Interspecies evolutionary dynamics mediated by public goods in bacterial quorum sensing

Eduardo J. Aguilar [1], Valmir C. Barbosa [2,*], Raul Donangelo [3,4] and Sergio R. Souza [4,5]

[1]*Instituto de Ciência e Tecnologia, Universidade Federal de Alfenas, Rodovia José Aurélio Vilela, 11999, 37715-400 Poços de Caldas, Minais Gerais, Brazil*

[2]*Programa de Engenharia de Sistemas e Computação, COPPE, Universidade Federal do Rio de Janeiro, Centro de Tecnologia, Sala H-319, 21941-914 Rio de Janeiro, Rio de Janeiro, Brazil*

[3]*Instituto de Física, Facultad de Ingeniería, Universidad de la República, Julio Herrera y Reissig 565, 11300 Montevideo, Uruguay*

[4]*Instituto de Física, Universidade Federal do Rio de Janeiro, Centro de Tecnologia, Bloco A, 21941-909 Rio de Janeiro, Rio de Janeiro, Brazil*

[5]*Departamento de Física, ICEx, Universidade Federal de Minas Gerais, Avenida Antônio Carlos, 6627, 31270-901 Belo Horizonte, Minais Gerais, Brazil*

Bacterial quorum sensing is the communication that takes place between bacteria as they secrete certain molecules into the intercellular medium that later get absorbed by the secreting cells themselves and by others. Depending on cell density, this uptake has the potential to alter gene expression and thereby affect global properties of the community. We consider the case of multiple bacterial species coexisting, referring to each one of them as a genotype and adopting the usual denomination of the molecules they collectively secrete as public goods. A crucial problem in this setting is characterizing the coevolution of genotypes as some of them secrete public goods (and pay the associated metabolic costs) while others do not but may nevertheless benefit from the available public goods. We introduce a network model to describe genotype interaction and evolution when genotype fitness depends on the production and uptake of public goods. The model comprises a random graph to summarize the possible evolutionary pathways the genotypes may take as they interact genetically with one another, and a system of coupled differential equations to characterize the behavior of genotype abundance in time. We study some simple variations of the model analytically and more complex variations computationally. Our results point to a simple trade-off affecting the long-term survival of those genotypes that do produce public goods. This trade-off involves, on the producer side, the impact of producing and that of absorbing the public good. On the nonproducer side, it involves the impact of absorbing the public good as well, now compounded by the molecular compatibility between the producer and the nonproducer. Depending on how these factors turn out, producers may or may not survive.

I. INTRODUCTION

Bacterial quorum sensing (QS) is the general denomination for a variety of mechanisms whereby bacteria communicate with one another by the secretion and uptake of molecules that get diffused in the intercellular medium. The absorption of such molecules by the cells is important for the regulation of some genes and thus affects survival and successful proliferation, as well as many functional traits, depending on cell density. These include a bacterial species' propensity to establish symbioses or to aggregate into biofilms, and also its virulence and motility. Although bacterial QS has been around as a research topic for about five decades, only much more recently has sufficient evidence accumulated. Interest in it is now widespread and a number of useful reviews are available [1–6]. Whereas initial interest was sparked by the curious phenomenon of bioluminescence arising in certain marine bacteria [7,8], the current focus is on the role of QS in the immunological effects of the mammalian gut microbiome [9,10] and in combating bacterial virulence [11,12].

When only one single species is involved in bacterial QS, a cell's uptake, though comprising mainly the molecules that the cell itself secretes, may also include those of the other cells nearly indistinguishably. Such molecules are called autoinducers, in allusion to the fact that they act as gene regulators for the species. In settings allowing for multiple species, it is in principle conceivable that the autoinducers secreted by members of a species be absorbable also by those of another. Depending on the species involved, the autoinducers they exchange can be expected to affect cells of different species differently, perhaps acting as proper autoinducers in some cases and simply as sources of nutrients or even toxicity in others. A cell's secretions, moreover, may include molecules other than the ones directly involved in autoinduction. Commonly, all secreted molecules are collectively referred to as public goods (cf., e.g., [6]). In evolutionary terms, and depending on the metabolic costs involved, the production and uptake of public goods can affect a cell's fitness significantly.

Despite the much earlier developments in the field, bacterial QS seems to have been the object of mathematical modeling for no longer than about two decades. Many of the modeling efforts have counted on the close participation of researchers involved with experimentation, resulting in detailed analytical accounts of specific QS processes [4], including the

---

*valmir@cos.ufrj.br

combinatorial response to intercellular signaling [13]. Others have taken a somewhat more distanced stance and addressed, e.g., the importance of stochastic elements in bacterial QS [14–20], the occurrence of transitions [21], and biofilm [22] and pattern [23,24] formation. Moreover, some have taken up the basic tenets of QS and sought to characterize higher-level phenomena occurring not only in bacteria but also in other systems in which QS can be said to play a central role. Such phenomena include synchronization in networks [25], oscillations in coupled systems [26], the appearance of dense aggregates of active particles [27], and phase separation in colloidal mixtures [20]. To the best of our knowledge, each of these models has considered one single bacterial species, taking diversity into account only insofar as it stems from a few of that species' strains.

Here we introduce a model of bacterial QS that focuses on how multiple species coexist and evolve in the presence of the public goods they produce. In order to avoid the many complications involved in bacterial taxonomy [28], we henceforth refer to genotypes as the basic units of diversity. Given two genotypes, one of them might be considered merely a strain of the other or they might both refer to totally distinct species. In our model, all possible genotypes of a given length are considered concomitantly. The complete QS system is represented by a random graph whose nodes stand for genotypes and whose edges exist with probabilities that aim to reflect the mutations that occur as cells undergo binary fission (the most common mechanism of bacterial proliferation, though several others exist [29,30]) and horizontal gene transfer (HGT, the transfer of genetic material between individuals [31]). These probabilities are parametrized so that sampling from the random graph may result in a graph with more or fewer edges and therefore more or less far-reaching effects of localized random variation.

In addition to the structural component given by this random graph, our model includes a system of coupled differential equations describing how the genotypes' abundances vary with time. These equations depend on the same parametrization as the random graph. Crucially, they also depend on which genotypes are producers of public goods, as well as on whether and how absorption affects each individual genotype. All these aspects of public-good production and uptake contribute to each genotype's fitness in the QS system and can be quantified in such a way that a great variety of global configurations can be studied.

This model continues earlier work [32] in which we added network structure to the so-called quasispecies models of the dynamics of prebiotic molecules and RNA viruses [33–38], and also work in which we modeled the network dynamics of autoimmunity and immunodeficiency [39,40] and of the evolution of eukaryotic cellular division [41]. We now proceed by giving the model's details in Sec. II and analyzing a special case thereof in Sec. III. Results are given in Sec. IV and discussed in Sec. V. We conclude in Sec. VI and give directions for further analysis in the Appendix.

II. MODEL

We represent each bacterial gene by a sequence of B binary digits (0's or 1's), each standing for a nucleotide. A bacterial genotype is represented by a sequence of L binary digits, provided $L \geqslant B$ and that L is a multiple of B. A genotype is then viewed as a sequence of L/B genes. Such representations are of course oversimplistic in more than one sense, but even so they provide the necessary means to model the two main sources of variation in bacterial genotypes, viz., mutations during binary fission (at the level of the nucleotide) and HGT (at the level of the gene). Henceforth, we use N to denote the set of all 2^L genotypes. Genotypes are numbered 0 through $2^L - 1$.

Our focus is on the effects of public-good production and consumption by the genotypes, which we model using parameters μ_j and σ_{ji}, as follows. For each genotype j, μ_j takes its value from $\{0, 1\}$ and indicates whether j is a producer of public goods ($\mu_j = 1$) or not ($\mu_j = 0$). For each genotype pair i, j, σ_{ji} takes its value from the interval $[-1, 1]$ and is used to indicate the degree to which consuming public goods produced by genotype j can impact the rate at which genotype i is capable of multiplying. This degree ranges from a full contribution toward a slower pace ($\sigma_{ji} = -1$) to a full contribution toward a faster pace ($\sigma_{ji} = 1$), with neutrality in between ($\sigma_{ji} = 0$). For $j = i$, the biological process in question includes that of bacterial autoinduction.

Whenever $\mu_j \sigma_{ji} \neq 0$, we model the consumption by i of the public goods produced by j as occurring with probability $u_{ji} = u^{J_{ij}}$, where $u < 1$ is a base probability of the model and J_{ij} is the number of homologous genes at which genotypes i and j differ. Expressing u_{ji} in this way takes into account the need for molecular compatibility between organisms of genotype j and those of genotype i, and the fact that such compatibility occurs at the level of the genes of j and i. Thus, lower values of J_{ij} indicate higher compatibility. $J_{ij} = 0$, in particular, holds when $j = i$, so autoinduction occurs with probability 1 whenever $\mu_j = 1$.

As will become apparent when we give the model's dynamical equations in Sec. II C, we make another strong simplifying assumption, now regarding the production and consumption of public goods by the genotypes. The assumption is that setting $\mu_j = 1$ for some genotype j makes the public goods it produces available to all genotypes in sufficient concentration at all times. That is to say, we make no provision to model the dynamics of public-good production, diffusion, or consumption. Our equations are concerned solely with the dynamics of genotype proliferation, given a static backdrop of overall public-good availability. We comment further on these and related issues in Sec. VI.

A. Fitness of a genotype

The quantification of an individual's fitness in studies of evolutionary dynamics can be challenging. Ideally, it should rely on the identification of traits exerting a quantifiable influence on gene spread through the population. This is rarely possible, so it is common for the difficulty to be sidestepped with the adoption of a proxy, as some form of distance to an agreed-upon wild type or the experimental counting of individuals as proliferation unfolds.

However, the case at hand is exceptional in that, clearly, a genotype can be said to be as fit as its total consumption of public goods is beneficial to it. That is, if genotype j is a

producer of public goods (i.e., $\mu_j = 1$), then its contribution to the fitness of genotype i should be an increasing function of $\sigma_{ji}u^{J_{ij}}$. It should also depend on how abundant genotype j is in the entire population, which immediately makes the fitness of i time dependent. Letting X_j denote the absolute abundance of genotype j at a certain instant and

$$x_j = \frac{X_j}{\sum_{k \in N} X_k} \qquad (1)$$

its relative abundance (so $\sum_{j \in N} x_j = 1$), we express the instantaneous fitness of i by an increasing function of the convex combination of $\mu_j \sigma_{ji} u^{J_{ij}}$ over N that uses x_j as the weight of genotype j. We denote this combination by d_i, so

$$d_i = \sum_{j \in N} \mu_j \sigma_{ji} u^{J_{ij}} x_j, \qquad (2)$$

and denote the fitness of genotype i by f_i, defining it as $f_i = 2^{d_i}$. Clearly, $-1 \leqslant d_i \leqslant 1$, from which $0.5 \leqslant f_i \leqslant 2$. For $d_i = 0$ (the combined public goods instantaneously available to i are neither beneficial nor detrimental to its ability to multiply), we have $f_i = 1$. This is convenient in view of how we incorporate fitness values into the model's dynamical equations (cf. Sec. II C).

B. Network structure

A further important element in our model is a substrate on which genotypes can transform into one another via the mechanisms afforded by mutations and HGT. The substrate we use is a random graph of node set N (the set of all 2^L distinct genotypes on L nucleotides) and undirected edges. The existence of an edge between any two genotypes i and j depends on two base probabilities, p and r, and has probability

$$\pi_{ij} = p^{H_{ij}} + r^{J_{ij}} - p^{H_{ij}}r^{J_{ij}}. \qquad (3)$$

In this expression, H_{ij} is the Hamming distance between genotypes i and j, i.e., the number of nucleotides at which they differ. J_{ij}, already introduced at the beginning of Sec. II, is the number of homologous B-nucleotide genes at which i and j differ.

Given the value of p, increasing H_{ij} makes it less likely for the edge joining i and j to exist, reflecting the fact that the transformation of i into j (or conversely) by co-occurring mutations at H_{ij} nucleotides becomes less likely as well. Similarly, given the value of r, increasing J_{ij} makes joining i and j by an edge less likely, now reflecting the also less likely transformation of i into j (or conversely) by the co-occurring horizontal transfer of J_{ij} genes. The expression in Eq. (3) comes from assuming that events of these two types in sample space (mutations and HGT) are independent though not mutually exclusive.

Regardless of the value of p or r, for $i = j$ we always have $\pi_{ii} = 1$ (assuming $0^0 = 1$). As a result, every undirected graph G sampled from the random graph given by Eq. (3) necessarily has self-loops at all genotypes, no matter how sparsely interconnected the genotypes may be as a function of p and r.

C. Network dynamics

For a fixed graph G sampled from the random graph introduced in Sec. II B, we describe the corresponding network dynamics as a set of 2^L coupled differential equations, one for each possible genotype. These equations bring together all the elements introduced in Sec. II A, which led to the definition of the fitness $f_i = 2^{d_i}$ of genotype i for d_i as in Eq. (2), and those introduced in Sec. II B, where the random graph giving rise to G was specified via the edge probability π_{ij} of Eq. (3) that an edge exists between genotypes i and j.

The probabilities $p^{H_{ij}}$ and $r^{J_{ij}}$ appearing in the expression for π_{ij} are, as a matter of principle, expected to be relevant also to our dynamical equations. If so, they must appear in the equations in normalized form, as follows. Letting N_i denote the set of genotypes (including i itself) to which genotype i is joined by an edge in graph G, probability $p^{H_{ij}}$ becomes

$$q_{j \to i} = \frac{p^{H_{ij}}}{\sum_{k \in N_j} p^{H_{kj}}}, \qquad (4)$$

so $\sum_{i \in N_j} q_{j \to i} = 1$, and similarly probability $r^{J_{ij}}$ becomes

$$s_{i \to j} = \frac{r^{J_{ij}}}{\sum_{k \in N_i} r^{J_{ik}}}, \qquad (5)$$

so $\sum_{j \in N_i} s_{i \to j} = 1$. Probability $q_{j \to i}$ is the probability that genotype j gives rise to genotype i during binary fission by undergoing mutation at H_{ij} of its nucleotides. Probability $s_{i \to j}$, in turn, is the probability that genotype i gives rise to genotype j when J_{ij} genes of genotype i undergo HGT.

The differential equation describing the evolution in time of the absolute abundance of genotype i, X_i, is

$$\dot{X}_i = \sum_{j \in N_i} q_{j \to i} f_j X_j + \lambda \sum_{j \in N_i} s_{i \to j} f_i X_i. \qquad (6)$$

The first summation on the right-hand side of this equation accounts for the total contribution to $\dot{X}_i$ from mutations affecting genotypes $j \in N_i$. The individual contribution from each such j depends on probability $q_{j \to i}$, on the fitness of genotype j, and on its abundance. The second summation is similar, now accounting for the total contribution to $\dot{X}_i$ when genes get transferred from genotype i to genotypes $j \in N_i$ via HGT. In Eq. (6), λ is a parameter that can be used to regulate the relative rate at which the two types of contribution to $\dot{X}_i$ occur. Because $\sum_{j \in N_i} s_{i \to j} = 1$, Eq. (6) can be immediately rewritten as

$$\dot{X}_i = \sum_{j \in N_i} q_{j \to i} f_j X_j + \lambda f_i X_i, \qquad (7)$$

where it becomes clear that the second summation on the right-hand side of Eq. (6) depends only on the fitness and abundance of genotype i. Thus, the probabilities $r^{J_{ij}}$ giving rise to $s_{i \to j}$ through Eq. (5) are relevant only insofar as they affect the sampling of graph G from the random graph.

In spite of the presumed function of parameter λ in Eqs. (6) and (7), the simplified equation for $\dot{X}_i$ given in Eq. (7) highlights the fact that the contribution of $f_i X_i$ to $\dot{X}_i$ is actually multiplied by $q_{i \to i} + \lambda$. Without λ weighing in like this, such factor would already be exponentially higher than that of any other $f_j X_j$ for $p < 1$; that is, for $j \neq i$ we would have

$q_{i \to i}/q_{j \to i} = (1/p)^{H_{ij}}$. It seems, therefore, that λ has no clear role to play, so henceforth we use $\lambda = 0$. The final form of the equation for $\dot{X}_i$ is then

$$\dot{X}_i = \sum_{j \in N_i} q_{j \to i} f_j X_j. \tag{8}$$

In order to avoid the unbounded growth of X_i that Eq. (8) may entail, we rewrite it for the genotypes' relative abundances instead. Such abundances are as given in Eq. (1); i.e., the relative abundance of genotype i is $x_i = X_i / \sum_{k \in N} X_k$, so $\sum_{i \in N} x_i = 1$ holds at all times. We obtain

$$\dot{x}_i = \frac{\dot{X}_i}{\sum_{k \in N} X_k} - x_i \frac{\sum_{k \in N} \dot{X}_k}{\sum_{k \in N} X_k}$$
$$= \sum_{j \in N_i} q_{j \to i} f_j x_j - x_i \sum_{k \in N} \sum_{j \in N_k} q_{j \to k} f_j x_j. \tag{9}$$

Letting

$$\phi = \sum_{k \in N} \sum_{j \in N_k} q_{j \to k} f_j x_j \tag{10}$$

yields

$$\dot{x}_i = \sum_{j \in N_i} q_{j \to i} f_j x_j - \phi x_i. \tag{11}$$

Note that, even though this equation embodies the new non-linearities implied by the term involving ϕ, the truly striking ones are those already present in Eq. (8), that is, those through which genotype fitnesses depend on relative abundances exponentially.

III. A SPECIAL CASE WITH ONE SINGLE PRODUCER

By Eq. (3), letting $p = r = 0$ implies $\pi_{ij} = [i = j]$.[1] That is, graph G has no edges other than the 2^L self-loops, so by Eq. (4) we have $q_{i \to i} = 1$ for every genotype i. It might seem that such a trivial topology, implying as it does that genotypes never undergo any form of random variation, would be unable to give rise to interesting dynamics. This is not necessarily so, however, because genotypes still influence one another through the close coupling that their abundance-dependent fitnesses provide.

Letting

$$\bar{f} = \sum_{k \in N} f_k x_k, \tag{12}$$

it follows from Eq. (10) that $\phi = \bar{f}$, thus leading, by Eq. (11), to

$$\dot{x}_i = (f_i - \bar{f}) x_i. \tag{13}$$

Here we look at the case in which $\mu_i = [i = 0]$ (genotype 0 is the only producer of public goods) and $\sigma_{ji} = [j = 0][i \neq 0]\sigma$ (all genotypes other than 0 are equally impacted by the public goods that genotype 0 produces) for some $\sigma \in [-1, 1]$.

---

[1] For a logical proposition P, the Iverson bracket $[P]$ equals 1 if P is true, 0 if P is false. This notation generalizes the Kronecker delta, since $[i = j] = \delta_{ij}$.

All σ_{ji}'s with $j \neq 0$ are irrelevant. (A similar special case, but with multiple producers of public goods, is analyzed in the Appendix.)

Assuming $x_i(0) = 2^{-L}$ for every genotype i, Eq. (13) implies that, in the long run, only the genotypes whose initial fitnesses are greatest survive. If this happens for more than one genotype, then all of them survive with the same abundance. Of all genotypes $i \neq 0$, identifying the fittest depends on the value of σ. For $\sigma > 0$, the fittest are the $(L/B)(2^B - 1)$ genotypes for which $J_{0i} = 1$. For $\sigma = 0$, all $2^L - 1$ genotypes have the same fitness. For $\sigma < 0$, the fittest are the $(2^B - 1)^{L/B}$ genotypes for which $J_{0i} = L/B$. It follows that survival is determined by how σ_{00} relates to σu (if $\sigma_{00} > 0$ and $\sigma > 0$), or to $\sigma u^{L/B}$ (if $\sigma_{00} < 0$ and $\sigma < 0$), or directly to σ (in all other cases). In general, we denote the greatest fitness among genotypes $i \neq 0$ by f_+. There are nine cases to be considered, discussed next. In this discussion, any conclusion that a group of genotypes $i \neq 0$ survives is to be extended to all genotypes $i \neq 0$ if $u = 1$, since in this case they all have the same fitness.

(C1) $\sigma_{00} > 0$ and $\sigma > 0$: $f_0 > 1$ and $f_+ > 1$, with three subcases.

(C1a) $\sigma_{00} > \sigma u$: $f_0 > f_+$; genotype 0 survives.

(C1b) $\sigma_{00} = \sigma u$: $f_0 = f_+$; genotype 0 survives, and so do all genotypes i such that $J_{0i} = 1$.

(C1c) $\sigma_{00} < \sigma u$: $f_0 < f_+$; genotypes i such that $J_{0i} = 1$ survive.

(C2) $\sigma_{00} > 0$ and $\sigma = 0$: $f_0 > 1$ and $f_+ = 1$; genotype 0 survives.

(C3) $\sigma_{00} > 0$ and $\sigma < 0$: $f_0 > 1$ and $f_+ < 1$; genotype 0 survives.

(C4) $\sigma_{00} = 0$ and $\sigma > 0$: $f_0 = 1$ and $f_+ > 1$; genotypes i such that $J_{0i} = 1$ survive.

(C5) $\sigma_{00} = 0$ and $\sigma = 0$: $f_0 = 1$ and $f_+ = 1$; all genotypes survive.

(C6) $\sigma_{00} = 0$ and $\sigma < 0$: $f_0 = 1$ and $f_+ < 1$; genotype 0 survives.

(C7) $\sigma_{00} < 0$ and $\sigma > 0$: $f_0 < 1$ and $f_+ > 1$; genotypes i such that $J_{0i} = 1$ survive.

(C8) $\sigma_{00} < 0$ and $\sigma = 0$: $f_0 < 1$ and $f_+ = 1$; genotypes $i \neq 0$ survive.

(C9) $\sigma_{00} < 0$ and $\sigma < 0$: $f_0 < 1$ and $f_+ < 1$, with three subcases.

(C9a) $\sigma_{00} > \sigma u^{L/B}$: $f_0 > f_+$; genotype 0 survives.

(C9b) $\sigma_{00} = \sigma u^{L/B}$: $f_0 = f_+$; genotype 0 survives, and so do all genotypes i such that $J_{0i} = L/B$.

(C9c) $\sigma_{00} < \sigma u^{L/B}$: $f_0 < f_+$; genotypes i such that $J_{0i} = L/B$ survive.

A careful examination of these cases reveals that a number of outcomes are possible in the long run, which seems remarkable as we consider that the only edges in graph G are the self-loops at all genotypes. Interestingly, these results provide useful insight even when lifting the $p = r = 0$ simplification, as we show in Sec. IV.

IV. RESULTS

All our computational results come from time stepping Eq. (11) from $x_i(0) = 2^{-L}$ for all genotypes $i \in N$. This is done on a fixed graph G, obtained by Monte Carlo

sampling as explained in Sec. II B. The expected steady-state value of each x_i is obtained by averaging over several such graphs.

A potentially problematic aspect of time stepping Eq. (11), depending on the case at hand, is the often very small absolute value of each of the terms contributing to d_i in Eq. (2). An issue to be considered is that several genotypes may have nearly neutral (close to 1) fitnesses for a long time even if eventually they are to diverge from one another substantially. In this case, convergence to the steady-state relative abundances can be slowed down significantly. Another issue is the effect of the unavoidable round-off errors that accompany operations on numbers very close to zero. Because of such errors, the value obtained for d_i may depend on the order used for the terms to enter the sum, which can break important symmetries whenever the equations dictate that i and j exist for which we should have $d_i = d_j$. To avoid this, at every step the terms that make up d_i are added in increasing order of absolute value. Thus, except for cases involving one single producer, a sorting operation is required at each step, thus making convergence to the steady state very slow indeed; see Sec. VI for more on this.

We sidestep these issues by reducing the size and variety of the scenarios we consider while at the same time retaining the possibility of interesting and diverse behavior. We thus consider only a small number of genotypes: we use $L = 9$ and $B = 3$, hence 512 genotypes, each with three 3-nucleotide genes. Moreover, all scenarios we consider generalize the special case of Sec. III by allowing $p, r > 0$ and therefore the appearance of nontrivial topologies (that is, those on which manifestations of the mutational and HGT-related aspects of the model can occur). All other simplifications of that section continue to be assumed, so $p = r$, genotype 0 is the only producer of public goods, and $\sigma_{0i} = \sigma$ for every genotype $i \neq 0$. Given these simplifications, the expected behavior of all

$$n_J = \binom{L/B}{J}(2^B - 1)^J \qquad (14)$$

genotypes that differ from genotype 0 at J genes is the same. A handy simplification when analyzing results is then to consider all such genotypes as a single group; that is, consider them through their mean relative abundances, denoted by x_J and given by

$$x_J = n_J^{-1} \sum_{i \in N \mid J_{0i} = J} x_i. \qquad (15)$$

Of course, $\sum_{J=0}^{L/B} n_J x_J = 1$.

Given the random graph used to represent our network of genotypes, the expected degree of a randomly chosen genotype i (the expected number of edges incident to it, including its self-loop) can be obtained by summing up π_{ij} (the probability that an edge exists between genotypes i and j) over $j \in N$. Based on Eq. (3), we write this summation in terms of the number H of nucleotides at which i and j may differ and likewise the number J of genes at which they may differ.
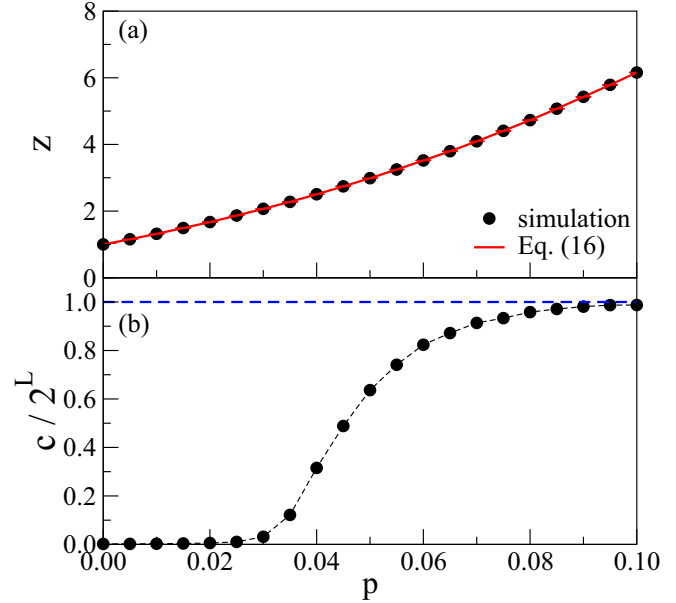

FIG. 1. (a) Expected degree of a randomly chosen genotype. (b) Expected fraction of 2^L corresponding to the connected component to which that genotype belongs. Both panels refer to $L = 9$ and $B = 3$.

Letting z denote the desired expected degree, we obtain

$$z = \sum_{H=0}^{L} \binom{L}{H} p^H + \sum_{J=0}^{L/B} n_J r^J - \sum_{J=0}^{L/B} \binom{L/B}{J} r^J \sum_{n_1=1}^{B}$$

$$\cdots \sum_{n_J=1}^{B} \binom{B}{n_1} \cdots \binom{B}{n_J} p^{n_1 + \cdots + n_J}$$

$$= (1 + p)^L + [1 + (2^B - 1)r]^{L/B}$$

$$- [1 + [(1 + p)^B - 1]r]^{L/B}. \qquad (16)$$

Clearly, increasing p (or r, should it be allowed to change independently of p) leads z to grow as well, and with it the expected number of genotypes that can be reached from genotype i through some undirected path in the graph.

Such genotypes constitute the so-called connected component to which i belongs. In the context at hand, this is the set of genotypes expected to relate to i via mutation or HGT. For $i = 0$, this means that probability p can be used to control the expected extent to which the behavior of the single producer is evolutionarily related to those of other genotypes. That is, while all 2^L genotypes partake of the public goods produced by genotype 0, the set of those that interact with the producer, directly or indirectly through mutation or HGT, can be shrunk or enlarged by controlling the value of p. This is illustrated in Fig. 1, where the expected number of genotypes in the connected component is denoted by c. Clearly, letting $p \leqslant 0.1$ suffices for all interesting cases of the connected component in question to be encompassed. This includes the case in which it only contains genotype 0, the case in which it contains a sizable fraction of all 2^L genotypes, and the transition in between as p grows from $p = 0$. All values of p are thus constrained henceforth. (Incidentally, in this case it can be
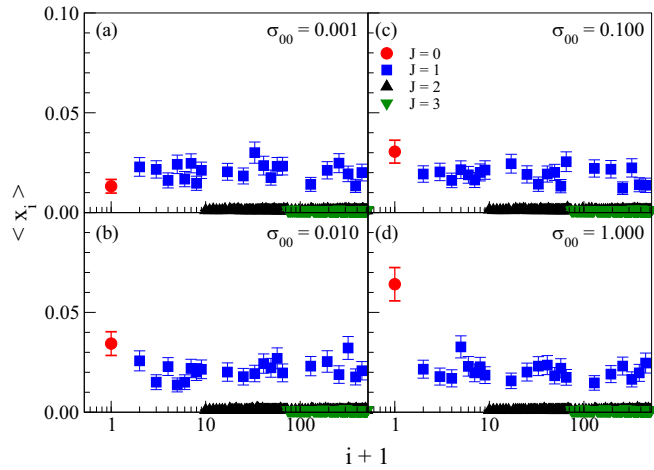
FIG. 2. Average steady-state individual relative abundances (x_i) for $L = 9$, $B = 3$, $p = 0.075$ ($z \approx 4.41$, $c \approx 478$), $u = 0.01$, and $\sigma = 1$. Values of σ_{00} are (a) 0.001, (b) 0.01, (c) 0.1, and (d) 1. See Fig. 3 for the corresponding mean relative abundances (x_J).

verified that relaxing the assumption $p = r$ only mildly, e.g., $0.95p \leqslant r \leqslant 1.05p$, causes z to vary mildly as well, about $\pm 5\%$ for $p = 0.1$, less for smaller p. This may help regard the $p = r$ constraint as not so stringent after all.)

We present our results in two groups of figures, the first relating to case C1, of Sec. III, with both $\sigma_{00} > 0$ and $\sigma > 0$, and the second relating to case C9, with both $\sigma_{00} < 0$ and $\sigma < 0$. Note that, in regard to cases C1–C9 of that section, C1 and C9 are the only nontrivial ones in that outcomes depend on how parameters other than σ_{00} and σ intervene. With only two exceptions, to be discussed shortly, all figures show the steady-state value of x_J for $J \in \{0, 1, 2, 3\}$. Note, in these figures, that because of their definition in Eq. (15) these four values of x_J do not in general add up to 1.

All figures in the first group, Figs. 2–5, are for $u = 0.01$ and $\sigma = 1$. Figures 2 and 3 are for $p = 0.075$, leading to $z \approx 4.41$ and $c \approx 478$. Figure 2 is one of the exceptions alluded
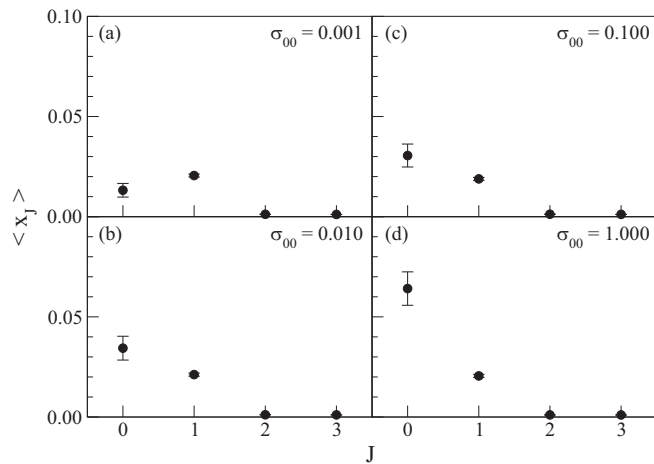
to above, since it shows the steady-state x_i's individually. A similar breakdown is of course possible in all other cases of this group and the second, but is omitted in most of them. Figures 4 and 5 complete the first group, respectively with $p = 0.045$ ($z \approx 2.74$ and $c \approx 250$) and $p = 0.0384$ ($z \approx 2.43$ and $c \approx 136$).

All figures in the second group, Figs. 6–9, are for $u = 0.1$ and $\sigma = -1$. They refer to the same values of p as the figures in the first group, in the same order. Figure 6 is the second exception alluded to earlier, since it pairs with Fig. 7 in that the former refers to individual relative abundances and the latter to mean relative abundances for the same value of p. Note that the value of u for all figures in this second group is greater than that of Figs. 2–5 by one order of magnitude. This is meant to avoid aggravating the first convergence issue mentioned at the beginning of this section even further. As discussed in Sec. V, the critical value of J_{0i} for Figs. 2–5 is $J_{0i} = 1$, whereas for Figs. 6–9 it is $J_{0i} = L/B = 3$, so values of σ_{00} closer to zero are needed.

FIG. 4. Average steady-state mean relative abundances (x_J) for $L = 9$, $B = 3$, $p = 0.045$ ($z \approx 2.74$, $c \approx 250$), $u = 0.01$, and $\sigma = 1$. Values of σ_{00} are (a) 0.001, (b) 0.01, (c) 0.1, and (d) 1.

FIG. 3. Average steady-state mean relative abundances (x_J) for $L = 9$, $B = 3$, $p = 0.075$ ($z \approx 4.41$, $c \approx 478$), $u = 0.01$, and $\sigma = 1$. Values of σ_{00} are (a) 0.001, (b) 0.01, (c) 0.1, and (d) 1. See Fig. 2 for the corresponding individual relative abundances (x_i).
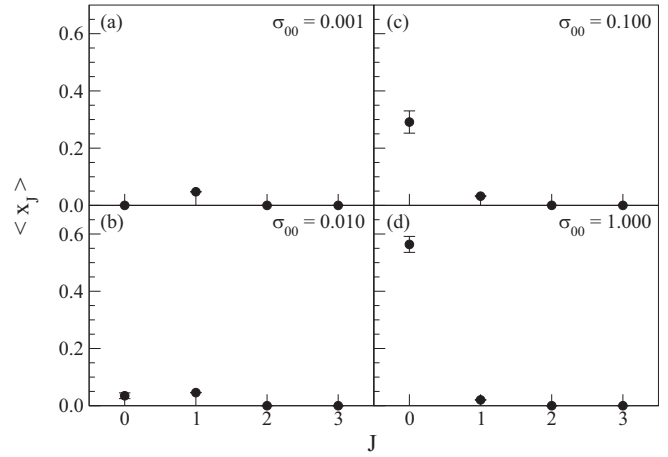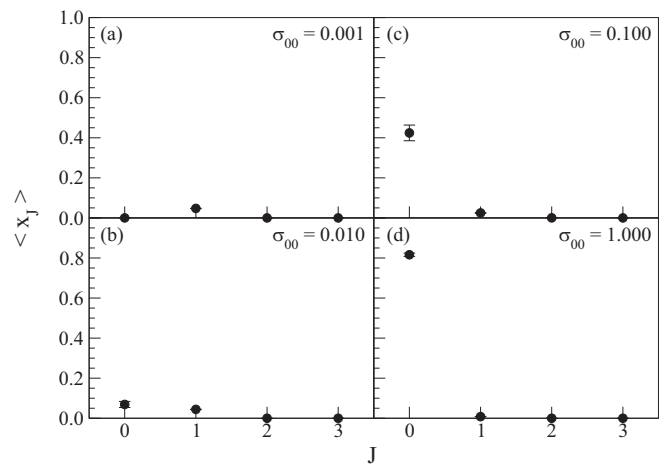

FIG. 5. Average steady-state mean relative abundances (x_J) for $L = 9$, $B = 3$, $p = 0.0384$ ($z \approx 2.43$, $c \approx 136$), $u = 0.01$, and $\sigma = 1$. Values of σ_{00} are (a) 0.001, (b) 0.01, (c) 0.1, and (d) 1.
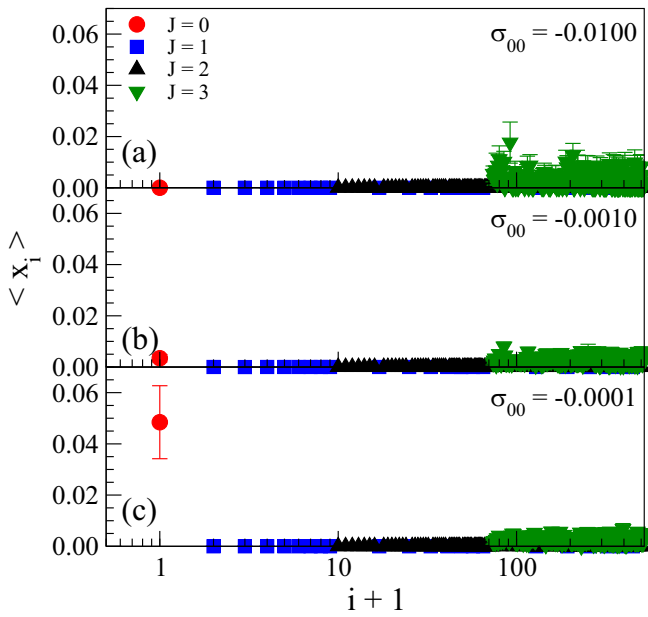
FIG. 6. Average steady-state individual relative abundances (x_i) for $L = 9$, $B = 3$, $p = 0.075$ ($z \approx 4.41$, $c \approx 478$), $u = 0.1$, and $\sigma = -1$. Values of σ_{00} are (a) -0.01, (b) -0.001, and (c) -0.0001. See Fig. 7 for the corresponding mean relative abundances (x_J).

V. DISCUSSION

The public goods produced by bacteria in QS can be greatly diverse, including for example enzymes [6] and vitamins [42], and are therefore essential for a number of cellular functions related to growth and metabolism. However, a cell's
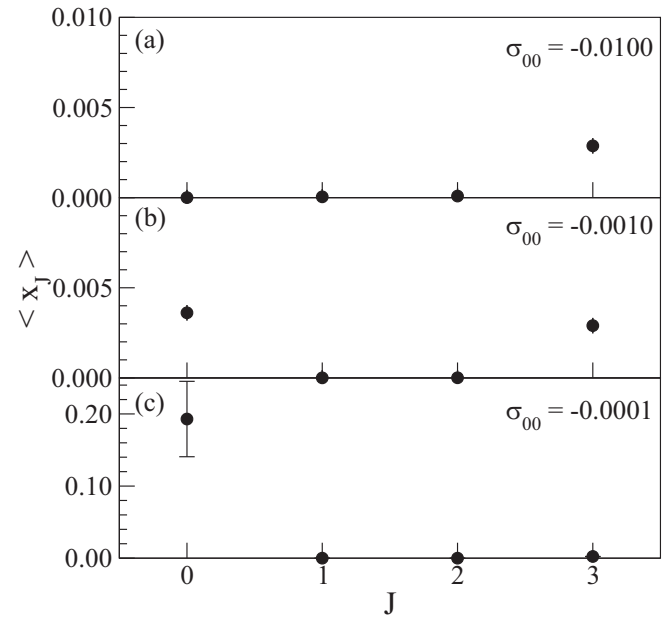

FIG. 8. Average steady-state mean relative abundances (x_J) for $L = 9$, $B = 3$, $p = 0.045$ ($z \approx 2.74$, $c \approx 250$), $u = 0.1$, and $\sigma = -1$. Values of σ_{00} are (a) -0.01, (b) -0.001, and (c) -0.0001. Note the different ordinate scales.

production of public goods has costs associated with it that have the potential to completely offset the benefits accrued to the cell itself by autoinduction. A central factor affecting the tipping of this balance one way or the other is the presence of certain mutants, commonly referred to as "cheaters," that do not join in the production of public goods but may never-
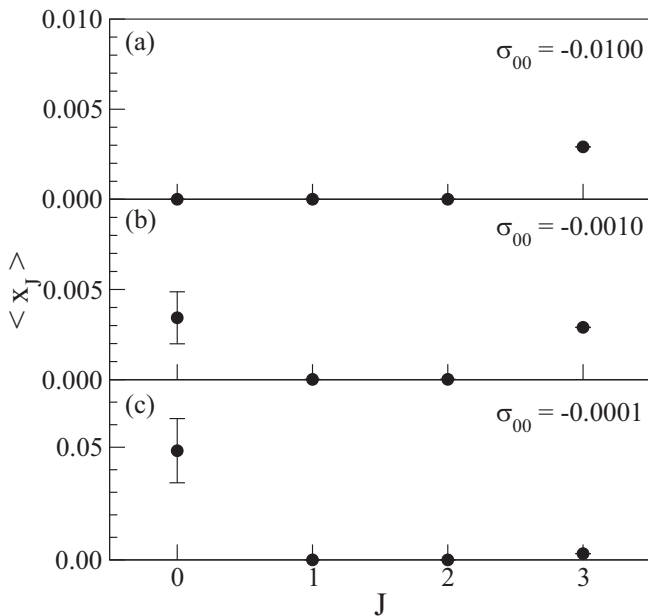

FIG. 7. Average steady-state mean relative abundances (x_J) for $L = 9$, $B = 3$, $p = 0.075$ ($z \approx 4.41$, $c \approx 478$), $u = 0.1$, and $\sigma = -1$. Values of σ_{00} are (a) -0.01, (b) -0.001, and (c) -0.0001. Note the different ordinate scales. See Fig. 6 for the corresponding individual relative abundances (x_i).
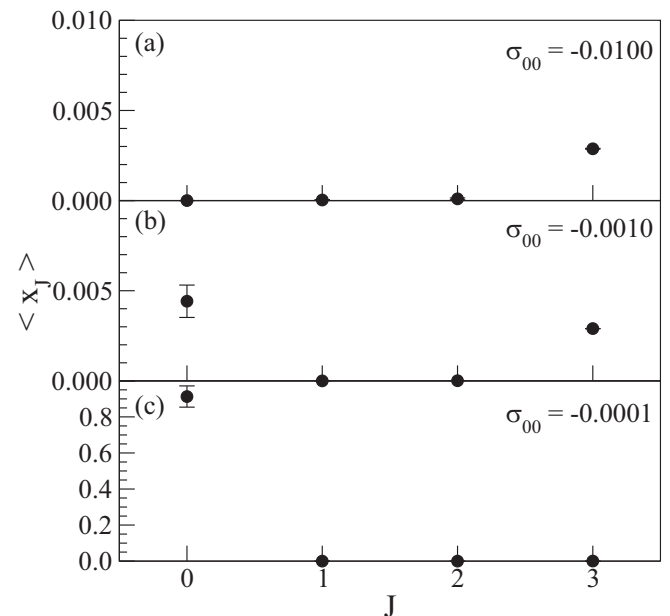

FIG. 9. Average steady-state mean relative abundances (x_J) for $L = 9$, $B = 3$, $p = 0.0384$ ($z \approx 2.43$, $c \approx 136$), $u = 0.1$, and $\sigma = -1$. Values of σ_{00} are (a) -0.01, (b) -0.001, and (c) -0.0001. Note the different ordinate scales.

theless benefit from them. Whether the presence of cheater genotypes can negatively impact the survival of genotypes that do produce public goods depends on a variety of issues. Ultimately, these issues can be summarized as two overall strategies, one highlighting positive effects on producers even if cheaters might be positively affected as well, the other highlighting negative effects on cheaters even if producers might be negatively affected as well. The first (strategy 1) is autoregulation of the costs versus benefits of public-good production; that is, public goods are only produced if the latter outweighs the former (cf., e.g., [43]). The second (strategy 2) is the exploitation of some weakening effect associated with public-good uptake by cheaters (e.g., [44,45]). All our computational results have genotype 0 as the sole producer of public goods and are therefore focused on the trade-offs affecting the survival of genotype 0 and all $2^L - 1$ cheaters.

The first group of results (depicted in Figs. 2–5) is for $\sigma_{00}, \sigma > 0$ and can be seen as modeling strategy 1, since $\sigma_{00} > 0$ can be taken as indicating that for genotype 0 public-good benefits outweigh production costs, and $\sigma > 0$ as indicating that cheaters benefit from the public goods without significant hindrance. These results are for $\sigma u = 0.01$, so in the absence of any mutational or HGT-related exchange between genotypes (as in Sec. III, case C1) we would have a sharp threshold for σ_{00} at 0.01 separating the exclusive survival of genotype 0 (for σ_{00} above the threshold) from the exclusive survival of the $n_1 = 21$ genotypes i having $J_{0i} = 1$ (for σ_{00} below the threshold). For σ_{00} precisely at the threshold the two groups would coexist.

Figures 2 and 3 both refer to scenarios in which, on average, genotype 0 undergoes genetic interactions, directly or indirectly, with about $(478 - 1)/(512 - 1) \approx 93.3\%$ of all other genotypes. Figure 2 shows individual relative abundances in the long run, Fig. 3 mean relative abundances. The value of σ_{00} is varied over three orders of magnitude from Figs. 2(a) and 3(a) through Figs. 2(d) and 3(d), leading genotype 0 from a situation of very low relative abundance for $\sigma_{00} = 0.001$ to one of clear preponderance for $\sigma_{00} = 1$. All along the competition is seen to be taking place between genotype 0 ($J = 0$) and those that differ from it at exactly $J = 1$ gene. All other genotypes are seen to be heading toward perishing. This progression along the increasing values of σ_{00} seems to lead to something resembling the outcome for case C1 of Sec. III, but there are important differences as well, all owing to the possibility of genetic interactions between genotype 0 and the others. One of these differences is that the $J = 1$ genotypes never vanish from the scene, not even as genotype 0 becomes the most abundant one. Another is that the sharp threshold at $\sigma_{00} = 0.01$ separating the demise of genotype 0 from the other regimes in Sec. III now seems spread over at least two orders of magnitude.

Figures 4 and 5 show the effect of restricting the possibilities of genetic interaction with genotype 0, first to about $(250 - 1)/(512 - 1) \approx 48.7\%$ of all other genotypes (Fig. 4), then to $(136 - 1)/(512 - 1) \approx 26.4\%$ (Fig. 5). While the overall trends relating the survival of genotype 0 to that of the $J = 1$ genotypes remain the same, it seems clear that genotype 0 comes ever closer to being the sole survivor as σ_{00} is increased and its possibilities of genetic interaction become limited to a smaller neighborhood in the graph.

The second group of results (depicted in Figs. 6–9) is for $\sigma_{00}, \sigma < 0$. This group can be regarded as modeling strategy 2, with $\sigma_{00} < 0$ reflecting the downside of the absence of autoregulation (public-good benefits are outweighed by production costs) and $\sigma < 0$ indicating that cheaters pay a heavy price for absorbing potentially beneficial public goods without joining in producing them. Results are now for $\sigma u^{L/B} = -0.001$, which like before would constitute a sharp threshold for σ_{00} at -0.001 should genotypes undergo no mutational or HGT-related interactions (as in Sec. III, case C9). Now the divide would be between the exclusive survival of genotype 0 (for σ_{00} above the threshold) and the exclusive survival of the $n_3 = 343$ genotypes i having $J_{i0} = 3$ (for σ_{00} below the threshold). Coexistence would ensue for σ_{00} precisely at the threshold.

In each of Figs. 6–9 the value of σ_{00} is increased by one order of magnitude from panel (a) to panel (b), then by one more from panel (b) to panel (c). Along these increases the competition for survival takes place between genotype 0 and those that differ from it at exactly $J = 3$ genes. All other genotypes are heading toward perishing. In all four figures the genotypes having $J = 3$ survive alone for $\sigma_{00} = -0.01$. For larger values of σ_{00} genotype 0 is seen to survive as well, becoming the most abundant one for $\sigma_{00} = -0.0001$. As σ_{00} is thus increased a transition similar to that of case C9 of Sec. III takes place, but once again this happens much less abruptly, along at least one order of magnitude.

As with the first group of figures, moving from Figs. 6 and 7 to Fig. 8 and then to Fig. 9 allows us to track the effects of letting genotype 0 interact genetically with progressively fewer other genotypes, from about 93.3% of them in Figs. 6 and 7, to about 48.7% in Fig. 8, then about 26.4% in Fig. 9. Clearly, as genotype 0 becomes more confined in its possibilities for genetic interaction, so does its prevalence relative to the $J = 3$ genotypes become more pronounced.

VI. CONCLUSION

We have studied the evolutionary dynamics of bacterial QS when multiple species (here referred to as genotypes) participate. We have focused on the production and uptake of public goods by the cells and on the trade-offs arising when cheater genotypes (those that have the potential to benefit from the uptake of public goods but do not join in producing them) are present. Our model has a number of parameters intended to allow for several scenarios to be considered. Two probability parameters (p, r) are related to how genotypes interact genetically, and another (u) is related to how molecular compatibility between genotypes influences the uptake by one genotype of public goods produced by another. Further parameters specify which genotypes are producers (the μ_i's) and whether (and how strongly) public-good uptake is detrimental or beneficial to a genotype's fitness (the σ_{ji}'s). Earlier studies of regulatory networks have revealed a fundamental link between the network and functional properties of a biological system and the leveraging of its intrinsic and environmental noise to achieve robustness [46]. In light of results such as this, the probability parameters p, r, and u can be regarded as helping characterize this important connection for bacterial QS as well. Interestingly, this

is fully in line with what happens in other complex systems in which noise has a positive role to play [14–16,47–49].

All our computational results refer to genotype 0 being the sole producer, all others being cheaters. All cheaters would be equally affected by the uptake of the public goods produced by genotype 0 if this only depended on the σ_{0i}'s (i.e., all σ_{0i}'s equal some fixed σ), but the action of parameter u causes significant differentiation. For fixed genotype and gene sizes, we have found that the long-term survival of genotype 0 depends chiefly on how σ_{00} stands relative to a function of σ and u (though parameters p and r also have a part to play). Our analytical results in the Appendix suggest that a similar conclusion may come to hold when multiple producers are considered. However, as suggested at the beginning of Sec. IV, tackling cases of multiple producers (and beyond these, substantially larger cases) is likely to require a complete reformulation of our current computational environment. This reformulation will have to harness the power of graphics processing units (GPUs), whose inherently massive parallelism is expected to lead computational performance to increase manyfold. Even so, the need for frequent sorting operations will still constitute a serious bottleneck and will have to be handled carefully. Doing this to good effect will depend on GPU architecture as well as on selecting and carefully adapting the sorting algorithm to be used. A detailed, recent review of the relevant issues involved is available [50].

Our model can be easily adapted, mainly by altering the modes of genetic interaction between genotypes, to the study of QS in other microorganisms (e.g., fungi [51] and fungal-bacterial systems [52]). In fact, to varying degrees the molecule-mediated communication that characterizes QS in microorganisms is observed also in other biological systems, from simple infectious agents such as viruses [53], to communities of so-called artificial cells [54], to precursor-cell clusters as they differentiate into specific organs and tissues (sometimes with surprisingly accurate global results; cf., e.g., [55]). We expect that also in some of these systems the analysis of long-term global behavior could benefit from the development of models closely related to the one we have presented.

We finalize by noting that, in spite of what we perceive as special novelties in our approach, such as a random network structure capable of accommodating a great many bacterial species, the bringing together of both bacterial evolution and QS into a single model, and the many ways in which this model can be fine tuned, it is still not testable. There are two reasons for this, the first being the main simplifying assumption mentioned at the end of the preamble to Sec. II, viz., that the public goods available for consumption constitute a static backdrop against which all else unfolds. Lifting this assumption will require further modeling that takes into account both the physical medium in question and the many facets of bacterial dissemination and public-good diffusion through the medium. The second reason is that, as with so many physical theories when they first appear, experimental data involving a large number of bacterial species as they trade public goods while coexisting are still unavailable. We expect that progress on these fronts will eventually make the theory truly testable.

APPENDIX: A SPECIAL CASE WITH MULTIPLE PRODUCERS

The special case of Sec. III can be extended to allow multiple producers of public goods while remaining tractable as far as predicting behavior in the long run is concerned. While studying multiple-producer scenarios computationally has proven infeasible given the resources we can presently access, the special case we analyze in this Appendix serves to illustrate a more complex situation in which the fate of producers and certain other genotypes continues to be determined, in the nontrivial cases, by a comparison between some producer k's σ_{kk} and a threshold.

The extension of Sec. III we now present preserves the $p = r = 0$ assumption of that section, so graph G continues to have only self-loops for edges and the dynamics continues to be given by Eq. (13). For uniform initial abundances, which we continue to assume, survival in the long run continues to occur for the initially fittest genotypes only, always with the same abundance.

Let P be the set of producers, i.e., $\mu_i = [i \in P]$. Assume $L/B \geqslant 2$ (at least two genes per genotype) and let g be a fixed gene out of a genotype's L/B genes. The multiple producers that we consider are the 2^B genotypes that differ from one another only at gene g. It follows that every genotype $i \notin P$ must agree with exactly one genotype $k_i \in P$ at gene g and differ from all the $2^B - 1$ other genotypes in P at this same gene. Therefore, for some $J_{ik_i} \in \{1, \dots, L/B - 1\}$, genotype i differs from producer k_i at J_{ik_i} genes and from all other producers at $J_{ik_i} + 1$ genes.

Similarly to how we proceeded in Sec. III, here we assume either $\sigma_{ji} = [j \in P][i \notin P]\sigma$ or $\sigma_{ji} = [j \in P][i \in P][i \neq j]\sigma_p$, with $\sigma, \sigma_p \in [-1, 1]$, depending on whether $i \in P$. That is, for $k \in P$, genotype $i \neq k$ is impacted by the public goods that genotype k produces either through $\sigma_{ki} = \sigma$ (if $i \notin P$) or through $\sigma_{ki} = \sigma_p$ (if $i \in P$). We also assume $\sigma_{kk} = \sigma_{pp}$ with $\sigma_{pp} \in [-1, 1]$ for all $k \in P$. All σ_{ji}'s with $j \notin P$ are irrelevant. It follows that every genotype $k \in P$ has the same value for x_k, d_k, and f_k, which we henceforth denote by x_p, d_p, and f_p. With these settings in place, by Eq. (2) we have

$$d_p = \sigma_{pp} x_p + (2^B - 1)\sigma_p u x_p \tag{A1}$$

and

$$d_i = \sigma u^{J_{ik_i}} x_p + (2^B - 1)\sigma u^{J_{ik_i}+1} x_p \tag{A2}$$

for $i \notin P$.

For $\sigma > 0$, the fittest genotypes $i \notin P$ are such that $J_{ik_i} = 1$. Comparing f_p to f_+ depends on how σ_{pp} relates to

$$\alpha = \left[1 - (2^B - 1)\left(\frac{\sigma_p}{\sigma} - u\right)\right]\sigma u, \qquad (A3)$$

which for $\sigma_{pp} > 0$ and $\sigma_p < [(2^B - 1)^{-1} + u]\sigma$ (to ensure $\alpha > 0$) extends case C1 of Sec. III. The resulting case is as follows.

(D1) $\sigma_{pp} > 0$, $\sigma > 0$, and $\sigma_p < [(2^B - 1)^{-1} + u]\sigma$: $f_p > 1$ and $f_+ > 1$, with three subcases.

(D1a) $\sigma_{pp} > \alpha$: $f_p > f_+$; genotypes $k \in P$ survive.

(D1b) $\sigma_{pp} = \alpha$: $f_p = f_+$; genotypes $k \in P$ survive, and so do all genotypes $i \notin P$ such that $J_{ik_i} = 1$.

(D1c) $\sigma_{pp} < \alpha$: $f_p < f_+$; genotypes $i \notin P$ such that $J_{ik_i} = 1$ survive.

Case C9 can be extended similarly by first noting that for $\sigma < 0$ the fittest genotypes $i \notin P$ are such that $J_{ik_i} = L/B - 1$. It follows that comparing f_p to f_+ depends on the relationship

between σ_{pp} and

$$\beta = \left[1 - (2^B - 1)\left(\frac{\sigma_p}{\sigma u^{L/B-2}} - u\right)\right]\sigma u^{L/B-1}. \qquad (A4)$$

The desired extension comes from assuming $\sigma_{pp} < 0$ and $\sigma_p > [(2^B - 1)^{-1} + u]\sigma u^{L/B-2}$ (to ensure $\beta < 0$). It is as follows.

(D9) $\sigma_{pp} < 0$, $\sigma < 0$, and $\sigma_p > [(2^B - 1)^{-1} + u]\sigma u^{L/B-2}$: $f_p < 1$ and $f_+ < 1$, with three subcases.

(D9a) $\sigma_{pp} > \beta$: $f_p > f_+$; genotypes $k \in P$ survive.

(D9b) $\sigma_{pp} = \beta$: $f_p = f_+$; genotypes $k \in P$ survive, and so do all genotypes $i \notin P$ such that $J_{ik_i} = L/B - 1$.

(D9c) $\sigma_{pp} < \beta$: $f_p < f_+$; genotypes $i \notin P$ such that $J_{ik_i} = L/B - 1$ survive.

The trivial cases of Sec. III (cases C2–C8) can be extended similarly. Of course, allowing for σ_p and σ to relate to each other differently will lead to further variations.

---

[1] M. B. Miller and B. L. Bassler, Annu. Rev. Microbiol. **55**, 165 (2001).

[2] L. A. Hawver, S. A. Jung, and W.-L. Ng, FEMS Microbiol. Rev. **40**, 738 (2016).

[3] K. Papenfort and B. L. Bassler, Nat. Rev. Microbiol. **14**, 576 (2016).

[4] J. Pérez-Velázquez, M. Gölgeli, and R. García-Contreras, Bull. Math. Biol. **78**, 1585 (2016).

[5] D. P. da Silva, M. C. Schofield, M. R. Parsek, and B. S. Tseng, Pathogens **6**, 51 (2017).

[6] M. Whiteley, S. P. Diggle, and E. P. Greenberg, Nature **551**, 313 (2017).

[7] K. H. Nealson, T. Platt, and J. Woodland Hastings, J. Bacteriol. **104**, 313 (1970).

[8] E. P. Greenberg, J. Woodland Hastings, and S. Ulitzur, Arch. Microbiol. **120**, 87 (1979).

[9] L. García-Bayona and L. E. Comstock, Science **361**, eaat2456 (2018).

[10] A. Heddi and A. Zaidman-Rémy, C. R. Biol. **341**, 290 (2018).

[11] S. T. Rutherford and B. L. Bassler, Cold Spring Harb. Perspect. Med. **2**, a012427 (2012).

[12] T. Defoirdt, Trends Microbiol. **26**, 313 (2018).

[13] D. M. Cornforth, R. Popat, L. McNally, J. Gurney, T. C. Scott-Phillips, A. Ivens, S. P. Diggle, and S. P. Brown, Proc. Natl. Acad. Sci. U.S.A. **111**, 4280 (2014).

[14] D. Valenti, L. Schimansky-Geier, X. Sailer, and B. Spagnolo, Eur. Phys. J. B **50**, 199 (2006).

[15] A. A. Dubkov and B. Spagnolo, Eur. Phys. J. B **65**, 361 (2008).

[16] D. Valenti, L. Tranchina, M. Brai, A. Caruso, C. Cosentino, and B. Spagnolo, Ecol. Modell. **213**, 449 (2008).

[17] M. Weber and J. Buceta, BMC Syst. Biol. **7**, 6 (2013).

[18] G. Cárcamo-Oyarce, P. Lumjiaktase, R. Kümmerli, and L. Eberl, Nat. Commun. **6**, 5945 (2015).

[19] N. Michelusi, J. Boedicker, M. Y. El-Naggar, and U. Mitra, in *Proceedings of the 49th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 2015* (IEEE, Piscataway, NJ, 2015), pp. 133–138.

[20] J. Singh and A. V. Anil Kumar, Phys. Rev. E **101**, 022606 (2020).

[21] T. I. Yusufaly and J. Q. Boedicker, Phys. Rev. E **94**, 062410 (2016).

[22] C. Abaurrea Velasco, M. Abkenar, G. Gompper, and T. Auth, Phys. Rev. E **98**, 022605 (2018).

[23] M. Rein, N. Heinß, F. Schmid, and T. Speck, Phys. Rev. Lett. **116**, 058102 (2016).

[24] M. Bauer and E. Frey, Phys. Rev. E **97**, 042307 (2018).

[25] G. Russo and J. J. E. Slotine, Phys. Rev. E **82**, 041919 (2010).

[26] U. K. Verma, S. S. Chaurasia, and S. Sinha, Phys. Rev. E **100**, 032203 (2019).

[27] A. Fischer, F. Schmid, and T. Speck, Phys. Rev. E **101**, 012601 (2020).

[28] R. Rosselló-Móra and R. Amann, Syst. Appl. Microbiol. **38**, 209 (2015).

[29] E. R. Angert, Nat. Rev. Microbiol. **3**, 214 (2005).

[30] S. Bulgherese, Environ. Microbiol. **18**, 2305 (2016).

[31] J. P. J. Hall, M. A. Brockhurst, and E. Harrison, Philos. Trans. R. Soc. B **372**, 20160424 (2017).

[32] V. C. Barbosa, R. Donangelo, and S. R. Souza, J. Theor. Biol. **312**, 114 (2012).

[33] M. Eigen, Naturwissenschaften **58**, 465 (1971).

[34] M. Eigen and P. Schuster, Naturwissenschaften **64**, 541 (1977).

[35] C. K. Biebricher and M. Eigen, in *Quasispecies: Concept and Implications for Virology*, Current Topics in Microbiology and Immunology Vol. 299, edited by E. Domingo (Springer, Berlin, 2006), pp. 1–31.

[36] E. Domingo, Contrib. Sci. **5**, 161 (2009).

[37] A. S. Lauring and R. Andino, PLoS Pathog. **6**, e1001005 (2010).

[38] A. Más, C. López-Galíndez, I. Cacho, J. Gómez, and M. A. Martínez, J. Mol. Biol. **397**, 865 (2010).

[39] V. C. Barbosa, R. Donangelo, and S. R. Souza, J. Stat. Mech. (2015) P01022.

[40] V. C. Barbosa, R. Donangelo, and S. R. Souza, J. Stat. Mech. (2016) 063501.

[41] V. C. Barbosa, R. Donangelo, and S. R. Souza, Phys. Rev. E **98**, 032409 (2018).

[42] O. M. Sokolovskaya, A. N. Shelton, and M. E. Taga, Science **369**, eaba0165 (2020).

[43] J. B. Xavier, W. Kim, and K. R. Foster, Mol. Microbiol. **79**, 166 (2011).

[44] A. A. Dandekar, S. Chugani, and E. P. Greenberg, Science **338**, 264 (2012).

[45] C. Majerczyk, E. Schneider, and E. P. Greenberg, eLife **5**, e14712 (2016).

[46] R. Silva-Rocha and V. de Lorenzo, Annu. Rev. Microbiol. **64**, 257 (2010).

[47] B. Spagnolo, D. Valenti, C. Guarcello, A. Carollo, D. Persano Adorno, S. Spezia, N. Pizzolato, and B. Di Paola, Chaos Soliton. Frac. **81**, 412 (2015).

[48] D. Valenti, L. Magazzù, P. Caldara, and B. Spagnolo, Phys. Rev. B **91**, 235412 (2015).

[49] B. Spagnolo, C. Guarcello, L. Magazzù, A. Carollo, D. Persano Adorno, and D. Valenti, Entropy **19**, 20 (2017).

[50] D. I. Arkhipov, D. Wu, K. Li, and A. C. Regan, arXiv:1709.02520.

[51] J. M. Hornby, E. C. Jensen, A. D. Lisec, J. J. Tasto, B. Jahnke, R. Shoemaker, P. Dussault, and K. W. Nickerson, Appl. Environ. Microbiol. **67**, 2982 (2001).

[52] E. F. Dixon and R. A. Hall, Cell. Microbiol. **17**, 1431 (2015).

[53] Z. Erez, I. Steinberger-Levy, M. Shamir, S. Doron, A. Stokar-Avihail, Y. Peleg, S. Melamed, A. Leavitt, A. Savidor, S. Albeck, G. Amitai, and R. Sorek, Nature **541**, 488 (2017).

[54] B. C. Buddingh', J. Elzinga, and J. C. M. van Hest, Nat. Commun. **11**, 1652 (2020).

[55] Y. Afek, N. Alon, O. Barad, E. Homstein, N. Barkai, and Z. Bar-Joseph, Science **331**, 183 (2011).

Correction: Some quantities were presented with incorrect indices and have been fixed: J_{ji} should read J_{ij} in the third paragraph of Sec. II. $J_{k_i i}$ should read J_{ik_i} in Eq. (A2)ff. $J_{J_{k_i i}}$ should read J_{ik_i} after Eq. (A2).