



**HAL**  
open science

# Spreading a Confirmed Rumor: A Case for Oscillatory Dynamics

Bartłomiej Dudek, Adrian Kosowski

► **To cite this version:**

Bartłomiej Dudek, Adrian Kosowski. Spreading a Confirmed Rumor: A Case for Oscillatory Dynamics. 2017. hal-01503359v1

**HAL Id: hal-01503359**

**<https://inria.hal.science/hal-01503359v1>**

Preprint submitted on 25 May 2017 (v1), last revised 28 Nov 2017 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Spreading a Confirmed Rumor: A Case for Oscillatory Dynamics

Bartłomiej Dudek<sup>1</sup> and Adrian Kosowski<sup>2</sup>

<sup>1</sup>Institute of Computer Science, University of Wrocław, Poland

<sup>2</sup>Inria Paris and IRIF, Université Paris Diderot, France

## Abstract

We consider an information spreading problem in which a population of  $n$  agents is to determine, through random pairwise interactions, whether an authoritative rumor source  $X$  is present in the population or not. The studied problem is a generalization of the rumor spreading problem, in which we additionally impose that the rumor should disappear when the rumor source no longer exists. It is also a generalization of the self-stabilizing broadcasting problem and has direct application to amplifying trace concentrations in chemical reaction networks.

We show that there exists a protocol such that, starting from any possible initial state configuration, in the absence of a rumor source all agents reach a designated “uninformed” state after  $O(\log^2 n)$  rounds w.h.p., whereas in the presence of the rumor source, at any time after at least  $O(\log n)$  rounds from the moment  $X$  appears, at least  $(1 - \varepsilon)n$  agents are in an “informed” state with probability  $1 - O(1/n)$ , where  $\varepsilon > 0$  may be arbitrarily fixed. The protocol uses a constant number of states and its operation relies on an underlying oscillatory dynamics with a closed limit orbit. On the negative side, we show that any system which has such an ability to “suppress false rumors” in sub-polynomial time must either exhibit significant and perpetual variations of opinion over time in the presence of the rumor source, or use a super-constant number of states.

**Key words:** Rumor-spreading, Epidemic process, Oscillator dynamics, Population protocols, Broadcasting, Distributed clock synchronization.

## 1 Introduction

Rumour-spreading is an epidemic process in which agents in a population change their states from “uninformed” to “informed” following the appearance of a rumor, originating, e.g., from a single informed agent. In pairwise interactions, uninformed agents become informed upon meeting a previously informed agent.

Such a rumour-spreading process, taking place through interactions between random pairs of agents, propagates information through a population of  $n$  agents in  $O(\log n)$  parallel rounds of interaction [19, 38]. In communication networks, this type of process was first studied in the context of the random phone call model [38]. Rumour spreading processes have since been considered in many scenarios, differing in scheduler properties and of engagement into interaction (e.g., push vs. push-pull [26]), the environment of propagation (complete interaction graphs vs. other topologies [17, 21, 41]), and fault-tolerance [15, 16].

When modeling social interactions, as well as in a number of other applications, a task inherently linked to rumour spreading is that of suppressing false or outdated rumours. Specifically, it may happen that a certain part of a population find themselves in an informed state before the original rumour

source is identified as a source of false information, a false rumor may be propagated accidentally because of an agent which previously changed state from “uninformed” to “informed” due to a fault or miscommunication, or the rumor may contain information which is no longer true.

Allowing a system to defend itself against this type of “false positive” error is an algorithmic challenge. One solution consists in designing protocols which propagate messages from the source towards other nodes with a form of time-to-live counter, propagated after each retransmission, where message retransmission does not continue for messages with an age of at least, e.g.,  $\Omega(\log n)$ . This type of solution is fundamental to broadcasting/flooding routines in communication networks and may also be applied to population dynamics if a sufficiently large number of possible states of an agent (e.g.,  $\Omega(\log n)$ ) is allowed [4]. In social networks, it would correspond to an agent being able to confidently distinguish primary news sources from secondary sources and to have a very well-calibrated estimation of trust for different sources. Another type of solution to false positives in rumor-spreading relies on an external clock which regularly, e.g., every  $O(\log n)$  rounds, resets all informed agents to an uninformed (or “informed but not propagating the rumor” state). This type of solution is, however, beyond most the capabilities of most models of decentralized interaction networks.

In this work we provide an efficient way of solving the problem of *spreading confirmed rumors*, i.e., spreading of rumors with a source and suppression of rumors lacking a primary source. Under a fair random scheduler, the protocol suppresses rumors without a source, moving all agents to “uninformed” states in  $O(\log^2 n)$  parallel rounds w.h.p., and in the case of a rumor with a source, after  $O(\log n)$  parallel rounds at any moment of time maintains  $(1 - \varepsilon)n$  nodes in “informed” states, with high probability. The protocol relies on an underlying oscillatory dynamics. It works from any starting configuration, and thus may be seen as having some form of self-stabilizing property (though the term should be used with caution, given that the protocol never actually stabilizes in the presence of a rumor source, modifying states perpetually). We also show that any fast constant-state solution to the problem of spreading confirmed rumors requires not only the possibility of one-sided error in moving nodes to the informed state, but also must admit perpetual and significant volatility in the representation of states in the population over time.

The paper carries two general messages. Broadly speaking, we highlight the possibility that permanent volatility of the proportions of some states (opinions) in a population may play a significant part in real-world mechanisms of spreading information. Thus, in some contexts, information (rumors) spreading through a system may be viewed as a form of a *periodic signal* (i.e., given in the frequency domain), rather than as a *state of the system* (i.e., a snapshot of the system at a fixed moment of time). More specifically, at a computational level, we show that distributed systems have a natural capacity to generate their own approximate “internal clock” (counting modulo a constant) through oscillatory dynamics, which can be easily controlled and serve as a replacement for algorithmic time-to-live mechanisms or for global clocks external to the system.

## 1.1 Motivation

The problem of spreading confirmed rumors is a natural generalization of the classical rumor spreading problem [23, 38]. It is also a generalization of the *self-stabilizing broadcasting* problem (see e.g. [10]) through the following reduction. Suppose a designated source agent wishes to transmit a one-bit message, 0 or 1, to a population which may be in an arbitrary starting configuration. Then, such a source may act as a rumor source if the message it wishes to spread is 1, and deactivate (stay silent) otherwise. As a result, the population adopts an informed or uninformed state, corresponding to the value spread by the source. The same approach can be used for broadcasting of messages of size  $O(1)$  by the deployment of several independent protocols running in parallel on the same system. Larger messages obviously cannot be broadcast in systems with  $O(1)$  states. Our protocol thus solves the self-stabilizing broadcasting problem in  $O(\log^2 n)$  expected time with constant one-sided error.

A specific application of the studied confirmed rumor spreading scenario pertains to chemical reaction networks (CRN-s) [20]. The objective of the proposed protocol is then to detect the existence of a particle  $X$  in a population (e.g., a trace pollution, or specific type of living cell) and to amplify this concentration to a level directly measurable, e.g., by changing the color of the entire studied solution. Since perfect initialization of such protocols is not possible, convergence to an “informed” or “uninformed” state is required from a possibly corrupted initial configuration. We should remark that interaction protocols of the type studied in this work, including oscillatory dynamics, are readily implementable in the framework of DNA computing [42].

## 1.2 Model and Results

**Schedulers and protocols.** The protocols studied in this work are stated in the population protocol framework of Angluin et al. [5–7], permitting protocols which are non-deterministic and have a super-constant number of states. (These extensions are not necessary for our main positive result, as we discuss later.)

A non-deterministic *population protocol* for a population of  $n$  agents is defined as a pair  $P = (K_n, R_n)$ , where  $K_n$  is the set of *states* and  $R_n$  is the set of *interaction rules*. We will simply write  $P = (K, R)$ , when considering a protocol which is *universal* (i.e., defined in the same way for each value of  $n$ ) or if the value of  $n$  is clear from the context. The set of rules  $R \subseteq K^4 \times [0, 1]$  is given so that each rule  $j \in R$  is of the form  $j = (i_1(j), i_2(j), o_1(j), o_2(j), q_j)$ , describing an interaction read as: “ $(i_1(j), i_2(j)) \rightarrow (o_1(j), o_2(j))$  with probability  $q_j$ ”. For all  $i_1, i_2 \in K$ , we define  $R_{i_1, i_2} = \{j \in R : (i_1(j), i_2(j)) = (i_1, i_2)\}$  as the set of rules acting on the pair of states  $i_1, i_2$ , and impose that  $\sum_{j \in R_{i_1, i_2}} q_j \leq 1$ .

In any configuration of the system, each of the  $n$  agents from the population is in one of states from  $K_n$ . The protocol is executed by an asynchronous scheduler, which runs in *steps*. In every step the scheduler uniformly at random chooses from the population a pair of distinct agents to interact: the *initiator* and the *receiver*. If the initiator and receiver are in states  $i_1$  and  $i_2$ , respectively, then the protocol executes at most one rule from set protocol  $R_{i_1, i_2}$ , selecting rule  $j \in R_{i_1, i_2}$  with probability  $q_j$ . If rule  $j$  is executed, the initiator then changes its state to  $o_1(j)$  and the receiver to  $o_2(j)$ . The rumor source has a special state, denoted  $X$ , which is never modified by any rule.

Time in the system is measured in the scheduler’s steps. The step-by-step setting is frequently used [3, 34] and more convenient than one based on parallel execution in rounds, since it does not require the scheduler to find matchings or perform conflict resolution between overlapping interacting pairs of nodes. For the sake of consistency with the parallel model, we sometimes equivalently phrase asymptotic time complexity in *parallel round*, where one parallel round corresponds to  $\Theta(n)$  steps.

**The positive result.** For a state  $A \in K$ , we denote the number of agents in state  $A$  as  $\#A$ , and the *concentration* of state  $A$  as  $a = \#A/n$ , and likewise for a set of states  $\mathcal{A}$ , we write  $\#\mathcal{A} = \sum_{A \in \mathcal{A}} \#A$ .

We are now ready to phrase our main positive result as follows.

**Theorem 1** (Fast rumor spreading and suppression). *For any  $\varepsilon > 0$ , there exists a protocol  $P_r$  with  $|K| = 13$  states, including a distinguished source state  $X$  and a distinguished set of output states  $\mathcal{Y}$ , which solves the problem of spreading confirmed rumors as follows:*

1. *For any starting configuration, in the presence of the source ( $\#X \geq 1$ ), after an initialization period of  $O(\log n)$  parallel rounds, at an arbitrary time step we have  $\#\mathcal{Y} \geq (1 - \varepsilon)n$ , with probability  $1 - O(1/n)$ .*
2. *For any starting configuration, in the absence of the source ( $\#X = 0$ ), the system always reaches a configuration such that  $\#\mathcal{Y} = 0$  for all subsequent time steps. Such a configuration is reached in  $O(\log^2 n)$  parallel rounds, with probability  $1 - O(1/n)$ .*

The main building block in the construction of protocol  $P_r$  is a 7-state sub-protocol  $P_o$  following oscillator dynamics, discussed in detail in Section 3, which we believe to be of independent interest. It has the property that in the absence of  $X$  it stops in a corner state of the phase space, in which only one of three possible states appears in the population, and otherwise regularly (every  $O(\log n)$  steps) moves sufficiently far away from all corner states (see Theorem 3 for a formalization of this property). The complete design of protocol  $P_o$  is shown in Fig. 1 and analyzed in Section 3. Simulation timelines shown in the Appendix illustrate the idea of its operation. For small values of  $\#X > 0$ , the protocol can be very roughly (and non-rigorously) viewed as cyclic composition of three dominant rumor spreading processes over three sets of states  $A_1, A_2, A_3$ , one converting states  $A_1$  to  $A_3$ , the next from  $A_3$  to  $A_2$ , and the last from  $A_3$  to  $A_1$ , which spontaneously take over at moments of time separated by  $O(\log n)$  parallel rounds. For other starting configurations, and especially for the case of  $\#X = 0$ , the 6-dimensional dynamics of the protocol is more involved to describe and analyze (see Subsection 3.2).

We remark on some subtleties of the theorem statement for protocol  $P_r$ . In the case of  $\#X > 0$ , the asymptotic analysis of the protocol resembles that of rumor spreading, except that the protocol fails to reach (at any given moment of time) a fraction of the population, which will, with high probability, be not more than any given small constant. The failure probability stated in this case as  $O(1/n)$  is immediately boosted to an arbitrary inverse polynomial in  $n$  by choosing appropriately the constants in the protocol. In the case of  $\#X = 0$ , the time for the protocol to stabilize is by the oscillator subroutine is given as  $O(\log n)$  rounds *in expectation* (or, with constant probability), and due to the fact that we exploit the variance of the random choices made by the scheduler in the analysis, it does not readily extend to a comparable w.h.p. analysis. The best w.h.p. bound we know (which appears tight) is  $O(\log^2 n)$  rounds with probability  $1 - O(1/n)$ , as stated in the theorem. On the positive side, the analysis we perform is largely insensitive to the details of the scheduler and may be performed similarly for other fair random scheduler models (including parallel schedulers).

In terms of the protocol model, the design of protocol  $P_o$  is universal. The design of protocol  $P_r$  is not universal, as it includes a forgetting (fade-out) rule with a transition probability depending on  $1/\log n$ . As we remark in Section 4,  $P_r$  can be made universal at the cost of a significant increase in the number of states and forsaking the practicality of the approach. All protocols are presented in the non-deterministic framework, however, the universal protocols considered here are amenable to a form of conversion into deterministic rules discussed in [3], which simulates randomness of rules by exploiting the inherent randomness of the scheduler in choosing interacting node pairs to distribute weakly dependent random bits around the system.

All protocols designed in this work are *initiator-preserving*, which means that for any rule  $j \in R$ , we have  $o_1(j) = i_1(j)$  (i.e., have all rules of the form  $A + B \rightarrow A + C$ , also more compactly written as  $A; B \rightarrow C$ ), which makes them relevant in a larger number of application. We note that the basic rumor spreading model is initiator-preserving and given simply as  $1; 0 \rightarrow 1$ . The protocols can also obviously be rewritten to act on unordered pairs of agents picked by the scheduler, rather than ordered pairs.

**The lower bound.** For convenience of notation, we identify a configuration of the population with a vector  $z = (z^{(1)}, \dots, z^{(k)}) \in \{0, 1, \dots, n\}^k = Z$ , where  $z^{(i)}$ , for  $1 \leq i \leq k$ , denotes the number of agents in the population having state  $i$ , and  $\|z\|_1 = n$ . Our main lower bound may now be stated as follows.

**Theorem 2** (Stabilization precludes fast rumor suppression). *Let  $\varepsilon_1 > 0$  be arbitrarily chosen, let  $P$  be any  $k$ -state protocol, and let  $z_0$  be a configuration of the system with at most  $n^{\varepsilon_0}$  agents in state  $X$ , where  $\varepsilon_0 \in (0, \varepsilon_1]$  is a constant depending only on  $k$  and  $\varepsilon_1$ . Let  $B$  be a subset of the state space around  $z_0$  such that the population of each state within  $B$  is within a factor of at most  $n^{\varepsilon_0}$  from that in  $z_0$  (for any  $z \in B$ , for all states  $i \in \{1, \dots, k\}$ , we have  $z_0^{(i)}/n^{\varepsilon_0} < z^{(i)} \leq n^{\varepsilon_0} \max\{1, z_0^{(i)}\}$ ).*

Suppose that in an execution of  $P$  starting from configuration  $z_0$ , with probability  $1 - o(1)$ , the configurations of the system in the next  $n^{2\varepsilon_1}$  parallel rounds are confined to  $B$ .

Then, an execution of  $P$  for  $n^{2\varepsilon_0}$  parallel rounds, starting from a configuration in which state  $X$  has been removed from  $z_0$ , reaches a configuration in a  $O(n^{6\varepsilon_1})$ -neighborhood of  $B$ , with probability  $1 - o(1)$ .

In the statement of the Theorem, for the sake of maintaining the size of the population, we interpret “removing state  $X$  from  $z_0$ ” as replacing the state of all agents in state  $X$  by some other state, chosen adversarially (in fact, this may be any state which has sufficiently many representatives in configuration  $z_0$ ). The  $O(n^{6\varepsilon_1})$ -neighborhood of  $B$  is understood in the sense of the 1-norm or, asymptotically equivalently, the total variation distance, reflecting configurations which can be converted into a configuration from  $B$  by flipping the states of  $O(n^{6\varepsilon_1})$  agents. Informally, if  $B$  represents the set of configurations of the considered protocol, which are understood as the protocol giving the answer “ $\#X > 0$ ”, then our theorem says that, with probability  $1 - o(1)$ , the vast majority of agents will not notice that  $\#X$  had been set to 0, even a polynomial number of steps after this has occurred.

Our lower bound is obviously applicable to any protocol which stabilizes to a fixed configuration for the case  $\#X > 0$  (e.g., one which eventually sets all agents in one informed state). More strongly, it also applies to any protocol in which the number of agents in any state does not change by at least a polynomially large factor during polynomially many rounds. For example, if an algorithm has the property that a configuration reached by the protocol for the case  $\#X > 0$  with high probability has  $\Theta(n)$  agents in some of its states (and persistently 0 agents in other states), then this also precludes an efficient solution to rumor suppression. The concentration of some state would regularly need to drop from  $\Theta(n)$  to  $o(n)$  to make this possible.

The lower bound is stated in reference to protocols with a constant number of states, however, it may be extended to protocols with a non-constant number of states  $k$ , showing that such protocols require  $n^{\exp[-O(\text{poly}(k))]}$  time to reach a desirable output. (This time is larger than polylogarithmic up to some threshold value  $k = O(\text{poly log log } n)$ .) The lower bound covers randomized protocols, including those in which rule probabilities depend on  $n$  (i.e., non-universal ones). Whereas we use the language of discrete dynamics, we informally remark that the protocols covered by the lower bound of Theorem 2 include those whose dynamics  $z_t/n$ , described in the continuous limit ( $n \rightarrow +\infty$ ), has only point attractors, repellers, and fixed points. In this sense, the use of oscillatory dynamics in our protocol seems inevitable.

The proof of Theorem 2 is provided in Section 5. It proceeds by a coupling argument between a process starting from  $z_0$  and a perturbed process in which state  $X$  has been removed. The analysis differently treats rules and states which are seldom encountered during the execution of the protocol from those that are encountered with polynomially higher probability (such a clear separation is only possible when  $k = O(\text{poly log log } n)$ ). Eventually, the probability of success of the coupling reduces to a two-dimensional biased random walk scenario, in which the coordinates represent differences between the number of times particular rules have been executed in the two coupled processes. An essential element of the analysis is that it works only when state  $X$  is removed in the perturbed process. Thus, there is nothing to prevent the dynamics from stabilizing even to a single point in the case of  $X = 0$ , which is indeed the case for our protocol  $P_r$ .

### 1.3 Related work

Our work fits into the line of research on rumor spreading, population protocols, and related interaction models.

Perhaps the most closely related result in terms of considered research questions is [10], which deals with the previously mentioned problem of self-stabilizing broadcasting, in which a rumor source always exists, but may transmit different values of a bit. A protocol solving their problem was presented, giving

a stabilized solution in  $\tilde{O}(\log n)$  time, using a super-constant number of states of agents, but exchanging messages of bit size  $O(1)$  (this assumption can be modeled in the population protocol framework as a restriction on the permitted rule set). Applying our protocol to the self-stabilizing broadcasting problem provides a solution using a constant number of states, but with one-sided error. In a similar way, variants of our protocol can also be used for solving more general problems studied in [10], in which different rumors are broadcast by competing sources. We do not know if a clear separation between oscillatory and non-oscillatory protocols, of the sort given by Theorem 1 and Theorem 2, also holds for self-stabilizing broadcasting.

In a work complementary to this paper [4], we look at applications of the confirmed rumor spreading problem in DNA computing, focusing on performance on protocols based on a time-to-live principle (with  $O(\log n)$  states) and on issues of fault tolerance in a real-world model with leaks.

Our work also touches on the issue of how distributed systems may spontaneously achieve some form of coordination with minimum agent capabilities. The basic work in this direction, starting with the seminal paper [29], focuses on synchronizing timers through asynchronous interprocess communication to allow processes to construct a total ordering of events. A separate interesting question concerns local clocks which, on their own, have some drift, and which need to synchronize in a network environment (cf. e.g. [31, 33], or [30] for a survey of open problems).

**Rumor spreading.** Rumor spreading protocols are frequently studied in a synchronous setting. In a synchronous protocol, in each parallel round, each vertex independently at random activates a local rule, which allows it either to spread the rumor (if it is already informed), or possibly also to receive it (if it has not yet been informed, as is the case in the push-pull model). The standard push rumor spreading model assumes that each informed neighbor calls exactly one uninformed neighbor. In the basic scenario, corresponding to the complete interaction network, the number of parallel rounds for a single rumor source to inform all other nodes is given as  $\log_2 n + \ln n + o(\log n)$ , with high probability [23, 38]. More general graph scenarios have been studied in [21] in the context of applications in broadcasting information in a network. Graph classes studied for the graph model include hypercubes [21], expanders [41], and other models of random graphs [22]. The push-pull model of rumor spreading is an important variation: whereas for complete networks the speedup due to the pull process is in the order of a multiplicative constant [26], the speed up turns out to be asymptotic, e.g., on preferential attachment graphs, where the rumor spreading time is reduced from  $\Theta(\log n)$  rounds in the push model to  $\Theta(\log n / \log \log n)$  rounds in the push-pull model [17], as well as on other graphs with a non-uniform degree distribution. The push-pull model often also proves more amenable to theoretical analysis. We note that asynchronous rumor spreading on graphs, in models closer to our random scheduler, has also been considered in recent work [24, 36], with [24] pointing out the tight connections between the synchronous (particularly push-pull) and asynchronous models in general networks.

**Population protocols.** Population protocols are a model which captures the way in which the complex behavior of systems (biological, sensor nets, etc.) emerges from the underlying local interactions of agents. The original model of Angluin et al. [5, 6] was motivated by applications in sensor mobility. Despite the limited computational capabilities of individual sensors, such protocols permit at least (depending on available extensions to the model) the computation of two important classes of functions: threshold predicates, which decide if the weighted average of types appearing in the population exceeds a certain value, and modulo remainders of similar weighted averages. The majority function, which belongs to the class of threshold functions, was shown to be stably computable for the complete interaction graph [5]; further results in the area of majority computation can be found in [6, 8, 9, 34]. A survey of applications and models of population protocols is provided in [8, 35]. An interesting line of research is related to studies of the algorithmic properties of dynamics of chemical reaction networks [20]. These

are as powerful as population protocols, though some extensions of the chemical reaction model also allow the population size to change in time.

We remark that a recent line of work in this area [3, 20] provides a powerful set of tools for proving lower bounds on the number of states (typically  $\Omega(\log \log n)$  states) for fast (typically polylogarithmic) population protocols for different problems, especially for the case of deterministic protocols. We were unable to leverage these results to prove our lower bound for the randomized scenario studied here, and believe our coupling analysis is complementary to their results.

**Nonlinearity in interaction protocols.** Linear dynamical systems, as well as many nonlinear protocols subjected to rigorous analytical study, have a relatively simple structure of point attractors and repellers in the phase space. The underlying continuous dynamics (in the limit of  $n \rightarrow +\infty$ ) of many interaction protocols defined for complete graphs would fit into this category: basic models of randomized rumour spreading [38]; models of opinion propagation (e.g. [1, 12]); population protocols for problems such as majority and thresholds [5, 6]; all reducible Markov chain processes, such as random walks and randomized iterative load balancing schemes.

Nonlinear dynamics with non-trivial limit orbits are fundamental to many areas of systems science, including the study of physical, chemical and biological systems, and to applications in control science. In general, population dynamics with interactions between pairs of agents are non-linear (representable as a set of quadratic difference equations) and have potentially complicated structure if the number of states is 3 or more. For example, the simple continuous Lotka-Volterra dynamics [32] gives rise to a number of discrete models, for example one representing interactions of the form  $A + B \rightarrow A + A$ , over some pairs  $A, B$  of states in a population (cf. [43] for further generalizations of the framework or [13] for a rigorous analysis in the random scheduler model). The model describes transient stability in a setting in which several species are in a cyclic predator-prey relation. Cyclic protocols of the type have been consequently identified as a potential mechanism for describing and maintaining biodiversity, e.g., in bacterial colonies [27, 28]. Cycles of length 3, in which type  $A_2$  attacks type  $A_1$ , type  $A_3$  attacks type  $A_2$ , and type  $A_1$  attacks type  $A_3$ , form the basis of the basic oscillator, also used as the starting point for protocols in this work, which is referred to as the RPS (rock-paper-scissors) oscillator or simply the 3-cycle oscillator, which we discuss further in Section 2. This protocol has been given a lot of attention in the statistical physics literature. The original analytical estimation method applied to RPS was based on approximation with the Fokker-Planck equation [40]. A subsequent analysis of cyclic 3- and 4-species models using Khasminskii stochastic averaging can be found in [14], and a mean field approximation-based analysis of RPS is performed in [37]. In [13], we have performed a study of some algorithmic implications of RPS, showing that the protocol may be used to perform randomized choice in a population, promoting minority opinions, in  $\tilde{O}(n^2)$  steps. All of these results provide a good qualitative understanding of the behavior of the basic cyclic protocols. We remark that the protocol used in this paper is directly inspired by the properties of RPS, as we discuss further on, but has a more complicated interaction structure (see Fig. 1).

We also remark that local interaction dynamics on arbitrary graphs (as opposed to the complete interaction graph) exhibit a much more complex structure of their limit behavior, even if the graph has periodic structure, e.g., that of a grid. Oscillatory behavior may be overlaid with spatial effects [43], or the system may have an attractor at a critical point, leading to simple dynamic processes displaying self-organized criticality (SOC, [39]).

## 2 Preliminaries: Discrete vs. Continuous Dynamics

**Notation.** For a configuration of a population protocol, we write  $z = (z^{(1)}, \dots, z^{(k)})$  to describe the number of agents in the  $k$  states of the protocol, and likewise use vector  $u = (u^{(1)}, \dots, u^{(k)}) = z/n$  to



describe their concentrations. The concentration of a state called  $A$  which is the  $i_A$ -th state in vector  $u$  is equivalently written as  $a \equiv a(u) \equiv u^{(i_A)}$ , depending on which notation is the easiest to use in a given transformation.

If vector  $u$  represents the current configuration of the protocol and  $u' := u'|u$  is the random variable describing the next configuration of the protocol after the execution of a single rule, we write  $\Delta u := u' - u$ . We also use the notation  $\Delta$  to functions of state  $u$ .

Next, we define the *continuous dynamics* associated with the protocol by the following vector differential equation:

$$\dot{u} \equiv \frac{du}{dt} := n\mathbb{E}(\Delta u)$$

and likewise, for each coordinate,  $\dot{a} = n\mathbb{E}(\Delta a)$  (we use the dot-notation and  $d/dt$  interchangeably for time differentials). This continuous description serves for the analysis only, and reflects the behavior of the protocol in the limit  $n \rightarrow \infty$ .<sup>1</sup>

**Warmup: the RPS oscillator.** Our oscillatory dynamics may be seen as an extension of the rock-paper-scissors (RPS) protocol (see Related work). This is a protocol with three states  $A_1, A_2, A_3$  and three rules:

$$A_i; \quad A_{i-1} \mapsto A_i \quad \text{with probability } p,$$

where  $p > 0$  is an arbitrarily fixed constant, and the indices of states  $A_i$  are always 1, 2, or 3, and any other values should be treated as  $\text{mod } 3$  in the given range. For  $i \in \{1, 2, 3\}$ , the change of concentration of agents of state  $A_i$  in the population in the given step can be expressed for the RPS protocol as:

$$\Delta a_i = \frac{1}{n} \cdot \Delta^{\#A_i} = \begin{cases} +1/n, & \text{with probability } pa_{i-1}a_i, \\ -1/n, & \text{with probability } pa_i a_{i+1}, \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

Thus, the corresponding continuous dynamics for RPS is given as:

$$\dot{a}_i = n\mathbb{E}(\Delta a_i) = pa_{i-1}a_i - pa_i a_{i+1},$$

for  $i = 1, 2, 3$ . The orbit of motion for this dynamics in  $\mathbb{R}^3$  is given by two constants of motion. First,  $a_1 + a_2 + a_3 = 1$  by normalization. Secondly, for any starting configuration with a strictly positive number of agents in each of the three states, the following function  $\phi$  of the configuration:

$$\phi = \ln(a_1 a_2 a_3) \quad (2)$$

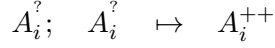
is easily verified to be constant over time  $\dot{\phi} = 0$ , hence  $\phi = \ln(a_1 a_2 a_3) = \text{const} < 0$  (or more simply,  $a_1 a_2 a_3 = \text{const}$ ). Thus, for the continuous dynamics, the initial product of concentrations completely determines its perpetual orbit, which is obtained by intersecting the appropriate curve  $a_1 a_2 a_3 = \text{const}$  with the plane  $a_1 + a_2 + a_3 = 1$ . As a matter of convention, the plane  $a_1 + a_2 + a_3 = 1$  with conditions  $a_i \geq 0$  is drawn as an equilateral triangle (we adopt this convention throughout the paper, for subsequent protocols). All of the orbits are concentric around the point  $(1/3, 1/3, 1/3)$ , which is in itself a point orbit maximizing the value of  $\phi = -\ln 27$ . The discrete dynamics follows a path of motion which typically resembles random-walk-type perturbations around the path of motion, until eventually, after  $\tilde{O}(n^2)$  steps it crashes into one of the sides of the triangle. Subsequently, if  $a_i = 0$ , for some  $i = 1, 2, 3$ , then no rule can make  $a_i$  increase. (If  $a_{i-1} > 0$ , in the next  $O(\log n)$  steps, all remaining agents of  $A_{i+1}$  will convert to  $A_{i-1}$ , and there will be only agents from  $A_{i-1}$  left.) Thus, the protocol will terminate in a corner of the state space.

A further discussion of the RPS dynamics can be found in [13, 25].

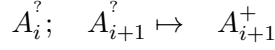
---

<sup>1</sup>We note that some of our results rely on the stochasticity of the random scheduler model, and do not immediately generalize to the continuous case.

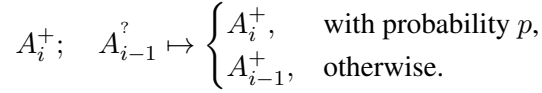
(1) Interaction with an initiator from the same species makes receiver aggressive:



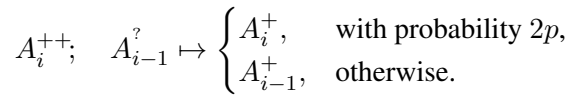
(2) Interaction with an initiator from a different species makes receiver lazy (case of no attack):



(3) A lazy initiator has probability  $p$  of performing a successful attack on its prey:



(4) An aggressive initiator has probability  $2p$  of performing a successful attack on its prey:



(5) The source converts any receiver into a lazy state of a uniformly random species:

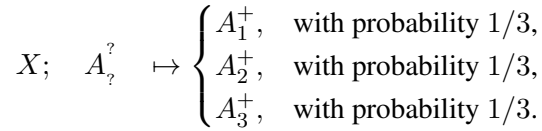


Figure 1: Rules of the basic oscillator protocol  $P_o$ . The adopted notation for initiator-preserving rules is of the form  $A; B \mapsto C$ , corresponding to transitions written as  $A + B \rightarrow A + C$  in the notation of chemical reaction networks or  $(A, B) \rightarrow (A, C)$  in the notation of population dynamics. All rules apply to  $i \in \{1, 2, 3\}$ , whereas a question mark  $?$  in a superscript or subscript denotes a wildcard, matching any permitted combination of characters, which may be set independently for each agent. Probability  $p > 0$  is any (sufficiently small) absolute constant.

### 3 Oscillator Dynamics for Rumor Spreading and Suppression

This section is devoted to the proof of Theorem 1. We start by designing protocol  $P_r$ . It has 7 states. The source of the rumor is denoted by  $X$ . Additionally, there are six states, called  $A_i^+$  and  $A_i^{++}$ , for  $i \in \{1, 2, 3\}$ . The rules of the basic process considered in this paper are presented in Figure 1.

The naming of states in the protocol is intended to maintain a direct connection with the RPS oscillator dynamics. In fact, we will retain the convention  $A_i = \{A_i^+, A_i^{++}\}$  and  $a_i = a_i^+ + a_i^{++}$ , and consider the two states  $A_i^+$  and  $A_i^{++}$  to be different flavors of the same species  $A_i$ , referring to the respective superscripts as either lazy (+) or aggressive (++) .

The RPS dynamics provides the basic oscillator mechanism which is still largely retained in our scenario. Most of the difficulty lies in controlling its operation as a function of the presence or absence of the rumor source. We do this by applying two separate mechanisms. The presence of rumor source  $X$  shifts the oscillator towards an orbit closer to the central orbit  $(A_1, A_2, A_3) = (1/3, 1/3, 1/3)$  through rule (5), which increases the value of potential  $\phi$ . Conversely, independent of the existence of rumor source  $X$ , a second mechanism is intended to reduce the value of potential  $\phi$ . This mechanism exploits the difference between the aggressive and lazy flavors of the species. Following rule (1), an agent belonging to a species becomes more aggressive if it meets another from the same species, and

subsequently attacks agents from its prey species with doubled probability following rule (4). This behavior somehow favors larger species, since they are expected to have (proportionally) more aggressive agents than the smaller species (in which pairwise interactions between agents of the same species are less frequent) — the fraction of agents in  $A_i$  which are aggressive would, in an idealized static scenario, be proportional to  $a_i$ . (This is, in fact, often far from true due to the interactions between the different aspects of the dynamics). As a very loose intuition, the destabilizing behavior of the considered rule on the oscillator is resemblant of the effect an eccentrically fitted weight has on a rotating wheel, pulling the oscillator towards more external orbits (with smaller values of  $\phi$ ).

The intuition for which the proposed dynamics works, and which we will formalize and prove rigorously until the end of the section, can now be stated as follows: in the presence of rumor source  $X$ , the dynamics will converge to a form of orbit on which the two effects, the stabilizing and destabilizing one, eventually compensate each other (in a time-averaged sense). The period of a single rotation of the oscillator around such an orbit is between  $O(1)$  and  $O(\log n)$ , depending on the concentration of  $X$ . In the absence of  $X$ , the destabilizing rule will prevail, and the oscillator will quickly crash into a side of the triangle.

The following Theorem captures the basic properties of the protocol  $P_o$ : it stops in a corner of the state space in the absence of source  $X$ , and moves away from the corners of the state space in the presence of  $X$  (typically entering into oscillatory motion when the concentration of  $X$  is small). From the perspective of the confirmed rumor spreading problem, an easy decoding of the output of such a system is possible using the extensions described in Section 4.

**Theorem 3.** *There exists a universal protocol  $P_o$  with  $|K| = 7$  states, including a distinguished source state  $X$ , which has the following properties.*

1. *For any starting configuration, in the absence of the source ( $\#X = 0$ ), the protocol always reaches a configuration such that:*
  - *all agents are in the same state;*
  - *no further state transitions occur after this time.*

*Such a configuration is reached in  $O(\log n)$  parallel rounds, with constant probability.*

2. *For any starting configuration, in the presence of the source ( $\#X \geq 1$ ), we have with probability  $1 - O(1/n)$ :*
  - *for each state  $i \in K$ , there exists a time step in the next  $O(\log n)$  parallel rounds when at least a constant fraction of all agents are in state  $i$ ;*
  - *during the next  $O(\log n)$  parallel rounds, at least a constant fraction of all agents change their state at least once.*

The rest of the section is devoted to the proof of Theorem 3. We start by noting some basic properties in Subsection 3.1, then prove the properties of the protocol for the case of  $X = 0$  (Subsection 3.2, and finally analyze (the somewhat less involved) case of  $X > 0$  (Subsection 3.3). For the case of  $X = 0$ , the proof is based on a repeated application of concentration inequalities for several potential functions (applicable in different portions of the 6-dimensional phase space). In two specific regions, in the  $O(1/\sqrt{n})$ -neighborhood of the center of the  $(A_1, A_2, A_3)$ -triangle and very close to its sides, we rely on stochastic noise to “push” the trajectory away from the center of the triangle, and also to push it onto one of its sides. Fortunately, each of these stages takes  $O(\log n)$  parallel rounds, with strictly positive probability. Overall, the  $O(\log n)$  parallel rounds bound for the case of  $X = 0$  is provided with constant probability; this translates into  $O(\log n)$  parallel rounds in expectation, since subsequent executions of the process for  $O(\log n)$  rounds have independently constant success probability, and the process has a geometrically decreasing tail over intervals of length  $O(\log n)$ .

### 3.1 Properties of the Oscillator

In the following, we define  $s = a_1 + a_2 + a_3 \in [0, 1]$ . Handling the case of  $s < 1$  allows us not only to take care of the fact that a fraction of the population may be taken up by rumor source  $X$ , but also allows for easier composition of  $P_o$  with other protocols (sharing the same population). We set  $p$  as a constant value independent of  $n$ , which is sufficiently small (e.g.,  $p = s^2/10^{12}$  is a valid choice; we make no efforts in the proofs to optimize constants, but the protocol appears in simulations to work well with much larger values of  $p$ ).

We will occasionally omit an explanation of index  $i$ , which will then implicitly mean “for all  $i = 1, 2, 3$ ”.

From the definition of the protocol one obtains the distribution of changes of the sizes of states in a step:

$$\Delta a_i = \begin{cases} +1/n, & \text{with probability } \frac{1}{3}x(s - a_i) + pa_i^+ a_{i-1} + 2pa_i^{++} a_{i-1}, \\ -1/n, & \text{with probability } \frac{2}{3}xa_i + pa_i a_{i+1}^+ + 2pa_{i+1}^{++} a_i, \\ 0, & \text{otherwise.} \end{cases}$$

$$\Delta a_i^{++} = \begin{cases} +1/n, & \text{with probability } a_i(a_i - a_i^{++}), \\ -1/n, & \text{with probability } xa_i^{++} + (s - a_i)a_i^{++}, \\ 0, & \text{otherwise.} \end{cases}$$

Taking the expectations of the above random variables, and recalling that  $a_i = a_i^+ + a_i^{++}$ , we obtain:

$$\begin{aligned} \dot{a}_i &= x(s/3 - a_i) + pa_i^+ a_{i-1} + 2pa_i^{++} a_{i-1} - pa_i a_{i+1}^+ - 2pa_{i+1}^{++} a_i \\ &= x(s/3 - a_i) + pa_{i-1}(a_i + a_i^{++}) - pa_i(a_{i+1} + a_{i+1}^{++}) \end{aligned} \quad (3)$$

$$\dot{a}_i^{++} = -xa_i^{++} + a_i(a_i - a_i^{++}) - (s - a_i)a_i^{++} = -xa_i^{++} + a_i^2 - sa_i^{++}. \quad (4)$$

Moreover, we have by a simple transformation:

$$\dot{\phi} = \sum \frac{\dot{a}_i}{a_i} = \frac{x}{3} \left( \left( \sum \frac{s}{a_i} \right) - 9 \right) + p \left( \sum a_i^{++} \left( \frac{a_{i-1}}{a_i} - 1 \right) \right) \quad (5)$$

### 3.2 Stopping in $O(n \log n)$ Sequential Steps in the Absence of a Source

Throughout this subsection we assume that  $x = 0$ . We consider first the case where  $a_i \neq 0$ , for  $i = 1, 2, 3$  (noting that as soon as  $a_i = 0$ , we can easily predict the subsequent behavior of the oscillator, as was the case for the RPS dynamics).

The dynamics of  $P_o$  is defined in such a way that that when  $x = 0$  and in the absence of the rules of the RPS oscillator, the value of  $a_i^{++}$  would be close to  $\frac{a_i^2}{s}$ . Consequently, we define  $\kappa_i$ ,  $i = 1, 2, 3$  as the appropriate normalized corrective factor:

$$\kappa_i = s \frac{a_i^{++}}{a_i} - a_i \quad \text{thus} \quad a_i^{++} = \frac{a_i}{s} (a_i + \kappa_i).$$

Note that as  $0 \leq a_i^{++} \leq a_i \leq 1$ , then  $-1 \leq \kappa_i \leq 1$ . Next, we introduce the following definitions:

$$\begin{aligned} \delta_i &= a_i - a_{i-1} \\ \delta &= \sqrt{\delta_1^2 + \delta_2^2 + \delta_3^2} \\ \kappa &= \sqrt{\kappa_1^2 + \kappa_2^2 + \kappa_3^2} \end{aligned}$$

We also reuse potential  $\phi$  from the original RPS oscillator. This time, it is no longer a constant of motion. By (5) and the definition of  $\kappa_i$ , for  $x = 0$  we upper-bound  $\dot{\phi}$  as:

$$\begin{aligned}
\dot{\phi} &= \frac{p}{s} \left( \sum a_i (a_i + \kappa_i) \left( \frac{a_{i-1}}{a_i} - 1 \right) \right) \\
&= \frac{p}{s} \left( \sum (a_i + \kappa_i) (a_{i-1} - a_i) \right) \\
&= \frac{p}{s} \left( -\frac{1}{2} \sum (a_i - a_{i-1})^2 + \sum \kappa_i (a_{i-1} - a_i) \right) \\
&\leq \frac{p}{s} \left( -\frac{1}{2} \delta^2 + \kappa \delta \right)
\end{aligned} \tag{6}$$

The above change  $\dot{\phi}$  of the potential is indeed negative when  $\kappa \approx 0$  (which is in accordance with our intention in designing the destabilizing rules for the oscillator).

The functions  $\delta$ ,  $\phi$  and  $\kappa$  are intricately dependent on each other. In general, we will try to show that  $\delta$  and  $\phi$  increase over time, while  $\kappa$  stays close to 0. This requires that we first introduce a number of auxiliary potentials based on these two functions.

First, for  $x = 0$ , we can rewrite (4) as:

$$a_i^{\dot{+}+} = a_i^2 - s a_i^{++} = -a_i \kappa_i.$$

Next, introducing the definition of  $\kappa_i$  to (3), we obtain for  $x = 0$ :

$$\begin{aligned}
\dot{a}_i &= p a_{i-1} (a_i + a_i^{++}) - p a_i (a_{i+1} + a_{i+1}^{++}) = p a_i (a_{i-1} - a_{i+1}) + p (a_i^{++} a_{i-1} - a_{i+1}^{++} a_i) = \\
&= p a_i \delta_{i-1} + p \left( \frac{1}{s} a_i (a_i + \kappa_i) a_{i-1} - \frac{1}{s} a_{i+1} (a_{i+1} + \kappa_{i+1}) a_i \right) = \\
&= p a_i \delta_{i-1} + \frac{p}{s} a_i (\kappa_i a_{i-1} - \kappa_{i+1} a_{i+1} + (a_{i-1} a_i - a_{i+1}^2)).
\end{aligned}$$

From the above, an upper bound on  $|\dot{a}_i|$  follows directly using elementary transformations:

$$\begin{aligned}
|\dot{a}_i| &\leq p a_i |\delta_{i-1}| + \frac{p}{s} a_i (|\kappa_i| a_{i-1} + |\kappa_{i+1}| a_{i+1} + |a_{i-1} a_i - a_{i+1}^2|) \leq \\
&\leq p a_i \delta + \frac{p}{s} a_i (\kappa (a_{i-1} + a_{i+1}) + |a_{i-1} a_i - a_{i+1}^2|) = \\
&= p a_i \delta + p a_i \kappa \frac{a_{i-1} + a_{i+1}}{s} + \frac{p}{s} a_i |(a_{i-1} - a_{i+1}) a_i - (a_{i+1} - a_i) a_{i+1}| \leq \\
&\leq p a_i (\delta + \kappa) + \frac{p}{s} a_i (|a_{i-1} - a_{i+1}| a_i + |a_{i+1} - a_i| a_{i+1}) \leq \\
&\leq p a_i (\delta + \kappa) + p a_i \delta \frac{a_i + a_{i+1}}{s} \\
&\leq p a_i (2\delta + \kappa).
\end{aligned}$$

We are now ready to estimate  $\dot{\kappa}_i$  for  $x = 0$ , using the definition of  $\kappa_i$  and the previously obtained formula for  $\dot{a}_i^{++}$ :

$$\begin{aligned}
\dot{\kappa}_i &= s \left( \frac{\dot{a}_i^{++}}{a_i} - \frac{a_i^{++}}{a_i^2} \dot{a}_i \right) - \dot{a}_i = -s \kappa_i - \left( s \frac{a_i^{++}}{a_i^2} + 1 \right) \dot{a}_i \\
&= -s \kappa_i - \left( s \frac{a_i^{++}}{a_i} - a_i \right) \frac{\dot{a}_i}{a_i} - 2 \dot{a}_i = -s \kappa_i - \frac{\dot{a}_i}{a_i} \kappa_i - 2 \dot{a}_i.
\end{aligned}$$

Next from the bound on  $|\dot{a}_i|$ :

$$\begin{aligned}
\dot{\kappa}_i^2 &= 2\kappa_i \dot{\kappa}_i = 2\left(-s\kappa_i^2 - \frac{\dot{a}_i}{a_i}\kappa_i^2 - 2\dot{a}_i\kappa_i\right) \\
&\leq 2\left(-s\kappa_i^2 + \frac{|\dot{a}_i|}{a_i}\kappa_i^2 + 2|\dot{a}_i|\kappa_i\right) \\
&\leq 2\left(-s\kappa_i^2 + p(2\delta + \kappa)(\kappa_i^2 + 2a_i|\kappa_i|)\right) \\
&\leq 2\left(-s\kappa_i^2 + p(2\delta + \kappa)(\kappa + 2\kappa)\right) \\
&= -2s\kappa_i^2 + p(12\delta\kappa + 6\kappa^2).
\end{aligned}$$

Next:

$$\begin{aligned}
\dot{\kappa} &= \frac{1}{2\kappa} \sum \dot{\kappa}_i^2 \leq \frac{1}{2\kappa} \left(-2s \sum \kappa_i^2 + 3p(12\delta\kappa + 6\kappa^2)\right) \leq \\
&\leq \frac{1}{2\kappa} \left(-2s\kappa^2 + p(36\delta\kappa + 18\kappa^2)\right) = \\
&= (-s + 9p)\kappa + 18p\delta \leq \\
&\leq -\frac{s}{2}\kappa + 18p\delta,
\end{aligned} \tag{7}$$

where in the final transformation we took into account that  $p \leq s/18$ .

Now, we define the potential  $\eta$  for any configuration with all  $a_i > 0$  as:

$$\eta = \left(\ln \frac{s^3}{27} - \phi\right)^{1/2} = \left(-\sum_{i=1}^3 \ln \frac{a_i}{s/3}\right)^{1/2}.$$

We remark that  $\eta$  is always well-defined when  $\min_i a_i > 0$ , and that  $\eta \geq 0$ .

**Overview of the proof.** The proof for the case of  $X = 0$  proceeds by following the trajectory of the discrete dynamics of  $P_o$ , divided into a number of stages. We define a series of time steps  $t_0, t_1, \dots, t_7$  by conditions on the configuration met at time  $t_i$ , and show that subject to these conditions holding, we have  $t_{j+1} \leq t_j + O(n \log n)$  (we recall that here time is measured in steps), with at least constant probability. Overall, it follows that the configuration at time  $t_7$ , which corresponds to having reached a corner state, is reached from  $t_0$ , which is any initial configuration with  $X = 0$ , in  $O(n \log n)$  time steps, with constant probability.

The intermediate time steps may be schematically described as follows (see Fig. 2). For configurations which start close to the center of the triangle ( $\delta \leq s/12$ ), we define a pair of potentials  $\psi^{(1)}, \psi^{(2)}$ , based on a linear combination of modified versions of  $\eta$  and  $\kappa$ . The dynamics will eventually escape from the area  $\delta \leq s/12$ ; however, first it may potentially reach a very small area of radius  $O(1/\sqrt{n})$  around the center of the triangle with  $\kappa \approx 0$  (Lemma 3, time  $t_1$ , reached in  $O(n \log n)$  steps by a multiplicative drift analysis on potential  $\psi^{(2)} < 0$ ), pass through the vicinity of center of the triangle, escaping it with  $\kappa \approx 0$  (Lemma 4, time  $t_2$ , reached in  $O(n \log n)$  steps with constant probability by a protocol-specific analysis of the scheduler noise, which with constant probability increases  $\eta$  without increasing  $\kappa$  too much), and eventually escapes completely to the area of  $\delta > s/12$  (Lemma 5, exponentially increasing value of potential  $\psi^{(1)} > 0$ ).

In the area of  $\delta > s/12$ , we define a new potential  $\psi$  based on  $\phi$  and  $\kappa$ . This increases (Lemma 8, additive drift analysis on  $\psi$  with bounded variance) until a configuration at time  $t_4$  with a constant number of agents of some species  $A_i$  is reached. This configuration then evolves towards a configuration at time  $t_5$  at which some species has  $O(1)$  agents, and additionally its predator species is a constant part

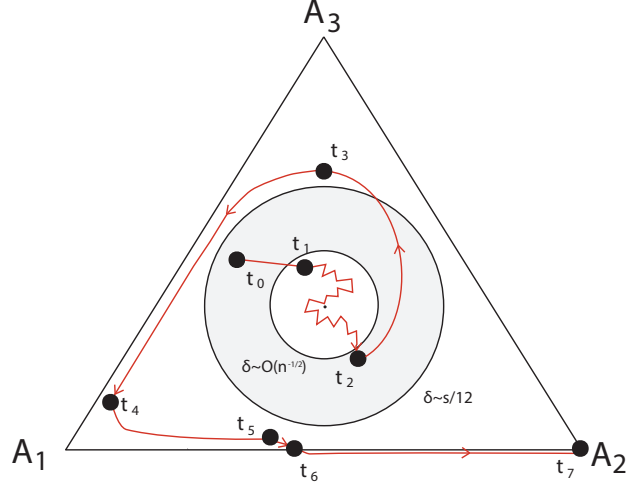


Figure 2: Schematic illustration of order of phases in the proof of stabilization of protocol  $P_0$  for  $X = 0$ .

of the population (Lemma 9, direct analysis of the process combined with analysis of potential  $\psi$  and a geometric drift argument). Then, the species with  $O(1)$  agents is eliminated in  $O(n)$  steps with constant probability ( $t_6$ , Lemma 10), and finally one more species is eliminated in another  $O(n \log n)$  steps (at time  $t_7$ , Lemma 11, straightforward analysis of the dynamics). At this point, the dynamics has reached a corner.

Throughout the proof, we make sure to define boundary conditions on the analyzed cases to make sure that the process does not fall back to a previously considered case with probability  $1 - o(1)$ .

**Phase with  $\delta \leq s/12$ .** We then have  $a_i \in [3s/12, 5s/12]$  and  $\frac{a_i}{s/3} \in [3/4, 5/4]$ , for  $i = 1, 2, 3$ . In this range, we have:

$$\frac{1}{3} \left( \frac{a_i}{s/3} - 1 \right)^2 < \left( \frac{a_i}{s/3} - 1 \right) - \ln \frac{a_i}{s/3} < \frac{3}{4} \left( \frac{a_i}{s/3} - 1 \right)^2.$$

Summing the above inequalities for  $i = 1, 2, 3$  and noting that  $\sum_{i=1}^3 \left( \frac{a_i}{s/3} - 1 \right) = 0$ , we obtain:

$$\frac{1}{3} \sum_{i=1}^3 \left( \frac{a_i}{s/3} - 1 \right)^2 < \eta^2 < \frac{3}{4} \sum_{i=1}^3 \left( \frac{a_i}{s/3} - 1 \right)^2.$$

Next, we have:

$$\sum_{i=1}^3 \left( \frac{a_i}{s/3} - 1 \right)^2 = \left( \frac{3}{s} \right)^2 \sum_{i=1}^3 (a_i - s/3)^2 = \frac{3\delta^2}{s^2}.$$

Combining the two above expressions gives the sought bound between  $\eta$  and  $\delta$  as:

$$\frac{\delta}{s} < \eta < \frac{3}{2} \frac{\delta}{s}$$

and equivalently

$$\delta \in \left( \frac{2}{3} s \eta, s \eta \right).$$

We have directly from (6) and from the relations between  $\eta$  and  $\delta$ :

$$\dot{\eta} = -\frac{\dot{\phi}}{2\eta} \geq \frac{p}{2s\eta} \left( \frac{1}{2} \delta^2 - \kappa \delta \right) = \frac{p}{4s\eta} \delta^2 - \frac{p}{2s\eta} \kappa \delta \geq \frac{ps}{9} \eta - \frac{p}{2} \kappa, \quad (8)$$

and from (7):

$$\dot{\kappa} \leq -\frac{s}{2} \kappa + 18p\delta \leq -\frac{s}{2} \kappa + 18ps\eta. \quad (9)$$

Moving to the discrete-time model, it is advantageous to eliminate the discontinuity of partial derivatives of  $\eta$  and  $\kappa$  at points with  $\eta = 0$  and  $\kappa = 0$  respectively, which is a side-effect of the applied square root transformation in the respective definitions of  $\eta$  and  $\kappa$ . We define the auxiliary functions  $\eta^*$  and  $\kappa^*$  by adding an appropriate corrective factor:

$$\eta^* = \sqrt{\eta^2 + \frac{1}{n}}$$

$$\kappa^* = \sqrt{\kappa^2 + \frac{1}{n}}$$

and derive accordingly from (8) and (9):

$$\dot{\eta}^* = \frac{\eta}{\eta^*} \dot{\eta} \geq \frac{ps}{9} \left( \eta - \frac{1}{\sqrt{n}} \right) - \frac{p}{2} \kappa \geq \frac{ps}{9} \eta^* - \frac{p}{2} \kappa^* - \frac{2ps}{9\sqrt{n}} \quad (10)$$

$$\dot{\kappa}^* = \frac{\kappa}{\kappa^*} \dot{\kappa} \leq -\frac{s}{2} \left( \kappa - \frac{1}{\sqrt{n}} \right) + 18ps \eta \leq -\frac{s}{2} \kappa^* + 18ps \eta^* + \frac{s}{\sqrt{n}}. \quad (11)$$

Let  $u$  be the 5-dimensional vector representing the current configuration of the system:  $u := (a_1^+, a_1^{++}, a_2^+, a_2^{++}, a_3^+) \equiv (u^{(1)}, \dots, u^{(5)})$ ; note that the last element  $a_3^{++}$  is determined as  $a_3^{++} = s - \sum_{i=1}^5 u^{(i)}$ .<sup>2</sup> The following lemma is obtained by a folklore application of Taylor's theorem.

**Lemma 1.** *Let  $f : \mathbb{R}^5 \rightarrow \mathbb{R}$  be a  $C^2$  function in a sufficiently large neighborhood of  $u$ , with  $\min_{1 \leq i \leq 5} u^{(i)} \geq 2/n$ . Then,  $|\mathbb{E} \Delta f(u) - \dot{f}| \leq \frac{2}{n^2} \max_{\|u^* - u\|_\infty \leq 1/n} D_f(u^*)$ , where  $D_f(u^*) := \max_{1 \leq i, j \leq 5} \left| \frac{\partial^2 f(u^*)}{\partial u^{(i)} \partial u^{(j)}} \right|$ .*

*Proof.* Let  $u'$  be the random variable representing the configuration of the system after its next transition from configuration  $u$ . Observe that in every non-idle step of execution of the protocol, exactly one agent changes its state, so  $\|u' - u\|_\infty \leq 1/n$ .

Applying Taylor approximation we have:

$$\begin{aligned} \mathbb{E}(\Delta f) &= \mathbb{E}(f(u')|u) - f(u) = \sum_{u'} (f(u') - f(u)) \Pr(u'|u) = \sum_{u'} (\nabla f(u) \cdot (u' - u) + R_2(u, u')) \Pr(u'|u) = \\ &= \nabla f(u) \cdot \sum_{u'} (u' - u) \Pr(u'|u) + R_2(u) = \nabla f(u) \cdot \frac{1}{n} (\dot{u}^{(1)}, \dots, \dot{u}^{(5)})^T + R_2(u) = \dot{f} + R_2(u), \end{aligned} \quad (12)$$

where  $\nabla f(u)$  is the gradient of  $f$  at  $u$ ,  $R_2(u, u') \in \mathbb{R}$  denotes the second-order Taylor remainder for function  $f$  expanded at point  $u$  along the vector towards point  $u'$ , and  $R_2(u) \in \mathbb{R}$  is subsequently an appropriately chosen value, satisfying:

$$|R_2(u)| \leq \frac{1}{n^2} \max_{\|u^* - u\|_\infty \leq 1/n} D_f(u^*).$$

□

The following lemma is obtained directly by computing and bounding all second order partial derivatives of functions  $\eta^*$  and  $\kappa^*$  with respect to variables  $(u^{(1)}, \dots, u^{(5)})$ .

<sup>2</sup>In principle it is also correct to represent  $u$  as a vector of dimension 6, i.e., including  $a_3^{++}$  in  $u$  as a free dimension. However, such a representation would lead to second-order partial derivatives  $\frac{\partial^2}{\partial u^{(i)} \partial u^{(j)}} \eta^*(u)$  which are too large for our purposes.



**Lemma 2.** *There exists a constant  $c_1 > 1$  depending only on  $s, p$ , such that, for any configuration  $u$  with  $\delta(u) \leq s/12$ :*

- $\max_{\|u^* - u\|_\infty \leq 1/n} D_{\eta^*}(u^*) < c_1 \sqrt{n}$ ,
- $\max_{\|u^* - u\|_\infty \leq 1/n} D_{\kappa^*}(u^*) < c_1 \sqrt{n}$ .

□

In view of the above lemmas, we obtain from (8) and (9), for an appropriately chosen constant  $c_2 = 2c_1 + s$ :

$$\begin{cases} \mathbb{E}\Delta\eta^* \geq \frac{1}{n} \left( \frac{ps}{18}\eta^* - \frac{p}{2}\kappa^* - \frac{c_2}{\sqrt{n}} \right), & \text{when } \delta \leq s/12, \\ \mathbb{E}\Delta\kappa^* \leq \frac{1}{n} \left( -\frac{s}{3}\kappa^* + 18ps\eta^* + \frac{c_2}{\sqrt{n}} \right), & \text{when } \delta \leq s/12. \end{cases}$$

For  $j = 1, 2$ , we now define two linear combinations of functions  $\eta^*$  and  $\kappa^*$ :

$$\psi^{(j)} = \eta^* - \frac{3jp}{s}\kappa^*.$$

When  $\delta \leq s/12$ , we have:

$$\begin{aligned} \mathbb{E}\Delta\psi^{(j)} &\geq \frac{1}{n} \left( \frac{ps}{18}\eta^* - \frac{p}{2}\kappa^* - \frac{c_2}{\sqrt{n}} + jp\kappa^* - 54jp^2\eta^* - \frac{3jp}{s}\frac{c_2}{\sqrt{n}} \right) \geq \frac{1}{n} \left( \frac{ps}{24}\eta^* + \frac{jp}{2}\kappa^* - \frac{2c_2}{\sqrt{n}} \right) \geq \\ &\geq \frac{ps}{24n} \left( \eta^* + \frac{3jp}{s}\kappa^* - \frac{48c_2}{ps\sqrt{n}} \right) \geq \frac{ps}{24n} \left( |\psi^{(j)}| - \frac{c_3}{\sqrt{n}} \right), \end{aligned}$$

where we denoted  $c_3 := \frac{48c_2}{ps}$  and used the fact that  $p < \frac{s}{72 \cdot 54 \cdot 2}$ .

We subsequently perform an analysis of  $\psi_t^{(j)} = \psi^{(j)}(u_t)$ ,  $j = 1, 2$ , treating them as stochastic processes. We remark that  $\psi_t^{(2)} \leq \psi_t^{(1)}$ , since  $\psi_t^{(1)} - \psi_t^{(2)} = \frac{3p}{s}\kappa^* \geq 0$ .

**Lemma 3.** *Let  $u_{t_0}$  be an arbitrary starting configuration of the system. Then, with constant probability, for some  $t_1 = t_0 + O(n \log n)$ , a configuration  $u_{t_1}$  is reached such that  $\psi_{t_1}^{(1)} \geq \psi_{t_1}^{(2)} \geq -\frac{2c_3}{\sqrt{n}}$ .*

*Proof.* W.l.o.g. assume  $t_0 = 0$ . We subsequently only analyze process  $\psi_t^{(2)}$ . Let  $t_1$  be the first time step such that  $\psi_{t_1}^{(2)} > -\frac{2c_3}{\sqrt{n}}$ . If  $t_1 \neq 0$ , then  $\psi_0^{(2)} < 0$ . Note that then  $\psi_t^{(2)} < 0$  for all  $t \leq t_1$ , from which it follows by a straightforward calculation from the definition of  $\psi$ ,  $\kappa$ , and  $\eta$ , that  $\delta_t < \frac{s}{12}$  for all  $t \leq t_1$ .

We now define the filtered stochastic process  $\psi_t^{*(2)}$  as  $\psi_t^{*(2)} := |\psi_t^{(2)}|$  for  $t < t_1$ , and put  $\Delta\psi_t^{*(2)} := 0$  for  $t \geq t_1$ . For all  $t \geq 0$ , we then have:

$$\mathbb{E}(\Delta\psi_t^{*(2)} | \psi_t^{*(2)} \neq 0) \leq \frac{ps}{48n} \psi_t^{*(2)}.$$

Since  $0 \leq \psi_t^{*(2)} < 9$  for all time steps, a direct application of multiplicative drift analysis (cf. [18]) gives:

$$\mathbb{E}t_1 \leq \frac{48n}{ps} \left( 1 + \ln \frac{9\sqrt{n}}{2c_3} \right),$$

and the claim follows by Markov's inequality. □

**Lemma 4.** *Let  $u_{t_1}$  be an arbitrary starting configuration of the system such that  $\psi_{t_1}^{(j)} \in [-\frac{2c_3}{\sqrt{n}}, \frac{4c_3}{\sqrt{n}}]$ , for  $j = 1, 2$ . Then, with constant probability, for some  $t_2 = t_1 + O(n)$ , a configuration  $u_{t_2}$  is reached such that  $\psi_{t_2}^{(1)} \geq \frac{4c_3}{\sqrt{n}}$ .*

*Proof.* W.l.o.g. assume that  $t_1 = 0$  and suppose that initially  $\psi_0^{(2)} \leq \psi_0^{(1)} < \frac{4c_3}{\sqrt{n}}$  (i.e., that  $t_2 \neq t_1$ ). Then, from the lower and upper bounds on  $\psi_0^{(1)}$  and  $\psi_0^{(2)}$  we obtain the following bounds on  $\kappa_0$  and  $\delta_0$ :

$$\begin{aligned} \frac{3p}{s} \kappa_0^* = \psi_0^{(1)} - \psi_0^{(2)} &\leq \frac{2c_3}{\sqrt{n}} + \frac{4c_3}{\sqrt{n}} \implies \kappa_0 \leq \frac{2c_3 s}{p\sqrt{n}}, \\ \eta_0^* = 2\psi_0^{(1)} - \psi_0^{(2)} &\leq 2\frac{4c_3}{\sqrt{n}} + \frac{2c_3}{\sqrt{n}} \implies \eta_0 \leq \frac{10c_3}{\sqrt{n}} \implies \delta_0 < \frac{10c_3 s}{\sqrt{n}}. \end{aligned}$$

It follows that, for  $i = 1, 2, 3$ ,  $a_{i,0} \in [\frac{s}{3} - \frac{10c_3 s}{\sqrt{n}}, \frac{s}{3} + \frac{10c_3 s}{\sqrt{n}}]$  and  $a_{i,0}^{++} = \frac{a_{i,0}}{s}(a_{i,0} + \kappa_{i,0}) \in [(\frac{1}{3} - \frac{10c_3}{\sqrt{n}})(\frac{s}{3} - \frac{10c_3 s}{\sqrt{n}} - \frac{2c_3 s}{p\sqrt{n}}), (\frac{1}{3} + \frac{10c_3}{\sqrt{n}})(\frac{s}{3} + \frac{10c_3 s}{\sqrt{n}} + \frac{2c_3 s}{p\sqrt{n}})]$ . For the sake of clarity of notation, we will simply write  $a_{i,0} = s/3 \pm O(1/\sqrt{n})$  and  $a_{i,0}^{++} = s/9 \pm O(1/\sqrt{n})$ , hence also  $a_{i,0}^+ = 2s/9 \pm O(1/\sqrt{n})$ .

We will consider now the sequence of exactly  $n$  transitions of the protocol, between time steps  $t = 0, 1, \dots, n$ .

For all  $t$  we have  $\mathbb{E}\Delta\psi_t^{(2)} \geq -\frac{c_3 p s}{24n^{3/2}}$ . Consider the Doob submartingale  $Y_t = \sum_{\tau=0}^{t-1} X_\tau$  with increments  $(X_t)$  given as:

$$X_t = \begin{cases} \Delta\psi_t^{(2)} + \frac{c_3 p s}{24n^{3/2}}, & \text{if } Y_t > -\frac{c_3}{\sqrt{n}} \\ 0, & \text{otherwise,} \end{cases}$$

Noting that  $|X_t| \leq \frac{9}{n}$ , an application of the Azuma inequality for submartingales to  $(Y_n)$  gives:  $\Pr[Y_n \leq -\frac{c_3}{\sqrt{n}}] \leq \exp[-c_3^2/162]$  (cf. e.g. [11][Thm. 16]). From here it follows directly that:

$$\Pr\left[\psi_n^{(2)} > -\frac{c_3}{\sqrt{n}} + \psi_0^{(2)} - n\frac{c_3 p s}{24n^{3/2}}\right] \geq 1 - \exp[-c_3^2/162] > 1/2.$$

Noting that  $\psi_0^{(2)} \geq -\frac{2c_3}{\sqrt{n}}$ , we have:

$$\Pr\left[\psi_n^{(2)} \geq -\frac{2c_3}{\sqrt{n}}\right] > 1/2. \quad (13)$$

We now describe the execution of transitions in the protocol for times  $t = 0, 1, \dots, n-1$  through the following coupling. First, we select the sequence of pairs of agents chosen by the scheduler. Let  $V_2^+$  (respectively,  $V_1^+$ ) denote the subsets of the set of  $n$  agents, having initial states  $A_2^+$  (resp.,  $A_1^+$ ) at time 0, respectively, which are involved in exactly one transition in the considered time interval, acting in it as the initiator (resp., receiver). Let  $S \subseteq \{0, 1, \dots, n-1\}$  denotes the subset of time steps at which the scheduler activates a transition involving an element of  $V_2^+$  as the initiator and an element of  $V_1^+$  as the receiver. The execution of the protocol is now given by:

- Phase  $P_A$ : Selecting the sequence of pairs of elements activated by the scheduler in time steps  $(0, 1, \dots, n-1)$ . This also defines set  $S$ . Executing the rules of the protocol in their usual order for time steps from set  $\{0, 1, \dots, n-1\} \setminus S$ .
- Phase  $P_B$ : Executing the rules of the protocol for time steps from set  $S$ .

Observe that since elements of pairs activated in time steps from  $S$  are activated only once throughout the  $n$  steps of the protocol, the above probabilistic coupling does not affect the distribution of outcomes.

Directly from (13), we obtain through a standard bound on conditional probabilities that at least a constant fraction of choices made in phase  $P_A$  leads to an outcome “ $\psi_n^{(2)} \geq -\frac{2c_3}{\sqrt{n}}$ ” with at least constant probability during phase  $P_B$ :

$$\Pr\left[P_A : \Pr\left[\psi_n^{(2)} \geq -\frac{2c_3}{\sqrt{n}} \mid P_A\right] > 1/4\right] \geq 1/3. \quad (14)$$

We now remark on the size of set  $S$ . The distribution of  $|S|$  depends only on  $a_{2,0}^+$ ,  $a_{1,0}^+$ , and the choices made by the random scheduler. We recall that  $a_{2,0}^+ = 2s/9 \pm O(1/\sqrt{n})$ . Since the expected number of isolated edges in a random multigraph on  $n$  nodes (representing the set of agents) and  $n$  edges (representing the set of time steps) is  $(1 \pm o(1))e^{-4}n$ , the number of such edges having the first endpoint in an agent in state  $A_2^+$  and the second endpoint in an agent in state  $A_1^+$  is  $(1 \pm o(1))\frac{4e^{-4}s^2}{81}n$ . A straightforward concentration analysis (using, e.g., the asymptotic correspondence between  $G(n, m)$  and  $G(n, p)$  random graph models and an application of Azuma's inequality for functions of independent random variables) shows that the bound  $|S| = (1 \pm o(1))\frac{4e^{-4}s^2}{81}n$  holds with very high probability. In particular, we have:

$$\Pr[|S| > c_4 n] = 1 - e^{\Omega(-n)}, \quad (15)$$

for some choice of constant  $c_4$  which depends only on  $s$ .

Relations (14) and (15) provide all the necessary information about phase  $P_A$  that we need. Subsequently, we will only analyse phase  $P_B$ , conditioning on a fixed execution of phase  $P_A$  such that the following event  $F_A$  holds:

$$\Pr\left[\psi_n^{(2)} \geq -\frac{2c_3}{\sqrt{n}} \mid P_A\right] > 1/4 \quad \wedge \quad |S| > c_4 n. \quad (16)$$

We remark that, by a union bound over (14) and (15),  $\Pr[F_A] \geq 1/3 - e^{\Omega(-n)} > 1/4$ .

In the remainder of our proof, our objective will be to show that:

$$\Pr\left[\psi_n^{(1)} \geq \frac{4c_3}{\sqrt{n}} \mid P_A\right] > c_5,$$

for some constant  $c_5 > 0$  depending only on  $s, p$ , for any choice of  $P_A$  for which event  $F_A$  holds. When this is shown, the claim of the lemma will follow directly, with a probability value given as at least  $c_5 \Pr[F_A] > c_5/4$  by the law of total probability.

We now proceed to analyze the random choices made during phase  $P_B$ . Each of the considered  $|S|$  interactions involves a pair of agents of the form  $(A_2^+, A_1^+)$ , and describes the following transition:

$$(A_2^+, A_1^+) \rightarrow \begin{cases} (A_2^+, A_2^+), & \text{with probability } p, \\ (A_2^+, A_2^+), & \text{with probability } 1 - p, \end{cases}$$

independently at random for each transition. The only state changes observed during this phase are from  $A_1^+$  to  $A_2^+$ , and we denote by  $B$  the number of such state changes. The value of random variable  $B$  completely describes the outcome of phase  $P_B$ .

We have  $\mathbb{E}B = p|S|$ , and by a standard additive Chernoff bound:

$$\Pr[|B - p|S|| \leq 2\sqrt{n} \mid P_A] \geq 1 - 2e^{-4} > 7/8. \quad (17)$$

Let  $\mathcal{B} \subseteq [p|S| - 2\sqrt{n}, p|S| + 2\sqrt{n}]$  be the subset of the considered interval containing values of  $B$  such that  $\left(\psi_n^{(1)} \mid P_A, B \in \mathcal{B}\right) \geq \frac{4c_3}{\sqrt{n}}$ . If  $\Pr[B \in \mathcal{B} \mid P_A] \geq 1/8$ , then the claim follows directly.

Otherwise, it follows from (16) and (17) that there must exist a value  $b \in [p|S| - 2\sqrt{n}, p|S| + 2\sqrt{n}] \setminus \mathcal{B}$ , such that:

$$\left(\psi_n^{(2)} \mid P_A, B = b\right) \geq -\frac{2c_3}{\sqrt{n}}.$$

Given that:

$$\left(\psi_n^{(1)} \mid P_A, B = b\right) \leq \frac{4c_3}{\sqrt{n}}.$$

and recalling that  $\psi_n^{(2)} \leq \psi_n^{(1)}$ , we obtain the following bound on  $\eta_n$ :

$$(\eta_n^* | P_A, B = b) = \left( 2\psi_n^{(1)} - \psi_n^{(2)} | P_A, B = b \right) \leq 2\frac{4c_3}{\sqrt{n}} + \frac{2c_3}{\sqrt{n}} = \frac{10c_3}{\sqrt{n}}. \quad (18)$$

We now consider lower bounds on the value of  $\psi_n^{(2)}$ , conditioned on  $P_A, B = b^+$  (respectively,  $P_A, B = b^-$ ), where  $b^+$  (resp.,  $b^-$ ) is a value arbitrarily fixed in the range  $b^+ \in [b + \frac{20c_3s}{\sqrt{n}}, b + \frac{21c_3s}{\sqrt{n}}]$  (resp.,  $b^- \in [b - \frac{21c_3s}{\sqrt{n}}, b - \frac{20c_3s}{\sqrt{n}}]$ ). The executions of the protocol with  $B = b^+$  and  $B = b^-$  differ with respect to the execution with  $B = b$  in the number of executed transitions from  $a_1^+$  to  $a_2^+$  by at least  $\frac{20c_3s}{\sqrt{n}}$ . Recalling that  $\delta_2 = a_2 - a_1$ , it follows that for some  $b' \in \{b^+, b^-\}$  we have after  $n$  steps:

$$(\delta_n | P_A, B = b') \geq (|\delta_{2,n}| | P_A, B = b') \geq \frac{20c_3s}{\sqrt{n}}.$$

Subsequently, we will assume that  $b' = b^+$ ; the case of  $b' = b^-$  is handled analogously. From the relation  $\eta > \delta/s$  and (18) we have:

$$(\eta_n^* | P_A, B = b^+) \geq \frac{20c_3}{\sqrt{n}} \geq (\eta_n^* | P_A, B = b) + \frac{10c_3}{\sqrt{n}}. \quad (19)$$

When comparing the value of  $\kappa_n^*$  in the two cases,  $B = b^+$  and  $B = b$ , it is convenient to consider  $\kappa^*$  as the length of the vector  $(\kappa_1, \kappa_2, \kappa_3, 1/\sqrt{n})$  in Euclidean space. For each of the coordinates  $\kappa_i$ ,  $i = 1, 2, 3$ , we have:

$$|(\kappa_{i,n} | P_A, B = b^+) - (\kappa_{i,n} | P_A, B = b)| < \frac{40c_3}{\sqrt{n}},$$

hence:

$$(\kappa_n^* | P_A, B = b^+) < (\kappa_n^* | P_A, B = b) + \frac{120c_3}{\sqrt{n}}. \quad (20)$$

Introducing (19) and (20) into the definition of  $\psi_n^{(1)}$ , we obtain directly:

$$\left( \psi_n^{(1)} | P_A, B = b^+ \right) > \left( \psi_n^{(1)} | P_A, B = b \right) + \frac{10c_3}{\sqrt{n}} - \frac{3p}{s} \frac{120c_3}{\sqrt{n}} \geq -\frac{2c_3}{\sqrt{n}} + \frac{10c_3}{n} - \frac{3p}{s} \frac{120c_3}{\sqrt{n}} > \frac{4c_3}{\sqrt{n}},$$

where we again used the fact that  $p$  is a sufficiently small constant w.r.t.  $s$ . We thus obtain:

$$\left( \psi_n^{(1)} | P_A, B \in [b + \frac{20c_3s}{\sqrt{n}}, b + \frac{21c_3s}{\sqrt{n}}] \right) > \frac{4c_3}{\sqrt{n}},$$

where by the definition of random variable  $B$  as a sum of i.i.d. binary random variables and the choice of value  $b$  in the direct vicinity of the expectation of  $B$ , the event  $B \in [b + \frac{20c_3s}{\sqrt{n}}, b + \frac{21c_3s}{\sqrt{n}}]$  holds with constant probability. The case of  $b' = b^-$  is handled analogously.  $\square$

**Lemma 5.** *Let  $u_{t_2}$  be an arbitrary starting configuration of the system such that  $\max\{\psi_{t_2}^{(1)}, \psi_{t_2}^{(2)}\} = \psi_{t_2}^{(1)} \geq \frac{4c_3}{\sqrt{n}}$ . Then, with constant probability, for some  $t_3 = t_2 + O(n \log n)$ , a configuration  $u_{t_3}$  is reached such that  $\delta_{t_3} > s/12$ .*

*Proof.* We subsequently consider only the process  $\psi_t^{(1)}$ . We start by showing the following claim.

*Claim.* Suppose  $\psi_0^{(1)} = A \geq \frac{4c_3}{\sqrt{n}}$ . Then, with probability at least  $1 - \exp[-A^2psn/46656]$ , for some time step  $t \leq \frac{72n}{ps}$  the process reaches a value  $\psi_t^{(1)} \geq 2A$ , or  $\delta_t > s/12$ .

*Proof (of claim).* Consider the Doob submartingale  $Y_t = \sum_{\tau=0}^{t-1} X_t$  with increments  $(X_t)$  given as:

$$X_t = \begin{cases} \Delta\psi_t^{(1)} - \frac{psA}{48n}, & \text{if } Y_t > \frac{A}{2} \text{ or } \delta_t > s/12, \\ 0, & \text{otherwise.} \end{cases}$$

Noting that  $|X_t| \leq \frac{9}{n}$ , an application of the Azuma inequality for submartingales (cf. e.g. [11][Thm. 16]) to  $(Y_T)$  with  $T = \frac{72n}{ps}$  gives:

$$\Pr[Y_T \leq -\frac{A}{2}] \leq \exp[-A^2psn/46656],$$

Moreover, assuming the barrier  $\delta_t > s/12$  was not reached, we have:

$$\left(\psi_T^{(1)} \mid Y_T > -\frac{A}{2}\right) = \psi_0^{(1)} + \frac{psA}{48n}T + Y_T > A + \frac{psA}{48n} \frac{72n}{ps} - \frac{A}{2} = 2A,$$

which completes the proof of the claim.

We now prove the lemma by iteratively applying the claim over successive intervals of time  $(\tau_0, \tau_1, \dots)$ , such that  $\tau_0 = t_2$  and  $\tau_{i+1}$  is the first time step not before  $\tau_i$  such that  $\psi_{\tau_{i+1}}^{(1)} \geq 2\psi_{\tau_i}^{(1)}$  or  $\delta_{\tau_{i+1}} \geq s/12$ . By the claim, we have:

$$\Pr\left[\tau_{i+1} - \tau_i \leq \frac{72n}{ps}\right] \geq 1 - \exp[-(\psi_{\tau_i}^{(1)})^2psn/46656].$$

Noting that  $c_3 > 48/(ps)$  by definition, and that before the barrier  $\delta > s/12$  is reached, we have  $\psi_{\tau_i}^{(1)} \geq \frac{4c_3}{\sqrt{n}}2^i \geq \frac{192}{ps\sqrt{n}}2^i$ , we obtain:

$$\Pr\left[\tau_{i+1} - \tau_i \leq \frac{72n}{ps}\right] > 1 - \exp[-4^{i+1}].$$

and further:

$$\Pr\left[\tau_{i+1} \leq \frac{72n}{ps}(i+1)\right] > \prod_{j=0}^i (1 - \exp[-4^{j+1}]) > 1 - \sum_{j=0}^i \exp[-4^{j+1}] > 0.98.$$

In particular, putting  $i = \log_2 n$ ,  $\Pr\left[\tau_i \leq \frac{72n \log_2 n}{ps}\right] > 0.98$ . Since for this value of  $i$ , we must have  $\delta_{\tau_i} \geq s/12$  (since otherwise we would have  $\psi_{\tau_i}^{(1)} = \omega(1)$ , which is impossible), the claim of the lemma follows.  $\square$

**Phase with  $\delta > s/12$ .** The second phase of convergence corresponds to configurations of the system which are sufficiently far from the center point  $(a_1, a_2, a_3) = (s/3, s/3, s/3)$ . Formally, we analyze a variant of potential  $\phi$  (with an additive corrective factor proportional to  $\kappa^2$ ) to show that, starting from a configuration with  $\delta > s/12$ , we will eliminate one of the three populations  $a_1, a_2, a_3$  in  $O(n \log n)$  sequential steps with constant probability, without approaching the center point too closely (a value of  $\delta = \Omega(1)$  will be maintained throughout).

For this part of the analysis, we define the considered potential as:

$$\psi = \eta^2 - \frac{4p}{s^2}\kappa^2 = \ln \frac{s^3}{27} - \phi - \frac{4p}{s^2}\kappa^2, \quad (21)$$

for any configuration  $u$  with  $\min_{i=1,2,3} a_i > 0$ .

We have directly from (6) and (7):

$$\begin{aligned}\dot{\psi} &= -\dot{\phi} - \frac{4p}{s^2} 2\kappa\dot{\kappa} \geq \frac{p}{s} \left( \frac{1}{2}\delta^2 - \kappa\delta + 4\kappa^2 - 144\frac{p}{s}\kappa\delta \right) = \\ &= \frac{p}{s} \left( \frac{1}{4}\delta^2 + \left( \frac{\delta}{2} - 2\kappa \right)^2 + \left( 1 - 144\frac{p}{s} \right) \kappa\delta \right) \geq \frac{1}{4} \frac{p}{s} \delta^2,\end{aligned}\quad (22)$$

where in the last transformation we took into account that  $p \leq s/144$ .

For the sake of technical precision in formulating the subsequent lemmas, we also consider the stochastic process  $\psi_t^*$ , given as  $\psi_t^* = \psi(u_t)$  for any  $t < t_d$ , where  $t_d$  is defined as the first time in the evolution of the system such that a configuration with  $\min_{i=1,2,3} a_{i,t_d} < c_6/n$  is reached, where  $c_6 = 313600/s$  is a constant depending only on  $s$ . For all  $t \geq t_d$ , we define  $\psi_t^* := \psi_{t-1}^* + \frac{1}{n}$ .

**Lemma 6.** *In any configuration  $u_t$  with  $\delta \geq s/20$  we have:  $\mathbb{E}\Delta\psi_t^* \geq \frac{1}{8} \frac{p\delta^2}{sn} > \frac{ps}{3600n}$ .*

*Proof.* We have:

$$\mathbb{E}\Delta\psi = -\mathbb{E}\Delta\phi - \frac{4p}{s^2} \mathbb{E}\Delta(\kappa^2) \quad (23)$$

Following the definition of  $\phi$  in Eq. (2), we have by linearity of expectation:

$$\mathbb{E}\Delta\phi = \mathbb{E} \left( \sum \ln(a_i + \Delta a_i) - \sum \ln a_i \right) = \sum \mathbb{E} \ln \left( 1 + \frac{\Delta a_i}{a_i} \right).$$

Next, using the bound  $\ln(1+b) \leq b$  which holds for  $b > -1$ , we have:

$$\sum \mathbb{E} \ln \left( 1 + \frac{\Delta a_i}{a_i} \right) \leq \sum \frac{\mathbb{E}\Delta a_i}{a_i} = \sum \frac{\dot{a}_i/n}{a_i} = \dot{\phi}/n$$

from which it follows that:

$$\mathbb{E}\Delta\phi \leq \frac{1}{n} \dot{\phi} \quad (24)$$

To analyze  $\mathbb{E}\Delta(\kappa^2)$ , we apply a variant of Lemma 1. A direct application of the lemma is not sufficient due to the singularity related to the  $a_i^{-1}$  term in the definition of  $\kappa_i$ ; however, this effect is compensated when we take into account that any change of the value of  $\kappa_i^2$  occurs in the considered protocol with probability at most proportional to  $a_i$ . For the specific case of  $\kappa_i^2$ , for fixed  $i = 1, 2, 3$ , we consider  $\kappa_i^2 : \mathbb{R}^2 \rightarrow \mathbb{R}$  as a function of the restricted configuration  $\bar{u} = (a_i^+, a_i^{++})$ , and we rewrite expression (12) as:

$$\begin{aligned}\mathbb{E}(\Delta\kappa_i^2) &= \sum_{u'=(a_i^{++}, a_i^{++}) \neq u} (\nabla f(u) \cdot (u' - u) + R_2(u, u')) \Pr(u'|u) \leq \\ &\leq \frac{\dot{\kappa}_i^2}{n} + \frac{1}{n^2} \max_{\|u^* - u\|_\infty \leq 1/n} D_{\kappa_i^2}(u^*) \sum_{u'=(a_i^{++}, a_i^{++}) \neq u} \Pr(u'|u) \leq \frac{\dot{\kappa}_i^2}{n} + \frac{1}{n^2} \max_{\|u^* - u\|_\infty \leq 1/n} D_{\kappa_i^2}(u^*) a_i.\end{aligned}$$

A straightforward computation from the definition of function  $\kappa_i$  shows that:

$$\max_{\|u^* - u\|_\infty \leq 1/n} D_{\kappa_i^2}(u^*) \leq \frac{8s^2}{a_i^2}.$$

It follows that

$$\mathbb{E}(\Delta\kappa_i^2) \leq \frac{\dot{\kappa}_i^2}{n} + \frac{a_i}{n^2} \frac{8s^2}{a_i^2} = \frac{1}{n} \left( \frac{\dot{\kappa}_i^2}{n} + \frac{8s^2}{a_i n} \right),$$

and so:

$$\mathbb{E}(\Delta\kappa^2) \leq \frac{1}{n} \left( \kappa^2 + \frac{24s^2}{\min_{i=1,2,3} a_i n} \right). \quad (25)$$

Introducing (24) and (25) into (23), we obtain:

$$\begin{aligned} \mathbb{E}\Delta\psi &= -\mathbb{E}\Delta\phi - \frac{4p}{s^2} \mathbb{E}\Delta(\kappa^2) \geq -\frac{1}{n} \dot{\phi} - \frac{1}{n} \frac{4p}{s^2} \left( \kappa^2 + \frac{24s^2}{\min_{i=1,2,3} a_i n} \right) = \frac{1}{n} \dot{\psi} - \frac{96p}{n^2 \min_{i=1,2,3} a_i} \geq \\ &\geq \frac{p}{4sn} \left( \delta^2 - \frac{392s}{\min_{i=1,2,3} a_i n} \right) \geq \frac{p}{8sn} \delta^2, \end{aligned}$$

where in the second-to-last transformation we used (22), and in the last transformation we used the relation  $\frac{392s}{\min_{i=1,2,3} a_i n} \leq \frac{\delta^2}{2}$  which holds when  $\delta \geq s/20$  and  $\min_{i=1,2,3} a_i \geq c_6/n$ .

The claim thus follows when  $\psi_t^* = \psi_t$  and  $\psi_{t+1}^* = \psi_{t+1}$ , i.e., for  $t < t_d$ . For larger values of  $t$ , the claim follows trivially from the definition of  $\psi_t^*$ .  $\square$

The above Lemma is used to show that, starting from any configuration with  $\delta > s/12$ , we quickly reach a configuration in which some species has a constant number of agents.

**Lemma 7.** *If  $\delta_t \geq s/20$ , we have:*

$$(i) \quad |\Delta\psi_t^*| \leq c_7,$$

$$(ii) \quad \text{Var}(\Delta\psi_t^*) \leq \frac{c_8}{n}.$$

where  $c_7 > 0$  and  $c_8 > 0$  are constants depending only on  $s$ . Moreover, in any configuration  $u$  with  $\min a_i \geq 2/n$ , we have:

$$(iii) \quad |\Delta\psi(u)| \leq \frac{c_7}{n \min a_i}.$$

$$(iv) \quad \text{Var}(\Delta\psi(u)) \leq \frac{c_8}{n^2 \min a_i}.$$

*Proof.* We first consider the case of a configuration with  $\min a_i \geq 2/n$ . Using the definition of  $\psi$  (and within it, of  $\phi$  and  $\kappa$ ). Consider any transition from a configuration  $u$  to a subsequent configuration  $u'$  and let  $S \subseteq \{1, 2, 3\}$  be defined as the set of indices of configurations changing between  $u$  and  $u'$  ( $S = \{i : a_i^+(u) \neq a_i^+(u') \vee a_i^{++}(u) \neq a_i^{++}(u')\}$ ). We verify that there exists an absolute constant  $c_7 > 0$  such that:

$$|\psi(u') - \psi(u)| \leq \frac{c_7}{n \min_{i \in S} a_i} \leq c_7.$$

Moreover, by the definition of the protocol a transition from  $u$  to  $u'$  occurs with probability  $\Pr(u'|u) \leq \min_{i \in S} a_i$ . Since there is only a constant number of possible successor configurations  $u_{t+1}$  for  $u_t$  (loosely bounding, not more than  $3^6$ ), it follows that:

$$\Pr \left[ |\Delta\psi(u)| > \frac{1}{b} \right] \begin{cases} < \frac{3^6 c_7 b}{n}, & \text{for any } b > 0 \\ = 0, & \text{for } b > \frac{c_7}{n \min_{i=1,2,3} a_i}. \end{cases}$$

The bounds on the variance of  $\text{Var}(\Delta\psi(u))$  and that of  $\Delta\psi_t^* = \Delta\psi(u)$  (for  $t < t_d$ ) with  $\min a_{i,t} \geq (c + 6 + 1)/n$  follow directly. The analysis of  $\Delta\psi_t^*$  when  $\min a_{i,t} = c_6/n$  and  $t < t_d$  is performed analogously, noting that if the succeeding configuration  $u' = u_{t+1}$  is such that  $\min_{i=1,2,3} a_i(u') < c_6/n$ , then  $\Delta\psi_t^* = \frac{1}{n}$ . Finally, for  $t \geq t_d$ , the result holds trivially by the definition of  $\psi_t^*$ .  $\square$

**Lemma 8.** *Let  $u_{t_3}$  be an arbitrary starting configuration of the system such that  $\delta_{t_3} > s/12$ . Then, with probability  $1 - O(1/n)$ , for some  $t_4 = t_3 + O(n \log n)$ , a configuration  $u_{t_4}$  is reached such that  $\min_{i=1,2,3} a_{i,t_4} = c_6$ .*

*Proof.* W.l.o.g. assume that  $t_3 = 0$ . First we remark that, by the relation between  $\eta$  and  $\delta$  for  $\delta \leq s/12$ , a process starting with  $\delta_0 > s/12$  satisfies:

$$\psi_0 = \psi_0 = \eta_0^2 - \frac{4p}{s^2} \kappa_0^2 > \frac{(s/12)^2}{s^2} - \frac{12p}{s^2} > \frac{1}{150}.$$

Moreover, for any configuration  $u'$  with  $\delta(u') \leq s/20$  we have:

$$\psi(u') = \eta^2(u') - \frac{4p}{s^2} \kappa^2 \leq \frac{(\frac{3}{2} \frac{s}{20})^2}{s^2} < \frac{1}{170}.$$

Thus, initially  $\psi_0 > \frac{1}{150}$  and as long as for all time steps  $t$  we have  $\psi_t \geq \frac{1}{170}$ , the barrier condition  $\delta_t \geq s/20$  has not been violated. Moreover, for  $\psi_t \in [\frac{1}{170}, \frac{1}{150}]$ , we have by Lemma 6 that  $\mathbb{E}\Delta\psi_t \geq 0$ . Moreover, by Lemma 7 (iii) and the fact that  $\delta_t < \frac{1}{144}$  which implies  $\min_{i=1,2,3} a_{i,t} > s/4$ , we have that  $|\Delta\psi_t| \leq \frac{4c_7}{sn}$ .

It follows from a standard application of Azuma's inequality for martingales (resembling the analysis of the hitting time of the random walk with step size  $O(1/n)$ , from one endpoint of a path of length  $\Theta(1)$  to the other) that:

$$\Pr \left[ \exists_{t < n^2/\ln n} \psi_t < \frac{1}{170} \right] = O(1/n),$$

hence also throughout the first  $n^2/\ln n$  steps of the process we have  $\delta > s/18$ , with probability  $1 - O(1/n)$ .

We are now ready to analyze the subsequent stages of the process, designing a Doob submartingale  $Y_t = \sum_{\tau=0}^{t-1} X_t$  with time increments  $(X_t)$  defined as:

$$X_t = \begin{cases} \Delta\psi_t^* - \frac{ps}{3600}, & \text{if } \psi_\tau > 1/170 \text{ and } \min_{i=1,2,3} a_{i,\tau} \geq c_6 \text{ for all } \tau \leq t \\ 0, & \text{otherwise.} \end{cases}$$

Using Lemma 7 (i) and (ii) and applying the Azuma-McDiarmid inequality<sup>3</sup> in the bounded variance version (cf. e.g. [11][Thm. 18]) to  $Y_t$  for  $t_c = c^3 n \ln n$ , for some sufficiently large constant  $c > 0$  depending only on  $s$ , we obtain:

$$\Pr[Y_t \leq -c^2 \ln n] \leq \exp \left[ -\frac{c^4 \ln^2 n}{2t \frac{c_8}{n} + \frac{2c_7 c^2 \ln n}{3}} \right] = \exp \left[ -\frac{c \ln n}{2c_8 + \frac{2}{3} \frac{c_7}{c}} \right] = 1 - O(1/n).$$

If the event  $X_t = \Delta\psi_t^* - \frac{ps}{3600}$  were to hold for all  $t < t_c$  with  $c = \frac{2 \cdot 3600}{ps}$  and if  $Y_{t_c} > -c^2 \ln n$ , then we would have  $\psi_{t_c}^* = \psi_0^* + Y_{t_c} + \frac{ps}{3600} t_c \geq 0 - c^2 \ln n + 3c^2 \ln n = 2c^2 \ln n$ , which would mean that  $\psi_{t_c}^* \neq \psi_{t_c}$ , since  $\psi \leq 3 \ln n + O(1)$  by definition. If  $\psi_{t_c}^* \neq \psi_{t_c}$ , then  $t_4 < t_c$ , and the proof is complete. (Indeed, to reach a configuration with  $\min_{i=1,2,3} a_i < c_6/n$ , the protocol has to pass through a configuration with  $\min_{i=1,2,3} a_i = c_6/n$ , since the size of each population changes by at most 1 in each transition.) Otherwise, we must have that at least one of the following events holds:  $Y_{t_c} \leq -c^2 \ln n$ , or  $\psi_\tau \leq 1/170$  for some  $\tau < t_c$ , or  $\min_{i=1,2,3} a_{i,\tau} < c_6$  for some  $\tau < t_c$ . We have established that each of the first two of these events holds with probability  $O(1/n)$ , whereas if the latter event holds, then  $t_4 < t_c$ . Thus,  $t_4 < t_c$  holds with probability  $1 - O(1/n)$  by a union bound.  $\square$

**Lemma 9.** *Let  $u_{t_4}$  be a starting configuration of the system such that  $\min_{i=1,2,3} a_{i,t_4} = c_6/n$ . Then, with constant probability, for some  $t_5 = t_4 + O(n \log n)$ , a configuration  $u_{t_5}$  is reached such that  $a_{j,t_5} \leq c_6/n$  and  $a_{j+1,t_5} > s/40$ , for some  $j \in \{1, 2, 3\}$ .*

<sup>3</sup>If our objective in the proof of the lemma were to show a bound on  $t_4$  which holds with constant probability (which would be sufficient for our purposes later on), rather than a w.h.p. bound, then this specific step of the proof can also be performed using Markov's inequality. In any case, we would need to make use of the bounded variance of  $\psi_t^*$  in the proof of the next Lemma.



*Proof.* W.l.o.g. assume that  $\arg \min_{i=1,2,3} a_{i,t_4} = 2$ . If  $a_{3,t_4} > s/40$ , then the claim follows immediately, putting  $t_5 = t_4$  and  $j = 2$ . Otherwise, we will show that with constant probability, the system will evolve so that  $a_2$  will increase over time until within  $O(n \log n)$  steps we will have a time step  $t_5$  with  $j = 3$  (i.e.,  $a_{3,t_5} \leq c_6/n$  and  $a_{1,t_5} > s/40$ ).

In the considered case, w.l.o.g. assume  $t_4 = 0$ . Next, let  $T = cn \ln n$  for a sufficiently large constant  $c$ ; we choose as  $c := 2 \log_2 \frac{1}{0.005ps}$  for convenience in later analysis. Intuitively, in view of Lemmas 6 and 7, the potential  $\psi_T^*$  will be further increased in the next steps: the random variable  $(\psi_T^* - \psi_0^* | u_0)$  has an expected value of  $\Theta(T/n) = \Theta(\log n)$ , with a standard deviation of  $\Theta(\sqrt{T/n}) = \Theta(\sqrt{\log n})$ .

By an application of the Azuma-McDiarmid inequality for martingales with bounded variance similar to that in the proof of Lemma 8, we obtain the following result:<sup>4</sup>

$$\Pr \left[ \forall_{t \leq T} \psi_t^* \geq \psi_0^* + \frac{pst}{3600n} - \varepsilon \frac{T}{n} \right] = 1 - n^{-\Omega(1)}, \quad \text{for any constant } \varepsilon > 0. \quad (26)$$

Observe that since  $a_{2,0} = c_6/n = O(1/n)$ , we have  $\psi_0^* \geq \ln n - O(1)$ . Taking this into account, for our purposes, a slightly weaker and simpler form of expression (26) will be more convenient:

$$\Pr \left[ \forall_{t \in [0.5n \log_2 n, T]} \psi_t^* \geq (1 + 10^{-4}ps) \ln n \right] = 1 - o(1). \quad (27)$$

The proof of the lemma is completed by a more fine-grained analysis of the considered protocol. In the initial configuration  $t_4 = 0$ , we have  $a_{2,0} = c_6/n$  (there are exactly  $c_6$  agents in state  $A_2$ ), and since  $t_4 \neq t_5$ , we have  $a_{3,0} < s/40$ . Consequently,  $a_{1,0} = s - s/40 - O(1/n) > 0.9s$ . Informally, since the prey of  $A_2$  (i.e.,  $A_1$ ) is more than twice more numerous than its predator (i.e.,  $A_3$ ), we should observe the increase in the size of population of  $A_2$ , regardless of the activities ( $A_i^+$  or  $A_i^{++}$ ) of the agents in the population. We consider the evolution of the system, finishing at the earliest time  $t_e$  when  $a_2(t_e) > s/100$ . The following relations are readily shown (apply e.g. Lemma 14 with  $i = 2$  and  $u = 0$ ):

$$\mathbb{E}(\Delta a_{2,t} | a_{3,t} < 0.05s, t < t_e) \geq \frac{0.05ps}{n} a_2 \quad (28)$$

$$\mathbb{E}(\Delta a_{3,t} | a_{3,t} < 0.05s, t < t_e) \leq -\frac{0.05ps}{n} a_3. \quad (29)$$

From (29), taking into account that  $|\Delta a_{i,t}| \leq \frac{1}{n}$  and  $a_{3,0} < s/40$ , an application of Azuma's inequality for martingales shows that:

$$\Pr[\forall_{t \leq \min\{t_e, T\}} a_{3,t} < 0.05s] = 1 - o(1). \quad (30)$$

Taking into account the above, by a straightforward geometric growth analysis (compare e.g. proof of Lemma 5), we obtain from (28):

$$\Pr[t_e < T] = 1 - o(1). \quad (31)$$

Moreover, since the speed of increase of  $a_2$  is bounded (even in the absence of predators) by that of a standard push rumor spreading process (formally,  $\mathbb{E}(\Delta a_{2,t}) \leq a_{2,t}$ ), we have (compare e.g. [19]):

$$\Pr[t_e > 0.5n \log_2 n] = 1 - o(1). \quad (32)$$

Now, we observe that with constant probability, the size of population  $A_2$  does not decrease in the time interval  $[0, t_e]$  below the value  $a_{2,0} = c_6/n$ , attained at the beginning of this interval:

$$\Pr \left[ \forall_{t \in [0, t_e]} a_{2,t} \geq c_6/n \right] = \Omega(1). \quad (33)$$

---

<sup>4</sup>Such an analysis can also be performed using Chebyshev's inequality, obtaining a slightly weaker expression in the probability bound.

Indeed, with constant probability the value  $a_{2,t}$  is initially non-decreasing: with constant probability, in the first  $O(n)$  rounds each of the  $c_6 = O(1)$  agents from  $A_2$  will be triggered by the scheduler  $O(1)$  times in total, and each interaction involving an agent from  $A_2$  will have this agent as the initiator, and an agent from the largest of the three populations,  $A_1$ , as the receiver (the prey). Thus, with constant probability, the number of agents in population  $A_2$  is increased to an arbitrary large constant (e.g.,  $1000c_6$ ). After this, we use the geometric growth property (28) to show that  $a_{2,t}$  reaches the barrier  $a_{2,t} > s/100$  (at time  $t_e$ ) before the event  $a_{2,t} < c_6/n$  occurs (cf. e.g. proof of Lemma 5, or standard analysis of variants of rumor-spreading processes in their initial phase [26]).

When the event from bound (33) holds, at least one of the following events must also hold:

- (A)  $\min_{i=1,2,3} a_{i,t} \geq c_6/n$ , for all  $t \leq t_e$ ,
- (B) or there exists a time step  $t < t_e$  such that  $a_{1,t} \leq c_6/n$ ,
- (C) or there exists a time step  $t < t_e$  such that  $a_{3,t} \leq c_6/n$ .

To complete the proof, we will show that each of the events (A) and (B) holds with probability  $o(1)$ . Indeed, then in view of (33), event (C) will necessarily hold with probability  $\Omega(1)$ . This means that, with probability  $\Omega(1)$ , there exists a time step  $t < t_e$  such that  $a_{3,t} < c_6/n$  and  $a_{2,t} < s/100$  (since  $t < t_e$ ), and so also  $a_{1,t} > s - s/100 - c_6/n > 0.98s > s/40$ ; thus, the claim of the lemma will hold with  $t_5 = t$  and  $j = 3$ .

To show that event (B) holds with probability  $o(1)$ , notice that  $a_{2,t} < s/100$  by definition of  $t_e$ , and moreover  $a_{3,t} < 0.05s$  with probability  $1 - o(1)$ , hence the event  $a_{1,t} < s - s/100 - 0.05s = 0.94s$  holds with probability  $o(1)$ .

To show that event (A) holds with probability  $o(1)$ , notice that, substituting in (27)  $t = t_e$ , by a union bound over (27), (31) and (32) we obtain:

$$\Pr [\psi_{t_e}^* \geq (1 + 10^{-4}ps) \ln n] = 1 - o(1).$$

This means that, with probability  $1 - o(1)$ , we have  $\psi_{t_e}^* \neq \psi_{t_e}$  or  $\psi_{t_e} \geq (1 + 10^{-4}ps) \ln n$ . In the first case, event (A) cannot hold. In the second case, observe that  $a_{2,t_e} = s/100 + O(1/n)$  by definition of  $t_e$ , so  $a_{1,t_e} > s - s/100 - O(1/n) - c_6/n > 0.98s$ , and it follows that  $\psi_{t_e} = \sum_{i=1}^3 \ln \frac{1}{a_{i,t_e}} + O(1) = \ln n + O(1)$ . Since the condition  $\psi_{t_e} \geq (1 + 10^{-4}ps) \ln n$  is not fulfilled, event (A) can only hold with probability  $o(1)$ .  $\square$

**Lemma 10.** *Let  $u_{t_5}$  be a starting configuration of the system such that  $a_{j,t_5} \leq c_6/n$  and  $a_{j+1,t_5} > s/40$ , for some  $j \in \{1, 2, 3\}$ . Then, with constant probability, for some  $t_6 = t_5 + O(n)$ , a configuration  $u_{t_6}$  is reached such that  $\min_{i=1,2,3} a_{i,t_6} = 0$ .*

*Proof.* We consider the pairs of interacting agents chosen by the scheduler in precisely the next  $n$  rounds after time  $t_5$ . Given that set  $A_{j,t_5}$  has constant size, and set  $A_{j+1,t_5}$  has linear size in  $n$ , it is straightforward to verify that with constant probability, the set of randomly chosen  $n$  pairs of agents has all of the following properties:

- Each agent from  $A_{j,t_5}$  belongs to exactly one pair picked by the scheduler, and is the receiver in this pair.
- Each agent interacting in a pair with an agent from  $A_{j,t_5}$  belongs to exactly one pair.
- Each agent interacting in a pair with an agent from  $A_{j,t_5}$  belongs to set  $A_{j+1,t_5}$ .

Conditioned on such a choice of interacting pairs by the scheduler, the protocol changes the state of all agents from set  $A_{j,t_5}$  to state  $j + 1$  with probability at least  $p^{|A_{j,t_5}|} \geq p^{c_6} = \Omega(1)$ . State  $j$  is then effectively eliminated.  $\square$

In the absence of species  $j$ , the interaction between species  $j - 1$  and  $j + 1$  collapses to a lazy predator-prey process, with transitions of the form  $(A_{j-1}, A_{j+1}) \rightarrow (A_{j-1}, A_{j-1})$  associated with a constant transition probability. A w.h.p. bound on the time of elimination of species  $j + 1$  follows immediately from the analysis of the push rumor spreading model, and we have the following Lemma.

**Lemma 11.** *Let  $u_{t_6}$  be a starting configuration of the system such that  $a_{j,t_6} = 0$ , for some  $j \in \{1, 2, 3\}$ . Then, with probability  $1 - O(1/n)$ , for some  $t_7 = t_6 + O(n \log n)$ , a configuration  $u_{t_7}$  is reached such that for all  $t \geq t_7$ ,  $a_{j,t} = a_{j+1,t} = 0$  and  $a_{j-1,t} = s$ .  $\square$*

After a further  $O(n \log n)$  steps after time  $t_7$ , the final configuration of all agents in the oscillator's population will be  $a_{j-1,t}^{++} = s$ .

### 3.3 Operation of the Oscillator in the Presence of a Source

In this section we prove properties of the oscillatory dynamics for the case  $\#X > 0$ . It is possible to provide a detailed analysis of the limit trajectories of the dynamics in this case, as a function of the concentration of  $x$ . Here, for the sake of compactness we only show the minimal number of properties of the oscillator required for the proof of Theorem 3. When the given configuration is such that  $\min_{i=1,2,3} a_i$  is sufficiently large, say  $\min_{i=1,2,3} a_i > 0.02s$ , then both the subclaims of Theorem 3(2) hold for the considered configuration. (The first subclaim hold directly; the second subclaim follows by a straightforward concentration analysis of the number of agents changing state in protocol  $P_o$  over the next  $0.01sn$  steps, since we will always have  $\min_{i=1,2,3} a_i \geq 0.01s$  during the considered time interval.) Otherwise, the considered configuration is close to one of the sides of the triangle. We will show that in the next  $O(n \log n)$  steps, with high probability, the protocol will either reach a configuration with  $\min_{i=1,2,3} a_i > 0.02s$ , or will visit successive areas around the triangle, as illustrated in Fig. 3. The following Lemmas show that within each area, an exponential growth process occurs, which propagates the agent towards the next area.

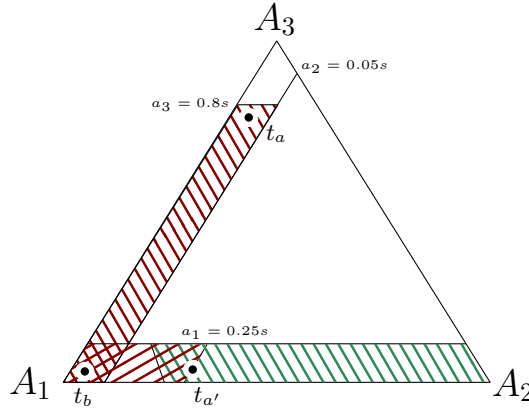


Figure 3: Configurations considered in Lemmas 12 and 13 (descending dashed slope) and Lemmas 14 and 15 (ascending dashed slope) for  $i = 1$  (dark red) and  $i = 2$  (green). Unless the process leaves the shaded area towards the center of the triangle, the shape of the time trajectory is counter-clockwise around the triangle, w.h.p.

**Lemma 12.** *If  $a_{i-1} < 0.8s$ ,  $a_{i+1} < 0.05s$  then  $\mathbb{E}\Delta a_{i-1} = \frac{\dot{a}_{i-1}}{n} \leq \frac{1}{n}(xs/3 - 0.05psa_{i-1})$ .*

*Proof.* From the assumptions we have that  $a_i > 0.15s$ . Starting from (3) we obtain:

$$\begin{aligned}
\dot{a}_{i-1} &= x(s/3 - a_{i-1}) + pa_{i+1}(a_{i-1} + a_{i-1}^{++}) - pa_{i-1}(a_i + a_i^{++}) \\
&\leq xs/3 + 2pa_{i-1}a_{i+1} - pa_{i-1}a_i \\
&\leq xs/3 + pa_{i-1}(2a_{i+1} - a_i) \\
&\leq xs/3 + pa_{i-1}(0.1s - 0.15s) = xs/3 - 0.05psa_{i-1}
\end{aligned}$$

□

From the above Lemma, the following Lemma follows directly by a standard concentration analysis (cf. e.g. proof of Lemma 5 for a typical analysis of this type of exponential growth process).

**Lemma 13.** *Let  $u_{t_a}$  be a starting configuration of the system such that  $a_{i-1,t_a} < 0.75s$  and  $a_{i+1,t_a} < 0.05s$ . Then, for some  $t_b \in [t_a, t_a + c_9 \log_2 n]$ , where  $c_9$  is a constant depending only on  $p$  and  $s$ , with probability  $1 - O(1/n)$ , the system reaches a configuration  $u_{t_b}$  such that exactly one of the following two conditions is fulfilled:*

- either  $\min_{j=1,2,3} a_{j,t_b} \geq 0.02s$ ,
- or  $a_{i+1,t_b} < 0.05s$  and  $a_{i-1,t_b} < 0.02s$ .

□

A similar analysis is performed for the next area.

**Lemma 14.** *If  $a_{i+1} < 0.25s$  and  $a_{i-1} < 0.05s$  then  $\dot{a}_{i+1} \geq xs/12 + 0.6psa_{i+1}$  and  $\dot{a}_{i-1} \leq xs/3 - 0.2psa_{i-1}$ .*

*Proof.* From assumptions we have that  $a_i > 0.7s$ . Starting from (3) we obtain:

$$\begin{aligned}
\dot{a}_{i+1} &= x(s/3 - a_{i+1}) + pa_i(a_{i+1} + a_{i+1}^{++}) - pa_{i+1}(a_{i-1} + a_{i-1}^{++}) \\
&\geq x(s/3 - a_{i+1}) + pa_{i+1}(a_i - 2a_{i-1}) \\
&\geq x(s/3 - 0.25s) + pa_{i+1}(0.7s - 2 \cdot 0.05s) \\
&= xs/12 + 0.6psa_{i+1}
\end{aligned}$$

$$\begin{aligned}
\dot{a}_{i-1} &= x(s/3 - a_{i-1}) + pa_{i+1}(a_{i-1} + a_{i-1}^{++}) - pa_{i-1}(a_{i+1} + a_{i+1}^{++}) \\
&\leq xs/3 + pa_{i-1}(2a_{i+1} - a_i) \\
&\leq xs/3 + pa_{i-1}(2 \cdot 0.25s - 0.7s) = xs/3 - 0.2psa_{i-1}.
\end{aligned}$$

□

Again, a concentration result follows directly.

**Lemma 15.** *Let  $u_{t_b}$  be a starting configuration of the system such that  $a_{i-1,t_b} < 0.02s$  and  $a_{i+1,t_b} < 0.02s$ . Then, for some  $t_{a'} \in [t_b, t_b + c_{10} \log_2 n]$ , where  $c_{10}$  is a constant depending only on  $p$  and  $s$ , with probability  $1 - O(1/n)$ , the system reaches a configuration  $u_{t_{a'}}$  such that exactly one of the following two conditions is fulfilled:*

- either  $\min_{j=1,2,3} a_{j,t_{a'}} \geq 0.02s$ ,
- or  $a_{i-1,t_{a'}} < 0.05s$ ,  $a_{i+1,t_{a'}} > 0.25s$ , and (consequently)  $a_{i,t_{a'}} < 0.75s$ .

□

An iterated application of Lemmas 13 and Lemmas 15 moves the process along time moments  $t_a, t_b, t'_a, \dots$ , where time moment  $t'_a$  is again fed to Lemma 13, considering the succeeding value of  $i$ . After a threefold application of both Lemmas, the process has w.h.p. in  $O(n \log n)$  steps either performed a complete rotation, passing through three moments of time designated as “ $t_a$ ”, rotated by one third of a full circle, or has reached at some time  $t'$  a point with  $\min_{j=1,2,3} a_{j,t'} \geq 0.02s$ . In either case, the claim of Theorem 3(2) follows directly.

## 4 Composition of Oscillatory Dynamics with Other Protocols

**Protocol composition.** In this Section, we show how to use protocol  $P_o$  as a component in the construction of other, more complex protocols. The *tensor composition* of two protocols defined for a population of size  $n$ , one protocol  $P_B$  using state set  $B = \{B_i : 1 \leq i \leq k_b\}$  and rule set  $R_B$ , the other protocol  $P_C$  using a state set  $C = \{C_i : 1 \leq i \leq k_c\}$ , disjoint from  $B$ , and rule set  $R_C$ , with a system of additional rules  $R_{BC}$  acting on input states from  $R_B \times R_C$  to output states from  $R_B \times R_C$ , is a new protocol, denoted  $(P_B \times P_C, R_{BC})$ , having state set  $B \times C$ . The rules are defined so that, for a selected pair of agents in states  $(B_i, C_j)$  and  $(B_{i'}, C_{j'})$ , each rule from  $R_B$ , acting on the pair of input states  $(B_i, B_{i'})$  with some probability  $p_b$  in protocol  $P_B$ , has probability  $p_b/4$  of being executed in the composition; likewise each rule from  $R_C$ , acting on the pair of input states  $(C_j, C_{j'})$  with some probability  $p_c$  in protocol  $P_C$ , has probability  $p_c/4$  of being executed in the composition; finally, each rule from  $R_{BC}$ , acting on the pair of input states  $(B_i, C_{j'})$  with some probability  $p_{bc}$  in rule set  $R_{BC}$ , has probability  $p_{bc}/2$  of being executed in the composition. (We always consider compositions which act from left to right, i.e., rules of the form  $(C, B)$  do not appear.) Here, an application of a rule to a product state is understood as an update of the respective component (the first component for protocol  $B$ , and the second component for protocol  $C$ ). The intuition of the composition is that a projection of the state onto its first component recovers protocol  $B$ , which may be seen as operating unaffected (under a coupling, which slows down its operation by a factor of 4 in expectation).

As a matter of naming convention, we name the states in the separate state sets of the composed protocols with distinct letters of the alphabet, together with their designated subscripts and superscripts. The rumor source  $X$  is treated specially and uses a separate letter (and may be seen as a one state protocol without any rules, on top of which all other protocols are composed; in particular, its state is never modified). The six remaining states of protocol  $P_o$  are named with the letters  $A_i^?$ , as usual in its definition. Subsequent protocols will use different letters, e.g.,  $L^?$ . The compositions we design will be such that they will never modify states from protocol  $P_o$ , i.e., states  $A_i^?$  will always act as initiators in initiator-preserving interaction rules between the two protocols.

**Design of protocol  $P_r$ .** To complete the proof of Theorem 1, we design protocol  $P_r$  as the composition of protocol  $P_o$  with an additional protocol. We start with a simpler design, which is not universal and guarantees that an uninformed state is reached in  $O(n \log^2 n)$  steps w.h.p. in the absence of the rumor source. Protocol  $P_o$  is composed with a protocol  $P_l$ , having two states,  $\{L^{on}, L^{off}\}$ . We informally refer to states  $L$  as lights. The composition is given in Fig. 4. Protocol  $P_l$  has a single rule pattern denoted as (7), and the additional rules for the protocol composition are given as (6).

To analyze the operation of the protocol, observe first that, in the absence of states  $A_i^+$ , all lights switch to an off state, given by rule (7), which implies exponential decay with a constant fraction of all lights switching off every  $O(n \log n)$ , in expectation. Overall, all lights switch off after  $O(n \log^2 n)$  steps, with high probability. By Theorem 3, in the absence of the rumor source all states  $L^{on}$  disappear from the system after  $O(n \log n)$  steps in expectation and in  $O(n \log^2 n)$  steps w.h.p.. Overall, all nodes reach a state having its  $L$ -component set to off after  $O(n \log^2 n)$  steps from the start of the process.

In the presence of the rumor source, we directly exploit Theorem 3(2), and in particular the fact

(6) A lazy agent turns the light on:

$$A_i^+; L^? \mapsto L^{on}$$

(7) Light eventually turns off:

$$L^?; L^{on} \mapsto L^{off}, \quad \text{with probability } q.$$

Figure 4: Protocol  $P_l$  composed with protocol  $P_o$ : Additional rules describing the mechanism of turning the light on and off. The value of  $q$  is set to  $c/\log n$ , where  $c > 0$  is a sufficiently small constant.

that states of the form  $A_i^+$  appear in the population with constant concentration every  $O(n \log n)$  steps, w.h.p. By a simple concentration analysis, the light switch is thus switched to on within the subsequent  $O(n)$  steps for a uniformly random subset of nodes, on a constant proportion of the population. By fine-tuning the decay probability  $q$  in rule (7), we can ensure that the expected rate of decay of enabled lights is smaller than the rate at which they are switched on, unless the proportion of switched on lights in the population is at least  $(1 - \varepsilon/2)n$ , for an arbitrarily chosen  $\varepsilon > 0$ . We eventually obtain that at any time  $(1 - \varepsilon)n$  agents have their light turned on, w.h.p.

Given the natural decoding of states as “informed” (having component  $L^{on}$ ) or “uninformed” (having component  $L^{off}$ ), the proof of Theorem 1 is complete. We remark that it is also possible to design a related protocol in which exactly one state is recognized as “informed” and exactly one state is recognized as “uninformed”; we omit the details of the construction. Likewise, a universal variant of protocol  $P_r$  can be designed which relies on a composition with the 3-state majority protocol of [6], which takes  $O(\log n)$  time to move from its metastable state, as a replacement for the rule (7) executed with inverse logarithmic probability.

## 5 The Lower Bound

This Section is devoted to the proof of Theorem 2. First, we restate some notation. We recall that the vector  $z = (z^{(1)}, \dots, z^{(k)}) \in \{0, 1, \dots, n\}^k = Z$  describes the number of agents having particular states, and  $\|z\|_1 = n$ . In this section we will identify the set of states with  $\{1, \dots, k\} = [1, k]$ . It is now also more convenient for us to work with a scheduler which selects unordered (rather than ordered) pairs of interacting agents; we note that both models are completely equivalent in terms of computing power under a fair random scheduler, since selecting an ordered pair of agents can be seen as selecting an unordered pair, and then setting their orientation through a coin toss. Indexing with integers  $\{1, 2, \dots, r\}$  the set of all distinct rules of the protocol, where  $r \leq k^4$ , for a rule  $j \equiv \{\{i_1(j), i_2(j)\} \rightarrow \{o_1(j), o_2(j)\}\}$ ,  $1 \leq j \leq r$ ,  $i_1(j), i_2(j), o_1(j), o_2(j) \in \{1, \dots, k\}$ , we will denote by  $q_j$ , the probability (selected by the protocol designer) that rule  $j$  is executed as the next interaction rule once the scheduler has selected  $(i_1, i_2)$  as the interacting pair, and by  $p_j(z)$  the probability that  $j$  is the next rule chosen in configuration  $z$  (we have  $p_j(z) = q_j \frac{z^{(i_1(j))} z^{(i_2(j))}}{n^2} (1 - O(1/n))$ , where the  $O(1/n)$  factor compensates the property of a scheduler which always selects a distinct pair of elements).

For any configuration  $z_0 \in Z$ , we define the  $d$ -box  $B_d(z_0)$  around  $z_0$  as the set of all states  $z \in Z$  such that  $z_0^{(i)}/d \leq z^{(i)} \leq d \max\{1, z_0^{(i)}\}$ , for all  $1 \leq i \leq k$ . We start the proof with the following property of boxes.

**Lemma 16.** *Fix  $k \in \mathbb{Z}^+$  and let  $0 < \varepsilon_1 < 0.001$  be arbitrarily fixed. There exists  $\varepsilon_0 = \varepsilon_0(k, \varepsilon_1)$ ,  $0 < \varepsilon_0 < \varepsilon_1$ , such that, for any interaction protocol  $P$  with  $k$  states and any configuration  $z_0 \in Z$ , there*

exists a value  $\varepsilon = \varepsilon(P, z_0) \in [\varepsilon_0, \varepsilon_1]$  such that, for any rule  $j$  of the protocol,  $1 \leq j \leq r$ , exactly one of the following bounds holds:

1. (i) for all  $z \in B_{n^{\varepsilon_0}}(z_0)$ ,  $p_j(z) \leq n^{\varepsilon-1}$ ,
2. (ii) for all  $z \in B_{n^{\varepsilon_0}}(z_0)$ ,  $p_j(z) \geq n^{24\varepsilon-1}$ .

and for any state  $i$ ,  $1 \leq i \leq k$ , exactly one of the following bounds holds:

1. (iii) for all  $z \in B_{n^{\varepsilon_0}}(z_0)$ ,  $z^{(i)} \leq n^\varepsilon$ ,
2. (iv) for all  $z \in B_{n^{\varepsilon_0}}(z_0)$ ,  $z^{(i)} \geq n^{24\varepsilon}$ .

*Proof.* Let  $k$  be fixed and let  $\varepsilon_0 = 96^{-(k+k^4+f+1)} \leq 96^{-(k+r+f+1)}$ , where  $f = \log_2(1/\varepsilon_1)$ . Consider the (multi)set  $M$  of real values  $M := \{\log_n \max\{n^{\varepsilon_0}, z_0^{(i)}\} : i \in \{1, \dots, r\}\} \cup \{\log_n \max\{n^{\varepsilon_0}, np_j(z_0)\} : j \in \{1, \dots, r\}\} \subseteq [0, 1]$ . Since  $|M| = k + r$ , by the pigeonhole principle, there must exist an interval  $I_l = [96^{-l}, 96^{-l+1})$ , for some  $l \in \{f, \dots, k + r + f\}$ , such that  $I_l \cap M = \emptyset$ . Now, we set  $\varepsilon = 2 \cdot 96^{-l} > 96\varepsilon_0$ , we also have  $\varepsilon < 2 \cdot 96^{-f} < \varepsilon_1$ . We immediately obtain that for any state  $i$ ,  $1 \leq i \leq k$ , we either have  $z_0^{(i)} \leq n^{\varepsilon/2}$  or  $z_0^{(i)} \geq n^{48\varepsilon}$ . Recalling that for any  $z \in B_{n^{\varepsilon_0}}(z_0)$ ,  $z_0^{(i)}/n^{\varepsilon_0} \leq z^{(i)} \leq n^{\varepsilon_0} \max\{1, z_0^{(i)}\}$ , claims (iii) and (iv) follow.

To show claims (i) and (ii), notice that if rule  $j$ ,  $j \in \{1, \dots, r\}$  is such that  $\min\{z_0^{(i_1(j))}, z_0^{(i_2(j))}\} \leq n^\varepsilon$ , then for all  $z \in B_{n^{\varepsilon_0}}(z_0)$  we have  $\min\{z^{(i_1(j))}, z^{(i_2(j))}\} \leq n^\varepsilon$  (by (iii) and (iv)), and so  $p_j(z) \leq n^{\varepsilon-1}$  by the properties of the random scheduler. Otherwise, we have  $\min\{z_0^{(i_1(j))}, z_0^{(i_2(j))}\} \geq n^{24\varepsilon}$ , and so  $\frac{1}{2n^{2\varepsilon_0}} \leq p_j(z_0)/p_j(z) \leq 2n^{2\varepsilon_0}$ , where we recall that  $\varepsilon > 96\varepsilon_0$ . Since we have  $p_j(z_0) \leq n^{\varepsilon/2}$  or  $p_j(z_0) \geq n^{48\varepsilon}$ , claims (i) and (ii) follow.  $\square$

Given any  $k$ -state protocol  $P$ , we will arbitrarily choose a value of  $\varepsilon$  for which the claim of the above Lemma holds (e.g., the smallest possible such value of  $\varepsilon$ ). Note that a similar analysis is also possible for protocols using a super-constant number of states in  $n$ , however, then the value of  $\varepsilon_0$  is dependent on  $n$ ; retracing the arguments in the proof, we can choose appropriately  $\varepsilon_0 \geq n^{\exp[-O(k^4)]}$ . (We make no effort to optimize the polynomial in  $k$  in the exponent.)

In what follows, let  $z_0$  be a fixed configuration of the protocol (admitting a certain property which we will define later). We will then consider a rule  $j$  to be a *low probability (LP) rule* (writing  $j \in LP$ ) in box  $B_{n^{\varepsilon_0}}(z_0)$  if it satisfies condition (i) of the Lemma, and a *high probability (HP) rule* in this box (writing  $j \in HP$ ) if it satisfies condition (ii). Note that  $LP \cup HP = \{1, \dots, r\}$ .

Likewise, for  $1 \leq i \leq k$ , we will classify  $i$  as a *low-representation (LR) state* (writing  $i \in LR$ ) in box  $B_{n^{\varepsilon_0}}(z_0)$  if  $i$  satisfies condition (iii) of the Lemma, and a *high representation (HR) state* (writing  $i \in HR$ ) in this box if  $i$  satisfies condition (iv). Note that  $LR \cup HR = \{1, \dots, k\}$ . Moreover, we define a set of *very high representation (VHR) states*,  $VHR \subseteq HR$ , as the set of all  $i$  such that for all  $z' \in B_{n^{\varepsilon_0}}(z_0)$ ,  $z'_i \geq n^{1-8\varepsilon}$ . Denoting  $HR' = HR \setminus VHR$ , we have by the definition of a box that for all  $i' \in HR'$ , for all  $z' \in B_{n^{\varepsilon_0}}(z_0)$ :  $z'_i \leq n^{1-8\varepsilon}/O(n^{2\varepsilon_0}) < n^{1-6\varepsilon}$ .

From now on, we assume that configuration  $z_0$  admits the following property: for  $T = n^{1+2\varepsilon}$ , an execution of the protocol starting from configuration  $z_0$  passes through a sequence of configurations  $z_t$ ,  $t = 1, 2, \dots, T$ , such that the configuration does not leave the box around  $z_0$  in any step with sufficiently large probability, lower-bounded by some absolute constant  $\Pi \in (0, 1]$ :

$$\Pr[\forall t < T \ z_t \in B \subseteq B_{n^{\varepsilon_0}}(z_0)] \geq \Pi, \quad (34)$$

where  $B$  is an arbitrarily fixed subset of  $B_{n^{\varepsilon_0}}(z_0)$ .

We now show that the above property has the following crucial implication: for an interacting pair involving selected high and very high representation states, a rule creating a low representation state can only be triggered with sufficiently small probability. Informally, it seldom happens that in the protocol a low representation state is created out of any high representation state.

**Lemma 17.** For a protocol having the property given by Eq. (34), for  $i_1 \in HR$  and  $i_2 \in VHR$ , let  $R_{i_1, i_2}$  be the set of rules of the form  $\{i_1, i_2\} \rightarrow \{o_1, o_2\}$ , taken over all  $o_1 \in LR, o_2 \in [1, r]$ . Then,  $\sum_{j \in R_{i_1, i_2}} q_j = O(n^{-14\varepsilon})$ .

*Proof.* Suppose, by contradiction, that  $\sum_{j \in R_{i_1, i_2}} q_j > 3n^{-14\varepsilon}$ .

Associate with process  $z_t$  a random variable  $J_t \in \{0, 1\}$ , defined as follows. For all  $t < t_e$ , where  $t_e$  is the first moment of time such that  $z_{t_e} \notin B_{n^{\varepsilon_0}}(z_0)$ , we put  $J_t = 1$  if a rule from  $R_{i_1, i_2}$  is used for the interaction made by the protocol in process  $z_t$  at time  $t$ , and set  $J_t = 0$  otherwise. For all  $t \geq t_e$ , we set  $J_t$  to 1. We have  $\mathbb{E}(J_t | z_1, \dots, z_t) \geq 2n^{2\varepsilon-1}$ ; indeed, for  $t < t_e$ , it holds that:

$$\begin{aligned} \mathbb{E}(J_t | z_1, \dots, z_t) &= \Pr[J_t = 1 | z_t] = \sum_{j \in R_{i_1, i_2}} p_j(z_t) = \sum_{j \in R_{i_1, i_2}} q_j \frac{z^{(i_1(j))} z^{(i_2(j))}}{n^2} (1 - O(1/n)) \geq \\ &\geq 3n^{-14\varepsilon} \frac{n^{24\varepsilon} n^{1-8\varepsilon}}{n^2} (1 - O(1/n)) > 2n^{2\varepsilon-1}. \end{aligned}$$

By a simple stochastic domination argument,  $(J_t)$  can be lower-bounded by a sequence of independent binomial trials with success probability  $n^{2\varepsilon-1}$ , hence by an application of a multiplicative Chernoff bound for  $T = n^{1+2\varepsilon}$ :

$$\Pr \left[ \sum_{t=1}^T J_t > n^{4\varepsilon} \right] = \Pr \left[ \sum_{t=1}^T J_t > \frac{1}{2} 2n^{2\varepsilon-1} T \right] = 1 - o(1),$$

where the  $o(1)$  factor is exponentially small in  $n$ .

We now show the following claim.

*Claim.* With probability  $\Pi - o(1)$ , the following event holds:  $z_t \in B$  for all  $t \in [0, T)$  and the total number of rule activations in the time interval  $[0, t)$  during which an agent changes state from a state in  $LR$  to a different state is at most  $O(kn^{3\varepsilon})$ .

*Proof (of claim).* Acting similarly as before, we associate with process  $z_t$  a random variable  $L_t \in \{0, 1\}$ , defined as follows. For all  $t < t_e$ , we put  $L_t = 1$  if a rule acting on at least one agent in a state from  $LR$  is made by the protocol in process  $z_t$  at time  $t$ , and set  $L_t = 0$  otherwise. For all  $t \geq t_e$ , we set  $L_t$  to a dummy variable set always to 0, i.a.r. We observe that:

$$\mathbb{E}(L_t | z_1, \dots, z_t) \leq 2k \frac{n^\varepsilon}{n},$$

since  $|LR| \leq k$ , and for  $t < t_e$ ,  $z_i < n^\varepsilon$  for any  $i \in LR$ , hence the scheduler selects an agent from a  $LR$  state into an interacting pair with probability at most  $2k \frac{n^\varepsilon}{n}$ . Applying an analogous argument as in the case of random variable  $L_t$ , this time for the upper tail, we obtain:

$$\Pr \left[ \sum_{t=1}^T L_t < 4kn^{3\varepsilon} \right] = \Pr \left[ \sum_{t=1}^T L_t > 2 \cdot 2kn^\varepsilon T \right] = 1 - o(1).$$

The claim follows directly.

Now, by a union bound we obtain:

$$\Pr \left[ \sum_{t=1}^T J_t > n^{4\varepsilon} \wedge \sum_{t=1}^T L_t < 4kn^{3\varepsilon} \right] = 1 - o(1).$$

Taking into account that  $t_e > T$  holds with probability  $\Pi = \Omega(1)$  by (34), we have by a union bound that with probability at least  $\Pi - o(1) = \Omega(1)$ , the following event holds:  $z_t \in B$  for all  $t \in [0, T]$ ,  $\sum_{t=1}^T J_t > n^{4\varepsilon}$ , and  $L_t < 4kn^{3\varepsilon}$ . However then  $\sum_{t=1}^T J_t - \sum_{t=1}^T L_t > n^{4\varepsilon} - kn^{3\varepsilon} > kn^\varepsilon$ , so there must exist at time  $T$  a state  $i \in LR$  such that  $z_T^{(i)} > n^\varepsilon$ . This is a contradiction with  $z_T \in B \subseteq B_{n^{\varepsilon_0}}(z_0)$  by Lemma 16(iii).  $\square$



In the rest of the proof, we consider the evolution of a protocol starting from configuration  $z_0$  and having property (34). We compare this evolution to the evolution of the same protocol, starting from a perturbed configuration  $z_0^*$ , such that:

$$(C1) \quad \|z_0 - z_0^*\|_1 \leq n^\varepsilon.$$

$$(C2) \quad \text{for all low representation states } i \in LR, \text{ we have } z_0^{*(i)} \leq z_0^{(i)}.$$

Intuitively, the perturbed state  $z_0^*$  may correspond to removing a small number of agents from  $z_0$  (and replacing them by high representation states for the sake of normalization), e.g., as in the case of the disappearance of a rumour source from a system which has already performed a rumor-spreading process.

Our objective will be to show that, with probability at least  $\Pi - o(1)$ , after  $T = n^{1+2\varepsilon}$  the process  $z_T^*$  is still not far from  $z_0$ , being constrained to a box in a similar way as process  $z_t$ . To achieve this, we define a coupling between processes  $z_t$  and  $z_t^*$  (knowing that process  $z_t$  is constrained to a box around  $z_0$  with probability  $\Pi$ ). Informally, the analysis proceeds as follows. We run the processes together for  $T = n^{1+2\varepsilon}$  steps. In most steps, the 1-norm distance  $\|z_t - z_t^*\|_1$  between the two processes remains unchanged, without exceeding  $O(n^{3\varepsilon})$ . Otherwise, exactly one of the two processes executes a rule (and the other pauses). With a frequency of roughly  $n^\varepsilon/n$  steps (i.e., roughly  $n^{3\varepsilon}$  times in total during the process), an LP rule is executed which increases the distance between these two states. We think of this type of “error” as unfixable, contributing to the  $O(n^{3\varepsilon})$  distance of the processes; however, such errors are relatively uncommon. With a higher frequency of roughly  $n^{3\varepsilon}/n$  steps (i.e., roughly once every  $n^{1-3\varepsilon}$  steps), a less serious “error” occurs, when some HP rule  $\iota$  increases the distance between the two states. The rate of such errors is too high to leave them unfixed, and we have a time window of about  $n^{1-3\varepsilon}$  steps to fix such an error (before the next such error occurs). We observe that since  $\iota$  is an HP rule, which is activated with probability at least  $n^{24\varepsilon-1}$ , rule  $\iota$  will still be activated frequently during this time window. The coupling of transitions of states  $z_t$  and  $z_t^*$  is in this case performed so as to force the two processes to execute rule  $\iota$  lazily, never at the same time. The number of executions of rule  $\iota$  in the ensuing time window by each of the two processes follows the standard coupling pattern of a pair of lazy random walks on a line, initially located at distance 1, until their next meeting (cf. e.g. [2]). During this part of the coupling, we allow the distance  $\|z_t - z_t^*\|_1$  to increase even up to  $n^{6\varepsilon}$  (as a result of executions of rule  $\iota$ ), but the entire contribution to the distance related to rule  $\iota$  is reduced to 0 before the next HP rule “error” occurs, with sufficiently high probability (in this case, with probability  $1 - O(n^{-6\varepsilon})$ ). Overall, the coupling is successful with probability  $\Pi - O(n^{-\varepsilon})$ .

We remark that we use the bound on the number of states  $k$  to enforce a sufficiently large polynomial separation between the frequencies of LR states and HR states, and likewise for LP rules and HP rules. We also implicitly assume that  $k = n^{o(1)}$ , throughout the process. The analysis also works for a choice of  $k = O(\log \log n)$ , with a sufficiently small hidden constant. The separation between LR/HR states and LP/HP rules is used in at least two places in the proof. First, it enforces that rules creating LR states from VHR states may appear in the definition of the protocol only with polynomially small probability (Lemma 17), which helps to maintain over time the invariant  $z_t^{*(i)} \leq z_t^{(i)}$ , for all LR states. Secondly, we use the separation of LP/HP rules in the analysis of the coupling to show that a fixable “error” caused by a HP rule can be sufficiently quickly repaired, before new errors occur.

In the formalization of the coupling, we make both processes  $z_t$  and  $z_t^*$  lazy, i.e., add to each process an additional independent coin-toss at each step, and enforce that with probability  $1/2$  no rule is executed in a given step (i.e., the step is skipped by the protocol). We assume a random scheduler which picks uniformly a random pair of nodes at each step. Thus, if the scheduler picks a pair of agents in states  $\{i_1, i_2\}$ , and  $j$  is a rule acting on this pair of states, the probability that the interaction corresponding to rule  $j$  will be  $q_j/2$ . (The laziness of the process here is a purely technical assumption for the analysis, and corresponds to using a measure of time which is scaled by a factor of  $2 \pm o(1)$  w.h.p.; this does not affect the asymptotic statement of the theorem.)

We will also find it convenient to apply an auxiliary notation for representing the evolution of a state. For process  $z_t$  (resp.,  $z_t^*$ ), we define  $\rho_t(j)$  (resp.  $\rho_t^*(j)$ ), for all  $j \in [1, r]$ , as the number of times rule  $j$  has been executed since time 0. Observe that the pair  $(z_0, (\rho_t(j) : j \in [1, t])$  completely describes the evolution of a state (i.e., the order in which the rules were executed is irrelevant). Moreover, since each execution of a rule changes the states of at most 4 agents, we have:

$$\|z_t - z_t^*\|_1 \leq 4 \sum_{j=1}^r |\rho_t(j) - \rho_t^*(j)| + \|z_0 - z_0^*\|_1 \leq 4 \sum_{j=1}^r |\rho_t(j) - \rho_t^*(j)| + n^\varepsilon.$$

### Definition of the coupling.

1. At each step  $t$ , we order the agents of configurations  $z_t$  and  $z_t^*$ , so that  $a_l(t)$  denotes the type of the  $l$ -th agent in  $z_t$  and  $a_l^*(t)$  is the type of the  $l$ -th agent in  $z_t^*$ . The orderings are such that  $|\{l : a_l(t) = a_l^*(t)\}|$  is maximized; in particular, for any state  $i$  such that  $z^{(i)}(t) \leq z^{*(i)}(t)$  (respectively,  $z^{*(i)}(t) \leq z^{(i)}(t)$ ) we have that if for some  $l$ ,  $a_l(t) = i$  (resp.,  $a_l^*(t) = i$ ), then  $a_l^*(t) = i$  (resp.,  $a_l(t) = i$ ).
2. The scheduler then picks a pair of distinct indices  $l_1, l_2 \in \{1, \dots, n\}$  as the pair of interacting agents.
  - 2.1. If  $a_{l_1}(t) = a_{l_1}^*(t)$  and  $a_{l_2}(t) = a_{l_2}^*(t)$ , then the same rule  $j = j^*$  acting on the pair of states  $(a_{l_1}(t), a_{l_2}(t))$  is chosen as the current interaction rule, with probability  $q_j$ .
  - 2.2. Otherwise, a pair of (clearly distinct) rules  $j$  and  $j^*$  are picked independently at random for  $z_t$  and  $z_t^*$  from among the rules available for state pairs  $(a_{l_1}(t), a_{l_2}(t))$  and  $(a_{l_1}^*(t), a_{l_2}^*(t))$ , with probabilities  $q_j$  and  $q_{j^*}$ , respectively.
3. The processes finally perform their coin tosses to decide which of the selected rules ( $j$  for  $z_t$  and  $j^*$  for  $z_t^*$ ) will be applied in the current step.
  - 3.1. If  $j = j^*$  and rule  $j$  has been executed exactly the same number of times in the history of the two processes ( $\rho_t(j) = \rho_t^*(j)$ ), then with probability  $1/2$  both of the processes execute rule  $j$ , and with probability  $1/2$  neither execute their rule.
  - 3.2. If  $j \neq j^*$ , or if  $j = j^*$  and rule  $j$  has been executed a different number of times in the history of the two processes ( $\rho_t(j) \neq \rho_t^*(j)$ ), then exactly one of the two processes performs its chosen rule and the other process waits, with the process performing the rule being chosen as  $z_t$  or  $z_t^*$ , with probability  $1/2$  each.

The correctness of the coupling (i.e., that the marginals  $z_t$  and  $z_t^*$  each correspond to a valid execution of the given protocol under a random scheduler) is immediate to verify.

**Lemma 18.** *Let  $z_t$  be a process satisfying property (34), and let  $z_0^*$  satisfy conditions (C1) and (C2). Then, for  $T = n^{1+2\varepsilon}$ , with probability  $\Pi - O(n^{-\varepsilon})$  we have  $\|z_T^* - z\|_1 = O(n^{6\varepsilon})$ , for some  $z \in B$ .*

*Proof.* To prove the claim, it suffices to show that with probability  $\Pi - O(n^{-\varepsilon})$  the provided coupling succeeds, i.e., it maintains a sufficiently small difference  $z_T(i) - z_T^*(i)$  for all states  $i$ , with  $z_T \in B$ .

In the analysis of the provided coupling, we will assume that the box condition  $z_t \in B$  holds always throughout the process (otherwise, we assume the coupling does not succeed). To state this formally, we work with auxiliary processes  $\bar{z}_t$  and  $\bar{z}_t^*$ , given as  $\bar{z}_t = z_t$  and  $\bar{z}_t^* = z_t^*$  for all  $t < t_e$ , where  $t_e$  is the first moment of time such that  $z_t \notin B_{n^{\varepsilon_0}}(z_0)$ , and set to the dummy value  $\bar{z}_t = \bar{z}_t^* = z_0$  for all  $t \geq t_e$ . At the end of the process, we will thus have  $\bar{z}_T = z_T$  and  $\bar{z}_T^* = z_T^*$  with probability at least  $\Pi$ . In the following, we silently assume that  $t < t_e - 1$  (in particular, that  $z_t \in B$  and  $z_{t+1} \in B$ ), and we will simply show

that the coupling of  $\bar{z}_t$  and  $\bar{z}_t^*$  is successful with probability  $1 - n^{-\varepsilon}$ . The condition of  $t \geq t_e - 1$  is trivially handled.

In addition to the box condition (which is now enforced) we try to maintain, with sufficiently high probability, throughout the first  $T$  steps of the process several invariants (all at a time), corresponding to the following events holding:

- $F_D(t)$ : for all states  $i \in LR$ ,  $\bar{z}_t^{*(i)} \leq \bar{z}_t^{(i)}$ . (LR domination condition)
- $F_{LR}(t)$ : for all states  $i \in LR$ ,  $\bar{z}_t^{*(i)} \leq \bar{z}_t^{(i)} \leq n^\varepsilon$ . (LR state condition)
- $F_{LP}(t)$ : for all rules  $j \in LP$ ,  $\max\{p_j(\bar{z}_t), p_j(\bar{z}_t^*)\} \leq 2n^{\varepsilon-1}$ . (LP rule condition)
- $F_{HR}(t)$ : for all states  $i \in HR$ ,  $\min\{\bar{z}_t^{(i)}, \bar{z}_t^{*(i)}\} \geq n^{24\varepsilon}/2$ . (HR state condition)
- $F_{HP}(t)$ : for all rules  $j \in HP$ ,  $\min\{p_j(\bar{z}_t), p_j(\bar{z}_t^*)\} \geq n^{24\varepsilon-1}/2$ . (HP rule condition)
- $F_{HR'}(t)$ : for all states  $i \in HR'$ ,  $\max\{\bar{z}_t^{(i)}, \bar{z}_t^{*(i)}\} \leq 2n^{1-6\varepsilon}$ . (HR' state condition)
- a family of possible events  $S_{w,d}(t)$ , for some  $d \in \{0, \dots, 4n^{3\varepsilon}\}$  and  $w \in \{0, \dots, n^{6\varepsilon}\}$ , with specific events defined as follows:
  - $S_{0,d}(t)$  holds if for all rules  $j \in HP$  we have  $\rho_t(j) = \rho_t^*(j)$ , and  $\sum_{j \in LP} |\rho_t(j) - \rho_t^*(j)| = d$ . This implies, in particular,  $\|\bar{z}_t^* - \bar{z}_t\|_1 \leq 4d + n^\varepsilon \leq 5n^{3\varepsilon}$ . (identical rate of HP execution)
  - $S_{w,d}(t)$  for  $w > 0$  holds if there exists a rule  $\iota \in HP$  such that for all rules  $j \in HP \setminus \{\iota\}$  we have  $\rho_t(j) = \rho_t^*(j)$ ,  $|\rho_t(\iota) - \rho_t^*(\iota)| = w$ , and moreover  $\sum_{i \in LP} |\rho_t(i) - \rho_t^*(i)| = d$ . This implies, in particular,  $\|\bar{z}_t^* - \bar{z}_t\|_1 \leq 4d + 4w + n^\varepsilon \leq 5n^{6\varepsilon}$ . (single HP execution difference)

We will call the coupling *successful* if for all  $t \leq T$ , all events  $F.(t)$  and some event  $S_{w,d}(t)$  holds, and we will say it is a *failure* otherwise. (We remark that condition  $F_D(t)$  is implied by condition  $F_{LR}(t)$ , but we retain both for convenience in discussion.)

The analysis of the coupled process is now the following. First, we remark that all of the given events  $F.(t)$  and event  $S_{0,0}(t)$  hold for  $t = 0$ .

If the process meets condition  $S_{0,d}$  at time  $t$  and all conditions  $F.(t)$ , then we have the following:

- With probability at least  $1 - O(n^{3\varepsilon-1})$ , the coupling will follow clauses 2.1 and 3.1 of its definition, and the two processes  $\bar{z}$  and  $\bar{z}^*$  will execute the same rule  $j$  (or both pause). Hence, we continue to step  $t + 1$  satisfying condition  $S_{0,d}$  and all of the conditions  $F.(t + 1)$ , making use of the box condition for process  $\bar{z}_t$ . (We note that, to show  $F_{LP}(t)$ , when considering the special case of a rule involving a state from  $LR$ , we can make use of  $F_{LR}(t)$  and note that the activation probability of such a rule is bounded by  $2n^{\varepsilon-1}$  due to the  $n^\varepsilon$  bound on the population of a LR state).
- With probability at most  $O(n^{3\varepsilon-1})$ , the coupling will, however, select distinct rules,  $j$  for  $\bar{z}_t$  and  $j^*$  for  $\bar{z}_t^*$ , and will select exactly one of them to execute, say  $j' \in \{j, j^*\}$ .
  - If  $j' \in LP$ , which happens in the current step of the process with probability at most  $2n^{\varepsilon-1}$  by  $F_{LP}$ , then the event  $S_{0,d+1}(t + 1)$  will hold in the next step (provided  $d + 1 \leq 4n^{3\varepsilon}$ ; otherwise, if  $d + 1 > 4n^{3\varepsilon}$ , we will say that the coupling has failed).
  - If  $j' \in HP$ , which happens in the coupling with probability  $O(n^{3\varepsilon-1})$  (as bounded due to clause 2.2), then the event  $S_{1,d}(t + 1)$  will hold in the next step. The condition  $F_D(t + 1)$  requires more careful consideration. Taking into account that  $F_D(t)$  holds, we need to consider two cases: either  $j' = j$  and the rule applied to  $\bar{z}_t$  changed at least one of the two interacting states  $\{i_1(j), i_2(j)\}$ , say  $i_1(j) \in LR$ , so that  $\bar{z}_t^{i_1(j)}(t) = \bar{z}_t^{*i_1(j)}(t)$  and

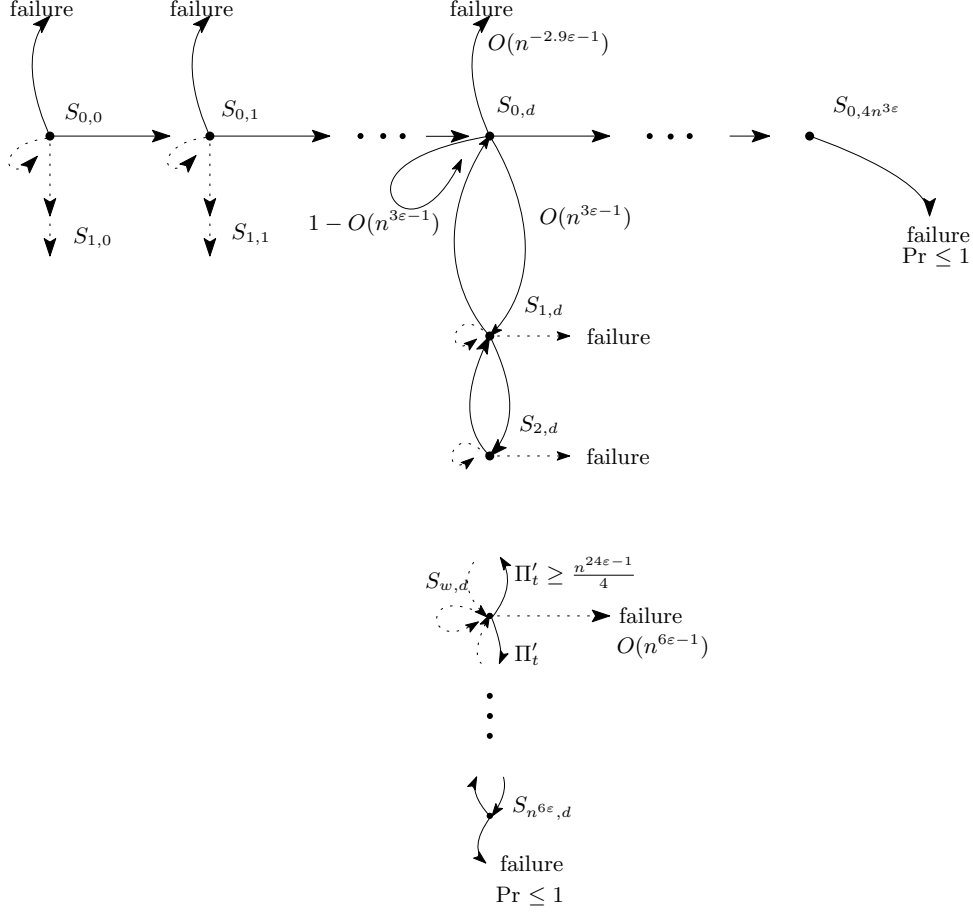


Figure 5: Illustration of transitions between states  $S_{w,d}$  for the coupling in the lower bound proof.

$\bar{z}^{i_1(j)}(t+1) \leq \bar{z}^{*i_1(j)}(t+1) - 1$ , or  $j' = j^*$  and the rule applied to  $\bar{z}_t^*$  created a pair of states  $\{o_1(j^*), o_2(j^*)\}$ , say  $o_1(j^*) \in LR$ . In the first case, by the description of the ordering given in clause 1 of the definition of the coupling, the problem occurs only if one of the agents picked by the scheduler belongs to an  $LR$  state, and the other agent is at a position in which the states of  $\bar{z}$  and  $\bar{z}^*$  differ in the ordering of the agents; hence, the probability that the coupling fails at this step is at most  $O(\frac{kn^\epsilon \cdot n^{3\epsilon}}{n^2}) \leq O(n^{5\epsilon-2})$ . In the second case, we likewise analyze the ordering of the agents considered by the scheduler, and note that the interacting agent, which belongs to the part of the ordering in which  $\bar{z}_t$  and  $\bar{z}_t^*$  differ, must be in a  $HR$  state, since the agents in a  $LR$  state in  $\bar{z}^*$  are matched by their counterparts in  $\bar{z}$  (as noted in clause 1 of the discussion of the coupling). If the other interacting agent is in a state from  $LR \cup HR'$ , then such an event occurs with probability  $O(\frac{n^{3\epsilon} \cdot kn^{1-6\epsilon}}{n^2}) \leq O(n^{-2.9\epsilon-1})$ , and we say that with this probability the coupling has failed. Finally, if the other interacting agent is in a state from  $VHR$ , then by Lemma 17, we have that the probability of picking a rule under which the coupling fails is at most  $O(n^{-14\epsilon})$ , conditioned on the event  $j \neq j^*$  holding, hence overall the probability of failure is  $O(n^{-14\epsilon} n^{3\epsilon-1}) = O(n^{-11\epsilon-1})$ . Overall, we obtain that  $F_D(t+1)$  holds with probability  $O(n^{-2.9\epsilon-1})$ . Given  $F_D(t+1)$ ,  $S_{1,d(t+1)}$ , and the box condition, the remaining conditions  $F(t+1)$  follow directly.

Overall, we obtain that following a time  $t$  satisfying  $S_{0,d}(t)$  and all conditions  $F(t)$ , we reach the following successor state (see Fig. 5):

$$\left\{ \begin{array}{ll} S_{0,d}(t+1) \wedge F.(t+1), & \text{with probability } 1 - O(n^{3\varepsilon-1}), \\ S_{0,d+1}(t+1) \wedge F.(t+1), & \text{with probability } \leq 2n^{\varepsilon-1}, \text{ if } d+1 \leq 4n^{3\varepsilon}, \\ S_{1,d}(t+1) \wedge F.(t+1), & \text{with probability } O(n^{3\varepsilon-1}), \\ \text{failure:} & \left\{ \begin{array}{ll} \text{with probability } O(n^{-2.9\varepsilon-1}), & \text{if } d+1 \leq 4n^{3\varepsilon}, \\ \text{with some probability } \leq 1, & \text{otherwise.} \end{array} \right. \end{array} \right.$$

At this point, before proceeding further, we can provide some intuition on the meaning of the respective events  $S$ . The coupling process can be seen as a walk along the path  $(S_{0,d} : d \leq 4n^{3\varepsilon})$ , starting from state  $S_{0,0}$ , and at each step, either staying in the current state  $S_{0,d}$ , moving on to the next state  $S_{0,d+1}$ , branching to a side branch  $S_{1,d}$  (which we will analyze later), or failing. The process also fails if it reaches the endpoint of its path ( $d = 4n^{3\varepsilon}$ ). Since the process is run for  $T = n^{1+2\varepsilon}$  steps, the probability that failure will occur before the end of the path is reached is  $O(n^{-0.9\varepsilon})$ , and the probability of reaching the end of the path and failing is exponentially small in  $n^\varepsilon$  by a Chernoff bound (in expectation, the process will progress halfway along the path). Hence, we have that the process succeeds with probability  $1 - O(n^{-0.9\varepsilon})$ , or otherwise may fail in a side branch  $S_{1,d}$ .

A side branch is entered with probability  $O(n^{\varepsilon-1})$ . To show that the coupling succeeds with the required probability, it suffices to show that we return from any state  $S_{1,d}$  to state  $S_{0,d}$  with probability at least  $1 - O(n^{-6\varepsilon})$ ; then, all (i.e., w.h.p. at most  $O(n^{1+2\varepsilon}n^{3\varepsilon-1}) = O(n^{5\varepsilon})$ ) excursions into side branches during the process will succeed with probability  $1 - O(n^{-\varepsilon})$ .

Consider now an excursion into a side branch  $S_{w,d}$  ( $w \geq 1$ ) associated with a rule  $\iota \in HP$ , which has been executed a different number of times in  $\bar{z}_t$  and  $\bar{z}_t^*$ . Now, if the process meets condition  $S_{w,d}$  at time  $t$  and all conditions  $F.(t)$ , then we have the following:

- With probability at least  $1 - O(n^{6\varepsilon-1})$ , the coupling will follow clause 2.1 of its definition, selecting a single rule  $j$ .
  - If  $j \neq \iota$ , then clause 3.1 will follow, and the two processes  $\bar{z}$  and  $\bar{z}^*$  will execute the same rule  $j$  (or both pause). Hence, at time  $t+1$ , all of the conditions  $F.(t+1)$  and condition  $S_{w,d}(t+1)$  is satisfied.
  - Else, the event  $j = \iota$  occurs. The probability of such an event is denoted  $\pi_t \in [p_\iota(\bar{z}_t) - O(n^{6\varepsilon-1}), p_\iota(\bar{z}_t)]$  (due to the conditioning performed in the first clause of the coupling); since  $p_\iota(\bar{z}_t) \geq n^{24\varepsilon-1}$  by the box condition for HP rules, it follows that  $2\pi_t \geq n^{24\varepsilon-1} - O(n^{6\varepsilon-1}) \geq n^{24\varepsilon-1}/2$ . Now, following clause 3.2 of the coupling, depending on which of the two processes  $\bar{z}_t, \bar{z}_t^*$  is chosen to execute the rule, with probability  $\pi_t/2 =: \pi'_t$  the system moves to  $S_{w-1,d}(t+1)$ , and with probability  $\pi'_t$  the system moves to  $S_{w+1,d}(t+1)$  (unless  $w+1 > n^{6\varepsilon}$ , in which case the coupling has failed). As before, given there was no failure, all conditions  $F.(t+1)$  are readily verified to be satisfied in the new time step.
- With probability at most  $O(n^{6\varepsilon-1})$ , for simplicity of analysis we assume the coupling has failed.

This time, for a time  $t$  satisfying  $S_{w,d}(t)$  for  $w \geq 1$  and all conditions  $F.(t)$ , we obtain the following distribution of successor states:

$$\left\{ \begin{array}{ll} S_{w-1,d}(t+1) \wedge F.(t+1), & \text{with probability exactly } \pi'_t \geq n^{24\varepsilon-1}/4, \\ S_{w+1,d}(t+1) \wedge F.(t+1), & \text{with probability exactly } \pi'_t, \text{ if } w+1 \leq n^{6\varepsilon}, \\ \text{failure:} & \left\{ \begin{array}{ll} \text{with probability } O(n^{6\varepsilon-1}), & \text{if } w+1 \leq n^{6\varepsilon}, \\ \text{with some probability } \leq 1, & \text{otherwise,} \end{array} \right. \\ S_{w,d}(t+1) \wedge F.(t+1), & \text{otherwise.} \end{array} \right.$$

The picture here corresponds to a lazy random walk along the side line  $S_{w,d}$  for  $w \in [0, n^{6\epsilon}]$ , with an additional failure probability at each step. The walk starts at  $w = 1$  and ends with a return to the primary line  $S_{0,d}$  if the endpoint  $w = 0$  is reached, or ends with failure if the other endpoint  $w \geq n^{6\epsilon} =: w_{\max}$  is reached. At each step, the walk is lazy (with probability of transition depending on the current step), but unbiased with respect to transitions to the left or to the right. Assuming that failure does not occur sooner, with probability  $1 - O(\frac{1}{w_{\max}}) = 1 - O(n^{-6\epsilon})$  the walk will reach point  $w = 0$  in  $O(w_{\max}^2) = O(n^{12\epsilon})$  moves (transitions along the line), without reaching the other endpoint of the line sooner. Since a move is made in each step  $t$  with probability  $\pi_t' \geq n^{24\epsilon-1}/4$ , by a straightforward Chernoff bound, the number of steps spent on this line is given w.h.p. as at most  $O(n^{12\epsilon}/n^{24\epsilon-1}) = O(n^{1-12\epsilon})$ . As the probability of failure in each of these steps is  $O(n^{6\epsilon-1})$ , the probability that the process fails during these steps is  $O(n^{-6\epsilon})$ . Overall, by a union bound, we obtain that the process successfully returns to  $S_{0,d}$  with probability  $1 - O(n^{-6\epsilon})$  (and within  $O(n^{1-12\epsilon})$  steps). In view of the previous observations, we have that with probability  $1 - O(n^{-\epsilon})$ , all conditions  $F$ . and some condition  $S_{w,d}$  hold at time  $T$ . Thus, with probability  $1 - O(n^{-\epsilon})$ , process  $z_T^*$  is sufficiently close to  $B$ , i.e., there exists a point  $z \in B$  such that  $\|z_T^* - z\|_1 = O(n^{6\epsilon})$ .  $\square$

**Acknowledgment.** We sincerely thank Dan Alistarh and Przemek Uznański for inspiring discussions, and Lucas Boczkowski for many detailed comments which helped to improve this manuscript.

## References

- [1] M. A. Abdullah and M. Draief. Majority consensus on random graphs of a given degree sequence. *CoRR*, abs/1209.5025, 2012.
- [2] D. Aldous and J. A. Fill. Reversible markov chains and random walks on graphs, 2002. Unfinished monograph, recompiled 2014, available at <http://www.stat.berkeley.edu/~aldous/RWG/book.html>.
- [3] D. Alistarh, J. Aspnes, D. Eisenstat, R. Gelashvili, and R. L. Rivest. Time-space trade-offs in population protocols. In *Proc. Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain*, pages 2560–2579, 2017.
- [4] D. Alistarh, B. Dudek, A. Kosowski, D. Soloveichik, and P. Uznanski. Robust detection in leak-prone population protocols, 2017. Accepted to DNA'17.
- [5] D. Angluin, J. Aspnes, Z. Diamadi, M. J. Fischer, and R. Peralta. Computation in networks of passively mobile finite-state sensors. *Distributed Computing*, 18(4):235–253, 2006.
- [6] D. Angluin, J. Aspnes, and D. Eisenstat. A simple population protocol for fast robust approximate majority. *Distributed Computing*, 21(2):87–102, 2008.
- [7] D. Angluin, J. Aspnes, D. Eisenstat, and E. Ruppert. The computational power of population protocols. *Distributed Computing*, 20(4):279–304, 2007.
- [8] J. Aspnes and E. Ruppert. An introduction to population protocols. *Bulletin of the EATCS*, 93:98–117, 2007.
- [9] L. Becchetti, A. E. F. Clementi, E. Natale, F. Pasquale, R. Silvestri, and L. Trevisan. Simple dynamics for majority consensus. *CoRR*, abs/1310.2858, 2013.
- [10] L. Boczkowski, A. Korman, and E. Natale. Minimizing message size in stochastic communication patterns: Fast self-stabilizing protocols with 3 bits. In *Proc. Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain*, pages 2540–2559, 2017.

- [11] F. R. K. Chung and L. Lu. Survey: Concentration inequalities and martingale inequalities: A survey. *Internet Mathematics*, 3(1):79–127, 2006.
- [12] C. Cooper, R. Elsässer, and T. Radzik. The power of two choices in distributed voting. In *Proc. 41st International Colloquium on Automata, Languages, and Programming, ICALP 2014, Copenhagen, Denmark, Part II*, pages 435–446, 2014.
- [13] J. Czyzowicz, L. Gasieniec, A. Kosowski, E. Kranakis, P. G. Spirakis, and P. Uznanski. On convergence and threshold properties of discrete Lotka-Volterra population protocols. In *Proc. 42nd International Colloquium on Automata, Languages, and Programming, ICALP 2015, Kyoto, Japan, Part I*, pages 393–405, 2015.
- [14] A. Dobrinevski and E. Frey. Extinction in neutrally stable stochastic Lotka-Volterra models. *Phys. Rev. E*, 85:051903, May 2012.
- [15] B. Doerr, C. Doerr, S. Moran, and S. Moran. Simple and optimal randomized fault-tolerant rumor spreading. *Distributed Computing*, 29(2):89–104, 2016.
- [16] B. Doerr and M. Fouz. Asymptotically optimal randomized rumor spreading. *Electronic Notes in Discrete Mathematics*, 38:297–302, 2011.
- [17] B. Doerr, M. Fouz, and T. Friedrich. Social networks spread rumors in sublogarithmic time. *Electronic Notes in Discrete Mathematics*, 38:303–308, 2011.
- [18] B. Doerr, D. Johannsen, and C. Winzen. Multiplicative drift analysis. *CoRR*, abs/1101.0776, 2011.
- [19] B. Doerr and M. Künnemann. Tight analysis of randomized rumor spreading in complete graphs. In *2014 Proc. Eleventh Workshop on Analytic Algorithmics and Combinatorics, ANALCO 2014, Portland, Oregon, USA, January 6, 2014*, pages 82–91, 2014.
- [20] D. Doty. Timing in chemical reaction networks. In *Proc. Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 772–784, 2014.
- [21] U. Feige, D. Peleg, P. Raghavan, and E. Upfal. Randomized broadcast in networks. In *Algorithms, International Symposium SIGAL '90, Tokyo, Japan, 1990*, pages 128–137, 1990.
- [22] N. Fountoulakis and K. Panagiotou. Rumor spreading on random regular graphs and expanders. *Random Struct. Algorithms*, 43(2):201–220, 2013.
- [23] A. M. Frieze and G. R. Grimmett. The shortest-path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10(1):57–77, 1985.
- [24] G. Giakkoupis, Y. Nazari, and P. Woelfel. How asynchrony affects rumor spreading time. In *Proc. 2016 ACM Symposium on Principles of Distributed Computing, PODC 2016, Chicago, IL, USA*, pages 185–194, 2016.
- [25] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.
- [26] R. M. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. Randomized rumor spreading. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 565–574, 2000.

- [27] B. Kerr, M. A. Riley, M. W. Feldman, and B. J. M. Bohannan. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature*, 418(6894):171–174, Jul 2002.
- [28] B. C. Kirkup and M. A. Riley. Antibiotic-mediated antagonism leads to a bacterial game of rock-paper-scissors in vivo. *Nature*, 428(6981):412–414, Mar 2004.
- [29] L. Lamport. Time, clocks, and the ordering of events in a distributed system. *Commun. ACM*, 21(7):558–565, 1978.
- [30] C. Lenzen, T. Locher, P. Sommer, and R. Wattenhofer. Clock synchronization: Open problems in theory and practice. In *Proc. 36th Conference on Current Trends in Theory and Practice of Computer Science, SOFSEM 2010, Spindleruv Mlýn, Czech Republic*, pages 61–70, 2010.
- [31] C. Lenzen, T. Locher, and R. Wattenhofer. Clock synchronization with bounded global and local skew. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 509–518, 2008.
- [32] A. J. Lotka. Contribution to the theory of periodic reactions. *The Journal of Physical Chemistry*, 14(3):271–274, 1909.
- [33] J. Lundelius and N. A. Lynch. A new fault-tolerant algorithm for clock synchronization. In *Proc. Third Annual ACM Symposium on Principles of Distributed Computing, Vancouver, B. C., Canada*, pages 75–88, 1984.
- [34] G. B. Mertzios, S. E. Nikolettseas, C. L. Raptopoulos, and P. G. Spirakis. Determining majority in networks with local interactions and very small local memory. *Distributed Computing*, 30(1):1–16, 2017.
- [35] O. Michail, I. Chatzigiannakis, and P. G. Spirakis. *New Models for Population Protocols*. Synthesis Lectures on Distributed Computing Theory. Morgan & Claypool Publishers, 2011.
- [36] K. Panagiotou and L. Speidel. Asynchronous rumor spreading on random graphs. *CoRR*, abs/1608.01766, 2016.
- [37] M. Parker and A. Kamenev. Extinction in the Lotka-Volterra model. *Phys. Rev. E*, 80:021129, 2009.
- [38] B. Pittel. On spreading a rumor. *SIAM J. Appl. Math.*, 47(1):213–223, Mar. 1987.
- [39] V. B. Priezzhev, D. Dhar, A. Dhar, and S. Krishnamurthy. Eulerian walkers as a model of self-organized criticality. *Phys. Rev. Lett.*, 77:5079–5082, Dec 1996.
- [40] T. Reichenbach, M. Mobilia, and E. Frey. Coexistence versus extinction in the stochastic cyclic Lotka-Volterra model. *Phys. Rev. E*, 74:051907, Nov 2006.
- [41] T. Sauerwald. On mixing and edge expansion properties in randomized broadcasting. *Algorithmica*, 56(1):51–88, 2010.
- [42] D. Soloveichik, G. Seelig, and E. Winfree. Dna as a universal substrate for chemical kinetics. In A. Goel, F. C. Simmel, and P. Sosík, editors, *DNA*, volume 5347 of *Lecture Notes in Computer Science*, pages 57–69. Springer, 2008.
- [43] A. Szolnoki, M. Mobilia, L. Jiang, B. Szczesny, A. M. Rucklidge, and M. Perc. Cyclic dominance in evolutionary games: A review. *CoRR*, abs/1408.6828, 2014.



## Appendix. Simulation Examples for Oscillator Dynamics

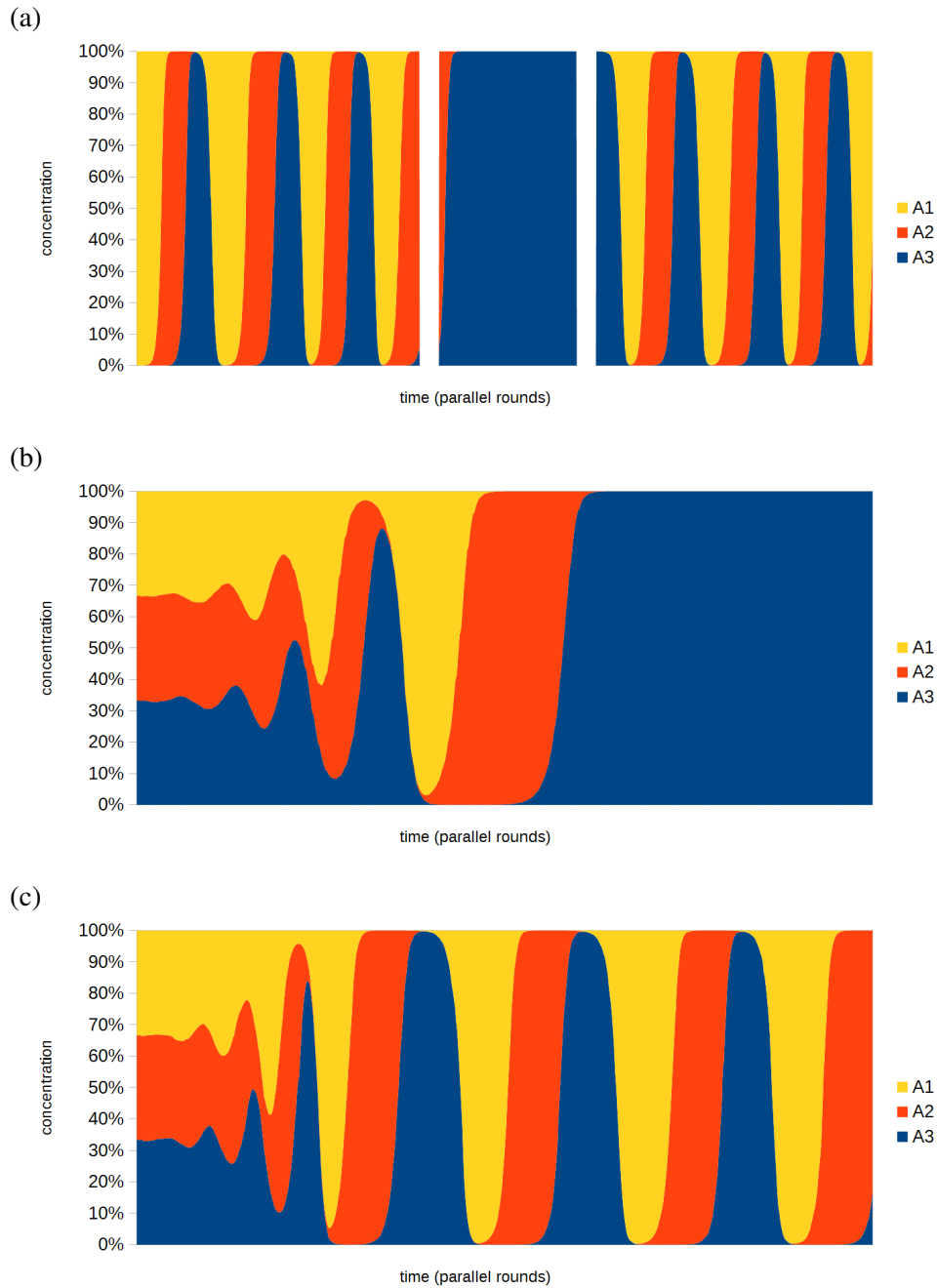


Figure 6: Illustration of concentration of species  $A_1$ ,  $A_2$ ,  $A_3$  as a function of time steps for a simulation of the protocol  $P_o$  for  $n = 10^6$ ,  $p = 6 \cdot 10^{-2}$ ,  $s = 1$ . (a) Initialization from a corner configuration ( $\#X = 1$ , left panel), further dynamics of the protocol after rumour source is removed ( $\#X = 0$ , middle panel), further dynamics of the protocol after rumour source is reinserted ( $\#X = 1$ , right panel). (b) Initialization from a configuration with  $A_1 = A_2 = A_3 = 1/3$  and  $\#X = 0$ . (c) As above, with  $\#X = 1$ . The range of the horizontal time scale corresponds to 2000 parallel rounds of the protocol.