

2019-12

# Statistics of correlated percolation in a bacterial community

*This work was made openly accessible by BU Faculty. Please [share](#) how this access benefits you.  
Your story matters.*

---

Version	Published version
Citation (published version):	Xiaoling Zhai, Joseph W Larkin, Kaito Kikuchi, Samuel E Redford, Ushasi Roy, Gürol M Süel, Andrew Mugler. 2019. "Statistics of correlated percolation in a bacterial community.." PLoS Comput Biol, Volume 15, Issue 12, pp. e1007508 <a href="https://doi.org/10.1371/journal.pcbi.1007508">https://doi.org/10.1371/journal.pcbi.1007508</a>

<https://hdl.handle.net/2144/40455>

*Boston University*

RESEARCH ARTICLE

# Statistics of correlated percolation in a bacterial community

Xiaoling Zhai<sup>1</sup>, Joseph W. Larkin<sup>2\*</sup>, Kaito Kikuchi<sup>2</sup>, Samuel E. Redford<sup>2</sup>, Ushasi Roy<sup>1</sup>, Gürol M. Süel<sup>2,3</sup>, Andrew Mugler<sup>1\*</sup>

**1** Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana, United States of America, **2** Division of Biological Sciences, University of California San Diego, La Jolla, California, United States of America, **3** San Diego Center for Systems Biology, University of California San Diego, La Jolla, California, United States of America

\* Current Address: Departments of Biology and Physics, Boston University, Boston, Massachusetts, United States of America

\* [amugler@purdue.edu](mailto:amugler@purdue.edu)



**OPEN ACCESS**

**Citation:** Zhai X, Larkin JW, Kikuchi K, Redford SE, Roy U, Süel GM, et al. (2019) Statistics of correlated percolation in a bacterial community. *PLoS Comput Biol* 15(12): e1007508. <https://doi.org/10.1371/journal.pcbi.1007508>

**Editor:** Ned S. Wingreen, Princeton University, UNITED STATES

**Received:** June 14, 2019

**Accepted:** October 22, 2019

**Published:** December 2, 2019

**Copyright:** © 2019 Zhai et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data and code are available at [https://github.com/Xiaoling1992/correlated\\_percolation.git](https://github.com/Xiaoling1992/correlated_percolation.git).

**Funding:** This work was supported by the San Diego Center for Systems Biology (NIH P50 GM085764, G.M.S.), the Howard Hughes Medical Institute-Simons Foundation Faculty Scholars program (G.M.S.), the National Institute of General Medical Sciences (R01 GM121888, G.M.S. and A.M.) and the Simons Foundation Mathematical Modeling of Living Systems Program (376198, A.M.). The funders had no role in study design, data

## Abstract

Signal propagation over long distances is a ubiquitous feature of multicellular communities, but cell-to-cell variability can cause propagation to be highly heterogeneous. Simple models of signal propagation in heterogeneous media, such as percolation theory, can potentially provide a quantitative understanding of these processes, but it is unclear whether these simple models properly capture the complexities of multicellular systems. We recently discovered that in biofilms of the bacterium *Bacillus subtilis*, the propagation of an electrical signal is statistically consistent with percolation theory, and yet it is reasonable to suspect that key features of this system go beyond the simple assumptions of basic percolation theory. Indeed, we find here that the probability for a cell to signal is not independent from other cells as assumed in percolation theory, but instead is correlated with its nearby neighbors. We develop a mechanistic model, in which correlated signaling emerges from cell division, phenotypic inheritance, and cell displacement, that reproduces the experimentally observed correlations. We find that the correlations do not significantly affect the spatial statistics, which we rationalize using a renormalization argument. Moreover, the fraction of signaling cells is not constant in space, as assumed in percolation theory, but instead varies within and across biofilms. We find that this feature lowers the fraction of signaling cells at which one observes the characteristic power-law statistics of cluster sizes, consistent with our experimental results. We validate the model using a mutant biofilm whose signaling probability decays along the propagation direction. Our results reveal key statistical features of a correlated signaling process in a multicellular community. More broadly, our results identify extensions to percolation theory that do or do not alter its predictions and may be more appropriate for biological systems.

## Author summary

Many multicellular systems send signals over long distances by relaying information over connected cell-to-cell paths. In physics, the statistics of connected path formation are

collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

described by percolation theory. We previously discovered that the statistics of electrical signal propagation in communities of the bacterium *Bacillus subtilis* are consistent with the predictions of percolation theory. However, we find experimentally that key features of this system go beyond the simple assumptions of basic percolation theory, which include site-to-site independence and spatial uniformity of the signaling probability. Why are the predictions of percolation theory still upheld? Using a computational model, we find that the cell-to-cell dependence does not change the predictions due to the universal nature of percolation theory near its critical point, and the spatial variability of the signaling probability actually expands the parameter range over which the predictions hold. We validate our findings using a mutant bacterial strain. Our work explores the robustness of percolation theory to its underlying assumptions, and the resulting consequences for long-range bacterial signaling.

## Introduction

Long-range signal transmission is central to the function of many multicellular communities. However, cell-to-cell variability within these communities [1, 2] can cause some cells not to participate in signaling, which may degrade or attenuate the signal [3–5]. In physics, signal transmission in the presence of non-propagating agents is the domain of percolation theory [6]. As a result, many investigators have turned to percolation theory to describe signal transmission in multicellular systems. In bacterial communities, percolation theory has been used to predict the scaling laws that result from signal disruption during quorum sensing [7]. In neuroscience, percolation theory has been used to describe (i) the transition from a fully connected to a disconnected electrical network in rat hippocampus cultures [8, 9], (ii) the spatiotemporal structure of viral propagation within astrocyte monolayers [10], and (iii) the transition from conscious to unconscious brain activities during general anesthesia [11]. In pancreatic islets, percolation theory has been used to understand the dependence of calcium wave propagation on the coupling strength of gap junctions between the islet cells [12]. In colonies of *Spirostomum* (an aquatic worm-like cell), percolation theory was recently shown to describe how the propagation of a hydrodynamic cell-to-cell trigger-wave depends on the colony density [13].

We recently demonstrated that the transmission of an electrical signal from the interior to the periphery of a biofilm of *Bacillus subtilis* bacteria is consistent with the predictions of percolation theory [5]. In this system, starvation of the interior cells causes release of intracellular potassium, which leads to depolarization and potassium release in neighboring cells, resulting in a cell-to-cell relay wave that propagates to the biofilm periphery [14–16]. The signal temporarily prevents peripheral cells from taking up nutrients and thus allows nutrients to diffuse to the interior cells, preserving biofilm viability and increasing its overall fitness [14]. However, it turns out that not all cells participate in the potassium release: we discovered that the fraction of participating cells is near the percolation threshold, and that clusters of participating cells have a size distribution that follows a power law with an exponent predicted by percolation theory [5]. Operating near the percolation threshold allows the biofilm to maintain successful signal transmission while minimizing the number of cells that undergo the costly potassium release [5].

Despite the success of percolation theory as a description of signal transmission within this system, it is reasonable to suspect that several key assumptions of percolation theory may require scrutiny in this and many similar multicellular systems [5]. First, percolation theory assumes that the probability for each cell to participate in signal transmission is independent

of other cells. However, in reality it may be that the participation probability of a cell is correlated with that of its neighbors. For example, if the molecular mechanism governing participation is heritable, then one expects the participation of a given cell to be correlated with other cells in its lineage, which are most likely to be nearby in the densely packed biofilm. Second, percolation theory assumes that the participation probability does not vary from one biofilm to another, or from location to location within a biofilm. However, in reality we know that there is variability across biofilms, and particular mutant strains have spatial variability in the participation fraction [5]. These considerations raise the question of when and how percolation theory remains a predictive description of signal transmission in biological systems. Conversely, they suggest a strategy by which deviations from percolation theory would give important insights about the ways in which a biological system differs from the model assumptions [17]. They also raise the broader question of which predictions of a model from statistical physics are dependent on the model details, and which predictions are universal.

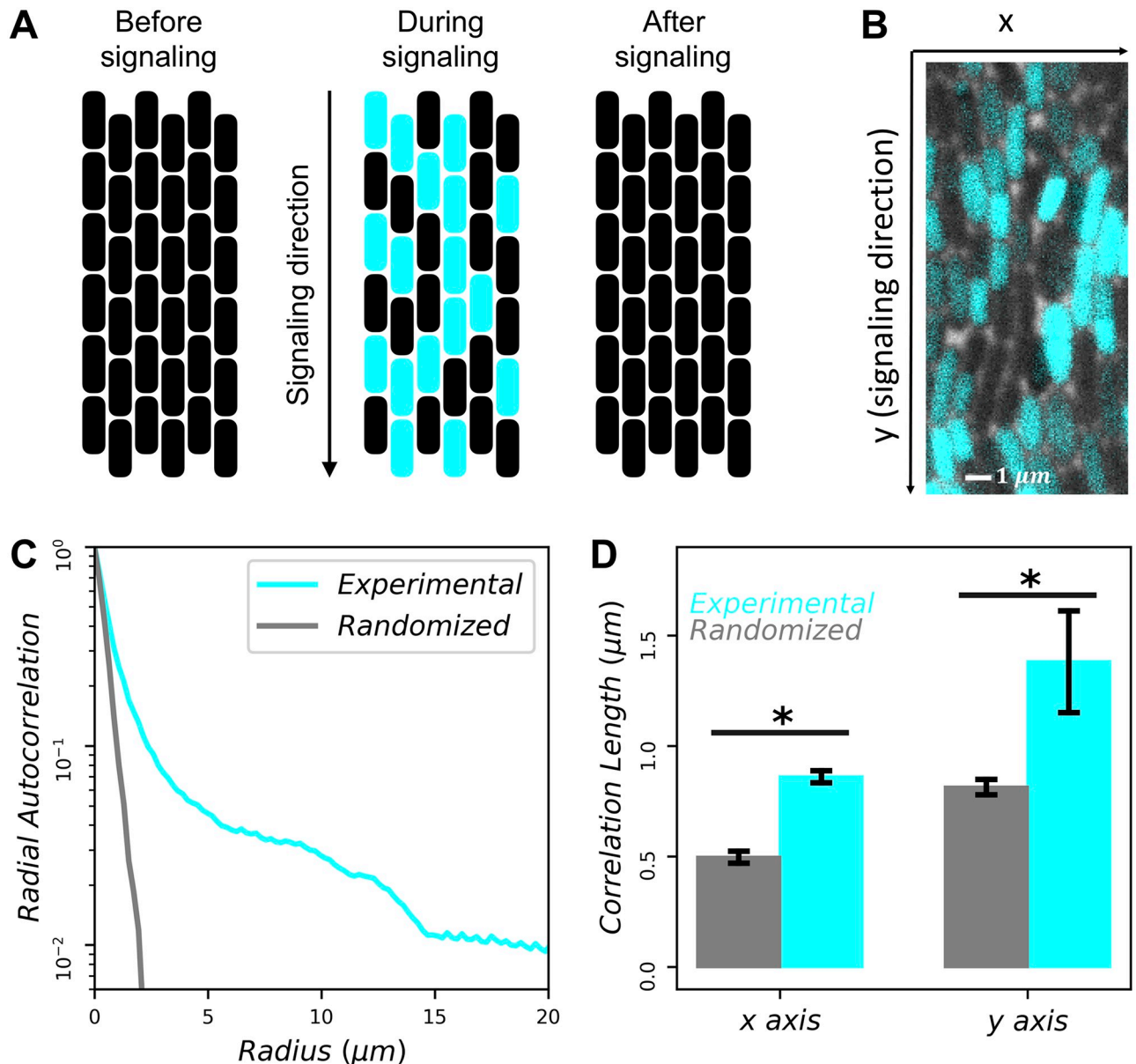
Here we use a combination of simulations and experiments to investigate the statistical properties of signal percolation in a bacterial biofilm. We find that signal correlations exist between cells, due to a combination of phenotypic inheritance and spatial proximity of a cell to its progeny. We find that while these correlations lower the percolation threshold, they are not sufficiently long-range to affect the cluster size statistics. Instead, we find that variability in the signaling fraction within and across biofilms affects the statistics by widening the range of fractions at which one observes the power-law distribution of cluster sizes. We validate our findings using a mutant biofilm whose participation fraction decays as a function of propagation distance. Our results demonstrate that certain community-level signaling properties are robust to cell-level features whereas others are not, and we discuss the implications for biofilm function.

## Results

We first review the key features of electrical signaling in the biofilm [5, 14–16], and those of percolation theory, as these features will motivate our present results. The electrical signal is transmitted by cells across the biofilm in a wave-like manner (Fig 1A). We measure the membrane potential of cells during the peak of signal transmission using a fluorescent dye (cyan in Fig 1B; see Materials and methods). We previously observed a bimodal distribution of dye intensity across cells [5], which provides a threshold above or below which we define cells as “on” (participating in the signal) or “off” (not participating in the signal), respectively. This observation motivates our use of percolation theory, as percolation theory describes the connectivity and spatial statistics of systems on a lattice in which each cell has a probability  $\phi$  to be on.

We alert the reader that in typical applications of percolation theory, one can measure both the input (the ability of each component to signal or not) and the output (whether or not each component actually participates in the signal as it propagates). Here, because we do not know the molecular mechanism that confers the ability to signal, we can only measure the output. Nonetheless, we observe in [5] that (i) isolated clusters participate in signaling, and (ii) the percolation threshold remains predictive of whether the signal propagates across the biofilm. Therefore, as in [5], we conclude that the signaling mechanism is sufficiently short-range that percolation is a relevant criterion for propagation, but sufficiently long-range that the output can be treated as a reasonably good proxy for the input.

Our experiments focus on a 2D cell monolayer at the edge of the biofilm (see Materials and methods). We previously found that cells are most likely to have six neighbors [5]. For an infinite 2D, six-neighbor lattice, percolation theory predicts that (i) a connected path of on-cells



**Fig 1. Signaling probability of each cell is correlated with neighboring cells.** (A) Cartoon illustrating electrical signaling wave transmitted across biofilm. Cyan represents cells that participate in signaling. (B) Zoomed-in snapshot of cells in biofilm during peak of signal transmission (actual experimental window is approximately 35 cells tall by 230 cells wide). Cyan indicates fluorescence intensity of ThT dye, proportional to membrane potential. (C) Correlation function is longer-range than that from randomized data ( $N = 3$  biofilms). (D) Correlations are significantly longer than random both perpendicular ( $x$ ) and parallel ( $y$ ) to the signaling direction ( $p < 0.001$  and  $p = 0.007$  assuming Gaussian errors, respectively).

<https://doi.org/10.1371/journal.pcbi.1007508.g001>

emerges above the critical value  $\phi_c = 1/2$ , and that (ii) at  $\phi_c$ , the distribution of on-cell cluster sizes  $P(n)$  becomes a power law [6].

In the experiments, we image a finite window of approximately 35 by 230 cells (see [Materials and methods](#)). Finite-size effects can change the value of  $\phi_c$  at which connectivity sets in, which we call  $\phi_c^{\text{conn}}$  [6]. Indeed, simulations predict that  $\phi_c^{\text{conn}} = 0.45$  in this finite geometry [5]. Finite-size effects should not change the value of  $\phi_c$  at which  $P(n)$  becomes a power law, which we call  $\phi_c^{\text{pow}} = 1/2$ , so long as  $\sqrt{n}$  is sufficiently below the smaller lattice dimension.

However, at larger  $n$  values the distribution will deviate from a power law, even at  $\phi_c^{\text{pow}}$ , due to finite-size effects.

We previously observed that the fraction of on-cells in the experiments is  $\phi = 0.43 \pm 0.02$  (mean  $\pm$  standard error), and that the distribution  $P(n)$  of on-cell cluster sizes is a power law over three decades [5]. The fact that  $\phi \approx \phi_c^{\text{conn}}$  suggests that the system sits at the connectivity threshold. However, the fact that  $\phi < \phi_c^{\text{pow}}$  raises the question of why a power law is observed, particularly one with no apparent finite-size effects at large  $n$ . To address this question, as well as the broader question of what features of percolating systems are expected to be robust to the underlying assumptions about the components, we now investigate the effects of signal correlations and of variability in the signaling fraction.

### Participation in signaling is spatially correlated

Percolation theory assumes that a fraction  $\phi$  of on-cells are situated randomly in space. However, in the biofilm one might expect that on-cells are spatially co-located, for example if participating in the signal is a heritable phenotype. To determine whether there are spatial correlations in on-cells, we measure the radial autocorrelation function

$$C(r) = \langle s_i s_j \rangle_r - \phi^2, \tag{1}$$

where  $s = 1$  for on-cells,  $s = 0$  for off-cells, and the average is taken over all pixels  $i$  and  $j$  whose separation is  $r$  (see [Materials and methods](#)). We find that  $C(r)$  is a decreasing function of  $r$ , as expected ([Fig 1C](#), cyan curve). We then compare  $C(r)$  to the autocorrelation function computed with the locations of on-cells randomized. Specifically, we retain the locations of all cells and the number of on-cells, but we randomize which cells are on (as would be the case in percolation theory). We see in [Fig 1C](#) that  $C(r)$  falls off more steeply in this case (gray curve). These results suggest that on-cells are more spatially correlated than expected from random placement.

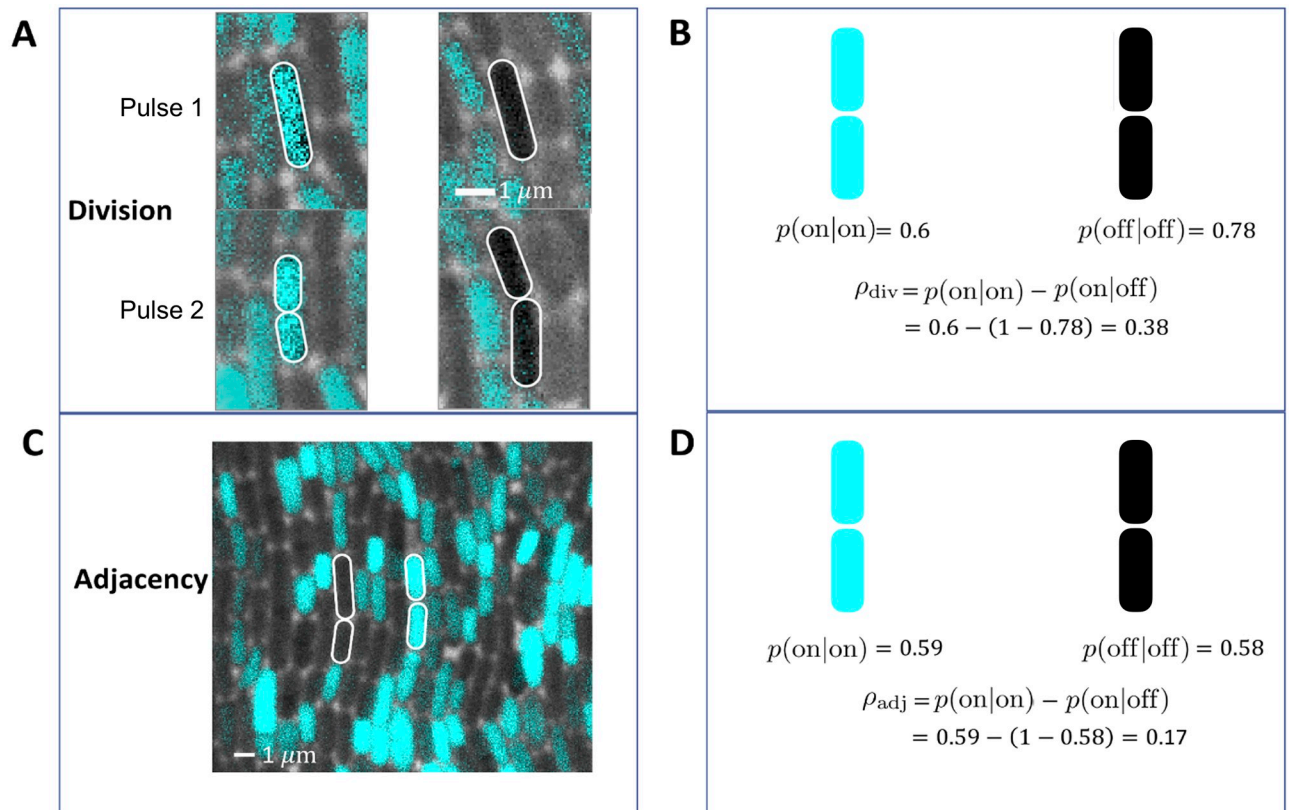
We next investigate the strength of correlation perpendicular ( $x$ ) and parallel ( $y$ ) to the direction of signal transmission ([Fig 1B](#)). We define the correlation lengths as  $\xi_x = \int dx C(x)$  and  $\xi_y = \int dy C(y)$ , where  $C(x)$  and  $C(y)$  are defined as in [Eq 1](#) but restricted to separations perpendicular ( $x$ ) or parallel ( $y$ ) to the signaling direction, and the integrals run from zero to the maximal separation values. Even in the randomized data, we see that the correlation length is larger in the  $y$  direction than in the  $x$  direction (compare the gray bars in [Fig 1D](#)) because cells are longer than they are wide, and the long axis of each cell is generally oriented in the signaling direction ([Fig 1B](#)). In the actual (non-randomized) data, the correlation lengths are 70% larger than random in both the  $x$  and  $y$  directions, and both differences are significant ( $p < 0.01$ ; [Fig 1D](#)). These results suggest that on-cells are significantly correlated both parallel and perpendicular to the signaling direction.

To quantify the correlation at the single-cell level, we consider the conditional probabilities  $p(\text{on}|\text{on})$  and  $p(\text{off}|\text{off})$ , where  $p(\text{on}|\text{on})$  is the probability that a cell is on given that the cell above it is also on, and similarly for  $p(\text{off}|\text{off})$ . We then calculate the order parameter

$$\rho = p(\text{on}|\text{on}) - p(\text{on}|\text{off}), \tag{2}$$

where  $p(\text{on}|\text{off}) = 1 - p(\text{off}|\text{off})$ . With no correlation, we have  $p(\text{on}|\text{on}) = p(\text{on}|\text{off}) = \phi$ , and therefore  $\rho = 0$ . With perfect correlation, we have  $p(\text{on}|\text{on}) = 1$  and  $p(\text{on}|\text{off}) = 0$ , and therefore  $\rho = 1$ . Thus,  $\rho$  quantifies the cell-to-cell correlation in the signaling direction on a scale from zero to one.

We estimate the conditional probabilities, and thus  $\rho$ , in two ways ([Fig 2](#)). First, because cell division is usually parallel to the signaling direction, we track division events that occur in



**Fig 2. Order parameter  $\rho$  quantifies degree of spatial correlations.** (A, B) Lineage-tracing experiments yield  $\rho_{\text{div}} = 0.38$  ( $N = 49$  division events). (C, D) Spatial analysis of the biofilm images yield  $\rho_{\text{adj}} = 0.17$  ( $N = 51$  cell pairs).

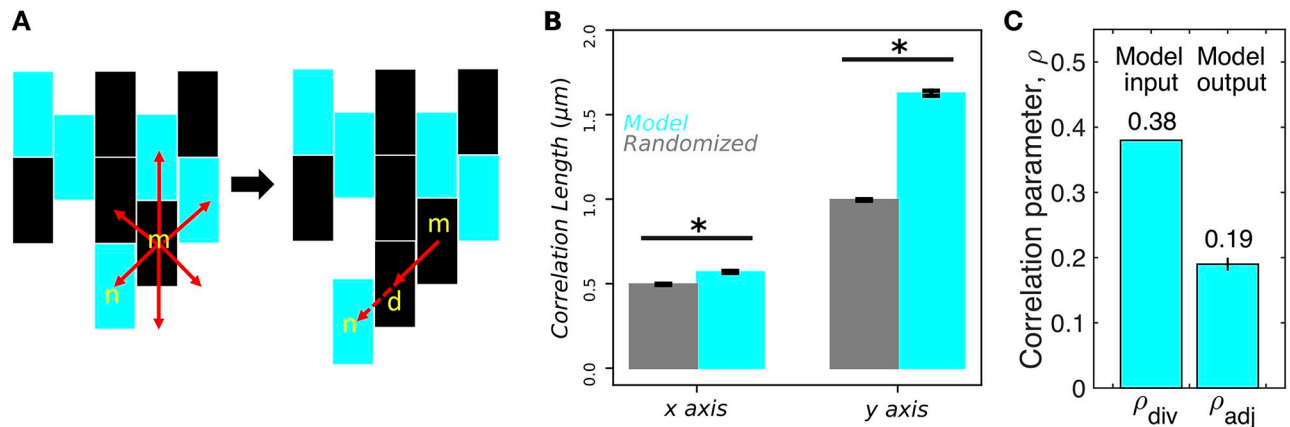
<https://doi.org/10.1371/journal.pcbi.1007508.g002>

between signal pulses (Fig 2A; see Materials and methods). We then count the number of times that the top cell has the same or different signaling state as the bottom cell. From this method we obtain  $\rho_{\text{div}} = 0.38$  (Fig 2B). Second, we estimate the conditional probabilities directly from pairs of cells that are adjacent to each other in the signaling direction during signaling (Fig 2C; see Materials and methods). From this method we obtain  $\rho_{\text{adj}} = 0.17$  (Fig 2D). These results confirm at the single-cell level that spatial correlations exist in the signaling direction ( $\rho_{\text{adj}} > 0$ ) but suggest that these correlations are less strong than those produced directly by division ( $\rho_{\text{adj}} < \rho_{\text{div}}$ ).

### Mechanistic model of correlated signaling

To understand the experimental results above, we propose a mechanistic model of spatially correlated cell signaling. We hypothesize that the signaling state is heritable during cell division with a certain probability, and that cell displacement can occur at the leading edge as the biofilm grows. The assumption that cells possess a signaling state variable is supported by the observation that a cell generally does not switch its on/off signaling behavior between successive pulses in the experiments [5].

Specifically, as shown in Fig 3A, we generate a 2D, six-neighbor lattice of rectangular cells with aspect ratio 2 (the approximate experimental value) in the following way. Each cell divides after a time  $\tau$  drawn from a Gaussian distribution with mean  $\bar{\tau}$  and standard deviation  $\delta\tau$ . The “mother” cell (m) retains its location and signaling state, while the “daughter” cell (d) occupies one of the eligible neighboring locations with equal probability. Eligibility requires that the



**Fig 3. Mechanistic model of correlated signaling captures experimental features.** (A) Mother cell (m) produces daughter cell (d) with correlated signaling state at any neighboring site at which a maximum of one neighbor cell (n) is displaced. Cyan indicates that cell has the ability to signal. (B) Correlations are significantly longer than random both perpendicular (x) and parallel (y) to the signaling direction ( $N = 10^4$  lattices;  $p < 0.001$  for both assuming Gaussian errors). Compare to experiments in Fig 1D. (C) Stochasticity in division times, neighbor selection, and cell displacement reduces correlation parameter from  $\rho_{div} = 0.38$  to  $\rho_{adj} = 0.19 \pm 0.01$ , close to experimentally measured  $\rho_{adj} = 0.17$  ( $N = 10^4$  lattices).

<https://doi.org/10.1371/journal.pcbi.1007508.g003>

neighboring location either be empty or be occupied by a neighboring cell (n) that, when displaced by the division along the same direction, would occupy an empty location (Fig 3A). Because the biofilm is growing downward, the eligible locations will most often be the location directly below and, with lower probability, the locations below-and-to-the-right and below-and-to-the-left. The signaling state of the daughter, given that of the mother, is determined from the division parameter  $\rho_{div}$  and the fraction of on-cells  $\phi$  according to

$$p(\text{on}|\text{on}) = \phi + \rho_{div} - \phi\rho_{div}, \quad (3)$$

$$p(\text{on}|\text{off}) = \phi - \phi\rho_{div}, \quad (4)$$

which follow from Eq 2 and the requirement that the fraction of on-cells remains  $\phi$  throughout the process (see Materials and methods). We produce a 100 by 230 lattice of cells by initializing the top row randomly and generating the next 99 rows according to the above mechanism. Then we remove the top 55 and bottom 10 rows, leaving a 35 by 230 cell window as in the experiments. This procedure allows the mechanism to achieve statistical steady state and focuses on the biofilm edge as in the experiments.

We find that the spatial statistics are not sensitive to the value of  $\delta\tau/\bar{\tau}$ , so long as it is greater than zero, and therefore we average our results over the range  $0 < \delta\tau/\bar{\tau} < 1$  (rejecting samples with  $\tau \leq 0$  for large  $\delta\tau$ ). We also find that allowing neighbor cell displacement is necessary to generate correlations in the x direction, but that allowing two or more levels of displacement does not qualitatively change the results. Thus, the only parameters in the model are  $\phi$  and  $\rho_{div}$  which we set from the experiments as  $\phi = 0.43$  [5] and  $\rho_{div} = 0.38$  (Fig 2B).

This model, with no free parameters, makes three predictions. Specifically, the model predicts that (i) the correlation length in the x direction is significantly different from random (Fig 3B), (ii) the correlation length in the y direction is significantly different from random (Fig 3B), and (iii) the spatial correlation parameter measured from adjacent cells in the y direction after the biofilm is generated is  $\rho_{adj} = 0.19 \pm 0.01$  (Fig 3C). The model output  $\rho_{div}$  is reduced from the model input  $\rho_{div} = 0.38$  due to the stochasticity in division times, neighbor selection, and cell displacement. Predictions (i) and (ii) are consistent with the experiments, as both the x and y correlation lengths were found to be significantly different than random

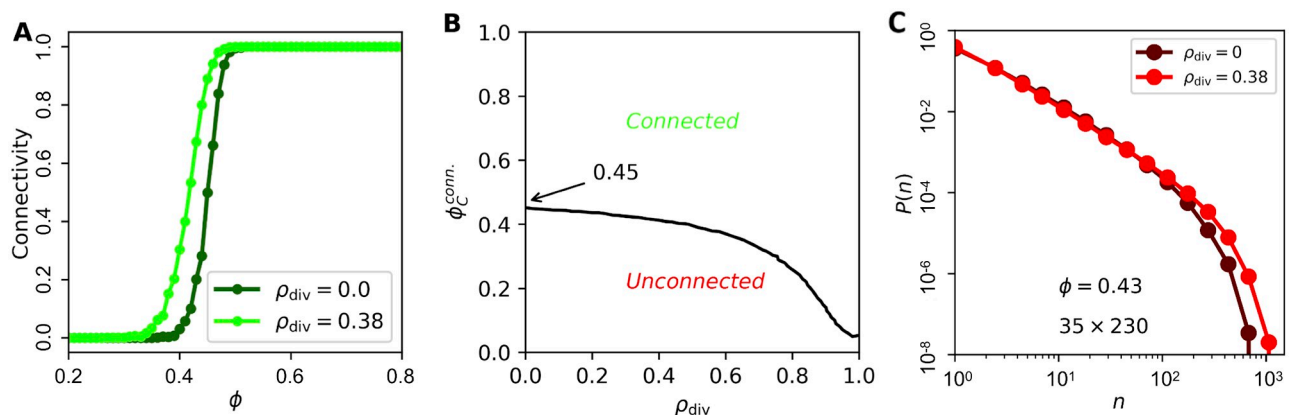


(Fig 1D). Prediction (iii) is also consistent with the experiments, as  $\rho_{\text{adj}}$  was measured to be 0.17 (Fig 2D), which is very close to  $0.19 \pm 0.01$ . We have also checked that these predictions remain unchanged when accounting for the fact that on-cells grow more slowly than off-cells [5] (see Materials and methods). The fact that all three predictions are validated by the experiments gives us confidence that the model captures the basic underlying mechanism, especially because it has no free parameters.

### Impact of correlations on spatial statistics

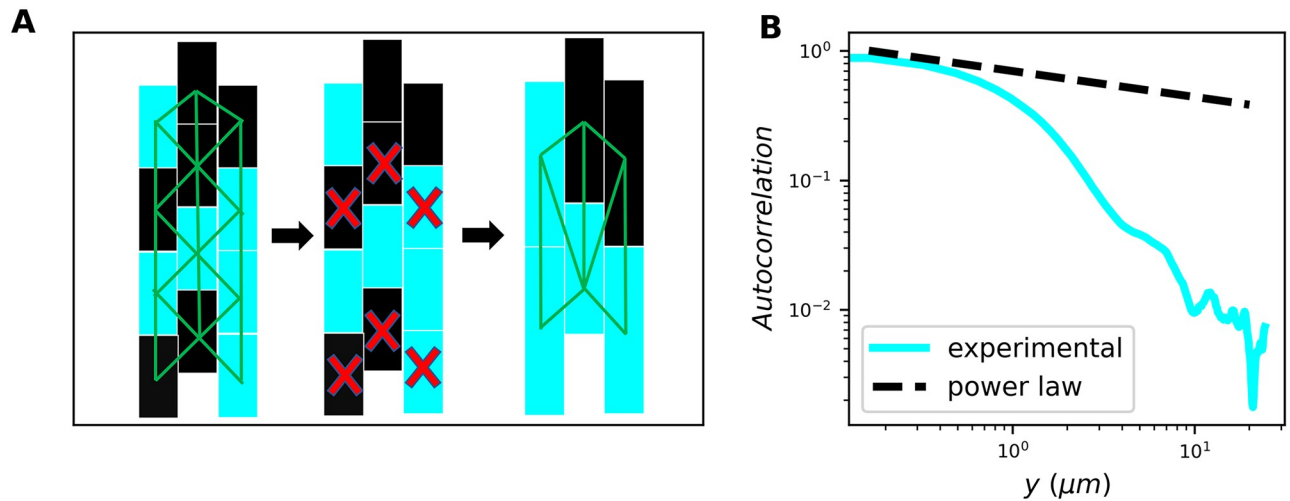
We now use our mechanistic model to investigate the impact of the spatial correlations on the statistical properties of the biofilm. First we focus on the connectivity: the probability, over an ensemble of simulated biofilms, that a connected path of on-cells exists from the top to the bottom of the lattice. The connectivity is expected to show a sharp transition from 0 to 1 at a critical fraction of on-cells  $\phi_c^{\text{conn}}$ . For an infinite lattice (in 2D with six neighbors),  $\phi_c^{\text{conn}} = 1/2$  [6]. Finite-size effects reduce the sharpness, but  $\phi_c^{\text{conn}}$  can still be defined as the value of  $\phi$  for which the connectivity is 50%. For a finite lattice of the approximate size of the experimental window (35 cells tall by 230 cells wide), without correlations, we previously found  $\phi_c^{\text{conn}} = 0.45$  [5] (Fig 4A, dark green curve). With correlations, using our mechanistic model with  $\rho_{\text{div}} = 0.38$ , we find  $\phi_c^{\text{conn}} = 0.4$  (Fig 4A, light green curve). More generally, the connectivity threshold is shown as a function of  $\rho_{\text{div}}$  in Fig 4B, and we see that as  $\rho_{\text{div}} \rightarrow 1$ ,  $\phi_c^{\text{conn}}$  becomes close to zero, even with the stochasticity inherent in the model. Thus, spatial correlations reduce the connectivity threshold. This makes sense, as correlations increase the probability of connected on-cells, particularly in the signaling direction, and this lowers the fraction of on-cells needed to create a connected path.

Second, we investigate the impact of correlations on the distribution of on-cell cluster sizes  $P(n)$ . The distribution is expected to become a power law at a critical fraction of on-cells  $\phi_c^{\text{pow}} = 1/2$  [6]. The experimental fraction of on-cells is  $\phi = 0.43 \pm 0.02$  [5], which is lower than  $\phi_c^{\text{pow}}$ . In simulations without correlations, at  $\phi = 0.43$ , we find that  $P(n)$  acquires a rolloff (when viewed on a log-log scale) at large  $n$  (Fig 4C, dark red curve). The rolloff indicates that the distribution is becoming more exponential, as expected for  $\phi < \phi_c^{\text{pow}}$ . However, in experiments, we find that  $P(n)$  maintains the power law dependence, with no rolloff, for three decades, i.e. out to  $n = 10^3$  [5]. Because we have seen that spatial correlations preserve connectivity at lower



**Fig 4. Spatial correlations increase connectivity but have little effect on cluster size distribution.** (A) Connectivity, defined as probability that a connected path of on-cells exists, occurs at lower on-cell fraction  $\phi$  as correlation parameter  $\rho_{\text{div}}$  increases ( $N = 10^3$  lattices). (B) Connectivity threshold  $\phi_c^{\text{conn}}$ , defined as  $\phi$  value for which connectivity is 50%, decreases with  $\rho_{\text{div}}$  ( $N = 10^3$  lattices). (C) Spatial correlations ( $\rho_{\text{div}} = 0.38$ ) have little effect on distribution, in particular not removing exponential rolloff at large  $n$  ( $N = 10^3$  lattices).

<https://doi.org/10.1371/journal.pcbi.1007508.g004>



**Fig 5. Short-range correlations do not affect critical properties.** (A) Illustration of the renormalization argument: upon site decimation, lattice remains triangular,  $\phi$  remains constant, and  $\rho$  vanishes. (B) Correlation function in experiments is short-range, i.e. sub-power-law ( $N = 3$  biofilms).

<https://doi.org/10.1371/journal.pcbi.1007508.g005>

$\phi$  (Fig 4A and 4B), we hypothesize that correlations may also preserve the power law dependence of  $P(n)$  at lower  $\phi$ , and thus explain the experimental observation. Surprisingly, using our mechanistic model, we find that the spatial correlations actually have little impact on  $P(n)$  (Fig 4C, light red curve): the rolloff is slightly shifted to larger  $n$ , but it is certainly still present over the three-decade range.

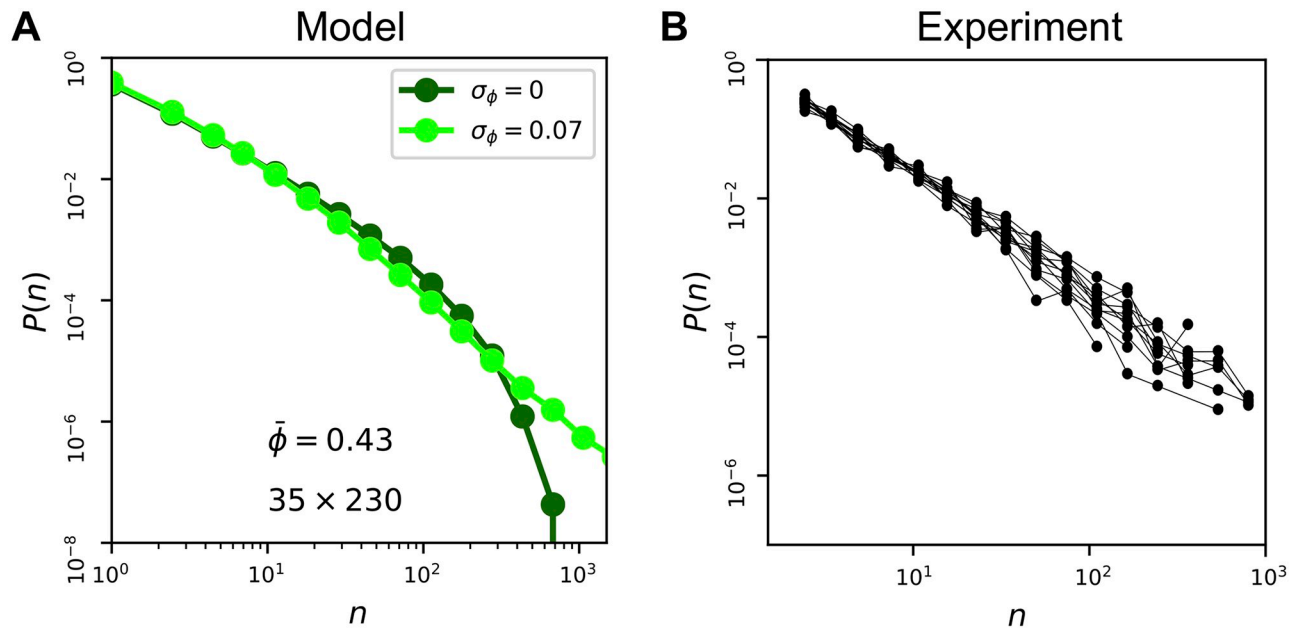
Why do correlations not change the distribution of cluster sizes? Renormalization-group arguments from statistical physics imply that correlations do not change the critical properties of percolation theory if the correlations are sufficiently short-range [18]. The intuitive reason can be seen from a site-decimation procedure [6], as illustrated in Fig 5A. We imagine decimating every other cell in each column (red X's), with each remaining cell expanding to fill the space below it. Fig 5A illustrates that the resulting lattice remains triangular (green lines). Furthermore, because the probability of any cell to be on is  $\phi$ , the fraction of on-cells remains  $\phi$  after decimation. Finally, the new conditional probabilities after one round of decimation are

$$p_1(\text{on}|\text{on}) = p(\text{on}|\text{on})p(\text{on}|\text{on}) + p(\text{on}|\text{off})p(\text{off}|\text{on}), \quad (5)$$

$$p_1(\text{on}|\text{off}) = p(\text{on}|\text{on})p(\text{on}|\text{off}) + p(\text{on}|\text{off})p(\text{off}|\text{off}), \quad (6)$$

which follow from the rules of probability and the assumption that the signaling state is spatially Markovian, i.e. the daughter is conditionally independent of the grandmother given the mother (see Materials and methods). As a result, the correlation parameter after one round of decimation is  $\rho_1 = p_1(\text{on}|\text{on}) - p_1(\text{on}|\text{off}) = [p(\text{on}|\text{on}) - p(\text{on}|\text{off})]^2 = \rho^2$ , where the first and last steps use the definition in Eq 2, and the middle step inserts the expressions in Eqs 5 and 6 and simplifies (see Materials and methods). Similarly, after  $j$  rounds of decimation we have  $\rho_j = \rho^{j+1}$ . Because  $\rho < 1$ , we see that  $\rho_j \rightarrow 0$  as  $j \rightarrow \infty$ . Thus, correlations vanish upon repeated rounds of decimation and renormalization. This means that correlations are not expected to change the critical properties of the distribution  $P(n)$ .

The above intuition only holds if the correlations are sufficiently short-range. Indeed, Eqs 5 and 6 assume that the correlations are minimally short-range, namely Markovian. In general, it has been shown that spatial correlations only affect the critical properties of percolation if they decay as a power law, specifically  $C(r) \sim r^{-a}$  with  $a > 3/2$  in 2D [18]. As seen in Fig 5B,



**Fig 6. Variability can lead to power-law cluster size distribution, even for  $\phi < \phi_c$ .** (A) In the model, variability ( $\sigma_\phi = 0.07$ ) removes rolloff, causing distribution to approach a power law over three decades ( $N = 10^3$  lattices). (B) In the experiments, cluster size distributions from individual biofilms are power laws without significant rolloff, consistent with the model and the fact that we find variability in  $\phi$  within each biofilm. Data are from [5] but processed individually for each biofilm.

<https://doi.org/10.1371/journal.pcbi.1007508.g006>

the correlations in the experimental data are much shorter-range than a power law. This suggests that the spatial correlations that we observe in the biofilm are not sufficiently long-range to affect the critical properties. Together with Fig 4C, we conclude that spatial correlations are not sufficient to explain the experimentally observed power law dependence of  $P(n)$  over three decades [5].

### Variability in signaling fraction

If spatial correlations cannot explain the experimentally observed power law, then what can? An important feature of the experiments that is not yet accounted for in the model is variability in the on-cell fraction  $\phi$ . In particular, we previously observed that the value of  $\phi$  is roughly Gaussian-distributed across 12 experiments with a mean of  $\bar{\phi} = 0.43$  and a standard deviation of  $\sigma_\phi = 0.07$  (from which the standard error of  $0.07/\sqrt{12} = 0.02$  comes) [5]. Furthermore, subdividing each of the 12 images into either 4 or 16 equal parts with the same aspect ratio as the original image, we find that the standard deviation of the on-cell fraction across parts (averaged over all images) is  $\sigma_\phi = 0.04$  (4 parts) or  $\sigma_\phi = 0.05$  (16 parts). Because these values are similar to  $\sigma_\phi = 0.07$ , we conclude that the variability within biofilms is similar to that across the biofilms in our experiments.

Some variability is expected from finite size effects. Specifically, in basic percolation theory, binomial statistics dictate that the standard deviation in the fraction of on-cells would be  $\sqrt{N\bar{\phi}(1-\bar{\phi})}/N = 0.006$  in a biofilm with  $N = 230 \times 35 = 8,050$  cells. In our mechanistic model with correlations, we find that the standard deviation is similarly small at 0.009. Because these values are much smaller than the observed value of  $\sigma_\phi = 0.07$ , we conclude that the experimental variability is not due to finite size effects alone, and that it is necessary to explicitly incorporate variability into the model.

To incorporate variability in the on-cell fraction, we draw  $\phi$  for each lattice from a Gaussian distribution with standard deviation  $\sigma_\phi$ . Because we have found that correlations have little effect on  $P(n)$ , we set  $\rho_{\text{div}} = 0$  from here on for simplicity. The results are shown in Fig 6A, and we see that  $\sigma_\phi$  has a significant effect on the distribution. In particular, for the experimental value  $\sigma_\phi = 0.07$  (light green curve), we see that the exponential rolloff at large  $n$  is removed, extending the range of the power law out to  $n \sim 10^3$  as observed in the experiments [5]. The intuitive reason is that a non-negligible fraction of lattices in the ensemble have  $\phi$  values that are equal to or greater than  $\phi_c^{\text{pow}} = 1/2$ . Because  $\phi$  is higher in these lattices, they are more likely to have large clusters. Therefore, these lattices dominate the distribution at large  $n$ , eliminating the rolloff. Thus, variability in  $\phi$  effectively widens the range of mean  $\bar{\phi}$  values at which a power law distribution is observed. We conclude that the experimental variability in  $\phi$  across biofilms is sufficient to explain the experimentally observed power-law distribution.

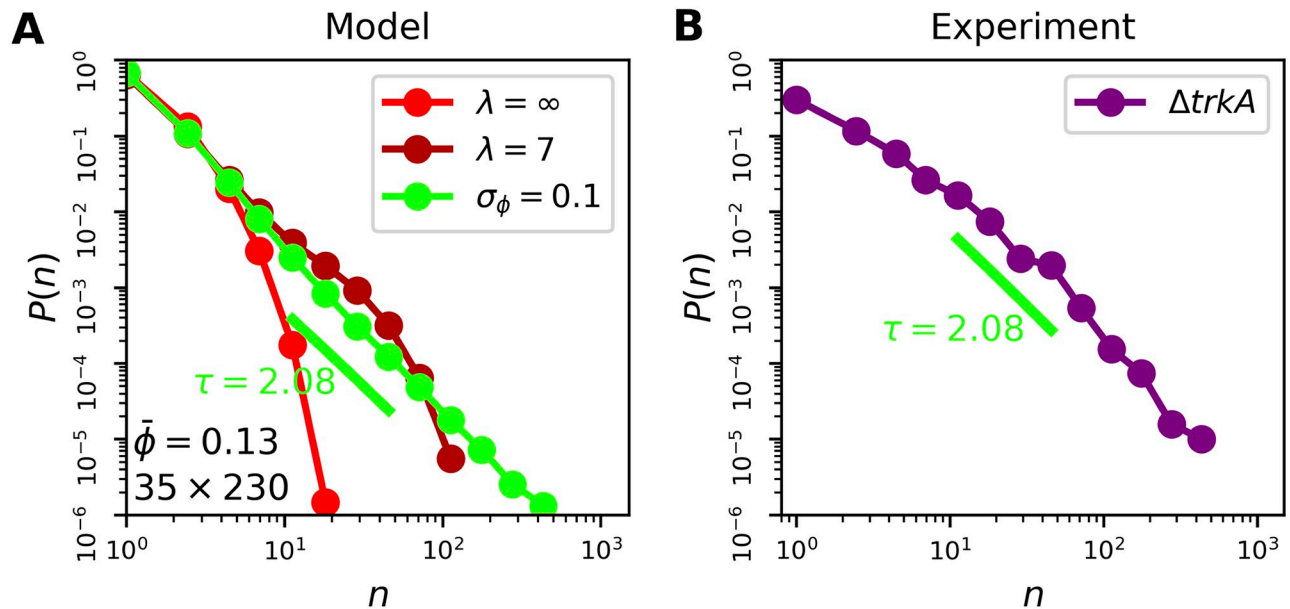
Given that we also observe variability in  $\phi$  within each biofilm, to a similar degree as across biofilms, our results suggest that the cluster size distribution from each biofilm individually should follow a power law without a significant rolloff. We test this hypothesis in Fig 6B by plotting the data from [5] separately for each biofilm. We see that indeed, the individual distributions follow a power law and do not exhibit significant rolloffs. This result suggests that the mechanism we identify above, in which variability widens the range of  $\bar{\phi}$  values at which a power law distribution is observed, also applies at the individual biofilm level. It also shows that the signaling statistics are reproducible from biofilm to biofilm and thus constitute a plausible feature that could be optimized for biological function, as suggested in our previous work [5].

### Model validation using mutant strain

How can our model be tested with further experiments? One approach is to investigate a system with a different fraction of on-cells and see if our model remains valid. We previously investigated mutant strains with different on-cell fractions, including the  $\Delta trkA$  strain with  $\bar{\phi} = 0.13$  and  $\sigma_\phi = 0.1$  [5]. As seen in Fig 7A (light red curve), basic percolation theory ( $\rho_{\text{div}} = 0$ ,  $\sigma_\phi = 0$ ) predicts that a system with an on-cell fraction of  $\phi = 0.13$  would have a distribution of cluster sizes  $P(n)$  that is entirely exponential because 0.13 is much lower than  $\phi_c^{\text{pow}} = 1/2$ .

However, the  $\Delta trkA$  strain differs from the wild-type strain in that the fraction of on-cells is not constant in space, but rather decreases along the signaling direction [15] with a characteristic lengthscale of approximately  $\lambda = 15 \mu\text{m}$ , or about 7 cell lengths [5]. The reason for this decrease is likely that the signal is dying out due to insufficient connectivity of the on-cells. Therefore, in the case of  $\Delta trkA$ , the on-cell fraction is sufficiently low that it is likely no longer fair to treat the ability to signal and the act of signaling as equivalent. Because the experiments measure the latter, we must incorporate the observed spatial decrease into the model. To do so, we allow the on-cell fraction to vary as  $\phi(y) = \phi_0 e^{-y/\lambda}$ , where  $\phi_0$  is set to ensure that the spatial average of  $\phi(y)$  is 0.13. We see in Fig 7A (dark red curve) that this feature extends the distribution to larger cluster sizes  $n$ . The reason is similar to that given above regarding variability: the portions of the lattice in which  $\phi$  is large contain large clusters, thereby enhancing the large- $n$  region of the distribution. Nonetheless, the distribution remains far from a power law in its shape. In particular, a clear exponential rolloff at large  $n$  is evident.

If our main finding above is correct, namely that variability in  $\phi$  across biofilms is a crucial determinant of the shape of  $P(n)$ , then we must also incorporate into our model the variability  $\sigma_\phi = 0.1$  observed for the  $\Delta trkA$  strain. Indeed, we find that doing so has a major effect on the distribution (Fig 7A, green curve). Specifically, it removes the exponential rolloff, resulting in a power-law distribution over almost three decades. This is a strong prediction, considering that



**Fig 7. Statistics of mutant  $\Delta trkA$  strain.** (A) We progressively incorporate into the model the on-cell fraction  $\phi = 0.13$  (light red), the exponential decay of  $\phi$  in space with lengthscale  $\lambda = 7$  cells (dark red), and the variability  $\sigma_\phi = 0.1$  across lattices (green);  $N = 10^3$  lattices for each. Resulting  $P(n)$  is a power law (green) despite the fact that 0.13 is far below the critical fraction  $\phi_c^{pow} = 1/2$ . (B)  $P(n)$  from  $\Delta trkA$  data is a power law whose exponent is consistent with the model ( $N = 7$  biofilms).

<https://doi.org/10.1371/journal.pcbi.1007508.g007>

$\bar{\phi} = 0.13$  is much lower than  $\phi_c^{pow} = 1/2$ , and that without variability the shape is far from a power law even after accounting for the spatial dependence of  $\phi$ .

To test this prediction, we measure the distribution of cluster sizes in the  $\Delta trkA$  biofilms (see [Materials and methods](#)). Remarkably, the result, shown in [Fig 7B](#), is a distribution that is roughly a power law over almost three decades, consistent with the model prediction. Indeed, the power-law exponent of 2.08 estimated from the model distribution via a maximum likelihood technique [19] ([Fig 7A](#), green line) is consistent with the slope of the experimental distribution ([Fig 7B](#), green line). This result validates our model. In particular, it supports the finding that variability of the signaling fraction across biofilms plays an important role in shaping the statistical properties of the system.

## Discussion

We have shown that experimentally observed features that go beyond the basic assumptions of percolation theory, including spatial correlations, variability, and non-uniformity, can have important consequences for signal propagation in a bacterial community. Using a mechanistic model that accounts for heritability in a cell's propensity to participate in signaling, we have found that signal correlations decrease the fraction of participating cells needed to create a connected path, but have little effect on the cluster statistics. In contrast, variability of the signaling fraction across samples has a significant effect on the statistics, in particular producing a power-law distribution of cluster sizes at signaling fractions lower than the expected critical fraction from percolation theory. We have validated our model using a mutant strain, in particular finding that both spatial decay and variability in the signaling fraction play a crucial role in shaping the signaling statistics.

While it is clear that key observations in this system are consistent with the predictions of percolation theory (the fraction of signaling cells is very close to the percolation threshold and

the cluster size distribution is a power law with the predicted exponent [5]), the deviations of other observations from either the assumptions or predictions of percolation are informative. In this sense we recognize that percolation theory is a toy model. Searching for deviations from the toy model has allowed us infer information about the biological mechanism, and then develop an enhanced model of percolation that is more appropriate. The approach of extending percolation to account for additional features has a long precedent in the literature, with variants including explosive percolation, fractional percolation, correlated percolation, bootstrap percolation, invasion percolation, and dynamical percolation [6, 20–22]. This approach is particularly suitable for biological systems, where it is natural to expect that the complexities of growth and variability may lead to observable departures from simple textbook models.

We found that incorporating the experimentally observed variability and non-uniformity of the signaling fraction into the model was necessary to explain the experimentally observed cluster statistics, whereas incorporating the experimentally observed spatial correlations in signaling was not necessary. This finding implies that certain underlying cell-level features are important in determining population-level statistical properties, whereas others are not. This categorization is consistent with approaches from statistical physics, particularly the renormalization group, which reflect the powerful notion that some microscopic details are relevant for macroscopic properties, whereas others are provably irrelevant [23]. Indeed, we explain our finding that spatial correlations do not affect the cluster statistics using a renormalization argument (Fig 5), as well as more rigorous known results from statistical physics [18]. It will be interesting to see what other cell-level features are relevant or irrelevant for capturing population-level phenomena in multicellular systems.

Finite-size effects play an important role in our results. In particular, our experimental observation window is sufficiently short in the signaling direction ( $\sim 35$  cells) that spatial correlations in the signaling propensity have a measurable effect on the connectivity (Fig 4). Yet, the window is wide perpendicular to signaling ( $\sim 230$  cells), and thus the window area is sufficiently large that the spatial correlations have little effect on the cluster size statistics. This choice of window size follows from experimental constraints and the desire to focus on the short and wide biofilm edge, where signaling is most important for function [15]. Nonetheless, it is an interesting open question how the finite size and aspect ratio of the system set distinct thresholds for the relevance of correlations to the connectivity and cluster statistics.

Dimensionality also plays an important role in our results. Because the biofilm edge is where cell growth is most pronounced, it is quasi-two-dimensional. In fact, biofilm formation itself can promote cell spreading via osmotic pressure gradients, reducing biofilm thickness at the edge [24]. This is part of the reason that our experiments have focused on 2D monolayers of cells. However, the properties of percolation theory depend critically on the dimensionality of the system [6]. In particular, the percolation threshold is generally smaller in 3D lattices than in 2D lattices [25] because there are more available paths for the signal to take. This observation suggests that a lower fraction of signaling cells is necessary in the bulk of the biofilm than at its edge. This prediction is currently difficult to test, as the 2D nature of our experiments is crucial for obtaining fluorescence data at the single-cell level.

The fact that spatial correlations lower the connectivity threshold in a finite system may help explain why the biofilm has an on-cell fraction of  $\phi = 0.43 \pm 0.02$  [5]. Naive percolation theory predicts a threshold of  $\phi_c^{\text{conn}} = 1/2$  [6], which the biofilm does not meet. Accounting for finite-size effects lowers the threshold to  $\phi_c^{\text{conn}} = 0.45$  [5] (Fig 4), which the biofilm barely meets. Accounting for correlations lowers the threshold further to  $\phi_c^{\text{conn}} = 0.4$  (Fig 4), which the biofilm meets comfortably. Thus, correlations provide some leeway between the necessary

and observed signaling fraction, which may enhance the reliability of signaling or make it robust to errors.

Although we observe spatial correlations in the signaling activity, and the results are consistent with a model that assumes inheritance of the signaling state, the inheritance mechanism is unknown. *B. subtilis* cells maintain phenotypic states through intracellular genetic networks that control the production of transcription factors [26]. Moreover, the inheritance of transcription factors and other proteins from parent cells to daughter cells can maintain specific cell types for several generations, leading to spatial correlation of cell types [27]. *B. subtilis* has even evolved the ability to control the number of generations over which certain phenotypic states are maintained [28]. Such a mechanism could drive the inheritance in signaling state that we observe here: the transcription factors regulating the expression or non-expression of ion channels, for example, could be passed from mother to daughter.

The effect of spatial correlations is a general question that is fundamental to understanding multicellular behaviors. The length scale of cell-to-cell signaling in quorum sensing bacterial communities depends on the establishment of spatial correlations [7, 29]. Moreover, the interplay of spatial heterogeneity and signaling lengthscale dictates the cooperativity of pathogenic *Pseudomonas aeruginosa* biofilms [30]. In eukaryotes, spatial correlations in cell-substrate interactions can drive collective cell migration [31], which is a fundamental multicellular process in tissue development [32] and wound healing [33].

Our study motivates further avenues of exploration in both statistical physics and cell biology. In statistical physics, our study motivates more general investigations of whether and how particular microscopic features affect macroscopic properties of percolation. The effects of spatial correlations in the site occupation probability are relatively well understood [20, 22, 34–37], whereas the effects of variability and non-uniformity in the site occupation probability are still relatively open questions [38, 39]. In cell biology, our study builds on previous work [5, 8–13] that demonstrates the utility of percolation theory as a quantitative and predictive description of multicellular phenomena. It will be interesting to see in what biological systems ideas from percolation theory will provide useful insights next.

## Materials and methods

### Experimental methods

**Microfluidics and experimental conditions.** Bacterial strains and growth conditions were as in [5]. We performed experiments in Y04D microfluidic plates using the CellASIC ONIX microfluidic system (EMD Millipore). Cells were imaged at the edge of biofilms and were confined to a single-cell layer by the PDMS structures of the microfluidic chamber. Each microscope field of view was roughly  $330\ \mu\text{m} \times 70\ \mu\text{m}$  and contained 8,000 – 10,000 cells. Every 5 minutes, we took phase contrast and fluorescence images on an Olympus IX83 inverted microscope with autofocus and a 40X, 0.6 NA air objective.

To probe membrane potential, we used the cationic fluorescent dye Thioflavin-T (ThT), which acts as a Nernstian voltage indicator [15]. When cells are hyperpolarized, they retain more of the dye and have a higher signal. ThT was present in the media at a concentration of  $10\ \mu\text{M}$ . We considered a cell to be an on-cell if its mean ThT signal exceeded a particular threshold during a signal pulse [5].

**Computation of correlation function.** To compute correlation functions, we first thresholded ThT images so that they were binary: biofilm regions above the ThT threshold would appear white and sub-threshold regions would appear black. We then applied a 2-pixel radius median filter to thresholded images so that clusters of on-cells became contiguous white regions. From this image, we created a 2D autocorrelation plot using the ImageJ command FD

Math. The resulting plot was mean-subtracted and normalized such that the origin had a value of 1 and decayed to 0 away from the origin (see source code for the Radially Averaged Auto-correlation ImageJ plugin for further details).

To compute the radial autocorrelation curves (Fig 1C), we took a radial average of this 2D correlation plot. For  $x$  and  $y$  correlation curves, we took profiles of the correlation plot along the  $x$  and  $y$  axes, respectively.

To construct randomized images for such correlation computations, we took segmented biofilm images and randomly assigned a fraction of cells to be on and made them white. We then computed the autocorrelation curve on these images the same way as with the experimental images.

**Lineage tracing for  $\rho_{\text{div}}$ .** To determine  $\rho_{\text{div}}$  we tracked individual cell lineages over time within biofilms using the mTrackJ imageJ plugin [40]. For each lineage, we recorded the firing state (i.e. on or off) of the parent cell and the daughter cells. Using many lineages, we computed the conditional probabilities  $p(\text{on}|\text{on})$ ,  $p(\text{on}|\text{off})$ ,  $p(\text{off}|\text{on})$ , and  $p(\text{off}|\text{off})$ . We then computed the order parameter  $\rho_{\text{div}}$  using Eq 2.

**Spatial analysis for  $\rho_{\text{adj}}$ .** To determine  $\rho_{\text{adj}}$ , we segmented cells in static images taken during signal pulses and determined the firing state of each cell (i.e. on or off). Because the electrical signal propagates in the direction of cell growth, cells are generally oriented along the signaling direction (Fig 1B). The adjacent cell in each case was defined as the cell whose bottom edge was closest to the given cell's top edge, and whose centroid was within half the average cell width. We then computed the conditional probabilities  $p(\text{on}|\text{on})$ ,  $p(\text{on}|\text{off})$ ,  $p(\text{off}|\text{on})$ , and  $p(\text{off}|\text{off})$  for the firing state of a cell given the state of the adjacent cell. We then computed the order parameter  $\rho_{\text{adj}}$  using Eq 2.

**Image analysis for  $\Delta\text{trkA}$ .** We evaluated the cluster size distribution for  $\Delta\text{trkA}$  biofilms in Fig 7B by first segmenting single biofilm cells in phase images using the Trainable Weka Segmentation plugin in ImageJ. We then thresholded the corresponding ThT images as described in the above section on computing correlation curves. Each contiguous white region in the thresholded image was a cluster of on-cells. We then counted how many segmented cells had the majority of their area within each cluster. The curve in Fig 7B plots the normalized histogram of these cluster sizes.

## Theoretical methods

**Mechanistic model.** To derive Eqs 3 and 4, we require that the fraction of on-cells is  $\phi$  at each step in the growth process. Specifically, the rules of probability state that

$$p(d) = \sum_m p(d, m) = \sum_m p(d|m)p(m), \tag{7}$$

where  $d$  is the signaling state (on, off) of the daughter, and  $m$  is the signaling state (on, off) of the mother. Taking  $d = \text{on}$  and requiring that  $p(\text{on}) = \phi$  and  $p(\text{off}) = 1 - \phi$ , Eq 7 becomes

$$\phi = p(\text{on}|\text{on})\phi + p(\text{on}|\text{off})(1 - \phi). \tag{8}$$

Solving for  $\phi$ , we obtain

$$\phi = \frac{p(\text{on}|\text{off})}{1 + p(\text{on}|\text{off}) - p(\text{on}|\text{on})}. \tag{9}$$

Combining this equation with Eq 2 and solving for the conditional probabilities, we obtain Eqs 3 and 4.



**Differential growth rates.** We previously observed that signal participation reduces the cell elongation rate [5], implying that on-cells grow more slowly than off-cells. Specifically, Fig 1B of [5] shows that the elongation rate is reduced by a factor of about 4 at peak signaling activity. On-cells signal for about 20 minutes (Fig 4E of [5]), whereas pulses occur every 80 minutes or so (Fig S4 of [5]). Therefore the net growth rate ratio of on-cells to off-cells is approximately  $\gamma = (1/4)(20/80) + (1)(60/80) \approx 80\%$ .

To incorporate this feature into the model, we take the mean division time to be  $\bar{\tau}$  and  $\bar{\tau}/\gamma$  for off-cells and on-cells, respectively. Differential growth rates change the resulting fraction of on-cells in the lattice, and therefore Eqs 3 and 4 must be modified to maintain this fraction at  $\phi$ . Specifically, using the shorthand  $q \equiv p(\text{on}|\text{on})$  and  $r \equiv p(\text{on}|\text{off})$  and recognizing that  $p(\text{off}|\text{on}) = 1 - q$  and  $p(\text{off}|\text{off}) = 1 - r$ , the deterministic dynamics of the number of on- and off-cells are

$$\dot{n}_{\text{on}} = \gamma q n_{\text{on}} + r n_{\text{off}}, \tag{10}$$

$$\dot{n}_{\text{off}} = \gamma(1 - q)n_{\text{on}} + (1 - r)n_{\text{off}}, \tag{11}$$

where time is scaled by  $\bar{\tau}$ . At long times, the larger of the two eigenvalues of this linear dynamical system is

$$\lambda_+ = \frac{1}{2} \left[ 1 + \gamma q - r + \sqrt{(1 + \gamma q - r)^2 - 4\gamma(q - r)} \right], \tag{12}$$

and the ratio of the two components of the corresponding eigenvector gives the ratio of  $n_{\text{on}}$  and  $n_{\text{off}}$ . Setting the ratio of  $n_{\text{on}}$  and  $n_{\text{on}} + n_{\text{off}}$  to  $\phi$  obtains

$$\phi = \frac{r}{\lambda_+ + r - \gamma q}. \tag{13}$$

Note that taking  $\gamma = 1$  makes  $\lambda_+ = 1$ , and Eq 13 recovers Eq 9. Combining Eqs 12 and 13 with  $\rho = q - r$  (Eq 2) and solving for  $q$  and  $r$  obtains

$$p(\text{on}|\text{on}) = \phi + \frac{\rho(1 - \phi)}{1 - (1 - \gamma)\phi}, \tag{14}$$

$$p(\text{on}|\text{off}) = \phi - \frac{\gamma\phi\rho}{1 - (1 - \gamma)\phi}. \tag{15}$$

These expressions replace Eqs 3 and 4, respectively. This derivation ignores the differential crowding effects in the simulation due to the differential growth rates, but for  $\phi = 0.43$ ,  $\rho = 0.38$ , and  $\gamma = 0.8$  we still find that the resulting fraction of on-cells in the lattice is  $0.428 \pm 0.009$ , which includes 0.43 within error.

In this model with differential growth rates, we find that all of the predictions of Fig 3 remain unchanged: the correlation lengths in both the  $x$  and  $y$  directions are significantly larger than random ( $p < 0.001$  for both), and  $\rho_{\text{adj}} = 0.19 \pm 0.02$ , which actually now agrees with the measured 0.17 within error.

**Renormalization argument.** To derive Eqs 5 and 6, we recognize that the conditional probability of the daughter given the mother after one round of decimation is the conditional probability of daughter given the grandmother before the decimation. Again using the rules of

probability, we write the latter as

$$p(d|g) = \sum_m p(d, m|g) = \sum_m p(d|m, g)p(m|g), \quad (16)$$

where  $g$  is the signaling state (on, off) of the grandmother. The spatial Markovian assumption states that  $d$  is conditionally independent of  $g$  given  $m$ . Therefore we have  $p(d|m, g) = p(d|m)$ , and Eq 16 becomes

$$p(d|g) = \sum_m p(d|m)p(m|g). \quad (17)$$

Setting  $d = \text{on}$  and  $g = \text{on}$  gives Eq 5. Setting  $d = \text{on}$  and  $g = \text{off}$  gives Eq 6.

To derive the relation  $\rho_1 = \rho^2$  below Eq 6, we insert Eqs 5 and 6 into the definition  $\rho_1 = p_1(\text{on}|\text{on}) - p(\text{on}|\text{off})$ . Again using the shorthand  $q \equiv p(\text{on}|\text{on})$  and  $r \equiv p(\text{on}|\text{off})$  and recognizing that  $p(\text{off}|\text{off}) = 1 - q$  and  $p(\text{off}|\text{off}) = 1 - r$ , this insertion obtains

$$\rho_1 = q^2 + r(1 - q) - qr - r(1 - r) = q^2 - 2qr + r^2 = (q - r)^2. \quad (18)$$

Because  $\rho = q - r$  (Eq 2), we see that  $\rho_1 = \rho^2$ .

## Acknowledgments

We thank Aleksandra Walczak and Jordi Garcia-Ojalvo for helpful discussions.

## Author Contributions

**Conceptualization:** Gürol M. Süel, Andrew Mugler.

**Formal analysis:** Xiaoling Zhai, Ushasi Roy, Andrew Mugler.

**Funding acquisition:** Gürol M. Süel, Andrew Mugler.

**Investigation:** Xiaoling Zhai, Joseph W. Larkin, Kaito Kikuchi, Samuel E. Redford, Andrew Mugler.

**Methodology:** Xiaoling Zhai, Joseph W. Larkin, Andrew Mugler.

**Software:** Xiaoling Zhai, Joseph W. Larkin.

**Supervision:** Gürol M. Süel, Andrew Mugler.

**Validation:** Xiaoling Zhai.

**Writing – original draft:** Xiaoling Zhai, Andrew Mugler.

**Writing – review & editing:** Xiaoling Zhai, Joseph W. Larkin, Ushasi Roy, Gürol M. Süel, Andrew Mugler.

## References

1. Li B, You L. Predictive power of cell-to-cell variability. *Quantitative Biology*. 2013; 1(2):131–139. <https://doi.org/10.1007/s40484-013-0013-3>
2. Symmons O, Raj A. What's Luck Got to Do with It: Single Cells, Multiple Fates, and Biological Nondeterminism. *Molecular Cell*. 2016; 62(5):788–802. <https://doi.org/10.1016/j.molcel.2016.05.023>.
3. Steinberg BE, Glass L, Shrier A, Bub G. The role of heterogeneities and intercellular coupling in wave propagation in cardiac tissue. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2006; 364(1842):1299–1311. <https://doi.org/10.1098/rsta.2006.1771>
4. Waxman SG. Axonal conduction and injury in multiple sclerosis: the role of sodium channels. *Nature Reviews Neuroscience*. 2006; 7(12):932. <https://doi.org/10.1038/nrn2023>

5. Larkin JW, Zhai X, Kikuchi K, Redford SE, Prindle A, Liu J, et al. Signal Percolation within a Bacterial Community. *Cell systems*. 2018; 7(2):137–145. <https://doi.org/10.1016/j.cels.2018.06.005>
6. Stauffer D, Aharony A. Introduction to percolation theory. Taylor & Francis; 1994.
7. Silva KPT, Yusufaly TI, Chellamuthu P, Boedicker JQ. Disruption of microbial communication yields a two-dimensional percolation transition. *Physical Review E*. 2019; 99(4):042409. <https://doi.org/10.1103/PhysRevE.99.042409>
8. Breskin I, Soriano J, Moses E, Tlusty T. Percolation in living neural networks. *Physical review letters*. 2006; 97(18):188102. <https://doi.org/10.1103/PhysRevLett.97.188102>
9. Eckmann JP, Feinerman O, Gruendlinger L, Moses E, Soriano J, Tlusty T. The physics of living neural networks. *Physics Reports*. 2007; 449(1-3):54–76. <https://doi.org/10.1016/j.physrep.2007.02.014>
10. Gönci B, Németh V, Balogh E, Szabó B, Dénes Á, Környei Z, et al. Viral epidemics in a cell culture: novel high resolution data and their interpretation by a percolation theory based model. *PloS one*. 2010; 5(12):e15571. <https://doi.org/10.1371/journal.pone.0015571>
11. Zhou DW, Mowrey DD, Tang P, Xu Y. Percolation model of sensory transmission and loss of consciousness under general anesthesia. *Physical review letters*. 2015; 115(10):108103. <https://doi.org/10.1103/PhysRevLett.115.108103>
12. Benninger RKP, Zhang M, Head WS, Satin LS, Piston DW. Gap Junction Coupling and Calcium Waves in the Pancreatic Islet. *Biophysical Journal*. 2008; 95(11):5048–5061. <https://doi.org/10.1529/biophysj.108.140863>
13. Mathijssen A, Culver J, Bhamla MS, Prakash M. Collective intercellular communication through ultra-fast hydrodynamic trigger waves. *Nature*. 2019; 571:560–564. <https://doi.org/10.1038/s41586-019-1387-9>
14. Liu J, Prindle A, Humphries J, Gabalda-Sagarra M, Asally M, yeon D Lee D, et al. Metabolic co-dependence gives rise to collective oscillations within biofilms. *nature*. 2015; 523:550–554. <https://doi.org/10.1038/nature14660>
15. Prindle A, Liu J, Asally M, Ly S, Garcia-Ojalvo J, Süel GM. Ion channels enable electrical communication in bacterial communities. *nature*. 2015; 527:59–63. <https://doi.org/10.1038/nature15709>
16. Humphries J, Xiong L, Liu J, Prindle A, Yuan F, Arjes HA, et al. Species-independent attraction to biofilms through electrical signaling. *Cell*. 2017; 168(1-2):200–209. <https://doi.org/10.1016/j.cell.2016.12.014>
17. Shao X, Mugler A, Kim J, Jeong HJ, Levin BR, Nemenman I. Growth of bacteria in 3-d colonies. *PLoS computational biology*. 2017; 13(7):e1005679. <https://doi.org/10.1371/journal.pcbi.1005679>
18. Weinrib A. Long-range correlated percolation. *Physical Review B*. 1984; 29(1):387. <https://doi.org/10.1103/PhysRevB.29.387>
19. Clauset A, Shalizi CR, Newman ME. Power-law distributions in empirical data. *SIAM review*. 2009; 51(4):661–703. <https://doi.org/10.1137/070710111>
20. Essam JW. Percolation theory. *Reports on progress in physics*. 1980; 43(7):833. <https://doi.org/10.1088/0034-4885/43/7/001>
21. Steif JE. A survey of dynamical percolation. In: *Fractal geometry and stochastics IV*. Springer; 2009. p. 145–174.
22. Saberi AA. Recent advances in percolation theory and its applications. *Physics Reports*. 2015; 578:1–32. <https://doi.org/10.1016/j.physrep.2015.03.003>
23. Goldenfeld N. Lectures on phase transitions and the renormalization group. CRC Press; 2018.
24. Seminara A, Angelini TE, Wilking JN, Vlamakis H, Ebrahim S, Kolter R, et al. Osmotic spreading of *Bacillus subtilis* biofilms driven by an extracellular matrix. *Proceedings of the National Academy of Sciences*. 2012; 109(4):1116–1121. <https://doi.org/10.1073/pnas.1109261108>
25. Wang J, Zhou Z, Zhang W, Garoni TM, Deng Y. Bond and site percolation in three dimensions. *Physical Review E*. 2013; 87(5):052107. <https://doi.org/10.1103/PhysRevE.87.052107>
26. Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. *Nature*. 2010; 467(7312):167. <https://doi.org/10.1038/nature09326>
27. Chen R, Guttenplan SB, Blair KM, Kearns DB. Role of the  $\sigma$  D-dependent autolysins in *Bacillus subtilis* population heterogeneity. *Journal of bacteriology*. 2009; 191(18):5775–5784. <https://doi.org/10.1128/JB.00521-09>
28. Norman TM, Lord ND, Paulsson J, Losick R. Memory and modularity in cell-fate decision making. *Nature*. 2013; 503(7477):481. <https://doi.org/10.1038/nature12804>
29. Silva KP, Chellamuthu P, Boedicker JQ. Signal destruction tunes the zone of activation in spatially distributed signaling networks. *Biophysical journal*. 2017; 112(5):1037–1044. <https://doi.org/10.1016/j.bpj.2017.01.010>

30. Darch SE, Simoska O, Fitzpatrick M, Barraza JP, Stevenson KJ, Bonnacaze RT, et al. Spatial determinants of quorum signaling in a *Pseudomonas aeruginosa* infection model. *Proceedings of the National Academy of Sciences*. 2018; 115(18):4779–4784. <https://doi.org/10.1073/pnas.1719317115>
31. Angelini TE, Hannezo E, Trepas X, Fredberg JJ, Weitz DA. Cell migration driven by cooperative substrate deformation patterns. *Physical review letters*. 2010; 104(16):168104. <https://doi.org/10.1103/PhysRevLett.104.168104>
32. Friedl P, Hegerfeldt Y, Tusch M. Collective cell migration in morphogenesis and cancer. *International Journal of Developmental Biology*. 2004; 48(5-6):441–449. <https://doi.org/10.1387/ijdb.041821pf>
33. Martin P, Parkhurst SM. Parallels between tissue repair and embryo morphogenesis. *Development*. 2004; 131(13):3021–3034. <https://doi.org/10.1242/dev.01253>
34. Prakash S, Havlin S, Schwartz M, Stanley HE. Structural and dynamical properties of long-range correlated percolation. *Physical Review A*. 1992; 46(4):R1724. <https://doi.org/10.1103/PhysRevA.46.R1724>
35. Sahimi M, Mukhopadhyay S. Scaling properties of a percolation model with long-range correlations. *Physical Review E*. 1996; 54(4):3870. <https://doi.org/10.1103/PhysRevE.54.3870>
36. Makse HA, Andrade JS, Batty M, Havlin S, Stanley HE, et al. Modeling urban growth patterns with correlated percolation. *Physical Review E*. 1998; 58(6):7054. <https://doi.org/10.1103/PhysRevE.58.7054>
37. Sahimi M. Non-linear and non-local transport processes in heterogeneous media: from long-range correlated percolation to fracture and materials breakdown. *Physics Reports*. 1998; 306(4-6):213–395. [https://doi.org/10.1016/S0370-1573\(98\)00024-6](https://doi.org/10.1016/S0370-1573(98)00024-6)
38. Kundu S, Manna S. Percolation model with an additional source of disorder. *Physical Review E*. 2016; 93(6):062133. <https://doi.org/10.1103/PhysRevE.93.062133>
39. Ikeda H. Percolation in Anisotropic Systems: —Real-Space Renormalization Group—. *Progress of Theoretical Physics*. 1979; 61(3):842–849. <https://doi.org/10.1143/PTP.61.842>
40. Meijering E, Dzyubachyk O, Smal I. Methods for cell and particle tracking. In: *Methods in enzymology*. vol. 504. Elsevier; 2012. p. 183–200.