \leq \kappa_j \right\}, \quad \Xi_j^{(i)} = \Xi_{j,1}^{(i)} \cup \Xi_{j,2}^{(i)}.$$

where $\delta_{j,1}, \delta_{j,2} \in (0, 1)$ will be chosen shortly. Finally, define the random phase

$$\tau_{\text{arm}} = \inf\{j \in \mathbb{N} : \mathbb{1}(\Xi_{j'}^{(i)}) = 0 \ \forall \ i \in [n], j' \in \{j, j+1, \ldots\}\}.$$

In words, τ_{arm} is the earliest phase such that, at all phases j thereafter, (1) the most played arms are $\delta_{j,1}$-close to best active arms and (2) all arms $\delta_{j,2}$-close to the best are played at least κ_j times.

As discussed above, Step 1 involves a tail bound for τ_{arm}. The analysis is based on [14, Theorem 2], which includes a tail bound showing that the most played arm is δ-close to the best, provided that $1/\delta^2$ samples have been collected from each of the δ-far arms. In our case, phase j lasts $A_j - A_{j-1} = \Theta(j^{\beta-1})$ time steps, so each of $S + 2$ active arms is played $\Theta(j^{\beta-1}/S)$ times on average. Hence, we can show the most played arm within the phase is $\delta_{j,1}$-close to the best if we choose

$\delta_{j,1} = \Theta(\sqrt{S/j^{\beta-1}})$, which allows us to bound $\mathbb{P}(\Xi_{j,1}^{(i)})$. Analogously, we choose $\delta_{j,2} = \Theta(1/\sqrt{\kappa_j})$ and show that $\delta_{j,2}$-close arms must be played $1/\delta_{j,2}^2 = \tilde{\Theta}(\kappa_j)$ times before they are distinguished as such, which allows us to bound $\mathbb{P}(\Xi_{j,2}^{(i)})$. Taken together, we can prove a tail bound for τ_{arm} (Lemma 9 in Appendix E.1) with these choices $\delta_{j,1} = \Theta(\sqrt{S/j^{\beta-1}})$ and $\delta_{j,2} = \tilde{\Theta}(1/\sqrt{\kappa_j})$.

7.2 Communicating frequently

Next, for any $i, i' \in [n]$ such that $(i, i') \in E_{\mathrm{hon}}$, let $\Xi_j^{(i \to i')} = \cap_{j'=\lfloor \theta_j \rfloor}^{j-2} \{H_{j'}^{(i')} \neq i\}$ denote the event that i did not send a recommendation to i' between phases $\lfloor \theta_j \rfloor$ and $j - 2$. Also define

$$\tau_{\mathrm{com}} = \inf\{j \in \mathbb{N} : \mathbb{1}(\cup_{i \to i' \in E_{\mathrm{hon}}} \Xi_{j'}^{(i \to i')}) = 0 \ \forall \ j' \in \{j, j+1, \ldots\}\}.$$

Here we abuse notation slightly; the union is over all (undirected) edges in E_{hon} but viewed as pairs of directed edges. Hence, at all phases $j \geq \tau_{\mathrm{com}}$, each honest agent i' receives a recommendation from each of its honest neighbors i at some phase j' between θ_j and $j - 2$.

Step 2 involves the tail bound for τ_{com} that was mentioned above (see Lemma 10 in Appendix E.2). The proof amounts to bounding the probability of $\Xi_j^{(i \to i')}$. Recall this event says i' did not contact i for a recommendation at any phase $j' \in \{\theta_j, \ldots, j - 2\}$. Clearly, this means i' did not *block* i at any such phase. Hence, in the worst case, i' blocked i just before θ_j, in which case i was un-blocked at $\theta_j^\eta = (j/3)^{\rho_1 \eta} \leq j/3$, where the inequality holds by assumption in Theorem 2. Hence, $\Xi_j^{(i \to i')}$ implies i' was not blocking i between phases $j/3$ and $j - 2$, so each of the $\Theta(j)$ neighbors that i' contacted in these phases was sampled uniformly from a set containing i, yet i was never sampled. The probability of this decays exponentially in j, which yields an exponential tail for τ_{com}.

7.3 Avoiding accidental blocking

Next, we show honest agents eventually stop blocking each other. Toward this end, we first note

$$\forall \ i \in [n], \quad \forall \ j \geq \tau_{\mathrm{arm}}, \quad \mu_{\min S_j^{(i)}} \leq \mu_{B_j^{(i)}} + \delta_{j,1} \leq \mu_{\min S_{j+1}^{(i)}} + \delta_{j,1} \tag{12}$$

where the first inequality uses the definition of τ_{arm}, and the second holds because $\min S_{j+1}^{(i)} \in \arg\max_{k \in S_{j+1}^{(i)}} \mu_k$ and $B_j^{(i)} \in S_{j+1}^{(i)}$ in Algorithm 1 (see Claim 13 in Appendix E.3 for details). In words, (12) says the best active arm can decay by at most $\delta_{j,1}$ at phase j. Applying iteratively and since there are K arms total, we then show (see Claim 14 in Appendix E.3)

$$\forall \ i \in [n], \quad \forall \ j' \geq j \geq \tau_{\mathrm{arm}}, \quad \mu_{\min S_{j'}^{(i)}} \geq \mu_{\min S_j^{(i)}} - (K-1) \sup_{j'' \in \{j, \ldots, j'\}} \delta_{j'',1}.$$

Combining the previous two inequalities, we conclude (see Corollary 3 in Appendix E.3)

$$\forall \ i \in [n], \quad \forall \ j' \geq j \geq \tau_{\mathrm{arm}}, \quad \mu_{B_{j'}^{(i)}} \geq \mu_{\min S_j^{(i)}} - K \sup_{j'' \in \{j, \ldots, j'\}} \delta_{j'',1}. \tag{13}$$

Now the key part of Step 3 is to use (13) to show (see Claim 15 in Appendix E.3)

$$\forall \ j \in \mathbb{N} \ s.t. \ \theta_j \geq \tau_{\mathrm{arm}}, \quad \forall \ i, i' \in [n] \ s.t. \ i \in \{H_{j'}^{(i')}\}_{j'=\theta_j}^{j-2}, \quad i' \notin P_j^{(i)} \setminus P_{j-1}^{(i)}. \tag{14}$$

In words, this result says that if j is sufficiently large, and if i has sent a recommendation to i' since phase θ_j, then i will not block i' at phase j. The proof is by contraction: if instead i blocks i' at j, then by Algorithm 4, i has not changed its best arm estimate k since phase θ_j, so it would have recommended k to i' at some phase $j' \geq \theta_j$. Therefore, $\mu_{\min S_{j'}^{(i')}} \geq \mu_k$. Additionally, since

$j' \geq \theta_j = \Omega(j^{\rho_1})$, we know that for any $j'' \geq j'$, the choice of $\delta_{j'',1}$ in Section 7.1 guarantees that

$$K\delta_{j'',1} \leq O(K\delta_{j',1}) = \tilde{O}\left(\sqrt{K^2 S/(j')^{\beta-1}}\right) \leq \tilde{O}\left(\sqrt{K^2 S/j^{\rho_1(\beta-1)}}\right).$$

Combining these observations and using (13) (with j' and j replaced by $j-1$ and j'), we then show

$$\mu_{B_{j-1}^{(i')}} \geq \mu_{\min S_{j'}^{(i')}} - K \sup_{j'' \geq j'} \delta_{j'',1} \geq \mu_k - \tilde{O}\left(\sqrt{K^2 S/j^{\rho_1(\beta-1)}}\right). \tag{15}$$

On the other hand, i blocking i' at phase j means i plays the recommended arm $B_{j-1}^{(i')}$ fewer than κ_j times by the end of phase j. Since $j \geq \theta_j \geq \tau_{\mathrm{arm}}$, this implies (by definition of τ_{arm}) that $\mu_k > \mu_{B_{j-1}^{(i')}} + \delta_{j,2}$, where $\delta_{j,2} = \tilde{\Theta}(1/\sqrt{\kappa_j})$ as in Section 7.1. Combined with (15) and the choice $\kappa_j = j^{\rho_2}/(K^2 S)$ from Theorem 2, we conclude $j^{\rho_1(\beta-1)} \leq \tilde{O}(j^{\rho_2})$. This contradicts the assumption $\rho_2 < \rho_1(\beta-1)$ in Theorem 2, which completes the proof of (14).

Finally, we use (14) to show honest agents eventually stop blocking each other entirely, i.e.,

$$\forall\, j \in \mathbb{N} \ s.t.\ \theta_j \geq \tau_{\mathrm{com}}, \theta_{\theta_j} \geq \tau_{\mathrm{arm}}, \quad P_j^{(i)} \cap [n] = \emptyset\ \forall\, i \in [n] \tag{16}$$

(see Lemma 11 in Appendix E.3). Intuitively, (16) holds because after new blocking stops (14), old blocking will eventually "wear off". The proof is again by contradiction: if i is blocking some honest i' at phase j, the blocking must have started at some $j' \geq j^{1/\eta}$ (else, it ends by $(j')^\eta < j$). Thus, by assumption $j^{1/\eta} \geq j^{\rho_1} \geq \theta_j \geq \tau_{\mathrm{com}}$, i blocked i' at phase $j' \geq \tau_{\mathrm{com}}$. But by definition of τ_{com}, i' would have contacted i at some phase $j'' \in \{\theta_{j'}, \ldots, j'\}$. Applying (14) (at phase j'; note that by the above inequalities, $\theta_{j'} \geq \theta_{j^{1/\eta}} \geq \theta_{\theta_j} \geq \tau_{\mathrm{arm}}$, as required by (14)), we obtain a contradiction.

7.4 Coupling with noisy rumor process

To begin, we define an equivalent way to sample $H_j^{(i)}$ in Algorithm 2.[5] This equivalent method will allow us to couple the arm spreading and noisy rumor processes through a set of primitive random variables. In particular, for each honest agent $i \in [n]$, let $\{v_j^{(i)}\}_{j=1}^\infty$ and $\{\bar{H}_j^{(i)}\}_{j=1}^\infty$ be i.i.d. sequences drawn uniformly from $[0, 1]$ and $N_{\mathrm{hon}}(i)$. Then choose $H_j^{(i)}$ according to two cases:

- If $P_j^{(i)} \cap [n] = \emptyset$, let $Y_j^{(i)} = \mathbb{1}(v_j^{(i)} \leq d_{\mathrm{hon}}(i)/|N(i) \setminus P_j^{(i)}|)$ and consider two sub-cases. First, if $Y_j^{(i)} = 1$, set $H_j^{(i)} = \bar{H}_j^{(i)}$. Second, if $Y_j^{(i)} = 0$, sample $H_j^{(i)}$ from $N_{\mathrm{mal}}(i) \setminus P_j^{(i)}$ uniformly.
- If $P_j^{(i)} \cap [n] \neq \emptyset$, sample $H_j^{(i)}$ from $N(i) \setminus P_j^{(i)}$ uniformly.

Next, we observe that since $\delta_{j,1} \to 0$ as $j \to \infty$ by the choice of $\delta_{j,1}$ in Section 7.1 and $\Delta_2 > 0$ by Assumption 2, we have $\delta_{j,1} < \Delta_2$ for large enough j. Paired with the definition of τ_{arm}, this allows us to show that for all large j and $i \in [n]$ with $1 \in S_j^{(i)}$ (i.e., with the best arm active), $B_j^{(i)} = 1$ (i.e., the best arm is played most). See Claim 17 in Appendix E.4 for the formal statement.

Finally, we observe that by (16), only the first case of the above sampling strategy occurs for large j. Moreover, in this case, $Y_j^{(i)}$ is Bernoulli with parameter

$$d_{\mathrm{hon}}(i)/|N(i) \setminus P_j^{(i)}| \geq d_{\mathrm{hon}}(i)/|N(i)| \triangleq d_{\mathrm{hon}}(i)/d(i) \geq \Upsilon,$$

where the second inequality holds by Definition 1. Hence, the probability that $Y_j^{(i)} = 1$, and thus the probability that i contacts the random honest neighbor $\bar{H}_j^{(i)}$ in the above sampling strategy, dominates the probability that i contacts $\bar{H}_j^{(i)}$ in the noisy rumor process of Definition 1. Additionally, by the previous paragraph, agents with the best arm active will recommend it (for large enough j).

---

[5]Claim 16 in Appendix E.4 verifies this equivalency (the proof is a straightforward application of the law of total probability).

Taken together, we can show that the probability of receiving the best arm in the arm spreading process dominates the probability of being informed of the rumor in the noisy rumor process. This allows us to prove a tail bound for τ_{spr} in terms of a tail bound for the random phase $\bar{\tau}_{\mathrm{spr}}$ from Definition 1, on the event that the tails of τ_{arm} and τ_{com} are sufficiently small (in the sense of (16); see Lemma 12 in Appendix E.4 for details).

7.5 Spreading the best arm

In summary, we prove tail bounds for τ_{arm} and τ_{com} (Sections 7.1 and 7.2) and show the tails of τ_{spr} are controlled by those of $\bar{\tau}_{\mathrm{spr}}$, provided the tails of τ_{arm} and τ_{com} are not too heavy (Sections 7.3 and 7.4). Combining and summing tails allows us to bound $\mathbb{E}[A_{\tau_{\mathrm{spr}}}]$ in terms of $C_\star$ (which accounts for the tails of τ_{arm} and τ_{com}) and $\mathbb{E}[A_{\bar{\tau}_{\mathrm{spr}}}]$ (which accounts for the tail of $\bar{\tau}_{\mathrm{spr}}$), as mentioned in the Theorem 2 proof sketch. See Theorem 3 and Corollary 4 in Appendix E.5 for details.

8 CONCLUSION

In this work, we showed that existing algorithms for multi-agent bandits with malicious agents fail to generalize beyond the complete graph. In light of this, we proposed a new blocking algorithm and showed it has low regret on any connected and undirected graph. This regret bound relied on the analysis of a novel process involving gossip and blocking. Our work leaves open several questions, such as whether our insights can be applied to multi-agent reinforcement learning.

ACKNOWLEDGMENTS

REFERENCES

[1] Animashree Anandkumar, Nithin Michael, Ao Kevin Tang, and Ananthram Swami. 2011. Distributed algorithms for learning and cognitive medium access with logarithmic regret. *IEEE Journal on Selected Areas in Communications* 29, 4 (2011), 731–745.

[2] Jean-Yves Audibert and Sébastien Bubeck. 2010. Best Arm Identification in Multi-Armed Bandits. In *COLT-23th Conference on Learning Theory-2010*. 13–p.

[3] Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. 2002. Finite-time analysis of the multiarmed bandit problem. *Machine learning* 47, 2-3 (2002), 235–256.

[4] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. 1995. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of IEEE 36th Annual Foundations of Computer Science*. IEEE, 322–331.

[5] Orly Avner and Shie Mannor. 2014. Concurrent bandits and cognitive radio networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 66–81.

[6] Baruch Awerbuch and Robert Kleinberg. 2008. Competitive collaborative learning. *J. Comput. System Sci.* 74, 8 (2008), 1271–1288.

[7] Yogev Bar-On and Yishay Mansour. 2019. Individual regret in cooperative nonstochastic multi-armed bandits. *Advances in Neural Information Processing Systems* 32 (2019), 3116–3126.

[8] Eugenio Bargiacchi, Timothy Verstraeten, Diederik Roijers, Ann Nowé, and Hado Hasselt. 2018. Learning to coordinate with coordination graphs in repeated single-stage multi-agent decision problems. In *International conference on machine learning*. PMLR, 482–490.

[9] Ilai Bistritz and Nicholas Bambos. 2020. Cooperative multi-player bandit optimization. *Advances in Neural Information Processing Systems* 33 (2020).

[10] Ilai Bistritz and Amir Leshem. 2018. Distributed multi-player bandits-a game of thrones approach. In *Advances in Neural Information Processing Systems*. 7222–7232.

[11] Ilija Bogunovic, Andreas Krause, and Jonathan Scarlett. 2020. Corruption-tolerant Gaussian process bandit optimization. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 1071–1081.

[12] Ilija Bogunovic, Arpan Losalka, Andreas Krause, and Jonathan Scarlett. 2021. Stochastic linear bandits robust to adversarial attacks. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 991–999.

[13] Etienne Boursier and Vianney Perchet. 2019. SIC-MMAB: Synchronisation Involves Communication in Multiplayer Multi-Armed Bandits. *Advances in Neural Information Processing Systems* 32 (2019), 12071–12080.

[14] Sébastien Bubeck, Rémi Munos, and Gilles Stoltz. 2011. Pure exploration in finitely-armed and continuous-armed bandits. *Theoretical Computer Science* 412, 19 (2011), 1832–1852.

[15] Swapna Buccapatnam, Jian Tan, and Li Zhang. 2015. Information sharing in distributed stochastic bandits. In *2015 IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2605–2613.

[16] Nicolo Cesa-Bianchi, Claudio Gentile, Yishay Mansour, and Alberto Minora. 2016. Delay and cooperation in non-stochastic bandits. In *Conference on Learning Theory*, Vol. 49. 605–622.

[17] Mithun Chakraborty, Kai Yee Phoebe Chua, Sanmay Das, and Brendan Juba. 2017. Coordinated Versus Decentralized Exploration In Multi-Agent Multi-Armed Bandits.. In *IJCAI*. 164–170.

[18] Ronshee Chawla, Abishek Sankararaman, Ayalvadi Ganesh, and Sanjay Shakkottai. 2020. The Gossiping Insert-Eliminate Algorithm for Multi-Agent Bandits. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*. 3471–3481.

[19] Ronshee Chawla, Abishek Sankararaman, and Sanjay Shakkottai. 2022. Multi-agent low-dimensional linear bandits. *IEEE Trans. Automat. Control* (2022).

[20] Flavio Chierichetti, Silvio Lattanzi, and Alessandro Panconesi. 2010. Almost tight bounds for rumour spreading with conductance. In *Proceedings of the forty-second ACM symposium on Theory of computing*. 399–408.

[21] Hiba Dakdouk, Raphaël Féraud, Romain Laroche, Nadège Varsier, and Patrick Maillé. 2021. Collaborative Exploration and Exploitation in massively Multi-Player Bandits. (2021).

[22] Abhimanyu Dubey et al. 2020. Cooperative multi-agent bandits with heavy tails. In *International Conference on Machine Learning*. PMLR, 2730–2739.

[23] Abhimanyu Dubey et al. 2020. Kernel methods for cooperative multi-agent contextual bandits. In *International Conference on Machine Learning*. PMLR, 2740–2750.

[24] Abhimanyu Dubey and AlexSandy' Pentland. 2020. Differentially-Private Federated Linear Bandits. *Advances in Neural Information Processing Systems* 33 (2020).

[25] Evrard Garcelon, Baptiste Roziere, Laurent Meunier, Jean Tarbouriech, Olivier Teytaud, Alessandro Lazaric, and Matteo Pirotta. 2020. Adversarial Attacks on Linear Contextual Bandits. *Advances in Neural Information Processing Systems* 33 (2020).

[26] Anupam Gupta, Tomer Koren, and Kunal Talwar. 2019. Better Algorithms for Stochastic Bandits with Adversarial Corruptions. In *Conference on Learning Theory*. 1562–1578.

[27] F Maxwell Harper and Joseph A Konstan. 2015. The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)* 5, 4 (2015), 1–19.

[28] Eshcar Hillel, Zohar S Karnin, Tomer Koren, Ronny Lempel, and Oren Somekh. 2013. Distributed exploration in multi-armed bandits. In *Advances in Neural Information Processing Systems*. 854–862.

[29] Kwang-Sung Jun, Lihong Li, Yuzhe Ma, and Jerry Zhu. 2018. Adversarial attacks on stochastic bandits. *Advances in Neural Information Processing Systems* 31 (2018), 3640–3649.

[30] Dileep Kalathil, Naumaan Nayyar, and Rahul Jain. 2014. Decentralized learning for multiplayer multiarmed bandits. *IEEE Transactions on Information Theory* 60, 4 (2014), 2331–2345.

[31] Varun Kanade, Zhenming Liu, and Bozidar Radunovic. 2012. Distributed non-stochastic experts. In *Advances in Neural Information Processing Systems*. 260–268.

[32] Hsu Kao, Chen-Yu Wei, and Vijay Subramanian. 2022. Decentralized cooperative reinforcement learning with hierarchical information structure. In *International Conference on Algorithmic Learning Theory*. PMLR, 573–605.

[33] Sayash Kapoor, Kumar Kshitij Patel, and Purushottam Kar. 2019. Corruption-tolerant bandit learning. *Machine Learning* 108, 4 (2019), 687–715.

[34] Ravi Kumar Kolla, Krishna Jagannathan, and Aditya Gopalan. 2018. Collaborative learning of stochastic bandits over a social network. *IEEE/ACM Transactions on Networking* 26, 4 (2018), 1782–1795.

[35] Nathan Korda, Balázs Szörényi, and Li Shuai. 2016. Distributed clustering of linear bandits in peer to peer networks. In *Journal of machine learning research workshop and conference proceedings*, Vol. 48. International Machine Learning Societ, 1301–1309.

[36] Anusha Lalitha and Andrea Goldsmith. 2021. Bayesian Algorithms for Decentralized Stochastic Bandits. *IEEE Journal on Selected Areas in Information Theory* 2, 2 (2021), 564–583.

[37] Peter Landgren, Vaibhav Srivastava, and Naomi Ehrich Leonard. 2016. Distributed cooperative decision-making in multiarmed bandits: Frequentist and Bayesian algorithms. In *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 167–172.

[38] Tor Lattimore and Csaba Szepesvári. 2020. *Bandit algorithms.* Cambridge University Press.

[39] Heath J LeBlanc, Haotian Zhang, Xenofon Koutsoukos, and Shreyas Sundaram. 2013. Resilient asymptotic consensus in robust networks. *IEEE Journal on Selected Areas in Communications* 31, 4 (2013), 766–781.

[40] Fang Liu and Ness Shroff. 2019. Data Poisoning Attacks on Stochastic Bandits. In *International Conference on Machine Learning*. 4042–4050.

[41] Junyan Liu, Shuai Li, and Dapeng Li. 2021. Cooperative Stochastic Multi-agent Multi-armed Bandits Robust to Adversarial Corruptions. *arXiv preprint arXiv:2106.04207* (2021).

[42] Keqin Liu and Qing Zhao. 2010. Distributed learning in multi-armed bandit with multiple players. *IEEE Transactions on Signal Processing* 58, 11 (2010), 5667–5681.

[43] Lydia T Liu, Horia Mania, and Michael Jordan. 2020. Competing bandits in matching markets. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 1618–1628.

[44] Thodoris Lykouris, Vahab Mirrokni, and Renato Paes Leme. 2018. Stochastic bandits robust to adversarial corruptions. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*. 114–122.

[45] Udari Madhushani and Naomi Leonard. 2021. When to call your neighbor? strategic communication in cooperative stochastic bandits. *arXiv preprint arXiv:2110.04396* (2021).

[46] Yishay Mansour, Aleksandrs Slivkins, and Steven Wu. 2018. Competing bandits: Learning under competition. In *9th Innovations in Theoretical Computer Science, ITCS 2018*. Schloss Dagstuhl-Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing, 48.

[47] David Martínez-Rubio, Varun Kanade, and Patrick Rebeschini. 2019. Decentralized cooperative stochastic multi-armed bandits. *Advances in Neural Information Processing Systems* (2019).

[48] Aritra Mitra, Hamed Hassani, and George Pappas. 2021. Exploiting Heterogeneity in Robust Federated Best-Arm Identification. *arXiv preprint arXiv:2109.05700* (2021).

[49] Conor Newton, AJ Ganesh, and Henry Reeve. 2021. Asymptotic Optimality for Decentralised Bandits. In *Reinforcement Learning in Networks and Queues, Sigmetrics 2021*.

[50] Boris Pittel. 1987. On spreading a rumor. *SIAM J. Appl. Math.* 47, 1 (1987), 213–223.

[51] Jonathan Rosenski, Ohad Shamir, and Liran Szlak. 2016. Multi-player bandits–a musical chairs approach. In *International Conference on Machine Learning*. 155–163.

[52] Abishek Sankararaman, Ayalvadi Ganesh, and Sanjay Shakkottai. 2019. Social learning in multi agent multi armed bandits. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 3, 3 (2019), 1–35.

[53] Shahin Shahrampour, Alexander Rakhlin, and Ali Jadbabaie. 2017. Multi-armed bandits in multi-agent networks. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2786–2790.

[54] Balázs Szörényi, Róbert Busa-Fekete, István Hegedűs, Róbert Ormándi, Márk Jelasity, and Balázs Kégl. 2013. Gossip-based distributed stochastic bandit algorithms. In *Journal of Machine Learning Research Workshop and Conference Proceedings*, Vol. 2. International Machine Learning Societ, 1056–1064.

[55] Cem Tekin and Mihaela Van Der Schaar. 2015. Distributed online learning via cooperative contextual bandits. *IEEE transactions on signal processing* 63, 14 (2015), 3700–3714.

[56] Daniel Vial, Sanjay Shakkottai, and R Srikant. 2021. Robust multi-agent multi-armed bandits. In *Proceedings of the Twenty-second International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*. 161–170.

[57] Po-An Wang, Alexandre Proutiere, Kaito Ariu, Yassir Jedra, and Alessio Russo. 2020. Optimal algorithms for multiplayer multi-armed bandits. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 4120–4129.

[58] Zhaowei Zhu, Jingxuan Zhu, Ji Liu, and Yang Liu. 2021. Federated bandit: A gossiping approach. In *Abstract Proceedings of the 2021 ACM SIGMETRICS/International Conference on Measurement and Modeling of Computer Systems*. 3–4.

A NOTES ON APPENDICES

The appendices are organized as follows. First, Appendix B contains the additional numerical results that were mentioned in Section 6. Next, we prove Theorem 1 in Appendix C and all results from Section 5 in Appendix D. We then provide a rigorous version of the proof sketch from Section 7 in Appendix E. Finally, Appendix F contains some auxiliary results – namely, Appendix F.1 records some simple inequalities, Appendix F.2 provides some bandit results that are essentially known but stated in forms convenient to us, and Appendices F.3-F.4 contain some tedious calculations.

For the analysis, we use C_i, C_i', etc. to denote positive constants depending only on the algorithmic parameters α, β, η, ρ_1, and ρ_2. Each is associated with a corresponding claim, e.g., C_1 with Claim 1. Within the proofs, we use C, C', etc. to denote constants whose values may change across proofs. Finally, $\mathbb{1}$ denotes the indicator function, $\mathbb{E}_j$ and $\mathbb{P}_j$ are expectation and probability conditioned on all randomness before the j-th communication period, and $A^{-1}(t) = \min\{j \in \mathbb{N} : t \leq A_j\}$ denotes the current phase at time $t \in \mathbb{N}$.
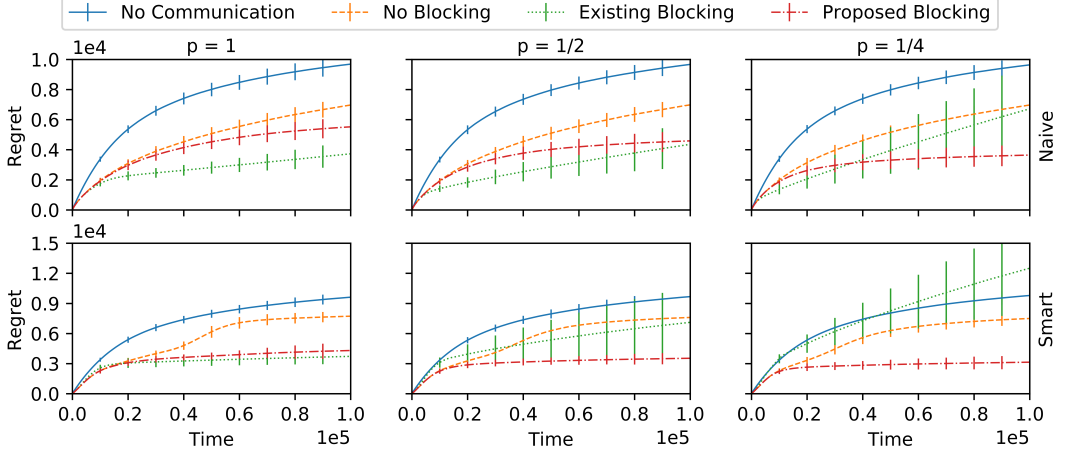
Fig. 4. Empirical results for real data. Rows of subfigures correspond to the malicious strategy, while columns correspond to the edge probability p for the $G(n + m, p)$ random graph.
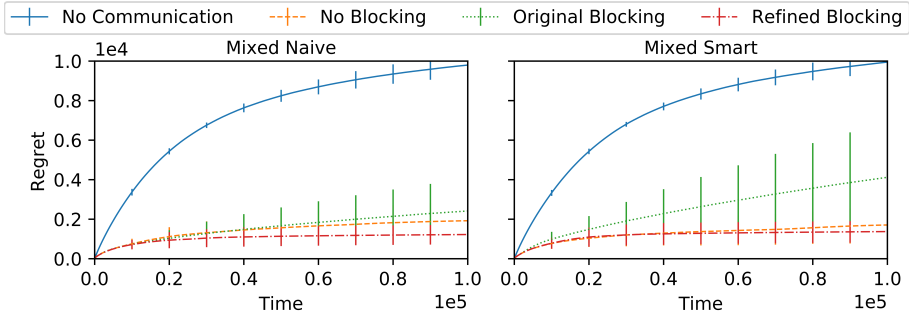

Fig. 5. Empirical results for synthetic data with $p = 1/2$ and mixed malicious strategies.

B ADDITIONAL EXPERIMENTS

As mentioned in Section 6, we also considered arm means derived from real data. The setup was the same as for the synthetic means, except for two changes (as in [56]): we choose $m = 15$ instead of $m = 10$, and we sample $\{\mu_k\}_{k=1}^K$ uniformly from a set of arm means derived from MovieLens [27] user film ratings via matrix completion; see [56, Section 6.2] for details. The results (Figure 4) are qualitatively similar to the synthetic case.

Finally, we repeated the synthetic data experiments from Section 6 with the intermediate $G(n + m, p)$ graph parameter $p = 1/2$ and two new malicious strategies called *mixed naive* and *mixed smart*. As discussed in Section 6, these approaches use a "mixed report" where the malicious agents more frequently recommend good arms – namely, the second best when the best is inactive and the naive or smart recommendation otherwise. Results are shown in Figure 5. They again reinforce the key message that the proposed rule adapts more gracefully to networks beyond the complete graph – in this case, our blocking rule has less than half of the regret of the existing one at the horizon T. Additionally, we observe that the no blocking algorithm from [18] has much lower regret in Figure 5 than Figure 3, though still higher than our proposed blocking algorithm. This suggests

that our algorithm remains superior even for "nicer" malicious strategies under which blocking is less necessary (in the sense that [18] has lower regret in Figure 5 than Figure 3).

C PROOF OF THEOREM 1

We first observe that since $h \leq 2^{h-1} \ \forall \ h \in \mathbb{N}$, we can lower bound the arm gap as follows:

$$\Delta_2 = \mu_1 - \mu_2 = 1 - \frac{13}{15} - \sum_{h=1}^{(n/2)-2} \left(\frac{1}{16}\right)^{2^{h-1}} > 1 - \frac{13}{15} - \sum_{h=1}^{\infty} \left(\frac{1}{16}\right)^h = \frac{1}{15}.$$

Next, we show that for phases $j \geq J_l$, agents aware of arms at least as good as $1 - l + n/2$ will (1) play such arms most often in phase j and (2) have such arms active thereafter. The proof is basically a noiseless version of a known bandit argument, specialized to the setting of Section 4.

LEMMA 1. *Under the assumptions of Theorem 1, for any $l \in [n/2]$, $j \geq J_l$, and $i \in [n]$ such that* $\min S_j^{(i)} \leq 1 - l + n/2$, *we have* $B_j^{(i)} \leq 1 - l + n/2$ *and* $\min S_{j'}^{(i)} \leq 1 - l + n/2 \ \forall \ j' \geq j$.

PROOF. First, we prove by contradiction that $B_j^{(i)} \leq 1-l+n/2$: suppose instead that $B_j^{(i)} \geq 2-l+n/2$. Let $k_1 = \min S_j^{(i)}$ and $k_2 = B_j^{(i)}$. Then $T_{k_2}^{(i)}(A_j) - T_{k_2}^{(i)}(A_{j-1}) \geq (A_j - A_{j-1})/3$; otherwise, since $|S_j^{(i)}| = S + 2 = 3$ by assumption and k_2 is the most played arm in phase j, we obtain

$$\sum_{k \in S_j^{(i)}} (T_k^{(i)}(A_j) - T_k^{(i)}(A_{j-1})) \leq 3(T_{k_2}^{(i)}(A_j) - T_{k_2}^{(i)}(A_{j-1})) < A_j - A_{j-1},$$

which is a contradiction. Furthermore, there clearly exists $t \in \{1 + A_{j-1}, \ldots, A_j\}$ such that

$$T_{k_2}^{(i)}(t-1) - T_{k_2}^{(i)}(A_{j-1}) = T_{k_2}^{(i)}(A_j) - T_{k_2}^{(i)}(A_{j-1}) - 1, \quad I_t^{(i)} = k_2.$$

Combining these observations, and since $T_{k_2}^{(i)}(A_{j-1}) \geq 0$, we obtain that

$$T_{k_2}^{(i)}(t-1) \geq T_{k_2}^{(i)}(A_j) - T_{k_2}^{(i)}(A_{j-1}) - 1 \geq \frac{A_j - A_{j-1}}{3} - 1, \quad I_t^{(i)} = k_2.$$

By the UCB policy and since $\alpha = 4$ by assumption, the previous expression implies

$$\mu_{k_1} < \mu_{k_1} + \sqrt{\frac{4 \log t}{T_{k_1}^{(i)}(t-1)}} \leq \mu_{k_2} + \sqrt{\frac{4 \log t}{T_{k_2}^{(i)}(t-1)}} \leq \mu_{k_2} + \sqrt{\frac{4 \log t}{\frac{A_j - A_{j-1}}{3} - 1}}. \tag{17}$$

Since $A_j = j^2$ by the assumption $\beta = 2$ and $t \leq A_j$, we also know

$$\frac{4 \log t}{\frac{A_j - A_{j-1}}{3} - 1} \leq \frac{8 \log j}{\frac{2j-1}{3} - 1} = \frac{12 \log j}{j - 2} = h(j), \tag{18}$$

where we define $h(j') = 12 \log(j')/(j' - 2) \ \forall \ j' > 2$. Note this function decreases on $[3, \infty)$, since

$$h'(j') = \frac{12(j' - 2 - j' \log j')}{j'(j' - 2)^2} \leq \frac{12(j' - 2 - j' \log 3)}{j'(j' - 2)^2} < \frac{-24}{j'(j' - 2)^2} < 0 \ \forall \ j' \geq 3,$$

where the second inequality is $e < 3$. Thus, since $j \geq J_l \geq J_1 \geq 3$, we know $h(j) \leq h(J_l)$. Combined with (17) and (18), we obtain $\mu_{k_1} < \mu_{k_2} + \sqrt{h(J_l)}$. Finally, recall $k_2 \geq 2 - l + n/2$ and $k_1 \leq 1 - l + n/2$, so $\mu_{k_1} - \mu_{k_2} \geq \mu_{1-l+n/2} - \mu_{2-l+n/2} > 0$. Combined with $\mu_{k_1} < \mu_{k_2} + \sqrt{h(J_l)}$, we conclude

$$(\mu_{1-l+n/2} - \mu_{2-l+n/2})^2 < h(J_l) = 12 \log(J_l)/(J_l - 2). \tag{19}$$

We now show that in each of three cases, (19) yields a contradiction.

- $l = 1$: By definition, the left side of (19) is $(\mu_{n/2} - \mu_{1+n/2})^2 = (13/15)^2$ and the right side is $12 \log(2^8)/(2^8 - 2)$, and one can verify that $12 \log(2^8)/(2^8 - 2) \leq (13/15)^2$.

- $1 < l < n/2$: Here $1 - l + n/2 > 1$ and $2 - l + n/2 < 1 + n/2$, i.e., both arms are mediocre. By definition, we thus have $(\mu_{1-l+n/2} - \mu_{2-l+n/2})^2 = 2^{-2^{l+1}}$, so to obtain a contradiction to (19), it suffices to show $h(J_l) \leq 2^{-2^{l+1}}$. We show by induction that this holds for all $l \in \{2, 3, \ldots\}$. For $l = 2$, note $J_2 = (J_1 + 2)^2 > J_1^2 + 2 = 2^{16} + 2$ (so $J_2 - 2 > 2^{16}$) and $J_2 = J_1^2 + 4J_1 + 4 = 2^{16} + 2^{10} + 2^2 < 2^{17}$ (so $\log J_2 < 17 \log 2 < 17$). Thus, $h(J_2) < 12 \cdot 17/2^{16} = 204/2^{16} < 2^{-8}$. Now assume $h(J_l) < 2^{-2^{l+1}}$ for some $l \geq 2$; we aim to show $h(J_{l+1}) < 2^{-2^{l+2}}$. Since $J_l \geq 2$, we have $J_{l+1} \leq (2J_l)^2 \leq J_l^4$; we also know $J_{l+1} - 2 = J_l^2 + 4J_l + 2 > J_l(J_l - 2)$. Thus, we obtain

$$h(J_{l+1}) < \frac{12 \log J_l^4}{J_l(J_l - 2)} = \frac{4}{J_l} \cdot h(J_l) < \frac{4}{J_l} \cdot 2^{-2^{l+1}} < \frac{4}{J_1^{2^{l-1}}} \cdot 2^{-2^{l+1}} = 2^{2 - 2^{l+2} - 2^{l+1}},$$

    where the inequalities follow from the previous paragraph, the inductive hypothesis, and (7) from Section 4.1, respectively. Since $2 < 2^{l+1}$, this completes the proof.

- $l = n/2$: Recall that in the previous case, we showed $h(J_l) < 2^{-2^{l+1}}$ for any $l \in \{2, 3, \ldots\}$. Therefore, $h(J_{n/2}) \leq 2^{-2^{(n/2)+1}} \leq 2^{-8}$ by assumption $n \in \{4, 6, \ldots\}$. Since $(\mu_{1-l+n/2} - \mu_{2-l+n/2})^2 = \Delta_2^2 = (1/15)^2 > (1/16)^2 = 2^{-8}$ in this case, we obtain a contradiction to (19).

Thus, we have established the first part of the lemma ($B_j^{(i)} \leq 1 - l + n/2$). To show $\min S_j^{(i)} \leq 1 - l + n/2$, we suppose instead that $\min S_{j'}^{(i)} > 1 - l + n/2$ for some $j' \geq j$. Then $j^\dagger = \min\{j' \geq j : \min S_{j'}^{(i)} > 1 - l + n/2\}$ is well-defined. If $j^\dagger = j$, then $\min S_j^{(i)} > 1 - l + n/2$, which violates the assumption of the lemma, so we assume $j^\dagger > j$. In this case, we know $\min S_{j^\dagger - 1}^{(i)} \leq 1 - l + n/2$ (since $j^\dagger$ is minimal) and $B_{j^\dagger - 1}^{(i)} > 1 - l + n/2$ (else, because $B_{j^\dagger - 1}^{(i)} \in S_{j^\dagger}^{(i)}$, we would have $\min S_{j^\dagger}^{(i)} \leq B_{j^\dagger - 1}^{(i)} \leq 1 - l + n/2$). But since $j^\dagger - 1 \geq J_l$ (by assumption $j^\dagger > j$ and $j \geq J_l$), this contradicts the first part of the lemma. □

Next, for each $l \in [n/2]$, we define the event

$$\mathcal{E}_l = \{l + n/2 \notin P_{J_l}^{(l+1+n/2)}\} \cap \cap_{j=1}^{J_l} \cap_{i=l+n/2}^{n} \{\min S_j^{(i)} > 1 - l + n/2, n + 1 \notin P_j^{(i)}\},$$

where $\{n \notin P_{J_{n/2}}^{(n+1)}\} = \Omega$ by convention (so $\mathcal{E}_{n/2}$ is well-defined). Thus, in words $\mathcal{E}_l$ says (1) $l + 1 + n/2$ is not blocking $l + n/2$ at phase J_l, (2) no honest agent $i \geq l + n/2$ has ever been aware of arms as good as $1 - l + n/2$ up to phase J_l, and (3) no such i has ever blocked the malicious agent $n + 1$. Point (2) will allow us to show that agent n does not become aware of the best arm until phase $J_{n/2}$ (when $\mathcal{E}_{n/2}$ holds). The other events in the definition of $\{\mathcal{E}_l\}_{l=1}^{n/2}$ will allow us to inductively lower bound their probabilities. The next lemma establishes the base of this inductive argument.

LEMMA 2. *Under the assumptions of Theorem 1, $\mathbb{P}(\mathcal{E}_1) \geq 3^{-2(J_1 - 1)}$.*

PROOF. We first observe that at all phases $j \in [J_1 - 1]$, only the second case of the malicious strategy – where the malicious agent recommends to avoid blocking – arises, which implies $n + 1 \notin P_j^{(i)} \ \forall j \in [J_1], i \in [n]$. Therefore, it suffices to show $\mathbb{P}(\mathcal{E}_1') \geq 3^{-2(J_1 - 1)}$, where we define

$$\mathcal{E}_1' = \{1 + n/2 \notin P_{J_1}^{(2+n/2)}\} \cap \cap_{j=1}^{J_1} \cap_{i=1+n/2}^{n} \{\min S_j^{(i)} > n/2\}.$$

To do so, we will show $\mathcal{E}_1' \supset \mathcal{F} \triangleq \cap_{j=1}^{J_1-1} \cap_{i=1+n/2}^{2+n/2} \{H_j^{(i)} = n + 1\}$ and $\mathbb{P}(\mathcal{F}) \geq 3^{-2(J_1 - 1)}$.

To show $\mathbb{P}(\mathcal{F}) \geq 3^{-2(J_1 - 1)}$, first note that by the law of total expectation, we have

$$\mathbb{P}(\mathcal{F}) = \mathbb{E}[\mathbb{1}(\cap_{j=1}^{J_1-2} \cap_{i=1+n/2}^{2+n/2} \{H_j^{(i)} = n + 1\})\mathbb{P}_{J_1-1}(\cap_{i=1+n/2}^{2+n/2}\{H_{J_1-1}^{(i)} = n + 1\})].$$

Now when $\cap_{j=1}^{J_1-2} \cap_{i=1+n/2}^{2+n/2} \{H_j^{(i)} = n+1\}$ occurs, the malicious strategy implies $P_{J_1-1}^{(i)} = \emptyset$, so $H_{J_1-1}^{(i)}$ is sampled from a set of three agents which includes $n+1$, for each $i \in \{1+n/2, 2+n/2\}$. Since this sampling is independent, we conclude that when $\cap_{j=1}^{J_1-2} \cap_{i=1+n/2}^{2+n/2} \{H_j^{(i)} = n+1\}$ occurs,

$$\mathbb{P}_{J_1-1}(\cap_{i=1+n/2}^{2+n/2} \{H_{J_1-1}^{(i)} = n+1\}) = 3^{-2}.$$

Thus, combining the previous two expressions with the definition of $\mathcal{F}$ and iterating, we obtain

$$\mathbb{P}(\cap_{j=1}^{J_1-1} \cap_{i=1+n/2}^{2+n/2} \{H_j^{(i)} = n+1\}) = \mathbb{P}(\mathcal{F}) = \mathbb{P}(\cap_{j=1}^{J_1-2} \cap_{i=1+n/2}^{2+n/2} \{H_j^{(i)} = n+1\})/3^2 = 3^{-2(J_1-1)}.$$

To show $\mathcal{E}_1' \supset \mathcal{F}$, first observe that when $\mathcal{F}$ occurs, $2+n/2$ does not contact $1+n/2$ at any phase $j \in [J_1 - 1]$, so $1+n/2 \notin P_{J_1}^{(2+n/2)}$. Thus, it only remains to show that $\mathcal{F}$ implies $\min S_j^{(i)} > n/2$ for all $j \leq J_1$ and $i > n/2$. Suppose instead that $\mathcal{F}$ holds and $\min S_j^{(i)} \leq n/2$ for some such j and i. Let $j^\dagger = \min\{j \leq J_1 : \min S_j^{(i)} \leq n/2 \text{ for some } i > n/2\}$ be the earliest j it occurs; note $j^\dagger > 1$ by assumption that $\min S_1^{(i)} > n/2$ for $i > n/2$. Let $i^\dagger$ be some agent it occurs for, i.e., $i^\dagger > n/2$ is such that $k^\dagger \in S_{j^\dagger}^{(i^\dagger)}$ for some $k^\dagger \leq n/2$. Since $j^\dagger$ is the earliest such phase and $j^\dagger > 1$, we know $k^\dagger$ was not active for $i^\dagger$ at the previous phase $j^\dagger - 1$, so it was recommended to $i^\dagger$ at this phase. By the malicious strategy and $j^\dagger - 1 \leq J_1 - 1$, this implies that the agent $i^\ddagger$ who recommended $k^\dagger$ to $i^\dagger$ is honest, so $k^\dagger$ was active for $i^\ddagger$ at the previous phase, which implies $i^\ddagger \leq n/2$ (else, we contradict the minimality of $j^\dagger$). From the assumed line graph structure, we must have $i^\dagger = 1+n/2$ and $i^\ddagger = n/2$, i.e., $1+n/2$ contacted $n/2$ at phase $j^\dagger - 1 \leq J_1 - 1$. But this contradicts the definition of $\mathcal{F}$, which stipulates that $1+n/2$ only contacts $n+1$ before J_1. $\qquad\square$

To continue the inductive argument, we lower bound $\mathbb{P}(\mathcal{E}_{l+1})$ in terms of $\mathbb{P}(\mathcal{E}_l)$.

LEMMA 3. *Under the assumptions of Theorem 1, for any $l \in [(n/2-1]$, we have $\mathbb{P}(\mathcal{E}_{l+1}) \geq 3^{-4}\mathbb{P}(\mathcal{E}_l)$.*

PROOF. The proof is somewhat lengthy and proceeds in four steps.
Step 1: Probabilistic arguments. First, we define the events

$$\mathcal{G}_1 = \cap_{i=l+n/2}^{l+2+n/2} \{H_{J_l}^{(i)} = n+1\}, \quad \mathcal{G}_2 = \{H_{J_l+1}^{(l+1+n/2)} = l+n/2\}, \quad \mathcal{G} = \mathcal{G}_1 \cap \mathcal{G}_2.$$

Then by the law of total expectation, we know that

$$\mathbb{P}(\mathcal{E}_l \cap \mathcal{G}) = \mathbb{E}[\mathbb{E}_{J_l+1}[\mathbb{1}(\mathcal{E}_l \cap \mathcal{G}_1)\mathbb{1}(\mathcal{G}_2)]] = \mathbb{E}[\mathbb{1}(\mathcal{E}_l \cap \mathcal{G}_1)\mathbb{P}_{J_l+1}(\mathcal{G}_2)]. \tag{20}$$

Now if $\mathcal{E}_l \cap \mathcal{G}_1$ occurs, then $l+n/2 \notin P_{J_l}^{(l+1+n/2)}$ (by $\mathcal{E}_l$) and $H_{J_l}^{(l+1+n/2)} = n+1$ (by $\mathcal{G}_1$); the latter implies $l+n/2 \notin P_{J_l+1}^{(l+1+n/2)} \setminus P_{J_l}^{(l+1+n/2)}$, so combined with the former, $l+n/2 \notin P_{J_l+1}^{(l+1+n/2)}$. Thus, $\mathcal{E}_l \cap \mathcal{G}_1$ implies that $H_{J_l+1}^{(l+1+n/2)}$ is sampled from a set of most three agents containing $l+n/2$, so $\mathbb{P}_{J_l+1}(\mathcal{G}_2) \geq 1/3$. Substituting into (20), and again using total expectation, we thus obtain

$$\mathbb{P}(\mathcal{E}_l \cap \mathcal{G}) \geq \mathbb{P}(\mathcal{E}_l \cap \mathcal{G}_1)/3 = \mathbb{E}[\mathbb{E}_{J_l}[\mathbb{1}(\mathcal{E}_l)\mathbb{1}(\mathcal{G}_1)]]/3 = \mathbb{E}[\mathbb{1}(\mathcal{E}_l)\mathbb{P}_{J_l}(\mathcal{G}_1)]/3.$$

Analogously, when $\mathcal{E}_l$ holds, $n+1 \notin P_{J_l}^{(i)} \forall i \in \{l+n/2, \ldots, l+2+n/2\}$, which by similar logic gives $\mathbb{P}_{J_l}(\mathcal{G}_1) \geq 3^{-3}$. Therefore, combining the previous two inequalities, we have shown $\mathbb{P}(\mathcal{E}_l \cap \mathcal{G}) \geq 3^{-4}\mathbb{P}(\mathcal{E}_l)$. Consequently, it suffices to show that $\mathcal{E}_l \cap \mathcal{G} \subset \mathcal{E}_{l+1}$.
Step 2: Event decomposition. For $l' \in \{l, l+1\}$, we decompose $\mathcal{E}_{l'} = \cap_{h=1}^4 \mathcal{H}_{l',h}$, where

$$\mathcal{H}_{l',1} = \{l'+n/2 \notin P_{J_{l'}}^{(l'+1+n/2)}\}, \quad \mathcal{H}_{l',2} = \cap_{j=1}^{J_{l'-1}} \cap_{i=l'+n/2}^{n} \{\min S_j^{(i)} > 1 - l'+n/2, n+1 \notin P_j^{(i)}\},$$

$$\mathcal{H}_{l',3} = \cap_{j=J_{l'-1}+1}^{J_{l'}} \cap_{i=l'+n/2}^{n} \{\min S_j^{(i)} > 1 - l'+n/2\}, \quad \mathcal{H}_{l',4} = \cap_{j=J_{l'-1}+1}^{J_{l'}} \cap_{i=l'+n/2}^{n} \{n+1 \notin P_j^{(i)}\}.$$

As a simple consequence of these definitions, we note that

$$\mathcal{H}_{l,2} \cap \mathcal{H}_{l,3} \cap \mathcal{H}_{l,4} = \cap_{j=1}^{J_l} \cap_{i=l+n/2}^{n} \{\min S_j^{(i)} > 1 - l + n/2, n + 1 \notin P_j^{(i)}\}$$
$$\subset \cap_{j=1}^{J_l} \cap_{i=l+1+n/2}^{n} \{\min S_j^{(i)} > 1 - (l+1) + n/2, n+1 \notin P_j^{(i)}\} = \mathcal{H}_{l+1,2}.$$

Hence, to prove $\mathcal{E}_l \cap \mathcal{G} \subset \mathcal{E}_{l+1}$, it suffices to show $\mathcal{E}_l \cap \mathcal{G} \subset \mathcal{H}_{l+1,1} \cap \mathcal{H}_{l+1,3} \cap \mathcal{H}_{l+1,4}$. For the remainder of the proof, we thus assume $\mathcal{E}_l \cap \mathcal{G}$ holds and argue $\mathcal{H}_{l+1,1} \cap \mathcal{H}_{l+1,3} \cap \mathcal{H}_{l+1,4}$ holds.

Step 3: Some consequences. We begin by deriving several consequences of $\mathcal{E}_l \cap \mathcal{G}$. First, note each $i \in \{l + 1 + n/2, l + 2 + n/2\}$ contacts $n + 1$ at phase J_l (by $\mathcal{G}_1$), who recommends $1 - l + n/2$ (by the malicious strategy). Since $\min S_{J_l}^{(i)} > 1 - l + n/2$ (by $\mathcal{H}_{l,3}$), this implies $1 - l + n/2 = \min S_{J_l+1}^{(i)}$, so $1 - l + n/2$ is most played in phase $J_l + 1$ (by Lemma 1). In summary, we have shown

$$H_{J_l}^{(i)} = n + 1, R_{J_l}^{(i)} = B_{J_l+1}^{(i)} = 1 - l + n/2 \; \forall \, i \in \{l + 1 + n/2, l + 2 + n/2\}. \tag{21}$$

Second, as a consequence of the above and Lemma 1, we can also write

$$1 - l + n/2 = \min S_{J_l+1}^{(i)} \geq \min S_{J_l+2}^{(i)} \geq \cdots \; \forall \, i \in \{l + 1 + n/2, l + 2 + n/2\}. \tag{22}$$

Third, we know $l + n/2$ contacts $n + 1$ at phase J_l (by $\mathcal{G}_1$), who responds with a currently active arm (by the malicious strategy), so since $\min S_{J_l}^{(l+n/2)} > 1 - l + n/2$ (by $\mathcal{H}_{l,3}$), we have

$$\min S_{J_l+1}^{(l+n/2)} > 1 - l + n/2. \tag{23}$$

As a consequence of (23), we see that when $l + 1 + n/2$ contacts $l + n/2$ at phase $J_l + 1$ (which occurs by $\mathcal{G}_2$), $l + n/2$ recommends some arm strictly worse than $1 - l + n/2$. On the other hand, by (22) and Lemma 1, we know the most played arm for $l + 1 + n/2$ in phase $J_l + 2$ has index at most $1 - l + n/2$. Taken together, the recommendation is not most played, so

$$l + n/2 \in P_j^{(l+1+n/2)} \; \forall \, j \in \{J_l + 2, \dots, (J_l + 2)^2 = J_{l+1}\}. \tag{24}$$

Step 4: Completing the proof. Using the above, we prove in turn that $\mathcal{H}_{l+1,4}$, $\mathcal{H}_{l+1,3}$, and $\mathcal{H}_{l+1,1}$ hold. For $\mathcal{H}_{l+1,4}$, we use proof by contradiction: if $\mathcal{H}_{l+1,4}$ fails, we can find $i \geq l + 1 + n/2$ and $j \in \{J_l + 1, \dots, J_{l+1}\}$ such that $n + 1 \in P_j^{(i)}$. Let $j^\dagger = \min\{j \in \{J_l + 1, \dots, J_{l+1}\} : n + 1 \in P_j^{(i)}\}$ be the minimal such j (for this i). Since $n + 1 \notin P_{J_l}^{(i)}$ (by $\mathcal{H}_{l,4}$) and $j^\dagger$ is minimal, we must have $n + 1 \in P_{j^\dagger}^{(i)} \setminus P_{j^\dagger-1}^{(i)}$, i.e., $n + 1$ was blocked for the recommendation it provided at $j^\dagger - 1$. If $i \geq l + 3 + n/2$, this contradicts the malicious strategy, since $j^\dagger - 1 \in \{J_l, \dots, J_{l+1} - 1\}$ and the strategy avoids blocking for such i and $j^\dagger$. A similar contradiction arises if $i \in \{l + 1 + n/2, l + 2 + n/2\}$ and $j^\dagger \geq J_l + 2$ (since $j^\dagger - 1 \in \{J_l + 1, \dots, J_{l+1} - 1\}$ in this case), so we must have $i \in \{l + 1 + n/2, l + 2 + n/2\}$ and $j^\dagger = J_l + 1$. But in this case, $n + 1 \in P_{j^\dagger}^{(i)} \setminus P_{j^\dagger-1}^{(i)} = P_{J_l+1}^{(i)} \setminus P_{J_l}^{(i)}$ contradicts (21).

Next, we show $\mathcal{H}_{l+1,3}$ holds. The logic is similar to the end of the Lemma 2 proof. If instead $\mathcal{H}_{l+1,3}$ fails, we can find $j \in \{J_l + 1, \dots, J_{l+1}\}$ and $i \in \{l + 1 + n/2, \dots, n\}$ such that $\min S_j^{(i)} \leq (n/2) - l$. Let $j^\dagger$ be the minimal such j and $i^\dagger \geq l + 1 + n/2$ an agent with $\min S_{j^\dagger}^{(i^\dagger)} = k^\dagger$ for some $k^\dagger \leq (n/2) - l$. Since $\min S_{J_l}^{(i^\dagger)} > 1 - l + n/2$ (by $\mathcal{H}_{l,3}$), $j^\dagger \geq J_l + 1$, and $j^\dagger$ is minimal, we know that $k^\dagger$ was recommended to $i^\dagger$ at phase $j^\dagger - 1 \in \{J_l, \dots, J_{l+1} - 1\}$. By the malicious strategy, this implies that the recommending agent (say, $i^\ddagger$) was honest. Therefore, $k^\dagger$ was active for $i^\ddagger$ at phase $j^\dagger - 1$, so since $j^\dagger$ is minimal, $i^\ddagger \leq l + n/2$. Hence, by the assumed graph structure, $i^\dagger = l + 1 + n/2$ contacted $i^\ddagger = l + n/2$ at phase $j^\dagger - 1$, who recommended $k^\dagger$. If $j^\dagger - 1 \in \{J_l, J_l + 2, \dots, J_{l+1} - 1\}$, this contact cannot occur, since $l + 1 + n/2$ instead contacts $n + 1$ at J_l (by $\mathcal{G}_1$) and does not contact $l + n/2$ at $J_l + 2, \dots, J_{l+1}$ (by (24)). Hence, we must have $j^\dagger - 1 = J_l + 1$, so $\min S_{J_l+1}^{(l+n/2)} \leq k^\dagger \leq (n/2) - l$, contradicting (23).

Finally, we prove $\mathcal{H}_{l+1,1}$. Suppose instead that $l+1+n/2 \in P_{J_{l+1}}^{(l+2+n/2)}$, i.e., $l+1+n/2$ is blocked at J_{l+1}. Then since $P_0^{(l+2+n/2)} = \emptyset$, we must have $l+1+n/2 \in P_j^{(l+2+n/2)} \setminus P_{j-1}^{(l+2+n/2)}$ for some $j \in [J_{l+1}]$. Let $j^\dagger$ be the maximal such j. Then $j^\dagger \geq \sqrt{J_{l+1}} = J_l + 2$; otherwise, if $j^\dagger < \sqrt{J_{l+1}}$, $l+1+n/2$ would have been un-blocked by phase J_{l+1}. Therefore, the blocking rule implies

$$B_{j^\dagger}^{(l+2+n/2)} \neq R_{j^\dagger-1}^{(l+2+n/2)} = B_{j^\dagger-1}^{(l+1+n/2)}. \tag{25}$$

By $j^\dagger \in \{J_l + 2, \ldots, J_{l+1}\}$, $\mathcal{H}_{l+1,3}$, and (22), we also know

$$-l + n/2 < \min S_{j^\dagger}^{(l+2+n/2)}, \min S_{j^\dagger-1}^{(l+1+n/2)} \leq 1 - l + n/2,$$

so $\min S_{j^\dagger}^{(l+2+n/2)} = \min S_{j^\dagger-1}^{(l+1+n/2)} = 1 - l + n/2$. Combined with (25), we must have $B_{j^\dagger+h-2}^{(l+h+n/2)} > \min S_{j^\dagger+h-2}^{(l+h+n/2)} = 1 - l + n/2$ for some $h \in \{1, 2\}$, which contradicts Lemma 1 (since $j^\dagger \geq J_l + 2$). $\quad\square$

Finally, we can prove the theorem. Define $\sigma = \min\{j \in \mathbb{N} : 1 \in S_j^{(n)}\}$. Then by definition, $I_t^{(n)} \neq 1$ for any $t \leq A_{\sigma-1}$. Hence, because $\Delta_2 = 1/15$ in the problem instance of the theorem, we obtain

$$\frac{A_{\sigma-1} \wedge T}{15} = \sum_{t=1}^{A_{\sigma-1} \wedge T} \frac{\mathbb{1}(I_t^{(n)} \neq 1)}{15} = \sum_{t=1}^{A_\sigma \wedge T} \sum_{k=2}^{K} \frac{\mathbb{1}(I_t^{(n)} = k)}{15} \leq \sum_{t=1}^{T} \sum_{k=2}^{K} \Delta_k \mathbb{1}(I_t^{(n)} = k).$$

Thus, by Claim 23 from Appendix F.1 and since $A_{\sigma-1} = (\sigma-1)^2$ by the choice $\beta = 2$, we can write

$$R_T^{(n)} = \mathbb{E}\left[\sum_{t=1}^{T} \sum_{k=2}^{K} \Delta_k \mathbb{1}(I_t^{(i)} = k) \right] \geq \frac{\mathbb{E}[A_{\sigma-1} \wedge T]}{15} = \frac{\mathbb{E}[(\sigma-1)^2 \wedge T]}{15}.$$

Let $l \in [n/2]$ be chosen later. Then $\sigma > J_l$ implies $\sigma - 1 \geq J_l$ (since $\sigma, J_l \in \mathbb{N}$). Thus, we can write

$$\mathbb{E}[(\sigma-1)^2 \wedge T] \geq \mathbb{E}[((\sigma-1)^2 \wedge T)\mathbb{1}(\sigma > J_l)] \geq (J_l^2 \wedge T)\mathbb{P}(\sigma > J_l).$$

By definition of σ and $\mathcal{E}_l$, along with Lemmas 2 and 3, we also know

$$\mathbb{P}(\sigma > J_l) \geq \mathbb{P}(\mathcal{E}_l) \geq 3^{-4(l-1)}\mathbb{P}(\mathcal{E}_1) \geq 3^{-4(l-1)} \cdot 3^{-2(J_1-1)} = 9^{3-2l-J_1}.$$

By (7) from Section 4.1, we know $J_l^2 \geq J_1^{2^l} = (2^8)^{2^l} = 2^{2^{l+3}}$. Combined with the previous three bounds, and letting C denote the constant $C = 9^{3-J_1}/15$, we thus obtain

$$R_T^{(n)} \geq (2^{2^{l+3}} \wedge T) \cdot 9^{3-2l-J_1}/15 = C \cdot 81^{-l} \cdot (2^{2^{l+3}} \wedge T) \; \forall \, l \in [n/2]. \tag{26}$$

We now consider three different cases, each with a different choice of l.

- If $T > 2^{2^{(n/2)+3}}$, choose $l = n/2$. Then (26) becomes $R_T^{(n)} \geq C \cdot 81^{-n/2} \cdot 2^{2^{(n/2)+3}}$. Observe that

$$81^{-n/2} \cdot 2^{2^{(n/2)+3}} \geq 16^{-n} \cdot 2^{2^{(n/2)+3}} = (2^4)^{2 \cdot 2^{n/2}-n} \geq (2^4)^{2^{n/2}} > \exp(2^{n/2})$$
$$= \exp(\exp(n \log(2)/2)) > \exp(\exp(n/3)),$$

  where the second inequality is $n \leq 2^{n/2}$ for $n \in \{2, 4, 8, \ldots\}$. On the other hand, Claim 1 below shows $R_T^{(n)} \geq \log(T)/C_1$ for some absolute constant $C_1 > 0$. Thus, we have shown

$$R_T^{(n)} = (R_T^{(n)}/2) + (R_T^{(n)}/2) \geq (C/2)\exp(\exp(n/3)) + \log(T)/(2C_1).$$

- If $T \in (2^8, 2^{2^{(n/2)+3}}]$, let $l = \lceil \log_2(\log_2(T)) - 3 \rceil$. Then $2^{2^{l+3}} \geq T$, so $2^{2^{l+3}} \wedge T = T$. Furthermore, we know $l \leq \log_2(\log_2(T)) - 2$, which implies

$$81^{-l} \geq 81^2 \cdot 81^{-\log_2(\log_2(T))} \geq 81^2 \cdot 2^{-7\log_2(\log_2(T))} = 81^2/\log_2^7(T) = 81^2\log^7(2)/\log^7(T).$$

Next, observe that $0 = \log_2(\log_2(2^8)) - 3 < \log_2(\log_2(T)) - 3 \leq n/2$ for this case of T, so $l \in [n/2]$. Thus, we can choose this l in (26) and combine with the above bounds to lower bound regret as $R_T^{(n)} \geq 81^2 \log^7(2) CT/\log^7(T)$.

• If $T \leq 2^8$, choose $l = 1$. Then $2^{2^{l+3}} = 2^{16} \geq T$, so (26) implies $R_T^{(n)} \geq CT/81$.

Hence, in all three cases, we have shown $R_T^{(n)} \geq C' \min\{\log(T) + \exp(\exp(n/3)), T/\log^7(T)\}$ for some absolute constant $C' > 0$. This establishes the theorem.

We return to state and prove the aforementioned Claim 1. We note the analysis is rather coarse; our only goal here is to establish a $\log T$ scaling (not optimize constants).

CLAIM 1. *Under the assumptions of Theorem 1, we have $R_T^{(n)} \geq \log(T)/C_1$, where $C_1 = 15 \log 99$.*

PROOF. If $T = 1$, the bound holds by nonnegativity. If $T \in \{2, \ldots, 99\}$, then since $\min S_1^{(n)} > n/2$ and $\Delta_2 \geq 1/15$ by assumption in Theorem 1, we know $\Delta_{I_1^{(n)}} \geq 1/15$, which implies $R_T^{(n)} \geq 1/15 \geq \log(T)/C_1$. Thus, only the case $T \geq 100$ remains. By Claim 23 from Appendix F.1 and $\Delta_2 \geq 1/15$,

$$R_T^{(n)} \geq \frac{\mathbb{E}[\sum_{t=1}^T \mathbb{1}(I_t^{(n)} \neq 1)]}{15} = \frac{\log(99)\mathbb{E}[\sum_{t=1}^T \mathbb{1}(I_t^{(n)} \neq 1)]}{C_1} > \frac{2\mathbb{E}[\sum_{t=1}^T \mathbb{1}(I_t^{(n)} \neq 1)]}{C_1}.$$

Thus, it suffices to show that $\sum_{t=1}^T \mathbb{1}(I_t^{(n)} \neq 1) \geq \log(T)/2$. Suppose instead that this inequality fails. Then since the left side is an integer, we have $\sum_{t=1}^T \mathbb{1}(I_t^{(n)} \neq 1) \leq \lfloor \log(T)/2 \rfloor$ by assumption. Therefore, we can find $t \in \{T - \lfloor \log(T)/2 \rfloor + 1, \ldots, T\}$ such that $I_t^{(n)} = 1$ (else, we violate the assumed inequality). By this choice of t and the assumed inequality, we can then write

$$T_1^{(n)}(t-1) = t - 1 - \sum_{s=1}^{t-1} \mathbb{1}(I_s^{(n)} \neq 1) \geq (T - \lfloor \log(T)/2 \rfloor) - (\lfloor \log(T)/2 \rfloor) \geq T - \log T.$$

We can lower bound the right side by $4 \log T$ (else, applying Claim 20 from Appendix F.1 with $x = T$, $y = 1$, and $z = 5$ yields $T < 100$, a contradiction), which is further bounded by $4 \log t$. Combined with the fact that rewards are deterministic, $\mu_1 = 1$, and $\alpha = 4$ in Theorem 1, we obtain

$$\hat{\mu}_1^{(n)}(t-1) + \sqrt{\alpha \log(t)/T_1^{(n)}(t-1)} = 1 + \sqrt{4 \log(t)/T_1^{(n)}(t-1)} \leq 2. \tag{27}$$

Next, let $k \in S_{A^{-1}(t)}^{(n)}$ be any other arm which is active for n at time t. Then clearly

$$T_k^{(n)}(t-1) \leq \sum_{t=1}^T \mathbb{1}(I_t^{(n)} = k) \leq \sum_{t=1}^T \mathbb{1}(I_t^{(n)} \neq 1) \leq \lfloor \log(T)/2 \rfloor \leq \log(T)/2$$

$$\Rightarrow \hat{\mu}_k^{(n)}(t-1) + \sqrt{\alpha \log(t)/T_k^{(n)}(t-1)} \geq \sqrt{8 \log(t)/\log(T)} = 2\sqrt{\log(t)/\log(\sqrt{T})}. \tag{28}$$

By (27), (28), the fact that $I_t^{(n)} = 1$, and the UCB policy, we conclude $t \leq \sqrt{T}$. Since $T \geq 4$, this further implies $t \leq T/2$. But we also know that $t \geq T - \lfloor \log(T)/2 \rfloor + 1 > T - \log(T)/2$. Combining these inequalities gives $T < \log T$, a contradiction. □

D PROOFS FROM SECTION 5
D.1 Proof of Theorem 2
Fix an honest agent $i \in [n]$. Let $\tau^{(i)} = \tau_{\text{spr}} \vee \tau_{\text{blk}}^{(i)}$, where we recall from the proof sketch that

$$\tau_{\text{spr}} = \inf\{j \in \mathbb{N} : B_{j'}^{(i')} = 1 \ \forall \ i' \in [n], j' \geq j\},$$

$$\tau_{\text{blk}}^{(i)} = \inf\{j \in \mathbb{N} : H_{j'-1}^{(i)} \in P_{j'}^{(i)} \setminus P_{j'-1}^{(i)} \ \forall \ j' \geq j \ s.t. \ R_{j'-1}^{(i)} \neq 1\}.$$

Let $\gamma_i \in (0, 1)$ be chosen later. Denote by $\overline{S}^{(i)} = \{2, \ldots, K\} \cap \hat{S}^{(i)}$ and $\underline{S}^{(i)} = \{2, \ldots, K\} \setminus \hat{S}^{(i)}$ the suboptimal sticky and non-sticky arms, respectively, for agent i. Then by Claim 23 from Appendix F.1, we can decompose regret as $R_T^{(i)} = \sum_{h=1}^{4} R_{T,h}^{(i)}$, where

$$R_{T,1}^{(i)} = \mathbb{E}\left[\sum_{t=1}^{A_{\tau_{spr}} \wedge T} \Delta_{I_t^{(i)}}\right], \quad R_{T,2}^{(i)} = \sum_{k \in \overline{S}^{(i)}} \Delta_k \mathbb{E}\left[\sum_{t=1+A_{\tau_{spr}}}^{T} \mathbb{1}(I_t^{(i)} = k)\right],$$

$$R_{T,3}^{(i)} = \sum_{k \in \underline{S}^{(i)}} \Delta_k \mathbb{E}\left[\sum_{t=1+A_{\tau_{spr}}}^{A_{\tau^{(i)} \vee \lceil T^{\gamma_i/\beta}\rceil} \wedge T} \mathbb{1}(I_t^{(i)} = k)\right], \quad R_{T,4}^{(i)} = \sum_{k \in \underline{S}^{(i)}} \Delta_k \mathbb{E}\left[\sum_{t=1+A_{\tau^{(i)} \vee \lceil T^{\gamma_i/\beta}\rceil}}^{T} \mathbb{1}(I_t^{(i)} = k)\right],$$

and where $\sum_{t=s_1}^{s_2} \mathbb{1}(I_t^{(i)} = k) = 0$ whenever $s_1 > s_2$ by convention. Thus, we have rewritten regret as the sum of four terms: $R_{T,1}^{(i)}$, which accounts for regret before the best arm spreads; $R_{T,2}^{(i)}$, the regret due to sticky arms after the best arm spreads; $R_{T,3}^{(i)}$, regret from non-sticky arms after the best arm spreads but before phase $\tau^{(i)} \vee \lceil T^{\gamma_i/\beta}\rceil$; and $R_{T,4}^{(i)}$, regret from non-sticky arms after this phase. The subsequent lemmas bound these terms in turn.

LEMMA 4. *Under the assumptions of Theorem 2, for any $i \in [n]$ and $T \in \mathbb{N}$, we have*

$$R_{T,1}^{(i)} \leq \mathbb{E}[A_{\tau_{spr}}] = O(S^{\beta/(\rho_1^2(\beta-1))} \vee (S\log(S/\Delta_2)/\Delta_2^2)^{\beta/(\beta-1)} \vee (\bar{d}\log(n+m))^{\beta/\rho_1} \vee nK^2 S) + \mathbb{E}[A_{2\bar{\tau}_{spr}}].$$

PROOF. Assumption 2 ensures $\Delta_k \leq 1$, so $R_{T,1}^{(i)} \leq \mathbb{E}[A_{\tau_{spr}}]$. The result follows from the bound on $\mathbb{E}[A_{\tau_{spr}}]$ discussed in Section 7.5 and formally stated as Corollary 4 in Appendix E.5. □

LEMMA 5. *Under the assumptions of Theorem 2, for any $i \in [n]$ and $T \in \mathbb{N}$, we have*

$$R_{T,2}^{(i)} \leq \sum_{k \in \overline{S}^{(i)}} \frac{4\alpha \log T}{\Delta_k} + \frac{4(\alpha-1)|\overline{S}^{(i)}|}{2\alpha - 3}.$$

PROOF. For any $k \in \overline{S}^{(i)}$, Claim 22 and Corollary 6 from Appendix F.2 imply

$$\mathbb{E}\left[\sum_{t=1+A_{\tau_{spr}}}^{T} \mathbb{1}\left(I_t^{(i)} = k\right)\right] = \mathbb{E}\left[\sum_{t=1+A_{\tau_{spr}}}^{T} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) < \frac{4\alpha \log t}{\Delta_k^2}\right)\right]$$

$$+ \mathbb{E}\left[\sum_{t=1+A_{\tau_{spr}}}^{T} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) \geq \frac{4\alpha \log t}{\Delta_k^2}\right)\right]$$

$$\leq \frac{4\alpha \log T}{\Delta_k^2} + \frac{4(\alpha-1)}{2\alpha - 3},$$

so multiplying by Δ_k, using $\Delta_k \leq 1$, and summing over $k \in \overline{S}^{(i)}$ completes the proof. □

LEMMA 6. *Under the assumptions of Theorem 2, for any $i \in [n]$, $\gamma_i \in (0, 1)$, and $T \in \mathbb{N}$, we have*

$$R_{T,3}^{(i)} \leq \sum_{k \in \underline{S}^{(i)}} \frac{4\alpha \log A_{\lceil T^{\gamma_i/\beta}\rceil}}{\Delta_k} + \frac{4(\alpha-1)|\underline{S}^{(i)}|}{2\alpha - 3} + \frac{4C_4\alpha K}{\Delta_2 \gamma_i} \log\left(\frac{C_4 K}{\Delta_2 \gamma_i}\right) + 1 + \mathbb{E}[A_{\tau_{spr}}].$$

PROOF. The proof is nontrivial; we defer it to the end of this sub-appendix. □

LEMMA 7. *Under the assumptions of Theorem 2, for any $i \in [n]$, $\gamma_i \in (0,1)$, and $T \in \mathbb{N}$, we have*

$$R_{T,4}^{(i)} \leq \frac{2\eta - 1}{\eta - 1} \max_{\tilde{S} \subset \underline{S}^{(i)}: |\tilde{S}| \leq d_{mal}(i)+2} \sum_{k \in \tilde{S}} \frac{4\alpha \log T}{\Delta_k} + \frac{8\alpha\beta \log_\eta(1/\gamma_i)(d_{mal}(i)+2)}{\Delta_2} + \frac{4(\alpha-1)|\underline{S}^{(i)}|}{2\alpha - 3}.$$

PROOF SKETCH. The proof follows the same logic as that of [56, Lemma 4], so we omit it. The only differences are (1) we replace m (the number of malicious agents connected to i for the complete graph) with $d_{mal}(i)$, and (2) we use Claim 19 from Appendix F.1 to bound the summation in [56, Lemma 4]. We refer the reader to the Theorem 2 proof sketch for intuition. □

Additionally, we note the sum $R_{T,3}^{(i)} + R_{T,4}^{(i)}$ can be naively bounded as follows.

LEMMA 8. *Under the assumptions of Theorem 2, for any $i \in [n]$ and $T \in \mathbb{N}$, we have*

$$R_{T,3}^{(i)} + R_{T,4}^{(i)} \leq \sum_{k \in \underline{S}^{(i)}} \frac{4\alpha \log T}{\Delta_k} + \frac{4(\alpha-1)|\underline{S}^{(i)}|}{2\alpha - 3}.$$

PROOF. The proof follows the exact same logic as that of Lemma 5 so is omitted. □

We can now prove the theorem. First, we use the regret decomposition $R_T^{(i)} = \sum_{h=1}^{4} R_{T,h}^{(i)}$, Lemmas 4-7, and the fact that $|\overline{S}^{(i)}| + |\underline{S}^{(i)}| \leq K$ to write

$$R_T^{(i)} \leq \frac{2\eta - 1}{\eta - 1} \max_{\tilde{S} \subset \underline{S}^{(i)}: |\tilde{S}| \leq d_{mal}(i)+2} \sum_{k \in \tilde{S}} \frac{4\alpha \log T}{\Delta_k} + \sum_{k \in \overline{S}^{(i)}} \frac{4\alpha \log T}{\Delta_k} + \sum_{k \in \underline{S}^{(i)}} \frac{4\alpha \log A_{\lceil T\gamma_i/\beta \rceil}}{\Delta_k} \quad (29)$$

$$+ \frac{8\alpha\beta \log_\eta(1/\gamma_i)(d_{mal}(i)+2)}{\Delta_2} + \frac{8(\alpha-1)K}{2\alpha - 3} + \frac{4C_4 \alpha K}{\Delta_2 \gamma_i} \log\left(\frac{C_4 K}{\Delta_2 \gamma_i}\right) + 1 + 2\mathbb{E}[A_{\tau_{\text{spr}}}]. \quad (30)$$

Now choose $\gamma_i = \Delta_2/(K\Delta_{S+d_{mal}(i)+4}) \in (0,1)$. Then

$$(30) = \tilde{O}\left((d_{mal}(i)/\Delta_2) \vee (K/\Delta_2)^2\right) + 2\mathbb{E}[A_{\tau_{\text{spr}}}]$$

$$= \tilde{O}\left((d_{mal}(i)/\Delta_2) \vee (K/\Delta_2)^2 \vee S^{\beta/(\rho_1^2(\beta-1))} \vee (S/\Delta_2^2)^{\beta/(\beta-1)} \vee \bar{d}^{\beta/\rho_1} \vee nK^2 S\right) + 4\mathbb{E}[A_{2\bar{\tau}_{\text{spr}}}],$$

where the second inequality is due to Lemma 4. Furthermore, by Claim 18 from Appendix F.1, we know that $\log(A_{\lceil T\gamma_i/\beta \rceil}) \leq \log(e^{2\beta}(T^{\gamma_i/\beta})^\beta) = 2\beta + \gamma_i \log T$. Combined with $\underline{S}^{(i)} \leq K$, $\Delta_k \geq \Delta_2 \, \forall \, k \in \underline{S}^{(i)}$, and the choice of γ_i, we can thus write

$$\sum_{k \in \underline{S}^{(i)}} \frac{4\alpha \log A_{\lceil T\gamma_i/\beta \rceil}}{\Delta_k} \leq \sum_{k \in \underline{S}^{(i)}} \frac{8\alpha\beta + \gamma_i \log T}{\Delta_k} \leq \frac{8\alpha\beta K + \gamma_i K \log T}{\Delta_2} = \frac{8\alpha\beta K}{\Delta_2} + \frac{\log T}{\Delta_{S+d_{mal}(i)+4}}.$$

Therefore, we can bound (29) as follows:

$$(29) \leq \frac{2\eta - 1}{\eta - 1} \max_{\tilde{S} \subset \underline{S}^{(i)}: |\tilde{S}| \leq d_{mal}(i)+2} \sum_{k \in \tilde{S}} \frac{4\alpha \log T}{\Delta_k} + \sum_{k \in \overline{S}^{(i)}} \frac{4\alpha \log T}{\Delta_k} + \frac{\log T}{\Delta_{S+d_{mal}(i)+4}} + \frac{8\alpha\beta K}{\Delta_2}$$

$$\leq \frac{2\eta - 1}{\eta - 1} \sum_{k=2}^{d_{mal}(i)+3} \frac{4\alpha \log T}{\Delta_k} + \sum_{k=d_{mal}(i)+4}^{S+d_{mal}(i)+4} \frac{4\alpha \log T}{\Delta_k} + \frac{8\alpha\beta K}{\Delta_2},$$

where the second inequality holds by $\Delta_2 \leq \cdots \leq \Delta_K$ and $|\overline{S}^{(i)}| \leq S$. Combining the above yields

$$R_T^{(i)} \leq 4\alpha \log(T) \left(\frac{2\eta - 1}{\eta - 1} \sum_{k=2}^{d_{\mathrm{mal}}(i)+3} \frac{1}{\Delta_k} + \sum_{k=d_{\mathrm{mal}}(i)+4}^{S+d_{\mathrm{mal}}(i)+4} \frac{1}{\Delta_k} \right) + 2\mathbb{E}[A_{2\tilde{\tau}_{\mathrm{spr}}}] \tag{31}$$
$$+ \tilde{O}\left((d_{\mathrm{mal}}(i)/\Delta_2) \vee (K/\Delta_2)^2 \vee S^{\beta/(\rho_1^2(\beta-1))} \vee (S/\Delta_2^2)^{\beta/(\beta-1)} \vee \bar{d}^{\beta/\rho_1} \vee nK^2S \right).$$

Alternatively, we can simply use Lemmas 4, 5, and 8 to write

$$R_T^{(i)} \leq 4\alpha \log(T) \sum_{k=2}^{K} \frac{1}{\Delta_k} + \frac{4(\alpha-1)K}{2\alpha-3} + \mathbb{E}[A_{\tau_{\mathrm{spr}}}]. \tag{32}$$

Therefore, combining the previous two expressions and again invoking Lemma 4 to bound the additive terms in (32) by those in (31), we obtain the desired bound.

Thus, it only remains to prove Lemma 6. We begin by using some standard bandit arguments recounted in Appendix F.2 to bound $R_{T,3}^{(i)}$ in terms of a particular tail of $\tau^{(i)}$.

CLAIM 2. *Under the assumptions of Theorem 2, for any $i \in [n]$, $\gamma_i \in (0,1)$, and $T \in \mathbb{N}$, we have*

$$R_{T,3}^{(i)} \leq \sum_{k \in \underline{S}^{(i)}} \frac{4\alpha \log A_{\lceil T^{\gamma_i/\beta} \rceil}}{\Delta_k} + \frac{4(\alpha-1)|\underline{S}^{(i)}|}{2\alpha-3} + \mathbb{E}[A_{\tau_{spr}}]$$
$$+ \frac{4\alpha K \log T}{\Delta_2} \mathbb{P}\left(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil, A_{\tau_{spr}} < \frac{4\alpha K \log T}{\Delta_2} \right). \tag{33}$$

PROOF. If $T = 1$, we can naively bound $R_{T,3}^{(i)} \leq 1$, which completes the proof. Thus, we assume $T > 1$ (which will allow us to divide by $\log T$ later). For any $k \in \underline{S}^{(i)}$, we first write

$$\mathbb{E}\left[\sum_{t=1+A_{\tau_{\mathrm{spr}}}}^{A_{\tau^{(i)} \vee \lceil T^{\gamma_i/\beta} \rceil} \wedge T} \mathbb{1}\left(I_t^{(i)} = k \right) \right] \leq \mathbb{E}\left[\sum_{t=1+A_{\tau_{\mathrm{spr}}}}^{T} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) \geq \frac{4\alpha \log t}{\Delta_k^2} \right) \right] \tag{34}$$

$$+ \mathbb{E}\left[\mathbb{1}(\tau^{(i)} \leq \lceil T^{\gamma_i/\beta} \rceil) \sum_{t=1+A_{\tau_{\mathrm{spr}}}}^{A_{\lceil T^{\gamma_i/\beta} \rceil} \wedge T} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) < \frac{4\alpha \log t}{\Delta_k^2} \right) \right] \tag{35}$$

$$+ \mathbb{E}\left[\mathbb{1}(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil) \sum_{t=1+A_{\tau_{\mathrm{spr}}}}^{A_{\tau^{(i)}} \wedge T} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) < \frac{4\alpha \log t}{\Delta_k^2} \right) \right]. \tag{36}$$

By Corollary 6 from Appendix F.2, (34) is bounded by $4(\alpha-1)/(2\alpha-3)$. By Claim 22 from Appendix F.2 and $\mathbb{1}(\cdot) \leq 1$, (35) is bounded by $4\alpha \log(A_{\lceil T^{\gamma_i/\beta} \rceil})/\Delta_k^2$. For (36), Claim 22 similarly gives

$$\sum_{t=1+A_{\tau_{\mathrm{spr}}}}^{A_{\tau^{(i)}} \wedge T} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) < \frac{4\alpha \log t}{\Delta_k^2} \right) \leq \frac{4\alpha \log T}{\Delta_k^2},$$

which clearly implies the following bound for (36):

$$\mathbb{E}\left[\mathbb{1}(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil) \sum_{t=1+A_{\tau_{\mathrm{spr}}}}^{A_{\tau^{(i)}} \wedge T} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) < \frac{4\alpha \log t}{\Delta_k^2} \right) \right] \leq \frac{4\alpha \log T}{\Delta_k^2} \mathbb{P}(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil).$$

For the remaining probability term, by Markov's inequality, we have

$$\mathbb{P}(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil) \leq \mathbb{P}\left(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil, A_{\tau_{\mathrm{spr}}} < \frac{4\alpha K \log T}{\Delta_2}\right) + \frac{\Delta_2 \mathbb{E}[A_{\tau_{\mathrm{spr}}}]}{4\alpha K \log T}.$$

Hence, combining the previous two inequalities, and since $\Delta_2 \leq \Delta_k$, we can bound (36) by

$$\frac{4\alpha \log T}{\Delta_2 \Delta_k} \mathbb{P}\left(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil, A_{\tau_{\mathrm{spr}}} < \frac{4\alpha K \log T}{\Delta_2}\right) + \frac{\mathbb{E}[A_{\tau_{\mathrm{spr}}}]}{K\Delta_k}.$$

Finally, combining these bounds, then multiplying by Δ_k, summing over $k \in \underline{S}^{(i)}$, and using $\Delta_k \leq 1$ and $|\underline{S}^{(i)}| \leq K$, we obtain the desired result.   □

To bound (33), we consider two cases defined in terms of the following inequalities:

$$4\alpha K \log(T)/\Delta_2 < \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor^{\beta}, \quad 4\alpha K \log(T)/\Delta_2 < \kappa_{\lceil T^{\gamma_i/\beta} \rceil}, \tag{37}$$

$$\frac{4\alpha \log A_j}{\Delta_2^2} < \lceil \kappa_j \rceil - 1 \; \forall j \geq \lceil T^{\gamma_i/\beta} \rceil, \quad \log(T) \sum_{j=\lceil T^{\gamma_i/\beta} \rceil}^{\infty} (\lceil \kappa_j \rceil - 1)^{3-2\alpha} < \frac{\Delta_2(2\alpha - 3)}{8\alpha K^2}. \tag{38}$$

Roughly speaking, when all of these inequalities hold, then T is large enough to ensure that the event $\{\tau^{(i)} = \Omega(\mathrm{poly}(T)), A_{\tau_{\mathrm{spr}}} = O(\log T)\}$ in (33) is unlikely. The next claim makes this precise.

CLAIM 3. *Under the assumptions of Theorem 2, for any $i \in [n]$, $\gamma_i \in (0, 1)$, and $T \in \mathbb{N}$ such that all of the inequalities in (37)-(38) hold, we have*

$$\frac{4\alpha K \log T}{\Delta_2} \mathbb{P}\left(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil, A_{\tau_{spr}} < \frac{4\alpha K \log T}{\Delta_2}\right) \leq 1.$$

PROOF. If $T = 1$, the left side is zero and the bound is immediate, so we assume $T > 1$. First note that if $A_{\tau_{\mathrm{spr}}} < 4\alpha K \log(T)/\Delta_2$, then since $A_{\tau_{\mathrm{spr}}} = \lceil \tau_{\mathrm{spr}}^{\beta} \rceil \geq \tau_{\mathrm{spr}}^{\beta}$ by definition and $4\alpha K \log(T)/\Delta_2 < \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor^{\beta} \wedge \kappa_{\lceil T^{\gamma_i/\beta} \rceil}$ by (37), $\tau_{\mathrm{spr}} < \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor \wedge \kappa_{\lceil T^{\gamma_i/\beta} \rceil}^{1/\beta}$. By definition $\theta_{\lceil T^{\gamma_i/\beta} \rceil} = (\lceil T^{\gamma_i/\beta} \rceil/3)^{\rho_1}$ with $\rho_1 \in (0, 1)$, this further implies $\tau_{\mathrm{spr}} < \lceil T^{\gamma_i/\beta} \rceil$. Thus, when $\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil$ and $A_{\tau_{\mathrm{spr}}} < 4\alpha K \log(T)/\Delta_2$, we have $\tau_{\mathrm{spr}} < \tau^{(i)}$, which by definition of $\tau^{(i)}$ implies $\tau^{(i)} = \tau_{\mathrm{blk}}^{(i)}$. Therefore, we can write

$$\mathbb{P}\left(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil, A_{\tau_{\mathrm{spr}}} < \frac{4\alpha K \log T}{\Delta_2}\right) \leq \mathbb{P}(\tau_{\mathrm{blk}}^{(i)} > \lceil T^{\gamma_i/\beta} \rceil, \tau_{\mathrm{spr}} < \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor \wedge \kappa_{\lceil T^{\gamma_i/\beta} \rceil}^{1/\beta}).$$

Now by definition, $\tau_{\mathrm{spr}} < \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor$ implies that $B_j^{(i)} = 1 \; \forall j \geq \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor$. Also by definition, $\tau_{\mathrm{blk}}^{(i)} > \lceil T^{\gamma_i/\beta} \rceil$ implies that for some $j \geq \lceil T^{\gamma_i/\beta} \rceil$ and $k > 1$, $R_{j-1}^{(i)} = k$ but $H_{j-1}^{(i)} \notin P_j^{(i)} \setminus P_{j-1}^{(i)}$. Thus,

$$\mathbb{P}(\tau_{\mathrm{blk}}^{(i)} > \lceil T^{\gamma_i/\beta} \rceil, \tau_{\mathrm{spr}} < \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor \wedge \kappa_{\lceil T^{\gamma_i/\beta} \rceil}^{1/\beta})$$

$$\leq \sum_{k=2}^{K} \sum_{j=\lceil T^{\gamma_i/\beta} \rceil}^{\infty} \mathbb{P}(R_{j-1}^{(i)} = k, H_{j-1}^{(i)} \notin P_j^{(i)} \setminus P_{j-1}^{(i)}, \tau_{\mathrm{spr}} < \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor \wedge \kappa_{\lceil T^{\gamma_i/\beta} \rceil}^{1/\beta}).$$

Now fix k and j as in the double summation. Again using $\tau_{\mathrm{spr}} < \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor \Rightarrow B_j^{(i)} = 1 \; \forall j \geq \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor$, the blocking rules implies that if $\tau_{\mathrm{spr}} < \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor$, $R_{j-1}^{(i)} = k$, and $H_{j-1}^{(i)} \notin P_j^{(i)} \setminus P_{j-1}^{(i)}$, then $T_k^{(i)}(A_j) > \kappa_j$. Since $T_k^{(i)}(A_j) \in \mathbb{N}$, this means $T_k^{(i)}(A_j) \geq \lceil \kappa_j \rceil$. Hence, there must exist some

$t \in \{\lceil \kappa_j \rceil, \ldots, A_j\}$ such that $T_k^{(i)}(t-1) \geq \lceil \kappa_j \rceil - 1$ and $I_t^{(i)} = k$. Thus, taking another union bound,

$$\mathbb{P}(R_{j-1}^{(i)} = k, H_{j-1}^{(i)} \notin P_j^{(i)} \setminus P_{j-1}^{(i)}, \tau_{\mathrm{spr}} < \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor \wedge \kappa_{\lceil T^{\gamma_i/\beta} \rceil}^{1/\beta})$$

$$\leq \sum_{t=\lceil \kappa_j \rceil}^{A_j} \mathbb{P}(T_k^{(i)}(t-1) \geq \lceil \kappa_j \rceil - 1, I_t^{(i)} = k, \tau_{\mathrm{spr}} < \kappa_{\lceil T^{\gamma_i/\beta} \rceil}^{1/\beta}).$$

Next, note $\tau_{\mathrm{spr}} < \kappa_{\lceil T^{\gamma_i/\beta} \rceil}^{1/\beta}$ implies that for any $j \geq \lceil T^{\gamma_i/\beta} \rceil$ and $t \geq \lceil \kappa_j \rceil$, we have $t \geq \lceil \kappa_{\lceil T^{\gamma_i/\beta} \rceil} \rceil \geq \lceil \tau_{\mathrm{spr}}^{\beta} \rceil = A_{\tau_{\mathrm{spr}}}$, so $1 \in S_{A^{-1}(t)}^{(i)}$ by definition of τ_{spr}. Therefore, for any $t \in \{\lceil \kappa_j \rceil, \ldots, A_j\}$,

$$\mathbb{P}(T_k^{(i)}(t-1) \geq \lceil \kappa_j \rceil - 1, I_t^{(i)} = k, \tau_{\mathrm{spr}} < \kappa_{\lceil T^{\gamma_i/\beta} \rceil}^{1/\beta}) \leq \mathbb{P}(T_k^{(i)}(t-1) \geq \lceil \kappa_j \rceil - 1, 1 \in S_{A^{-1}(t)}^{(i)}, I_t^{(i)} = k).$$

Now let $k_1 = k$, $k_2 = 1$, and $\ell = \lceil \kappa_j \rceil - 1$. Then $\mu_{k_2} - \mu_{k_1} = \Delta_k \geq \sqrt{4\alpha \log(t)/(\lceil \kappa_j \rceil - 1)}$ by definition and (38), respectively. Therefore, we can use Corollary 5 from Appendix F.2 to obtain

$$\mathbb{P}(T_k^{(i)}(t-1) \geq \lceil \kappa_j \rceil - 1, 1 \in S_{A^{-1}(t)}^{(i)}, I_t^{(i)} = k) \leq 2t^{2(1-\alpha)}.$$

Combining the above five inequalities, then using Claim 19 from Appendix F.1 and (38), we obtain

$$\mathbb{P}\left(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil, A_{\tau_{\mathrm{spr}}} < \frac{4\alpha K \log T}{\Delta_2}\right) \leq 2K \sum_{j=\lceil T^{\gamma_i/\beta} \rceil}^{\infty} \sum_{t=\lceil \kappa_j \rceil}^{\infty} t^{2(1-\alpha)}$$

$$\leq \frac{2K}{2\alpha - 3} \sum_{j=\lceil T^{\gamma_i/\beta} \rceil}^{\infty} (\lceil \kappa_j \rceil - 1)^{3-2\alpha} \leq \frac{\Delta_2}{4\alpha K \log T},$$

so multiplying both sides by $4\alpha K \log(T)/\Delta$ completes the proof. □

On the other hand, when (37)-(38) fails, we can show that T is bounded, and thus we bound the $\log T$ term in (33) by a constant and the probability term by 1, as shown in the following claim.

Claim 4. *Under the assumptions of Theorem 2, there exists a constant $C_4 > 0$ such that, for any $i \in [n]$, $\gamma_i \in (0, 1)$, and $T \in \mathbb{N}$ for which any of the inequalities in (37)-(38) fails, we have*

$$\frac{4\alpha K \log T}{\Delta_2} \mathbb{P}\left(\tau^{(i)} > \lceil T^{\gamma_i/\beta} \rceil, A_{\tau_{spr}} < \frac{4\alpha K \log T}{\Delta_2}\right) \leq \frac{4C_4 \alpha K}{\Delta_2 \gamma_i} \log \frac{C_4 K}{\Delta_2 \gamma_i}.$$

Proof. By Claims 33-36 in Appendix F.4, we can set $C_4 = \max_{i \in \{33, \ldots, 36\}} C_i$ to ensure that, if any of the inequalities in (37)-(38) fail, then $\log T \leq (C_4/\gamma_i) \log(C_4 K/(\Delta_2 \gamma_i))$. The claim follows after upper bounding the probability by 1. □

Finally, Lemma 6 follows by combining the previous three claims.

D.2 Proof of Corollary 1

As discussed in the proof sketch, we will couple with the noiseless process. We define this process as follows: let $\{\underline{H}_j^{(i)}\}_{j=1}^{\infty}$ be i.i.d. Uniform($N_{\mathrm{hon}}(i)$) random variables for each $i \in \mathbb{N}$, and

$$\underline{I}_0 = \{i^\star\}, \quad \underline{I}_j = \underline{I}_{j-1} \cup \{i \in [n] \setminus \underline{I}_{j-1} : \underline{H}_j^{(i)} \in \underline{I}_{j-1}\} \forall j \in \mathbb{N}.$$

For the coupling, we first define

$$\sigma_0 = 0, \quad \sigma_l = \inf \left\{ j > \sigma_{l-1} : \min_{i \in [n]} \sum_{j'=1+\sigma_{l-1}}^{j} \bar{Y}_{j'}^{(i)} \geq 1 \right\} \forall l \in \mathbb{N}.$$

Next, for each $i \in [n]$ and $l \in \mathbb{N}$, let $Z_l^{(i)} = \min\{j \in \{1 + \sigma_{l-1}, \ldots, \sigma_l\} : \bar{Y}_j^{(i)} = 1\}$. Note this set is nonempty, and since $Z_l^{(i)}$ is a deterministic function of $\{\bar{Y}_j\}_{j=1}^{\infty}$, which is independent of $\{\bar{H}_j^{(i)}\}_{j=1}^{\infty}$, $\bar{H}_{Z_l^{(i)}}^{(i)}$ is Uniform($N_{\text{hon}}(i)$) for each $l \in \mathbb{N}$. Hence, we can set

$$\underline{H}_j^{(i)} = \begin{cases} \bar{H}_{Z_l^{(i)}}^{(i)}, & \text{if } j = Z_l^{(i)} \text{ for some } l \in \mathbb{N} \\ \text{Uniform}(N_{\text{hon}}(i)), & \text{if } j \notin \{Z_l^{(i)}\}_{l=1}^{\infty} \end{cases}$$

without changing the distribution of $\{\underline{I}_j\}_{j=0}^{\infty}$. This results in a coupling where the noiseless process dominates the noisy one, in the following sense.

CLAIM 5. *For the coupling described above, $\underline{I}_j \subset \bar{I}_{\sigma_j}$ for any $j \geq 0$.*

PROOF. We use induction on j. For $j = 0$, we simply have $\underline{I}_j = \{i^\star\} = \bar{I}_j = \bar{I}_{\sigma_j}$. Now assume $\underline{I}_{j-1} \subset \bar{I}_{\sigma_{j-1}}$; we aim to show that if $i \in \underline{I}_j$, then $i \in \bar{I}_{\sigma_j}$. We consider two cases, the first of which is straightforward: if $i \in \underline{I}_{j-1}$, then $i \in \bar{I}_{\sigma_{j-1}}$ by the inductive hypothesis, so since $\sigma_{j-1} < \sigma_j$ by definition and $\{\bar{I}_{j'}\}_{j'=0}^{\infty}$ increases monotonically, we obtain $i \in \bar{I}_{\sigma_j}$, as desired.

For the second case, we assume $i \in [n] \setminus \underline{I}_{j-1}$ and $\underline{H}_j^{(i)} \in \underline{I}_{j-1}$. Set $j' = Z_j^{(i)}$ and recall $j' \in \{1 + \sigma_{j-1}, \ldots, \sigma_j\}$ by definition. From the coupling above, we know $\bar{Y}_{j'}^{(i)} = 1$ and $\bar{H}_{j'}^{(i)} = \underline{H}_j^{(i)}$. Since $\underline{H}_j^{(i)} \in \underline{I}_{j-1}$ in the present case, we have $\bar{H}_{j'}^{(i)} \in \underline{I}_{j-1}$ as well. Hence, because $\underline{I}_{j-1} \subset \bar{I}_{\sigma_{j-1}}$ by the inductive hypothesis, $j' - 1 \geq \sigma_{j-1}$ by definition, and $\{\bar{I}_{j''}\}_{j''=0}^{\infty}$ is increasing, we obtain $\bar{H}_{j'}^{(i)} \in \bar{I}_{j'-1}$. We have thus shown $\bar{Y}_{j'}^{(i)} = 1$ and $\bar{H}_{j'}^{(i)} \in \bar{I}_{j'-1}$, so $i \in \bar{I}_{j'}$ by Definition 1 Finally, using $j' \leq \sigma_j$ and again appealing to monotonocity, we conclude that $i \in \bar{I}_{\sigma_j}$.   □

We can now relate the rumor spreading times of the two processes. In particular, let $\bar{\tau}_{\text{spr}} = \inf\{j \in \mathbb{N} : \bar{I}_j = [n]\}$ (as in Definition 1) and $\underline{\tau}_{\text{spr}} = \inf\{j \in \mathbb{N} : \underline{I}_j = [n]\}$. We then have the following.

CLAIM 6. *For any $j \in \{3, 4, \ldots\}$ and $\iota > 1$, we have $\mathbb{P}(\bar{\tau}_{spr} > \iota j \log(j)/\Upsilon) \leq \mathbb{P}(\underline{\tau}_{spr} > j) + 27nj^{1-\iota}$.*

PROOF. Let $h(0) = 0$ and $h(j') = \iota j' \log(j)/\Upsilon$ for each $j' \in \mathbb{N}$. Then clearly

$$\mathbb{P}(\bar{\tau}_{\text{spr}} > \iota j \log(j)/\Upsilon) = \mathbb{P}(\bar{\tau}_{\text{spr}} > h(j), \sigma_j \leq h(j)) + \mathbb{P}(\bar{\tau}_{\text{spr}} > h(j), \sigma_j > h(j)). \tag{39}$$

For the first term in (39), by definition of $\bar{\tau}_{\text{spr}}$ and Claim 5, we have

$$\{\bar{\tau}_{\text{spr}} > h(j), \sigma_j \leq h(j)\} \subset \{\bar{\tau}_{\text{spr}} > \sigma_j\} = \{\bar{I}_{\sigma_j} \neq [n]\} \subset \{\underline{I}_j \neq [n]\} = \{\underline{\tau}_{\text{spr}} > j\}. \tag{40}$$

For the second term in (39), we first observe that for any $j' \in \mathbb{N}$,

$$\lfloor h(j') \rfloor - \lceil h(j'-1) \rceil - 1 \geq h(j') - h(j'-1) - 3 = (\iota \log(j)/\Upsilon) - 3 > 0,$$

where the last inequality holds by assumption on j and ι. Thus, by the union bound, we can write

$$\mathbb{P}(\sigma_{j'} > h(j'), \sigma_{j'-1} \leq h(j'-1)) \leq \sum_{i=1}^{n} \mathbb{P}(\cap_{j''=1+\lceil h(j'-1) \rceil}^{\lfloor h(j') \rfloor} \{\bar{Y}_{j''}^{(i)} = 0\}) = n(1-\Upsilon)^{\lfloor h(j') \rfloor - \lceil h(j'-1) \rceil - 1}$$

$$\leq n(1-\Upsilon)^{(\iota \log(j)/\Upsilon)-3} \leq n \exp(-\iota \log(j) + 3\Upsilon) < 27nj^{-\iota}.$$

Hence, because $\sigma_0 = 0 = h(0)$, we can iterate this argument to obtain that for any $j' \in \mathbb{N}$,

$$\mathbb{P}(\sigma_{j'} > h(j')) \leq \mathbb{P}(\sigma_{j'} > h(j'), \sigma_{j'-1} \leq h(j'-1)) + \mathbb{P}(\sigma_{j'-1} > h(j'-1))$$

$$\leq 27nj^{-\iota} + \mathbb{P}(\sigma_{j'-1} > h(j'-1)) \leq \cdots \leq 27nj'j^{-\iota}.$$

In particular, $\mathbb{P}(\sigma_j > h(j)) \leq 27nj^{1-\iota}$. Combining with (39) and (40) completes the proof.   □

To bound the tail of $\underline{\tau}_{\mathrm{spr}}$, we will use the following result.

CLAIM 7 (LEMMA 19 FROM [18]). *Under the assumptions of Corollary 1, there exists an absolute constant $C_7 > 0$ such that, for any $h \in \mathbb{N}$, we have $\mathbb{P}(\underline{\tau}_{spr} \geq C_7 h \log(n)/\phi) \leq n^{-4h}$.*

Using the previous two claims, we can prove a tail bound for $\bar{\tau}_{\mathrm{spr}}$.

CLAIM 8. *Under the assumptions of Corollary 1, there exists an absolute constant $C_8 > 0$ such that, for any $h \in \{3, 4, \ldots\}$, we have $\mathbb{P}(\bar{\tau}_{spr} \geq \xi(h)) \leq 56 \cdot 2^{-h}$, where we define*

$$\xi(h) = C_8 (\log n)^2 h^3 \log(C_8 \log(n)/\phi)/(\phi \Upsilon).$$

PROOF. Since G_{hon} is d-regular with $d \geq 2$ by assumption, we have $n \geq 2$ as well. Therefore, setting $C_8 = (2C_7) \vee ((e+1)/\log(2))$, we know $\log(C_8 \log(n)/\phi) \geq 1$, which implies

$$(h-1)\log(C_8 \log(n)/\phi) \geq h - 1 \geq \log h \quad \Rightarrow \quad h \log(C_8 \log(n)/\phi) \geq \log(C_8 h \log(n)/\phi).$$

Consequently, if we define $\iota = h \log n$ and $j = \lfloor C_8 h \log(n)/\phi \rfloor$, we can write

$$\xi(h) = (h \log n)(C_8 h \log(n)/\phi)(h \log(C_8 \log(n)/\phi))/\Upsilon$$
$$\geq (h \log n)(C_8 h \log(n)/\phi) \log(C_8 h \log(n)/\phi)/\Upsilon \geq \iota j \log(j)/\Upsilon.$$

Because $C_8 \geq (e+1)/\log(2)$, $h \geq 3$, and $n \geq 2$, we also know

$$j \geq C_8 \cdot h \cdot \log(n) - 1 \geq ((e+1)/\log(2)) \cdot 3 \cdot \log(2) - 1 = 3(e+1) - 1 > 3e > e^2. \tag{41}$$

Hence, $j \in \{3, 4, \ldots\}$; combined with $\iota \geq 3 \log 2 > 1$, we can apply Claim 6 to obtain

$$\mathbb{P}\left(\bar{\tau}_{\mathrm{spr}} > \xi(h)\right) \leq \mathbb{P}\left(\bar{\tau}_{\mathrm{spr}} > \iota j \log(j)/\Upsilon\right) \leq \mathbb{P}\left(\underline{\tau}_{\mathrm{spr}} > j\right) + 27 n j^{1-\iota}. \tag{42}$$

On the other hand, (41) implies $C_8 h \log(n)/\phi \geq 2$, so by definition of C_8,

$$j \geq (C_8 h \log(n)/\phi) - 1 \geq (C_8 h \log(n)/\phi)/2 = (C_8/2) h \log(n)/\phi \geq C_7 h \log(n)/\phi.$$

Therefore, by Claim 7, we know that

$$\mathbb{P}\left(\underline{\tau}_{\mathrm{spr}} > j\right) \leq \mathbb{P}\left(\underline{\tau}_{\mathrm{spr}} > C_7 h \log(n)/\phi\right) \leq n^{-4h} < n^{1-h}.$$

Finally, notice that $\iota = h \log n \geq 3 \log 2 > 2$, so $1 - \iota < -\iota/2$, thus by (41),

$$27 n j^{1-\iota} < 27 n j^{-\iota/2} = 27 n \sqrt{j}^{-\iota} < 27 n \exp(-\iota) = 27 n^{1-h}.$$

Hence, substituting the previous two inequalities into (42) and using $n \geq 2$ completes the proof. □

We now bound $\mathbb{E}[A_{2\bar{\tau}_{\mathrm{spr}}}]$. First, we define

$$\chi = \lceil C_8 (\log n)^2 \log(C_8 \log(n)/\phi)/(\phi \Upsilon) \rceil, \quad C = 8\beta/\log 2, \quad h_\star = \lceil C \log C \rceil \vee 3.$$

Notice that for any $h \geq h_\star \geq C \log C$, we have $2^{-h} \leq h^{-4\beta}$ (else, we can invoke Claim 20 from Appendix F.1 with $x = h$, $y = 1$, and $z = C/2$ to obtain $h < C \log C$, a contradiction). We write

$$\sum_{j=1}^{\infty} (A_j - A_{j-1}) \mathbb{P}(\bar{\tau}_{\mathrm{spr}} \geq j/2) \leq A_{2\chi h_\star^3} + \sum_{h=h_\star}^{\infty} \sum_{j=2\chi h^3+1}^{2\chi(h+1)^3} (A_j - A_{j-1}) \mathbb{P}(\bar{\tau}_{\mathrm{spr}} \geq j/2). \tag{43}$$

Now for any $h \geq h_\star$ and $j \geq 2\chi h^3 + 1$, we use $2^{-h} \leq h^{-4\beta}$, the definition of χ, and Claim 8 to write

$$\mathbb{P}(\bar{\tau}_{\mathrm{spr}} \geq j/2) \leq \mathbb{P}(\bar{\tau}_{\mathrm{spr}} \geq \chi h^3) \leq \mathbb{P}(\bar{\tau}_{\mathrm{spr}} \geq \xi(h)) \leq 56 \cdot 2^{-h} \leq 56 h^{-4\beta}.$$

Therefore, for any such h, we obtain

$$\sum_{j=2\chi h^3+1}^{2\chi(h+1)^3} (A_j - A_{j-1})\mathbb{P}(\bar{\tau}_{\text{spr}} \geq j/2) \leq 56 h^{-4\beta} \sum_{j=2\chi h^3+1}^{2\chi(h+1)^3} (A_j - A_{j-1}) \leq 56 h^{-4\beta} A_{2\chi(h+1)^3}.$$

Furthermore, by Claim 18 in Appendix F.1 and $h \geq 1$, we know that

$$A_{2\chi(h+1)^3} \leq e^{2\beta}(2\chi(h+1)^3)^\beta \leq e^{2\beta}(2\chi(2h)^3)^\beta = (4e)^{2\beta}\chi^\beta h^{3\beta} \; \forall \, h \geq 1.$$

Therefore, by the previous two inequalities and (43), and since $h_\star \geq 3$, we have shown

$$\sum_{j=1}^\infty (A_j - A_{j-1})\mathbb{P}(\bar{\tau}_{\text{spr}} \geq j/2) \leq (4e)^{2\beta}\chi^\beta \left(h_\star^{3\beta} + 56 \sum_{h=h_\star}^\infty h^{-\beta} \right) \leq (4e)^{2\beta}\chi^\beta \left(h_\star^{3\beta} + \frac{56}{\beta - 1} \right),$$

where the second inequality follows from Claim 19 from Appendix F.1, $h_\star \geq 3$, and $\beta > 1$. Because $h_\star$ is a constant, the right side is $O(\chi^\beta)$. Therefore, we have shown

$$\mathbb{E}[2_{\bar{\tau}_{\text{spr}}}] = \mathbb{E}\left[\sum_{j=1}^\infty (A_j - A_{j-1})\mathbb{1}(2\bar{\tau}_{\text{spr}} \geq j) \right] = \sum_{j=1}^\infty (A_j - A_{j-1})\mathbb{P}(\bar{\tau}_{\text{spr}} \geq j/2) = O(\chi^\beta).$$

Hence, by definition of χ, we obtain $\mathbb{E}[A_{2\bar{\tau}_{\text{spr}}}] = O(((\log n)^2 \log(\log(n)/\phi)/(\phi\Upsilon))^\beta)$. Combining this bound with Theorem 2 completes the proof of Corollary 1.

D.3 Proof of Corollary 2

Similar to the analysis in Appendix D.1, we can use the decomposition $R_T^{(i)} = \sum_{h=1}^4 R_{T,h}^{(i)}$, along with Lemmas 4 and 5, to bound regret as follows:

$$R_T^{(i)} \leq \sum_{k \in \overline{S}^{(i)}} \frac{4\alpha \log T}{\Delta_k} + \frac{4(\alpha - 1)|\overline{S}^{(i)}|}{2\alpha - 3} + R_{T,3}^{(i)} + R_{T,4}^{(i)} + \mathbb{E}[A_{\tau_{\text{spr}}}]. \tag{44}$$

Next, for each $k \in \underline{S}^{(i)}$, let $Y_k = \mathbb{1}(\cup_{j=\tau_{\text{spr}}}^\infty \{k \in S_j^{(i)}\})$ be the indicator that k was active after $A_{\tau_{\text{spr}}}$. Then as in the proof of Lemma 5, we can use Claim 22 and Corollary 6 from Appendix F.2 to write

$$\mathbb{E}\left[\sum_{t=1+A_{\tau_{\text{spr}}}}^T \mathbb{1}\left(I_t^{(i)} = k \right) \right] \leq \mathbb{E}[Y_k]\frac{4\alpha \log T}{\Delta_k^2} + \frac{4(\alpha - 1)}{2\alpha - 3}. \tag{45}$$

(The only difference from the proof of Lemma 5 is that, when applying Claim 22, we write

$$\sum_{t=1+A_{\tau_{\text{spr}}}}^T \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t - 1) < \frac{4\alpha \log t}{\Delta_k^2} \right) \leq Y_k \frac{4\alpha \log T}{\Delta_k^2},$$

where we can multiply by Y_k because the left side is also zero when $Y_k = 0$.) Combining (45) with the definitions of $R_{T,3}^{(i)}$ and $R_{T,4}^{(i)}$ and using $\Delta_k \leq 1$, we thus obtain

$$R_{T,3}^{(i)} + R_{T,4}^{(i)} \leq \mathbb{E}\left[\sum_{k \in \underline{S}^{(i)}} Y_k \frac{4\alpha \log T}{\Delta_k} \right] + \frac{4(\alpha - 1)|\underline{S}^{(i)}|}{2\alpha - 3}.$$

We claim, and will return to prove, that when $d_{\text{mal}}(i) = 0$,

$$\sum_{k \in \overline{S}^{(i)}} \frac{1}{\Delta_k} + \sum_{k \in \underline{S}^{(i)}} \frac{Y_k}{\Delta_k} \leq \sum_{k=2}^{S+2} \frac{1}{\Delta_k}. \tag{46}$$

Assuming (46) holds, we can combine the previous two inequalities and substitute into (44) to obtain $R_T^{(i)} \leq 4\alpha \log(T) \sum_{k=2}^{S+2} \Delta_k^{-2} + O(K) + \mathbb{E}[A_{\tau_{\mathrm{spr}}}]$. Bounding $\mathbb{E}[A_{\tau_{\mathrm{spr}}}]$ as in Lemma 4 yields the sharper version of Theorem 2, and further bounding $\mathbb{E}[A_{2\bar{\tau}_{\mathrm{spr}}}]$ as in Appendix D.2 sharpens Corollary 2.

To prove (46), we first show $\sum_{k \in \underline{S}^{(i)}} Y_k \leq 1 + \mathbb{1}(1 \in \hat{S}^{(i)})$. Suppose instead that $\sum_{k \in \underline{S}^{(i)}} Y_k \geq 2 + \mathbb{1}(1 \in \hat{S}^{(i)}) \triangleq H$. Then we can find H distinct arms $k_1, \ldots, k_H \in \underline{S}^{(i)}$, and H corresponding phases $j_h \geq \tau_{\mathrm{spr}}$, such that k_h was active at phase j_h for each $h \in [H]$. Without loss of generality, we can assume each j_h is minimal, i.e., $j_h = \min\{j \geq \tau_{\mathrm{spr}} : k_h \in S_j^{(i)}\}$. We consider two cases (which are exhaustive since $j_h \geq \tau_{\mathrm{spr}}$) and show that both yield contradictions.

- $j_h = \tau_{\mathrm{spr}} \; \forall \; h \in [H]$: We consider two further sub-cases.
  - $1 \in \hat{S}^{(i)}$, i.e., the best arm is sticky. Then $H = 3$, so $k_1, \ldots, k_3$ are all active at phase τ_{spr}. But all of these arms are non-sticky and only two such arms are active per phase.
  - $1 \notin \hat{S}^{(i)}$. Here k_1, k_2 are both active at phase τ_{spr}, as is 1 (by definition of τ_{spr}). But since k_1 and k_2 are suboptimal, we again have three non-sticky active arms.
- $\max_{h \in [H]} j_h > \tau_{\mathrm{spr}}$: We can assume (after possibly relabeling) that $j_1 > \tau_{\mathrm{spr}}$. Thus, by minimality of j_1, k_1 was not active at phase $j_1 - 1$ but became active at j_1, so it was recommended by some neighbor i' at $j_1 - 1$. But since $d_{\mathrm{mal}}(i) = 0$, i' is honest, and since $j_1 - 1 \geq \tau_{\mathrm{spr}}$, the best arm was most played for i' in phase $j_1 - 1$, so i' would not have recommended k_1.

Thus, $\sum_{k \in \underline{S}^{(i)}} Y_k \leq 1 + \mathbb{1}(1 \in \hat{S}^{(i)})$ holds. Combined with the fact that $|\overline{S}^{(i)}| = S - \mathbb{1}(1 \in \hat{S}^{(i)})$ by definition, at most $S + 1$ terms are nonzero in the summations on the left side of (46). Since $\Delta_2 \leq \cdots \leq \Delta_K$ by the assumed ordering of the arm means, this completes the proof.

D.4 Coarse analysis of the noisy rumor process

For completeness, we provide a coarser though more general bound for $\mathbb{E}[A_{2\bar{\tau}_{\mathrm{spr}}}]$ than the one derived in Appendix D.2. To begin, let $\mathbb{P}'_j$ and $\mathbb{E}'_j$ denote probability and expectation conditioned on $\{\bar{Y}_{j'}^{(i)}, \bar{H}_{j'}^{(i)}\}_{j'=1}^{j-1}$. For each $h \in [n]$, define the random phase $\bar{\tau}(h) = \inf\{j \in \mathbb{N} : |\bar{\mathcal{I}}_j| = h\}$. Note that $\bar{\tau}(1) = 1$ and $\bar{\tau}_{\mathrm{spr}} = \bar{\tau}(n)$. We then have the following tail bound.

CLAIM 9. *For any $l, j \in \mathbb{N}$, we have $\mathbb{P}(\bar{\tau}(l) > lj) \leq l(1 - \Upsilon/\bar{d}_{hon})^j$.*

PROOF. We use induction on l. For $l = 1$, $\bar{\tau}(1) = 1$ ensures $\mathbb{P}(\bar{\tau}(l) > lj) = 0$ for any $j \in \mathbb{N}$, so the bound is immediate. Next, assume the bound holds for $l \in \mathbb{N}$. We first write

$$\mathbb{P}(\bar{\tau}(l+1) > (l+1)j) \leq \mathbb{P}(\bar{\tau}(l+1) > (l+1)j, \bar{\tau}(l) \leq lj) + \mathbb{P}(\bar{\tau}(l) > lj).$$

Thus, by the inductive hypothesis, it suffices to bound the first term by $(1 - \Upsilon/\bar{d}_{\mathrm{hon}})^j$. We first write

$$\mathbb{P}(\bar{\tau}(l+1) > (l+1)j, \bar{\tau}(l) \leq lj) = \mathbb{E}[\mathbb{1}(\bar{\tau}(l) \leq lj)\mathbb{P}'_{lj+1}(\bar{\tau}(l+1) > (l+1)j)].$$

Now suppose $\bar{\tau}(l) \leq lj$. By Assumption 1, we can find $i \in \bar{\mathcal{I}}_{lj}$ and $i' \notin \bar{\mathcal{I}}_{lj}$ such that $i \in N_{\mathrm{hon}}(i')$. Then for $\bar{\tau}(l+1) > (l+1)j$ to occur, it must be the case that, for each $j' \in \{lj+1, \ldots, (l+1)j\}$, the event $\{\bar{H}_{j'}^{(i')} = i, \bar{Y}_{j'}^{(i')} = 1\}$ did *not* occur. Therefore, we have

$$\mathbb{P}'_{lj+1}(\bar{\tau}(l+1) > (l+1)j) \leq \mathbb{P}'_{lj+1}(\cap_{j'=lj+1}^{(l+1)j} \{\bar{H}_{j'}^{(i')} = i, \bar{Y}_{j'}^{(i')} = 1\}^C).$$

By the law of total expectation, we can write

$$\mathbb{P}'_{lj+1}(\cap_{j'=lj+1}^{(l+1)j} \{\bar{H}_{j'}^{(i')} = i, \bar{Y}_{j'}^{(i')} = 1\}^C)$$
$$= \mathbb{E}'_{lj+1}[\mathbb{1}(\cap_{j'=lj+1}^{(l+1)j-1} \{\bar{H}_{j'}^{(i')} = i, \bar{Y}_{j'}^{(i')} = 1\}^C)(1 - \mathbb{P}'_{(l+1)j}(\bar{H}_{(l+1)j}^{(i')} = i, \bar{Y}_{(l+1)j}^{(i')} = 1))].$$

Since $\bar{H}^{(i')}_{(l+1)j}$ is Uniform($N_{\text{hon}}(i)$) and $\bar{Y}^{(i')}_{(l+1)j}$ is Bernoulli(Υ), we have

$$\mathbb{P}'_{(l+1)j}(\bar{H}^{(i')}_{(l+1)j} = i, \bar{Y}^{(i')}_{(l+1)j} = 1) = \Upsilon/d_{\text{hon}}(i) \geq \Upsilon/\bar{d}_{\text{hon}}.$$

Therefore, combining the previous two expressions and iterating, we obtain

$$\mathbb{P}'_{lj+1}(\cap_{j'=lj+1}^{(l+1)j}\{\bar{H}^{(i')}_{j'} = i, \bar{Y}^{(i')}_{j'} = 1\}^C) \leq \mathbb{P}'_{lj+1}(\cap_{j'=lj+1}^{(l+1)j-1}\{\bar{H}^{(i')}_{j'} = i, \bar{Y}^{(i')}_{j'} = 1\}^C)(1 - \Upsilon/\bar{d}_{\text{hon}})$$
$$\leq \cdots \leq (1 - \Upsilon/\bar{d}_{\text{hon}})^j. \qquad \square$$

Next, we have a simple technical claim.

CLAIM 10. *Let $h_\dagger = (8\beta\bar{d}_{hon}/\Upsilon)\log(8\beta\bar{d}_{hon}n/\Upsilon)$. Then for any $h \geq h_\dagger$, $\exp(-h\Upsilon/\bar{d}_{hon}) \leq h^{-2\beta}/n$.*

PROOF. If the claimed bound fails, we have $h < (2\beta\bar{d}_{\text{hon}}/\Upsilon)\log(h) + (\bar{d}_{\text{hon}}/\Upsilon)\log(n)$. Then since $(\bar{d}_{\text{hon}}/\Upsilon)\log(n) < h_\dagger/2 \leq h/2$, we obtain $h < (4\beta\bar{d}_{\text{hon}}/\Upsilon)\log h$. Applying Claim 20 from Appendix F.1 with $x = h$, $y = 1$, and $z = 4\beta\bar{d}_{\text{hon}}/\Upsilon$, we obtain $h < 2z\log(2z) \leq h_\dagger$, a contradiction. $\qquad \square$

We can now bound $\mathbb{E}[A_{2\bar{\tau}_{\text{spr}}}]$. The analysis is similar to Appendix D.2. We first write

$$\mathbb{E}[A_{2\bar{\tau}_{\text{spr}}}] \leq A_{2n\lceil h_\dagger\rceil} + \sum_{h=\lceil h_\dagger\rceil}^{\infty} \mathbb{P}(\bar{\tau}_{\text{spr}} > nh) \sum_{j=2nh+1}^{2n(h+1)} (A_j - A_{j-1}).$$

By the previous two claims, we know

$$\mathbb{P}(\bar{\tau}_{\text{spr}} > nh) = \mathbb{P}(\bar{\tau}(n) > nh) \leq n(1 - \Upsilon/\bar{d}_{\text{hon}})^h \leq n\exp(-h\Upsilon/\bar{d}_{\text{hon}}) \leq h^{-2\beta}.$$

By Claim 18 from Appendix F.1, we also have

$$A_{2n\lceil h_\dagger\rceil} \leq (2e)^{2\beta}(nh_\dagger)^\beta, \quad A_{2n(h+1)} \leq e^{2\beta}(2n(h+1))^\beta \leq e^{2\beta}(2n(2h))^\beta = (2e)^{2\beta}(nh)^\beta \ \forall \ h \in \mathbb{N}.$$

Therefore, combining the previous three expressions, we obtain

$$\mathbb{E}A_{2\bar{\tau}_{\text{spr}}} \leq (2e)^{2\beta}n^\beta \left(h_\dagger^\beta + \sum_{h=\lceil h_\dagger\rceil}^{\infty} h^{-\beta}\right) \leq (2e)^{2\beta}n^\beta \left(h_\dagger^\beta + \frac{1}{\beta-1}\right),$$

where the second inequality uses $\beta > 1$, $h_\dagger \geq 2$, and Claim 19 from Appendix F.1. Hence, we have shown $\mathbb{E}[A_{2\bar{\tau}_{\text{spr}}}] = O((nh_\dagger)^\beta) = \tilde{O}((n\bar{d}_{\text{hon}}/\Upsilon)^\beta)$. Note this bound cannot be improved in general – for example, if G_{hon} is a line graph, it becomes $\tilde{O}((n/\Upsilon)^\beta)$, so since $\mathbb{E}[\bar{\tau}_{\text{spr}}]^\beta = O(\mathbb{E}[A_{\bar{\tau}_{\text{spr}}}])$, we have $\mathbb{E}[\bar{\tau}_{\text{spr}}] = \tilde{O}(n/\Upsilon)$, which is the correct scaling (up to log terms) in Definition 1.

E DETAILS FROM SECTION 7

In this appendix, we formalize the analysis that was discussed in Section 7. In particular, the subsequent five sub-appendices provide details on the respective five subsections of Section 7.

E.1 Details from Section 7.1

Let $\psi' = (\rho_2 + \beta(2\alpha - 1))/(2\alpha\beta)$. Note that since $\alpha > 2$ and $\rho_2 \in (0, \beta)$ by assumption,

$$0 < 1 - 1/(2\alpha) = \beta(2\alpha - 1)/(2\alpha\beta) < \psi' < (\beta + \beta(2\alpha - 1))/(2\alpha\beta) = 1,$$

so $\psi = \sqrt{\psi'}$ is well-defined and $\psi \in (0, 1)$. Next, for any $j \in \mathbb{N}$, define

$$\delta_{j,1} = \sqrt{\frac{4\alpha\log A_j}{(\frac{A_j - A_{j-1}}{S+2} - 1) \vee 1}}, \quad \delta_{j,2} = \sqrt{\alpha\log(A_{j-1} \vee 1)}\left(\frac{1-\psi}{\sqrt{\kappa_j}} - \frac{2}{\sqrt{(\frac{A_j - A_{j-1}}{S+2} - 1) \vee 1}}\right) \vee 0.$$

Since $A_j - A_{j-1} = \Theta(j^{\beta-1})$, $\psi < 1$, and $\rho_2 < \beta - 1$, we are guaranteed that $A_j - A_{j-1} \geq 2(S+2)$ and $\delta_{j,2} > 0$ for large j, so the following is well-defined:

$$J_1^\star = \min \left\{ j \in \mathbb{N} : A_{j'} - A_{j'-1} \geq 2(S+2), \delta_{j',2} > 0 \; \forall \; j' \geq j \right\}. \tag{47}$$

Also note $J_1^\star \geq 2$ (since $A_1 - A_0 = 1$). Next, recall from Section 7.1 that

$$\Xi_{j,1}^{(i)} = \left\{ B_j^{(i)} \notin G_{\delta_{j,1}}(S_j^{(i)}) \right\}, \quad \Xi_{j,2}^{(i)} = \left\{ \min_{w \in G_{\delta_{j,2}}(S_j^{(i)})} T_w^{(i)}(A_j) \leq \kappa_j \right\}, \quad \Xi_j^{(i)} = \Xi_{j,1}^{(i)} \cup \Xi_{j,2}^{(i)}.$$

Hence, if we let $\mathcal{S}^{(i)} = \{W \subset [K] : |W| = S+2, \hat{S}^{(i)} \subset W\}$ denote the possible active sets for agent i (i.e., $S_j^{(i)} \in \mathcal{S}^{(i)}$ for any phase j), we can write

$$\Xi_j^{(i)} = \cup_{W \in \mathcal{S}^{(i)}} ((\Xi_{j,1}^{(i)} \cap \{S_j^{(i)} = W\}) \cup ((\Xi_{j,1}^{(i)})^C \cap \Xi_{j,2}^{(i)} \cap \{S_j^{(i)} = W\})).$$

Consequently, by the union bound, we obtain

$$\mathbb{P}(\Xi_j^{(i)}) \leq \sum_{W \in \mathcal{S}^{(i)}} \left(\mathbb{P}(\Xi_{j,1}^{(i)} \cap \{S_j^{(i)} = W\}) + \mathbb{P}((\Xi_{j,1}^{(i)})^C \cap \Xi_{j,2}^{(i)} \cap \{S_j^{(i)} = W\}) \right). \tag{48}$$

The next two claims bound the two summands on the right side.

CLAIM 11. *Under the assumptions of Theorem 2, for any $i \in [n]$, $j \geq J_1^\star$, and $W \in \mathcal{S}^{(i)}$, we have*

$$\mathbb{P}(\Xi_{j,1}^{(i)} \cap \{S_j^{(i)} = W\}) \leq 4S(j-1)^{\beta(3-2\alpha)}/(2\alpha - 3).$$

PROOF. If $W \setminus G_{\delta_{j,1}}(W) = \emptyset$, the claim is immediate. Otherwise, $\Xi_{j,1}^{(i)} \cap \{S_j^{(i)} = W\}$ implies $B_j^{(i)} = w$ for some $w \in W \setminus G_{\delta_{j,1}}(W)$. Thus, by the union bound,

$$\mathbb{P}(\Xi_{j,1}^{(i)} \cap \{S_j^{(i)} = W\}) \leq \sum_{w \in W \setminus G_{\delta_{j,1}}(W)} \mathbb{P}(B_j^{(i)} = w, \min W \in S_j^{(i)}). \tag{49}$$

Fix $w \in W \setminus G_{\delta_{j,1}}(W)$. Then $B_j^{(i)} = w$ implies $T_w^{(i)}(A_j) - T_w^{(i)}(A_{j-1}) \geq (A_j - A_{j-1})/(S+2)$ (else, by definition of $B_j^{(i)}$, $\sum_{k \in S_j^{(i)}} (T_k^{(i)}(A_j) - T_k^{(i)}(A_{j-1})) < A_j - A_{j-1}$). Since $A_j - A_{j-1} \geq S+2$ (by $j \geq J_1^\star$), we conclude $T_w^{(i)}(A_j) - T_w^{(i)}(A_{j-1}) \geq 1$, so there exists $t \in \{1 + A_{j-1}, \ldots, A_j\}$ such that

$$T_w^{(i)}(t-1) - T_w^{(i)}(A_{j-1}) = T_w^{(i)}(A_j) - T_w^{(i)}(A_{j-1}) - 1, \quad I_t^{(i)} = w.$$

Combining and using the union bound and with $T_w^{(i)}(A_{j-1}) \geq 0$ by definition, we obtain

$$\mathbb{P}(B_j^{(i)} = w, \min W \in S_j^{(i)}) \leq \sum_{t=1+A_{j-1}}^{A_j} \mathbb{P}\left(T_w^{(i)}(t-1) \geq \frac{A_j - A_{j-1}}{S+2} - 1, \min W \in S_j^{(i)}, I_t^{(i)} = w \right). \tag{50}$$

Now fix t as in the summation. Observe that since $w \in W \setminus G_{\delta_{j,1}}(W)$ and $j \geq J_1^\star$, we have

$$\mu_{\min W} - \mu_w > \delta_{j,1} = \sqrt{\frac{4\alpha \log A_j}{\frac{A_j - A_{j-1}}{S+2} - 1}} \geq \sqrt{\frac{4\alpha \log t}{\frac{A_j - A_{j-1}}{S+2} - 1}}.$$

Therefore, for any such t, we can apply a basic bandit tail (namely, Corollary 5 from Appendix F.2 with the parameters $k_1 = w$, $k_2 = \min W$, and $\ell = (A_j - A_{j-1})/(S+2) - 1$) to obtain

$$\mathbb{P}\left(T_w^{(i)}(t-1) \geq \frac{A_j - A_{j-1}}{S+2} - 1, \min W \in S_j^{(i)}, I_t^{(i)} = w \right) \leq 2t^{2(1-\alpha)}.$$

Substituting into (50) and using Claim 19 from Appendix F.1 (which applies since $\alpha > 2$), we obtain

$$\mathbb{P}(B_j^{(i)} = w, \min W \in S_j^{(i)}) \leq 2 \sum_{t=1+A_{j-1}}^{\infty} t^{2(1-\alpha)} \leq \frac{2A_{j-1}^{3-2\alpha}}{2\alpha - 3} \leq \frac{2(j-1)^{\beta(3-2\alpha)}}{2\alpha - 3}.$$

Substituting into (49) and using $|W \setminus G_{\delta_{j,1}}(W)| \leq |W| - 1 = S + 1 \leq 2S$ completes the proof. □

CLAIM 12. *Under the assumption of Theorem 2, for any $i \in [n]$, $j \geq J_1^{\star}$, and $W \in \mathcal{S}^{(i)}$,*

$$\mathbb{P}((\Xi_{j,1}^{(i)})^C \cap \Xi_{j,2}^{(i)} \cap \{S_j^{(i)} = W\}) \leq 6 \cdot 2^{\beta} S(j-1)^{\beta(3-2\alpha)}/(2\alpha - 3).$$

PROOF. By definition, we have

$$(\Xi_{j,1}^{(i)})^C \cap \Xi_{j,2}^{(i)} \cap \{S_j^{(i)} = W\} = \left\{ B_j^{(i)} \in G_{\delta_{j,1}}(W), \min_{w \in G_{\delta_{j,2}}(W)} T_w^{(i)}(A_j) \leq \kappa_j, S_j^{(i)} = W \right\}.$$

As in the proof of Claim 11, we know that

$$T_{B_j^{(i)}}^{(i)}(A_j) \geq T_{B_j^{(i)}}^{(i)}(A_j) - T_{B_j^{(i)}}^{(i)}(A_{j-1}) \geq \frac{A_j - A_{j-1}}{S+2} > \left(\frac{A_j - A_{j-1}}{S+2} - 1 \right) \vee 1 > \kappa_j,$$

where the final inequality holds since $\delta_{j,2} > 0$ by assumption $j \geq J_1^{\star}$, which implies

$$\frac{(\frac{A_j - A_{j-1}}{S+2} - 1) \vee 1}{\kappa_j} > \left(\frac{2}{1-\psi} \right)^2 > 1.$$

Thus, $(\Xi_{j,1}^{(i)})^C \cap \Xi_{j,2}^{(i)} \cap \{S_j^{(i)} = W\}$ implies $B_j^{(i)} \notin \arg\min_{w \in G_{\delta_{j,2}}(W)} T_w^{(i)}(A_j)$, so by the union bound,

$$\mathbb{P}((\Xi_{j,1}^{(i)})^C \cap \Xi_{j,2}^{(i)} \cap \{S_j^{(i)} = W\}) \tag{51}$$

$$\leq \sum_{w_1 \in G_{\delta_{j,1}}(W), w_2 \in G_{\delta_{j,2}}(W) \setminus \{w_1\}} \mathbb{P}\left(T_{w_1}^{(i)}(A_j) - T_{w_1}^{(i)}(A_{j-1}) \geq \frac{A_j - A_{j-1}}{S+2}, T_{w_2}^{(i)}(A_j) \leq \kappa_j, S_j^{(i)} = W \right).$$

Now fix w_1, w_2 as in the double summation. Then similar to the proof of Claim 11,

$$\mathbb{P}\left(T_{w_1}^{(i)}(A_j) - T_{w_1}^{(i)}(A_{j-1}) \geq \frac{A_j - A_{j-1}}{S+2}, T_{w_2}^{(i)}(A_j) \leq \kappa_j, S_j^{(i)} = W \right) \tag{52}$$

$$\leq \sum_{t=1+A_{j-1}}^{A_j} \mathbb{P}\left(T_{w_1}^{(i)}(t-1) \geq \frac{A_j - A_{j-1}}{S+2} - 1, T_{w_2}^{(i)}(A_j) \leq \kappa_j, w_2 \in S_j^{(i)}, I_t^{(i)} = w_1 \right)$$

$$\leq \sum_{t=1+A_{j-1}}^{A_j} 2\kappa_j t^{1-2\alpha\psi^2} = 2 \sum_{t=1+A_{j-1}}^{A_j} (\kappa_j t^{-\rho_2/\beta}) t^{2-2\alpha},$$

where the second bound follows from applying Claim 21 from Appendix F.2 with $k_1 = w_1$, $k_2 = w_2$, $\ell = (A_j - A_{j-1})/(S+2) - 1$, $u = \kappa_j$, and $\iota = \psi$; note this claim applies since by assumption $j \geq J_1^{\star}$,

$$\mu_{w_2} - \mu_{w_1} \geq \mu_{w_2} - \mu_{\min W} \geq -\delta_{j,2} \geq \sqrt{\alpha \log t} \left(\frac{2}{\sqrt{\frac{A_j - A_{j-1}}{S+2} - 1}} - \frac{1-\psi}{\sqrt{\kappa_j}} \right).$$

Next, observe that for any $t \geq A_{j-1} \geq (j-1)^{\beta}$, by definition of κ_j, $j \geq J_1^{\star} \geq 2$, and $\rho_2 < \beta - 1$,

$$\kappa_j t^{-\rho_2/\beta} \leq \kappa_j (j-1)^{-\rho_2} = (1 - 1/j)^{-\rho_2}/(K^2 S) \leq 2^{\beta-1}/S.$$

Similar to the proof of Claim 11, we can then use Claim 19 to obtain

$$2 \sum_{t=1+A_{j-1}}^{A_j} (\kappa_j t^{-\rho_2/\beta}) t^{2-2\alpha} \leq \frac{2^\beta}{S} \sum_{t=1+A_{j-1}}^{A_j} t^{2-2\alpha} \leq \frac{2^\beta (j-1)^{\beta(3-2\alpha)}}{S(2\alpha-3)}.$$

Combining with (51) and (52) completes the proof, since

$$|G_{\delta_{j,1}}(W)||G_{\delta_{j,2}}(W) \setminus \{w_1\}| < (S+2)(S+1) \leq (3S)(2S) = 6S^2. \qquad \square$$

Finally, we provide the tail for $\tau_{\text{arm}} = \inf\{j \in \mathbb{N} : \mathbb{1}(\Xi_{j'}^{(i)}) = 0 \; \forall \; i \in [n], j' \geq j\}$.

LEMMA 9. *Under the assumptions of Theorem 2, for any $j \geq J_1^\star \vee 3$,*

$$\mathbb{P}(\tau_{arm} > j) \leq \frac{(6^\beta+2)nK^2 S(j-2)^{\beta(3-2\alpha)+1}}{(2\alpha-3)(\beta(2\alpha-3)-1)}.$$

PROOF. By (48), Claims 11 and 12, $|\mathcal{S}^{(i)}| = \binom{K}{2} < \frac{K^2}{2}$, and $\beta \geq 1$, we can write

$$\mathbb{P}(\Xi_{j'}^{(i)}) \leq \frac{(6 \cdot 2^\beta + 4)K^2 S(j'-1)^{\beta(3-2\alpha)}}{2(2\alpha-3)} \leq \frac{(6^\beta+2)K^2 S(j'-1)^{\beta(3-2\alpha)}}{2\alpha-3}.$$

Thus, because $\tau_{\text{arm}} > j$ implies $\mathbb{1}(\cup_{i=1}^n \Xi_{j'}^{(i)}) = 1$ for some $i \in [n]$ and $j' \geq j$, the union bound gives

$$\mathbb{P}(\tau_{\text{arm}} > j) \leq \sum_{j'=j}^\infty \sum_{i=1}^n \mathbb{P}(\Xi_{j'}^{(i)}) \leq \frac{(6^\beta+2)nK^2 S}{2\alpha-3} \sum_{j'=j}^\infty (j'-1)^{\beta(3-2\alpha)}.$$

Finally, use Claim 19 (which applies since $\beta(2\alpha-3) > 1$) to bound the sum. $\qquad \square$

E.2 Details from Section 7.2

Recall $\theta_j = (j/3)^{\rho_1}$, where $\rho_1 \in (0, 1/\eta]$ and $\eta > 1$. Hence, for all large j, we have

$$1 \leq \lfloor \theta_j \rfloor \leq j-2, \quad \lceil \lfloor \theta_j \rfloor^\eta \rceil \leq \theta_j^\eta + 1 \leq (j/3) + 1 \leq (j-2) - j/3.$$

Thus, the following is well-defined:

$$J_2^\star = \min\left\{ j \in \mathbb{N} : 1 \leq \lfloor \theta_{j'} \rfloor \leq j'-2, j'/3 \leq (j'-2) - \lceil \lfloor \theta_{j'} \rfloor^\eta \rceil \; \forall \; j' \geq j \right\}. \tag{53}$$

Now recall from Section 7.2 that $\Xi_j^{(i \to i')} = \cap_{j'=\lfloor \theta_j \rfloor}^{j-2} \{H_{j'}^{(i')} \neq i\}$, and

$$\tau_{\text{com}} = \inf\{j \in \mathbb{N} : \mathbb{1}(\cup_{(i,i') \in E_{\text{hon}}} \Xi_{j'}^{(i \to i')}) = 0 \; \forall \; j' \in \{j, j+1, \ldots\}\}.$$

The next lemma provides a tail bound for this random phase.

LEMMA 10. *Under the assumptions of Theorem 2, for any $j \geq J_2^\star$,*

$$\mathbb{P}(\tau_{com} > j) \leq 3(n+m)^3 \exp(-j/(3\bar{d})).$$

PROOF. We first use the union bound to write

$$\mathbb{P}(\tau_{\text{com}} > j) \leq \sum_{j'=j}^\infty \sum_{i \to i' \in E_{\text{hon}}} \mathbb{P}(\Xi_{j'}^{(i \to i')}). \tag{54}$$

Fix $i \to i' \in E_{\text{hon}}$ and $j' \geq j$. Suppose $\Xi_{j'}^{(i \to i')}$ holds. Then $i \notin P_{j''}^{(i')} \setminus P_{j''-1}^{(i')} \; \forall \; j'' \in \{\lfloor \theta_{j'} \rfloor+1, \ldots, j'-1\}$; else, we can find $j'' \in \{\lfloor \theta_{j'} \rfloor+1, \ldots, j'-1\}$ such that $H_{j''-1}^{(i')} = i$ (i.e., $H_{j''}^{(i')} = i$ for $j'' \in \{\lfloor \theta_{j'} \rfloor, \ldots, j'-2\}$), contradicting $\Xi_{j'}^{(i \to i')}$. Hence, we have two cases: $i \notin P_{j''}^{(i')} \setminus P_{j''-1}^{(i')} \; \forall \; j'' \in [j'-1]$, or $i \in P_{j''}^{(i')} \setminus P_{j''-1}^{(i')}$

for some $j'' \in [j'-1]$ and $\max\{j'' \in [j'-1] : i \in P_{j''}^{(i')} \setminus P_{j''-1}^{(i')}\} \leq \lfloor \theta_{j'} \rfloor$. In the former case, $i \notin P_{j''}^{(i')} \ \forall \ j'' \in [j'-1]$; in the latter, $i \notin P_{j''}^{(i')} \ \forall \ j'' \in \{\lceil \lfloor \theta_{j'} \rfloor^\eta \rceil + 1, \dots, j'-1\}$. Thus,

$$\mathbb{P}(\Xi_{j'}^{(i \to i')}) \leq \mathbb{P}(\cap_{j''=\lceil \lfloor \theta_{j'} \rfloor^\eta \rceil + 1}^{j'-2} \{i \notin P_{j''}^{(i')}, H_{j''}^{(i')} \neq i\})$$
$$= \mathbb{E}[\mathbb{1}(\cap_{j''=\lceil \lfloor \theta_{j'} \rfloor^\eta \rceil + 1}^{j'-3} \{i \notin P_{j''}^{(i')}, H_{j''}^{(i')} \neq i\}) \mathbb{1}(i \notin P_{j'-2}^{(i')}) \mathbb{P}_{j'-2}(H_{j'-2}^{(i)} \neq i)].$$

Now given that $i \notin P_{j'-2}^{(i')}$, $H_{j'-2}^{(i)}$ is sampled uniformly from a set of at most $\bar{d}$ elements which includes i, so $\mathbb{P}_{j'-2}(H_{j'-2}^{(i)} \neq i) \leq (1 - 1/\bar{d})$. Substituting above and iterating yields

$$\mathbb{P}(\Xi_{j'}^{(i \to i')}) \leq (1 - 1/\bar{d})^{j'-2-\lceil \lfloor \theta'_j \rfloor^\eta \rceil} \leq (1 - 1/\bar{d})^{j'/3}, \tag{55}$$

where the final inequality uses $j' \geq j \geq J_2^\star$. Combining (54) and (55) and computing a geometric series, we obtain

$$\mathbb{P}(\tau_{\text{com}} > j) \leq \frac{|E_{\text{hon}}|(1 - 1/\bar{d})^{j/3}}{1 - (1 - 1/\bar{d})^{j/3}} \leq \frac{|E_{\text{hon}}|(1 - 1/\bar{d})^{j/3}}{1 - (1 - 1/\bar{d})^{1/3}}$$

Finally, using $|E_{\text{hon}}| \leq n^2 < (n + m)^2$, $1 - x \leq e^{-x} \ \forall \ x \in \mathbb{R}$, $(1 + x)^r \leq 1 + rx$ for any $r \in (0, 1)$ and $x \geq -1$, and $\bar{d} \leq m + n$, we obtain the desired bound. $\square$

E.3 Details from Section 7.3

We begin with some intermediate claims.

CLAIM 13. *If the assumptions of Theorem 2 hold, then for any $i \in [n]$ and $j \geq \tau_{arm}$, we have*

$$\mu_{\min S_j^{(i)}} \leq \mu_{B_j^{(i)}} + \delta_{j,1} \leq \mu_{\min S_{j+1}^{(i)}} + \delta_{j,1}.$$

PROOF. The first inequality holds by definition of τ_{arm} and assumption $j \geq \tau_{\text{arm}}$. The second holds since $\min S_{j+1}^{(i)}$ is the best arm in $S_{j+1}^{(i)}$ and $B_j^{(i)} \in S_{j+1}^{(i)}$ in the algorithm. $\square$

CLAIM 14. *If the assumptions of Theorem 2 hold, then for any $i \in [n]$ and $j' \geq j \geq \tau_{arm}$,*

$$\mu_{\min S_{j'}^{(i)}} \geq \mu_{\min S_j^{(i)}} - (K - 1) \sup_{j'' \in \{j, \dots, j'\}} \delta_{j'',1}.$$

PROOF. If $j = j'$ or $\mu_{\min S_{j'}^{(i)}} \geq \mu_{\min S_j^{(i)}}$, the bound is immediate, so we assume $j' > j$ and $\mu_{\min S_{j'}^{(i)}} < \mu_{\min S_j^{(i)}}$ for the remainder of the proof. Under this assumption, there must exist a phase $j'' \in \{j + 1, \dots, j'\}$ at which the mean of the best active arm reaches a new strict minimum since j, i.e., $\mu_{\min S_{j''}^{(i)}} < \mu_{\min S_j^{(i)}}$. Let m denote the number of phases this occurs and $j^{(1)}, \dots, j^{(m)}$ these (ordered) phases; formally,

$$j^{(0)} = j, \quad j^{(l)} = \min\left\{ j'' \in \{j^{(l-1)} + 1, \dots, j'\} : \mu_{\min S_{j''}^{(i)}} < \mu_{\min S_{j^{(l-1)}}^{(i)}} \right\} \ \forall \ l \in [m].$$

The remainder of the proof relies on the following three inequalities:

$$m \leq K - 1, \quad \mu_{\min S_{j^{(m)}}^{(i)}} \leq \mu_{\min S_{j'}^{(i)}}, \quad \mu_{\min S_{j^{(l-1)}}^{(i)}} \leq \mu_{\min S_{j^{(l)}-1}^{(i)}} \ \forall \ l \in [m]. \tag{56}$$

The first inequality holds since $\mu_{\min S_{j^{(0)}}^{(i)}} > \cdots > \mu_{\min S_{j^{(m)}}^{(i)}}$ by definition, so $\min S_{j^{(0)}}^{(i)}, \dots, \min S_{j^{(m)}}^{(i)}$ are distinct arms; since there are $m + 1$ of these arms and K in total, $m + 1 \leq K$. For the second, we have $\mu_{\min S_{j^{(m)}}^{(i)}} = \mu_{\min S_{j'}^{(i)}}$ when $j^{(m)} = j'$ and $\mu_{\min S_{j^{(m)}}^{(i)}} \leq \mu_{\min S_{j'}^{(i)}}$ when $j^{(m)} < j'$ (if the latter

fails, we contradict the definition of m). For the third, note $j^{(l)} \geq j^{(l-1)} + 1$ by construction, so if $j^{(l)} = j^{(l-1)} + 1$, the bound holds with equality; else, $j^{(l)} - 1 \geq j^{(l-1)} + 1$, so if the bound fails,

$$j^{(l)} - 1 \in \left\{ j'' \in \{ j^{(l-1)} + 1, \ldots, j' \} : \mu_{\min S^{(i)}_{j''}} < \mu_{\min S^{(i)}_{j^{(l-1)}}} \right\},$$

which is a contradiction, since $j^{(l)}$ is the minimal element of the set at right. Hence, (56) holds. Combined with Claim 13 (note $j^{(l)} - 1 \geq j^{(0)} = j \geq \tau_{\mathrm{arm}} \ \forall \ l \in [m]$, as required), we obtain

$$\mu_{\min S^{(i)}_j} - \mu_{\min S^{(i)}_{j'}} = \sum_{l=1}^{m} \left(\mu_{\min S^{(i)}_{j^{(l-1)}}} - \mu_{\min S^{(i)}_{j^{(l)}}} \right) + \mu_{\min S^{(i)}_{j^{(m)}}} - \mu_{\min S^{(i)}_{j'}}$$

$$\leq \sum_{l=1}^{m} \left(\mu_{\min S^{(i)}_{j^{(l-1)}}} - \mu_{\min S^{(i)}_{j^{(l)}}} \right) \leq \sum_{l=1}^{m} \delta_{j^{(l)}-1,1} \leq (K-1) \sup_{j'' \in \{ j, \ldots, j' \}} \delta_{j'',1},$$

where the last inequality uses $j = j^{(0)} \leq j^{(l)} - 1 < j' \ \forall \ l \in [m]$. □

As a simple corollary of the previous two claims, we have the following.

COROLLARY 3. *If the assumptions of Theorem 2 hold, then for any $i \in [n]$ and $j' \geq j \geq \tau_{arm}$,*

$$\mu_{B^{(i)}_{j'}} \geq \mu_{\min S^{(i)}_j} - K \sup_{j'' \in \{ j, \ldots, j' \}} \delta_{j'',1}.$$

PROOF. Since $j' \geq j \geq \tau_{\mathrm{arm}}$, we can use Claims 13 and 14, respectively, to obtain

$$\mu_{B^{(i)}_{j'}} \geq \mu_{\min S^{(i)}_{j'}} - \delta_{j',1} \geq \mu_{\min S^{(i)}_{j'}} - \sup_{j'' \in \{ j, \ldots, j' \}} \delta_{j'',1} \geq \mu_{\min S^{(i)}_j} - K \sup_{j'' \in \{ j, \ldots, j' \}} \delta_{j'',1}. \quad □$$

Next, inspecting the analysis in Section 7.3, we see that $\delta_{j,2} \geq (K+1) \sup_{j' \in \{ \lfloor \theta_j \rfloor, \ldots, j \}} \delta_{j',1}$ for large j. Thus, the following is well-defined:

$$J_3^{\star} = \min \left\{ j \in \mathbb{N} : \delta_{j',2} \geq (K+1) \sup_{j'' \in \{ \lfloor \theta_{j'} \rfloor, \ldots, j' \}} \delta_{j'',1} \ \forall \ j' \geq j \right\}. \tag{57}$$

As discussed in Section 7.3, we can now show that no *new* accidental blocking occurs at late phases, at least among pairs of honest agents that have recently communicated.

CLAIM 15. *Under the assumptions of Theorem 2, if $j \geq J_3^{\star}$, $\lfloor \theta_j \rfloor \geq \tau_{arm}$, and $H^{(i')}_{j'} = i$ for some $i, i' \in [n]$ and $j' \in \{ \lfloor \theta_j \rfloor, \ldots, j - 2 \}$, then $i' \notin P^{(i)}_j \setminus P^{(i)}_{j-1}$.*

PROOF. Suppose instead that $i' \in P^{(i)}_j \setminus P^{(i)}_{j-1}$. Then by the algorithm,

$$B^{(i)}_j = \cdots = B^{(i)}_{\lfloor \theta_j \rfloor}, \quad T^{(i)}_{R^{(i)}_{j-1}}(A_j) \leq \kappa_j, \quad R^{(i)}_{j-1} = B^{(i')}_{j-1} \in S^{(i)}_j. \tag{58}$$

Since $j \geq \lfloor \theta_j \rfloor \geq \tau_{\mathrm{arm}}$, this implies $B^{(i')}_{j-1} \notin G_{\delta_{j,2}}(S^{(i)}_j)$. We then observe the following:

- Since $B^{(i')}_{j-1} \in S^{(i)}_j \setminus G_{\delta_{j,2}}(S^{(i)}_j)$, the definition of $G_{\delta_{j,2}}(S^{(i)}_j)$ implies $\mu_{B^{(i')}_{j-1}} < \mu_{\min S^{(i)}_j} - \delta_{j,2}$.
- Again using $j \geq \tau_{\mathrm{arm}}$, Claim 13 implies $\mu_{\min S^{(i)}_j} \leq \mu_{B^{(i)}_j} + \delta_{j,1}$.
- Since $\lfloor \theta_j \rfloor \leq j' \leq j$, (58) implies $B^{(i)}_j = B^{(i)}_{j'}$, so $\mu_{B^{(i)}_j} = \mu_{B^{(i)}_{j'}}$.
- Since $H^{(i')}_{j'} = i$, the algorithm implies $B^{(i)}_{j'} \in S^{(i')}_{j'+1}$, so $\mu_{B^{(i)}_{j'}} \leq \mu_{\min S^{(i')}_{j'+1}}$.
- Since $\tau_{\mathrm{arm}} \leq \lfloor \theta_j \rfloor < j' + 1 < j$, Corollary 3 implies $\mu_{\min S^{(i')}_{j'+1}} \leq \mu_{B^{(i')}_{j-1}} + K \sup_{j'' \in \{ \lfloor \theta_j \rfloor, \ldots, j \}} \delta_{j'',1}$.

Stringing together these inequalities, we obtain

$$\mu_{B_{j-1}^{(i')}} < \mu_{B_{j-1}^{(i')}} + K \sup_{j'' \in \{\lfloor \theta_j \rfloor, \ldots, j\}} \delta_{j'',1} - \delta_{j,2} + \delta_{j,1} \le \mu_{B_{j-1}^{(i')}} + (K+1) \sup_{j'' \in \{\lfloor \theta_j \rfloor, \ldots, j\}} \delta_{j'',1} - \delta_{j,2},$$

i.e., $\delta_{j,2} < (K+1) \sup_{j'' \in \{\lfloor \theta_j \rfloor, \ldots, j\}} \delta_{j'',1}$, which contradicts $j \ge J_3^\star$. □

Finally, we prove that honest agents eventually stop blocking each other.

LEMMA 11. *Under the assumptions of Theorem 2, if $j \ge J_2^\star$, $\lfloor \theta_j \rfloor \ge \tau_{com} \vee J_3^\star$, and $\lfloor \theta_{\lfloor \theta_j \rfloor} \rfloor \ge \tau_{arm}$, then for any $i \in [n]$ and $j' \ge j$, $P_{j'}^{(i)} \cap [n] = \emptyset$.*

PROOF. Fix $i \to i' \in E_{hon}$ and $j' \ge j$; we aim to show $i' \notin P_{j'}^{(i')}$. This clearly holds if $i' \notin P_{j''}^{(i)} \setminus P_{j''-1}^{(i)} \ \forall j'' \le j'$. Otherwise, $j_{B,1} = \max\{j'' \le j' : i' \in P_{j''}^{(i)} \setminus P_{j''-1}^{(i)}\}$ (the latest phase up to and including j' that i blocked i') is well-defined. We consider two cases of $j_{B,1}$.

The first case is $j_{B,1} \le \lfloor \theta_j \rfloor$. Let $j_{B,2} = \min\{j'' > j' : P_{j''}^{(i)} \setminus P_{j''-1}^{(i)}\}$ denote the first phase after j' that i blocked i'. Combined with the definition of $j_{B,1}$, the algorithm implies

$$i' \notin P_{j''}^{(i)} \ \forall j'' \in \{\lceil j_{B,1}^\eta \rceil + 1, \ldots, j_{B,2}\}. \tag{59}$$

Since $j \ge J_2^\star$, the definition of $J_2^\star$ implies

$$\lceil \lfloor \theta_j \rfloor^\eta \rceil + 1 \le (j-2) - (j/3) + 1 < (j-2) + 1 = j - 1,$$

so $\lceil j_B^\eta \rceil + 1 < j - 1 < j'$ as well. Combined with $j' \le j_{B,2} - 1$ by definition, $i' \notin P_{j'}^{(i')}$ holds by (59).

The second case is $j_{B,1} > \lfloor \theta_j \rfloor$. By assumption, $j_{B,1} > \lfloor \theta_j \rfloor \ge \tau_{com}$. Hence, by definition of τ_{com}, $H_{j_C}^{(i')} = i$ for some $j_C \in \{\lfloor \theta_{j_{B,1}} \rfloor, \ldots, j_{B,1} - 2\}$. Note that $j_{B,1} \ge \lfloor \theta_j \rfloor \ge J_3^\star$ and $\lfloor \theta_{j_{B,1}} \rfloor \ge \lfloor \theta_{\lfloor \theta_j \rfloor} \rfloor \ge \tau_{arm}$ by assumption (and by monotonicity of $\{\lfloor \theta_{j''} \rfloor\}_{j'' \in \mathbb{N}}$ in the latter case). Hence, we can apply Claim 15 (with $j = j_{B,1}$ and $j' = j_C$ in the claim) to obtain $i' \notin P_{j_{B,1}}^{(i)} \setminus P_{j_{B,1}-1}^{(i)}$. This is a contradiction. □

E.4 Details from Section 7.4

We first verify that the sampling strategy in Section 7.4 is identical to the one in Algorithm 2.

CLAIM 16. *Suppose we replace the sampling of $H_j^{(i)}$ in Algorithm 2 with the sampling of Section 7.4, and recall $\mathbb{P}_j$ denotes probability conditioned on all randomness before this sampling occurs. Then*

$$\mathbb{P}_j(H_j^{(i)} = i') = \mathbb{1}(i' \in N(i) \setminus P_j^{(i)}) / |N(i) \setminus P_j^{(i)}| \ \forall i \in [n], i' \in [n+m], j \in \mathbb{N}.$$

PROOF. Since $\mathbb{P}_j$ conditions on $P_j^{(i)}$, we can prove the identity separately in the cases $P_j^{(i)} \cap [n] \ne \emptyset$ and $P_j^{(i)} \cap [n] = \emptyset$. The identity is immediate in the former case. For the latter, we have

$$\mathbb{P}_j(H_j^{(i)} = i') = \mathbb{P}_j(H_j^{(i)} = i' | Y_j^{(i)} = 1) \mathbb{P}_j(Y_j^{(i)} = 1) + \mathbb{P}_j(H_j^{(i)} = i' | Y_j^{(i)} = 0) \mathbb{P}_j(Y_j^{(i)} = 0)$$

$$= \frac{\mathbb{1}(i' \in N_{hon}(i))}{d_{hon}(i)} \frac{d_{hon}(i)}{|N(i) \setminus P_j^{(i)}|} + \frac{\mathbb{1}(i' \in N_{mal}(i) \setminus P_j^{(i)})}{|N_{mal}(i) \setminus P_j^{(i)}|} \left(1 - \frac{d_{hon}(i)}{|N(i) \setminus P_j^{(i)}|} \right)$$

$$= \frac{\mathbb{1}(i' \in N_{hon}(i))}{|N(i) \setminus P_j^{(i)}|} + \frac{\mathbb{1}(i' \in N_{mal}(i) \setminus P_j^{(i)})}{|N_{mal}(i) \setminus P_j^{(i)}|} \frac{|N(i) \setminus P_j^{(i)}| - d_{hon}(i)}{|N(i) \setminus P_j^{(i)}|}$$

$$= \frac{\mathbb{1}(i' \in N_{hon}(i)) + \mathbb{1}(i' \in N_{mal}(i) \setminus P_j^{(i)})}{|N(i) \setminus P_j^{(i)}|} = \frac{\mathbb{1}(i' \in N(i) \setminus P_j^{(i)})}{|N(i) \setminus P_j^{(i)}|}. \quad \square$$

Next, note that since $\delta_{j',1} \to 0$ as $j' \to \infty$, the following is well-defined:

$$J_4^\star = \min\{j \in \{2, 3, \ldots\} : \delta_{j',1} < \Delta_2 \ \forall \ j' \geq \lfloor j/2 \rfloor\}. \tag{60}$$

As in Section 7.4, we let $\mathcal{I}_j = \{i \in [n] : 1 \in S_j^{(i)}\}$ be the agents with the best arm active at phase j.

CLAIM 17. *Under the assumptions of Theorem 2, if $j \geq J_4^\star$ and $\lfloor j/2 \rfloor \geq \tau_{arm}$, then $B_{j'}^{(i)} = 1 \ \forall \ j' \geq \lfloor j/2 \rfloor, i \in \mathcal{I}_{j'}$.*

PROOF. Suppose instead that $B_{j'}^{(i)} \neq 1$ for some $j' \geq \lfloor j/2 \rfloor$ and $i \in \mathcal{I}_{j'}$. Since $j' \geq \lfloor j/2 \rfloor \geq \lfloor J_4^\star/2 \rfloor$, we know $\delta_{j',1} < \Delta_2$. Hence, because $1 \in S_{j'}^{(i)}$ by definition of $\mathcal{I}_{j'}$, we have $G_{\delta_{j',1}}(S_{j'}^{(i)}) = \{1\}$. Combined with $B_{j'}^{(i)} \neq 1$, we get $B_{j'}^{(i)} \in S_{j'}^{(i)} \setminus G_{\delta_{j',1}}(S_{j'}^{(i)})$, which contradicts $j' \geq \lfloor j/2 \rfloor \geq \tau_{arm}$. □

Finally, recall $\tau_{spr} = \inf\{j \in \mathbb{N} : B_{j'}^{(i)} = 1 \ \forall \ i \in [n], j' \geq j\}$ and $\bar{\tau}_{spr} = \inf\{j \in \mathbb{N} : \bar{\mathcal{I}}_j = [n]\}$, where $\{\bar{\mathcal{I}}_j\}_{j=1}^\infty$ is the noisy rumor process from Definition 1.

LEMMA 12. *Under the assumptions of Theorem 2, if $j \geq J_4^\star$, $\lfloor j/2 \rfloor \geq J_2^\star$, and $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_3^\star$, then*

$$\mathbb{P}(\tau_{com} \leq \lfloor \theta_{\lfloor j/2 \rfloor} \rfloor, \tau_{arm} \leq \lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor, \tau_{spr} > j) \leq \mathbb{P}(\bar{\tau}_{spr} > j/2).$$

PROOF. Let $\mathcal{E}_j = \{\tau_{com} \leq \lfloor \theta_{\lfloor j/2 \rfloor} \rfloor, \tau_{arm} \leq \lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor\}$ and

$$\tilde{\mathcal{I}}_{\lfloor j/2 \rfloor} = \{i^\star\}, \ \tilde{\mathcal{I}}_j = \bar{\mathcal{I}}_{j-1} \cup \{i \in [n] \setminus \tilde{\mathcal{I}}_{j-1} : \bar{Y}_j^{(i)} = 1, \bar{H}_j^{(i)} \in \tilde{\mathcal{I}}_{j-1}\} \ \forall \ j \in \{\lfloor j/2 \rfloor + 1, \lfloor j/2 \rfloor + 2, \ldots\}. \tag{61}$$

Then it suffices to prove the following:

$$\mathbb{P}_{\lfloor j/2 \rfloor}(\mathcal{E}_j \cap \{\tau_{spr} > j\}) \leq \mathbb{P}_{\lfloor j/2 \rfloor}(\mathcal{E}_j \cap \{\mathcal{I}_j \neq [n]\}) \leq \mathbb{P}_{\lfloor j/2 \rfloor}(\tilde{\mathcal{I}}_j \neq [n]) = \mathbb{P}(\bar{\tau}_{spr} > j/2). \tag{62}$$

For the first inequality in (62), we begin by proving

$$\mathcal{E}_j \cap \{\tau_{spr} > j\} \cap \{\mathcal{I}_j = [n]\} = \emptyset. \tag{63}$$

To do so, we show $\mathcal{E}_j, \mathcal{I}_j = [n] \Rightarrow \tau_{spr} \leq j$. Assume $\mathcal{E}_j$ and $\mathcal{I}_j = [n]$ hold; by definition of τ_{spr}, we aim to show $B_{j'}^{(i)} = 1 \ \forall \ i \in [n], j' \geq j$. We use induction. The base of induction ($j' = j$) holds by $\mathcal{I}_j = [n]$, $\mathcal{E}_j \subset \{\tau_{arm} \leq \lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \leq \lfloor j/2 \rfloor\}$, and Claim 17. Given the inductive hypothesis $B_{j'}^{(i)} = 1 \ \forall \ i \in [n]$, we have $1 \in S_{j'+1}^{(i)}$ by the algorithm and $\mathcal{I}_{j'+1} = [n]$ by definition, so we again use the assumption that $\mathcal{E}_j$ holds and Claim 17 to obtain $B_{j'+1}^{(i)} = 1 \ \forall \ i \in [n]$. Hence, (63) holds, so

$$\mathbb{P}_{\lfloor j/2 \rfloor}(\mathcal{E}_j \cap \{\tau_{spr} > j\}) = \mathbb{P}_{\lfloor j/2 \rfloor}(\mathcal{E}_j \cap \{\tau_{spr} > j\} \cap \{\mathcal{I}_j \neq [n]\}) \leq \mathbb{P}_{\lfloor j/2 \rfloor}(\mathcal{E}_j \cap \{\mathcal{I}_j \neq [n]\}).$$

For the second inequality in (62), we claim, and will return to prove, the following:

$$\mathcal{E}_j \subset \cap_{j'=\lfloor j/2 \rfloor}^j \{\tilde{\mathcal{I}}_{j'} \subset \mathcal{I}_{j'}\}. \tag{64}$$

Assuming (64) holds, we obtain

$$\mathbb{P}_{\lfloor j/2 \rfloor}(\mathcal{E}_j \cap \{\mathcal{I}_j \neq [n]\}) \leq \mathbb{P}_{\lfloor j/2 \rfloor}(\tilde{\mathcal{I}}_j \subset \mathcal{I}_j, \mathcal{I}_j \neq [n],) \leq \mathbb{P}_{\lfloor j/2 \rfloor}(\tilde{\mathcal{I}}_j \neq [n]).$$

Hence, it only remains to prove (64). We show by induction on j' when $\mathcal{E}_j$ holds, $\tilde{\mathcal{I}}_{j'} \subset \mathcal{I}_{j'}$ for each $j' \in \{\lfloor j/2 \rfloor, \ldots, j\}$. For $j' = \lfloor j/2 \rfloor$, recall $1 \in S_{\lfloor j/2 \rfloor}^{(i^\star)}$ by Assumption 3, so $\mathcal{I}_{\lfloor j/2 \rfloor} \supset \{i^\star\} = \tilde{\mathcal{I}}_{\lfloor j/2 \rfloor}$. Now assume $\tilde{\mathcal{I}}_{j'-1} \subset \mathcal{I}_{j'-1}$ for some $j' \in \{\lfloor j/2 \rfloor + 1, \ldots, j\}$. Let $i \in \tilde{\mathcal{I}}_{j'}$; we aim to show that $i \in \mathcal{I}_{j'}$, i.e., that $1 \in S_{j'}^{(i)}$. By (61), we have two cases to consider:

- $i \in \tilde{\mathcal{I}}_{j'-1}$: By the inductive hypothesis, $i \in \mathcal{I}_{j'-1}$ as well, so $1 \in S_{j'-1}^{(i)}$ by definition. Combined with $j' - 1 \geq \lfloor j/2 \rfloor$ and Claim 17 (recall $j \geq J_4^\star$ and $\lfloor j/2 \rfloor \geq \lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq \tau_{arm}$ when $\mathcal{E}_j$ holds, so the claim applies), this implies $B_{j'-1}^{(i)} = 1$, so $1 \in S_{j'}^{(i)}$ by the algorithm.

- $i \in [n] \setminus \tilde{I}_{j'-1}, \bar{Y}_{j'}^{(i)} = 1, \bar{H}_{j'}^{(i)} \in \tilde{I}_{j'-1}$: First observe that since $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_3^\star$ and $j' \geq \lfloor j/2 \rfloor \geq J_2^\star$, we can apply Lemma 11 on the event $\mathcal{E}_j$ (with j replaced by $\lfloor j/2 \rfloor$ in that lemma) to obtain $P_{j'}^{(i)} \cap [n] = \emptyset$. On the other hand, recall that $\bar{Y}_{j'}^{(i)}$ is Bernoulli(Υ) in Definition 1 and $v_{j'}^{(i)}$ is Uniform[0, 1] for the Section 7.4 sampling, so we can realize the former as $\bar{Y}_{j'}^{(i)} = \mathbb{1}(v_{j'}^{(i)} \leq \Upsilon)$. Hence, by assumption $\bar{Y}_{j'}^{(i)} = 1$ and definition $\Upsilon = \min_{i \in [n]} d_{\mathrm{hon}}(i)/d(i)$, we obtain

$$1 = \bar{Y}_{j'}^{(i)} = \mathbb{1}(v_{j'}^{(i)} \leq \Upsilon) \leq \mathbb{1}\left(v_{j'}^{(i)} \leq \frac{d_{\mathrm{hon}}(i)}{d(i)}\right) \leq \mathbb{1}\left(v_{j'}^{(i)} \leq \frac{d_{\mathrm{hon}}(i)}{|N(i) \setminus P_j^{(i)}|}\right) = Y_j^{(i)}.$$

In summary, we have shown that $P_{j'}^{(i)} \cap [n] = \emptyset$ and $Y_j^{(i)} = 1$. Hence, by the Section 7.4 sampling, we conclude $H_j^{(i)} = \bar{H}_j^{(i)}$. Let $i' = H_j^{(i)}$ denote this honest agent. Then by the inductive hypothesis, we know that $i' \in \tilde{I}_{j'-1} \subset I_{j'-1}$, i.e., that $1 \in S_{j'-1}^{(i')}$. By Claim 17, this implies $B_{j'-1}^{(i')} = 1$, so by the algorithm, $1 = B_{j'-1}^{(i')} = R_{j'-1}^{(i)} \in S_{j'}^{(i)}$.

Finally, for the equality in (62), note that $\tilde{I}_{j'}$ is independent of the randomness before $\lfloor j/2 \rfloor$. By $\tilde{I}_{\lfloor j/2 \rfloor} = \bar{I}_0 = \{i^\star\}$, the fact that $\bar{Y}_{j'}^{(i)}$ and $\bar{Y}_{j'-\lfloor j/2 \rfloor}^{(i)}$ are both Bernoulli(Υ) random variables and $\bar{H}_{j'}^{(i)}$ and $\bar{H}_{j'-\lfloor j/2 \rfloor}^{(i)}$ are both sampled uniformly from $N_{\mathrm{hon}}(i)$, $\tilde{I}_{j'}$ has the same distribution as $\bar{I}_{j'-\lfloor j/2 \rfloor}$. Also, note that $\tilde{I}_{j'}$ is independent of the randomness before $\lfloor j/2 \rfloor$. Finally, by definition of $\bar{\tau}_{\mathrm{spr}}$, $\bar{I}_{j-\lfloor j/2 \rfloor} \neq [n]$ implies that $\bar{\tau}_{\mathrm{spr}} > j - \lfloor j/2 \rfloor \geq j/2$. These observations successively imply

$$\mathbb{P}_{\lfloor j/2 \rfloor}(\tilde{I}_j \neq [n]) = \mathbb{P}(\tilde{I}_j \neq [n]) = \mathbb{P}(\bar{I}_{j'-\lfloor j/2 \rfloor} \neq [n]) \leq \mathbb{P}(\bar{\tau}_{\mathrm{spr}} > j/2). \qquad \square$$

E.5 Details from Section 7.5

Combining the lemmas of the above sub-appendices, we can bound the tail of the spreading time.

THEOREM 3. *Under the assumptions of Theorem 2, for any $j \geq J_\star$, where $J_\star$ is defined in Claim 32 from Appendix F.3, we have*

$$\mathbb{P}(\tau_{spr} > j) \leq \left(\frac{84^{\rho_1^2(\beta(2\alpha-3)-1)}(6^\beta + 2)nK^2 S}{(2\alpha - 3)(\beta(2\alpha - 3) - 1)} + 3\right) j^{\rho_1^2(\beta(3-2\alpha)+1)} + \mathbb{P}\left(\bar{\tau}_{spr} > \frac{j}{2}\right).$$

PROOF. For any $j \geq J_\star$, we can use Claim 32 to obtain

$$j \geq J_4^\star, \quad \lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_2^\star \vee J_3^\star, \quad \lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq J_1^\star \vee (2 + j^{\rho_1^2}/84) \geq 3 \qquad (65)$$

$$(n + m)^3 \exp(-\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor/(3\bar{d})) \leq j^{\rho_1^2(\beta(3-2\alpha)+1)}. \qquad (66)$$

In particular, $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq J_1^\star \vee 3$ implies that we can use Lemma 9 with j replaced by $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor$. Combined with (65) (namely, $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor - 2 \geq j^{\rho_1^2}/84$; this yields

$$\mathbb{P}(\tau_{\mathrm{arm}} > \lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor) \leq \frac{84^{\rho_1^2(\beta(2\alpha-3)-1)}(6^\beta + 2)nK^2 S}{(2\alpha - 3)(\beta(2\alpha - 3) - 1)} j^{\rho_1^2(\beta(3-2\alpha)+1)}.$$

Since $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_2^\star$ by (65), we can use Lemma 10 (with j replaced by $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor$) and (66) to obtain

$$\mathbb{P}(\tau_{\mathrm{com}} > \lfloor \theta_{\lfloor j/2 \rfloor} \rfloor) \leq 3(n + m)^2 \exp\left(-\frac{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor}{3\bar{d}}\right) \leq 3 j^{\rho_1^2(\beta(3-2\alpha)+1)}.$$

Furthermore, using the bounds $j \geq J_4^\star$, $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_3^\star$, and $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_2^\star$ from (65) (the last of which implies $\lfloor j/2 \rfloor \geq J_2^\star$, since $\theta_{\lfloor j/2 \rfloor} = (\lfloor j/2 \rfloor/3)^{\rho_1} \leq \lfloor j/2 \rfloor$ by $\rho_1 \in (0, 1)$), we can use Lemma 12 to get

$$\mathbb{P}(\tau_{\mathrm{arm}} \leq \lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor, \tau_{\mathrm{com}} \leq \lfloor \theta_{\lfloor j/2 \rfloor} \rfloor, \tau_{\mathrm{spr}} > j) \leq \mathbb{P}(\bar{\tau}_{\mathrm{spr}} > j/2).$$

Finally, by the union bound, we have

$$\mathbb{P}(\tau_{\text{spr}} > j) \leq \mathbb{P}(\tau_{\text{arm}} > \lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor) + \mathbb{P}(\tau_{\text{com}} > \lfloor \theta_{\lfloor j/2 \rfloor} \rfloor)$$
$$+ \mathbb{P}(\tau_{\text{arm}} \leq \lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor, \tau_{\text{com}} \leq \lfloor \theta_{\lfloor j/2 \rfloor} \rfloor, \tau_{\text{spr}} > j),$$

so combining the previous four inequalities yields the desired result. □

Finally, as a corollary, we can bound $\mathbb{E}[A_{\tau_{\text{spr}}}]$.

COROLLARY 4. *Under the assumptions of Theorem 2, we have*

$$\mathbb{E}[A_{\tau_{spr}}] \leq (J_\star)^\beta + \left(\frac{84^{\rho_1^2(\beta(2\alpha-3)-1)}(6^\beta + 2)}{(2\alpha-3)(\beta(2\alpha-3)-1)} + 3 \right) \frac{2\beta n K^2 S}{\rho_1^2(\beta(2\alpha-3)-1)-\beta} + \mathbb{E}[A_{2\bar{\tau}_{spr}}]$$
$$= O\left(S^{\beta/(\rho_1^2(\beta-1))} \vee (S\log(S/\Delta_2)/\Delta_2^2)^{\beta/(\beta-1)} \vee (\bar{d}\log(n+m))^{\beta/\rho_1} \vee nK^2 S \right) + \mathbb{E}[A_{2\bar{\tau}_{spr}}],$$

where $J_\star$ is defined as in Claim 32 from Appendix F.3.

PROOF. We first observe

$$\mathbb{E}[A_{\tau_{\text{spr}}}] = \sum_{j=1}^{\infty} (A_j - A_{j-1})\mathbb{P}(\tau_{\text{spr}} \geq j) \leq A_{J_\star - 1} + \sum_{j=J_\star}^{\infty} (A_j - A_{j-1})\mathbb{P}(\tau_{\text{spr}} \geq j). \tag{67}$$

For the first term in (67), using Claim 18 from Appendix F.1, we compute

$$A_{J_\star - 1} = A_{J_\star} - (A_{J_\star} - A_{J_\star - 1}) = \lceil (J_\star)^\beta \rceil - (A_{J_\star} - A_{J_\star - 1}) \leq (J_\star)^\beta + 1 - 1 = (J_\star)^\beta.$$

For the second term in (67), define the constant

$$C = \frac{84^{\rho_1^2(\beta(2\alpha-3)-1)}(6^\beta + 2)}{(2\alpha-3)(\beta(2\alpha-3)-1)} + 3.$$

Then by Theorem 3, we have

$$\sum_{j=J_\star}^{\infty} (A_j - A_{j-1})\mathbb{P}(\tau_{\text{spr}} \geq j) \leq \sum_{j=J_\star}^{\infty} (A_j - A_{j-1})\mathbb{P}(\bar{\tau}_{\text{spr}} > j/2) \tag{68}$$

$$+ CnK^2 S \sum_{j=J_\star}^{\infty} (A_j - A_{j-1}) j^{\rho_1^2(\beta(3-2\alpha)+1)}. \tag{69}$$

For (68), we simply use nonnegativity to write

$$\sum_{j=J_\star}^{\infty} (A_j - A_{j-1})\mathbb{P}(\bar{\tau}_{\text{spr}} > j/2) \leq \sum_{j=1}^{\infty} (A_j - A_{j-1})\mathbb{P}(2\bar{\tau}_{\text{spr}} \geq j) = \mathbb{E}[A_{2\bar{\tau}_{\text{spr}}}].$$

For the second term in (69), we use Claim 18 and $\beta > 1$ and $\rho_1^2(\beta(2\alpha-3)-1) > \beta$ by the assumptions of Theorem 2 to write

$$CnK^2 S \sum_{j=J_\star}^{\infty} (A_j - A_{j-1}) j^{\rho_1^2(\beta(3-2\alpha)+1)} < 2\beta C n K^2 S \sum_{j=J_\star}^{\infty} j^{-1-(\rho_1^2(\beta(2\alpha-3)-1)-\beta)}$$

$$\leq 2\beta C n K^2 S \int_{j=1}^{\infty} j^{-1-(\rho_1^2(\beta(2\alpha-3)-1)-\beta)} dj = \frac{2\beta C n K^2 S}{\rho_1^2(\beta(2\alpha-3)-1)-\beta}.$$

Finally, combining the above bounds completes the proof. □

F   OTHER PROOFS

F.1   Basic inequalities

In this sub-appendix, we prove some simple inequalities used frequently in the analysis.

CLAIM 18. *For any $j \in \mathbb{N}$, we have*

$$j^\beta \le A_j \le j^{2\beta}, \quad \beta(j-1)^{\beta-1} - 1 < A_j - A_{j-1} < \beta j^{\beta-1} + 1, \quad A_j - A_{j-1} \ge 1,$$

and for any $z \ge 1$ and $l \in \mathbb{N}$, we have $A_{l\lceil z \rceil} \le e^{2\beta}(lz)^\beta$.

PROOF. For the first pair of inequalities, observe $A_j = j^\beta = j^{2\beta} = 1$ when $j = 1$, and for $j \ge 2$,

$$j^\beta \le A_j \le j^\beta + 1 \le 2j^\beta \le j^{1+\beta} \le j^{2\beta}.$$

For the second pair of inequalities, we first observe

$$A_j - A_{j-1} > j^\beta - (j-1)^\beta - 1 = \beta x^{\beta-1} - 1 \ge \beta(j-1)^{\beta-1} - 1, \tag{70}$$

where the equality holds for some $x \in [j-1, j]$ by the mean value theorem and the second inequality is $x \ge j - 1$. By analogous reasoning, one can also show $A_j - A_{j-1} < \beta j^{\beta-1} + 1$, so the second pair of inequalities holds. The third inequality holds with equality when $j = 1$, and for $j \ge 2$, the lower bound in (70) and $\beta > 1$ imply $A_j - A_{j-1} > 0$, so since A_j and A_{j-1} are integers, $A_j - A_{j-1} \ge 1$. Finally, using $z \ge 1$, $\beta > 1$, and $2 < e$, we can write

$$A_{l\lceil z \rceil} < (l(z+1))^\beta + 1 < (2lz)^\beta + (2lz)^\beta = 2^{\beta+1}(lz)^\beta < e^{2\beta}(zl)^\beta. \qquad \square$$

CLAIM 19. *For any $j \in \mathbb{N}$ and $c > 1$, $\sum_{i=j+1}^{\infty} i^{-c} \le j^{1-c}/(c-1)$.*

PROOF. Since $j \ge 1$ and $c > 1$, we can write

$$\sum_{i=j+1}^{\infty} i^{-c} = \sum_{i=j+1}^{\infty} \int_{x=i-1}^{i} i^{-c} dx \le \sum_{i=j+1}^{\infty} \int_{x=i-1}^{i} x^{-c} dx = \int_{x=j}^{\infty} x^{-c} dx = \frac{j^{1-c}}{c-1}. \qquad \square$$

CLAIM 20. *For any $x, y, z > 0$ such that $x^y \le z \log x$, $x < ((2z/y) \log(2z/y))^{1/y} < (2z/y)^{2/y}$.*

PROOF. Multiplying and dividing the right side of the assumed inequality by $y/2$, we obtain $x^y \le (2z/y) \log x^{y/2}$. We can then loosen this bound to get $x^y < (2z/y)x^{y/2}$, or $x < (2z/y)^{2/y}$. Plugging into the log term of the assumed inequality yields $x^y < (2z/y) \log(2z/y)$. Raising both sides to the power $1/y$ establishes the first bound. The second bound follows by using $\log(2z/y) < 2z/y$. $\square$

REMARK 14. *We typically apply Claim 20 with y constant but z not. It allows us to invert inequalities of the form $x^y \le z \log x$ to obtain $x = \tilde{O}(z^{1/y})$.*

F.2   Bandit inequalities

Next, we state and prove some basic bandit inequalities. The proof techniques are mostly modified from existing work (e.g., [3]), but we provide the bounds in forms useful for our setting.

CLAIM 21. *Suppose that $k_1, k_2 \in [K]$, $t \in \mathbb{N}$, $\ell, u > 0$, and $\iota \in (0, 1]$ satisfy*

$$\mu_{k_2} - \mu_{k_1} \ge \sqrt{\alpha \log t} \left(\frac{2}{\sqrt{\ell}} - \frac{1-\iota}{\sqrt{u}} \right).$$

Let $j \in \mathbb{N}$ be such that $t \in \{1 + A_{j-1}, \ldots, A_j\}$, i.e., $j \in A^{-1}(t)$. Then for any $i \in [n]$, we have

$$\mathbb{P}(T_{k_1}^{(i)}(t-1) \ge \ell, T_{k_2}^{(i)}(A_j) \le u, k_2 \in S_j^{(i)}, I_t^{(i)} = k_1) \le 2(\lfloor u \rfloor \wedge t)t^{1-2\alpha\iota^2}$$

Proof. For $k \in [K]$, let $\{X_k(s)\}_{s=1}^{\infty}$ be an i.i.d. sequence distributed as ν_k, and for $s \in \mathbb{N}$, let

$$\hat{\mu}_k^{(i)}(s) = \frac{1}{s} \sum_{s'=1}^{s} X_k(s'), \quad U_k^{(i)}(t, s) = \hat{\mu}_k^{(i)}(s) + \sqrt{\frac{\alpha \log t}{s}}$$

denote the empirical mean and UCB index of if i has pulled the k-th arm s times before t. Then Algorithm 1 implies that if $k_2 \in S_j^{(i)}$ and $I_t^{(i)} = k_1$, we must have

$$U_{k_1}^{(i)}(t, T_{k_1}^{(i)}(t-1)) \geq U_{k_2}^{(i)}(t, T_{k_2}^{(i)}(t-1)). \tag{71}$$

Next, note that if $T_{k_2}^{(i)}(A_j) \leq u$, then by monotonicity, $T_{k_2}^{(i)}(t-1) \leq u$ as well. Combined with the fact that $T_{k_2}^{(i)}(t-1) \in [t]$ by definition, we conclude that $T_{k_2}^{(i)}(A_j) \leq u$ implies $T_{k_2}^{(i)}(t-1) \leq \lfloor u \rfloor \wedge t$. Similarly, $T_{k_1}^{(i)}(t-1) \geq \ell$ implies $T_{k_1}^{(i)}(t-1) \geq \lceil \ell \rceil$ (since $T_{k_1}^{(i)}(t-1) \in \mathbb{N}$). Combined with (71), we obtain that if the event in the statement of the claim occurs, it must be the case that

$$\max_{s_1 \in \{\lceil \ell \rceil, \dots, t\}} U_{k_1}^{(i)}(t, s_1) \geq \min_{s_2 \in \lfloor \lfloor u \rfloor \wedge t \rfloor} U_{k_2}^{(i)}(t, s_2).$$

Therefore, by the union bound, we obtain

$$\mathbb{P}(T_{k_1}^{(i)}(t-1) \geq \ell, T_{k_2}^{(i)}(A_j) \leq u, k_2 \in S_j^{(i)}, I_t^{(i)} = k_1) \leq \sum_{s_1 = \lceil \ell \rceil}^{t} \sum_{s_2 = 1}^{\lfloor u \rfloor \wedge t} \mathbb{P}(U_{k_1}^{(i)}(t, s_1) \geq U_{k_2}^{(i)}(t, s_2)). \tag{72}$$

Now fix s_1 and s_2 as in the double summation. We claim $U_{k_1}^{(i)}(t, s_1) \geq U_{k_2}^{(i)}(t, s_2)$ implies

$$\hat{\mu}_{k_1}^{(i)}(s_1) \geq \mu_{k_1} + \sqrt{\alpha \log(t)/s_1} \quad \text{or} \quad \hat{\mu}_{k_2}^{(i)}(s_2) \leq \mu_{k_2} - \iota \sqrt{\alpha \log(t)/s_2}.$$

Indeed, if instead both inequalities fail, then by choice of s_1, s_2 and the assumption of the claim,

$$U_{k_1}^{(i)}(t, s_1) < \mu_{k_1} + 2\sqrt{\alpha \log(t)/s_1} \leq \mu_{k_1} + 2\sqrt{\alpha \log(t)/\ell}$$
$$\leq \mu_{k_2} + (1 - \iota)\sqrt{\alpha \log(t)/u} \leq \mu_{k_2} + (1 - \iota)\sqrt{\alpha \log(t)/s_2} < U_{k_2}^{(i)}(t, s_2),$$

which is a contradiction. Thus, by the union bound, Hoeffding's inequality, and $\iota \in (0, 1)$, we obtain

$$\mathbb{P}(U_{k_1}^{(i)}(t, s_1) \geq U_{k_2}^{(i)}(t, s_2)) \leq \mathbb{P}(\hat{\mu}_{k_1}^{(i)}(s_1) \geq \mu_{k_1} + \sqrt{\alpha \log(t)/s_1}) + \mathbb{P}(\hat{\mu}_{k_2}^{(i)}(s_2) \leq \mu_{k_2} - \iota \sqrt{\alpha \log(t)/s_2})$$
$$\leq e^{-2\alpha \log t} + e^{-2\alpha \iota^2 \log t} = t^{-2\alpha} + t^{-2\alpha \iota^2} \leq 2t^{-2\alpha \iota^2},$$

so plugging into (72) completes the proof. $\square$

Corollary 5. *Suppose that $k_1, k_2 \in [K]$, $t \in \mathbb{N}$, and $\ell > 0$ satisfy $\mu_{k_2} - \mu_{k_1} \geq \sqrt{4\alpha \log(t)/\ell}$. Let $j \in \mathbb{N}$ be such that $t \in \{1 + A_{j-1}, \dots, A_j\}$, i.e., $j = A^{-1}(t)$. Then for any $i \in [n]$, we have*

$$\mathbb{P}(T_{k_1}^{(i)}(t-1) \geq \ell, k_2 \in S_j^{(i)}, I_t^{(i)} = k_1) \leq 2t^{2(1-\alpha)}.$$

Proof. Using $T_{k_2}^{(i)}(A_j) \leq A_j$ by definition and applying Claim 21 with $u = A_j$ and $\iota = 1$,

$$\mathbb{P}(T_{k_1}^{(i)}(t-1) \geq \ell, k_2 \in S_j^{(i)}, I_t^{(i)} = k_1) = \mathbb{P}(T_{k_1}^{(i)}(t-1) \geq \ell, T_{k_2}^{(i)}(A_j) \leq A_j, k_2 \in S_j^{(i)}, I_t^{(i)} = k_1)$$
$$\leq 2(\lfloor A_j \rfloor \wedge t)t^{1-2\alpha} = t^{2(1-\alpha)}. \qquad \square$$

Corollary 6. *For any $i \in [n]$, $k \in [K]$, and $T \in \mathbb{N}$, we have*

$$\mathbb{E}\left[\sum_{t=A_{\tau_{spr}}+1}^{T} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) \geq \frac{4\alpha \log t}{\Delta_k^2} \right) \right] \leq \frac{4(\alpha - 1)}{2\alpha - 3}.$$

PROOF. First observe that since $1 \in S^{(i)}_{A^{-1}(t)}$ whenever $t \geq A_{\tau_{\text{spr}}} + 1$ by definition, we have

$$\sum_{t=A_{\tau_{\text{spr}}}+1}^{T} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) \geq \frac{4\alpha \log t}{\Delta_k^2}\right) \leq \sum_{t=1}^{\infty} \mathbb{1}\left(T_k^{(i)}(t-1) \geq \frac{4\alpha \log t}{\Delta_k^2}, 1 \in S^{(i)}_{A^{-1}(t)}, I_t^{(i)} = k\right) \quad (73)$$

Next, let $k_1 = k$, $k_2 = 1$, and $\ell = 4\alpha \log(t)/\Delta_k^2$. Then by definition, we have $\mu_{k_2} - \mu_{k_1} = \Delta_k = \sqrt{4\alpha \log(t)/\ell}$. Therefore, we can use Corollary 5 to obtain

$$\mathbb{P}(T_k^{(i)}(t-1) \geq 4\alpha \log(t)/\Delta_k^2, 1 \in S^{(i)}_{A^{-1}(t)}, I_t^{(i)} = k) \leq 2t^{2(1-\alpha)}. \quad (74)$$

Hence, taking expectation in (73), then plugging in (74) to the right side and using Claim 19, yields

$$\mathbb{E}\left[\sum_{t=A_{\tau_{\text{spr}}}+1}^{T} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) \geq \frac{4\alpha \log t}{\Delta_k^2}\right)\right] \leq 2\left(1 + \sum_{t=2}^{\infty} t^{2(1-\alpha)}\right) \leq \frac{4(\alpha-1)}{2\alpha-3}. \qquad \square$$

CLAIM 22. *For any $i \in [n]$, $t_1, t_2 \in \mathbb{N}$ such that $t_1 < t_2$, and $\{\ell_t\}_{t=t_1}^{t_2} \subset (0, \infty)$,*

$$\sum_{t=t_1}^{t_2} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) < \ell_t\right) \leq \max_{t \in \{t_1, \ldots, t_2\}} \ell_t.$$

PROOF. Set $\ell = \max_{t \in \{t_1, \ldots, t_2\}} \ell_t$. Then clearly

$$\sum_{t=t_1}^{t_2} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) < \ell_t\right) \leq \sum_{t=t_1}^{t_2} \mathbb{1}\left(I_t^{(i)} = k, T_k^{(i)}(t-1) < \ell\right).$$

Now suppose the right strictly exceeds ℓ. Then since the right side is an integer, we can find $\lceil \ell \rceil$ times $t \in \{t_1, \ldots, t_2\}$ such that $I_t^{(i)} = k$ and $T_k^{(i)}(t-1) < \ell$. Let $\bar{t}$ denote the largest such t. Because $I_t^{(i)} = k$ occurred at least $\lceil \ell \rceil - 1$ times before t, we know $T_k^{(i)}(\bar{t}-1) \geq \lceil \ell \rceil - 1$. But since $T_k^{(i)}(\bar{t}-1) < \ell$, this implies $\ell + 1 > \lceil \ell \rceil$, which is a contradiction. $\qquad \square$

Finally, we recall a well-known regret decomposition.

CLAIM 23. *The regret $R_T^{(i)}$ defined in (5) satisfies $R_T^{(i)} = \sum_{k=2}^{K} \Delta_k \mathbb{E}[\sum_{t=1}^{T} \mathbb{1}(I_t^{(i)} = k)]$.*

PROOF. See, e.g., the proof of [38, Lemma 4.5]. $\qquad \square$

F.3 Calculations for the early regret

In this sub-appendix, we assume α, β, η, θ_j, κ_j, ρ_1, and ρ_2 are chosen as in Theorem 2. Recall C_i, C_i', etc. denote constants associated with Claim i that only depend on α, β, η, ρ_1, and ρ_2.

CLAIM 24. *There exists $C_{24}, C_{24}' > 0$ such that $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq C_{24} j^{\rho_1}$ and $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq C_{24} j^{\rho_1^2} \; \forall j \geq C_{24}'.$*

PROOF. This follows from the choice $\theta_{j'} = (j'/3)^{\rho_1} \; \forall j' \in \mathbb{N}$ in Theorem 2. $\qquad \square$

CLAIM 25. *There exists $C_{25} > 0$ such that $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_2^\star \; \forall j \geq C_{25}.$*

PROOF. This follows from Claim 24 and the fact that $J_2^\star$ is a constant by definition (53). $\qquad \square$

CLAIM 26. *There exists $C_{26}, C_{26}' > 0$ such that for any $j \geq C_{26} S^{1/(\beta-1)}$,*

$$\left(\frac{A_j - A_{j-1}}{S+2} - 1\right) \vee 1 = \frac{A_j - A_{j-1}}{S+2} - 1 \geq \frac{j^{\beta-1}}{C_{26}' S}.$$

Proof. By Claim 18, we can find $C, C' > 0$ depending only on β such that $A_j - A_{j-1} \geq Cj^{\beta-1}$ whenever $j \geq C'$. Hence, for any $j \geq (6S/C)^{1/(\beta-1)} \vee C'$, we know $A_j - A_{j-1} \geq Cj^{\beta-1} \geq 6S$, so

$$\frac{A_j - A_{j-1}}{S+2} - 1 \geq \frac{Cj^{\beta-1} - 3S}{3S} \geq \frac{Cj^{\beta-1} - Cj^{\beta-1}/2}{3S} = \frac{Cj^{\beta-1}}{6S} \geq 1,$$

where we also used $S \geq 1$. The claim follows if we set $C_{26} = (6/C)^{1/(\beta-1)} \vee C'$ and $C'_{26} = C/6$. □

CLAIM 27. *There exists $C_{27}, C'_{27} > 0$ such that for any $j \geq C_{27}S^{1/(\beta-1)}$,*

$$\log(A_{j-1} \vee 1) \geq \beta \log(j)/2 > 0, \quad \frac{1-\psi}{\sqrt{\kappa_j}} - \frac{2}{\sqrt{(\frac{A_j - A_{j-1}}{S+2} - 1) \vee 1}} > \sqrt{\frac{C'_{27}K^2S}{j^{\rho_2}}}. \tag{75}$$

Proof. By Claim 26, we can find constants $C_{26}, C'_{26} > 0$ such that for any $j \geq C_{26}S^{1/(\beta-1)}$,

$$2/\sqrt{([(A_j - A_{j-1})/(S+2)] - 1) \vee 1} \leq \sqrt{4C'_{26}S/j^{\beta-1}} \leq \sqrt{4C'_{26}K^2S/j^{\beta-1}},$$

where we also used $K \geq 1$. Furthermore, since $\rho_2 < \beta - 1$ by assumption, we can find $C > 0$ depending only on C'_{26}, ψ, β, and ρ_2, such that for any $j \geq C$, we have $4C'_{26}/j^{\beta-1} < (1-\psi)^2/(4j^{\rho_2})$. Combined with the previous inequality and the choice $\kappa_j = j^{\rho_2}/(K^2S)$ in Theorem 2, we obtain

$$\frac{1-\psi}{\sqrt{\kappa_j}} - \frac{2}{\sqrt{(\frac{A_j - A_{j-1}}{S+2} - 1) \vee 1}} > \sqrt{\frac{(1-\psi)^2K^2S}{4j^{\rho_2}}} \ \forall \, j \geq (C_{26} \vee C)S^{1/(\beta-1)}.$$

Hence, if we set $C_{27} = C_{26} \vee C \vee 4$ and $C'_{27} = (1-\psi)^2/4$, the second inequality in (75) holds for $j \geq C_{27}S^{1/(\beta-1)}$. Finally, define $h(j) = j - 1 - \sqrt{j} \ \forall \, j \in \mathbb{N}$. Then $h(4) = 1$ and $h'(j) = 1 - 1/(2\sqrt{j}) > 0 \ \forall \, j \geq 4$, so $h(j) \geq 0 \ \forall \, j \geq 4$. Thus, for any $j \geq C_{27}S^{1/(\beta-1)} \geq 4$, we know $j - 1 \geq \sqrt{j}$, so by Claim 18, $\log(A_{j-1}) \geq \log((j-1)^\beta) \geq \log(\sqrt{j}^\beta) = \beta \log(j)/2$, i.e., the first inequality in (75) holds. □

CLAIM 28. *There exists $C_{28} > 0$ such that for any $j \geq C_{28}S^{1/(\rho_1^2(\beta-1))}$, $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq J_1^\star$.*

Proof. By Claims 26 and 27, we can set $C = C_{26} \vee C_{27}$ to ensure that for $j \geq CS^{1/(\beta-1)}$, $A_j - A_{j-1} \geq 2(S+2)$ and $\delta_{j,2} > 0$. Hence, $J_1^\star \leq CS^{1/(\beta-1)}$ by definition (47). On the other hand, by Claim 24, we know $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq C_{24}j^{\rho_1^2}$ for $j \geq C'_{24}$. Thus, if we set $C_{28} = (C/C_{24})^{1/\rho_1^2} \vee C'_{24}$, then for any $j \geq C_{28}S^{1/(\rho_1^2(\beta-1))}$, we obtain $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq C_{24}j^{\rho_1^2}$ (since $j \geq C'_{24}$) and $C_{24}j^{\rho_1^2} \geq CS^{1/(\beta-1)}$ (since $j \geq (C/C_{24})^{1/\rho_1^2}S^{1/(\rho_1^2(\beta-1))} = ((C/C_{24})S^{1/(\beta-1)})^{1/\rho_1^2}$), which implies $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq CS^{1/(\beta-1)} \geq J_1^\star$. □

CLAIM 29. *There exists $C_{29} > 0$ such that for any $j \geq C_{29}S^{1/(\rho_1^2(\beta-1))}$, $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_3^\star$.*

Proof. We first upper bound $J_3^\star$. By Claim 24, we can find $C_{24}, C'_{24} > 0$ such that $\lfloor \theta_j \rfloor \geq C_{24}j^{\rho_1}$ when $j \geq C'_{24}$. Let $j \geq C'_{24} \vee ((C_{26}/C_{24})S^{1/(\beta-1)})^{1/\rho_1}$, where C_{26} is from Claim 26, and $j' \in \{\lfloor \theta_j \rfloor, \ldots, j\}$. Then $j' \geq \lfloor \theta_j \rfloor \geq C_{24}j^{\rho_1} \geq C_{26}S^{1/(\beta-1)}$, so we can find $C'_{26} > 0$ such that

$$(K+1)\delta_{j',1} \leq (K+1)\sqrt{\frac{4\alpha \log(A_{j'})C'_{26}S}{(j')^{\beta-1}}} \leq \sqrt{\frac{(2K)^2 4\alpha \log(j^{2\beta})C'_{26}}{(C_{24}j^{\rho_1})^{\beta-1}}} = \sqrt{\frac{CK^2S\log j}{j^{\rho_1(\beta-1)}}},$$

where the first inequality uses Claim 26; the second inequality uses $K \geq 1$, $j' \geq \lfloor \theta_j \rfloor \geq C_{24}j^{\rho_1}$, and Claim 18; and the equality defines $C = 32\alpha\beta C'_{26}/C_{24}^{\beta-1}$. On the other hand, if $j \geq C_{27}S^{1/(\beta-1)}$, where C_{27} is from Claim 27, we can find $C'_{27} > 0$ such that

$$\delta_{j,2} > \sqrt{\alpha\beta \log(j)C'_{27}K^2S/(2j^{\rho_2})} = \sqrt{C'K^2S\log(j)/j^{\rho_2}},$$

where the inequality uses Claim 27 and the equality defines $C' = \alpha\beta C'_{27}/2$. Finally, by assumption $\rho_2 < \rho_1(\beta - 1)$, we can find $C'' > 0$ such that, for any $j \geq C''$, we have $C'/j^{\rho_2} \geq C/j^{\rho_1(\beta-1)}$. Combined with the previous two inequalities, we obtain that for $j \geq C'_{24} \vee ((C_{26}/C_{24})S^{1/(\beta-1)})^{1/\rho_1} \vee (C_{27}S^{1/(\beta-1)}) \vee C''$ and any $j' \in \{\lfloor\theta_j\rfloor, \ldots, j\}$, $\delta_{j,2} > (K+1)\delta_{j',1}$. Therefore, by definition of $J_3^\star$ (57), we conclude that $J_3^\star \leq \tilde{C}S^{1/(\rho_1(\beta-1))}$, where $\tilde{C} = C'_{24} \vee (C_{26}/C_{24}) \vee C_{27} \vee C''$. Therefore, if we set $C_{29} = C'_{24} \vee (\tilde{C}/C_{24})^{1/\rho_1}$, we obtain that for any $j \geq C_{29}S^{1/(\rho_1^2(\beta-1))}$, $\lfloor\theta_{\lfloor j/2\rfloor}\rfloor \geq C_{24}j^{\rho_1}$ (since $j \geq C'_{24}$) and $C_9 j^{\rho_1} \geq \tilde{C}S^{1/(\rho_1(\beta-1))}$ (since $j \geq (\tilde{C}/C_{24})^{1/\rho_1}S^{1/(\rho_1^2(\beta-1))} = (\tilde{C}S^{1/(\rho_1(\beta-1))}/C_{24})^{1/\rho_1}$), so stringing together the inequalities, we conclude $\lfloor\theta_{\lfloor j/2\rfloor}\rfloor \geq C_{24}j^{\rho_1} \geq \tilde{C}S^{1/(\rho_1(\beta-1))} \geq J_3^\star$. $\qquad\square$

CLAIM 30. *There exists $C_{30} > 0$ such that $J_4^\star \leq C_{30}(S\log(C_{30}S/\Delta_2^2)/\Delta_2^2)^{1/(\beta-1)}$.*

PROOF. Let $\tilde{C}_{30} = 16C'_{26}\alpha\beta/(\beta-1)$ and $C_{30} = \tilde{C}_{30} \vee (3\tilde{C}_{30}^{1/(\beta-1)}) \vee (3C_{26}) \vee 16$, where C_{26} and C'_{26} are the constants from Claim 26. Also define $J_4^\dagger = C_{30}(S\log(C_{30}S/\Delta_2^2)/\Delta_2^2)^{1/(\beta-1)}$. Then by definition of $J_4^\star$ (60), it suffices to show $\delta_{j,1} < \Delta_2 \, \forall \, j \geq \lfloor\lfloor J_4^\dagger\rfloor/2\rfloor$. Fix such a j and suppose instead that $\delta_{j,1} \geq \Delta_2$. Since $C_{30} \geq 16$, we know $J_4^\dagger \geq C_{30}(\log C_{30})^{1/(\beta-1)} \geq 16$, so

$$j \geq (\lfloor J_4^\dagger\rfloor/2) - 1 \geq ((J_4^\dagger - 1)/2) - 1 = (J_4^\dagger/2) - (3/2) > J_4^\dagger/3. \tag{76}$$

Hence, because $C_{30} \geq 3C_{26}$, we have $j \geq J_4^\dagger/3 \geq C_{26}S^{1/(\beta-1)}$, so by Claims 26 and 18, respectively,

$$\delta_{j,1}^2 \leq 4\alpha\log(A_j)C'_{26}S/j^{\beta-1} \leq 4\alpha\log(j^{2\beta})C'_{26}S/j^{\beta-1} = 8C'_{26}\alpha\beta S\log(j)/j^{\beta-1}.$$

Rearranging and using the assumption $\delta_{j,1} \geq \Delta_2$, this implies $j^{\beta-1} \leq 8C'_{26}\alpha\beta S\log(j)/\Delta_2^2$. Hence, applying Claim 20 with $x = j$, $y = \beta - 1$, and $z = 8C'_{26}\alpha\beta S/\Delta_2^2$, we obtain

$$j \leq (16C'_{26}\alpha\beta S\log(16C'_{26}\alpha\beta S/((\beta-1)\Delta_2^2))/((\beta-1)\Delta_2^2))^{1/(\beta-1)} = (\tilde{C}_{30}S\log(\tilde{C}_{30}S/\Delta_2^2)/\Delta_2^2)^{1/(\beta-1)}.$$

But since $C_{30} \geq \tilde{C}_{30} \vee (3\tilde{C}_{30}^{1/(\beta-1)})$, we have shown $j \leq J_4^\dagger/3$, which contradicts (76). $\qquad\square$

CLAIM 31. *There exists $C_{31} > 0$ such that, for any $j \geq (C_{31}\bar{d}\log(C_{31}\bar{d}(n+m)))^{1/\rho_1}$,*

$$(n+m)^3\exp(-\lfloor\theta_{\lfloor j/2\rfloor}\rfloor/(3\bar{d})) \leq j^{\rho_1^2(\beta(3-2\alpha)+1)}. \tag{77}$$

PROOF. Let $\tilde{C}_{31} = 12\rho_1(\beta(2\alpha-3)-1)/C_{24}$ and $C_{31} = C'_{24} \vee 3 \vee (18/C_{24}) \vee \tilde{C}_{31}$, and suppose instead that (77) fails for some $j \geq (C_{31}\bar{d}\log(C_{31}\bar{d}(n+m)))^{1/\rho_1}$. Then we can write

$$\lfloor\theta_{\lfloor j/2\rfloor}\rfloor < 9\bar{d}\log(n+m) + 3\bar{d}\rho_1^2(\beta(2\alpha-3)-1)\log j. \tag{78}$$

Since $C_{31} \geq C'_{24} \vee 3$ and $\rho_1 \in (0,1)$, we know that $j \geq (C'_{24}\log(3))^{1/\rho_1} > C'_{24}$, so $\lfloor\theta_{\lfloor j/2\rfloor}\rfloor \geq C_{24}j^{\rho_1}$ by Claim 24. Since $C_{31} \geq (18/C_{24}) \vee 1$ and $\bar{d} \geq 1$, we also have $j \geq ((18/C_{24})\log(n+m))^{1/\rho_1}$, or $C_{24}j^{\rho_1}/2 \geq 9\bar{d}\log(n+m)$. Combining these two bounds with (78), we conclude

$$j^{\rho_1} < (6\rho_1^2(\beta(2\alpha-3)-1)/C_{24})\bar{d}\log j = \tilde{C}_{31}(\rho_1/2)\bar{d}\log j.$$

Applying Claim 20 with $x = j$, $y = \rho_1$, and $z = \tilde{C}_{31}(\rho_1/2)\bar{d}$, we obtain $j < (\tilde{C}_{31}\bar{d}\log(\tilde{C}_{31}\bar{d}))^{1/\rho_1}$. But since $C_{31} \geq \tilde{C}_{31}$, this contradicts the assumed lower bound on j. $\qquad\square$

CLAIM 32. *Define $C_{32} = C_{25} \vee C_{28} \vee C_{29}$ and*

$$J_\star = (C_{32}S^{1/(\rho_1^2(\beta-1))}) \vee (C_{30}(S\log(C_{30}S/\Delta_2^2)/\Delta_2^2)^{1/(\beta-1)}) \vee (C_{31}\bar{d}\log(C'_{31}(n+m)))^{1/\rho_1}.$$
$$= \Theta\left(S^{1/(\rho_1^2(\beta-1))} \vee (S\log(S/\Delta_2)/\Delta_2^2)^{1/(\beta-1)} \vee (\bar{d}\log(n+m))^{1/\rho_1}\right).$$

Then for any $j \in \mathbb{N}$ such that $j \geq J_\star$, we have

$$j \geq J_4^\star, \quad \lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_2^\star \vee J_3^\star, \quad \lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq J_1^\star \vee (2 + j^{\rho_1^2}/84) \geq 3$$

$$(n+m)^3 \exp(-\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor / (3\bar{d})) \leq j^{\rho_1^2(\beta(3-2\alpha)+1)}.$$

PROOF. The first bound holds by $j \geq C_{30}((S/\Delta_2^2) \log(C'_{30} S/\Delta_2^2))^{1/(\beta-1)}$ and Claim 30. The second holds since $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_2^\star$ by $j \geq C_{25}$ and Claim 25, and since $\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor \geq J_3^\star$ by $j \geq C_{29} S^{1/(\rho_1^2(\beta-1))}$ and Claim 29. The third holds since $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq J_1^\star$ by $j \geq C_{28} S^{1/(\rho_1^2(\beta-1))}$ and Claim 28, and since $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \geq (2 + j^{\rho_1^2}/48)$ for large enough C'_{24}. The fourth holds since $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor > 2$ (by the third) and $\lfloor \theta_{\lfloor \theta_{\lfloor j/2 \rfloor} \rfloor} \rfloor \in \mathbb{N}$. The fifth holds by $j \geq (C_{31} \bar{d} \log(C'_{31}(n+m)))^{1/\rho_1}$ and Claim 31. □

F.4 Calculations for the later regret

In this sub-appendix, we assume $\alpha, \beta, \eta, \theta_j, \kappa_j, \rho_1,$ and ρ_2 are chosen as in Theorem 2. Recall $C_i, C'_i,$ etc. denote constants associated with Claim i that only depend on $\alpha, \beta, \eta, \rho_1,$ and ρ_2.

CLAIM 33. *There exists $C_{33} > 0$ such that, for any $\gamma_i \in (0, 1)$,*

$$\lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor^\beta \leq 4\alpha K \log(T)/\Delta_2 \quad \Rightarrow \quad \log T \leq (C_{33}/\gamma_i) \log(C_{33} K/(\Delta_2 \gamma_i)). \tag{79}$$

PROOF. Similar to Claim 24 from Appendix F.3, we can find constants $C, C' > 0$ such that for any $j \geq C'$, $\lfloor \theta_j \rfloor \geq Cj^{\rho_1}$. If $\lceil T^{\gamma_i/\beta} \rceil < C'$, then $T^{\gamma_i} \leq \lceil T^{\gamma_i/\beta} \rceil^\beta < (C')^\beta$, so $\log T \leq (\beta/\gamma_i) \log(C')$, and the right side of (79) will hold for $C_{33} \geq \beta \vee C'$. If instead $\lceil T^{\gamma_i/\beta} \rceil \geq C'$, then $\lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor \geq C \lceil T^{\gamma_i/\beta} \rceil^{\rho_1} \geq CT^{\gamma_i \rho_1/\beta}$, so if the left side of (79) holds, we can write

$$C^\beta T^{\gamma_i \rho_1} = (CT^{\gamma_i \rho_1/\beta})^\beta \leq \lfloor \theta_{\lceil T^{\gamma_i/\beta} \rceil} \rfloor^\beta \leq 4\alpha \log(T)/\Delta_2.$$

Hence, applying Claim 20 with $x = T$, $y = \gamma_i \rho_1$, and $z = 4\alpha/(C^\beta \Delta_2)$, we obtain

$$\log T \leq \log(8\alpha/(\Delta_2 \gamma_i C^\beta \rho_1))^{2/(\gamma_i \rho_1)} \leq (C_{33}/\gamma_i) \log(C_{33} K/(\Delta_2 \gamma_i)),$$

where the last inequality holds for any $C_{33} \geq (8\alpha/(C^\beta \rho_1)) \vee (2/\rho_1)$. □

CLAIM 34. *There exists $C_{34} > 0$ such that, for any $\gamma_i \in (0, 1)$,*

$$\kappa_{\lceil T^{\gamma_i/\beta} \rceil} \leq 4\alpha K \log(T)/\Delta_2 \quad \Rightarrow \quad \log T \leq (C_{34}/\gamma_i) \log(C_{34} K/(\Delta_2 \gamma_i)).$$

PROOF. Recall $\kappa_j = j^{\rho_2}/(K^2 S)$ in Theorem 2. Hence, because $S \leq K$, we know that $\kappa_{\lceil T^{\gamma_i/\beta} \rceil} \geq T^{\gamma_i \rho_2/\beta}/K^3$. Rearranging and using the assumed bound, we obtain

$$T^{\gamma_i \rho_2/\beta} \leq K^3 \cdot \kappa_{\lceil T^{\gamma_i/\beta} \rceil} \leq K^3 \cdot 4\alpha K \log(T)/\Delta_2 = (4\alpha K^4/\Delta_2) \log T \leq (4\alpha K/\Delta_2)^4 \log T,$$

where the last inequality uses $\alpha \geq 1$ and $\Delta_2 \in (0, 1)$. Applying Claim 20 with $x = T$, $y = \gamma_i \rho_2/\beta$, and $z = (4\alpha K/\Delta_2)^4$, and noting that $2/y \leq (2/y)^4$ (since $\gamma_i \in (0, 1)$ and $\rho_2 \in (0, \beta - 1)$), we obtain

$$\log T \leq \log(8\alpha\beta K/(\Delta_2 \gamma_i \rho_2))^{8\beta/(\gamma_i \rho_2)} \leq (C_{34}/\gamma_i) \log(C_{34} K/(\Delta_2 \gamma_i)),$$

where the second inequality holds for $C_{34} \geq 8\alpha\beta/\rho_2$. □

CLAIM 35. *There exists $C_{35} > 0$ such that, for any $\gamma_i \in (0, 1)$,*

$$\exists j \geq \lceil T^{\gamma_i/\beta} \rceil \text{ s.t. } \lceil \kappa_j \rceil \leq 1 + 4\alpha \log(A_j)/\Delta_2^2 \quad \Rightarrow \quad \log T \leq (C_{35}/\gamma_i) \log(C_{35} K/(\Delta_2 \gamma_i)).$$

Proof. Fix $j \geq \lceil T^{\gamma_i/\beta} \rceil$ such that $\lceil \kappa_j \rceil \leq 1 + 4\alpha \log(A_j)/\Delta_2^2$. Note that if $j = 1$, then $1 \geq \lceil T^{\gamma_i/\beta} \rceil \geq T^{\gamma_i/\beta}$, so since $T \in \mathbb{N}$, we must have $T = 1$. This implies $\log T = 0$, so the claimed bound is immediate. Hence, we assume $j \geq 2$ moving forward. We first observe

$$j^{\rho_2} \leq K^3 \lceil \kappa_j \rceil \leq K^3 \left(1 + \frac{4\alpha \log A_j}{\Delta_2^2} \right) \leq K^3 \left(1 + \frac{8\alpha\beta \log j}{\Delta_2^2} \right) \leq K^3 \left(\frac{16\alpha\beta \log j}{\Delta_2^2} \right) = \frac{16\alpha\beta K^3 \log j}{\Delta_2^2}.$$

where the inequalities use $\lceil \kappa_j \rceil = \lceil j^{\rho_2}/(K^2 S) \rceil \geq j^{\rho_2}/K^3$, the assumed upper bound on $\lceil \kappa_j \rceil$, Claim 18, and $8\alpha\beta \log(j)/\Delta_2^2 \geq 1$ (since $\alpha, \beta \geq 1$, $j \geq 2$, and $\Delta_2 \in (0,1)$), respectively. Applying Claim 20 with $x = j$, $y = \rho_2$, and $z = 16\alpha\beta K^3/\Delta_2^2$, and noting that $2z/y \leq (32\alpha\beta K/(\Delta_2\rho_2))^3$ (since $\alpha \geq 1$, $\rho_2 \in (0, \beta - 1)$, and $\Delta_2 \in (0,1)$), we obtain $j \leq (32\alpha\beta K/(\Delta_2\rho_2))^{6/\rho_2} \leq (C_{35}K/(\Delta_2\gamma_i))^{6/\rho_2}$ for any $C_{35} \geq 32\alpha\beta/\rho_2$. Therefore, by assumption $j \geq \lceil T^{\gamma_i/\beta} \rceil$, we obtain that for any $C_{35} \geq 32\alpha\beta/\rho_2$,

$$T \leq \lceil T^{\gamma_i/\beta} \rceil^{\beta/\gamma_i} \leq j^{\beta/\gamma_i} \leq (C_{35}K/(\Delta_2\gamma_i))^{6\beta/(\rho_2\gamma_i)} \leq (C_{35}K/(\Delta_2\gamma_i))^{C_{35}/\gamma_i}. \qquad \square$$

CLAIM 36. *There exists $C_{36} > 0$ such that, for any $\gamma_i \in (0,1)$,*

$$\frac{\Delta_2(2\alpha - 3)}{8\alpha K^2} \leq \log(T) \sum_{j=\lceil T^{\gamma_i/\beta} \rceil}^{\infty} (\lceil \kappa_j \rceil - 1)^{3-2\alpha} \implies \log T \leq (C_{36}/\gamma_i) \log(C_{36}K/(\Delta_2\gamma_i)). \quad (80)$$

Proof. We first eliminate the corner case where $\min\{T^{\gamma_i/\beta}, \kappa_{\lceil T^{\gamma_i/\beta} \rceil}\} < 2$. In this case, one of $T^{\gamma_i/\beta} < 2$ and $\kappa_{\lceil T^{\gamma_i/\beta} \rceil} < 2$ must hold. If the former holds, then $\log T < (\beta/\gamma_i) \log 2$, and if the latter holds, then $2 > \kappa_{\lceil T^{\gamma_i/\beta} \rceil} = \lceil T^{\gamma_i/\beta} \rceil^{\rho_2}/(K^2 S) \geq T^{\gamma_i\rho_2/\beta}/K^3$, so $\log T \leq (\beta/(\gamma_i\rho_2)) \log(2K^3)$. In both cases, we can clearly find $C_{36} > 0$ satisfying the right side of (80).

Next, we assume $\kappa_{\lceil T^{\gamma_i/\beta} \rceil} \geq 2$ and $T^{\gamma_i/\beta} \geq 2$. By monotonicity, the former implies $\kappa_j \geq 2$ for any $j \geq \lceil T^{\gamma_i/\beta} \rceil$. For any such j, by definition and $S \leq K$, we can then write

$$\lceil \kappa_j \rceil - 1 \geq \kappa_j - 1 \geq \kappa_j/2 = j^{\rho_2}/(2K^2 S) \geq j^{\rho_2}/(2K^3).$$

Therefore, since $3 - 2\alpha < 0$ by assumption in Theorem 2, we obtain

$$\sum_{j=\lceil T^{\gamma_i/\beta} \rceil}^{\infty} (\lceil \kappa_j \rceil - 1)^{3-2\alpha} \leq 2^{2\alpha-3} K^{3(2\alpha-3)} \sum_{j=\lceil T^{\gamma_i/\beta} \rceil}^{\infty} j^{\rho_2(3-2\alpha)}.$$

For the summation at right, we use $T^{\gamma_i/\beta} \geq 2$ (which implies $\lceil T^{\gamma_i/\beta} \rceil - 1 \geq T^{\gamma_i/\beta} - 1 \geq T^{\gamma_i/\beta}/2$) and $\rho_2(2\alpha - 3) > 1$ by assumption in Theorem 2, along with Claim 19, to write

$$\sum_{j=\lceil T^{\gamma_i/\beta} \rceil}^{\infty} j^{\rho_2(3-2\alpha)} \leq \frac{(T^{\gamma_i/\beta}/2)^{1+\rho_2(3-2\alpha)}}{\rho_2(2\alpha - 3) - 1} = \frac{2^{\rho_2(2\alpha-3)} T^{\gamma_i(1+\rho_2(3-2\alpha))/\beta}}{2(\rho_2(2\alpha - 3) - 1)}.$$

Using $\rho_2(2\alpha - 3) > 1$ (by assumption), we also know

$$T^{\gamma_i(1+\rho_2(3-2\alpha))/\beta} \log(T) = \frac{2\beta T^{\gamma_i(1+\rho_2(3-2\alpha))/\beta} \log(T^{\gamma_i(\rho_2(2\alpha-3)-1)/(2\beta)})}{\gamma_i(\rho_2(2\alpha - 3) - 1)} \leq \frac{2\beta T^{\gamma_i(1+\rho_2(3-2\alpha))/(2\beta)}}{\gamma_i(\rho_2(2\alpha - 3) - 1)}.$$

Combining the previous three inequalities, we then obtain

$$\log(T) \sum_{j=\lceil T^{\gamma_i/\beta} \rceil}^{\infty} (\lceil \kappa_j \rceil - 1)^{3-2\alpha} \leq \frac{2^{(\rho_2+1)(2\alpha-3)}\beta}{(\rho_2(2\alpha - 3) - 1)^2} \frac{K^{3(2\alpha-3)}}{\gamma_i} T^{\gamma_i(1+\rho_2(3-2\alpha))/(2\beta)}.$$

Therefore, if the left side of (80) holds, we are guaranteed that

$$\frac{\Delta_2(2\alpha - 3)}{8\alpha K^2} \leq \frac{2^{(\rho_2+1)(2\alpha-3)}\beta}{(\rho_2(2\alpha - 3) - 1)^2} \frac{K^{3(2\alpha-3)}}{\gamma_i} T^{\gamma_i(1+\rho_2(3-2\alpha))/(2\beta)},$$

or, after rearranging, then using $\alpha > 1$ and $\Delta_2, \gamma_i \in (0, 1)$,

$$T^{\gamma_i(\rho_2(2\alpha-3)-1)/(2\beta)} \leq \frac{8 \cdot 2^{(\rho_2+1)(2\alpha-3)}\alpha\beta}{(2\alpha-3)(\rho_2(2\alpha-3)-1)^2} \frac{K^{6\alpha-7}}{\Delta_2\gamma_i} \leq (C_{36})^{6\alpha} \frac{K^{6\alpha-7}}{\Delta_2\gamma_i} \leq \left(\frac{C_{36}K}{\Delta_2\gamma_i}\right)^{6\alpha},$$

where the second inequality holds for large C_{36} and the third uses $\alpha \geq 1$ and $\Delta_2, \gamma_i \in (0, 1)$. Taking logarithms and choosing C_{36} appropriately in terms of ρ_2, α, and β yields the right side of (80). □