PHYSICAL REVIEW E **94**, 062410 (2016)

Spatial dispersal of bacterial colonies induces a dynamical transition from local to global quorum sensing

Tahir I. Yusufaly[1] and James Q. Boedicker[1,2]
[1]*Department of Physics and Astronomy, University of Southern California, Los Angeles, California 90089, USA*
[2]*Department of Biological Sciences, University of Southern California, Los Angeles, California 90089, USA*

(Received 13 March 2016; revised manuscript received 2 December 2016; published 21 December 2016)

Bacteria communicate using external chemical signals called autoinducers (AI) in a process known as quorum sensing (QS). QS efficiency is reduced by both limitations of AI diffusion and potential interference from neighboring strains. There is thus a need for predictive theories of how spatial community structure shapes information processing in complex microbial ecosystems. As a step in this direction, we apply a reaction-diffusion model to study autoinducer signaling dynamics in a single-species community as a function of the spatial distribution of colonies in the system. We predict a dynamical transition between a local quorum sensing (LQS) regime, with the AI signaling dynamics primarily controlled by the local population densities of individual colonies, and a global quorum sensing (GQS) regime, with the dynamics being dependent on collective intercolony diffusive interactions. The crossover between LQS to GQS is intimately connected to a trade-off between the signaling network's latency, or speed of activation, and its throughput, or the total spatial range over which all the components of the system communicate.

DOI: 10.1103/PhysRevE.94.062410

I. INTRODUCTION

Multicellular communities, such as colonies of bacteria, communicate with each other to coordinate changes in their collective group behavior. This communication usually takes the form of the production and secretion of extracellular signaling molecules called autoinducers (AI), as illustrated in Fig. 1. Released autoinducers diffuse through the environment, and each cell senses the local concentration of signal to inform changes in gene regulation. This intercellular signaling network, known as quorum sensing (QS), is crucial for a wide array of important microbial processes, including biofilm formation, regulation of virulence, and horizontal gene transfer [1–3].

Decades of research have advanced our knowledge of QS, but several subtleties remain unresolved. In particular, AI signals may convey information about many aspects of the cellular network and local environment beyond simply the total number of cells in the system [4]. Far from being reducible to homogeneous, uniform density populations, microbial communities are typically characterized by high spatial heterogeneity [5]. As a result, several new phenomena emerge due to crosstalk between spatially segregated populations [6]. Consequently, it appears that AI molecules can be an indicator of increased local population density, and can also be proxies of other variables, such as population dispersal [7–11].

In recent years, advances in the ability to experimentally probe the properties of cellular populations at the single-cell level [12] have resulted in a growing community of theoretical physicists working to catalog the different classes of collective behavior found in interacting communities of organisms [13–17]. This approach has already successfully yielded insight into a wide variety of ecological problems, with notable recent examples including the effects of invasion in cooperative populations [18], optimal foraging strategies in sheep herds [19], and the properties of microbial signal transduction networks [20–22] and their relationship to long-range spatial pattern formation [23,24]. However, to our knowledge, there have been few, if any, such studies on the specific question of exactly how spatiotemporal autoinducer profiles are influenced by the underlying heterogeneity of the microbial community that produces them.

Accordingly, in this article, using a reaction-diffusion model, we identify how emergent spatiotemporal AI signaling patterns depend on community spatial structure. Via numerical simulations, we show that for a community of single-species bacteria, there is a transition in the activation dynamics as a function of community dispersal, or equivalently, the spatial heterogeneity of the cell density. At low dispersal, corresponding to a situation in which the cells aggregate into a few large colonies, the activation is triggered by the local size and density of the individual colonies, and therefore, can be described as local QS (LQS). At high dispersal, corresponding to a situation in which the cells are spread thinly among many small colonies, activation is instead a collective, global effect, mediated by mutual interactions between spatially disconnected colonies, and thus, this state is best described as one of global QS (GQS).

II. METHODS

The dynamics of the autoinducer concentration c_{AI} are governed by the spatiotemporal cell density profile c_{cell} via the reaction-diffusion equation

$$\left(\frac{\partial}{\partial t} - D\nabla^2 + \gamma\right)c_{\mathrm{AI}} = r(c_{\mathrm{AI}})c_{\mathrm{cell}}(\vec{x},t), \tag{1}$$

where D and γ are, respectively, the AI signal diffusivity and decay, and r is the local, AI-concentration-dependent signal production rate of the cells. In this work, we restrict ourselves to a regime where the time scale of sigmoidal quorum sensing activation is taken to be much faster than the time scale for signal production and diffusion, which implies that in the limit of coarse-grained, long-time integrations, the activation can

©2016 American Physical Society

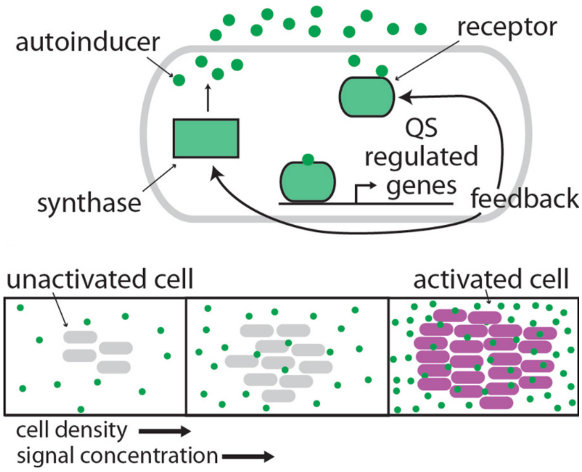

FIG. 1. The paradigmatic quorum sensing regulatory circuit consists of a synthase that produces autoinducer signals and a receptor that senses the local concentration of the same autoinducer. In the absence of a high density of bacteria, the synthase genes are expressed at a low basal level, secreting a small amount of autoinducer into the environment. Once the colonies have grown to beyond a critical population density, the collective concentration of autoinducer secreted is enough to activate quorum sensing.

be approximated by an instantaneous all-or-none step function [for a specific recent study analyzing the range of validity of this approximation in quorum sensing systems, we refer the reader to [30], and for further validation data comparing results of the Heaviside activation to a more realistic Hill curve, we refer to Fig. 5(a) in the Appendix]. Thus, we can approximate the activation to be instantaneous, modeling it via the Heaviside theta function θ:

$$r(c_{\mathrm{AI}}) = r_b\{1 + f\theta[c_{\mathrm{AI}}(\vec{x},t) - c_{\mathrm{AI}_{\mathrm{crit}}}]\}. \tag{2}$$

Here, r_b is the basal AI production rate of the cells in the absence of any activation, while $c_{\mathrm{AI}_{\mathrm{crit}}}$ is the threshold AI concentration for cells to transition to an activated state, with a fold change f ratio increase in synthase gene expression, and thus AI production rate.

Additionally, we take the cell density c_{cell} to be a Boolean function, adopting either a constant average $\bar{c}_{\mathrm{colony}}$ for points occupied by a colony, or 0 otherwise. With this setup, the spatiotemporal density of the cellular community c_{cell} can be interpreted as an input driving signal that generates the spatiotemporal AI profile c_{AI} as an output response signal.

We design and perform a series of computational experiments to isolate the effects of spatial dispersal on AI signaling. At time $t = 0$, we instantaneously colonize a quasi-two-dimensional (quasi-2D) region of space with an initial population of n_0 cells, and have the cells exponentially grow until they have multiplied to a final total carrying capacity value n_{max}, at which point the growth is saturated.

We perform several replicates of this simulation. In the first replicate, the cells are perfectly localized into a circular colony at the center of the quasi-2D region, corresponding to a situation of minimal dispersal, with only one homogeneous colony and no spatially distributed subpopulations.

TABLE I. The representative set of parameters used in this work. Although these quantities possess substantial statistical uncertainty and condition-dependent variability, reported values in the literature typically fall within the same characteristic orders of magnitude.

Quantity	Units	Value
Colony cell density ($\bar{c}_{\mathrm{colony}}$)	cells/μm^3	1[a]
Fold change (f)	Unitless	6[b]
Signal diffusivity (D)	μm^2/s	160[c]
Signal decay rate (γ)	molecules/h	0.01[d]
Activation threshold ($c_{\mathrm{AI}_{\mathrm{crit}}}$)	molecules/μm^3	100[e]
Basal production rate (r_b)	molecules/h/cell	500[f]
Doubling time	min	30[g]
Initial no. of cells (n_0)	10^3 cells	7
Max. no. of cells (n_{max})	10^3 cells	7000

[a] Taken directly from [8].

[b] Selected as a lower limit from the reported range of 5-fold to 300-fold change in [25], for numerical convenience.

[c] Taken with slight alteration for computational convenience, from [25].

[d] Taken directly from [26].

[e] Taken as a computationally convenient number corresponding to 160 nM concentration, an intermediate between the 10–70 nM reported in [27] and the 450 nM reported in [28].

[f] Adapted from information in [28]. In particular, point production intensity is (0.43 nM/h)×l, where l is the luminescence, ranging between 0 and 3000 (normalized), thus corresponding to an intensity range from about 0 to 1300 nM/h. For a 1 μm^3 cell, this corresponds to approximately 0–1000 molecules per hour per cell, with the midpoint value of 500 thus being a reasonable estimate of a typical rate.

[g] Rounded up from mean values reported in [29].

The cells produce, sense, and activate AI signals undergoing diffusion and decay according to Eqs. (A18) and (A19), with the representative parameter values shown in Table I. The system is numerically integrated to yield a discretized output spatiotemporal AI signaling profile $c_{\mathrm{AI}}(\vec{x},t)$. Calculations are performed with the aid of the BSIM software package [31]; for additional details on the setup of the simulation, we refer the reader to the Appendix.

For subsequent replicates of the simulation, we keep these details the same, except that we divide the original single colony into various discrete numbers of colonies N_{colonies}, and confine the colonies to occupy a specified square area with side length $L_{\mathrm{confinement}}$. Then, different sized confinement squares tune the packing density of the colonies, or the ratio of the total area of all colonies to the area in which they are confined. This setup is displayed in Fig. 2. In all cases, the initial number of cells n_0 and final number of cells at full growth, n_{max}, are the same as in the single-colony simulation. Consequently, increasing values of N_{colonies} gradually tune the system from a state of a few large colonies, to one of many smaller colonies, with a corresponding increase in the concentration of interacting patches of spatially extended subpopulations. Additionally, increasing areas of confinement gauges the importance of intercolony interactions, because as the packing density decreases, so does the average intercolony spacing, and thus, the magnitude of diffusive crosstalk between spatially separated colonies.

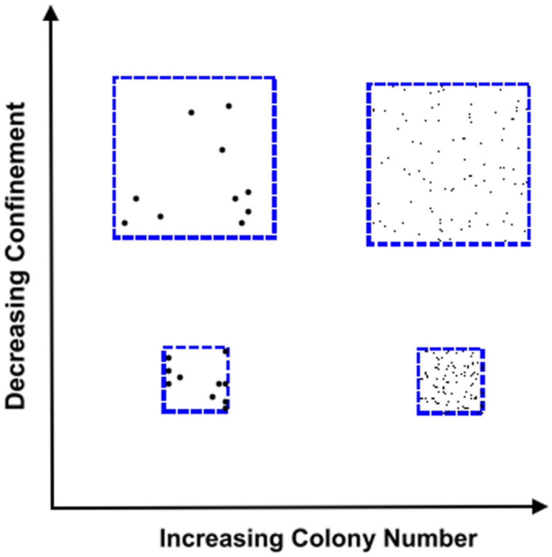

FIG. 2. In this study, we divide an initial population of n_0 cells into N_{colonies} equally sized colonies, coexisting within a square region of length $L_{\text{confinement}}$. These two parameters, N_{colonies} and $L_{\text{confinement}}$, serve as control parameters of our model. The number of colonies N_{colonies} tunes the *local* properties of individual colonies, while the confinement tunes the strength of *global* intercolony communication through molecular diffusion.

III. RESULTS AND DISCUSSION

We define a region of space as being activated if its concentration of AI signal is above the quorum sensing threshold $c_{\text{AI}_{\text{crit}}}$. Figure 3(a) illustrates the fraction of space that gets activated, at a sample intermediate time of 10 h, for the different spatial geometries studied. At low colony number N_{colonies}, corresponding to a small number of highly clustered colonies, the system is in a regime of LQS, where regions near the individual colonies receive enough locally produced signal to enable rapid activation, in a manner that is robust to the activity of neighboring colonies. This robustness can be seen by observing how the activation of the system changes as $L_{\text{confinement}}$ increases, corresponding to further separation between colonies. We see that for $N_{\text{colonies}} < N_c = 18$, the activation does not vanish but approaches a finite limiting value. Meanwhile, for $N_{\text{colonies}} > N_c$, the system is in a regime of GQS, where activation is necessarily communal in nature, triggered by the collective sharing of AI signal between all colonies. Here, as $L_{\text{confinement}}$ increases, the activation level vanishes because individual colonies are too small to self-activate without crosstalk from neighboring colonies. These arguments are further corroborated in Fig. 3(b), which shows asymptotic long-time coverage in the deconfined limit, $L_{\text{confinement}} \to \infty$, where the colonies are completely noninteracting. In this regime, it is even more apparent that there is a discontinuous transition in activation at N_c. In the Appendix, our numerical results are compared to an analytical mean-field model which captures the essential physics governing the transition.

Furthermore, working in nondimensionalized units (see the Appendix for more details; for the purposes of continuity, we stick to conventional units throughout the main article) demonstrates that, despite the significant number of input quantities in the model, the qualitative behavior of the system is invariant to most of the detailed parameter values, and instead, only depends on the basal activity level of colonies, as determined by r_b, the fold change f upon activation, and the maximal cell number $n_{\max}$. The existence of a critical threshold N_c separating these two regimes is independent of the fold change f, as activation of an isolated colony depends only on its own basal activity and geometry, and is not influenced by any possible activated neighboring colonies. In addition, variations in the total cell number $n_{\max}$ and basal activity level r_b simply shift N_c. In particular, increasing the value of $n_{\max}$ simply increases N_c by increasing the number of divisions required to fine-grain individual colonies to a size below the self-activation limit. Meanwhile decreasing r_b simply decreases N_c by decreasing the basal intensity per unit area, and thus the total basal intensity for a given colony radius.

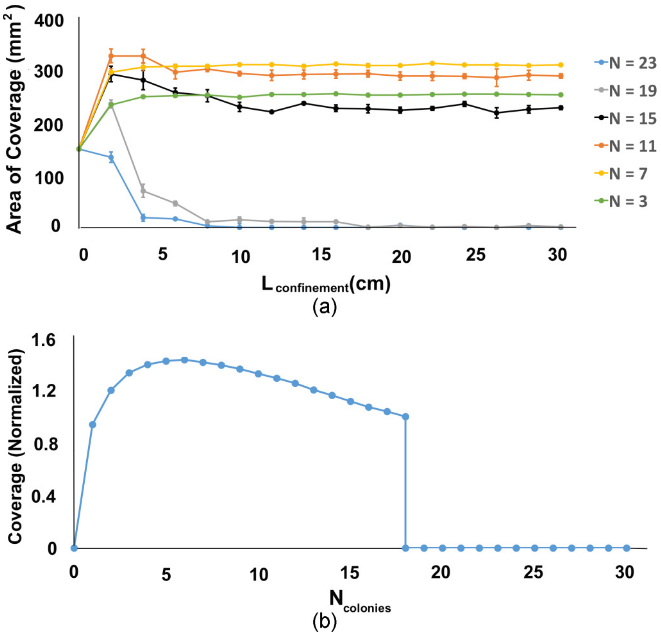

FIG. 3. (a) The average coverage, or area of space with above-threshold signal concentration, at a representative time point of 10 h, plotted as a function of the dimension of confinement $L_{\text{confinement}}$ for varying numbers of colonies N_{colonies}. Residuals represent standard errors over five simulation replicates each. (b) Asymptotic coverage level as a function of N_{colonies} in the limit of deconfined, or infinitely separated colonies. There is a clear transition from nonzero to zero levels between $N_{\text{colonies}} = 17$ and 18, illustrating a crossover from LQS to GQS.

This transition from LQS to GQS is related to trade-offs between the signaling network's latency, or speed of activation, and its throughput, or the total spatial range over which all the components of the system communicate [32]. Communities in the LQS regime have a reduced time to activation, but are restricted to short-range communication. The large colony size leads to production of signal near the colony center beyond what is needed for local activation. The slow rate of signal diffusion coupled with larger intercolony distances does not allow the signal to effectively communicate with neighboring colonies. On the contrary, although communities in the GQS

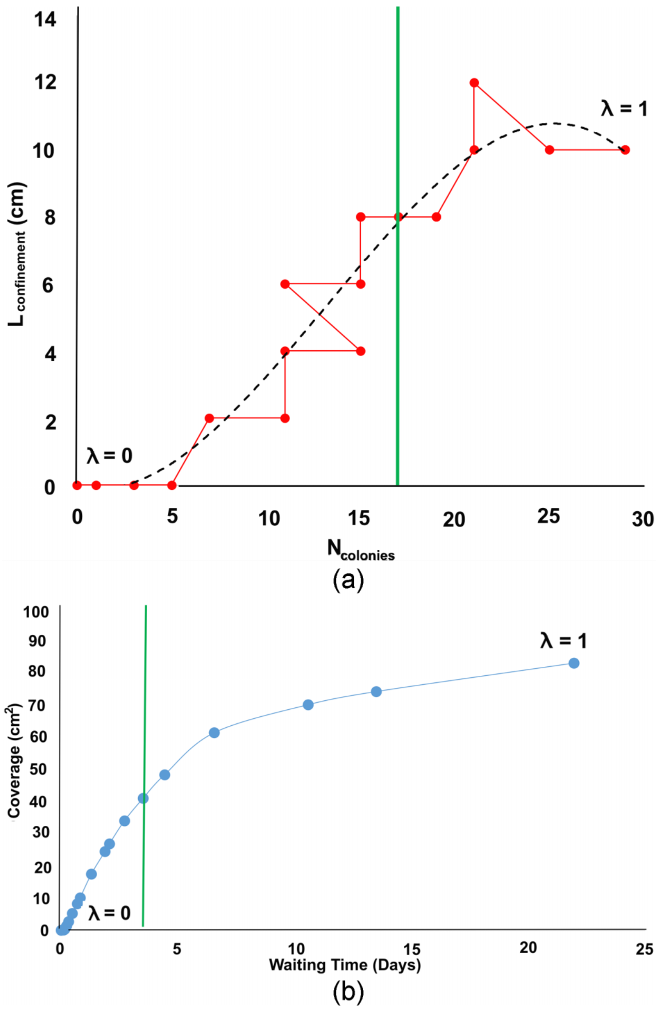

FIG. 4. (a) A Pareto optimal curve of control parameters $[N_{\text{optimal}}(\lambda),\ L_{\text{optimal}}(\lambda)]$, parametrized by an affine parameter λ ranging from 0 to 1. Each value of λ maps on to an associated Pareto point in $(N_{\text{colonies}},\ L_{\text{confinement}})$ parameter space, such that, for a given waiting time from the moment of initial colonization $\tau(\lambda)$ the overall coverage C (defined in Fig. 3) at τ is maximized (within the resolution of the numerical simulations)—quantitatively, $C_{\max}(\lambda) = C(\tau(\lambda), N_{\text{optimal}}(\lambda), L_{\text{optimal}}(\lambda)) \geqslant C(\tau(\lambda), N_{\text{colonies}}, L_{\text{confinement}})$ for all N_{colonies} and $L_{\text{confinement}}$. The red dots correspond to the observed discrete values of $(N_{\text{colonies}},\ L_{\text{confinement}})$, while the black dotted line is a resulting best-fit interpolation curve. Displayed below this optimal curve $[N_{\text{optimal}}(\lambda),\ L_{\text{optimal}}(\lambda)]$ is (b) a plot of the corresponding maximal coverage $C_{\max}(\lambda)$ vs the waiting time $\tau(\lambda)$. The green lines indicate a crossing from an LQS to GQS regime of quorum sensing, resulting in a rapid onset of "diminishing returns" in throughput, relative to the latency cost.

regime take longer to initiate activity, the broadly distributed colonies of smaller size enable long-range communication throughout the network of cells.

These trade-offs are illustrated in Fig. 4 describing a Pareto optimal contour in parameter space, with each point on the contour corresponding to a choice of parameters that, for a given waiting time from the moment of initial colonization τ, optimizes the overall coverage at that time, which we denote $C(\tau)$. The results show that if one is primarily interested in rapid initiation, it is preferable to be in the LQS regime, but that if one is interested in having the maximum possible fraction of space receive the signal in a reasonable amount of time, it is preferable to be near the GQS regime.

These characteristics of the optimal contour are, to a large degree, independent of slight variations in parameter values. As has already been discussed, changes in n_{cells} and r_b simply shift N_c, and thus, only show up as smooth deformations of the optimal contour, with the qualitative form of the peak, as well as its dynamic variation, still intact. Furthermore, slight variations in the remaining free parameter, the fold change f, simply amount to changes in the steepness of the contour, by stretching or compressing the magnitude of separation between active and inactive coverage levels. However, the fact remains that in the GQS regime, $N_{\text{colonies}} > N_c$, there comes a point of "diminishing returns," where coverage eventually decreases as $L_{\text{confinement}}$ increases, due to activation depending on the distance between neighboring colonies. For additional numerical evidence validating these physical arguments, we refer the reader to Figs. 5(b) and 5(c) of the Appendix.

IV. CONCLUSION

In summary, the results of this work clearly point to the prediction of an experimentally observable transition between LQS and GQS behavior. This transition is observed to be intimately connected to a tradeoff between short-term latency, optimized in the LQS regime, and long-term throughput, optimized in the GQS regime. The qualitative results are independent of specific parameter values, being primarily a consequence of how coverage dynamics scale as colony number and confinement are varied. A natural follow-up to this work would be exploring how the behavior of the trade-offs is influenced by heterogeneity, for example, due to the presence of multiple species and single-cell stochasticity.

ACKNOWLEDGMENTS

This work was funded by the Office of Naval Research under Grant No. N00014-15-1-2573. We thank Sylvian Hemus for early help with this project, and Osman Kahraman and Christoph Haselwandter for reading and commenting on the manuscript. We would also like to thank all the referees who reviewed our manuscript, whose suggestions greatly improved the work.

APPENDIX

1. Continuous formulation of reaction-diffusion equations

The dynamics of our system are governed by the reaction-diffusion equation

$$\left(\frac{\partial}{\partial t} - D\nabla^2 + \gamma\right) c_{\mathrm{AI}} = r(c_{\mathrm{AI}})\, c_{\mathrm{cell}}(\vec{x},t), \tag{A1}$$

where D and γ are, respectively, the AI signal diffusivity and decay, and r is the local, AI-concentration-dependent signal production rate of the cells, given by

$$r(c_{\mathrm{AI}}) = r_b\{1 + f\theta[c_{\mathrm{AI}}(\vec{x},t) - c_{\mathrm{AI}_{\mathrm{crit}}}]\}. \tag{A2}$$

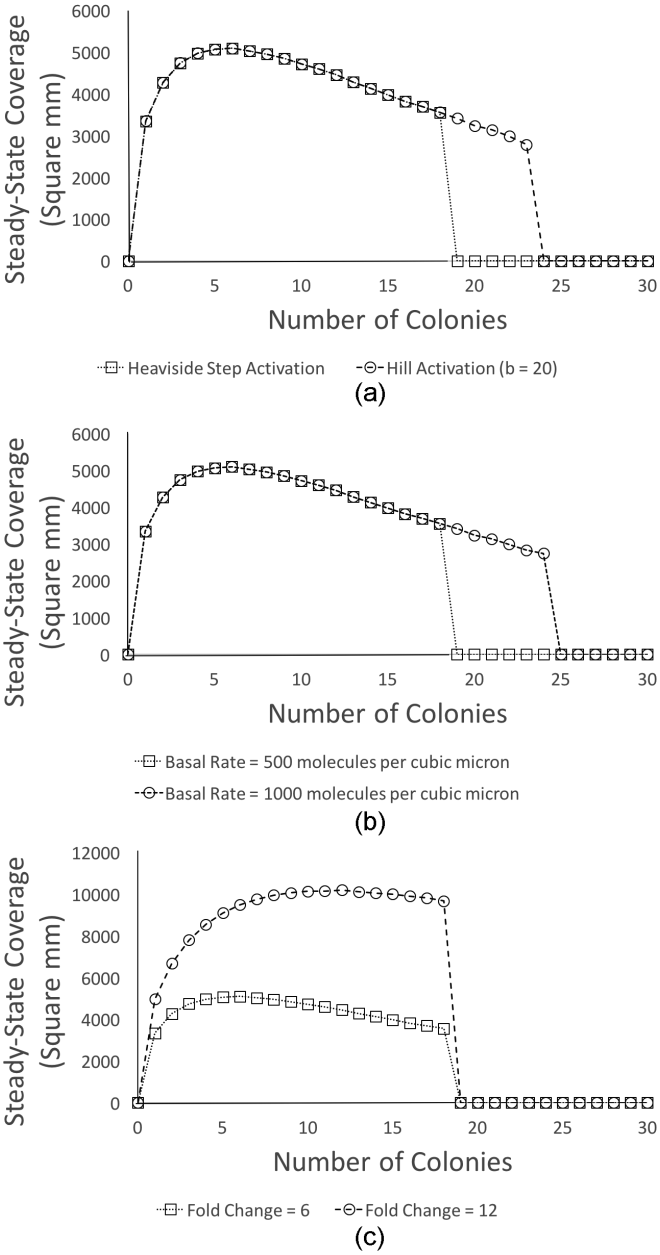

FIG. 5. Changes in the asymptotic coverage level curve as a function of N_{colonies}, in the limit of deconfined, or infinitely separated colonies for variations in (a) sharpness of the activation (relaxing the instantaneous Heaviside-step function with a Hill curve with coefficient $b = 20$), (b) basal production rate r_b, and (c) fold change f.

Here, r_b is the basal AI production rate of the cells in the absence of any activation, while $c_{\text{AI}_{\text{crit}}}$ is the threshold AI concentration for cells to transition to an activated state, with a fold change f ratio increase in synthase gene expression, and thus AI production rate. We take the activation to be instantaneous, modeling it via the Heaviside theta function θ, which is a reasonable approximation to the smoother Hill function as long as the time scale of the activation reaction is much faster than that of diffusion or signal production. With this setup, the time-dependent cell density profile $c_{\text{cell}}(\vec{x},t)$ serves as a tunable input driving signal, generating an output $c_{\text{AI}}(\vec{x},t)$.

As discussed in the text, the cell density is divided into a discrete number of equally sized building-block pieces centered at an equivalent number of colony center points, with each building block representing a separate colony. The building blocks are taken to have a given cell density $\bar{c}_{\text{colony}}$, representing the number of cells per unit area in a colony, and a time-dependent radius $R(t)$, representing colony growth. Thus, a given building-block density profile, for a colony centered at $\vec{x}_0$, is

$$c_{\text{colony}}(\vec{x},t;\vec{x}_0,R(t)) = \begin{cases} \bar{c}_{\text{colony}}, & |\vec{x}-\vec{x}_0| \leqslant R(t) \\ 0, & |\vec{x}-\vec{x}_0| > R(t). \end{cases} \tag{A3}$$

With this notation, if we confine our cells to a square area $A_{\text{confinement}} = L_{\text{confinement}} \times L_{\text{confinement}}$, then a general cell density profile c_{cell} can be written as a linear superposition of all colonies in the system

$$c_{\text{cell}}(\vec{x},t) = \sum_{i=1}^{N_{\text{colonies}}} c_{\text{colony}}(\vec{x},t;R(t),\vec{x}_{0i}(L_{\text{conf.}})), \tag{A4}$$

where N_{colonies} is the total number of colonies, $\{\vec{x}_{0i}\}$ is the set of colony centers, and $R(t)$ is the colony radii, which is constant for all colonies. As shown in Eq. (A4), the possible center positions are functionally constrained by the confinement area. We will discuss the choice of the $\{\vec{x}_{0i}\}$ momentarily, but first we address determination of $R(t)$. We proceed by noting that the total number of cells n_{cells} in the system starts at a given initial number n_0, then exponentially grows to a given maximum level $n_{\max}$, with a given time constant $t_{1/2}$, at which point the cells immediately stop dividing. The time $t_{\max}$ at which the population reaches this maximum level can easily be calculated by setting $n_{\max} = n_0 e^{t_{\max}/t_{1/2}}$, which gives

$$t_{\max} = t_{1/2}\ln(n_{\max}/n_0). \tag{A5}$$

Thus, the explicit time-dependent profile for the total number of cells is

$$n_{\text{cells}}(t) = \begin{cases} n_0 e^{t/t_{1/2}}, & 0 \leqslant t \leqslant t_{\max} \\ n_{\max}, & t_{\max} < t. \end{cases} \tag{A6}$$

To relate this to colony radii, we note that the total cell density profile n_{cell}, for all replicates, must satisfy the normalization condition

$$\int_{-L/2}^{L/2}\int_{-L/2}^{L/2} d^2\vec{x}\; c_{\text{cell}}(\vec{x},t) = N_{\text{cells}}(t). \tag{A7}$$

Thus, if the total number of cells is equally divided between N_{colonies} number of colonies, each with uniform cell density $\bar{c}_{\text{colonies}}$, then it follows that the radius $R(t)$ of each colony must be equal to

$$R(t) = \frac{1}{N_{\text{colonies}}}\sqrt{\frac{n_{\text{cells}}(t)}{\pi\,\bar{c}_{\text{colonies}}}}. \tag{A8}$$

We note that it follows immediately from this result that the radius of a single colony at full growth is given by

$$R_{\text{colony}} = \frac{1}{N_{\text{colonies}}}\sqrt{\frac{n_{\max}}{\pi\,\bar{c}_{\text{colonies}}}}. \tag{A9}$$

With this information, we now proceed to identify a range of possible colony centers—in particular, we would like to calculate, for a given number of colonies N_{colonies}, each with radius at full growth R_{colony}, and confined to an $L_{\text{confinement}} \times L_{\text{confinement}}$ square, the probability distribution over the N_{colonies}-dimensional random vector of colony centers $\vec{X}_0 = (\vec{x}_{01}, \vec{x}_{02}, \ldots, \vec{x}_{0N_{\text{colonies}}})$. The range of allowed center positions is set by two constraints: (1) no center can be closer than a distance R_{colony} to the edge of the square of confinement, to ensure that no cells leave the confinement boundary, and (2) colony centers must be separated by a distance of at least $2R_{\text{colony}}$ in order to avoid overlap. Thus, the probability distribution $p(\vec{X}_0)$ is simply a uniform distribution over all colony center combinations that remain within the prescribed boundaries and which do not contain pairs of points closer than the cutoff separation. While an analytical form of this distribution is, in general, untenable, it is straightforward to implement numerically.

2. Nondimensionalization

We can reduce the total number of independent parameters by transforming to unitless dimensions of space, time, molecular, and cellular concentrations. We start with our governing equation,

$$\left(\frac{\partial}{\partial t} - D\nabla^2 + \gamma\right)c_{\text{AI}} = r(c_{\text{AI}})c_{\text{cell}}(\vec{x},t), \tag{A10}$$

with

$$r(c_{\text{AI}}) = r_b\left\{1 + f\theta\left[c_{\text{AI}}(\vec{x},t) - c_{\text{AI}_{\text{crit}}}\right]\right\}. \tag{A11}$$

If we make a change of variables to unitless quantities, via the prescription

$$(\tilde{\vec{x}},\tilde{t},\tilde{c}_{\text{AI}},\tilde{c}_{\text{cell}}) = \left(\sqrt{\frac{\gamma}{D}}\vec{x}, \gamma t, \frac{c_{\text{AI}}}{c_{\text{AI}_{\text{crit}}}}, \frac{c_{\text{cell}}}{\bar{c}_{\text{colony}}}\right), \tag{A12}$$

then Eq, (A18) adopts the nondimensionalized form

$$\left(\frac{\partial}{\partial \tilde{t}} - \tilde{\nabla}^2 + 1\right)\tilde{c}_{\text{AI}} = \tilde{r}(\tilde{c}_{\text{AI}})\tilde{c}_{\text{cell}}(\tilde{\vec{x}},\tilde{t}). \tag{A13}$$

Thus, $\tilde{c}_{\text{cell}}$ is now a unitless source field that adopts discrete Boolean values of either 0 (corresponding to no cells present at a particular point in space and moment in time) or 1 (corresponding to cells being present with density $\bar{c}_{\text{colony}}$). Furthermore, the activation function (A11) has been transformed into the rescaled, unitless form

$$\tilde{r}(\tilde{c}_{\text{AI}}) = \left(\frac{r_b\bar{c}_{\text{colony}}}{\gamma c_{\text{AI}_{\text{crit}}}}\right)[1 + f\theta(\tilde{c}_{AI} - 1)]. \tag{A14}$$

The first set of parameters in parentheses on the right-hand side is nothing more than the unitless steady-state basal intensity of a point source of density $\bar{c}_{\text{colony}}$, as can be seen by setting the rate of production equal to the rate of decay, $r_b\bar{c}_{\text{colony}} = \gamma c_{\text{SS}}$, from which it is easily shown that

$$\tilde{c}_{\text{SS}} = \frac{c_{\text{SS}}}{c_{\text{AI}_{\text{crit}}}} = \frac{r_b\bar{c}_{\text{colony}}}{\gamma c_{\text{AI}_{\text{crit}}}}. \tag{A15}$$

Summarizing, we have shown that, through judicious choice of units, the behavior of the system depends on significantly fewer parameters than we started out—in particular, only the basal activity level $\tilde{c}_{\text{SS}}$ and the fold change upon activation f tune the output autoinducer response $\tilde{c}_{\text{AI}}(\tilde{\vec{x}},\tilde{t})$ for a given input cell source profile $\tilde{c}_{\text{cell}}(\tilde{\vec{x}},\tilde{t})$.

3. Discretization and computation

Discretization of the continuous reaction-diffusion system for numerical computation is fairly straightforward. We perform calculations in a simulation box of dimensions $L_x^{\text{box}} \times L_y^{\text{box}} \times L_z^{\text{box}}$ where L_x^{box} and L_y^{box} are set to be $\Delta L = 5$ cm greater than the confinement length $L_x^{\text{box}} = L_y^{\text{box}} = L_{\text{confinement}} + \Delta L$, in order to reduce finite-size numerical artifacts, and $L_z^{\text{box}} = 1$ μm. Continuous functions are discretized into an $N_x^{\text{grid}} \times N_y^{\text{grid}} \times N_z^{\text{grid}}$ mesh of rectangular units, each unit consequently having dimensions $\frac{L_x^{\text{box}}}{N_x^{\text{grid}}} \times \frac{L_y^{\text{box}}}{N_y^{\text{grid}}} \times \frac{L_z^{\text{box}}}{N_z^{\text{grid}}}$. Mesh units are chosen such that $\frac{L_x^{\text{box}}}{N_x^{\text{grid}}} = \frac{L_y^{\text{box}}}{N_y^{\text{grid}}} = 500$ μm, with $N_z^{\text{grid}} = 1$. In addition, time is discretized into windows of length $\delta t = 60$ min, which for a simulation over a time period $t^{\text{sim}} = 30$ days results in $t^{\text{sim}}/\delta t$ evenly sampled time points. Thus, the continuous functions $c_{\text{AI}}(\vec{x},t)$ and $c_{\text{cell}}(\vec{x},t)$ are each replaced by an $N_x^{\text{grid}} \times N_y^{\text{grid}} \times N_z^{\text{grid}} \times \delta t$ array of numbers in the x, y, z, and t coordinates, respectively. In the quasi-2D limit studied in this work, $N_z^{\text{grid}} = 1$, and thus we can refer to the discretized version of a function by $f(\vec{x},t) \to f[i_x,i_y,i_t]$, where i_x and i_y represent the mesh of points in the x and y directions, i_t represents a time point, and it is implicitly understood that we are fixed at the sole lattice point in the z direction.

The input cell density profile array is constructed as a superposition of the corresponding colony building-block arrays, as shown in Eq. (A4), with the colony centers chosen at random to avoid overlap and remain within confinement boundaries, as described previously. Five different replicates were run for each value of N_{colonies} and $L_{\text{confinement}}$, and the mean and standard error are calculated (note that in the limit of either a single colony or unconfined, infinitely separated colonies, there is no variation between different replicates, so the standard error is trivially zero). For explanatory purposes, it suffices to describe how to construct a single colony.

Each colony discretely builds up the density using units such that at time $t = 0$, one point has been added, at the mesh unit associated with the corresponding colony center. Then, the colony will grow over the time window $0 < t < t_{\max}$, via the sequential addition of discrete points, until $N_{\max}/N_0$ points have been added.

When building up a system consisting of N_{colonies} total colonies, in order to ensure a final total intensity of cell number $\frac{n_{\text{cells}}(t)}{N_{\text{colonies}}}$ in each individual colony, we build up any individual points in discrete intensity values of $\frac{n_0}{n_{\max}}\frac{n_{\text{cells}}(t)}{N_{\text{colonies}}}$.

With this information, then, for a colony, centered at a point (x_0,y_0), the coordinates of each successive point (x,y) are parametrized by a discrete Archimedean spiral,

$$(x[n],y[n]) = (x_0,y_0) + \frac{1}{N_{\text{colonies}}}\sqrt{n} \times (\cos(2\pi\sqrt{n}),\sin(2\pi\sqrt{n}))\ \mu\text{m}, \tag{A16}$$

where $n = 1,2,\ldots,\frac{n_{\max}}{n_0}$. This explicit form of the parametrization guarantees that upon completion of growth, the radius of the spiral disk is approximately equal to the desired maximum radius $R_{\max}$. This completes the necessary information required to construct a density profile for individual colonies, and by superposition, entire cellular communities.

4. Numerical tests of model and parameter robustness

As discussed earlier, nondimensionalization reduces the number of independent parameters to three: the basal rate r_b, the total number of cells $n_{\max}$, and the fold change f. Additionally, the major assumption of our model is the approximation of a smooth sigmoidal activation function with a Boolean Heaviside-step function. Specifically, a more realistic representation of the activation is given by a Hill function with Hill coefficient b:

$$r_{\text{Hill}}(c_{\text{AI}}) = r_b\left(1 + \frac{f}{1 + \left(\frac{c_{\text{AI}}(\vec{x},t)}{c_{\text{AI}_{\text{crit}}}}\right)^{-b}}\right). \tag{A17}$$

In the limit that $b \to \infty$, Eq. (A17) reduces to the numerically simpler approximation of Eq. (A2).

A change of cell number $n_{\max}$ will simply shift the number of colony divisions required for individual colonies to reach the LQS/GQS threshold. In particular, if with $n_{\max}$ cells, the critical colony number is N_{crit}, corresponding to a point where each colony has $\frac{n_{\max}}{N_{\text{crit}}}$ cells individually. Then, if we scale $n_{\max}$ by a factor a, individual colonies will now simply reach the critical cell number per colony at a scaled aN_{crit}.

To assess the robustness of our qualitative results to changes in the remaining parameters and the Heaviside approximation, we collect data for steady-state coverage in the limit of unconfined, infinitely separated colonies [Fig. 3(b) in the main text]. In Fig. 5, we compare the results generated with the default parameters used in the main text with changes in basal rate, fold change, and sharpness of the activation function. In all cases, while parameter variations shift transition points or sharpness of separation between activated and inactivated coverage levels, the existence of these transition points and separations remains intact, rendering the prediction of an LQS/GQS boundary insensitive to model parameter variations.

5. Analytical mean-field theory

In this section, we will derive an analytical mean-field theory to predict the LQS-GQS phase diagram, and compare these analytical results with the numerical simulations. Recall that, for long times, our time-independent, steady-state solutions will satisfy

$$(\gamma - D\nabla^2)c_{\text{AI}} - r(c_{\text{AI}})c_{\text{cell}}(\vec{x}) = 0, \tag{A18}$$

where D and γ are, respectively, the AI signal diffusivity and decay, and r is the local, AI-concentration-dependent signaling rate of the cells, given by

$$r(c_{\text{AI}}) = r_b + (r_a - r_b)\theta\left[c_{\text{AI}}(\vec{x}) - c_{\text{AI}_{\text{crit}}}\right]. \tag{A19}$$

For the system to be completely activated, we require that the concentration of c_{AI} at any point with a nonzero c_{cell} density be above the threshold. Note that we are keeping the total number of cells $n_{\text{cells}}^{\text{tot}}$ fixed, but dividing it out into ever more colonies N_{colonies}, such that there are $n_{\text{cells}}^{\text{tot}}/N_{\text{colonies}}$ cells per colony. And since we are keeping cell density of colony, $\bar{c}_{\text{colony}}$, fixed, the number of colonies N_{colonies} automatically sets the long-time radius of each colony

$$R_{\text{colony}} = \left(\frac{n_{\text{cells}}^{\text{tot}}/N_{\text{colonies}}}{\pi\,\bar{c}_{\text{colony}}}\right)^{1/2}. \tag{A20}$$

The remaining ambiguity concerns a choice of explicit parametrization of the long-term form of the steady-state cell density, $n_{\text{cell}}(\vec{x})$. In this work, we characterize n_{cell} completely with the above-mentioned parameters, plus one other: $L_{\text{confinement}}$, defined as the side length of a square region of space that the cells are allowed to occupy. Then, for a given boundary size $L_{\text{confinement}}$, given colony radius R_{colony}, and given total number of colonies N_{colonies}, we consider the ensemble of all possible nonoverlapping colony disk configurations that lie completely within the $L_{\text{confinement}}$ square region.

Thus, a given system realization is characterized by the following eight independent parameters: γ, D, r_b, $c_{\text{AI}_{\text{crit}}}$, $\bar{c}_{\text{colony}}$, R_{colony}, N_{colonies}, and $L_{\text{confinemenet}}$. The two related questions that I now pose are these (for given values of the first five parameters γ, D, r_b, $c_{\text{AI}_{\text{crit}}}$, $\bar{c}_{\text{colony}}$):

(1) What is the critical colony radius $R_{\text{LQS}}^{\text{crit}}$ beyond which an individual colony, with $R_{\text{colony}} > R_{\text{LQS}}^{\text{crit}}$, will *self-activate via LQS* in the limit of no neighboring influence?

(2) For subcritical colony radii $R_{\text{colony}} < R_{\text{LQS}}^{\text{crit}}$, how do the parameters N_{colonies} and $L_{\text{confinement}}$ set limits on the additional range of radii that allow the entire system to *group-activate via GQS*?

a. LQS threshold

We begin with setting a bound on the LQS regime by calculating the steady-state concentration at the center of an isolated colony with radius R_{colony} and cell density $\bar{c}_{\text{colony}}$. To start, let us just imagine that all cells are only producing at the basal rate, so that the governing equation becomes a simpler, linear equation

$$(\gamma - D\nabla^2)c_{\text{AI}} = r_b c_{\text{cell}}(\vec{x}). \tag{A21}$$

We assume that a single-cell point source at the origin can be modeled as a Dirac delta function,

$$c_{\text{cell}}^{\text{point}}(\vec{x}) = \delta(\vec{x}). \tag{A22}$$

Then, the exact solution in two dimensions is

$$c_{\text{AI}}^{\text{point}}(\vec{x}) = \frac{r_b}{2\pi D}K_0\left(\frac{|\vec{x}|}{\sqrt{D/\gamma}}\right), \tag{A23}$$

where K_0 is the zeroth-order Bessel function of the second kind. Thus, by the superposition principle, for an arbitrary cell density distribution n_{cell},

$$c_{\text{AI}}(\vec{x}) = \frac{r_b}{2\pi D}\int d^2\vec{x}'c_{\text{cell}}(\vec{x}')K_0\left(\frac{|\vec{x} - \vec{x}'|}{\sqrt{D/\gamma}}\right). \tag{A24}$$

Thus, for an isolated colony of uniform cell density $\bar{c}_{\text{colony}}$, centered at the origin, with a radius R_{colony}, the AI concentration at the center in the absence of any activation would be

$$c_{\text{AI}}^{\text{isolated}}(0) = \left(\frac{r_b\bar{c}_{\text{colony}}}{2\pi D}\right)(2\pi)\int_0^{R_{\text{colony}}} dr\, r K_0\left(\frac{r}{\sqrt{D/\gamma}}\right), \tag{A25}$$

which evaluates to

$$c_{\mathrm{AI}}^{\mathrm{isolated}}(0) = \frac{r_b \bar{c}_{\mathrm{colony}}}{\sqrt{D\gamma}} \left[\sqrt{\frac{D}{\gamma}} - R_{\mathrm{colony}} K_1 \left(\frac{R_{\mathrm{colony}}}{\sqrt{D/\gamma}} \right) \right]. \tag{A26}$$

Therefore, to identify the critical self-activation (via LQS) radius $R_{\mathrm{LQS}}^{\mathrm{crit}}$, we can set the above equation equal to the critical AI concentration $c_{\mathrm{AI}_{\mathrm{crit}}}$ and solve for the radius:

$$c_{\mathrm{AI}_{\mathrm{crit}}} = \frac{r_b \bar{c}_{\mathrm{colony}}}{\sqrt{D\gamma}} \left[\sqrt{\frac{D}{\gamma}} - R_{\mathrm{LQS}}^{\mathrm{crit}} K_1 \left(\frac{R_{\mathrm{LQS}}^{\mathrm{crit}}}{\sqrt{D/\gamma}} \right) \right]. \tag{A27}$$

More instructively, we can write this relation in the form

$$1 - \left(\frac{c_{\mathrm{AI}_{\mathrm{crit}}} \gamma}{r_b \bar{c}_{\mathrm{colony}}} \right) = \frac{R_{\mathrm{LQS}}^{\mathrm{crit}}}{\sqrt{D/\gamma}} K_1 \left(\frac{R_{\mathrm{LQS}}^{\mathrm{crit}}}{\sqrt{D/\gamma}} \right). \tag{A28}$$

The parameters we used in this work were $c_{\mathrm{AI}_{\mathrm{crit}}} = 100\,\mathrm{molecules}/\mu\mathrm{m}^3, \gamma = 0.01/\mathrm{h}, r_b = 500\,\mathrm{molecules/cell/h}$, $\bar{c}_{\mathrm{colony}} = 1$ cell$/\mu\mathrm{m}^3$, $D = 160$ $\mu\mathrm{m}^2/\mathrm{s}$, and $n_{\mathrm{cells}}^{\mathrm{tot}} = 7000 \times 10^3$ cells. Plugging in, we find $R_{\mathrm{LQS}}^{\mathrm{crit}} \approx 230$ μm, which yields $n_{\mathrm{LQS}}^{\mathrm{crit}} \approx 40$.

b. GQS boundary

Now, we turn to the question of limits on GQS activation in the case where $R_{\mathrm{colony}} < R_{\mathrm{LQS}}^{\mathrm{crit}}$. To answer this, we must add to the AI concentration at the center of an individual colony the contribution arising from the $N_{\mathrm{colonies}} - 1$ neighboring colonies. This contribution is distributed over a range of all possible values that arise from the ensemble of nonoverlapping disk configurations occupying the square region of size $L_{\mathrm{confinement}}$. For simplicity, let us in this mean-field theory assume that we can coarse-grain any neighbors into a continuous *metapopulation* cell density, equal to the average number of cells in the confined square area,

$$\bar{c}_{\mathrm{meta}} = \frac{n_{\mathrm{cells}}^{\mathrm{tot}}}{L_{\mathrm{confinement}}^2}. \tag{A29}$$

Let us also for simplicity assume that $L_{\mathrm{confinement}} \ggg R_{\mathrm{colony}}$, and that the Bessel source function decays rapidly enough such that we can effectively take $L_{\mathrm{confinement}} \approx \infty$ for the purposes of calculation.

Then, we must modify Eq. (A31) to include not just the single-colony contribution at the center $c_{\mathrm{AI}}(0)$, but also the neighbor contribution

$$c_{\mathrm{AI}}^{\mathrm{neighbors}}(0) \approx \left(\frac{r_b \bar{c}_{\mathrm{meta}}}{2\pi D} \right) (2\pi) \int_{R_{\mathrm{colony}}}^{\infty} dr\, r K_0 \left(\frac{r}{\sqrt{D/\gamma}} \right), \tag{A30}$$

which evaluates to

$$c_{\mathrm{AI}}^{\mathrm{neighbors}}(0) \approx \frac{r_b \bar{c}_{\mathrm{meta}}}{\sqrt{D\gamma}} \left[R_{\mathrm{colony}} K_1 \left(\frac{R_{\mathrm{colony}}}{\sqrt{D/\gamma}} \right) \right]. \tag{A31}$$

So now, if we modify our threshold condition, Eq. (A27), such that the sum $c_{\mathrm{AI}}^{\mathrm{isolated}} + c_{\mathrm{AI}}^{\mathrm{neighbors}}$ equals $c_{\mathrm{AI}_{\mathrm{crit}}}$, we get a set of solutions $(R_{\mathrm{crit}}, \bar{c}_{\mathrm{meta}_{\mathrm{crit}}})$ that satisfy

$$\begin{aligned} c_{\mathrm{AI}_{\mathrm{crit}}} = & \frac{r_b \bar{c}_{\mathrm{colony}}}{\sqrt{D\gamma}} \left[\sqrt{\frac{D}{\gamma}} - R^{\mathrm{crit}} K_1 \left(\frac{R^{\mathrm{crit}}}{\sqrt{D/\gamma}} \right) \right] \\ & + \frac{r_b \bar{c}_{\mathrm{meta}_{\mathrm{crit}}}}{\sqrt{D\gamma}} \left[R_{\mathrm{crit}} K_1 \left(\frac{R_{\mathrm{crit}}}{\sqrt{D/\gamma}} \right) \right]. \end{aligned} \tag{A32}$$

Solving for this critical set $(R_{\mathrm{crit}}, \bar{c}_{\mathrm{meta}_{\mathrm{crit}}})$, and converting the results to the equivalent in $(N_{\mathrm{crit}}, L_{\mathrm{confinement}_{\mathrm{crit}}})$ parameter space, we get the "phase diagram" shown in Fig. 6.

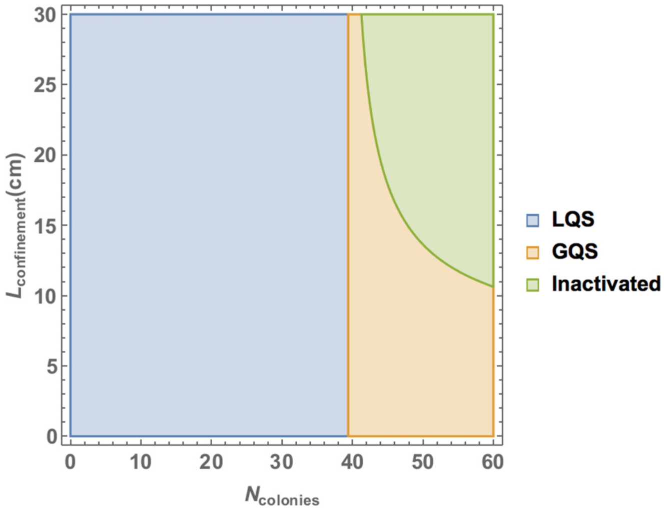

FIG. 6. Mean-field phase diagram calculated as described in this Appendix. The analytical results qualitatively reproduce the numerical results on the LQS-GQS transition, while the quantitative discrepancy may be attributed to the unrealistic treatment of the "near-field" source region, namely, approximating it as a Dirac delta point source.

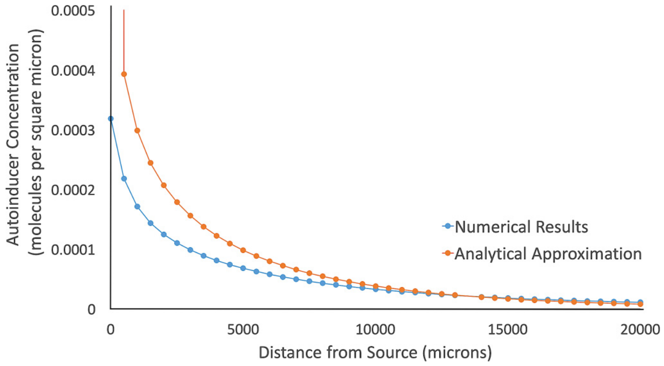

FIG. 7. Here, we compare the actual numerical results for the steady-state concentration due to a single cell with the results predicted from our analytical "point-source" Green's function approximation. The unrealistic approximation of the physics close to the cell (the "near-field" region) explains the quantitative discrepancy between the two curves, and by extension, the corresponding quantitative disagreement between our analytically predicted and numerically observed phase diagrams.

c. Discussion

As seen in the phase diagram, the essential qualitative results of the LQS-GQS transition are captured by the analytical mean-field model. However, the quantitative agreement between the analytical theory and numerical simulations is not as strong—in particular, the mean-field theory appears to overestimate the LQS boundary by a factor of 2, with the resulting GQS boundaries being similarly distorted. The origin of this discrepancy can be attributed to the simplifying approximation used in Eq. (A23), namely, treating a finite-radius cell as a localized Dirac delta function point source. In Fig. 7, we compare the actual numerical results for the steady-state concentration due to a single cell with the analytical prediction of the point-source Green's function. As is seen, the two curves agree well at very long distances from the source, or in the "far-field" region. However, the analytical approximation increasingly overestimates the concentration closer to the source, or in the "near-field" region. Thus, a tentative physical interpretation of our results would be that the qualitative system behavior is well described by a far-field analytical mean field theory, with "effective" phase boundaries arising from the effects of near-field corrections. A natural future extension of this research would be in developing more accurate analytical models that quantitatively explain how near-field effects renormalize the phase diagram boundaries.

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

[2] P. Williams, K. Winzer, W. C. Chan, and M. Cámara, Philos. Trans. R. Soc., B **362**, 1119 (2007).

[3] A. Jayaraman and T. K. Wood, Ann. Rev. Biomed. Eng. **10**, 145 (2008).

[4] H. Youk and W. A. Lim, Science **343**, 1242782 (2014).

[5] M. E. Hibbing, C. Fuqua, M. R. Parsek, and S. B. Peterson, Nat. Rev. Microbiol. **8**, 15 (2010).

[6] C. Ratzke and J. Gore, Nat. Microbiol. **1**, 16022 (2016).

[7] R. J. Redfield, Trends Microbiol. **10**, 365 (2002).

[8] B. A. Hense, C. Kuttler, J. Muller, M. Rothballer, A. Hartmann, and J.-U. Kreft, Nat. Rev. Microbiol. **5**, 230 (2007).

[9] P. D. Prez, J. T. Weiss, and S. J. Hagen, BMC Syst. Biol. **5**, 1 (2011).

[10] J. Q. Boedicker, M. E. Vincent, and R. F. Ismagilov, Angew. Chem., Int. Ed. Engl. **48**, 5908 (2009).

[11] S. A. West, K. Winzer, A. Gardner, and S. P. Diggle, Trends Microbiol. **20**, 586 (2012).

[12] M. Wu, J. W. Roberts, S. Kim, D. L. Koch, and M. P. DeLisa, Appl. Environ. Microbiol. **72**, 4987 (2006).

[13] N. Goldenfeld and C. Woese, Ann. Rev. Condens. Matter Phys. **2**, 375 (2011).

[14] E. Kussell and M. Vucelja, Rep. Prog. Phys. **77**, 102602 (2014).

[15] T. Mora and W. Bialek, J. Stat. Phys. **144**, 268 (2011).

[16] H. Celiker and J. Gore, Nat. Commun. **5**, 4643 (2014).

[17] C. J. Kastrup, J. Q. Boedicker, A. P. Pomerantsev, M. Moayeri, Y. Bian, R. R. Pompano, T. R. Kline, P. Sylvestre, F. Shen, S. H. Leppia, W. J. Tang, and R. F. Ismagilov, Nat. Chem. Biol. **4**, 742 (2008).

[18] K. S. Korolev, Phys. Rev. Lett. **115**, 208104 (2015).

[19] F. Ginelli, F. Peruani, M.-H. Pillot, H. Chate, G. Theraulaz, and R. Bon, Proc. Natl. Acad. Sci. USA **112**, 12729 (2015).

[20] S. Payne and L. You, Adv. Biochem. Eng./Biotechnol. **146**, 97 (2013).

[21] P. Mehta, S. Goyal, B. L. Bassler, and N. S. Wingreen, Mol. Syst. Biol. **5**, 325 (2009).

[22] T. Long, K. C. Tu, Y. Wang, P. Mehta, N. P. Ong, B. L. Bassler, and N. S. Wingreen, PLoS Biol. **7**, e1000068 (2009).

[23] G. Dilanji, J. Langebrake, P. De Leenheer, and S. J. Hagen, J. Am. Chem. Soc. **134**, 5618 (2012).

[24] S. Huang, J. K. Srimani, A. J. Lee, Y. Zhang, A. J. Lopatkin, K. W. Leong, and L. You, Biomaterials **61**, 239 (2015).

[25] A. Tamsir, J. J. Tabor, and C. A. Voigt, Nature (London) **469**, 212 (2010).

[26] J. R. Chandler, S. Heilmann, J. E. Mittler, and P. Greenberg, ISME J. **6**, 2219 (2012).

[27] B. O. Emerenini, B. A. Hense, C. Kuttler, and H. J. Eberl, PLoS ONE **10**, e0132385 (2015).

[28] J. B. Langebrake, G. E. Dilanji, S. J. Hagen, and P. De Leenheer, J. Theor. Biol. **363**, 53 (2014).

[29] E. O. Powell, J. Gen. Microbiol. **15**, 492 (1956).

[30] J. Ferkinghoff-Borg and T. Sams, Mol. BioSyst. **10**, 103 (2014).

[31] T. E. Gorochowski, A. Matyjaszkiewicz, T. Todd, N. Oak, K. Kowalska, S. Reid, K. T. Tsaneva-Atanasova, N. J. Savery, C. S. Grierson, and M. di Bernardo, PLoS ONE **7**, e42790 (2012).

[32] E. Grochowski, R. Ronen, J. Shen, and H. Wang, Proc. Int. Conf. Comput. Des. **2004**, 236 (2004).