



**HAL**  
open science

## Reaching Agreement in Competitive Microbial Systems

Victoria Andaur, Janna Burman, Matthias Függer, Manish Kushwaha, Bilal Manssouri, Thomas Nowak, Joel Rybicki

► **To cite this version:**

Victoria Andaur, Janna Burman, Matthias Függer, Manish Kushwaha, Bilal Manssouri, et al.. Reaching Agreement in Competitive Microbial Systems. 2022. hal-03545888

**HAL Id: hal-03545888**

**<https://hal.inrae.fr/hal-03545888>**

Preprint submitted on 27 Jan 2022

**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.

# Reaching Agreement in Competitive Microbial Systems

**Victoria Andaur** · victoria.andaaur@student-cs.fr  
Université Paris-Saclay, CNRS

**Janna Burman** · janna.burman@lri.fr  
Université Paris-Saclay, CNRS

**Matthias Függer** · mfuegger@lsv.fr  
CNRS, LSV, ENS Paris-Saclay, Université Paris-Saclay, Inria

**Manish Kushwaha** · manish.kushwaha@inrae.fr  
Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute

**Bilal Manssouri** · bilal.manssouri@student-cs.fr  
Université Paris-Saclay, CNRS

**Thomas Nowak** · thomas.nowak@lri.fr  
Université Paris-Saclay, CNRS

**Joel Rybicki** · joel.rybicki@ist.ac.at  
Institute of Science and Technology Austria

**Abstract.** In this work, we consider distributed agreement tasks in microbial distributed systems under stochastic population dynamics and competitive interactions. We examine how competitive exclusion can be used to solve distributed agreement tasks in the microbial setting. To this end, we develop a new technique for analyzing the time to reach competitive exclusion in systems with several competing species under biologically realistic population dynamics. We use this technique to analyze a protocol that exploits competitive interactions to solve approximate majority consensus efficiently in synthetic microbial systems.

We show that *direct* competition dynamics reach majority consensus with high probability when the initial gap between the species is small, i.e.,  $\Omega(\sqrt{n \log n})$ , where  $n$  is the initial population size of the majority species. In contrast, we show that *indirect* competition alone is not efficient: for example, solving majority consensus with high probability requires an initial gap of  $\Omega(n)$ . To corroborate our analytical results, we use computer simulations to show that these consensus dynamics occur within practical time scales.

## 1 Introduction

Computing permeates all areas of engineering, science, and nature. In the past few decades, theoretical computer scientists have started to investigate computation in various biological systems; e.g., [16, 20, 21, 32]. Microbiology in particular provides several fascinating examples of large-scale dynamic distributed systems: bacterial species are known to use various molecular communication mechanisms, such as quorum sensing, to coordinate their collective metabolic actions in order to influence and shape their environments [10].

The computational power of these tiny, self-replicating biological devices has been put into action by the discipline of synthetic biology, an area residing at the intersection of biology and engineering. For over two decades, synthetic biologists have shown how to engineer and re-program existing species to perform computation, yielding increasingly complex biological circuits. Early success stories in the area showed how to build synthetic genetic toggle switches [22] and biological oscillators [33, 37]. Recently, the area has rapidly shifted towards distributed

computing [28, 34, 35]: instead of engineering complex metabolic pathways within a single cell, the biological circuit is distributed across a consortium of multiple bacterial strains or species.

Naturally, this decentralization comes at a cost: computation across different parts of the biological circuit has to be coordinated somehow. While distributed computing theory has dealt with these types of issues for decades, most existing models of distributed computation do not account for the unique features exhibited by microbial systems. For example, microbial communities are subject to stochastic population dynamics, individual bacterial cells reproduce and die at a fast pace, and species exhibit complex ecological interactions [23].

## 1.1 Our Focus: Computation in Competitive Microbial Systems

In this work, we consider distributed computation and coordination in microbiological systems governed by stochastic population dynamics and ecological interactions. We focus on competitive interactions between species and we examine how the *principle of competitive exclusion* can be utilized to efficiently solve *approximate majority* and *consensus* – two fundamental tasks in distributed computing – in microbial populations.

In the consensus problem, the system is initialized with two input values, and the goal is to reach an output configuration where all participants agree on a single value. In the majority problem, the goal is to output the value that had the majority support initially. In this work we consider *majority consensus*, which is achieved if the initial minority population dies out.

The principle of competitive exclusion states that two complete competitors, i.e., with overlapping niches, cannot stably coexist in a single population [26]. Indeed, due to stochastic fluctuations in the population sizes, one of the species will eventually gain an upper hand and drive the other species into extinction. We study how this principle can be used to solve consensus efficiently in the microbial setting.

**Competitive interactions.** Bacterial communities exhibit a wide range of competitive interactions [23]. In order to outcompete other species, bacteria have evolved various types of methods of competition in the course of a biological arms race [25]. Broadly speaking, there are two categories of competitive interactions:

- (1) *Exploitative competition:* The species and individuals indirectly compete for shared resources, e.g., nutrients or space. This is an indirect means of competition, where competition is mediated by the shared resource. The species that better utilize available resources eventually prevail over inferior competitors.
- (2) *Interference competition:* The individuals use direct means to hinder or fight against competing species. Bacteria may secrete molecules that inhibit the performance of competing species, but they also employ a diverse selection of more lethal methods for interference competition. These range from mechanical weapons that can be used to puncture cell membranes of nearby competitors to using diffusive molecular weapons, such as toxins and antibiotics, that wipe out other species [25]. In some environments, e.g., in a mammalian gut, some bacterial species are even known to use elaborate schemes to trigger host immune responses to attack against competing strains.

## 1.2 Contributions

We investigate distributed models of microbiological computation that capture the effects of stochastic population dynamics and various competitive interactions. Formally, the population dynamics and protocols are expressed in stochastic the *biological reaction network* (BRN) model, which generalizes *chemical reaction networks* (CRN) [24].

**A model for microbial dynamics.** We introduce a new modeling framework that allows us to incorporate more realistic biological interactions; in particular in cases where (1) the growth of the population is limited by the availability of resources, and (2) the individuals follow stochastic birth-death dynamics. Moreover, we can model populations consuming multiple different types of resources. For example, the introduction of two distinct resources allows for the modeling of two-phase growth behavior, which is observed in experiments [31]. Our framework also allows for modeling of open systems with continuous in-flow of resources and out-flow of resources and cells. We emphasize that these type of biological dynamics are not captured by the standard population protocol [9] and chemical reaction network models studied in the distributed computing literature.

**Analytical results.** As our main technical contribution, we give a new coupling technique that allows us to bound the time to competitive exclusion in a (possibly complex) continuous-time two-species model via a simple, discrete-time birth-death process. We apply this technique to analytically bound the time to reach competitive exclusion. We model direct interference competition, where the two competing species fight each other: on every interaction consisting of two individuals of different species, one of these are killed. We show that in this model competitive exclusion occurs and, if the initial discrepancy between the two species is  $\Omega(\sqrt{n \log n})$ , then the majority species outcompetes the minority species with high probability.

We note that a similar model was studied recently by Cho et al. [13] who investigated approximate majority consensus with direct competitive interactions under birth dynamics. However, their model did not incorporate resource or individual death dynamics (beyond that given by competitive interactions), and thus, the model assumes unbounded population growth.

We further show the inefficiency of indirect exploitative competition by showing that the probability of a species outcompeting the other is equal to its relative initial population in many cases.

**Simulation results.** Using simulations, we observe that the protocol using exploitative competition reaches consensus slower than the protocols relying on interference competition. From a biological perspective, these results confirm the intuition that direct interference competition is more effective when dealing with rival species, as expected. However, direct interference competition tends to be more costly for the individuals, as they have to dedicate their energy in producing suitable bacterial weaponry.

From an algorithmic viewpoint, we see that exploiting interference competition can yield faster algorithms. In particular, this suggests that it may be beneficial to harness existing mechanisms for direct competition, as developed by bacterial species, also in the context of synthetic microbial consortia, despite their additional metabolic cost.

### 1.3 Outline of the Paper

We first review related work in Section 2. In Section 3 we introduce biological reaction networks (BRNs) as a basic modeling framework for microbial population dynamics. We introduce a new protocol in this formalism, called the *mutual annihilation protocol*, which serves as our main example of microbial dynamics with interference competition between two species.

In Section 4 we first establish an upper bound on the number of steps by coupling the protocol with an easier-to-analyze discrete-time Markov chain. Using this coupling, we show our main result: the mutual annihilation protocol achieves majority consensus with high probability when the initial gap between the two competing cell counts is large enough.

In Section 5, we prove that indirect competition alone is not sufficient to achieve majority consensus with high probability in many cases. Specifically, in symmetric systems without in-flow of new resources, the probability of reaching majority consensus is equal to the initial relative population of the majority species.

Finally, in Section 6 we provide simulation results that complement our analytical results. Using biologically realistic birth and death rates for bacteria, we simulate and compare two scenarios consisting of two competing strains of *E. coli* that use exploitative competition and interference competition, respectively.

## 2 Related Work

We now briefly overview related work on majority consensus in molecular and biological models of computation. In particular, we focus on the population protocol model and the chemical reaction network model.

### 2.1 Population Protocols

Population protocols were designed to study the computational power of an ensemble of indistinguishable agents that interact in an unpredictable manner [6, 7]. In this model, individuals are represented by finite state automata that communicate via pairwise interactions: in every step, two individuals are chosen to interact. During this interaction, they read each other’s states and update their own local states according to a predefined state transition function.

Recently, several papers have investigated fundamental space-time complexity trade-offs in the population protocol model. Space complexity is given by maximum number of local states used by the protocol. Time complexity in this model is studied under a stochastic scheduler that picks two individuals to interact uniformly at random. The (parallel) time complexity is given by the expected number of steps to reach a stable configuration divided by the total population size  $n$ . A long series of papers have investigated how fast majority can be computed by protocols that use a given number of states [11, 12, 17, 18]. Many of the recent results and techniques in this area are surveyed by Elsässer and Radzik [19] and Alistarh and Gelashvili [1].

The main difference to our setting is that in the stochastic population protocol model the total population size  $n$  is assumed to be static and all pairwise interactions occur at the same rate. In contrast, we consider systems, where the population sizes fluctuate and interaction rates may depend on the current configuration of the system. This allows us to more accurately model biological dynamics occurring in microbial populations.

**Approximate vs. exact majority.** The majority problem has been studied from two perspectives by investigating protocols that compute either *approximate* or *exact* majority. Both types of protocols typically guarantee consensus, but approximate majority protocols give only probabilistic guarantees on stabilization to the majority opinion, provided that the initial discrepancy between the two initial opinions is sufficiently large. In contrast, exact majority protocols guarantee that majority consensus is always eventually reached.

**Protocols for approximate majority.** Angluin et al. [8] analyzed a simple 3-state protocol for solving *approximate* majority in the clique in  $O(\log n)$  time provided that the discrepancy between majority and minority value is  $\omega(\sqrt{n} \log n)$ . Despite its simplicity, the protocol is challenging to analyze. Later, Mertzios et al. [29] studied protocols for approximate agreement when the interactions are restricted to occur on general interaction graphs. In a recent work, Alistarh et al. [5] study a generalized approximate majority problem where both initial opinions may have small initial counts. They present logarithmic space and time protocols that are self-stabilizing and withstand spurious faulty reactions.

**Protocols for exact majority.** Besides approximate majority, also *exact* majority has been considered in the population protocol setting: regardless of the initial gap, the final outputs should stabilize to the initial majority value. Draief and Vojnović gave a simple 4-state protocol

that solves exact majority in any connected interaction graph [18]. In the clique, the protocol stabilizes in  $O(n \log n)$  parallel time.

Alistarh et al. [2] gave the first protocols with expected converge time of the order polylog  $n$  using polylog  $n$  states. This result has been gradually improved [3, 11, 12], and recently, Doty et al. [17] gave protocols that solve majority in  $O(\log n)$  time using  $O(\log n)$  states. In the spatial setting, Alistarh et al. [4] showed how to solve exact majority in polylog  $n$  parallel time using polylog  $n$  states per node in regular graphs with constant conductance.

## 2.2 Other Models of Molecular and Biological Dynamics

Czyzowicz et al. [15] investigated the convergence of discrete Lotka–Volterra dynamics in the population protocol model. These dynamics can be used to model, for example, predator-prey type dynamics. However, they operate in the population protocol model, so while the proportions of the different species can change, the total population size is static throughout.

Condon et al. [14] investigated approximate agreement using multimolecular reactions in the chemical reaction network model. Similarly to our result, they show that if the gap is  $\Omega(\sqrt{n \log n})$ , then the majority opinion will with high probability win. However, their work does not consider biological population dynamics.

Closely related to our work, Cho et al. [13] showed how to use competitive interactions to solve approximate agreement in a two-species bacterial model. However, the studied model ignores resource limitations by assuming unbounded (exponential) growth and assuming that death only occurs during competitive interactions between individuals of different species. While unbounded growth may be a satisfactory model for short-term population dynamics in a cell culture, it is less so for mid- and long-term population dynamics.

In our model, we incorporate individual death reactions, and resource-consumer dynamics that limit the growth of the species. This leads to a models that better reflect microbiological population dynamics. To analyze our models, similarly as Cho et al. [13], we use a coupling argument to bound the behavior of the two-species dynamics. However, our coupling argument is more general, and it allows us to handle a larger, more realistic, class of models.

## 3 Model

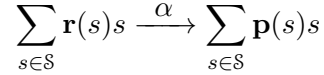
We start with some basic notation. We write  $\mathbb{N} = \{0, 1, \dots\}$  for the set of non-negative integers. For functions  $f, g : X \rightarrow \mathbb{R}$ , we write  $f \preceq g$  if  $f(x) \leq g(x)$  for all  $x \in X$ . For two nonempty sets  $A$  and  $B$ , we write  $A^B$  to denote the set of functions of the form  $B \rightarrow A$ . When convenient, we treat  $\mathbf{c} \in A^B$  as a vector with  $|B|$  elements.

### 3.1 Biological Reaction Networks

We model the stochastic behavior of bacteria over time by a *biological reaction network* (BRN) that comprises of a set of *species*  $\mathcal{S}$  a subset  $\mathcal{T} \subseteq \mathcal{S}$  of which are growing *cell types*, a set of *reactions*  $\mathcal{R}$  that act on these counts, and a volume  $v \in \mathbb{R}_0^+$ . For simplicity we assume  $v = 1$ , unless specified otherwise. While close to chemical reaction networks (CRNs), our model is a generalization with kinetics that do not necessarily follow a mass-action law.

A *configuration* of the BRN is an element from the set of configurations  $\mathcal{C} = \mathbb{N}^{\mathcal{S}}$ , where  $\mathbf{c}(s)$  denotes the count of species  $s$  in configuration  $\mathbf{c}$ . Intuitively a configuration is a snapshot of the current state of the BRN. A *reaction* from  $\mathcal{R}$  is a triple  $(\mathbf{r}, \mathbf{p}, \alpha)$  where  $\mathbf{r}, \mathbf{p} \in \mathbb{N}^{\mathcal{S}}$  and  $\alpha$ , called the *propensity function* of the reaction, is a function from  $\mathcal{C} \rightarrow \mathbb{R}_0^+$ . The *propensity* of the reaction in configuration  $\mathbf{c} \in \mathcal{C}$  is  $\alpha(\mathbf{c})$ . The species  $s$  with  $\mathbf{r}(s) > 0$  are called the reaction’s *reactants* and

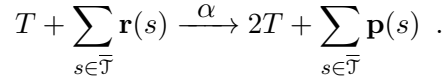
those with  $\mathbf{p}(s) > 0$  are called *products*. We will also use the common notation



for a reaction  $(\mathbf{r}, \mathbf{p}, \alpha)$ . A reaction  $(\mathbf{r}, \mathbf{p}, \alpha)$  is *applicable* to configuration  $\mathbf{c}$  if  $\mathbf{r}(s) \leq \mathbf{c}(s)$  for all  $s \in \mathcal{S}$ . We call  $\sum_{s \in \mathcal{S}} \mathbf{r}(s)$ , i.e., the number of involved reactants, the *order* of the reaction. For the moment assume that reactions are at least of order 1; we will later argue that this is without loss of generality.

We assume that there are two kinds of propensity functions: *growth reactions* and *mass-action reactions*. They are specified as follows:

- (1) *Growth reactions*: Let  $\bar{\mathcal{T}} = \mathcal{S} \setminus \mathcal{T}$  be the species that are not cell types. A growth reaction for cell type  $T \in \mathcal{T}$  is of the form



Intuitively the reactants from  $\bar{\mathcal{T}}$  account for resources and waste products in the growth medium that are (potentially) consumed and that determine the cell's growth rate. Assuming that cells grow independently of each other, we have that

$$\alpha(\mathbf{c}) = \mathbf{c}(T) \cdot \gamma(\mathbf{c}) ,$$

where  $\gamma(\mathbf{c})$  is the individual cell growth rate. We further assume that  $\gamma(\mathbf{c})$  is bounded by a maximal growth rate  $\Gamma$ . This is motivated by the observation that cells do not increase their growth rate arbitrarily with the number of available resources in the medium.

- (2) *Mass-action reactions*: For reactions other than growth reactions, we assume the classical mass-action kinetics to hold. Their propensity function  $\alpha$  is given as

$$\alpha(\mathbf{c}) = \frac{\xi}{v^{o-1}} \prod_{s \in \mathcal{S}} \binom{\mathbf{c}(s)}{\mathbf{r}(s)} ,$$

where  $o$  is the order of the reaction,  $\binom{\mathbf{c}(s)}{\mathbf{r}(s)}$  denotes the binomial coefficient of  $\mathbf{c}(s)$  and  $\mathbf{r}(s)$ , and  $\xi \geq 0$  is the *rate constant* of the reaction. We follow the convention that the binomial coefficient is 1 if  $\mathbf{r}(s) = 0$ , and it is 0 if  $\mathbf{r}(s) > \mathbf{c}(s)$ . By slight abuse of notation we write  $A + \dots \xrightarrow{\xi} B + \dots$  instead of  $A + \dots \xrightarrow{\alpha} B + \dots$  for mass-action reactions.

**BRNs as Markov chains.** The *stochastic kinetics* of a BRN from some *initial configuration* of the BRN is a continuous-time Markov chain defined as follows. The states of the Markov chain are the configurations of the BRN with the initial state being the initial configuration of the BRN. Given that the BRN is in configuration  $\mathbf{c}$ , the new configuration after an applicable reaction  $(\mathbf{r}, \mathbf{p}, \alpha)$  is equal to  $\mathbf{c}' = \mathbf{c} - \mathbf{r} + \mathbf{p}$ .

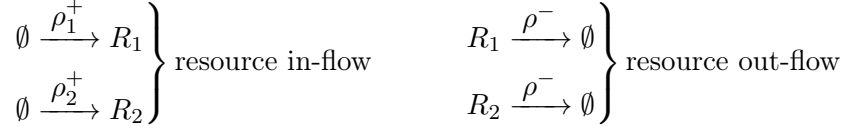
We will use the notation  $Q(x, y)$  for the propensity of the transition from state  $x$  to state  $y$  in a continuous-time Markov chain. For each continuous-time Markov chain, we can associate a corresponding a discrete-time Markov chain that only keeps track of the sequence of state changes, but not of their timing. We use  $P(x, y)$  for the transition probability from state  $x$  to state  $y$  in the discrete-time chain. We have the formula

$$P(x, y) = Q(x, y) / \sum_z Q(x, z) .$$

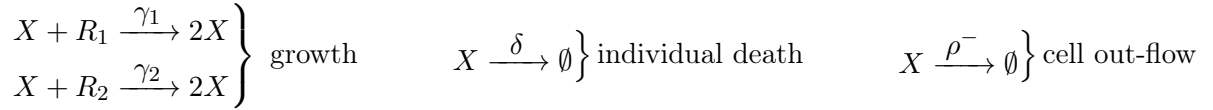
Finally note that BRNs with reactions of order 0 can be rewritten into ones with reactions of order at least 1 such that both BRNs behave identically on the original BRN's species. This is achieved by introducing a dummy species as a reactant and a product.

### 3.2 Two-Resource Protocols with Cell Types

Cells replicate dependent on the resources within the medium and die with a certain rate. Typically such resource-dependent behavior is modeled with cell growth rates that depend on one or two limiting resources [31]. In our model such a system is a BRN whose species are partitioned into proliferating cells with birth and death reactions, and a set of passive resources. For that purpose, we define: Let  $\gamma_1, \gamma_2$  be propensity functions and  $\delta, \rho_1^+, \rho_2^+, \rho^- \geq 0$  be rate constants. A *two-resource protocol with cell types*  $\mathcal{T}$ , is a BRN with set of species  $\mathcal{S} = \{R_1, R_2\} \cup \mathcal{T}$ , cell types  $\mathcal{T}$ , and reactions



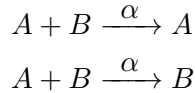
and, for every cell type  $X \in \mathcal{T}$ , the reactions



All reactions except the growth reactions are mass-action reactions.

In this work we analyze a specific two-resource protocol with two competing cell types: the *mutual annihilation protocol*. The protocol is exemplary for interference competition between two bacteria types  $A$  and  $B$ . While this protocol can be seen as an instance of naturally occurring direct attacks between  $A$  and  $B$ , such a protocol also lend itself to be implemented in a system of synthetic bacteria. For example, a setup where the two bacterial types carry a plasmid that is lethal if introduced within the other bacterial type via conjugation upon contact.

Formally, the mutual annihilation protocol is a two-resource protocol with cell types  $\mathcal{T} = \{A, B\}$  with the additional mass-action reactions



where  $\alpha > 0$  is the annihilation rate constant. This parameter is used to model the effects of direct interference competition between two species. In particular this can be used to model, for example, mechanical attacks that puncture the cell membranes of opposing species, but also mechanisms relying on bacterial conjugation.

Intuitively this protocol should strongly amplify any difference between the initial concentrations of  $A$  and  $B$  until finally majority consensus is achieved. The following sections are devoted to show that this is indeed the case. The proof is in two major parts: In Section 4.1 we prove a bound on the number of steps that change species counts, e.g., growth events and resource out-flow, until consensus is achieved. Section 4.3 then uses this result to prove our main result of the protocol achieving majority consensus with high probability if the initial population gap between  $A$  and  $B$  is large enough.

## 4 Analysis of the Mutual Annihilation Protocol

In this section, we analyse the mutual annihilation protocol. First, we bound the number of steps until consensus (competitive exclusion). Afterwards, we show that if the initial gap is large enough, then majority consensus is reached with high probability



## 4.1 Bounding the Number of Steps Until Consensus

In this section, we upper-bound the number of steps that a reaction is applied in the mutual annihilation protocol, resulting in a change of species counts, until consensus is achieved. This step upper bound is the main ingredient for our probability bound to achieve majority consensus.

Towards this goal we introduce three Markov chains that gradually abstract the behavior of the mutual annihilation protocol while maintaining central stochastic properties: the  $\mathcal{S}$ -chain, the lower-bounding  $\mathcal{S}$ -chain, and the  $M$ -chain.

### 4.1.1 The $\mathcal{S}$ -Chain of the Mutual Annihilation Protocol

Let  $\mathcal{S}$  be a set of species of a BRN. Consider a discrete-time Markov process  $(\mathbf{S}(k))_{k \geq 0}$  on the state space of the BRN's configurations  $\mathcal{C} = \mathbb{N}^{\mathcal{S}}$  such that  $\mathbf{S}_s(k)$  is a random variable denoting the number of individuals of species  $s \in \mathcal{S}$ . A configuration  $\mathbf{c} \in \mathcal{C}$  gives the quantity of each species so that  $c_s$  is the number of individuals of species  $s$ . We call any such process an  $\mathcal{S}$ -chain.

The mutual annihilation protocol includes the cell types  $A$  and  $B$  as well as the two resource species  $R_1$  and  $R_2$ . Its stochastic behavior is fully described via a corresponding  $\mathcal{S}$ -chain with species  $\mathcal{S} = \{A, B, R_1, R_2\}$ .

### 4.1.2 The Lower-Bounding $\mathcal{S}$ -Chain

We next define an abstraction of an  $\mathcal{S}$ -chain, for the particular case of the  $\mathcal{S}$ -chain of the mutual annihilation protocol. Let  $\Gamma$  be a uniform upper bound on the sum of the birth rates  $\gamma_1(r_1)$  and  $\gamma_2(r_2)$  of cells  $A$  and  $B$ , respectively, which exists by the assumption on BRN growth reactions.

Given an initial state  $(a, b, r_1, r_2)$  of the  $\mathcal{S}$ -chain of the mutual annihilation protocol with  $A(0) = a > b = B(0)$ , we define the *lower-bounding  $\mathcal{S}$ -chain* as a two-species chain with initial state  $(a, b)$ . In this chain, the majority species  $A$  has birth rate 0 and the minority species  $B$  has birth rate  $\Gamma$ . There are no resource species in the lower-bounding chain.

It is straightforward to show that the probability of reaching majority consensus in the lower-bounding  $\mathcal{S}$ -chain is a lower bound on the probability of reaching majority consensus in the  $\mathcal{S}$ -chain of the mutual annihilation protocol.

### 4.1.3 The $M$ -Chain

Towards the goal of abstracting  $\mathcal{S}$ -chains, and in particular the lower-bounding  $\mathcal{S}$ -chain of the mutual annihilation protocol, we introduce the  $M$ -chain which tracks the population size of the cell type that is currently the minimum among the cell types.

Formally, an  $M$ -chain is a discrete-time birth-death Markov process  $(M(k))_{k \geq 0}$  on the state space  $\mathbb{N}$  with birth probability function  $p' : \mathbb{N} \rightarrow [0, 1]$  and death probability function  $q' : \mathbb{N} \rightarrow [0, 1]$ , where  $p'(m) + q'(m) \leq 1$  for all  $m \in \mathbb{N}$ : a population of size  $m \geq 0$  increases by one in the next step with probability  $p'(m)$  and decreases in the next step with probability  $q'(m)$ .

Consider an  $\mathcal{S}$ -chains with the two cell types  $A$  and  $B$ . For such an  $\mathcal{S}$ -chain define the sequence of random variables

$$\text{Min}(k) = \min \{\mathbf{S}_A(k), \mathbf{S}_B(k)\}$$

and set  $p, q: \mathcal{C} \rightarrow [0, 1]$  as follows:

$$\begin{aligned} p(\mathbf{c}) &= \Pr[\text{Min}(k+1) = \text{Min}(k) + 1 \mid \mathbf{S}(k) = \mathbf{c}] \\ q(\mathbf{c}) &= \Pr[\text{Min}(k+1) = \text{Min}(k) - 1 \mid \mathbf{S}(k) = \mathbf{c}] . \end{aligned}$$

That is,  $p(\mathbf{c})$  gives the probability of the process transitioning from state  $\mathbf{c}$  to a state  $\mathbf{c}'$ , where the minimum of  $c_A$  and  $c_B$  increases by one in the next step. Analogously,  $q(\mathbf{c})$  gives the probability

that the minimum decreases by one during the next transition. Note that the probability of  $\text{Min}(k)$  increasing or decreasing may depend on the entire configuration  $\mathbf{S}(k)$ .

We say an  $M$ -chain *dominates* the  $\mathcal{S}$ -chain if for all  $\mathbf{c} \in \mathbb{N}^{\mathcal{S}}$  with  $m = \min\{\mathbf{c}_A, \mathbf{c}_B\}$ , functions  $p', q'$  satisfy

$$p(\mathbf{c}) \leq p'(m) \quad (1)$$

$$q(\mathbf{c}) \geq q'(m) \quad (2)$$

$$p(\mathbf{c}) \leq 1 - q'(m + 1) . \quad (3)$$

#### 4.1.4 Coupling the $\mathcal{S}$ -Chain to the $M$ -Chain

In the following we construct a coupling  $(\widehat{\mathbf{S}}, \widehat{M})$  of the two  $\mathcal{S}$ -chain and a dominating  $M$ -chain. Analogously to the variable  $\text{Min}$ , we define the variable

$$\widehat{\text{Min}}(k) = \min \left\{ \widehat{\mathbf{S}}_A(k), \widehat{\mathbf{S}}_B(k) \right\}$$

for all  $k \geq 0$  on the coupling. Moreover, we use  $p$  and  $q$  to denote the transition-probability functions for the  $\mathbf{S}$ -chain, and  $p'$  and  $q'$  the transition probability functions for a dominating  $M$ -chain. We will next define the discrete time processes  $\widehat{\mathbf{S}}$  and  $\widehat{M}$  inductively.

**Initially:** Set  $\widehat{\mathbf{S}}(0) = \mathbf{S}(0)$  and  $\widehat{M}(0) = M(0)$ .

**Step:** Let  $(\xi(k))_{k \in \mathbb{N}}$  be a sequence of i.i.d. random values sampled uniformly from the unit interval  $[0, 1)$ . Let  $k \in \mathbb{N}$ . Assuming  $(\widehat{\mathbf{S}}(k), \widehat{M}(k)) = (\mathbf{c}, m)$ , we set  $(\widehat{\mathbf{S}}(k+1), \widehat{M}(k+1))$  as follows:

**Minimum increases.** If  $\xi(k) \in [0, p(\mathbf{c}))$ , then sample  $\widehat{\mathbf{S}}(k+1)$  according to the conditional distribution

$$\mu(\mathbf{c}') = \Pr \left[ \widehat{\mathbf{S}}(k+1) = \mathbf{c}' \mid \widehat{\mathbf{S}}(k) = \mathbf{c} \text{ and } \widehat{\text{Min}}(k+1) = \widehat{\text{Min}}(k) + 1 \right] .$$

Respectively, if  $\xi(k) \in [0, p'(m))$ , then sample  $\widehat{M}(k+1)$  according to the conditional distribution

$$\mu(m') = \Pr \left[ \widehat{M}(k+1) = m' \mid \widehat{M}(k) = m \text{ and } \widehat{M}(k+1) = \widehat{M}(k) + 1 \right] .$$

**Minimum decreases.** If  $\xi(k) \in [1 - q(\mathbf{c}), 1)$ , then sample  $\widehat{\mathbf{S}}(k+1)$  according to the conditional distribution

$$\mu(\mathbf{c}') = \Pr \left[ \widehat{\mathbf{S}}(k+1) = \mathbf{c}' \mid \widehat{\mathbf{S}}(k) = \mathbf{c} \text{ and } \widehat{\text{Min}}(k+1) = \widehat{\text{Min}}(k) - 1 \right] .$$

Respectively, if  $\xi(k) \in [1 - q'(m), 1)$ , then sample  $\widehat{M}(k+1)$  according to the conditional distribution

$$\mu(m') = \Pr \left[ \widehat{M}(k+1) = m' \mid \widehat{M}(k) = m \text{ and } \widehat{M}(k+1) = \widehat{M}(k) - 1 \right] .$$

**Minimum does not change.** Otherwise, sample  $\widehat{\mathbf{S}}(k+1)$  according to the conditional distribution

$$\mu(\mathbf{c}') = \Pr \left[ \widehat{\mathbf{S}}(k+1) = \mathbf{c}' \mid \widehat{\mathbf{S}}(k) = \mathbf{c} \text{ and } \widehat{\text{Min}}(k+1) = \widehat{\text{Min}}(k) \right] .$$

Respectively, sample  $\widehat{M}(k+1)$  according to the conditional distribution

$$\mu(m') = \Pr \left[ \widehat{M}(k+1) = m' \mid \widehat{M}(k) = m \text{ and } \widehat{M}(k+1) = \widehat{M}(k) \right] .$$

Observe that by construction, the marginal distributions of  $\widehat{\mathbf{S}}(k)$  and  $\widehat{M}(k)$  equal the distributions of  $\mathbf{S}(k)$  and  $M(k)$  for all  $k \in \mathbb{N}$ .

We will next show that  $\widehat{\text{Min}} \preceq \widehat{M}$  in the coupled process under certain dominance conditions of the transition probabilities in the original  $\mathcal{S}$ -chain and a dominating  $M$ -chain.

**Lemma 1.** *Given an  $\mathcal{S}$ -chain and a dominating  $M$ -chain,  $\text{Min}(0) \leq M(0)$  implies  $\widehat{\text{Min}} \preceq \widehat{M}$ .*

*Proof.* We show by induction that  $\widehat{\text{Min}}(k) \leq \widehat{M}(k)$  for all  $k \in \mathbb{N}$ . The base case  $k = 0$  is vacuous. For the inductive step, suppose  $\widehat{\text{Min}}(k) \leq \widehat{M}(k)$  holds for some  $k \geq 0$ . Observe that since  $|\widehat{\text{Min}}(k+1) - \widehat{\text{Min}}(k)| \leq 1$  and  $|\widehat{M}(k+1) - \widehat{M}(k)| \leq 1$ , the claim follows if:

$$\widehat{\text{Min}}(k+1) < \widehat{\text{Min}}(k) \quad \text{or} \quad (4)$$

$$\widehat{M}(k+1) > \widehat{M}(k) . \quad (5)$$

Let  $\mathbf{c} = \widehat{\mathbf{S}}(k)$  and  $m = \min\{\mathbf{c}_A, \mathbf{c}_B\}$ . We distinguish two cases:

(1) Suppose  $\widehat{\text{Min}}(k) = \widehat{M}(k) = m$ . To show that  $\widehat{\text{Min}}(k+1) \leq \widehat{M}(k+1)$ , we consider the following subcases:

(1) If  $\widehat{\text{Min}}(k+1) = m+1$ , then  $\xi(k) \in [0, p(\mathbf{c})] \subseteq [0, p'(m))$ , by Assumption (1) of the lemma. Now the update rule of the coupling yields  $\widehat{M}(k+1) = m+1 > \widehat{M}(k)$ ; i.e., (5) is fulfilled and the claim follows.

(2) If  $\widehat{M}(k+1) = m-1$ , then  $\xi(k) \in [1-q'(m), 1] \subseteq [1-q(\mathbf{c}), 1)$ , by Assumption (2) of the lemma. Thus, the update rule implies  $\widehat{\text{Min}}(k+1) = m-1 < \widehat{\text{Min}}(k)$ ; i.e., (4) is fulfilled and the claim follows.

(3) Otherwise,  $\widehat{\text{Min}}(k+1) \leq m$  and  $\widehat{M}(k+1) \geq m$ ; the claim follows in this case.

Thus, in all three cases the claim follows.

(2) Otherwise, by the induction hypothesis,  $\widehat{\text{Min}}(k) < \widehat{M}(k)$ . If  $\widehat{M}(k) - \widehat{\text{Min}}(k) > 1$ , then the claim follows immediately, since both variables can change by at most one per step and thus no reordering can happen.

Hence, suppose that  $\widehat{M}(k) = \widehat{\text{Min}}(k) + 1$  holds. The only remaining case is the event where  $\widehat{\text{Min}}(k+1) = \widehat{\text{Min}}(k) - 1 = m$  and  $\widehat{M}(k+1) = \widehat{M}(k) + 1 = m+1$ . This implies that  $\xi(k) \in [0, p(\mathbf{c})] \cap [1-q'(m+1), 1)$ . Thus,  $p(\mathbf{c}) > 1-q'(m+1)$ , contradicting Assumption (3) of the lemma. Therefore, the case where  $\widehat{\text{Min}}$  increments and  $\widehat{M}$  decrements does not occur.  $\square$

#### 4.1.5 An $M$ -Chain for the Lower-Bounding $\mathcal{S}$ -Chain

We first define  $p'(m)$  by setting

$$p'(m) = \frac{m\Gamma}{2m\rho^- + m\Gamma + 2m\delta + 2m^2\alpha}$$

The maximum of  $p'$  is achieved for  $m = 1$ . This maximum is strictly smaller than 1 because  $p'(m) < 1$  for all  $m \in \mathbb{N}$ . Call the maximum  $P$ .

We then define  $q'(m)$  by setting

$$q'(m) = \min \left\{ 1 - P, \frac{m^2\alpha}{2m\rho^- + m\Gamma + 2m\delta + 2m^2\alpha} \right\} .$$

We observe that  $p'(m) = O(1/m)$  and  $q'(m) = \Omega(1)$ .

**Lemma 2.** *The functions  $p'$  and  $q'$  define an  $M$ -chain that dominates the lower-bounding chain of the mutual annihilation protocol.*

*Proof.* We first prove Condition (1). Assume  $m = b < a$ . Then we have:

$$\begin{aligned} p(a, b) &= \frac{b\Gamma}{(a+b)\rho^- + b\Gamma + (a+b)\delta + 2ab\alpha} \\ &\leq \frac{b\Gamma}{2b\rho^- + b\Gamma + 2b\delta + 2b^2\alpha} = p'(b) = p'(m) \end{aligned}$$

If  $m = a \leq b$ , then  $p(a, b) = 0$  and the inequality trivially holds.

We next prove Condition (2). Assume  $m = b < a$ . Then we have:

$$\begin{aligned} q(a, b) &= \frac{b\rho^- + b\delta + ab\alpha}{(a+b)\rho^- + b\Gamma + (a+b)\delta + 2ab\alpha} \\ &\geq \frac{ab\alpha}{(a+b)\rho^- + b\Gamma + (a+b)\delta + 2ab\alpha} \\ &\geq \frac{ab\alpha}{2a\rho^- + a\Gamma + 2a\delta + 2ab\alpha} \\ &= \frac{b\alpha}{2\rho^- + \Gamma + 2\delta + 2b\alpha} \\ &= \frac{b^2\alpha}{2b\rho^- + b\Gamma + 2b\delta + 2b^2\alpha} \geq q'(b) = q'(m) \end{aligned}$$

If  $m = a \leq b$ , then:

$$\begin{aligned} q(a, b) &= \frac{a\rho^- + a\delta + ab\alpha}{(a+b)\rho^- + b\Gamma + (a+b)\delta + 2ab\alpha} \\ &\geq \frac{ab\alpha}{(a+b)\rho^- + b\Gamma + (a+b)\delta + 2ab\alpha} \\ &\geq \frac{ab\alpha}{2b\rho^- + b\Gamma + 2b\delta + 2ab\alpha} \\ &= \frac{a\alpha}{2\rho^- + \Gamma + 2\delta + 2a\alpha} \\ &= \frac{a^2\alpha}{2a\rho^- + a\Gamma + 2a\delta + 2a^2\alpha} \geq q'(a) = q'(m) \end{aligned}$$

Lastly, Condition (3) easily follows from Condition (1) and the definition of  $q'(m)$  since

$$p(a, b) + q'(m+1) \leq p'(m) + (1-P) \leq P + (1-P) = 1$$

where  $m = \min\{a, b\}$ . □

#### 4.1.6 Number of Steps in the $M$ -Chain

We now show that the number of steps until extinction in the  $M$ -chain is at most linear in the initial population, both in expectation (Lemma 3) and with high probability (Lemma 4). Because the  $M$ -chain dominates the  $\mathcal{S}$ -chain, the upper bound also holds for the number of steps until consensus in the lower-bounding  $\mathcal{S}$ -chain.

**Lemma 3.** *The expected number of steps until extinction of any  $M$ -chain with  $p'(m) = O(1/m)$  and  $q'(m) = \Omega(1)$  is  $O(M)$  where  $M$  is the initial state.*

*Proof.* Let  $C \geq 1$  be a big-oh constant for  $p'(m)$ , i.e.,  $p'(m) \leq C/m$  for all  $m \geq 1$ . Let  $D > 0$  be a big-omega constant for  $q'(m)$ , i.e.,  $q'(m) \geq D$  for all  $m \geq 1$ . From known results for discrete-time birth-death processes [27], setting  $\alpha = C/D$ , we get that the expected number of steps until extinction from initial state  $M$  is equal to

$$\begin{aligned} \sum_{j=1}^M \sum_{k=j-1}^{\infty} \frac{p'(j) \cdots p'(k)}{q'(j) \cdots q'(k+1)} &\leq \sum_{j=1}^M \sum_{k=j-1}^{\infty} \frac{C^{k-j+1} (j-1)!}{D^{k-j+2} k!} = \frac{1}{D} \sum_{j=1}^M \sum_{k=j-1}^{\infty} \alpha^{k-j+1} \frac{(j-1)!}{k!} \\ &\leq \frac{1}{D} \sum_{j=1}^M \sum_{k=j-1}^{\infty} \alpha^{k-j+1} \frac{1}{(k-j+1)!} = \frac{1}{D} \sum_{j=1}^M \sum_{k=0}^{\infty} \alpha^k \frac{1}{k!} \\ &= \frac{1}{D} \sum_{j=1}^M e^{\alpha} = O(M). \end{aligned}$$

Here, we used the inequality  $\frac{(j-1)!}{k!} \leq \frac{1}{(k-j+1)!}$ , which is equivalent to  $\binom{k}{j-1} \geq 1$ .  $\square$

**Lemma 4.** *The number of steps until extinction of any  $M$ -chain with  $p'(m) = O(1/m)$  and  $q'(m) = \Omega(1)$  is  $O(M)$  with probability  $1 - O(1/\sqrt{M})$  where  $M$  is the initial state.*

*Proof.* We distinguish two phases: the first from states  $M$  to  $\Theta(\sqrt{M})$  and the second from  $\Theta(\sqrt{M})$  to 0.

For the first phase, we start by bounding the number of birth reactions and then the number of stuttering steps to show that enough death reactions occur. Let  $D > 0$  be a big-omega constant for  $q'(m)$ , i.e.,  $q'(m) \geq D$  for all  $m \geq 1$ . The probability of an individual step being a stuttering step is at most  $\beta = 1 - D < 1$ . Let  $C \geq 1$  be a big-oh constant for  $p'(m)$ , i.e.,  $p'(m) \leq C/m$  for all  $m \geq 1$ . We pose  $KM$  as an upper bound on the number of steps of the first phase where  $K \geq \frac{2}{1-\beta}$ .

The expected number of stuttering steps in the first  $KM$  steps is at most  $\mu \leq \beta KM$ . Setting  $\delta = \frac{1-\beta}{2\beta}$ , by the Chernoff bound, the probability of having more than  $(1+\delta)\beta KM = \frac{1+\beta}{2} KM$  stuttering steps in the first  $KM$  steps is upper-bounded by  $e^{-\delta^2 \beta KM/3} = e^{-\Omega(M)}$ . By the choice of  $K$ , the same bound holds for the probability that there are less than  $M$  non-stuttering steps in the first  $KM$  steps.

The first phase ends when a state  $\leq 4C\sqrt{M}$  is reached. We have  $m \geq \sqrt{M}$  and thus  $p'(m) \leq C/\sqrt{M}$  in particular in the first phase. Let  $E$  be the event that a state  $\leq 4C\sqrt{M}$  is reached in the first  $M$  non-stuttering steps. The event that the number  $b$  of births in the first  $M$  non-stuttering steps is at most  $2C\sqrt{M}$  implies event  $E$ . Therefore, the inverse event  $\neg E$  implies  $b > 2C\sqrt{M}$ . By the Chernoff bound, the probability of  $\neg E$  is bounded by

$$\Pr[\neg E] \leq \Pr[b \geq 2C\sqrt{M}] = \Pr[b \geq (1+\delta)\mu] \leq e^{-\delta^2 \mu/3}$$

where  $\mu \leq C\sqrt{M}$  and  $\delta = 1$ . We thus have:

$$\Pr[\neg E] = e^{-\Omega(\sqrt{M})}$$

In the second phase, denote by  $L$  the number of events until extinction. By Lemma 3, the expected value of  $L$  is upper-bounded by  $\mathbb{E} L = O(\sqrt{M})$ . By Markov's inequality we thus have:

$$\Pr[L > M] \leq \frac{\mathbb{E} L}{M} = O(1/\sqrt{M})$$

Combining the analyses of both phases shows that extinction happens in the first  $(K+1)M$  steps with high probability.  $\square$

## 4.2 Number of Steps Until Consensus in the Lower-Bounding $\mathcal{S}$ -Chain

We can now prove our probability bound for majority consensus with the mutual annihilation protocol.

**Lemma 5.** *The number of steps until consensus in the lower-bounding  $\mathcal{S}$ -chain is  $O(n)$  with high probability where  $n$  is the total initial population size.*

*Proof.* The number of steps until consensus with a smaller initial gap and the same total population stochastically dominates the number of steps until consensus with a larger initial gap. It is hence sufficient to prove the lemma for a constant initial gap. But then  $a = \Omega(n)$  and  $b = \Omega(n)$  initially. The corresponding  $M$ -chain thus has initial state  $m = \Theta(n)$ . An invocation of Lemma 4 concludes the proof.  $\square$

## 4.3 Probability of Reaching Majority Consensus

Equipped with the step upper bound from Lemma 5, we can now prove our probability bound for majority consensus.

**Theorem 6.** *The mutual annihilation protocol achieves majority consensus with high probability whenever the initial gap is  $\Omega(\sqrt{n \log n})$  where  $n$  is the total initial population size.*

*Proof.* Without loss of generality, let  $A_0 > B_0$ . We define the discrete-time stochastic process  $X_k = A_k - B_k$  from the lower-bounding  $\mathcal{S}$ -chain. By hypothesis, we have  $X_0 = \Omega(\sqrt{n \log n})$ . As long as  $X_k \geq 0$ , the process  $(X_k)$  is a sub-martingale. We will use Azuma's inequality [30] to prove the theorem.

With high probability, the number  $K$  of steps until consensus is  $O(n)$  by Lemma 5. The maximum step size of the process  $(X_k)$  is bounded by  $|X_k - X_{k-1}| \leq 1$ . Setting  $\varepsilon$  equal to the initial gap, i.e.,  $\varepsilon = \Omega(\sqrt{n \log n})$ , Azuma's inequality thus gives:

$$\begin{aligned} \Pr[X_K \leq 0] &\leq \exp\left(\frac{-\varepsilon^2}{2 \sum_{k=1}^K 1}\right) \leq \exp\left(\frac{-\Omega(n \log n)}{2K}\right) \\ &= \exp\left(\frac{-\Omega(n \log n)}{O(n)}\right) = e^{-\Omega(\log n)} \\ &= 1/n^{\Omega(1)} \end{aligned}$$

This concludes the proof of the theorem.  $\square$

## 5 Inefficiency of Purely Indirect Competition

In this section, we evaluate the performance of the naïve protocol, which does not employ any direction interactions between the bacterial species and relies on indirect competition via shared resources only, for majority consensus.

For simplicity, we will first assume equal birth rates and death rates for both species  $A$  and  $B$ . These rates can change after each transition and depend on the current population counts. We denote by  $\gamma_{k,A,B}$  the birth rate for each bacterium after the  $k^{\text{th}}$  transition, and by  $\delta_k$  the death rate for each bacterium after the  $k^{\text{th}}$  transition.

Denote by  $P_k(A, B)$  the probability of species  $B$  being extinct before species  $A$  with the naïve protocol, starting with the populations  $(A, B)$  right after the  $k^{\text{th}}$  transition. Ultimately we are interested in the case  $k = 0$ , i.e., the probability  $P_0(A, B)$ . Almost-sure consensus can be achieved only if one of the species gets extinct almost surely. This requirement translates into a condition on the sequence of birth and death rates. In a two-resource model, it is the case in particular if there is no resource in-flow.

**Lemma 7.** *If the species  $A$  and  $B$  are symmetric and get extinct almost surely, then we have  $P_k(A, B) = \frac{A}{A+B}$  whenever  $A + B \geq 1$ .*

*Proof.* The probabilities  $P_k(A, B)$  are bounded between 0 and 1 and satisfy the recurrence

$$\begin{aligned} P_k(A, B) &= \frac{\gamma_{k,A,B}}{\gamma_{k,A,B} + \delta_{k,A,B}} \frac{A}{A+B} P_{k+1}(A+1, B) + \frac{\gamma_{k,A,B}}{\gamma_{k,A,B} + \delta_{k,A,B}} \frac{B}{A+B} P_{k+1}(A, B+1) \\ &\quad + \frac{\delta_{k,A,B}}{\gamma_{k,A,B} + \delta_{k,A,B}} \frac{A}{A+B} P_{k+1}(A-1, B) + \frac{\delta_{k,A,B}}{\gamma_{k,A,B} + \delta_{k,A,B}} \frac{B}{A+B} P_{k+1}(A, B-1) \end{aligned} \quad (6)$$

for  $A \geq 1$  and  $B \geq 1$  with the boundary conditions  $P_k(0, B) = 0$  and  $P_k(A, 0) = 1$ .

It is straightforward to verify that  $P_k(A, B) = \frac{A}{A+B}$  satisfies this recurrence: We have

$$\frac{A}{A+B} \cdot \frac{A+1}{A+B+1} + \frac{B}{A+B} \cdot \frac{A}{A+B+1} = \frac{A^2 + A + AB}{(A+B)(A+B+1)} = \frac{A}{A+B}$$

as well as

$$\frac{A}{A+B} \cdot \frac{A-1}{A+B-1} + \frac{B}{A+B} \cdot \frac{A}{A+B-1} = \frac{A^2 - A + AB}{(A+B)(A+B-1)} = \frac{A}{A+B}$$

which shows that the right-hand side of (6) is indeed equal to:

$$\frac{\gamma_{k,A,B}}{\gamma_{k,A,B} + \delta_{k,A,B}} \cdot \frac{A}{A+B} + \frac{\delta_{k,A,B}}{\gamma_{k,A,B} + \delta_{k,A,B}} \cdot \frac{A}{A+B} = \frac{A}{A+B}$$

To prove uniqueness, let  $P_k(A, B)$  and  $\hat{P}_k(A, B)$  be two solutions of the recurrence that are bounded between 0 and 1, and set  $\Delta_k(A, B) = P_k(A, B) - \hat{P}_k(A, B)$ . We will prove  $\Delta_k(A, B) = 0$  by showing  $|\Delta_k(A, B)| \leq \varepsilon$  for all  $\varepsilon > 0$ .

Let  $\varepsilon > 0$ . We write  $\mathcal{T} = \{A \nearrow, A \searrow, B \nearrow, B \searrow\}$  for the set of possible transitions, i.e., birth/death of  $A$  and birth/death of  $B$ . The differences  $\Delta_k(A, B)$  satisfy the same recurrence as  $P_k(A, B)$ , but with the boundary condition  $\Delta_k(0, B) = \Delta_k(A, 0) = 0$ . We can rewrite the recurrence as

$$\Delta_k(A, B) = \sum_{\tau \in \mathcal{T}} \Pr[\tau_{k+1} = \tau \mid A(k) = A, B(k) = B] \cdot \Delta_{k+1}(\tau(A, B)) \quad (7)$$

for all  $A, B \geq 1$ , where we denoted the  $(k+1)^{\text{th}}$  transition by  $\tau_{k+1}$  and used the transition  $\tau$  as an operator on the bacterial species counts to update them to their new values after the transition. That is, we define  $\tau(A, B) = (A+1, B)$  for the case  $\tau = A \nearrow$  and analogously for the three other transitions.

Using the recurrence (7) multiple times leads to the formula

$$\Delta_k(A, B) = \sum_{\substack{\sigma \in \mathcal{T}^\ell \\ \sigma \text{ is live}}} \Pr\left[(\tau_r)_{r=k+1}^{k+\ell} = \sigma \mid A(k) = A, B(k) = B\right] \cdot \Delta_{k+\ell}(\sigma(A, B))$$

where we call a sequence  $\sigma \in \mathcal{T}^\ell$  *live* if neither species  $A$  nor  $B$  is extinct after any of the transitions in  $\sigma$  when starting with the populations  $(A, B)$  right after the  $k^{\text{th}}$  transition. All terms of the sum that do not correspond to live sequences are zero because of the boundary condition for  $\Delta_{k+\ell}$ .

Because of the almost-sure extinction hypothesis we have:

$$\lim_{\ell \rightarrow \infty} \Pr\left[(\tau_r)_{r=k+1}^{k+\ell} \text{ is live}\right] = 0$$

There hence exists an  $\ell \geq 1$  such that the probability of  $(\tau_r)_{r=k+1}^{k+\ell}$  being live is less or equal to  $\varepsilon$ . Since the bounds on the solutions guarantee  $|\Delta_{k+\ell}(\sigma(A, B))| \leq 1$ , we get:

$$\begin{aligned} |\Delta_k(A, B)| &\leq \sum_{\substack{\sigma \in \mathcal{T}^\ell \\ \sigma \text{ is live}}} \Pr\left[(\tau_r)_{r=k+1}^{k+\ell} = \sigma \mid A(k) = A, B(k) = B\right] \cdot |\Delta_{k+\ell}(\sigma(A, B))| \\ &\leq \Pr\left[(\tau_r)_{r=k+1}^{k+\ell} \text{ is live}\right] \leq \varepsilon \end{aligned}$$

We can thus conclude  $\Delta_k(A, B) = 0$ , which proves that  $P_k(A, B) = \frac{A}{A+B}$  is the unique solution of recurrence (6) that satisfies the boundary conditions and is bounded between 0 and 1.  $\square$

To generalize to non-symmetric species, we denote by  $\gamma_{k,A,B}^{(X)}$  and  $\delta_{k,A,B}^{(X)}$  the birth and death rates of species  $X \in \{A, B\}$  after the  $k^{\text{th}}$  transition in configuration  $(A, B)$ , respectively. We say that species  $X$  *dominates* species  $Y$  if  $\gamma_{k,A,B}^{(X)} \geq \gamma_{k,A,B}^{(Y)}$  and  $\delta_{k,A,B}^{(X)} \leq \delta_{k,A,B}^{(Y)}$ .

**Theorem 8.** *If the species  $A$  and  $B$  are symmetric or if one species dominates the other, without resource in-flow, the naïve protocol cannot guarantee majority consensus to be achieved with a probability larger than the relative initial population of the majority species.*

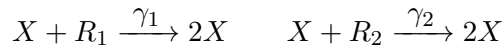
*Proof.* The lack of resource in-flow guarantees almost-sure extinction of both species. If the species are symmetric, then the claim hence follows from Lemma 7. If one species dominates the other, starting the system in a configuration in which the dominating species is in the minority, the probability of the dominated, majority species winning is upper-bounded by the probability of it winning with in a system with symmetric species.  $\square$

## 6 Simulations

In this section, we complement our analytical results with simulations to validate that the asymptotics shown in the previous sections can be already observed in realistic biological settings.

### 6.1 Setup

We modeled a culture of two bacterial types,  $A$  and  $B$ , that grow in a closed system of volume  $1 \mu\text{l}$ . Both bacterial types have identical growth behavior, feeding on two resources  $R_1$  and  $R_2$  that model components of the medium with different nutrient efficiencies. The duplication reactions for  $X \in \{A, B\}$  are thus



where we set the growth rate functions to

$$\gamma_1(R_1) = R_1/R_1(0) \cdot 1/20 \text{ min}^{-1} \quad \text{and} \quad \gamma_2(R_2) = R_2/R_2(0) \cdot 1/20 \cdot 0.08 \text{ min}^{-1} .$$

This corresponds to an initial expected duplication time of 20 min with the help of resource  $R_1$ , and only 8% of that duplication rate using resource  $R_2$ . As initial resource concentrations we chose  $[R_1] = 2 \cdot 10^6 \text{ ml}^{-1}$  and  $[R_2] = 10^8 \text{ ml}^{-1}$  resulting in a carrying capacity of about  $10^8$  bacteria per ml. Observe that the growth rates are bounded, as required by our model, since resources are not regenerated and there is no resource inflow.

Further, bacteria die with an individual death rate that was set to  $\delta = 10^{-4} \text{ min}^{-1}$ ; within the order of  $0.43 \text{ d}^{-1}$ , which was measured [36] for *E. coli*. We modeled two types of systems: one following exploitative competition, and the other following interference competition. For the second one, we assumed that each of the bacteria carries a respective plasmid that, if introduced



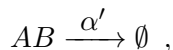
Parameter	Value	Note
Cell death rate const.	$\delta = 10^{-4} \text{ min}^{-1}$	order as in [36]
Cell growth rate for $R_1$	$\gamma_1(R_1) = R_1/R_1(0) \cdot 1/20 \text{ min}^{-1}$	20 min duplication time
Cell growth rate for $R_2$	$\gamma_2(R_2) = R_2/R_2(0) \cdot 1/20 \cdot 0.08 \text{ min}^{-1}$	reduced nutrition efficiency
Resource in-flow	$\rho_1^+ = \rho_2^+ = 0$	closed system
Resource/cell out flow	$\rho^- = 0$	closed system
Conjugation rate const.	$\alpha = 5 \cdot 10^{-10} \text{ ml}^{-1} \text{ min}^{-1}$	from [38]
Increased death rate const.	$\alpha' = 10^{-1} \text{ min}^{-1}$	death after 10 min

Table 1: Parameters of the simulation model.

via conjugation into the other bacterial type upon an interaction, leads to its death. For a more realistic setting, instead of immediate death as assumed in the previous sections, we consider a shortly lived intermediate bacterial type  $AB$  with increased death rate for bacteria that carry both plasmids. We thus have



as well as



where we chose  $\alpha = 5 \cdot 10^{-10} \text{ ml}^{-1} \text{ min}^{-1}$  in accordance with measurements of conjugating *E. coli* [38] and an increased death rate of  $\alpha' = 10^{-1} \text{ min}^{-1}$ . Table 1 summarizes the parameterization of the model.

## 6.2 Simulation Results

We ran stochastic simulations of two competing bacteria populations,  $A$  and  $B$ , with initial concentrations of  $[A] + [B] = 3 \cdot 10^5 \mu\text{l}^{-1}$ . We compared the performance of the mutual annihilation protocol to the case of resource-consumer dynamics without direct interference competition.

**Interference competition as an amplifier.** Figure 1 shows a single stochastic trajectory during the first 1400 min of simulated time with a total initial population of  $A + B = 3 \cdot 10^5$ . Here, the initial population counts of  $A$  and  $B$  have been set to differ by 2000 with  $A$  being the majority. We can clearly see that the initial difference between the two population sizes is amplified over time: already after 1400 minutes, the count of the minority species has decreased by three orders of magnitude.

Figure 2 shows the prevalence of type  $A$  after 60 min and 120 min as a function of *initial* fraction of type  $A$  individuals. Here, the initial fraction of type  $A$  varies from 0 to 1. Observe the steep s-shaped behavior that is typical for a large amplification away from the midpoint of equal concentrations. Figure 3 zooms into the middle of the s-shaped curve, with the top abscissa showing the difference  $A - B$ .

**Comparison with resource-consumer dynamics.** The dynamics of the mutual annihilation protocol show that direct *interference* competition quickly amplifies the differences between the two populations. We also compared this scenario to a setting with only indirect *exploitative* competition. To this end, we set the conjugation rate parameter  $\alpha = 0$  so that there is no direct competitive interactions. In this case, competition is mediated only by consumption of shared resources.

Figures 4 and 5 show the obtained results after 60 minutes of simulated time. We can see that compared to the case of interference competition, there is little to no amplification of

the differences under exploitative competition during the first 60 minutes of simulated time. Moreover, this holds even after a single day (1440 min), as shown by Figure 6.

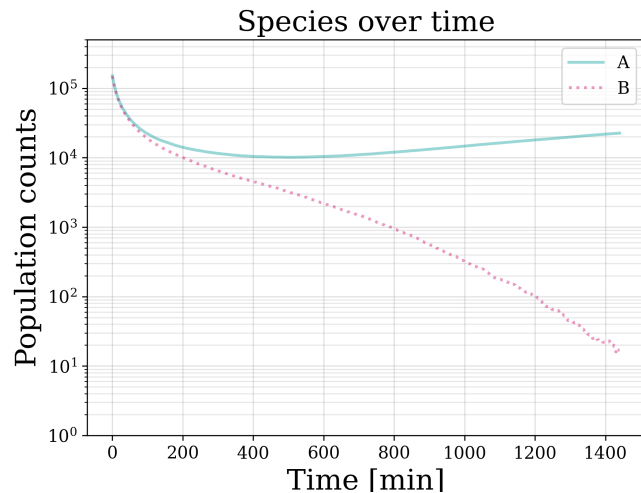


Figure 1: Stochastic simulation with initial population counts  $A = 151000$  and  $B = 149000$  over 1 day. Population counts are per  $\mu\text{l}$  and shown on a logarithmic scale.

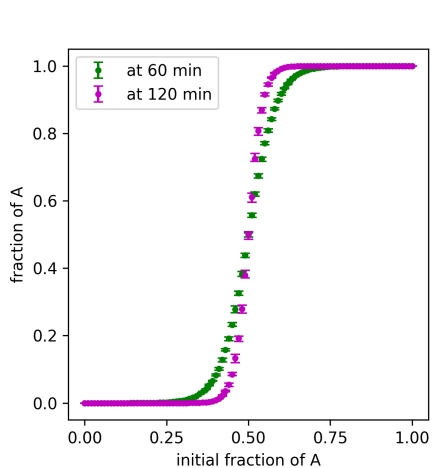


Figure 2: Fraction of  $A$  in the bacterial population after 60 min and 120 min.  $N = 10$  simulations per initial fraction. Error bars indicate maximum and minimum, markers average fractions.

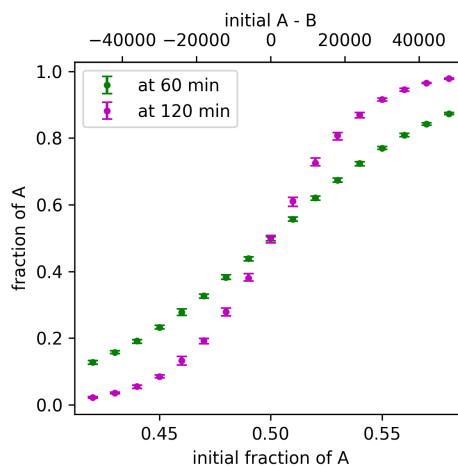


Figure 3: Zoomed version with initial population difference on the top abscissa.

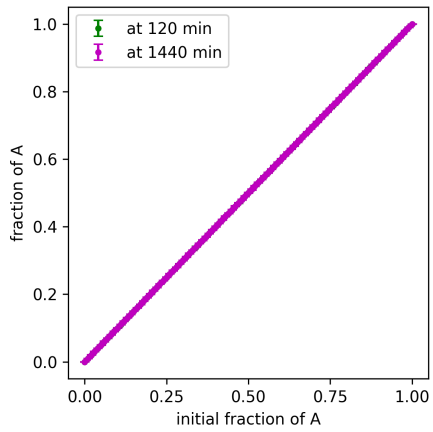


Figure 4: Same setting as in Figure 2, but with  $\alpha = 0$  and snapshot after one day instead of 60 min.

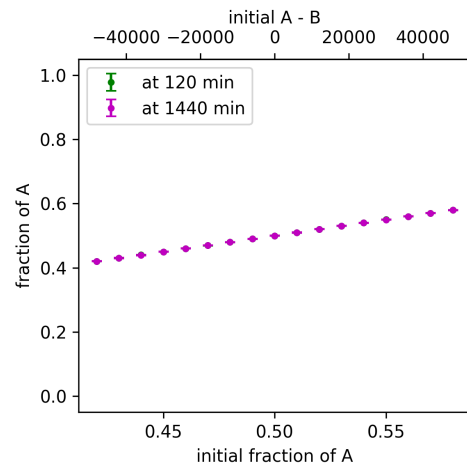


Figure 5: Zoomed version corresponding to Figure 3, but with  $\alpha = 0$  and snapshot after one day instead of 60 min.

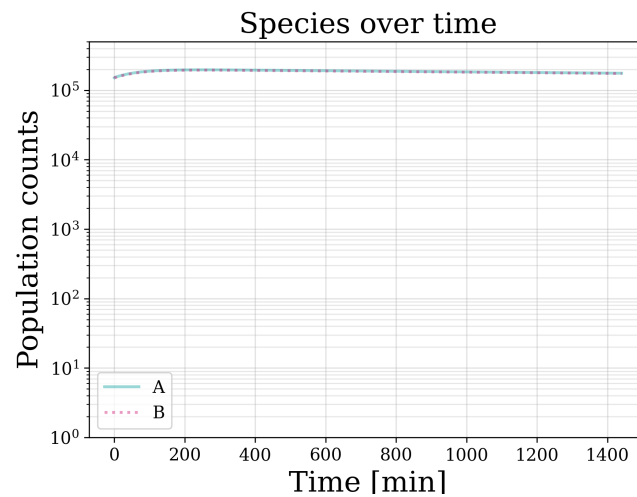


Figure 6: Stochastic simulation as in Figure 1, but with  $\alpha = 0$  and shown over 1 day.

## References

- [1] Dan Alistarh and Rati Gelashvili. Recent algorithmic advances in population protocols. *SIGACT News*, 49(3):63–73, 2018. doi:[10.1145/3289137.3289150](https://doi.org/10.1145/3289137.3289150).
- [2] Dan Alistarh, Rati Gelashvili, and Milan Vojnović. Fast and exact majority in population protocols. In *Proc. 34th ACM Symposium on Principles of Distributed Computing (PODC 2015)*, pages 47–56, 2015. doi:[10.1145/2767386.2767429](https://doi.org/10.1145/2767386.2767429).
- [3] Dan Alistarh, James Aspnes, David Eisenstat, Rati Gelashvili, and Ronald L Rivest. Time-space trade-offs in population protocols. In *Proc. 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2560–2579, 2017.
- [4] Dan Alistarh, Rati Gelashvili, and Joel Rybicki. Fast graphical population protocols. Manuscript, 2021. URL <https://arxiv.org/abs/2102.08808>.
- [5] Dan Alistarh, Martin Töpfer, and Przemysław Uznański. Robust comparison in population protocols, 2021.
- [6] Dana Angluin, James Aspnes, Zoë Diamadi, Michael J. Fischer, and René Peralta. Computation in networks of passively mobile finite-state sensors. *Distributed Computing*, 18(4): 235–253, 2006. doi:[10.1007/s00446-005-0138-3](https://doi.org/10.1007/s00446-005-0138-3).
- [7] Dana Angluin, James Aspnes, David Eisenstat, and Eric Ruppert. The computational power of population protocols. *Distributed Computing*, 20(4):279–304, 2007. doi:[10.1007/s00446-007-0040-2](https://doi.org/10.1007/s00446-007-0040-2).
- [8] Dana Angluin, James Aspnes, and David Eisenstat. A simple population protocol for fast robust approximate majority. *Distributed Computing*, 21(2):87–102, 2008.
- [9] James Aspnes and Eric Ruppert. An introduction to population protocols. In *Middleware for Network Eccentric and Mobile Applications*, pages 97–120. Springer, 2009. doi:[10.1007/978-3-540-89707-1\\_5](https://doi.org/10.1007/978-3-540-89707-1_5).
- [10] Bonnie L Bassler and Richard Losick. Bacterially speaking. *Cell*, 125(2):237–246, 2006. doi:[10.1016/j.cell.2006.04.001](https://doi.org/10.1016/j.cell.2006.04.001).
- [11] Stav Ben-Nun, Tsvi Kopelowitz, Matan Kraus, and Ely Porat. An  $O(\log^{3/2} n)$  parallel time population protocol for majority with  $O(\log n)$  states. In *Proc. 2020 ACM Symposium on Principles of Distributed Computing*, pages 191–199, 2020. doi:[10.1145/3382734.3405747](https://doi.org/10.1145/3382734.3405747).
- [12] Petra Berenbrink, Robert Elsässer, Tom Friedetzky, Dominik Kaaser, Peter Kling, and Tomasz Radzik. A population protocol for exact majority with  $O(\log^{5/3} n)$  stabilization time and  $\Theta(\log n)$  states. In *Proc. 32nd International Symposium on Distributed Computing (DISC 2018)*, pages 10:1–10:18, 2018. doi:[10.4230/LIPIcs.DISC.2018.10](https://doi.org/10.4230/LIPIcs.DISC.2018.10).
- [13] Da-Jung Cho, Matthias Függer, Corbin Hopper, Manish Kushwaha, Thomas Nowak, and Quentin Soubeyran. Distributed computation with continual population growth. In *Proc. 34th International Symposium on Distributed Computing (DISC)*, pages 7:1–7:17, 2020. doi:[10.4230/LIPIcs.DISC.2020.7](https://doi.org/10.4230/LIPIcs.DISC.2020.7).
- [14] Anne Condon, Monir Hajiaghayi, David Kirkpatrick, and Ján Maňuch. Approximate majority analyses using tri-molecular chemical reaction networks. *Natural Computing*, 19: 249–270, 2020.

- [15] Jurek Czyzowicz, Leszek Gasiniec, Adrian Kosowski, Evangelos Kranakis, Paul G. Spirakis, and Przemysław Uznański. On convergence and threshold properties of discrete Lotka-Volterra population protocols. In *Proc. 42nd International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 393–405, 2015. doi:[10.1007/978-3-662-47672-7\\_32](https://doi.org/10.1007/978-3-662-47672-7_32).
- [16] David Doty. Theory of algorithmic self-assembly. *Communications of the ACM*, 55(12): 78–88, 2012. doi:[10.1145/2380656.2380675](https://doi.org/10.1145/2380656.2380675).
- [17] David Doty, Mahsa Eftekhari, and Eric Severson. A stable majority population protocol using logarithmic time and states, 2020. URL <https://arxiv.org/abs/2012.15800>.
- [18] Moez Draief and Milan Vojnović. Convergence speed of binary interval consensus. *SIAM Journal on Control and Optimization*, 50(3):1087–1109, 2012.
- [19] Robert Elsässer and Tomasz Radzik. Recent results in population protocols for exact majority and leader election. *Bulletin of the EATCS*, 126, 2018. URL <http://bulletin.eatcs.org/index.php/beatcs/article/view/549/546>.
- [20] Ofer Feinerman and Amos Korman. Theoretical distributed computing meets biology: A review. In *Proc. 9th International Conference on Distributed Computing and Internet Technology (ICDIT)*, pages 1–18, 2013. doi:[10.1007/978-3-642-36071-8\\_1](https://doi.org/10.1007/978-3-642-36071-8_1).
- [21] Jasmin Fisher, David Harel, and Thomas A. Henzinger. Biology as reactivity. *Communications of the ACM*, 54(10):72–82, 2011. doi:[10.1145/2001269.2001289](https://doi.org/10.1145/2001269.2001289).
- [22] Timothy S. Gardner, Charles R. Cantor, and James J. Collins. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403(6767):339–342, January 2000. doi:[10.1038/35002131](https://doi.org/10.1038/35002131).
- [23] Melanie Ghoul and Sara Mitri. The ecology and evolution of microbial competition. *Trends in microbiology*, 24(10):833–845, 2016. doi:[10.1016/j.tim.2016.06.011](https://doi.org/10.1016/j.tim.2016.06.011).
- [24] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977. doi:[10.1021/j100540a008](https://doi.org/10.1021/j100540a008).
- [25] Elisa T Granato, Thomas A Meiller-Legrand, and Kevin R Foster. The evolution and ecology of bacterial warfare. *Current biology*, 29(11):R521–R537, 2019. doi:[10.1016/j.cub.2019.04.024](https://doi.org/10.1016/j.cub.2019.04.024).
- [26] Garrett Hardin. The competitive exclusion principle. *Science*, 131(3409):1292–1297, 1960. doi:[10.1126/science.131.3409.1292](https://doi.org/10.1126/science.131.3409.1292).
- [27] Samuel Karlin and Howard M. Taylor. *A First Course in Stochastic Processes*. Academic Press, New York, 2 edition, 1975.
- [28] Javier Macía, Francesc Posas, and Ricard V. Solé. Distributed computation: the new wave of synthetic biology devices. *Trends in Biotechnology*, 30(6):342 – 349, 2012. doi:[10.1016/j.tibtech.2012.03.006](https://doi.org/10.1016/j.tibtech.2012.03.006).
- [29] George B Mertzios, Sotiris E Nikolettseas, Christoforos L Raptopoulos, and Paul G Spirakis. Determining majority in networks with local interactions and very small local memory. *Distributed Computing*, 30(1):1–16, 2017. doi:[10.1007/s00446-016-0277-8](https://doi.org/10.1007/s00446-016-0277-8).
- [30] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press, Cambridge, 2nd edition, 2017.

- [31] Jacques Monod. The growth of bacterial cultures. *Annual Review of Microbiology*, 3(1): 371–394, 1949.
- [32] Saket Navlakha and Ziv Bar-Joseph. Distributed information processing in biological and computational systems. *Communications of the ACM*, 58(1):94–102, 2014. doi:[10.1145/2678280](https://doi.org/10.1145/2678280).
- [33] Béla Novák and John J. Tyson. Design principles of biochemical oscillators. *Nature Reviews Molecular Cell Biology*, 9(12):981–991, October 2008. doi:[10.1038/nrm2530](https://doi.org/10.1038/nrm2530).
- [34] Priscilla EM Purnick and Ron Weiss. The second wave of synthetic biology: from modules to systems. *Nature reviews Molecular cell biology*, 10(6):410–422, 2009. doi:[10.1038/nrm2698](https://doi.org/10.1038/nrm2698).
- [35] Sergi Regot, Javier Macia, Núria Conde, Kentaro Furukawa, Jimmy Kjellén, Tom Peeters, Stefan Hohmann, Eulàlia de Nadal, Francesc Posas, and Ricard Solé. Distributed biological computation with multicellular engineered networks. *Nature*, 469(7329):207–211, 2010. doi:[10.1038/nature09679](https://doi.org/10.1038/nature09679).
- [36] Severin J Schink, Elena Biselli, Constantin Ammar, and Ulrich Gerland. Death rate of e. coli during starvation is set by maintenance cost and biomass recycling. *Cell Systems*, 9(1): 64–73, 2019.
- [37] Jesse Stricker, Scott Cookson, Matthew R. Bennett, William H. Mather, Lev S. Tsimring, and Jeff Hasty. A fast, robust and tunable synthetic gene oscillator. *Nature*, 456(7221): 516–519, October 2008. doi:[10.1038/nature07389](https://doi.org/10.1038/nature07389).
- [38] Zhenmao Wan, Joseph Varshavsky, Sushma Teegala, Jamille McLawrence, and Noel L Goddard. Measuring the rate of conjugal plasmid transfer in a bacterial population using quantitative PCR. *Biophysical Journal*, 101(1):237–244, 2011.