Limits on reliable information flows through stochastic populations

Lucas Boczkowski[1], Emanuele Natale[2], Ofer Feinerman[3], Amos Korman[1]

1 CNRS, IRIF, Université Paris Diderot, Paris, France
2 Algorithms and Complexity Department, Max-Planck-Institut für Informatik, Saarbrücken, Germany
3 Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel

Abstract

Biological systems can share and collectively process information to yield emergent effects, despite inherent noise in communication. While man-made systems often employ intricate structural solutions to overcome noise, the structure of many biological systems is more amorphous. It is not well understood how communication noise may affect the computational repertoire of such groups. To approach this question we consider the basic collective task of rumor spreading, in which information from few knowledgeable sources must reliably flow into the rest of the population. We study the effect of communication noise on the ability of groups that lack stable structures to efficiently solve this task. We present an impossibility result which strongly restricts reliable rumor spreading in such groups. Namely, we prove that, in the presence of even moderate levels of noise that affect all facets of the communication, no scheme can significantly outperform the trivial one in which agents have to wait until directly interacting with the sources - a process which requires linear time in the population size. Our results imply that in order to achieve efficient rumor spread a system must exhibit either some degree of structural stability or, alternatively, some facet of the communication which is immune to noise. We then corroborate this claim by providing new analyses of experimental data regarding recruitment in *Cataglyphis niger* desert ants. Finally, in light of our theoretical results, we discuss strategies to overcome noise in other biological systems.

Introduction

Systems composed of tiny mobile components must function under conditions of unreliability. In particular, any sharing of information is inevitably subject to communication noise. The effects of communication noise in distributed living systems appears to be highly variable. While some systems disseminate information efficiently and reliably despite communication noise [1–5], others generally refrain from acquiring social information, consequently losing all its potential benefits [6–8]. It is not well understood which characteristics of a distributed system are crucial in facilitating noise reduction strategies and, conversely, in which systems such strategies are bound to fail. Progress in this direction may be valuable towards better understanding the constraints that govern the evolution of cooperative biological systems.

Computation under noise has been extensively studied in the computer science community. These studies suggest that different forms of error correction (*e.g.,* redundancy) are highly useful in maintaining reliability despite noise [9–12]. All these, however, require the ability to transfer significant amount of information over stable communication channels. Similar redundancy methods may seem biologically plausible in systems that enjoy stable structures, such as brain tissues.

The impact of noise in stochastic systems with ephemeral connectivity patterns is far less understood. To study these, we focus on *rumor spreading* - a fundamental information dissemination task that is a prerequisite to almost any distributed system [13–16]. The literature on *rumor spreading* is quite vast and encompasses different disciplines over the last decades [17, 18]. For a succinct overview as for theoretical computer science, see Section *Related works in computer science* in the Supplementary Information [48].

A successful and efficient rumor spreading process is one in which a large group manages to quickly learn information initially held by one or a few informed individuals. Fast information flow to the whole group dictates that messages be relayed between individuals. Similar to the game of Chinese Whispers, this may potentially result in runaway buildup of noise and loss of any initial information [19]. It currently remains unclear what are the precise conditions that enable fast rumor spreading. On the one hand, recent works indicate that in some models of random noisy interactions, a collective coordinated process can in fact achieve fast information spreading [20, 21]. These models, however, are based on *push* operations that inherently include a certain reliable component (see more details in Section *Separation between $\mathcal{PUSH}$ and $\mathcal{PULL}$*). On the other hand, other works consider computation through noisy operations, and show that several distributed tasks require significant running time [22]. The tasks considered in these works (including the problem of learning the input bits of all processors, or computing the parity of all the inputs) were motivated by computer applications, and may be less relevant for biological contexts. Moreover, they appear to be more demanding than basic tasks, such as rumor spreading, and hence it is unclear how to relate bounds on the former problems to the latter ones.

In this paper we take a general stance to identify limitations under which reliable and fast rumor spreading cannot be achieved. Modeling a well-mixed population, we consider a passive communication scheme in which information flow occurs as one agent observes the cues displayed by another. If these interactions are perfectly reliable, the population could achieve extremely fast rumor spreading [16]. In contrast, here we focus on the situation in which messages are noisy. Informally, our main theoretical result states that fast rumor spreading through large populations can only be achieved if either

- the system exhibits some degree of structural stability, or
- some facet of the pairwise communication is immune to noise.

In fact, our lower bounds hold even when individuals are granted unlimited computational power and even when the system can take advantage of complete synchronization. In light of these theoretical results, we then turn to discuss several examples of information sharing in distributed biological systems. We provide new analyses of the efficiency of information dissemination during recruitment by desert ants. These suggest that this system lacks reliability in all its communication components, and its deficient performances qualitatively validate our predictions. Finally, we revisit existing rumor spreading solutions in large biological systems and discuss different strategies for confronting noise.

The problem

An intuitive description of the model follows. For more precise definitions, see, *The models* Section in the Supplementary Information [48].

Consider a population of n *agents*. Thought of as computing entities, assume that each agent has a discrete internal *state*, and can execute randomized algorithms - by internally flipping coins. In addition, each agent has an *opinion*, which we assume for simplicity to be binary, *i.e.,* either 0 or 1. A small number, s, of agents play the role of *sources*. Source agents are aware of their role and share the same opinion, referred to as the *correct opinion*. The goal of all agents is to have their opinion coincide with the correct opinion.

To achieve this goal, each agent continuously displays one of several *messages* taken from some finite alphabet Σ. Agents interact according to a random pattern, termed as the *parallel-$\mathcal{PULL}$* model: In each round $t \in \mathbb{N}^+$, each agent u observes the message currently displayed by another agent v, chosen independently and uniformly at random from all agents. Importantly, communication is noisy, hence the message observed by u may differ from that displayed by v. More precisely, for any $m, m' \in \Sigma$, let $P_{m,m'}$ be the probability that, any time some agent u observes an agent v holding some message $m \in \Sigma$, u actually receives message m'. The probabilities $P_{m,m'}$ define the entries of the noise-matrix P [21], which does not depend on time.

The noise is characterized by a *noise parameter* $\delta > 0$. Our model encapsulates a large family of noise distributions, making our bounds highly general.

Specifically, the noise distribution can take *any* form, as long as it satisfies the following criterion.

Definition 1 (Noise ellipticity parameter δ) *We say that the noise has ellipticity δ if $P_{m,m'} \geq \delta$ for any $m, m' \in \Sigma$.*

When messages are noiseless, it is easy to see that the number of rounds that are required to guarantee that all agents hold the correct opinion with high probability is $\mathcal{O}(\log n)$ [16]. In what follows, we aim to show that when the δ-uniform noise criterion is satisfied, the number of rounds required until even one non-source agent can be moderately certain about the value of the correct opinion is very large. Specifically, thinking of δ and s as constants independent of the population size n, this number of rounds is at least $\Omega(n)$.

To prove the lower bound, we will bestow the agents with capabilities that far surpass those that are reasonable for biological entities. These include:

- Unique identities: Agents have unique identities in the range $\{1, 2, \ldots n\}$. When observing agent v, its identity is received without noise.
- Complete knowledge of the system: Agents have access to all parameters of the system (including n, s, and δ) as well as to the full knowledge of the initial configuration except, of course, the correct opinion and the identity of the sources. In addition, agents have access to the results of random coin flips used internally by all other agents.
- Full synchronization: Agents know when the execution starts, and can count rounds.

We show that even given this extra computational power, fast convergence cannot be achieved. All the more so, fast convergence is impossible under more realistic assumptions.

Results

The purpose of this work is to identify limitations under which efficient rumor spreading would be impossible. Our main result is theoretical and, informally, states that when all components of communication are noisy fast rumor spreading through large populations is not feasible. In other words, our results imply that fast rumor spreading can only be achieved if the system either exhibits some degree of structural stability or that some facet of its communication is immune to noise. These results in hand, a next concern is how far our highly theoretical analysis can go in explaining actual biological systems.

Theoretical results with a high degree of generality may hold relevance to a wider range of biological systems. Lower bound and impossibility results follow this approach. Indeed, impossibility results from physics and information theory have previously been used to further the understanding of several biological systems [23, 24]. The results we present here are, similarly, in the form of lower

bounds but, this time, they are derived from the realm of distributed computation. As such, our theorems are general enough to constrain the performances of a vast class of computational systems regardless of their particulars or the specific computational algorithms which they apply. This generality stretches over to biology and can provide us with fundamental lessons regarding the limitations faced by distributed biological systems [24–26].

While the generality of our lower bound results makes them relevant to a large number of biological systems it also constitutes a weakness. Namely, the assumptions on which such theorems are based are not tailored to describe a particular system. This implies that comparisons between the model assumptions and the actual details of a specific system will not be perfect. Nevertheless, we show how our theoretical results can shed light on some non-trivial behaviors in a specific biological system whose characteristics are close enough to the underlying theoretical assumptions (see Section *Recruitment in desert ants*). Particularly, we empirically show that when desert ants communicate information regarding a new food source they are subject to limitations which are similar to those assumed by our model. We then demonstrate a non-trivial slowdown in the speed at which information spreads through the system as a function of group size. Despite the non-perfect matching between the theoretical assumption and the biological system, this non-trivial result stands in direct accordance with our theoretical lower bounds.

Distributed computing provides an effective means of studying biological groups [27–30]. However, to the best of our knowledge, there are no examples in which algorithmic lower bounds, one of distributed computing most powerful tools, have been applied to a particular living system. This work uses lower bounds to provide insights into non-trivial dynamics observed during ant recruitment behavior.

Theoretical results

In all the statements that follow we consider the parallel-$\mathcal{PULL}$ model satisfying the δ-uniform noise criterion, with $cs/n < \delta \leq 1/|\Sigma|$ for some sufficiently large constant c, where the upper bound follows from the criterion given in Definition 1. Hence, the previous lower bound on δ implies a restriction on the alphabet size, specifically, $|\Sigma| < n/(cs)$.

Theorem 1.1 *Any rumor spreading protocol cannot converge in less than* $\Omega(\frac{n\delta}{s^2(1-\delta|\Sigma|)^2})$ *rounds.*

Observe that the lower bound we present loses relevance when s is of order greater than $\sqrt{n}$, as our proof technique becomes uninformative in presence of a large number of sources (see *Remark 2* in the Supplementary Information [48]). message, which is, in turn, subject to noise. We also consider the case in which an agent can reliably identify a source when it observes one (that is, this information is not noisy). For this case, the following lower bound, which is weaker than the previous one but still polynomial, apply (see also the Supplementary Information [48], Detectable Sources Section):

Corollary 1.1 *Assume that sources are reliably detectable. There is no rumor spreading protocol that converges in less than* $\Omega((\frac{n\delta}{s^2(1-\delta|\Sigma|)^2})^{1/3})$ *rounds.*

Our results suggest that, in contrast to systems that enjoy stable connectivity, structureless systems are highly sensitive to communication noise (see Figure 1). More concretely, the two crucial assumptions that make our lower bounds work are: distinguishable signals only when necessary.

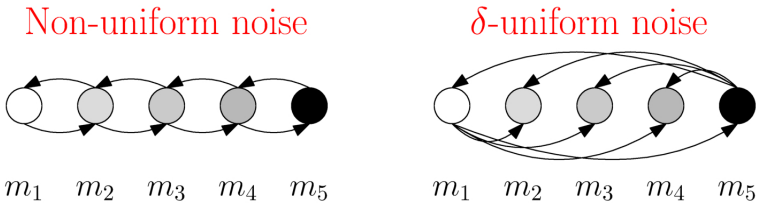


Figure 1: **Non-uniform noise vs. uniform noise.** On the left, we consider an example with non-uniform noise. Assume that the message vocabulary consists of 5 symbols, that is, $\Sigma = \{m_1, m_2, m_3, m_4, m_5\}$, where $m_1 = 0$ and $m_5 = 1$, represent the opinions. Assume that noise can occur only between consecutive messages. For example, m_2 can be observed as either m_2, m_3 or m_1, all with positive constant probability, but can never be viewed as m_4 or m_5. In this scenario, the population can quickly converge on the correct opinion by executing the following. The sources always display the correct opinion, *i.e.,* either m_1 or m_5, and each other agent displays m_3 unless it has seen either m_1 or m_5 in which case it adopts the opinion it saw and displays it. In other words, m_3 serves as a default message for non-source agents, and m_1 and m_5 serve as attracting sinks. It is easy to see that the correct opinion will propagate quickly through the system without disturbance, and within $\mathcal{O}(\log n)$ number of rounds, where n is the size of the population, all agents will hold it with high probability. In contrast, as depicted on the right picture, if every message can be observed as any other message with some constant positive probability (for clarity, some of the arrows have been omitted from the sketch), then convergence cannot be achieved in less than $\Omega(n)$ rounds, as Theorem 1.1 dictates.

Recruitment in desert ants

Our theoretical results assert that efficient rumor spreading in large groups could not be achieved without some degree of communication reliability. An example of a biological system whose communication reliability appears to be deficient in all of its components is recruitment in *Cataglyphis niger* desert ants. In this species, when a forager locates an oversized food item, she returns to the nest to recruit other ants to help in its retrieval [31, 32].

In our experimental setup, summarized in Figure 2, recruitment occurs in the small area of the nest's entrance chamber (S2a Fig). We find that within

this confined area, the interactions between ants follow a near uniformly random meeting pattern [33]. In other words, ants seem to have no control over which of their nest mates they will meet next (S2b Fig). This random meeting pattern approximates the first main assumption of our model. Another of the model's assumptions is that ants interact in parallel. This implies that the interaction rate per ant be constant and independent of group size. Indeed, the empirical rate of interaction during the recruitment process was measured to be 0.82 ± 0.07 (mean $\pm$ sem, $N = 44$) interactions per minute per ant and induces a small increase with group size: 0.62 ± 0.13 for two ants ($N = 8$) and 1 ± 0.2 for a group sizes of 9-10 ($N = 5$).

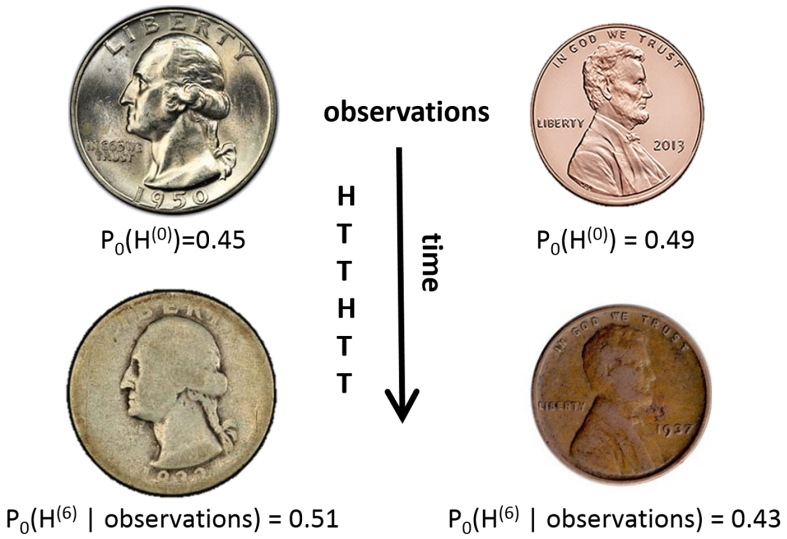


Figure 2: **Unreliable communication and slow recruitment by desert ant (*Cataglyphis niger*). a.** The experimental setup. The recruiter ant (circled) returns to the nest's entrance chamber (dark, 9cm diameter, disc) after finding the immobilized food item (arrow). Group size is ten. **b.** A *pdf* of the number of interactions that an ant experiences before meeting the same ant twice. The *pdf* is compared to uniform randomized interaction pattern. Data summarizes $N = 671$ interactions from seven experiments with a group size of 6 ants. **c.** Interactions of stationary ants with moving ants were classified into three different messages ('a' to 'c') depending on the ants' speed. The noise at which messages were confused with each other was estimated according to the response recipient, initially stationary, ants (see Materials and Methods). Gray scale indicates the estimated overlap between every two messages $\delta(i, j)$. Note $\delta = \min(\delta(i, j)) \approx 0.3$. Data collected over $N = 278$ interactions. **d.** The mean time it takes an ant that is informed about the food to recruit two nest-mates to exit the nest is presented for two group size ranges. Error bars represent standard error of the means over $N = 24$ experiments.

It has been shown that recruitment in *Cataglyphis niger* ants relies on rudimentary alerting interactions [34, 35] which are subject to high levels of noise [32]. Moreover, the information an ant passes in an interaction can be attributed solely to her speed before the interaction [32]. Binning ant speeds into three discrete messages and measuring the responses of stationary ants to these messages, we can estimate the probabilities of one message to be mistakenly perceived as another one (see Section *Methods*). We find that this communication is extremely noisy which complies with the uniform-noise assumption with a δ of approximately 0.3 (S2c Fig). While artificially dividing the continuous speed signals into a large number of discrete messages (thus creating a larger alphabet) would inevitably decrease δ, this is not supported by our empirical data (see Section *Methods*).

Finally, the interaction scheme, as exhibited by the ants, can be viewed somewhere in-between the noisy-push and the noisy-pull models. Moving ants tend to initiate more interaction [32] and this may resemble, at first glance, a noisy-push interaction scheme. However, the ants' interactions actually share characteristics with noisy-pull communication. Mainly, ants cannot reliably distinguish an ant that attempts to transmit information from any other non-communicating individual [32]. The fact that a receiver ant cannot be certain that a message was indeed communicated to her coincides with the lack of reliability in information transmission in line with our theoretical assumptions (see more details on this point in the Section *Separation between $\mathcal{PUSH}$ and $\mathcal{PULL}$*).

Given the coincidence between the communication patterns in this ant system and the requirements of our lower bound we expect long delays before any uninformed ant can be relatively certain that a recruitment process is occurring. We therefore measured the time it takes an ant, that has been at the food source, to recruit the help of two nest-mates regardless of total group size. One might have expected this time to be independent of the group size or even to decrease as two ants constitute a smaller fraction of larger groups. To the contrary, we find that the time until the second ant is recruited increases with group size ($p < 0.05$ Kolmogorov-Smirnov test over $N = 24$ experiments, see S2d Fig).

Our theoretical results set a lower bound on the minimal time it takes uninformed ants to be recruited. Note that our lower bounds actually correspond to the time until *any* individual can be sure with more than 2/3 probability of the rumor. In the context of the ant recruitment experiment this means that if an ant goes out of the nest only if she is sure with some probability that there is a reason to exit, then the lower bounds correspond to the time until the first ant (other than the recruiter) exits.

Our lower bound is linear in the group size (Theorem 1.1). Note that this does not imply that the ants' biological algorithm matches the lower bound and must be linear as well. Rather, our theoretical results qualitatively predict that as group size grows, recruitment times must eventually grow as well. This stands in agreement with S2d Fig. Thus, in this system, inherently noisy interactions on the microscopic level have direct implications on group level performance.

Overview of the main lower bound proof

Here, we provide the intuition for our main theoretical result, Theorem 1.1. For a formal proof please refer to the Supplementary Information [48], The Lower Bounds Section. The proof can be broken into three parts and, below, we refer to each of them separately.

Part I. From parallel-$\mathcal{PULL}$ to broadcast-$\mathcal{PULL}$. Consider an efficient protocol $\mathcal{P}$ for the parallel-$\mathcal{PULL}$ setting. The first part of the proof shows how $\mathcal{P}$ can be used to produce a protocol $\mathcal{P}'$ that operates in another model, called *broadcast-$\mathcal{PULL}$*. In this latter model, at each time step $t \in \mathbb{N}^+$ one agent is chosen u.a.r. and all agents observe it, receiving the same noisy sample of its message. The running time of the resulting protocol $\mathcal{P}'$ will be n times the running time of $\mathcal{P}$. The construction of $\mathcal{P}'$ builds on the permissive assumptions we employ regarding the power of computation of agents and their unique identities in $\{1, 2, \dots n\}$. In $\mathcal{P}'$, agents divide time steps in the broadcast-$\mathcal{PULL}$ model into *rounds*, each composed of precisely n time steps. For an integer i, where $1 \leq i \leq n$, during the i-th step of each round, all agents receive an observation, but $n-1$ of them ignore it. Specifically, only agent $(i \bmod n)+1$ keeps the observation. The agent will then wait until the end of the round to actually process this observation according to $\mathcal{P}$. This ensures that when a round is completed, each agent processes precisely one independent uniform sample from the configuration of the previous round, as it would in a round of the parallel-$\mathcal{PULL}$ model. This draws a precise injection from rounds in broadcast-$\mathcal{PULL}$ and rounds in parallel-$\mathcal{PULL}$. This construction implies that to prove Theorem 1.1 it is enough to prove that there is no rumor spreading protocol in the broadcast-$\mathcal{PULL}$ model that converges in less than $\Omega(\frac{n^2\delta}{s^2(1-\delta|\Sigma|)^2})$ rounds.

Part II. From broadcast-$\mathcal{PULL}$ to a statistical inference problem. To establish the desired lower bound, we next show how the rumor spreading problem in the broadcast-$\mathcal{PULL}$ model relates to a statistical inference test. That is, from the perspective of a given agent, the rumor spreading problem can be understood as the following: Based on a sequence of noisy observations, the agent should be able to tell whether the correct opinion is 0 or 1. We formulate this problem as a specific task of distinguishing between two random processes, one originated by running the protocol assuming the correct opinion is 0 and the other assuming it is 1.

One of the main difficulties lies in the fact that these processes may have a memory. At different time steps, they do not necessarily consist of independent draws of a given random variable. In other words, the probability distribution of an observation not only depends on the correct opinion, on the initial configuration and on the underlying randomness used by agents, but also on the previous noisy observation samples and (consequently) on the messages agents themselves choose to display on that round. An intuitive version of this problem is the task of distinguishing between two (multi-valued) biased coins, whose bias

changes according to the previous outcomes of tossing them (*e.g.,* due to wear). See Figure 3 for an illustration.

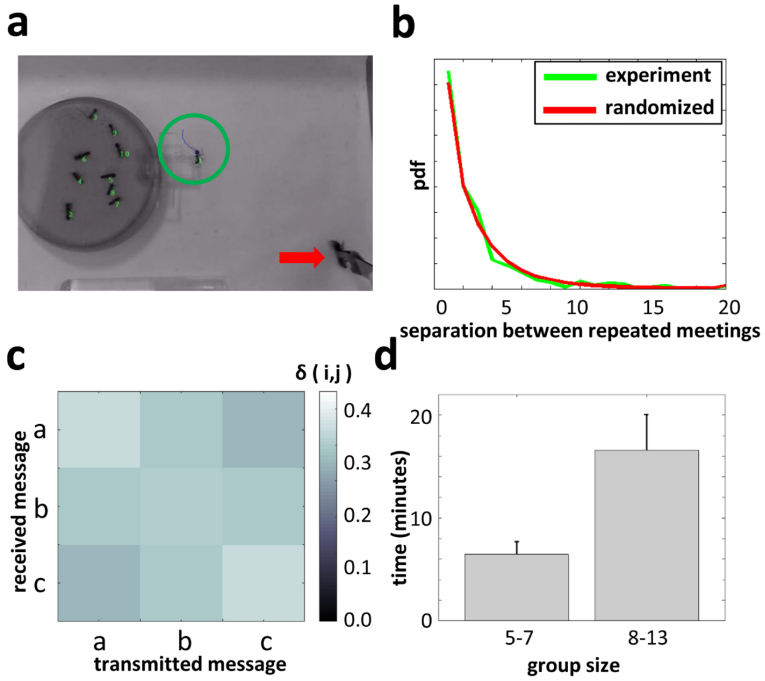


Figure 3: **Distinguishing between two types of coins.** On the top there are two possible coins with slightly different distributions for yielding a head (H) or a tail (T). (We depicted two possible outcomes but our model can account for more.) Given a sequence of observations (corresponding to the random outcomes of coin tosses), the goal of the observer is to guess the coin type being used (either 0 or 1). The wear induced by tossing the coins may, with time, change the probability that they land on either heads or tails in a way that depends on the coin type as well as on the previous toss outcomes (observations). In particular, notice that without a change in the probability of heads in our example we would not obtain a posterior probability 0.51 starting with a prior of 0.45 after six coin flips involving two heads. $P_j(H^{(t)} \mid \text{observations})$ for $j \in \{0,1\}$ denotes the probability of Coin of type j to yield H given the particular sequence of t observations. Here, this sequence is H,T,T,H,T,T. At the beginning of the next round, *i.e.,* the 7th round, $|\varepsilon^{(t=6)}|_1$ measures how "far" the the H vs T distribution generated by the worn down coin 1 is from the same distribution as generated by the worn down coin 2. More precisely, for the case above, $|\varepsilon^{(6)}|_1 = |P_0(H \mid \text{observations}) - P_1(H \mid \text{observations})| + |P_0(T \mid \text{observations}) - P_1(T \mid \text{observations})| = |0.51 - 0.43| + |0.49 - 0.57| = 0.16$. The parameter ε bounds all possible $|\varepsilon^{(t)}|_1$ from above.

Despite this apparent complexity, we show that the difficulty of this distinguishing task can be captured by two scalar parameters, denoted ε and δ. Parameter δ lower bounds the probability for any observation to be attained (given any sequence of observations). Parameter ε captures the extent to which the processes are 'similar'. More specifically, at round t, given a sequence of previous observations, denoted $x^{(<t)}$, the next observation has the same probability to be attained in each process, up to an ε additive term (see Figure S3). A crucial observation is that ε is very small, precisely, $\varepsilon = \Theta(s(1-\delta|\Sigma|)/n)$. This follows from the fact that given $x^{(<t)}$, the behavior of non-source agents in the two processes is the same, regardless of the value of the correct opinion. Indeed, internally, an agent is only affected by its initial knowledge, the randomness it uses, and the sequence of observations it sees. This means that at round t, the processes would differ only if the agent to be observed on that round happens to be a source (which happens with probability s/n) and, on top of that, the observed message is not the same due to noise (which accounts for the factor $(1-\delta|\Sigma|)$). However, a small value of ε is not enough to ensure slow running time. Indeed, even though the t'th observation may be distributed almost the same, if it happens that some observation can be attained only in one process, then seeing such an observation would immediately allow the observer to distinguish the two processes. A sufficiently large δ prevents the aforementioned scenario.

Part III. A lower bound for the statistical test problem. The last step of the proof shows that at least $\Omega(\delta/\varepsilon^2)$ samples are required in order to solve any distinguishing task with parameters δ and ε. The proof involves notions from Statistical Hypothesis Testing such as the Kullback-Leibler (KL) divergence (see, *e.g.,* Chapter 5 in [36]). For example, generalizing known results, we show that, if $P_0^{(\leq t)}$ and $P_1^{(\leq t)}$ are the two distributions of observations up to time t, any distinguishing algorithm must satisfy that the error probability is at least $1-\sqrt{KL(P_0^{(\leq t)}, P_1^{(\leq t)})}$. Hence, for the probability of error to be small, the term $KL(P_0^{(\leq t)}, P_1^{(\leq t)})$ must be large. To calculate the KL-divergence one often uses a tensorization lemma, but this could not be used in our case since the observations in different rounds are not independent. Instead, we use the more general Chain Rule identity (see [36]). This allows us to focus on the KL-divergence of every round separately rather than of the whole sequence. In contrast to the fully independent case, we also condition on the previous draws, on the randomness used by agents, and on the initial configuration. Finally, we obtain: $KL(P_0^{(\leq t)}, P_1^{(\leq t)}) = \mathcal{O}(\frac{t\varepsilon^2}{\delta})$. This implies that the number of observations t needs to be of order δ/ε^2 to make the error less than, say $1/3$. This bound translates to a lower bound of $\Omega\left(n^2\delta/(s^2(1-\delta|\Sigma|)^2)\right)$ steps for the broadcast-$\mathcal{PULL}$ model and hence a lower bound of $\Omega\left(n\delta/(s^2(1-\delta|\Sigma|)^2)\right)$ rounds for the parallel-$\mathcal{PULL}$ model.

Generalizations

Several of the assumptions discussed earlier for the parallel-$\mathcal{PULL}$ model were made for the sake of simplicity of presentation. In fact, our results can be shown to hold under more general conditions, that include: 1) different rate for sampling a source, and 2) a more relaxed noise criterion. In addition, our theorems were stated with respect to the parallel-$\mathcal{PULL}$ model. In this model, at every round, each agent samples a single agent u.a.r. In fact, for any integer k, our analysis can be applied to the model in which, at every round, each agent observes k agents chosen u.a.r. In this case, the lower bound would simply reduce by a factor of k. Our analysis can also apply to a sequential variant, in which in each time step, two agents u and v are chosen u.a.r from the population and u observes v. In this case, our lower bounds would multiply by a factor of n, yielding, for example, a lower bound of $\Omega(n^2)$ in the case where δ and s are constants. Observe that the latter increase is not surprising as each round in the parallel-$\mathcal{PULL}$ model consists of n observations, while the sequential model consists of only one observation in each time step. See more details in the Supplementary Information [48].

Discussion

Exponential Separation between $\mathcal{PUSH}$ and $\mathcal{PULL}$

Our lower bounds on the parallel-$\mathcal{PULL}$ model (where agents observe other agents) should be contrasted with known results in the parallel-$\mathcal{PUSH}$ model (this is the push equivalent to parallel-$\mathcal{PULL}$ model, where in each round each agent may or may not actively push a message to another agent chosen u.a.r.). Although never proved, and although their combination is known to achieve more power than each of them separately [16], researchers often view the parallel-$\mathcal{PULL}$ and parallel-$\mathcal{PUSH}$ models as very similar on complete communication topologies. Our lower bound result, however, undermines this belief, proving that in the context of noisy communication, there is an exponential separation between the two models. Indeed, when the noise level is constant for instance, convergence (and in fact, a much stronger convergence than we consider here) can be achieved in the parallel-$\mathcal{PUSH}$ using only logarithmic number of rounds [20, 21], by a simple strategy composed of two stages. The first stage consists of providing all agents with a guess about the source's opinion, in such a way that ensures a non-negligible bias toward the correct guess. The second stage then boosts this bias by progressively amplifying it. A crucial aspect in the first stage is that agents remain silent until a certain point in time that they start sending messages continuously, which happens after being contacted for the first time. This prevents agents from starting to spread information before they have sufficiently reliable knowledge. This allows to control the dynamics of the information spread in a balanced manner. More specifically, marking an edge corresponding to a message received for the first time by an agent, the set of marked edges forms a spanning tree of low depth, rooted at the source.

The depth of such tree can be interpreted as the deterioration of the message's reliability. On the other hand, as shown here, in the parallel-$\mathcal{PULL}$ model, even with the synchronization assumption, rumor spreading cannot be achieved in less than a linear number of rounds.

Perhaps the main reason why these two models are often considered similar is that with an extra bit in the message, a $\mathcal{PUSH}$ protocol can be *approximated* in the $\mathcal{PULL}$ model, by letting this bit indicate whether the agent in the $\mathcal{PUSH}$ model was aiming to push its message. However, for such a strategy to work, this extra bit has to be reliable. Yet, in the noisy $\mathcal{PULL}$ model, no bit is safe from noise, and hence, as we show, such an approximation cannot work. In this sense, the extra power that the noisy $\mathcal{PUSH}$ model gains over the noisy $\mathcal{PULL}$ model, is that the very fact that one node attempts to communicate with another is reliable. This, seemingly minor, difference carries significant consequences.

Strategies to overcome noise in biological systems

Communication in man-made computer networks is often based on reliable signals which are typically transferred over highly defined structures. These allow for ultra-fast and highly reliable calculations. Biological networks are very different from this and often lack reliable messaging, well defined connectivity patterns or both. Our theoretical results seem to suggest that, under such circumstances, efficient spread of information would not be possible. Nevertheless, many biological groups disseminate and share information, and, often, do so reliably. Next, we discuss information sharing in biological systems within the general framework of our lower-bounds.

The correctness of the lower bounds relies on two major assumptions: 1) stochastic interactions, and 2) uniform noise. Communication during desert ant recruitment complies with both these assumptions (see S2b-c Fig) and indeed the speed at which messages travel through the group (see S2d Fig) is low. Below, we discuss several biological examples where efficient rumor spreading is achieved. We expect that, in these examples, at least one of the assumptions mentioned above should break adding some degree of reliability to the overall communication. The group can then utilize this reliability and follow one of the strategies mentioned in Section *Theoretical results*, in order to yield reliable collective performance. We begin by discussing examples that violate the first assumption, namely, that of stochastic interactions, and then discuss examples that violate the second assumption, namely, uniform noise.

Stable connectivity as a means to overcome noise

Synaptic connectivity in the mammalian brain is known to be highly noisy [37]. However, this cellular-level noise has little effect on the global function of brains which are highly reliable. A major factor that allows this restoration of reliability are the structural stability and redundancies that characterize brain connectivity. It is well understood how such properties can allow for fast and

efficient propagation of electrical signals through large neuronal populations [1, 2].

Animal groups can also benefit from stable connectivity to enhance the reliability of rumor spreading. An example comes from house-hunting rock ants [38]. When these ants commence their move between nest sites, the information regarding the location of the target nest is held by only a few scout ants, that then disseminate it to the entire colony. To communicate this information, the ants engage in prolonged interactions which include the formation of stable pairs that walk in tandem towards the target nest and frequently contact each other [39]. These tandem run interactions allow for highly reliable communication and permit the follower ant to lead a subsequent interaction such that the rumor can continue to spread efficiently. While this redundancy of multiple interaction allows for the efficient flow of information, it comes at the inevitable price of long interaction durations [40].

Non-uniform noise

When the physical structure of a group is not well defined, the importance of reliable messaging schemes grows. In flocks of birds and schools of fish, changes in the behavior of a single individual can be relayed across a series of local interactions [41], and generate a response wave that travels across the entire group [3–5]. It has been shown that information can travel faster in flocks that display a higher level of alignment [42]. Within the context of our analysis, as a group becomes more ordered it becomes easier to distinguish a sudden directional change by an informed individual from the random velocity fluctuations of uninformed birds. Besides a possibly reduced level of uniformity in the interactions implied by the additional spatial structure, the reliability of information transfer violates the δ-uniform noise assumption and allows for fast and reliable directional changes on the collective scale.

As noted above, push-type communication is another route which may potentially add sufficient reliability to support rumor spreading. In this sense, what distinguishes push from pull is the trait by which a non-message cannot be confused with a message. An example for the usefulness of an active push behavior comes from alarm behavior in ants. A single ant sensing danger can actively excrete discrete volatile alarm pheromones that are sensed by a large number of group members and elicit panic or attack responses [43, 44]. Conversely, no ant would secrete these distinct pheromones unless she directly perceived danger or sensed the alarm signal. Therefore, when an ant senses an alarm pheromone the only possibility is that one of her nest mates has sensed danger. As indicated by our theoretical lower bound, if alarm messages would be confused with non-alarming messages then such fast and reliable information spread would not be possible.

The difficulty of spreading information fast, as indicated by our theoretical results, is further consistent with the fact that, even in fully-cooperative groups, such as ants or bees, an animal that receives information from a conspecific will often not transfer it further before obtaining its own independent first-hand

knowledge [40, 45–47, 49].

Finally, we note that given the aforementioned discussion, our insight regarding the difficulty of functioning under uniform noise can serve an evolutionary explanation for the emergence of new communication signals (*e.g.,* alarming signal) that would be distinct from other signals, and prevent confusion.

Methods

All experimental results presented in this manuscript are re-analyses of data obtained in *Cataglyphis niger* recruitment experiments [32]. In short, ants in the entrance chamber of an artificial nest were given access to a tethered food item just outside the nest's entrance (S2a Fig). The inability of the ants to retrieve the food induced a recruitment process [32].
The reaction of the ants to this manipulation was filmed and the locations, speeds and interactions of all participating ants were extracted from the resulting videos.

Calculation of δ. To estimate the noise parameter δ we used interactions between ants moving at three different speed ranges (measured in cm/sec), namely, 'a': 1-10, 'b': 10-20, and 'c': over 20 and "receiver" ants. Only interactions in which the receiver ant was initially stationary were used as to ensure that the state of these ants before the interaction is as similar as possible. The message alphabet is then assumed to be $\Sigma = \{a, b, c\}$. The response of a stationary ant v to the interaction was quantified in terms of her speed after the interaction.

An alphabet of three messages was used since the average responses of v to any two messages were significantly different (all p-values smaller than 0.01) justifying the fact that these are not artificial divisions of a continuous speed signal into a large number of overlapping messages. On the other hand, dividing the bins further (say, each bin divided into 2 equal bins) yielded statistically indistinguishable responses from the receiver (all p-values larger than 0.11). Therefore, our current data best supports a three letter alphabet.

Assuming equal priors to all messages in Σ, and given specific speed of the receiver ant, v, the probability that it was the result of a specific message $i \in \Sigma$ was calculated as $p_i(v) = p(v \mid i)/\sum_{k\in\Sigma} p(v \mid k)$, where $p(v \mid j)$ is the probability of responding in speed v after "observing" j. The probability $\delta(i, j)$ that message i was perceived as message j was then estimated as the weighted sum over the entire probability distribution measured as a response to j: $\delta(i, j) = \sum_v p(v \mid j) \cdot p_i(v)$. The parameter δ can then be calculated using $\delta = \min\{\delta(i, j) \mid i, j \in \Sigma\}$.

References

1. Abeles M. Corticonics: Neural circuits of the cerebral cortex. Cambridge Univ. Press; 1991.

2. Feinerman O, Rotem A, Moses E. Reliable neuronal logic devices from patterned hippocampal cultures. Nature physics. 2008;4(12):967–973.

3. Cavagna A, Cimarelli A, Giardina I, Parisi G, Santagati R, Stefanini F, et al. Scale-free correlations in starling flocks. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(26):11865–11870.

4. Marras S, Batty R, Domenici P. Information transfer and antipredator maneuvers in schooling herring. Adaptive Behavior. 2012;20(1):44–56.

5. Rosenthal SB, Twomey CR, Hartnett AT, Wu HS, Couzin ID. Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(15):4690–4695.

6. Giraldeau LA, Valone TJ, Templeton JJ. Potential disadvantages of using socially acquired information. Philosophical Transactions of the Royal Society of London B: Biological Sciences. 2002;357(1427):1559–1566.

7. Rieucau G, Giraldeau LA. Persuasive companions can be wrong: the use of misleading social information in nutmeg mannikins. Behavioral Ecology. 2009; p. 1217–1222.

8. Templeton JJ, L A G. Patch assessment in foraging flocks of European starlings: evidence for the use of public information. Behavioral Ecology. 1995;6(1):65–72.

9. Alon N, Braverman M, Efremenko K, Gelles R, Haeupler B. Reliable Communication over Highly Connected Noisy Networks. In: PODC; 2016. p. 165–173.

10. Gamal AE, Kim AYH. Network Information Theory. Cambridge Univ. Press; 2011.

11. Xu A, Raginsky M. Information-Theoretic Lower Bounds for Distributed Function Computation. IEEE Trans Information Theory. 2017;63(4):2314–2337.

12. von Neumann J. Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components. Automata Studies. 1956; p. 43–98.

13. Boczkowski L, Korman A, Natale E. Minimizing Message Size in Stochastic Communication Patterns: Fast Self-Stabilizing Protocols with 3 bits. In: SODA; 2017. p. 2540–2559.

14. Censor-Hillel K, Haeupler B, Kelner JA, Maymounkov P. Global computation in a poorly connected world: fast rumor spreading with no dependence on conductance. In: STOC; 2012. p. 961–970.

15. Demers A, Greene D, Hauser C, Irish W, Larson J, Shenker S, et al. Epidemic algorithms for replicated database maintenance. In: PODC; 1987.

16. Karp RM, Schindelhauer C, Shenker S, Vöcking B. Randomized Rumor Spreading. In: FOCS; 2000. p. 565–574.

17. Noymer A. The Transmission and Persistence of 'urban Legends': Sociological Application of Age-structured Epidemic Models. The Journal of mathematical sociology. 2001;25(3):1–98.

18. Nekovee M, Moreno Y, Bianconi G, Marsili M. Theory of rumour spreading in complex social networks. Physica A: Statistical Mechanics and its Applications. 2007;374(1):457–470. doi:10.1016/j.physa.2006.07.017.

19. Bikhchandani S, Hirshleifer D, Welch I. Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. The Journal of Economic Perspectives. 1998;12(3):151–170.

20. Feinerman O, Haeupler B, Korman A. Breathe before speaking: efficient information dissemination despite noisy, limited and anonymous communication. In: PODC, 2014;.

21. Fraigniaud P, Natale E. Noisy Rumor Spreading and Plurality Consensus. In: PODC; 2016. p. 127–136.

22. Goyal N, Kindler G, Saks ME. Lower Bounds for the Noisy Broadcast Problem. SIAM Journal of Computing. 2008;37(6):1806–1841.

23. Bialek W. Physical limits to sensation and perception. Annual review of biophysics and biophysical chemistry. 1987;16(1):455–478.

24. Feinerman O, Korman A. Theoretical distributed computing meets biology: A review. In: ICDCIT. Springer; 2013. p. 1–18.

25. Korman A, Greenwald E, Feinerman O. Confidence Sharing: An Economic Strategy for Efficient Information Flows in Animal Groups. PLoS Computational Biology. 2014;10(10). doi:10.1371/journal.pcbi.1003862.

26. Feinerman O, Korman A. The ANTS problem. Distributed Computing. 2017;30(3):149–168. doi:10.1007/s00446-016-0285-8.

27. Afek Y, Alon N, Barad O, Hornstein E, Barkai N, Bar-Joseph Z. A biological solution to a fundamental distributed computing problem. Science. 2011;331(6014):183–185.

28. Bonifaci V, Mehlhorn K, Varma G. Physarum can compute shortest paths. In: SODA; 2012. p. 233–240.

29. Grosu R, Lukina A, Smolka SA, Tiwari A, Yang J. Distributed V-Formation with Dynamic Neighborhood Sizing; 2017. BDA Workshop.

30. Singh S, Rashid S, Long Z, Navlakha S, Salman H, Oltvai ZN, et al. Distributed Gradient Descent in Bacterial Food Search. In: RECOMB; 2016. p. 233–240.

31. Amor F, Ortega P, Cerdá X, Boulay R. Cooperative prey-retrieving in the ant Cataglyphis floricola: An unusual short-distance recruitment. Insectes Sociaux. 2010;57(1).

32. Razin N, Eckmann JP, Feinerman O. Desert ants achieve reliable recruitment across noisy interactions. Journal of the Royal Society Interface; 10(20170079). 2013;.

33. Musco C, Su H, Lynch NA. Ant-Inspired Density Estimation via Random Walks: Extended Abstract. In: PODC; 2016. p. 469–478.

34. Dornhaus A, Chittka L. Food Alert in Bumblebees (Bombus terrestris): Possible Mechanisms and Evolutionary Implications. Behavioral Ecology and Sociobiology. 2001;50(6):570–576.

35. Hölldobler B. Recruitment behavior in Camponotus socius (Hym. Formicidae). Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology. 1971;75(2):123–142. doi:10.1007/BF00335259.

36. Rigollet P. High Dimensional Statistics. Lecture notes for course 18S997. 2015;.

37. Stevens C, Wang Y. Changes in reliability of synaptic function as a mechanism for plasticity. Nature. 1994;371:2O.

38. Beckers R, Goss S, Deneubourg J, Pasteels J. Colony size, communication, and ant foraging strategy. Psyche. 1989;96(3-4):239–256.

39. Franks NR, Richardson T. Teaching in tandem-running ants. Nature. 2006;439(7073):153–153.

40. Pratt SC, Sumpter DJT, Mallon EB, Franks NR. An agent-based model of collective nest choice by the ant Temnothorax albipennis. Animal Behaviour. 2005;70(5):1023–1036.

41. Ballerini M, Cabibbo N, Candelier R, Cavagna A, Cisbani E, Giardina I, et al. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(4):1232–1237.

42. Attanasi A, Cavagna A, Del Castello L, Giardina I, Grigera TS, Jelić A, et al. Information transfer and behavioural inertia in starling flocks. Nature Physics. 2014;10(9):691–696.

43. Blum MS. Alarm pheromones. Annual Review of Entomology. 1969;14(1):57–80.

44. Wilson EO, Regnier Jr FE. The evolution of the alarm-defense system in the formicine ants. The American Naturalist. 1971;105(943):279–289.

45. Seeley TD. Honeybee democracy. Princeton University Press; 2010.

46. Beckers R, Deneubourg J, Goss S. Trail laying behaviour during food recruitment in the antLasius niger (L.). Insectes Sociaux. 1992;39(1):59–72.

47. Carthy JD. The orientation of two allied species of British ant, II. Odour trail laying and following in acanthomyops (Lasius) fuliginosus. Behaviour. 1951; p. 304–318.

48. Boczkowski, Lucas and Natale, Emanuele and Feinerman, Ofer and Korman, Amos Supplementary Information for "Limits on Reliable Information Flows through Stochastic Populations" 2018 `https://hal.archives-ouvertes.fr/hal-01848636`hal-01848636

49. Beckers R, Deneubourg J, Goss S. Modulation of trail laying in the antLasius niger (Hymenoptera: Formicidae) and its role in the collective selection of a food source. Journal of Insect Behavior. 1993;6(6):751–759.