Signal Amplification in Synthetic Bacterial Communication

Thesis by
James M. Parkin

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2021
Defended February 4th, 2021

James M. Parkin
ORCID: 0000-0002-4058-2338

ABSTRACT

Synthetic biology will one day enable embedded control of a variety of chemical and biological contexts, from the human gastrointestinal tract to crop roots. Groups of engineered organisms, also known as synthetic consortia, can inhabit niches of interest while monitoring and intervening according to their genetic design. However, the spatial structure of the deployment environments can obstruct coordination between cosortia members. The mechanisms engineered bacteria use to communicate must contend with these adversarial conditions to maximize group performance.

Coordination between synthetic bacteria is typically achieved using small molecules that can traverse cell membranes through passive transport. Cell communicate by producing and sensing these small molecules. In cell-cell signaling relationships composed of a sender population and a receiver population, the concentration of signaling molecule sensed by the receiver cells depends on the spatial patterning of the two groups, the geometry of the diffusive environment, and the sender population's signal secretion rate.

To make sender-receiver communication more robust to these environmental features, we introduce a third consortium strain that transiently amplifies local signaling molecule concentrations. These amplifier cells employ a synchronized pulse-generating circuit built using Lux-type quorum sensing components and an IFFL transcriptional architecture. When applied to sender-receiver consortia growing on semi-solid media, these amplifier cells respond to sender-secreted signaling molecules by contributing a small amount themselves. The support of amplifier cells enables communication over longer distances than can be achieved by sender cells alone and can partially recover coordination in small consortia where the sender population is too small to successfully signal its receiver population alone. We extend these results using simulation to investigate the benefit that amplifier cells confer to consortia of varying complexity.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

Research in synthetic biology is an exercise in constant humility. Every breakthrough in this field demands a comparison to the astounding brilliance of the natural work it is imitating. It is like designing the first moon lander while seeing aliens pass through our solar system on intergalactic spacecraft without saying so much as "hello".

Such is the relationship of a synthetic biologist to all living things. The majority of research and commercial applications in this field rely on domesticated cell lines to express synthetic gene circuits without significantly altering the genomic programming for the cell line's key biological features, such as self-replication and metabolic homeostasis. In a 2020 comment paper published in Nature Communications, Voigt highlighted six commercial products that were manufactured using genetically modified oranganisms. Three were products produced by engineered cells and the other three were themselves engineered cells (Voigt (2020)). The commercial success of these commodities derived from engienered cells illustrates that bioengineered cellular factories represent a significant industrial source of valuable chemicals and materials, and the living commodities show that the future is bright for applications of engineered organisms.

While these indicate synthetic biology's growing relevance in the commercial sector, the article notes that the next phase of industrial synthetic biology would need to develop on engineering cells as embedded controllers for biosynthesis and biochemical environments of interest. Consider for example one of the living commodities highlighted in Voigt's article: a rhizosphere bacteria strain engineered to overexpress nitrogen-fixing genes marketed under the name PROVEN (Willits (2020), Temme (2019)). The modification alters a conditionally-expressed gene cluster to instead be persistently active, thereby increasing soil nitrogen and crop yields while reducing the need for chemical fertilizers. Other known benefits of bacteria-plant symbiosis, however, cannot be achieved by unconditional overexpression of the relevant biosynthesis pathways (Sarma et al. (2015)). Species of the genus *Rhizobium* associate with

legume roots and secrete biocontrol agents that suppress pathogens, but the same mechanisms of action that target pathogenic bacteria also antagonize other Rhizobia (Avis et al. (2008)). Improper balancing of strain demographics in the root microbiome or adjustments to the control circuits of antagonistic genes could spoil the potential benefit provided to the plant (Jain et al. (2012)). Action by the beneficial microbial consortia associated with the root system must also properly respond to chemical signals emitted by the plant itself in order to be effective (Avis et al. (2008)). The next generation of engineered microbial interventions for complex deployment scenarios must include the capability to sense environmental signals and respond accordingly.

One avenue for expanding the functional capabilities of synthetic bacterial devices is through engineering bacterial communities. As opposed to populations of synthetic bacteria made up of a single genotype, synthetic bacterial consortia include multiple genetically distinct strains that cooperate in completing a shared task. Strains that compose a synthetic microbial consortium can be specialized in different subroutines of the overall task. Dividing the requisite labor between component strains and optimizing each strain individually reduces the overall difficulty of the engineering task in comparison to optimizing a single strain to perform all functions. Division of labor has been successfully applied in various bioproduction applications to increase yield and titer of biochemicals (Tsoi et al. (2018), Saini et al. (2016), Zhang and Stephanopoulos (2016)). Furthermore a consortium may include multiple species to broaden the natural mechanisms in the consortium (H. J. Kim, Du, and Ismagilov (2011)). A unique challenge in consortia engineering, however, is making the intended function robust to variations in the population balance and spatial patterning of the component strains (Johns et al. (2016), Zomorrodi and Segrè (2016)).

Coordination between members of a synthetic bacterial consortium is most often achieved by modified quorum sensing systems. Quorum sensing refers to a positive-feedback gene regulatory motif, commonly found in bacteria, that is composed of a signal synthase, a signal receptor protein, and a receptor-controlled promoter. Communities of wild microbes use this motif to coordinate group behaviors in a manner that is conventionally understood to be density-dependent. The cell-cell signaling chemicals of quorum sensing systems are acyl-homoserine lactones (AHLs); the composition of the acyl chain varies between instances of this motif. AHL signaling molecules are created by the signal

synthase protein and can diffuse freely through cell membranes. The receptor proteins bind to AHL molecules, forming a complex that activates transcription at associated promoters containing the receptor's binding sequence. These components are a popular choice in the design of synthetic bacterial consortia due to their simplicity and portability. Because AHL molecules undergo passive transport through cell membranes, AHL-mediated communication can be implemented using only the two protein components and one AHL-induced promoter.

Synthetic bacterial consortia can make use of modified quorum sensing circuits to autonomously balance their strain composition and activity. In gene circuits engineered to limit or balance strain populations, a positive-feedback loop similar to natural quorum sensing circuits is connected to genes that lead to autolysis or expression of antibiotic compounds. Cells expressing these circuits limit their own population by implementing self-killing measures in a density-dependent fashion (Scott et al. (2017)). Consortia may also make use of mutualistic or antagonistic relationships between strains. These interactions are effected by metabolic relationships or targeted antibiotic interactions to achieve programmatic strain balancing (Balagaddé et al. (2008), Kong et al. (2018), Taillefumier et al. (2017)). The emergent behavior can also be made more robust without population-limiting circuits by encoding density-dependence through cell-cell signaling circuits. Chen et al. (2015) show in simulation that a two-strain relaxation oscillator is made more robust to demographic variation by the addition of a negative feedback loop used by one strain to attenuate its own activity in a density-dependent fashion. These examples demonstrate several approaches to engineering consortia that autonomously balance their strain demographics in well-mixed media.

Cell-cell communication circuits that apply spatiotemporal control over strain composition or behavior in media that is not well mixed are called pattern forming circuits. In diffusive environments, the absence of convection or turbulence allows for the formation of chemical gradients and non-homogeneous spatial patterning of the cells making up a consortium. As a result, cells in a consortium may sense different chemical signals depending on their position in these gradients. Pattern forming systems exploit spatial gradients in signaling molecules, cell density, and nutrients to generate complex spatial patterns out of growing consortia.

Ring-forming systems have been described several times in the literature (C. Liu et al. (2011), Schaerli et al. (2014), X. Xue, C. Xue, and Tang (2018), Basu et al. (2005), Cao et al. (2016), Potvin-Trottier et al. (2016)). Given the radial symmetry of bacterial colonies that grow from a small number of close founding cells, ring-forming systems are the first choice for demonstrating that a cell-cell communication system can be used to program spatially heterogeneous behaviors. Cao et al. (2016) and C. Liu et al. (2011) describe ring-forming bacteria that coordinate using quorum sensing components. On the other hand, Potvin-Trottier et al. (2016) describe a system that relies on a transcriptional oscillator circuit with the remarkable feature that, over many generations of cell division, the oscillations in daughter cells remain synchronized. As a colony of these oscillators grows, then, their synchronized oscillations give rise to uniform rings.

While these examples achieve similar spatial patterns through drastically different mechanisms, they all rely on the spatial heterogeneity of nutrient availability as a key component of pattern formation. C. Liu et al. (2011) employ a combination of AHL-mediated cell-cell signaling and synthetic chemotaxis to program a consortium to form stripes of high and low cell density. However, monotonically decreasing nutrient availability at each position is key to this mechanism: falling resource concentrations eventually fix cells in place. Indeed, the fact that cells struggle to express transgenic circuits when nutrients are low is a common feature in these ring-forming systems. Cao et al. (2016) identify that the gradient in gene expression capacity along the radius of a colony is critical to achieving the scale-free ring patterns observed in their experiments. Gene programs that create synchronized oscillations in time at the colony edge will produce oscillations in space along the colony radius as a result of the difference in gene expression capacity between the colony's interior and exterior.

These pattern formation systems underscore the fact that spatially heterogeneous gene expression is inherent to consortia growing in diffusive environments. While nutrient-dependent growth and gene expression can be leveraged in pattern-forming systems, these factors can be an enormous obstacle to coordination within multi-component consortia. The distance separating two consortium components, components that together form a signaling or metabolic relationship, dramatically impacts their emergent behavior (S. Gupta et al.

(2020), Langebrake et al. (2014), Macia et al. (2016)). A study investigating the capability of a sender-receiver pair composed of engineered *Pseudomonas putida* strains to communicate within the rhizosphere found that the introduced cells grow in sparsely-distributed groups of dozens of cells that can communicate reliably only over tens of microns (Gantner et al. (2006)). The limited signaling distances and the impact of nutrient availability highlight the need for communication networks in engineered bacterial consortia that can overcome not only spatial heterogeneity in strain demographics, but in cellular resources and relevant environmental events as well.

Agents in nature utilize traveling waves of signaling activity to share localized information over longer distances than can be achieved by diffusion from an isolated source. Slime molds such as *Dictyostelium* emit waves of cAMP, a nucleotide derivative, as they approach starvation. As neighboring amoebae join in the signaling activity, the emergent behavior appears to be an election of a community leader that serves as the target for chemotaxis by local members (Goldbeter (2006), Noorbakhsh et al. (2015)). *Bacillus subtilis* propagate waves of potassium signaling to coordinate resource sharing between the interior and exterior of growing colonies (Larkin et al. (2018), Prindle, J. Liu, et al. (2015)). Cheng and Ferrell (2018) describes a self-regenerating front of apoptotic activity that travels through cell-free extracts of Xenopus eggs in response to localized initiation. The authors of this study demonstrate a nearly 5-fold range in wave speeds under a variety of perturbations to the feedback loops supporting the traveling wave of apoptosis-related activity. In each of these articles, the researchers discuss a well-known result in traveling wave phenomena: positive feedback and local tethering are key to long-distance chemical signaling (Gelens, Anderson, and Ferrell (2014), Oleinik, Kolmogorov, and Piskunov (2019)).

The theoretical conditions necessary for traveling wave phenomena in cell-cell signaling circuits can be derived from mathematical models. For their 2020 eLife article, authors Dieterle et al. (2020) constructed reaction-diffusion models describing signaling molecule behavior in various wave-generating microbial consortia wherein a small "initiating colony" elicits a wave of signaling activity through a semi-infinite region of "relay cells". In each scenario considered, the authors derived relationships between the model parameters, such as cell density and signal emission rate, and characteristics of the traveling wave propagated by the relay cells such as velocity and signal concentration profile.

These relationships were derived for consortia employing switch-like, pulsatile, or Hill-like activation functions defining their signal emission rates as a function of the local signal concentration. Just as in the natural examples described above, the authors found each of the activation functions to be capable of producing traveling waves. Furthermore, the signaling fronts propagated by relay cells yielded higher signaling concentrations and traveled faster than fronts produced from initiating colony alone. The benefit of relay cells was more pronounced in scenarios where the dimensionality of the diffusive medium was greater than that of the consortium (e.g., signal molecules diffusing in three dimensions while the consortium cells exist in a two-dimensional plane. Other studies also suggest that local amplification in cell-cell signaling circuits could enable long-distance signaling in synthetic microbial consortia and may overcome environmental obstacles to group consensus (Langebrake et al. (2014), Holzer, Doelman, and Kaper (2013)).

Self-propagating signaling fronts that include a motif of positive-feedback and local tethering have been demonstrated experimentally in both synthetic bacteria and active chemical media.(‸The earliest described active chemical media being the inorganic Belousov-Zhabotinsky reaction (Belousov 1959, p. .) Cell-free approaches have also demonstrated that positive feedback circuits elicit traveling waves in response to localized initiation. By exploiting the precise control over the chemical composition of cell-free active media, these studies provide further validation and context for the theoretical results relating wave characteristics to reaction dynamics (Gines et al. (2017), Tayar et al. (2015)). Synthetic bacteria expressing synchronized oscillator circuits generate traveling waves when grown in a low-turbulence microfluidic device (Danino et al. (2010)). Much like the leader selection performed by *Dictyostelium*, synchronized oscillators tend towards a out-of-equilibrium state in which a minority initiates oscillations that trigger wave propagation through neighboring cells (Garcia-Ojalvo, Elowitz, and Strogatz (2004), Dalchau et al. (2018), Watts and Strogatz (1998)). These results suggest that a similar mechanism could enable sender cells to generate non-oscillatory traveling pulses through nearby propagator cells. Traveling pulse circuits could be used by consortium members to share local information, thereby enabling well-informed group decision-making from spatially heterogeneous environmental conditions.

The research presented in this thesis introduces an approach to cell-cell commu-

nication networks that supports sender-receiver relationships when the spatial patterning of deployed consortium members cannot be pre-determined. In a consortium, a sender-receiver pairing implies three subpopulations: sender, receiver, and bystander. Augmenting the bystander strains with a signal amplifier circuit enables them to generate traveling waves in response to initiating signals released by a sender population. These traveling waves would enable a consortium to share local information over longer distances than by the action of the sender population alone. Amplifier activity could also compensate for variations in the amount of sender cells by increasing the overall signal molecule concentration to counteract attenuated emission from a diminished sender population. The amplifier gene circuit in this approach is pulsatile, rather than bistable or oscillatory, which allows for repeated amplifier activity in response to periodic initiation from sender cells. Together, these features improve signaling from senders to receivers over variations in the spatial patterning of consortia and geometry of the diffusive environment.

This approach is investigated in the context of pulsatile amplifier cells and sender cells growing together on the surface of agarose hydrogels. We demonstrate that the amplifier strain generates a traveling wave of signaling activity that propagates messages from a sender population more quickly than by passive transport alone. This principle is then investigated in consortia founded by a small number of sender cells and many amplifier cells. We demonstrate that, without active signal propagation from the amplifier cells, the sender population could not marshal a response from its neighboring receiver cells. On the other hand, consortia with pulsatile activation responded quickly to their sparse sender populations. A mathematical model was developed to match the observed behavior of these sender-amplifier consortia and was used to investigate the behavior of hypothetical consortia *in silico* applied to a two-dimensional diffusive environment.

Chapter 2 describes the composition of the pulsatile amplifier circuit, its characterization through liquid-culture experiments, and applications of the amplifier strain to extend communication distances within a spatially-structured sender-receiver consortium growing in semi-solid media. In Chapter 3, the same consortium is investigated at a smaller length scale and with random spatial organization. Chapter 4 introduces a finite-differences approach to simulating the amplifier strain in various diffusive environments and ends with

an *in silico* investigation of how amplifier strains could benefit computation within hypothetical engineered consortia.

Chapter 2

LONG-DISTANCE SIGNALING USING PULSATILE AMPLIFICATION

2.1 Introduction

Quorum sensing circuits, auto-inductive genetic circuits used by wild bacteria for collective decision-making, are a popular platform for designing synthetic communication circuits. Conventionally, quorum sensing circuits are understood to allow groups of cells to estimate their local density. The signaling molecule promotes expression of its associated synthase protein, forming a positive-feedback loop that enables a rapid transition from low to high signal production rate once cell density passes a "quorum" threshold (Dockery and Keener (2001) Redfield (2002)). These natural systems have been applied relatively intact to synthetic applications, such as a circuit that delays expression of an engineered biosynthetic pathway until a quorum has been achieved (A. Gupta et al. (2017)). With small alterations to the feedback structure, these components can yield complex dynamical behaviors in multi-strain consortia (Youk and Lim (2014)). Two recent examples involve bidirectional signaling between two strains in a co-culture, in one case to achieve sustained oscillations in fluorescent protein expression and in another to recreate the classical predatory-prey dynamical system in an *E. coli* co-culture through AHL-controlled suicide and toxin-rescue circuits (Chen et al. (2015), Balagaddé et al. (2008)). Transcriptional oscillators that include a quorum sensing-like positive feedback loop tend towards phase synchronization within consortia. Depending on the diffusive environment and diffusion rate of the chemical species shared between cells, this synchronization tendency leads to different emergent behaviors. Low diffusion coefficients of signal chemicals or rapid extracellular signal loss rates produce traveling waves (~1.2 mm/hr) through dense consortia (Danino et al. (2010)). Higher diffusion coefficients or low rates of extracellular signal loss can enable synchronous oscillations over long distances (~2 mm) (Prindle, Samayoa, et al. (2012), J. K. Kim et al. (2019)).

Other applications, however, call for aperiodic yet repeated cell-cell communication and cannot rely on the circuits mentioned above. Consider a consortium

composed of sensor and actuator microbes embedded in an environment that restricts cell movement, such as intestinal mucus in the mammalian gut. The consortium is intended to persist in this environment and monitor it for a rare chemical event. In response to each event detected by the sensor population, the actuator strain should perform a discrete action, such as release a therapeutic small molecule. Without well-mixed conditions, individual sensor cells will differ in exposure to environmental variables of interest and actuators may perceive different concentrations of signal molecule produced by the sensors. These factors reduce the fraction of microbes that participate in signaling, therefore limiting the efficacy of the actuator strains' impact on the environment.

To overcome the obstacles to group consensus that are inherent to unmixed environments, microbes must collaborate in propagating signals even when they are not the signal's intended recipient. This is because, without sophisticated self-patterning programs or direct consortia printing, there is no guarantee the intended recipient population will receive a signal. Furthermore, amplification activity must be ephemeral to preclude steady states with high secretion rate and non-oscillatory in order to match propagated signals to individual initiation events. Signaling via traveling pulses, like action potentials through nervous tissue, achieves these two requirements (Holzer, Doelman, and Kaper (2013)).

Figure 2.1 shows a simplified circuit diagram of a pulsatile amplifier and two simulations demonstrating its behavior in well-mixed and diffusive environments. Pulsatile communication requires only ephemeral investment in the protein components and signal molecules used in cell-cell signaling, making it a relatively low-burden approach to coordination compared to bistable or oscillatory circuits. When designing a multi-strain bacterial device, or any consortium that may experience a spatially heterogeneous front of chemical signals, these features are necessary for reliable, long-distance cell-cell signaling. This system would enable cells to initiate and transmit communications without permanently altering their internal transcriptional state or spontaneously generating signals at regular intervals.

2.2 Characterizing genetic components in liquid culture

Lux-type quorum sensing systems include a transcription activator (receptor), an acyl-homoserine lactone (AHL) signal molecule[1], and an enzyme that

[1]This abbreviation is used to mean the cognate signal molecule of the quorum sensing molecules stated in context.

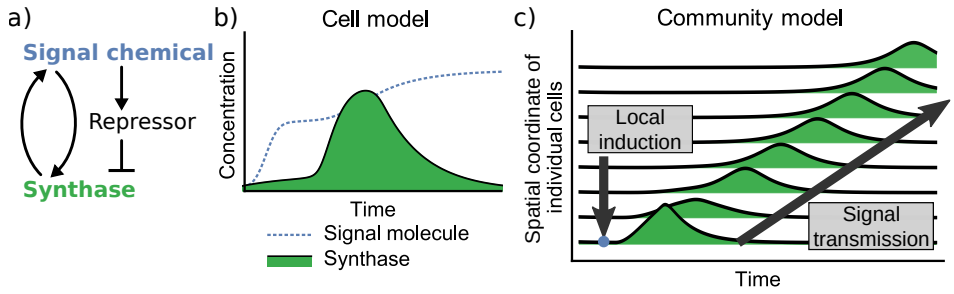

Figure 2.1: Abstracted interaction map of the species comprising the pulsatile gene circuit as well as two simulations of its behavior, one in well-mixed conditions and the other in a diffusive environment. The circuit schematic (a) depicts the positive-feedback relationship between the signal chemical and its synthase as well as the negative feedforward arm connecting the same signal chemical to a repressor that shuts off synthase production. In this model of a pulsatile gene circuit, a large expression delay is included that is not a part of the other models described in this document. This delay was included to make the spike-in signal chemical and cell-generated signal chemical distinct bumps in the under the "Cell model" figure, (b). The "community model" in (c) represents a scenario in which a 1D region of cells, each expressing the pulsatile circuit, is perturbed by a local addition of signal molecule. The ensemble effect of each cell amplifying local AHL concentration results in a traveling pulses of signaling activity that moves away from the perturbation where they initiated.

synthesizes the signal molecule (synthase). In these systems, the receptor binds to its cognate signal molecule and promotes transcription of the synthase to form a positive feedback loop. Cells expressing this type of gene circuit exhibit switch-like behavior between a low-output state and a high-output state. Stability analysis shows the low-output state to have a narrower margin of stability than the high-output state, which is stabilized by the structure of the feedback loop as well as local accumulation of AHL (Doelman (2019), Dockery and Keener (2001)).

To engineer pulsatile release of signaling molecules, we augment a Lux-type positive feedback system with a negative feedforward arm to eliminate persistence of the high-output state. [2] The combined regulatory arms form a well-studied architecture known as a Type-I incoherent feedforward loop (IFFL) that generates pulses in response to increases in the input (Goentoro et al. (2009), Alon (2006)). This negative regulation is accomplished by a transcrip-

[2]This circuit architecture, and implementations of it, is referred to as "pulsatile signaling circuit" and "amplifier circuit" interchangeably in this document.

tional repressor that, like the synthase, is regulated by the receptor. When exposed to AHL, the cell begins to express both synthase and repressor. The promoter controlling synthase expression therefore contains binding sites for the activating and repressing transcription factors. Soon after, the accumulated repressor protein prevents further expression of the synthase, which allows for a gradual return to a low-expression state. Repressor protein will continue to be expressed, however, as long as the AHL concentration is sufficient. The concentration of repressor must be reduced through degradation or dilution from cell growth before the circuit can generate a pulse of synthase expression again. This architecture elicits pulses in a similar manner as neurons: rapid excitation followed by a long refractory period (Holzer, Doelman, and Kaper (2013)).

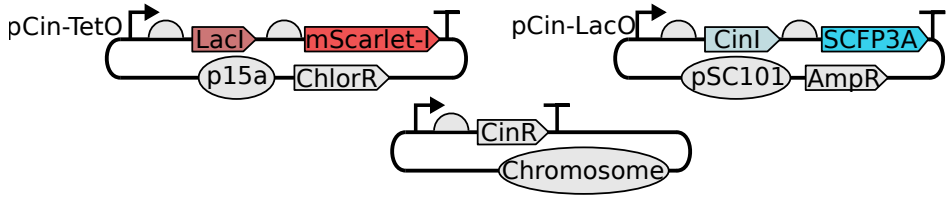

Figure 2.2: The genetic components of the pulsatile communication circuit. Quorum sensing components are drawn from the CinRI system from the species *Rhizobium leguminosarum*. The synthase protein, CinI, produces N-(3-hydroxy-7-*cis*-tetradecanoyl)-L-homoserine lactone (referred to as AHL in the context of the Cin system for convenience) (Lithgow et al. (2000)). CinR is expressed constitutively, AHL-bound CinR promotes expression from pCin, and LacI represses CinI transcription when bound to the LacO site. AHL freely diffuses through cell walls.

IFFL gene circuit generates pulses of synthase protein

The positive and negative regulation arms were cloned into different plasmids. CinI, the synthase protein used in this circuit, was cloned onto a low-copy pSC101 vector and the repressor, LacI, onto a p15a backbone. By dividing the circuit between two plasmids, a larger pool of strain candidates could be generated via combinations of a handful of plasmid candidates. As shown in Figure 2.2, co-transcriptional fluorescent reporters were appended to both the LacI and CinI coding sequences. Including these co-transcriptional reporters provides an unique marker of transcription activity for the two circuit promoters. In all experiments, plasmids were transformed into CY026, an *E. coli* strain that constitutively expresses CinR (Chen et al. (2015)).

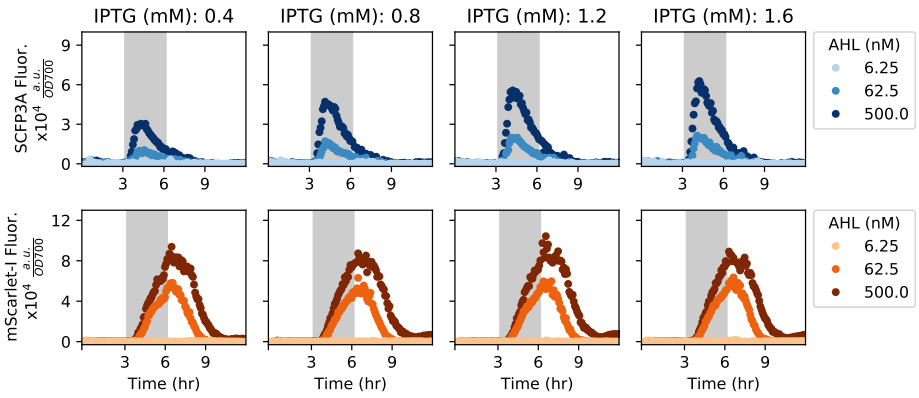

Figure 2.3: A series of experiments were performed wherein one inducer, IPTG, was kept at a constant concentration while another inducer, AHL, was introduced and then removed. The dark gray region indicates the period of time when AHL was included in the growth media. OD-normalized fluorescence data from these experiments. This figure only shows data collected from the candidate strain that demonstrated pulsatile response to AHL and showed the fastest doubling time.

We performed liquid culture *in vivo* experiments screening plasmid variants that differed in ribosomal binding site sequences 5' of the LacI and CinI coding regions. Two rounds of characterization were performed to optimize circuit behavior. In the first round, strains harboring repressor plasmid candidates were subjected to an induction ladder of the signal chemical. Plasmid candidates that did not fluoresce or that significantly reduced doubling time were removed from further consideration. In the following round, both repressor and synthase plasmid variants were co-transformed and strains were subjected to a grid of conditions varying in both AHL and IPTG concentrations. Candidates in this round of characterization were screened for pulsatile expression of the synthase-associated fluorescent reporter and robust cell growth. The behavior of the best-performing strain is shown in Figure 2.3. The data in this figure show that the IFFL architecture produces pulsatile synthase expression only in the presence of both AHL and IPTG. Growing cells expressing the pulsatile amplifier circuit do not produce fluorescent protein when grown in high-IPTG media, indicating that leaky expression does not result in spurious activation of the positive-feedback loop, thus generating a spurious signal, or the negative feedforward arm. During the period of AHL induction the synthase reporter shows a pulsatile response while the repressor reporter increases unceasingly. Following stimulation, when cells are grown in low-AHL and high-IPTG media,

there are no subsequent pulses or increases in the synthase reporter trace. The selected candidate shows well-tuned pulsatile behavior in response to exogenous AHL.

IPTG tunes synthase pulse amplitude but not repressor expression Both inducers, AHL and IPTG, are necessary for significant expression of the synthase co-transcriptional fluorescent reporter. This suggests that, before inducers are added, cells have accumulated enough LacI to block AHL-activated expression of the synthase. Alleviating LacI-mediated repression by including IPTG in the growth media allows for modulating the pulse height of synthase expression, a useful feature for experiments involving cell-cell signaling. Figure 2.4b) shows that, without added IPTG, no synthase expression is permitted.

While the circuit succeeded in eliciting a pulse of expression, the data did not show a clear relationship between amplitude of the synthase pulse and an increase in AHL concentration in the growth media. Activity from the LacI cistron was not highly correlated with IPTG concentration, meaning that synthase expression did not lead to increased AHL-mediated gene expression. Figure 2.4a) shows the impact of varying IPTG concentration at three AHL concentrations. At 0 nM, IPTG correlates with higher endpoint fluorescence ($\rho = 0.834$), though its effect is relatively negligible. Sensitivity to changes in AHL concentration should be greatest near a promoter's half-maximum induction concentration (IC50). The half-maximum AHL induction concentration of the negative cistron's promoter is roughly 40 nM. Even if the contribution of AHL from the amplifier cells were small, such as on the order of 5 nM, that would likely result in a detectable change in mScarlet-I fluorescence. However, unlike the simulation shown in Figure 2.1, this implementation of the amplifier circuit failed to noticeably increase AHL concentrations. Above the twice the IC50 concentration, promoters are less sensitive to changes in inducer concentrations but may still result in detectable changes in the accumulated fluorescence after a long enough time. The fact that the endpoint fluorescence from high-IPTG samples fall within the range of values from no-IPTG samples suggests that, even when activity synthase activity is driven to its maximum, the change in AHL concentration remained constant over the course of the experiments. The short pulse width, well-mixed media, and low cell density during synthase expression act to minimize signal retention within the cells. While chemical quantification procedures could provide insight on the absolute

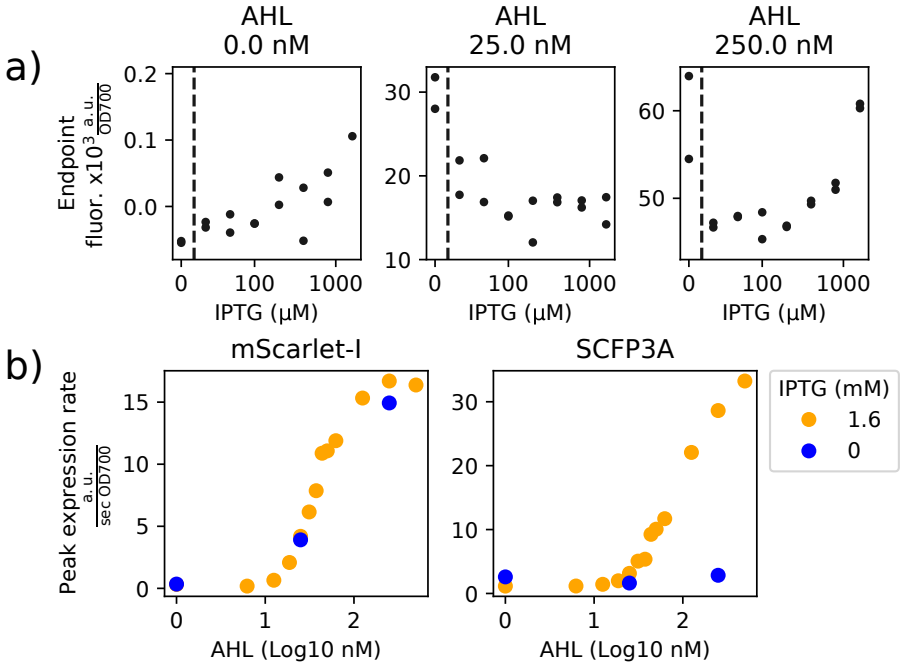

Figure 2.4: (a) mScarlet-I endpoint fluorescence from experimental wells varying in AHL and IPTG concentration. Fluorescence traces are normalized by OD700 and endpoints were taken to be the median value from timepoints with OD700 greater than 0.9. The x-axes are logarithmic in scale, the dashed vertical line separates the logarithmic scale from the zero IPTG concentration data. Pearson's correlation coefficients for data presented in each plot are 0.834, -0.369, 0.564, respecting plot order. (b) Maximum expression rate values from both mScarlet-I and SCFP3A channels at varying AHL and IPTG concentrations. These correspond to expression activity from repressor and synthase arms of the transcriptional network, respectively, and the data show that IPTG is necessary for synthase activity and that IPTG does drastically impact repressor expression.

amount of signaling molecule secreted by the amplifier cells during an experiment, the purpose for this strain is to facilitate AHL-based communication between cells (Yates et al. (2002)). When applied to well-mixed conditions, the amplifier cells were not be able to amplify their own response.

2.3 Extending signaling distance in semi-solid media

This chapter describes the time-lapse microscopy experiments performed to investigate the capability of the amplifier strain to support cell-cell signaling over long distances for consortia growing on semi-solid media. While the amplifier strain appeared unable to support cell-cell signaling within well-mixed liquid cultures, cells have greater leverage over their local chemical environments in unmixed conditions. In these diffusive environments, transport of small chemical species is achieved only due to diffusion, and therefore secreted chemicals accrue more rapidly in comparison to liquid media with convection-mediated mixing. Wild microbes in diffusive environments transmit chemical signals over long distances by the combined action of positive feedback loops expressed by dense communities of cells (Goldbeter (2006), Noorbakhsh et al. (2015), Larkin et al. (2018), Prindle, J. Liu, et al. (2015)). Here, we apply the pulsatile amplifier circuit characterized in Section 2.2 in an attempt to recreate the same emergent behavior.

Recent work investigating the difference between diffusive waves generated by passive or active propagation of chemical fronts shape our expectations for the pulse amplifier circuit. Passive propagation is the case in which sender agents are the sole source of signal chemicals. The diffusive wave is shaped by the diffusive environment, production rate, and degradation rate of the signal chemical. Both theoretical and experimental studies show that passive propagation can result in either constant-velocity fronts, diminishing-velocity fronts, or asymptotic fronts that approach a steady state concentration profile (Reátegui et al. (2017), Dieterle et al. (2020), Gines et al. (2017), Alon (2006), Wolpert (1969)). Adding a mechanism for active front propagation, often a bistable positive feedback loop, increases transmission velocity and yields a fixed-velocity diffusive wave when the release rate of the amplifiers is sufficient (Gines et al. (2017), Dieterle et al. (2020), Cheng and Ferrell (2018)). However, the scenario of interest here is complicated by the logistic growth and time-varying protein expression capacity of cells in the synthetic bacterial consortia. It is difficult to predict given the data from liquid culture characterization what

impact the amplifier strain will have on the diffusive wave dynamics, given the difficulty in analyzing the scenario mathematically.

Quantifying the extent to which the amplifier circuit extends communication distances requires as direct a comparison as possible between indicators of local AHL concentrations through the semi-solid media. That is accomplished by composing two consortia, one composed of a signal source and signal reporter and the other composed of a signal source and the signal amplifier strain. Each consortia is applied to the same scenario: the sender population is localized to one end of a long, rectangular agarose pad and its partner strain and its partner strain occupies the complementary region of the pad. After inoculating the surface of an agarose pad, the inducer chemicals embedded in the agarose stimulate the sender component to produce synthase protein and release signaling molecules. Following this, the signaling molecules diffuse to the partner strain, which is either the reporter or amplifier strain. If the amplifier cells can appreciably increase local signaling molecule concentrations, then the sender-amplifier consortium should produce a larger activated region of the pad than the sender-receiver consortium. Because the amplifier cells strictly increase AHL amounts, it is expected that the sender-amplifier consortium will elicit a higher degree of activation than the sender-reporter consortium.

Here, the role of a reporter component in a consortium is to produce fluorescent protein in response to signal molecule. Consider the amplifier strain's two fluorescent proteins, mScarlet-I and SCFP3A. Transcription of SCFP3A is controlled both by AHL-mediated activation and LacI-mediated repression, therefore it is not a reliable reporter of local AHL concentrations. mScarlet-I, on the other hand, is under singular control by AHL-activated expression and can be used as an AHL reporter. It would be convenient if the fluorescent reporters in the sender-reporter and sender-amplifier consortia could be assumed to have identical activity as a function of AHL.

Fortunately, the sender-amplifier and sender-reporter consortia can be formed from the same pair of engineered strains. As demonstrated in the previous section, IPTG is required for synthase expression yet does not impact mScarlet-I expression (see Figure 2.4. When deployed to pads lacking IPTG, the amplifier cells cannot signal to their neighbors and therefore take the role of signal reporter. This is an ideal scenario, as the genetic and kinetic details of transcription factor binding sites as well as the metabolic load on the host

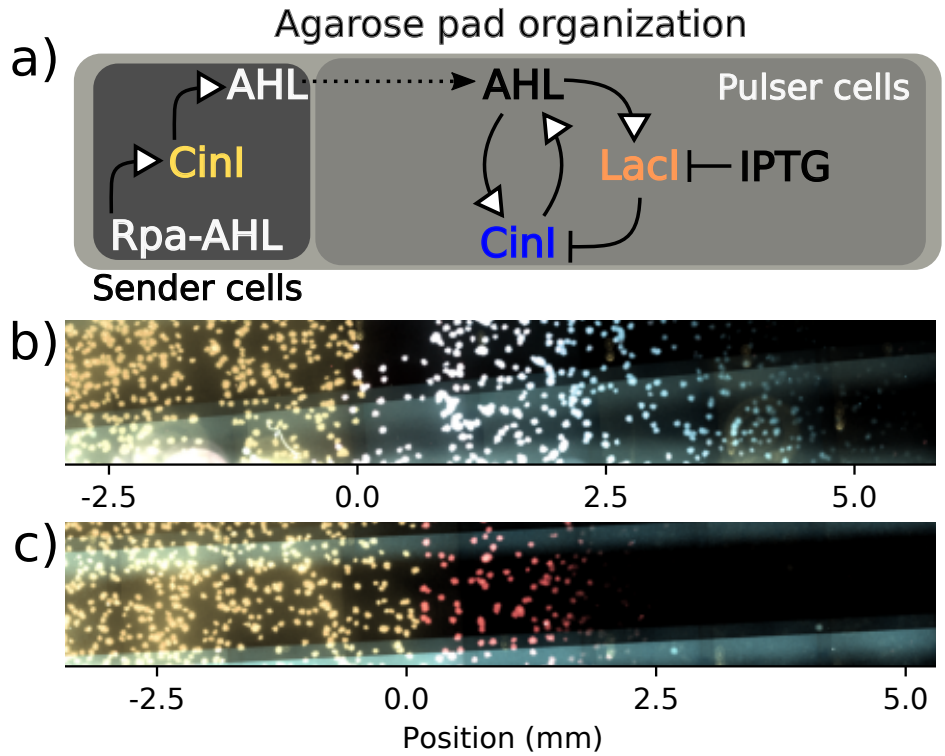

Figure 2.5: Schematic (a) summarizes the experimental setup described in the text, with shaded regions corresponding to the portions of the agar pad occupied by the sender and pulser cell strains. This subfigure also depicts abbreviated transcriptional networks of the sender and pulser strains. Subfigure (b) includes an image from a time-lapse microscopy experiment of a sender-amplifier consortium. Subfigures (b-c) are fluorescence micrographs where yellow indicates sender activity, red indicates repressor activity, and blue indicates synthase activity. Subfigure (c) is a sender-receiver consortium from the same experiment imaged at the same time, 4 hours after pad inoculation. These images show the initiation of cell-cell communication, when sender cells have secreted sufficient AHL to elicit a response from their consortium partners. Amplifier cells fluoresce in both SCFP3A and mScarlet-I channels, which can appear white in the composite image. Position values are relative to the interior end of the sender population.

cell from propagating transgenic plasmids are kept constant between the two consortia.

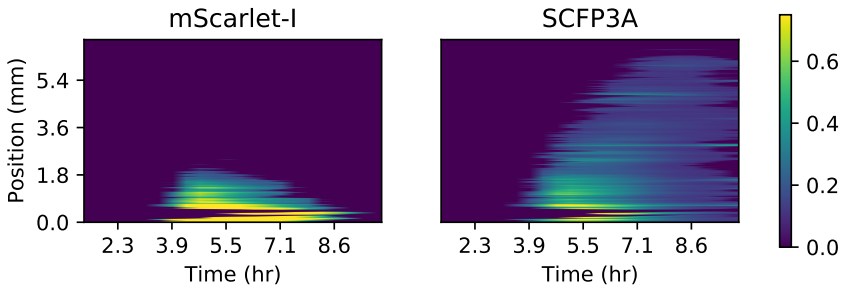

Figure 2.6: The two plots depict mean fluorescence values along the sender-amplifier consortium at each acquisition time. Images were first rotated such that the agarose pad's long axis was parallel to the image's column axis. Mean fluorescence values were determined by averaging the pixels above a threshold value at each image column. Stacking these vectors for each time point results in the heatmaps shown.

During agarose pad time-lapse microscopy experiments, cells are sandwiched between the agarose pad and a glass-bottomed dish. A depiction of consortium organization is shown in Figure 2.5. This restricts their movement while the signaling molecules they secrete can diffuse through the aqueous growth media contained in the pad. As a result, a small group of signal-secreting cells can locally accumulate a significant concentration of signal molecule more easily than in a well-mixed media of the same volume. The geometry and organization of cells on the agarose pads were both selected to maximize the rate at which sender cells locally accumulate signaling molecule. The rate at which diffusion draws signal chemicals away from their origin is significantly increased with each dimension permitting transport (Dieterle et al. (2020)). By isolating the sender cell population to one end of a pad, diffusion mainly acts to transport molecules in one direction: towards the interior. This reduces the impact of diffusion and leads to more rapid accumulation.

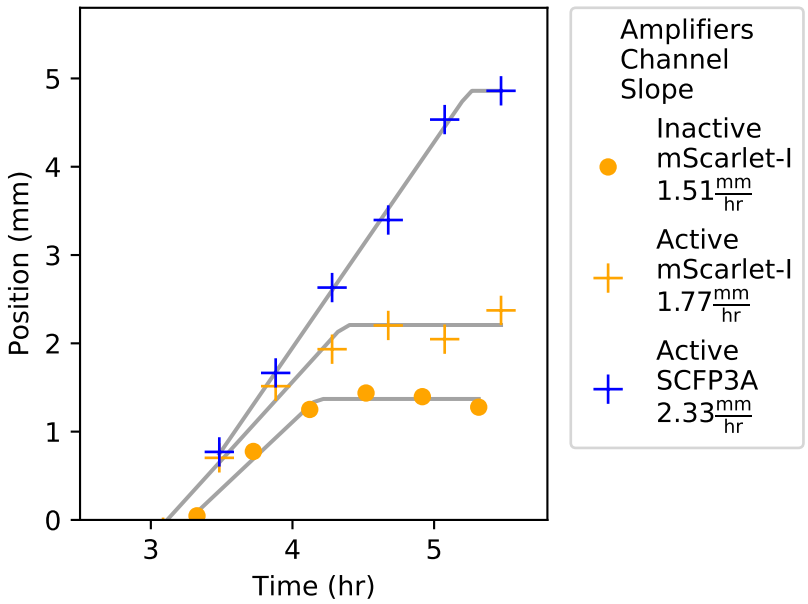

Figure 2.7: This plot shows the progression of SCFP3A and mScarlet-I fluorescence profiles recorded along agarose pads organized as in Figure 2.5. To determine the position of the signaling front, a threshold was applied to the fluorescence values at each frame. The points depicted in the plot correspond to the position farthest from the sender population that surpasses the threshold value. Threshold values were selected for each experiment to best capture the movement of the signaling front. Best-fit lines of the form $\mathrm{Min}(at+b, c)$, where t is time, were fit by least squares. Estimates of propagation velocity (a) and maximum propagation distance (c) taken from these best-fit lines.

A depiction of consortium organization is shown in Figure 2.5. Sender cells are the CY026 strain of *E. coli* co-transformed with two plasmids. One plasmid is a clone of the amplifier strain's negative regulation plasmid but with a large deletion in the LacI CDS. This loss-of-function deletion preserves the genetic context surrounding the AHL-associated promoter and the downstream reporter CDS in order to preserve its behavior as a reporter. Similarly, the other plasmid was designed to resemble the amplifier strain's positive feedback plasmid. It also uses a pSC101 backbone and codes for a bicistronic transcript of CinI and a reporter protein, but the reporter is sfYFP and the promoter is pRpa, a quorum-sensing promoter from the Rpa operon of the species *Rhodepsuedomonas palustris*. This sender plasmid also contains a constitutive source of RpaR, the receptor for p-courmaroyl-HSL (Rpa-AHL).

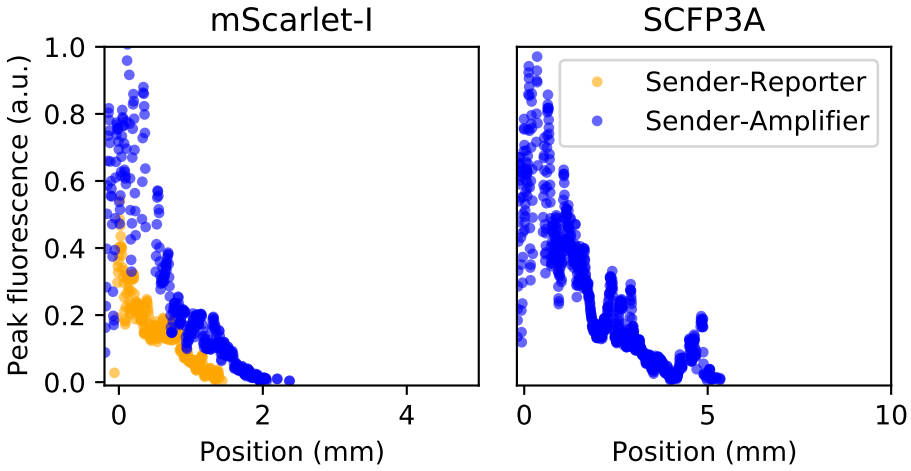

Figure 2.8: Peak fluorescence values were determined by calculating the average above-threshold fluorescence value at each distance and time, in the same manner that produced the heatmaps in Figure 2.6, then finding the maximum value attained at each distance over the course of the experiment.

Pulsatile signal amplification extends communication distance along one-dimensional, segrated sender-receiver consortia

Each time-lapse microscopy experiment includes four agarose pads mixed with a different set of inducer chemicals. The inducer sets are the four combinations of the presence or absence of Rpa-AHL and of IPTG. Rpa-AHL is used to activate synthase production within the sender cells and IPTG is used to enable the amplifier circuit in the amplifier cells. Table 2.1 shows the state of the sender and amplifier cells under each condition. Pads lacking inducers should result in no fluorescence from any cell. Images from these pads provides data on the background fluorescence from the agarose pad itself and the dividing cells. Pads with only IPTG included test the stability of the pulse circuit's "off" state, when sender cells are absent or inactive. While sender cells are inactive, it is not uncommon for asymmetric plasmid distribution during cell division to result in daughter cells lacking the negative regulation arm of the amplifier circuit, reducing it to a positive feedback loop around the synthase protein (Halleran, Flores-Bautista, and Murray (2019)). Cells lacking the negative regulation plasmid quickly lose cytosolic LacI, at which point they can rapidly switch to a high synthase expression state. This experimental control is necessary to monitor the effect of these errant positive feedback loops; spurious pulses generated by mutants would indicate that the signal amplifier is unreliable.

With only Rpa-AHL, senders are spurred to transmit signaling molecules to the inactive amplifier cells. Unable to produce synthases themselves, the receiver cells' repressor arm reports on the local abundance of signaling molecules generated by the sender cells. The propagation of this fluorescence front provides negative control of passive signal transport.

Table 2.1: Agarose pad experimental conditions

Condition	Senders	Amplifiers	Inducers included
Autofluorescence blank	Inactive	Inactive	None
Sender leak and amplifier self-activation	Inactive	Active	Rpa-AHL
Sender-Reporter	Active	Inactive	IPTG
Sender-Amplifier	Active	Active	Rpa-AHL + IPTG

Subfigures (b) and (c) in Figure 2.5 show false-colored images produced from the fluorescence micrographs of the sender-amplifier and sender-reporter agarose pads, respectively. From these images, it is clear that the sender cells on both pads can communicate to their partner strains that are nearby. Threshold-crossing points depicted in Figures 2.7 show that the region of mScarlet-I activation extends farther from the sender population in the sender-amplifier consortium than in the sender-reporter consortium. Fluorescence output is analyzed by manually selecting a threshold value and eliminating data that falls below this value. The position and average fluorescence of the response profile can then be determined by the above-threshold pixels. Figure 2.7 shows the positions of the leading edge of the traveling pulse over time for each experimental condition and channel.

Data from the sender-amplifier pad shows SCFP3A activity extending farther into the interior of the pad than mScarlet-I. This matches the observation from the liquid-culture experiments that mScarlet-I detection is delayed relative to SCFP3A. However, translation of the mScarlet-I front halts before the SCFP3A front. The promoters in the negative and positive regulation arms have similar AHL activation curves (see Figure 2.4); where amplifier cells can express SCFP3A, they should also be able to express mScarlet-I. Figure 2.6

shows the mean fluorescence from the sender-amplifier consortium along the direction of propagation. Expression in the mScarlet-I channel appears to halt abruptly after about 2.5mm of translation. Activation beyond that distance does not remain above the threshold value.

The time-lapse microscopy data show that the amplifier cells fail to maintain pulse amplitude over the propagation direction. Figure 2.8 shows the peak fluorescence value achieved at each position over the course of the experiment. In both channels, peak fluorescence drops continuously as signaling activity invades the center of the pad. The SCFP3A channel exhibits a linear trend while the mScarlet-I decline appears more similar to an exponential decay. Active amplifier cells significantly improve the response in the mScarlet-I channel relative to the sender-receiver consortium by increasing recruitment and by increasing peak response near overall. However, the diminishing amplitude and truncated mScarlet-I propagation relative to SCFP3A propagation through the amplifier population both suggest that, after a certain amount of time, the growth rate and protein expression capacity of cells on the agarose slows to the point that host cells can no longer contribute resources to transgene expression.

Increased sender emission rate extends sender-receiver signaling distance

Both sender-amplifier and sender-receiver consortia exhibited fixed-velocity propagation of signaling activity followed by an abrupt stalling and cessation of propagation. This raises the question of whether sender-receiver consortia can extend signal propagation without the aid of amplifier cells by simply increasing their signal production rate. To that end, a variant of the sender strain with a greater signal secretion rate was cloned. This strain was applied to a similar agarose pad organization as described above and the propagation of receiver activity was observed.

Figure 2.9 shows both a fluorescence micrograph of the experimental agarose pad as well as a plot depicting threshold-crossing events within the receiver cells. The threshold-crossing events show a fixed-velocity propagation that is similar in speed to that produced by the weaker sender strain. Its transmission distance, however, is much greater. This experiment shows that, along one-dimensional consortia and diffusive environments, sender cells can produce diffusive waves that are as rapid and travel as far as the sender-amplifier consortium.

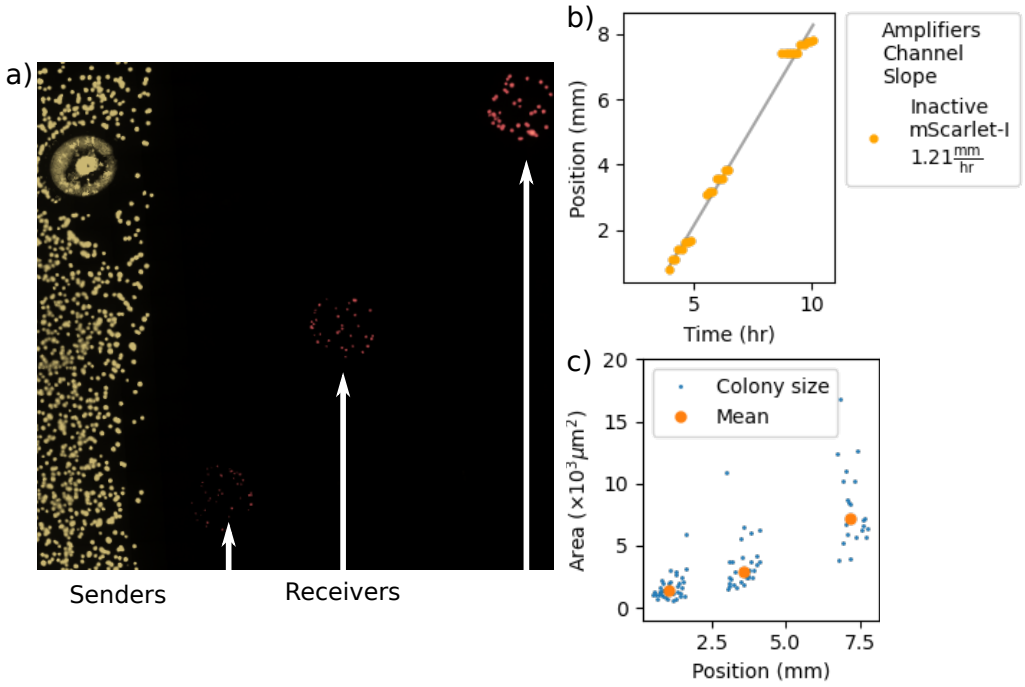

Figure 2.9: Measuring propagation distance and velocity from a stronger sender cell variant. (a) shows a fluorescence micrograph of sfYFP and mScarlet-I fluorescence. sfYFP indicates synthase expression within sender cells and mScarlet-I indicates AHL-induced expression within receiver cells. The sender population spans the width of the agarose pad while receiver cells were deposited at three different positions. White arrows direct to the positions of the deposited receivers. (b) depicts threshold-crossing events determined from fluorescence of the receiver cells. The three clusters of data points correspond to the three receiver depositions in (a). (c) shows the areas of each microcolony and mean areas at each deposition after 13 hours.

The microcolony sizes diminish with proximity to the sender population. Theoretical studies suggest that, in the case of a homogeneous and static consortium, pulsatile amplification produces a flat concentration profile over the region traversed by a propagated diffusive wave (Dieterle et al. (2020)). Persistent emission from a sender population, however, will result in an exponential concentration gradient over space. The microcolony area data depicted in Figure 2.9c) indicate that the AHL concentration near the sender population drives receiver cells to over-express LacI and mScarlet-I to the point that cell growth slows down. Receiver cells that are farther away are larger due to the decreasing AHL gradient a delay before the diffusive wave arrives. The growth penalty in the receiver strains shows that, while the diffusive wave is rapid and

long-ranged, the steep concentration gradient can be toxic to receiver cells.

Bistable amplifier strain

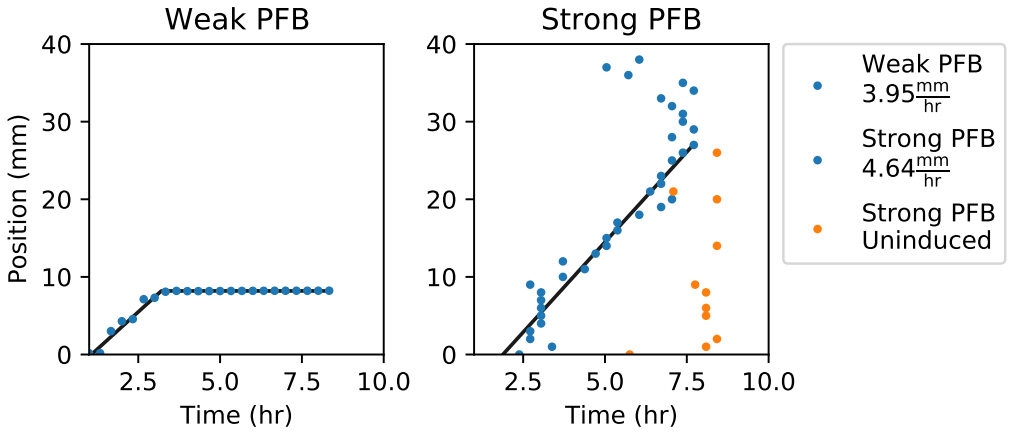

Figure 2.10: Comparing signaling propagation through bistable amplifier cells with different 5' UTR sequences controlling ribosomal binding to mRNA transcripts. Threshold-crossing events are grouped into 1mm-wide bins. The data points with smallest time value from each bin are plotted and used in fitting trend lines. Trend lines are piecewise, with a linear and flat components. Both "weak" and "strong" positive feedback circuits uniformly deposited over a long agarose pad. In control samples, cells were grown without inducer chemicals. In experimental samples, a 25nL droplet of 50uM C4-HSL was deposited at the 0 mm position and allowed to dry before imaging.

The pulse amplifier circuit limits the maximum expression rate through the promoter controlling synthase mRNA transcription by LacI-mediated repression and by the metabolic investment in LacI and mScarlet-I. It is well-known that signal production rate scales with the propagation velocity in diffusive waves (Dieterle et al. (2020), Oleinik, Kolmogorov, and Piskunov (2019)). Engineering an amplifier strain that propagates signaling activity more quickly could make use of a bistable rather than pulse-generating circuit by reducing the number of protein components and the removal of the transcriptional repressor.

A bistable amplifier strain was engineered to express a positive feedback loop controlling synthase production, similar to natural quorum sensing motifs. This circuit used components from the Rhl quorum sensing system from the species *Pseudomonas aeruginosa*. The AHL variety associated with the Rhl system is C4-HSL, a homoserine lactone with a four-carbon acyl chain (Pearson, Van Delden, and Iglewski (1999)). This molecule is much smaller than C14-HSL, the

AHL molecule associated with the CinRI system used in the pulsatile amplifier circuit, and therefore has a larger diffusion coefficient. While the transcription dynamics of the Rhl quorum sensing components are likely different from those of the Cin quorum sensing system, the lack of negative regulation on the synthase promoter, reduced metabolic load, and smaller AHL molecule could allow for an upper-bound estimate of the propagation velocity of AHL-mediated signal propagation along one dimension.

Figure 2.10 traces cell-cell signaling propagation through amplifier cells expressing either a "weak" or "strong" bistable quorum sensing circuit. The "weak" positive feedback appears to be insufficient for generating traveling waves and reaches only ~10mm into the interior of the pad. The "strong" circuit, however, shows a fixed-velocity traveling wave of signaling activity that traverses ~30mm. Data from the uninduced control of the "strong" positive feedback circuit shows a synchronous self-activation at around 8 hours, demonstrating that the low-expression state becomes unstable as cell density increases. A spurious signal can be seen in the induced case as well, initiating at position 40mm and 5 hours and moving towards the interior, eventually meeting the induced signaling wave at position 30mm and 7 hours. These results suggest a possible velocity upper-bound to AHL-mediated signaling propagation over one dimension and highlight a difficulty in engineering feedback within growing synthetic bacterial consortia.

2.4 Methods

Liquid culture experiments were performed in a Biotek Synergy H1F plate reader using M9CA minimal media (Teknova product code M8010-06) with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. Starter cultures were inoculated into M9CA media from single colonies picked from an agar plate. Inoculated cultures were shaken and incubated at 37°C until the optical density at 700 nm measured ~0.3. The experimental wells were then prepared by diluting starter cultures 1:20 into a final volume of 500μL in a 96-well glass-bottom plate. When necessary, inducer chemicals were added to wells using an Echo 525 acoustic liquid handler before the addition of cell culture and broth.

At two points during the liquid culture experiments, inducer concentrations were altered by diluting and washing experimental cell cultures. Washes were performed by first pelleting the full 500μL of culture from an experimental well,

discarding the supernatant, and then resuspending in 15mL of PBS. These wash-discard-resuspend cycles were repeated twice more for a total of three wash cycles. 500 μL of M9CA broth was used in the final resuspension step to return the culture to approximately the same density as before the wash steps. Subsequent experimental wells were prepared by first depositing inducer chemicals as needed with the Echo 525 liquid handler, followed by cell culture and sterile M9CA broth in a 1:20 ratio.

Analysis of data collected from plate reader experiments was performed using custom Python scripts. Background fluorescence and OD700 were determined by measuring these quantities in wells prepared without inducer chemicals, which previous experiments had shown to be identical in fluorescence to cells without fluorescent reporters. These background signals were determined to be time-varying, resulting from either background fluorescence from growing cells or from broth oxidation. OD700-normalized fluorescence values (fluorescence divided by OD700) were calculated using background-subtracted data.

Microscopy experiments were performed using agar pads prepared according to the protocol described in (Young et al. (2012)). Pads were prepared with inducers introduced to the molten agar, when necessary, and 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. When added to molten agar, final IPTG concentration was 1 mM and final C4-HSL concentration was 10 μM. Images were acquired using an Olympus IX81 inverted microscope through a UPlanFl10XPh objective and Chroma filters 31040, 41027, and 310442V2 for sfYFP, mScarlet-I and SCFP3A, respectively. Sample temperature was held at 37°C for the duration of the time-lapse microscopy.

Analysis of data collected from microscopy experiments was performed using custom Python scripts.

Chapter 3

SIGNAL PROPAGATION FROM ISOLATED SENDER CELLS

3.1 Introduction

The previous chapters considered signal amplification in a context that is unlikely for a consortium of engineered bacteria to encounter outside of a man-made arena: segregated sender and amplifier populations confined to a narrow, rectangular surface. A goal of synthetic consortia research is to deploy synthetic bacterial communities to natural environments such as the mammalian gut, soils, and rhizosphere. Introducing synthetic cells to these spaces could enable long-term monitoring of the chemical and biological agents (Mimee et al. (2015), Hsiao et al. (2016), Daeffler et al. (2017)) or intervention in support of human health (Steidler, Rottiers, and Coulie (2009), Christian et al. (2014)) or crop yield (Voigt (2020), Jin et al. (2020), Avis et al. (2008)). Even when synthetic colonies can persist and maintain their individual engineered functions in the face of competition from the natural microbiome, the physical organization of consortium members cannot be determined ahead of time. The growth context of cells in these settings are porous or high-viscosity media that constrain cell motility and spatial partitioning that results in patchy colonization (Donaldson, Lee, and Mazmanian (2015), Gantner et al. (2006)). While the spatial arrangement and density of soil bacteria that interact with root systems, a context known as the rhizosphere, is the subject of active research, *in situ* imaging of synthetic cells associated with the rhizosphere and with roots directly depict clumped or sparse cell distributions, where cells either are in closely-packed groups of tens to hundreds or separated by distances much longer the length of a single bacterium (Gantner et al. (2006), Tecon et al. (2018), Bulgarelli et al. (2012), Nunan et al. (2006)). Bacteria inhabiting the mammalian gastrointestinal track likely form long-term communities by invading the mucus which protects intestinal and colonic epithelial cells, those that are not quickly washed out with feces, that is (Donaldson, Lee, and Mazmanian (2015), Berlec et al. (2015)). Generally, there is no guarantee that strains in sender-receiver pairs will be close enough to enable efficient communication.

The geometric characteristics of the agarose pads and sender-receiver consortia described in Chapter 2 were designed to support cell-cell signaling activity and observation of the diffusive wave dynamics. The spatial scales of bacterial communities in deployment environments are much smaller than the consortia of these propagation experiments. The aim of this chapter is to evaluate the performance of the same consortia in a patterning scenario that is more adversarial to cell-cell signaling, the case of sparse sender cells in a small colony of dense amplifiers. By growing consortia from a smaller seeding population with a relatively larger diffusive environment, the role of diffusion in diminishing local signaling concentrations is more pronounced. Analytical results suggest an approach to nondimensionalization that normalizes the spatiotemporal pattern of gene activity to the size and signal secretion rate of the sender population. This approach is used to normalize data from each experimental sample for comparison of the signaling behavior between consortia.

3.2 Amplifying communication from a single sender cell

The efficacy of cell-cell signaling via diffusive molecules depends on the geometry of the diffusive environment, the volume that permits diffusion of molecular species and its characteristics, and on the characteristics of the wave generated by the microbial consortium. In time-lapse microscopy experiments, for example, cells live in a thin aqueous layer between an agarose pad and an impermeable glass plane. The experiments described in the previous chapter involved cells spread uniformly over the surface of agarose pads fashioned to be extended in length and short in width and height. The uniformity across with pad width ensures one less dimension along which diffusion can occur, thus encouraging cell-cell communication. Under the assumption that diffusion is fast enough that signaling molecules do not form a significant gradient along the pad height, the concentration of signaling molecules in the agarose pad takes on a gradient only over the pad's length, making the diffusive environment and the diffusive wave both one-dimensional. Greater differences between the dimensionality of the unoccupied environment and that of the diffusive signaling wave imparts a greater obstruction to cell-cell communication (Dieterle et al. (2020)). The case described here, therefore, is the most supportive of cell-cell signaling via diffusive waves.

The experiments described in this section apply a different geometry to present the consortium cells with a diffusive environment more adversarial to cell-cell

signaling. Rather than deposit cells uniformly on the surface of the agarose pad, a single 25nL droplet is placed at the center of a square pad. As a result, the consortium cells form a disk of roughly 0.7mm in diameter on the surface of an agarose pad that is 4.5mm to a side and 2mm tall. The seed population contains on the order of 100 cells with a roughly 1:100 ratio of sender cells to amplifier or reporter cells. Under these conditions, there is no symmetry that ensures flat concentration profiles along any axes. Therefore, diffusion dilutes local signaling concentrations faster than in the extended-length agarose pad experiments. Figure 3.1 contains a schematic of the cell organization on the agarose pads as well as composite images of the consortia at various times.

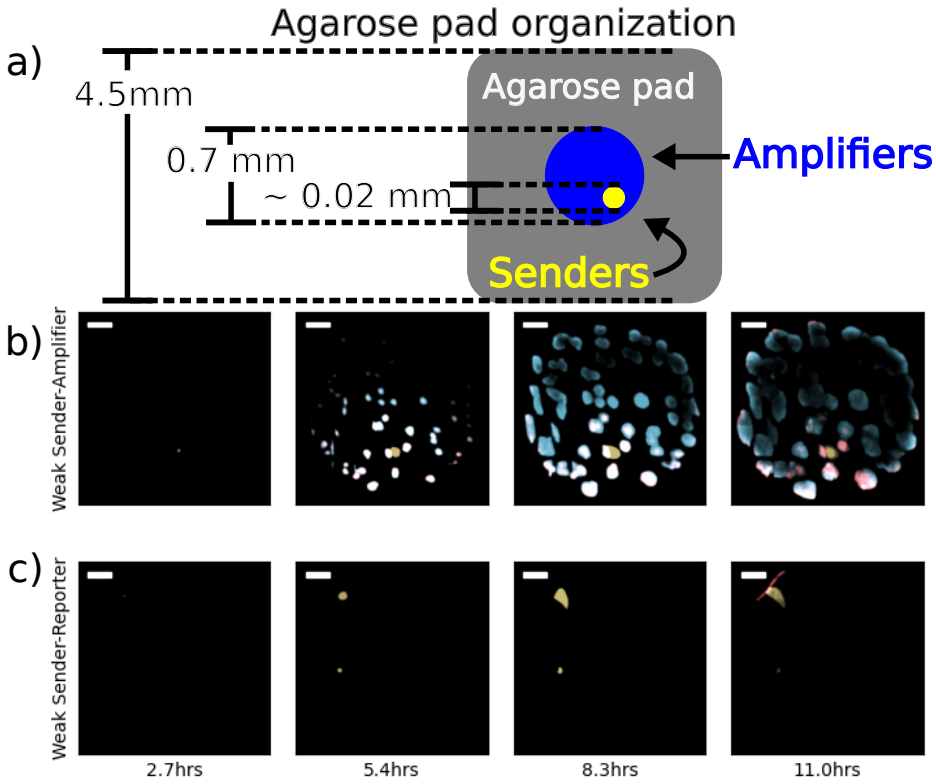

Figure 3.1: Fluorescence microscopy images of sender-amplifier consortia from two different agarose pads. The cartoon in (a) shows the size of the agarose pad and the consortium growing on it. The scenario depicted is one including a single sender population. Subfigures (b) and (c) show selected frames from experiments on sender-amplifier and sender-reporter consortia, respectively. The "Active" pad contains IPTG, thus enabling amplifier activity, while the "Inactive" pad does not. Each image is a composite of sfYFP, mScarlet-I, and SCFP3A expression. Scale bar is 100μm.

The sender-amplifier and sender-reporter consortia were applied to this scenario.

Performance of each consortium was evaluated by two metrics: 1) how quickly a sender cell can trigger fluorescent protein production in its neighbors, and 2) the fraction of cells that have increased in fluorescence by the end of the observation period. These metrics are respectively termed response time and recruitment fraction. As before, the same pair of *E. coli* strains can be assayed with and without pulsatile signal amplification by the presence or absence of IPTG (see Section 2.3). While data from both the mScarlet-I and SCFP3A channels are included, only amplifier cells on pads including IPTG can produce SCFP3A. Otherwise, these cells are called "reporters" and can only express mScarlet-I. Figure 3.2 shows the response times for each condition and Figure 3.3 shows the recruitment fractions.

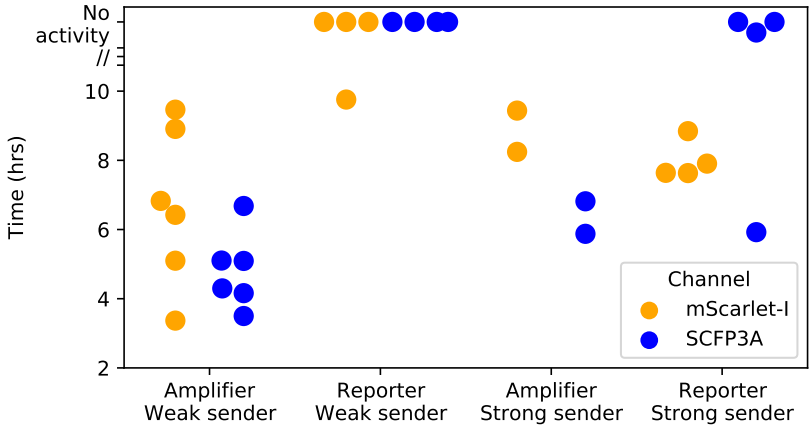

Figure 3.2: Initiation time is the duration from beginning of experiment to first threshold-crossing event from amplifier cells. These plots show initation time values calculated for each consortium and fluorescence channels. mScarlet-I fluorescence indicates repressor expression and SCFP3A indicates synthase expression. In the "Reporter/Strong sender" condition, there is a data point showing a SCFP3A threshold-crossing event at roughly 6 hours. This is the result of a putative mutant microcolony in the amplifier strain that resulted in persistent SCFP3A expression. The fact that only one such microcolony was observed out of the many sender-reporter consortia replicates suggests this is a rare event.

Comparisons between values associated with mScarlet-I fluorescence suggest that pairing amplifier cells with a "weak" sender can dramatically improve both response time and recruitment fraction relative to the case of passively signal chemical transport. The consortium containing the "strong" sender, on the other hand, show similar response times and recruitment fractions between

active and passive signal propagation. However, those consortia with weak sender cells and active signal propagation exhibit the fastest responses and greater recruitment than all other experimental conditions.

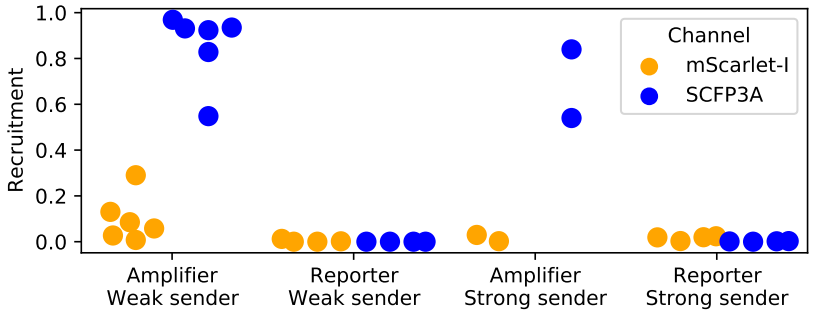

Figure 3.3: Recruitment fractions are calculated by dividing the above-threshold area by the total cell-occupied area. The area occupied by cells is determined using a low threshold on the SCFP3A fluorescence data that separates agarose autofluorescence and bacterial autofluorescence. As mentioned in the caption to Figure 3.2, one of the "Reporter/Strong sender" replicates produced a putative mutant amplifier microcolony that exhibited strong SCFP3A fluorescence. These data show that this mutant did result in a significant recruitment fraction.

Deducing the reasons for these differences in response time and recruitment fraction is confounded by the differences in the growth patterns of each consortium. These considerations include the number and arrangement of the founding population and how the microcolonies impact one another's growth rates and protein expression capacity. As can be seen in Figure 3.1, growing microcolonies can interfere with each others' growth by crowding. Furthermore, after 10 hours of growth it appears that only cells near microcolony edges can continue to express fluorescent protein, perhaps as a result of limited nutrient availability in the microcolony interior. To enable direct comparisons between experimental samples in a way that accounts for these details, a mathematical model was developed to formalize the expected diffusive signaling front from an isolated sender colony, considering size and growth rate.

3.3 Mathematical modeling of chemical front propagation from isolated sender cells

We will consider a disk of sender cells of radius r that is sandwiched between an impermeable boundary and a porous medium of infinite height and radius.

At time $t = 0$, the sender cells are induced to express synthase protein. The cells are assumed to reach equilibrium cellular concentrations of synthase immediately, allowing for the total amount of synthase to be proportional to the cell number. Therefore, synthase increases exponentially with rate constant λ. The production rate per unit area is therefore

$$f(x,t) = c_0\Theta(r-x)\Theta(t)e^{\lambda t}, \tag{3.1}$$

where c_0 is the AHL emission rate at the moment of induction and Θ is the Heaviside function. These features closely resemble the experimental scenarios, to a point. Once amplifier cells secrete AHL or nutrient limitations attenuate cell growth and protein expression, the model assumptions no longer describe consortia behavior. However, we will see that the diffusive waves propagated in sender-amplifier consortia appear to match those predicted by the math model. This will enable us in Section 3.4 to compare consortia on the basis of their apparent sender strengths as represented by the emission rate parameter c_0.

We will use the method of fundamental solutions to solve for $a(x,t)$, the concentration of AHL at a point x at time t. We begin with the differential equation governing signal diffusion in the agarose pad, including a source term representing secretion from the sender population, the inhomogeneous heat equation

$$\partial_t a(x,t) = \nabla^2 a(x,t) + f(x,t),$$

where the rate of change in concentration is a combination of diffusion and production terms. The solution to the inhomogeneous heat equation can be found by convolution of the fundamental solution with the inhomogeneous term. Here the heat kernel g must respect the no-flux boundary condition. The placement of sender cells on the boundary provides a convenient symmetry: the heat kernel for an unbounded 3D domain is multiplied by a factor of 2, effectively superimposing the portion of the distribution that would extend past the boundary back onto available volume. This yields the function

$$G(x,t) = \frac{1}{4(\pi D(t-\tau))^{3/2}} e^{\frac{-p(x)^2}{4D(t-\tau)}},$$

where $p(x)$ represents the distance from the point x and the center of the kernel function. The convolution integral will use polar coordinates with the origin located at the center of the disk of sender cells and vary over time and the area of the disk to determine the concentration at the space-time

coordinate (x, t). Radial symmetry allows the position x to be restricted to $\{(R, \theta)|\theta = 0, R > r, R \in \mathbb{R}^+\}$ without loss of generality. To calculate the distance p at position x, we use

$$p(R, \theta, x) = \sqrt{(R\,\mathrm{Cos}(\theta) - x)^2 + (R\,\mathrm{Sin}(\theta))}\,.$$

Composing the convolution integral yields

$$a(x,t) = \int_0^{2\pi} \mathrm{d}\theta \int_0^{\infty} \mathrm{d}R \int_0^{t} \mathrm{d}\tau\, f(x,\tau) \frac{R e^{\frac{-p(R,\theta,x)^2}{4D(t-\tau)}}}{4(\pi D(t-\tau))^{3/2}}\ ,$$

which becomes

$$a(x,t) = \int_0^{\pi} \mathrm{d}\theta \int_0^{r} \mathrm{d}R \int_0^{t} \mathrm{d}\tau\, c_0 e^{\lambda t} \frac{R e^{\frac{-p(R,\theta,x)^2}{4D(t-\tau)}}}{2(\pi D(t-\tau))^{3/2}}$$

after substituting $f(x, \tau)$, restricting the integrals to non-zero values of the Heaviside function, and applying symmetry across the x-axis to reduce the integral over θ to the range $[0, \pi]$. Next, a change of variables $\psi_x(R, \theta)$ is applied to re-center the polar coordinates from the center of the sender population to the point x:

$$\mathbf{u} = (R, \theta),$$
$$\psi_x(\mathbf{u}) = (p(R, \theta, x), \mathrm{ArcTan}(\frac{R\,\mathrm{Sin}(\theta)}{R\,\mathrm{Cos}(\theta) + x}),$$

$$|\det(D\psi)(\mathbf{u})| = \frac{p(R, \theta, x)}{R}.$$

The result of this operation is a more tractable integrand and θ integration bounds that depend on R:

$$\int_{x-r}^{x+r} \mathrm{d}R \int_{\mathrm{ArcCos}(\frac{r^2 - x^2 - R^2}{2xR})}^{\pi} \mathrm{d}\theta \int_0^{t} \mathrm{d}\tau\, c_0 e^{\lambda t} \frac{R e^{\frac{-R^2}{4D(t-\tau)}}}{2(\pi D(t-\tau))^{3/2}}.$$

From here, the integrals are resolved individually, beginning with τ. The τ integral can be performed by applying a change of variables $\tilde{\tau} = t - \tau$ and results in a sum of complementary error functions

$$\int_{x-r}^{x+r} \mathrm{d}R \int_{\mathrm{ArcCos}(\frac{r^2-x^2-R^2}{2xR})}^{\pi} \mathrm{d}\theta \frac{c_0}{2D\pi} e^{\lambda t - R\sqrt{\frac{\lambda}{D}}} (\mathrm{Erfc}(\frac{R}{2\sqrt{Dt}} - \sqrt{\frac{\lambda}{D}}t) + e^{2R\sqrt{\frac{\lambda}{D}}} \mathrm{Erfc}(\frac{R}{2\sqrt{Dt}} + \sqrt{\frac{\lambda}{D}}t)) \quad . \tag{3.2}$$

At this stage, we apply estimates for the parameters r, D, λ, and the variable t. Figures 3.1 and 3.2, and knowledge of the AHL signaling molecules, suggest the following:

1. The earliest response occurs after roughly 3 hours.
2. Sending colonies are roughly 20μm in diameter.
3. x is bounded from above by 1mm.
4. $D \approx 200 \frac{\mu \mathrm{m}}{\mathrm{s}^2}$ (Dilanji et al. (2012)).
5. Synthase accumulation should roughly match cell growth, or $\lambda \approx 10^{-3}/\mathrm{s}$.

These approximations motivate the following simplifying assumptions:

1. $\frac{R}{2\sqrt{Dt}} \approx 0$
2. $\sqrt{\frac{\lambda}{D}}t \gg 0$
3. $\lambda t - R\sqrt{\frac{\lambda}{D}} \approx \lambda t$

With these assumptions, the sum of the Erfc terms within the parenthesis of Equation (3.2) evaluates to 2 and all dependence on R within the integrand disappears, leaving

$$\begin{aligned} a(x,t) &= \int_{x-r}^{x+r} \mathrm{d}R \int_{\mathrm{ArcCos}(\frac{r^2-x^2-R^2}{2xR})}^{\pi} \mathrm{d}\theta \ \frac{c_0}{D\pi} e^{\lambda t} \\ &= \frac{c_0}{D\pi} e^{\lambda t} \int_{x-r}^{x+r} \mathrm{d}R \ \ \pi - \mathrm{ArcCos}(\frac{r^2-x^2-R^2}{2xR}). \end{aligned}$$

Here we apply a lower-bound approximation

$$\pi - \mathrm{ArcCos}(\frac{\mathrm{r}^2 - \mathrm{x}^2 - \mathrm{R}^2}{2\mathrm{xR}}) \geq \mathrm{ArcCsc}(x/r)\frac{2}{\pi}$$

and ultimately recover

$$a(x,t) = \frac{4c_0}{D\pi^2} r \text{ArcCsc(x/r)} e^{\lambda t}. \tag{3.3}$$

As the sender colony radius is a fixed parameter in this case, we conclude from Equation (3.3) that front propagation occurs rapidly and concentration grows exponentially at all positions.

3.4 Analyzing nondimensionalized wave propagation

Equation (3.3) suggests a nondimensionalization of time and position according to the arguments for the trigonometric and exponential functions, x/r and λt. This enables comparisons between consortia in a way that reduces the impact of variability in sender microcolony position, size, and growth. The term c_0 can further be nondimensionalized by normalizing to the minimum activation concentration for each channel's promoter, forming the nondimensional term $\frac{4c_0 r}{c_{th} D\pi^2}$.

When applying the nondimensionalization, it is essential to also consider the assumption of exponential accumulation of synthase molecule. The sender cells' bicistronic expression of CinI and YFP allows for the identification of the period of exponential accumulation of CinI. Restricting curve fitting to only data from this period duplicates the results of Figure 3.2: reporter components in the "weak sender"-reporter consortia show almost zero response to signaling activity. The only sample out of four which shows any reporter activity is depicted in Figure 3.1.c), and fluorescence in this case began after the period of exponential CinI accumulation. This suggests that senders are emitting AHL at a very slow rate, possibly less than 3000 molecules per second per square micron, at peak CinI accumulation.

Consortia with active signal propagation, on the other hand, have best-fit lines for their SFP3A threshold-crossing events that largely fall atop one another, showing agreement in behavior despite the variability in the founding population. It is also notable that these best-fit lines appear to represent the data well, despite violating the model assumption restricting AHL emission to the sender colony. These estimates of c_0, therefore, can provide a metric for how quickly a consortium propagates signal fronts away from sender cells, with larger values implying a faster response. Figure 3.5 shows estimates from each threshold-crossing event as well as those correspond to the fitted lines for each experimental pad. These results highlight the benefit provided by

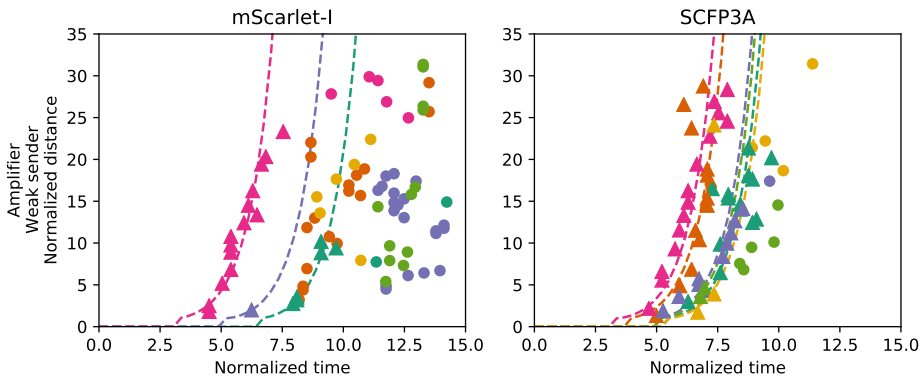

Figure 3.4: These plots depict threshold-crossing events using nondimensionalized position and time values along with best-fit lines to the propagation model. Triangular points correspond to threshold-crossing events that occurred while the sender microcolonies were rapidly growing and circular points represent data from after cell growth slowed. Only triangular data points were used in model fitting. Data from the same pad are colored similarly between the channel plots.

signal amplification in scenarios where the diffusive environment is adversarial to cell-cell communication within a consortium.

3.5 Methods

Cells were deposited in droplets onto agarose pads and allowed to dry. The droplets were either 1uL in volume and placed by pipette onto agarose pads 10mm by 10mm or 25nL in volume and transferred via acoustic liquid handler onto pads 5mm by 5mm. Once dried, the seeding colonies were roughly 2mm in diameter in the former case and 700um in the latter. The agarose pads were observed via fluorescence microscopy and incubated at 37 degrees Celsius.

Agarose pads contained chemical inducers that enabled different consortia behavior. In one set, pads contained both Rpa-AHL and IPTG to enable sender and amplifier activity. In the other set, pads contained only Rpa-AHL. The Rpa-AHL-only pads would allow amplifier cells to express mScarlet-I in response to signaling activity, but not contribute to it.

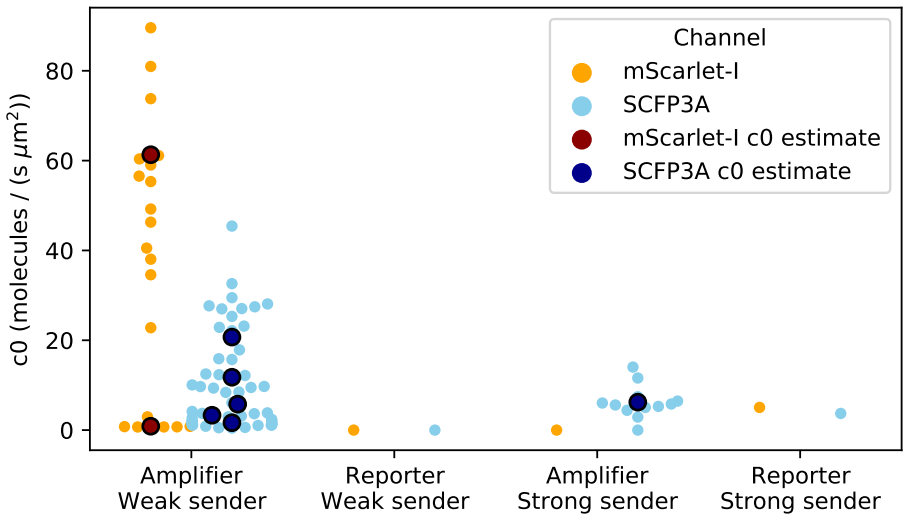

Figure 3.5: These data are estimates of initial signal emission rate density, c_0. The color dots correspond to estimates from an individual threshold-crossing event and white/black dots correspond to model fits from a single agarose pad as depicted in Figure 3.4. Channel-specific threshold concentrations were $10/\mu\text{m}^3$ and $20/\mu\text{m}^3$ for the promoters of SCFP3A and mScarlet-I, respectively, for model-fitting. Fits were not performed when there were fewer than three threshold-crossing events available.

Chapter 4

MODEL DEVELOPMENT

4.1 Introduction

The time-lapse microscopy experiments presented in Chapters 2 and 3 demonstrated that amplifier cells can support communication between sender-receiver signaling pairs. The question then becomes can amplifier cells benefit communication within more complex consortia? While the amplifier cells succeeded in increasing the signaling distance of sender cells, the data make clear that the organization of consortia members, environment geometry, and timing are critical to its performance. The timing can be broken down into three phases. In the first phase, sender cells grow and accumulate synthase protein. During the second phase, the consortia density has increased to the point that sender cells can initiate a propagating wave through the amplifier cells. This propagation ends at the third phase, when cell growth and protein production slow down such that the amplifier cells cannot produce enough synthase protein to sustain signal transmission. Cells in a consortium will transition between these phases at different times, compounding the spatially heterogeneity of the consortium's demographic patterning to the gene expression capacity.

This chapter walks through the development of a finite differences method (FDM) approach to simulating a reaction-diffusion model of sender-amplifier consortia. These simulations help to determine the applicability of amplifier strains to collaborative computations under the restrictions of the three phases outlined above. First, a set of ordinary differential equations (ODEs) describing the behavior of the amplifier strain in liquid culture is presented. While the goal is to simulate new engineered consortia in semisolid media, development begins with parameter identification on these ODEs using data from liquid culture experiments. These equations are adapted to a reaction-diffusion model and are supplemented with equations describing sender cell growth and signal release. While the gene circuit parameters inferred from liquid-culture experiments are maintained, microscope data are used to determine parameters governing nutrient-dependent growth and protein production. Finally, reaction-diffusion simulations of the sender-amplifier consortia are validated against experimental

data described in Chapters 2 and 3 and then extended to predict the behavior of complex hypothetical consortia.

4.2 Parameter inference from liquid-culture experiments

Well-mixed Media Model

The experiments described in Section 2 produced the the data used to infer gene circuit parameters. We here describe the ordinary differential equations model for pulse-generator circuit behavior in liquid culture. The intention is for the model to capture the observed fluorescence dynamics using as simple a model as possible, not to make mechanistic inferences about hidden cell state variables. This liquid-culture model will then be adapted to a spatiotemporal PDE problem, where any simplifying assumption made at in the liquid culture model will facilitate the PDE model's implementation and execution.

In the model, expression activation is modeled using Hill functions of the form

$$\mathcal{H}(A, \lambda, K) = \frac{A^\lambda}{K^\lambda + A^\lambda}.$$

In this form, the parameter K represents the IC50, the concentration of inducer chemical A which generates expression rate at half the maximal value, and the parameter λ controls how sharp the transition is from low to high output around the IC50 point. Lower values correspond to more gradual transitions and higher values. Protein quantity is also diminished through dilution as cells grow and divide. In the ODE, the protein concentration diminishes proportionally to the variable rate term

$$r_c(1 - \frac{c}{C_m}),$$

where c is the cell density, r_c is the maximal cell growth rate, and C_m is the carrying capacity of the amplifier strain. Co-transcribed proteins are expressed proportionally to one another, the proportionality being dependent on their genetic context and RBS sequence. Identical LVA-ssrA degradation tags were appended to the coding sequences of the synthase, repressor, and reporter proteins to promote equal degradation rates, represented in the model as ρ with subscripts indicating the associated model species. As a result, for each protein-reporter pair, the expression rates are proportional and autodegradation rates can be treated as identical in the model. This permits the assumption that the ratio of co-transcriptional protein amounts is constant, a useful assumption when performing parameter inference (See Appendix A).

The combined production and degradation terms result in the governing equation of the model's protein species. For LacI, the transcriptional repressor, that equation is

$$\frac{\mathrm{d}R}{\mathrm{d}t} = x_R\,\mathcal{H}(\mathrm{A}, \lambda_R, K_R) - (\rho_R + r_c(1 - \frac{c}{C_m}))R,$$

where x_R represents the maximal expression rate as a function of AHL, A, and the Hill function parameters λ_R and K_R. The synthase, CinI, is dependent on IPTG (model species I), AHL, and LacI concentration within the cell. Each chemical inducer is associated with a Hill function and the product of the three defines the production term. A decreasing Hill function

$$\mathcal{H}_n(R, \lambda, K) = \frac{K^\lambda}{K^\lambda + R^\lambda}$$

is used to model the dependence of synthase expression on repressor protein concentration. LacI-mediated expression, however, is modulated by IPTG concentration. IPTG binds to LacI, preventing it from acting on operator regions in promoter sequences. To model the effect of IPTG, we scale the maximum synthase expression rate by an an increasing Hill function. This implies that the action of IPTG is independent of LacI concentration. While this is an unconventional approach to modeling this well-studied repression system, it was necessary to prevent simulations from producing small pulses in synthase concentration when AHL induction is low and IPTG is absent in order to match the observed circuit behavior. Under a more conventional set of gene regulation approximations, a simulation without IPTG and minor AHL would allow low levels of both synthase and repressor expression. The slow expression of repressor gives the simulated gene circuit time to accumulate a non-negligible amount of synthase, which would cause problems for parameter inference.

Letting x_S be the maximal expression rate of the synthase protein, the governing equation for CinI is

$$\frac{\mathrm{d}S}{\mathrm{d}t} = x_S\,\mathcal{H}(\mathrm{I}, \lambda_I, K_I)\mathcal{H}(\mathrm{A}, \lambda_S, K_S)\mathcal{H}_n(R, \lambda_C, K_C) - (\rho_S + r_c(1 - \frac{c}{C_m}))S.$$

According to these equations, cells should begin producing fluorescent proteins the instant AHL is added to the growth media. The data, however, show a delay between inducer spike-in and the response in fluorescence, as shown in Figure

4.1. AHL transport into the cell, transcription, translation, and fluorescent protein maturation all contribute to this delay. We will rely on equations governing AHL transport to enable the model to reproduce these delays. Other methods of accounting for this delay include introducing species representing mRNA and immature fluorescent reporters or employing delay-differential equations. Both approaches are common choices in the literature and could result in the observed delay. However, these would complicate the numerical solution of the finite-difference approximation of the reaction-diffusion PDE used to simulate microbial consortia described in Subsection 4.3. Therefore, the model relies a modified Fick's first law to create these post-induction delays.

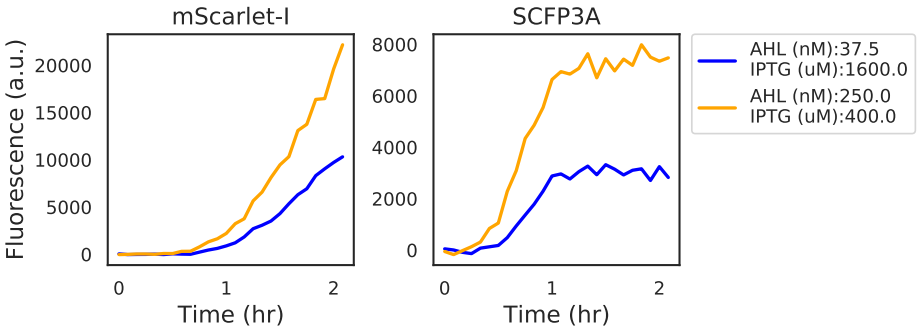

Figure 4.1: Fluorescence data from two experimental wells with high expression activity show a marked delay in the initial accumulation of mScarlet-I fluorescence in comparison to SCFP3A fluorescence.

AHL transport is modeled by Fick's first law with a bounded transport rate, leading to the conditional equation

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \begin{cases} t_A(A_{\mathrm{Ext}} - A) & \text{if } |t_A(A_{\mathrm{Ext}} - A)| < t_M, \\ \mathrm{Sign}(A_{\mathrm{Ext}} - A)t_M & \text{otherwise.} \end{cases} . \tag{4.1}$$

The parameter t_A is the rate constant defining transport into the cell as a proportion to the difference in signal molecule concentrations between the external media and bacterial cytosol. Investigations into trans-membrane transport of AHL molecules have found that larger quorum sensing molecules rely on active transport. Applying an upper bound t_M to the transport rate is to consider saturation of the active transport proteins. The impact of transport on the external AHL concentration is ignored, as the volume fraction of the bacterial population is much less than the volume of the growth media.

A_{Ext}, therefore, is a constant, equal to the concentration of AHL added at the beginning of an experiment. While the synthase could contribute to the concentration of AHL, we do not model this effect. As discussed in Section 2.2, increasing synthase expression did not result in increased mScarlet-I expression.

The full set of differential equations

$$
\begin{aligned}
\frac{\mathrm{d}A}{\mathrm{d}t} &= \begin{cases} t_A(A_{\text{Ext}} - A) & \text{if } |t_A(A_{\text{Ext}} - A)| < t_M \\ \text{Sign}(A_{\text{Ext}} - A)t_M & \text{else} \end{cases} \\
\frac{\mathrm{d}R}{\mathrm{d}t} &= x_R\, \mathcal{H}(\mathrm{A}, \lambda_R, K_R) - (\rho_R + r_c(1 - \frac{c}{C_m}))R \\
\frac{\mathrm{d}S}{\mathrm{d}t} &= x_S\, \mathcal{H}(\mathrm{I}, \lambda_I, K_I)\mathcal{H}(\mathrm{A}, \lambda_S, K_S)\mathcal{H}_{-}(R, \lambda_C, K_C) - (\rho_S + r_c(1 - \frac{c}{C_m}))S \\
\frac{\mathrm{d}c}{\mathrm{d}t} &= r_c\, c\,(1 - \frac{c}{C_m})
\end{aligned}
$$

are applied as a generative model used during Bayesian parameter inference described in Appendix A.

4.3 Reaction-diffusion Model

The model described in the previous section was adapted to a reaction-diffusion model to describe the behavior of cells growing on the surface of agarose pads. This simulation approach was adapted from one used in a related project (Doong, Parkin, and Murray (2017)). A reaction-diffusion model is a system of partial differential equations of the form

$$\partial_t u = D\nabla^2 u + f(u).$$

The function f is called the reaction term; it describes the dynamics resulting from interactions between model species. The diffusion term is $D\nabla^2 u$, which confers linear, isotropic diffusion at rates $D_{i,i}$. D is an $n \times n$ real-valued matrix and u is a vector of n terms.

Evolution equations from the well-mixed media model become the components of the reaction term $f(u)$, with a few modifications. Instead of logistic growth, the bacterial division rate varies with the local value of a model species representing nutrient concentration. Both cell growth and protein production depend on nutrient availability. Just as the relationship between chemical inducers and protein expression rate is defined by a Hill function, the nutrient

species scales cell growth and protein production generally; an approach that is adapted from C. Liu et al. (2011). Each evolution equation governing these species includes a production term that is scaled by a

$$\mathcal{N}(t, x, y) = \mathcal{H}(n(t, x, y), \lambda_n, K_n),$$

a term representing nutrient availability. While protein production and cell growth depend on nutrient availability, only cell growth actually consumes nutrients. This feature is anchored in the assumption that the metabolic investment in the synthetic circuit's proteins is minor compared to host functions.

The reaction terms for the model species representing the sender and amplifier strains are

$$\frac{\mathrm{d}C_{\text{sender}}}{\mathrm{d}t} = r_c \, \mathcal{N} C_{\text{sender}},$$
$$\frac{\mathrm{d}C_{\text{pulser}}}{\mathrm{d}t} = r_c \, \mathcal{N} C_{\text{pulser}}$$

and the resulting change in nutrient concentration is defined by

$$\frac{\mathrm{d}n}{\mathrm{d}t} = -\rho_n \left(C_{\text{sender}} + C_{\text{pulser}}\right) \mathcal{N},$$

where ρ_n is the nutrient consumption rate. The initial condition of the nutrient concentration is 100, uniform over the simulation space. 100 is a value selected out of convenience and the species is unitless. Nutrient and cell density are exchanged at a rate of r_c/ρ_n, meaning the total amount of cell produced during a simulation is bounded by $100 r_c / rho_n$.

The reaction terms corresponding to the protein species have a production and degradation term, just as in the well-mixed media model. Degradation terms, however, must be adapted according to the growth model. In place of the logistic function is the $\mathcal{N}$ scaling term and a masking function to limit protein expression to the appropriate regions. Masking is accomplished by applying a Heaviside function on the concentration of cell species, biased negatively by a small bias term ϵ_c. When simulations incorporate non-zero diffusion rates for the cell species, cells may invade nearby space. Masking ensures protein expression occurs only in regions occupied by cells of the correct identity. The small bias ensures that, in the case of linear diffusion, the rapidly expanding cell mass does not result in an equally rapid expansion of protein expression. It is a well-known fact from transport theory that linear isotropic diffusion results in infinite range expansion; any local perturbation of a continuously-valued,

diffusing species instantly produces a global effect. While this is not the case in a finite difference scheme, the range does increase with each time step and the bias term prevents expression from regions of negligible cell concentration. With these alterations, the reaction terms governing protein species evolution are

$$\frac{\mathrm{d}R}{\mathrm{d}t} = (x_R\ \mathcal{H}(\mathrm{A}, \lambda_R, K_R)\Theta(C_{\mathrm{pulser}} - \epsilon_c) - r_c R)\mathcal{N} - \rho_R R$$

for the repressor species and

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \left(x_S\, \mathcal{H}\left(\mathrm{A}, \lambda_S, K_S\right) \mathcal{H}_-\left(R, \lambda_C, K_C\right) \Theta\left(C_{\mathrm{pulser}} - \epsilon_c\right) + x_O\, \Theta\left(C_{\mathrm{sender}} - \epsilon_c\right) - r_c S\right) \mathcal{N} - \rho_S S$$

for the synthase species. Finally, AHL synthesis is emitted at a rate proportional to the synthase quantity at a location and degrades according to the autodegradation term ρ_A:

$$\frac{\mathrm{d}A}{\mathrm{d}t} = x_a\, S\left(C_{\mathrm{pulser}} + C_{\mathrm{sender}}\right) - \rho_A A\ .$$

The finite difference method (FDM) is an approach to approximating partial differential equations using difference equations. These difference equations are designed to approximate the PDE system on a set of grid points, in this case over the space and time variables. In this model, the difference scheme applies second-order second central difference equations to approximate the second-order spatial derivatives. The grid used in the FDM scheme is conceptually similar to a compartmentalization of the agarose arena into cubic chambers. The model species take on a single value within each chamber and evolve according to the ODE model equations and diffusion. For a parameter set p and state vector u, the finite-difference approximation to the reaction-diffusion system results in the difference equation

$$\frac{u_{n+1} - u_n}{\Delta t} = \frac{1}{\Delta x^2} A(p) \cdot u_n + f(u_n; p).$$

$A(p)$ is a matrix that applies the second-order central differences approximation of the Laplace operator, resulting in difference equations of the form

$$\mathcal{L}(u(t, x, y)) = \sum_{\substack{1=|(i,j)-(x,y)| \\ (i,j)\in \mathrm{P}}} \frac{1}{\Delta x^2}(u_{i,j} - u_{x,y}) \tag{4.2}$$

that describes diffusion at each position (x, y) given the values at neighboring positions in the grid P. The finite difference schemes applied to this problem take equally spaced grids in all spatial dimensions. In Equation (4.2), there are two spatial dimensions, so $\Delta x = \Delta y$. In the case of linear, isotropic diffusion, the matrix $A(p)$ is a constant, band matrix. The band structure in $A(p)$ applies the second-order central differences while D, from the PDE, is diagonal and scales the Laplacian terms according to their diffusion rates. The matrix $A(p)$ also provides no-flux boundary conditions to ensure diffusing species do not escape the agarose pads at its boundaries. Note that the size of matrix $A(p)$ is $n \times n$ for $u_n \in \mathbb{R}^n$; though the simulated arena is two-dimensional, the values of the model species over the two-dimensional spatial grid is represented as a vector and the diffusion operator as a two-dimensional matrix.

Accounting for out-of-plane diffusion

It would be convenient to find a difference scheme that can produce simulated data that is similar to experimental observations without explicitly tracking AHL concentration within the interior of the agarose pad. However, scaling laws described by Dieterle et al. (2020) suggest that the most natural treatments, either of an infinitely tall pad or a negligibly tall pad, are not appropriate given the similarity in the measured propagation velocity (in the area of 1-2 mm/hr) and the height of the pad (roughly 1.5mm). Here, the agarose pad organization described in Section 2.3 is applied to determine the impact of AHL diffusing out of the plane of interest through simulation.

The organization of strain components on the agarose pads applied in Section 2.3 to investigate signaling propagation along one dimension was designed to minimize rate at which diffusion carried AHL away from consortium cells. Sender strains were separated from their consortium partner strains, either amplifiers or reporters, and isolated to a narrow region at one end while their partner strain occupied the larger portion. As a result, the experimental setup was symmetrical in one dimension. Thus, simulations of this setup can consider only two dimensions, one in the direction of signal propagation and one perpendicular to the surface the cells live in.

Simulations of the model described in Section 4.3 were performed that matched the cell patterning and agarose dimensions applied described in Section 2.3. The simulation region described a surface 2mm high and 24mm long, matching the

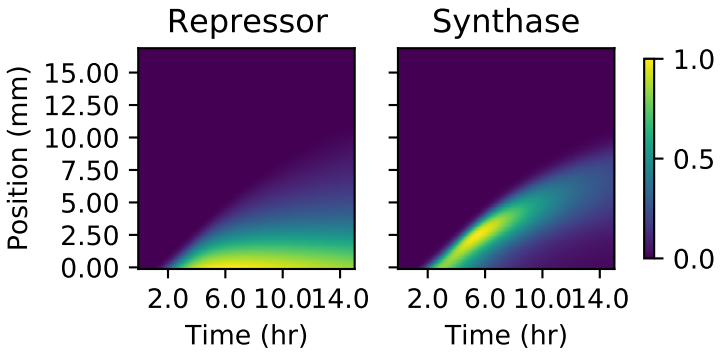

Figure 4.2: The two plots depict simulated protein quantities (in arbitrary units) along the sender-amplifier consortium at each simulation time step. Quantities are determined by multiplying the model species of cell density by those representing intracellular protein concentrations. The intensity values depicted in the heatmaps are normalized to the maximum value in each image.

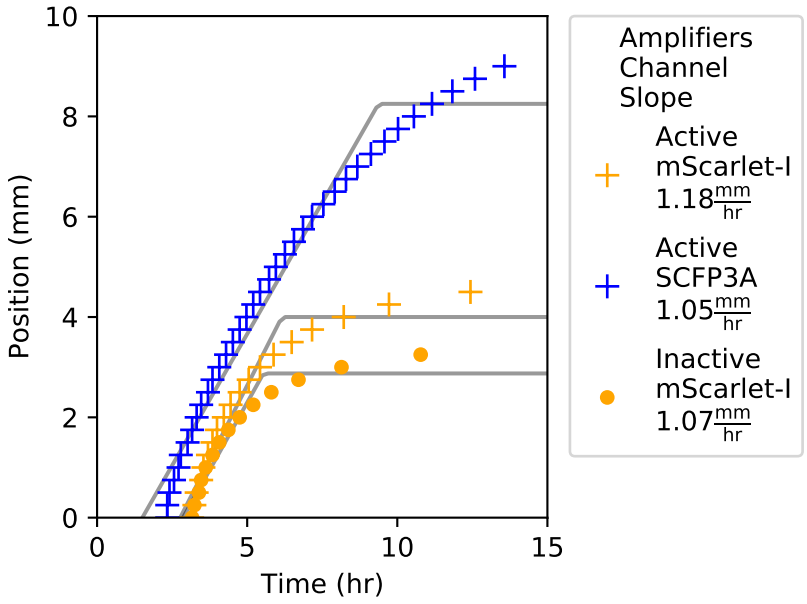

Figure 4.3: This plot shows the propagating front of signaling activity as represented by the quantities of repressor and synthase species. To determine the position of the signaling front, a threshold was applied to the fluorescence data presented in Figure 4.2. The points depicted in the plot correspond to the position farthest from the sender population that surpasses the threshold value at each time step. Threshold values were determined by the Otsu thresholding method. Propagation velocities and maximum signaling distances were determined by fitting a line of the form $\mathrm{Min}(at+b, c)$, where t is simulation time, a is velocity, and c is maximum calling distance, were fit by least squares.

dimensions of the agarose pads used during time-lapse microscopy experiments. In the initial data used for numerical solution of the initial value problem

$$u_{n+1} = A(p)u_n + f(u_n),$$

cells were confined to the bottom row of a two-dimensional grid that spanned the pad height with grid spacing $\Delta x = \frac{1}{16}mm$. Diffusion was considered only for the model species representing AHL. Due to the uniform cell occupation across the pad length, no spatial variation could occur in nutrient or cell density. Diffusion rates for the protein species were likewise set to zero. The simulation ran for a simulated time of 15 hours.

Figure 4.2 depicts protein quantities over the layer of cells and time grid from a simulated sender-amplifier consortium. Just as in the experimental data depicted in Figure 2.6, after a period of a few hours, the amplifier cells respond to AHL secreted by the near by sender cells (which occupy the grid positions between 0 and -4mm) and produce a propagating front of signaling activity that travels away from the sender population. The position of this signaling front over time is shown in Figure 4.3. The front produced by the simulated consortia travels farther and slightly slower than the experimental consortium (see Figure 2.7). However, the three phases of pulse propagation remain: pre-initiation growth, constant-velocity propagation, and nutrient depletion, and the behavior is remarkably similar considering the model parameters were identified from data generated by liquid culture experiments.

Figure 4.4 shows the concentration drop in the out-of-plane direction, measured at each grid point in the cell layer. The more-negative values coincide with the synthase peaks in Figure 4.2, while the concentration profile everywhere is significantly less steep. This suggests that, as might be expected, out-of-plane diffusion is most significant near isolated, high-AHL emission regions and at the leading edge of the propagation front in particular. Away from these areas, diffusion smooths the out-of-plane concentration gradient and reduces the impact of diffusion along that vector. A reasonable approach to capturing the effect of out-of-plane diffusion without explicitly modeling signal transport within the interior of the agarose pad, therefore, would be reducing the expression rate of synthase protein or the synthesis rate of AHL. This reduction in expression rate accounts for the quantity of signaling molecule that is effectively sequestered in the interior of the pad by diffusion. This is more

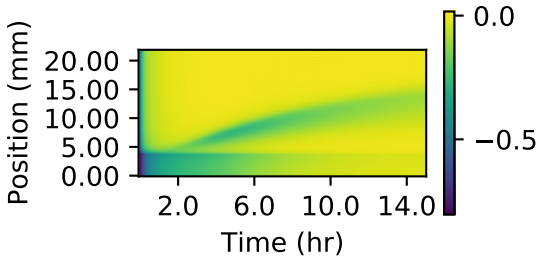

Figure 4.4: The difference between AHL concentrations in the cell layer and the layer immediately above it was calculated at each gridpoint within the cell layer, then divided by the value in the cell layer. The value in the cell layer was subtracted from the value in the higher layer. Negative values indicate a concentration drop in the out-of-plane direction, which results in out-of-plane diffusion. Values closer to zero indicate a flat profile where diffusion has less of an impact. The interface between the sender and amplifier regions is at 4mm, with sender cells below that line and amplifier cells above it.

appropriate account for out-of-plane diffusion by increasing the autodegradation rate, which would lead to rapid concentration declines even away from high-emission regions where the out-of-plane diffusion rate should be small. One unrealistic result of this assumption is that it cannot account for the gradual accumulation of signaling molecules in high-synthesis regions. As seen in Figure 4.4, the concentration gradient over the sender cell population grows more shallow over time as signaling molecules accumulate above it. As this occurs, AHL accumulation within the cell layer will accelerate. However, this takes place well behind the propagating signaling front and has little impact on it. In order to capture the impact of out-of-plane diffusion at the leading edge of the propagating signaling front, the simplest appropriate adjustment was made: reducing the protein production rate of the sender synthase and the signal synthesis rate by half.

Finite difference scheme for growing microbial consortia in two dimensions

The purpose of developing this model of cell-cell communication within microbial consortia is to accurately reproduce observations made during time-lapse microscopy experiments and then extend its use to hypothetical consortia. These data show heterogeneous responses in fluorescence and growth at the

sub-millimeter level. It is critical to determine whether these observations are natural consequences of the well-characterized growth and synthetic gene circuit models; if not, then our understanding of the synthetic consortia is insufficient to extend use of this model to more complicated consortia.

Unfortunately, the PDE model includes processes that proceed at drastically different time scales. This can make numerical solution problematic. It was the case that applying common general-purpose numerical solvers for stiff problems, such as Runge-Kutta 4(5) and LSODA, failed in integrating the nutrient-dependent growth model combined with the cell-cell signaling circuit with spatially heterogeneous microbial consortia. A fixed-step numerical integration method was therefore developed. While generally less computational efficient and accurate than variable time step methods, fixed-step methods can be customized to satisfy *a priori* accuracy constraints and guarantee completion of the requested integration. This subsection describes the fixed-step operator-splitting approach used in simulating complex, growing consortia on 2D surfaces.

It is likely that the "stiffness" at issue in this problem arises from the expanding microcolonies and the masking performed to restrict protein production to grid points that are sufficiently dense in cells. By splitting the system evolution into two steps and calculating the effect of the diffusion and reaction terms separately, each can be efficiently calculated according to their different time scales.

The diffusion step, due to the linearity of its differential operator, can be efficiently resolved using a second-order implicit method that ensures good performance while maintaining a relatively large fixed time step to speed integration. Relative to the diffusion term, the reaction term is both non-linear and more difficult to apply to an implicit approach. Instead, the reaction contribution is calculated using a simple explicit form, iterated several times over a smaller time step.

The diffusion step is performed by a Crank-Nicholson step. First, the diffusion step is proposed as an implicit difference equation

$$\tilde{u}_{n+1,i} = u_n + \frac{\Delta t}{2} A(p)(\tilde{u}_{n+1,i} + u_n)$$

and is then rearranged to form the matrix equation

$$(I - \frac{\Delta t}{2} A(p))\tilde{u}_{n+1} = (I + \frac{\Delta t}{2} A(p))u_n \,,$$

which permits approximation of the mid-step value $\tilde{u}_{n+1,1}$ from the known value u_n. Python scripts executing the simulation use SciPy's generalized minimum residual method to quickly approximate the value of the unknown. Following this, the mid-step value $\tilde{u}_{n+1,1}$ is used as the initial data for a series of forward Euler steps of the form

$$\tilde{u}_{n+1,i+1} = \frac{\Delta t}{4} f(\tilde{u}_{n+1,i}) + \tilde{u}_{n+1,i}$$

that employ a time step $\Delta t/4$. The forward Euler steps are repeated four times to match the time step taken during the diffusion calculation, resulting in the final step value $\tilde{u}_{n+1,5} = u_{n+1}$ at the completion of the time step. The time step for the implicit portion is selected according to the constraint $\Delta t < \frac{\Delta x^2}{6D_a}$ to prevent oscillations in the diffusion step.

All model species have positive diffusion coefficients in the linear diffusion operator $A(p)$. While *E. coli* can engage in random walks through flagellar motion and modeled as such using linear diffusion, the bacteria under the agarose pad move only as the result of crowding and pressure from microcolony formation. Nonlinear diffusion equations such as the permeable membrane equation would be a more appropriate descriptor of cell movement in this scenario. This is not pursued here, as linear diffusion enables more efficient integration. The nutrient-based growth model, however, achieves a similar range expansion effect that is achieved by slow diffusion of nutrient and cell species. While linear diffusion results in an immediate response across the spatial domain in the PDE, the low diffusion coefficients of nutrient and cell species creates a sharp transition from high to low density at the edge of cell-occupied regions.

Using this finite difference scheme and operator-splitting approach to time stepping, simulations were performed to match the conditions present in the experiments described in Chapter 3. Selected images representing the simulation state are depicted in Figure 4.5. The spatial grid in the scheme had $\Delta x = \frac{1}{16}$mm over a region 9mm $\times$ 9mm and a Crank-Nicholson timestep of $\Delta t = 3$seconds was applied. Initial cell data were sparse boolean arrays, where the small number of high-density grid points were selected randomly from a 0.7mm disk at the center of the spatial domain. For each simulation, a single high-density sender position was selected and 100 high-density amplifier or reporter positions were selected. The integration proceeded for 18000 time steps such that the

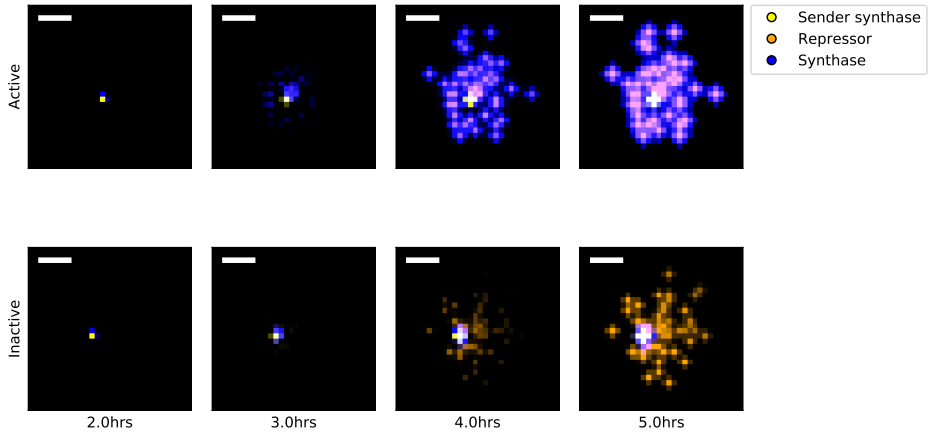

Figure 4.5: Simulations of sender-amplifier and sender-reporter consortia were performed, designed to recreate the experiments described in Chapter 3. The two rows of images show protein amounts from the sender-amplifier (Active) and sender-reporter (Inactive) consortia at selected time points. Protein concentrations are represented in the false-color images according to their identity and their associated cell type. Black indicates absence, and brightness correlates with quantity. A color key is indicated by the figure legend. The isolated point in the frames corresponding to 2 hours is the location of the single sender microcolony. Scale bar is 100μm.

time grid spanned 0 to 15 hours. Ten such simulations were performed for each consortium.

While the simulated sender-amplifier consortia exhibit a marginally more rapid response to signals originated by the sender population than the sender-reporter consortia, the simulated consortia generally appeared to respond more rapidly than observed in experiment. Figure 4.6 shows the response time values calculated for each simulation. Compared to the analogous plot of experimentally observed response times in Figure 3.2, the simulated consortia activate in roughly in roughly half the time. Simulated sender-reporter consortia, where signal amplification is inactive, also showed a significant response. This can be seen in Figure 4.5, where the images in the "Inactive" row show a clear response near the sender colony that is absent from the experimental data in Figure 3.1. In contrast, experimental data showed almost no reporter activity. This suggests that the adjustments described in Section 4.3 did not sufficiently account for the impact of out-of-plane diffusion in this scenario. In order to

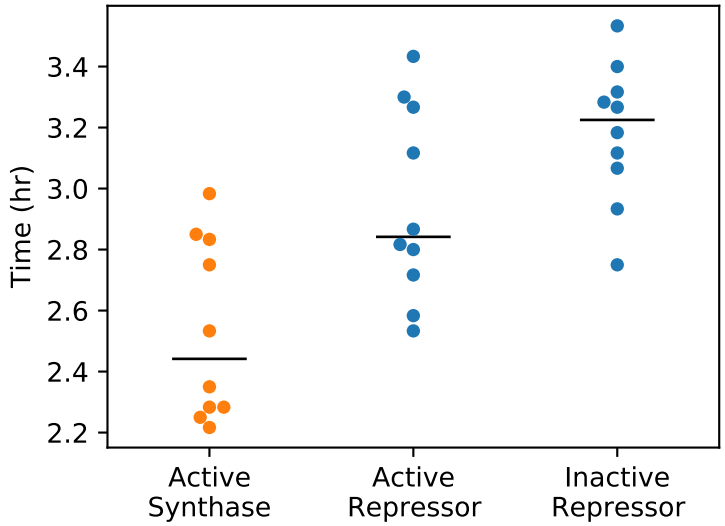

Figure 4.6: Response times observed in simulation. Each point represents the time at which the first reporter or amplifier grid point value crossed a threshold value. Threshold values were determined for each protein species by applying the Otsu method to the values generated from sender-amplifier consortia simulations. The black horizontal lines indicate the sample median.

improve this model's ability to accurately recreate experimentally observed phenomena, the parameters controlling synthase expression in sender cells and synthase activity rate could be fit to the model data.

4.4 Evaluating the benefit of amplification in hypothetical consortia

Simulations were designed to apply the inferred parameters governing cell growth, sender activity, and amplifier activity to predict the behavior of hypothetical consortia to scenarios matching the growth and diffusive contexts described in Chapter 3: a consortium growing within a disk of 0.7mm diameter on an agarose pad 4.5mm by 4.5mm by 2mm. The consortia applied to these simulations express cell-cell signaling and transcriptional programs selected to explore the potential for collective decision-making in random founding cell patterns and spatiotemporally heterogeneous sender activity.

Chained sender-receiver cascade

The first consortium assayed is a chained sender-receiver cascade. Like the party game "telephone", a cascade is initiated by one strain of sender cells secreting signals that are received by the next strain in the chain. This population then expresses synthase protein as activated by the initial signals, producing signal molecules that activate the following component of the chain. In total the chain contains five strains, the initiator and four subsequent chain components. Added to these chained sender-receiver relationships are amplifier cells that are triggered by each of the sender-receiver signaling species to emit a pulse of the same species. Figure 4.7a) conceptually represents this consortium structure, though not the founding cell patterning. The composition of these simulated consortia vary over the ratio of cascade cells to amplifier cells as well as toggling the amplifier cells between inactive, in which case they act as bystanders, or active amplifiers for all sender-receiver interactions. Varying the composition in this manner allows for the investigation of the relative benefit of amplifier cells as the cascade cells increase in density. Figure 4.7c) suggests that, with or without amplifier activity, the chained cascade produces negligible activity past the first interaction. However, the amplifiers succeed in boosting that response in all cases. This is likely due to nutrient depletion, which appears to enforce a time limit on effective consortium-wide communication at the simulated arena size, simulation time, and the assumption of an infinitely-deep diffusive environment.

Coincidence detector

The second consortium investigated was composed of amplifiers (or receivers when inactive), two sender populations, and a receiver. Simulating this consor-

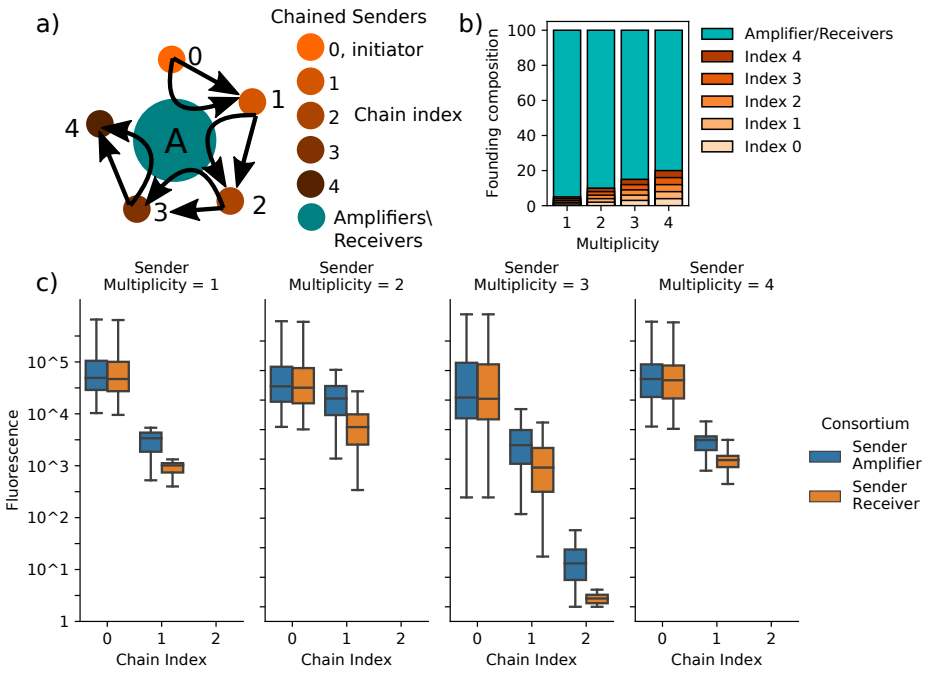

Figure 4.7: Hypothetical consortia composed of amplifiers/receivers and chained sender cells are simulated to evaluate how amplifying cell-cell signaling can benefit information propagation. (a) depicts a schematic representing the cell-cell signaling relationships represented in these simulations. The chained senders are connected by unidirectional, sender-receiver relationships in which chain index i emits signal molecules that activate synthase expression in chain index $i + 1$. The index 0 population are constitutively active. Each sender-receiver relationship employs unique, orthogonal signaling species Amplifiers, however, can amplifify the signaling activity at each index. Simulations vary the bystander cells' status as amplifier or receiver and the composition of the founding consortia populations, which is described in (b). Sender multiplicity is the number of founding cells of each of the sender populations. Because the founding population is set at 100 individuals, higher multiplicity reduces the total number of amplifier/receiver cells in the founding population. The benefit provided by amplification is measured by comparing the maximum synthase amount in each sender population achieved during simulations including either amplifiers or receivers as partners to the chained senders. Receivers do not respond to signaling activity, but they occupy space and consume nutrients. (c) depicts, in boxplots, the distributions of sender activity at each index of the sender chain obtained over ten simulations of each consortium. Simulations of chain/amplifier and chain/receiver consortia utilize the same initial conditions such that the comparisons of the distributions are not confounded. In all assayed conditions, utilizing parameters governing gene networks and cell growth in B.3, the amplifiers increase the response in chain indices greater than 0. However, the response in all multiplicities except 2 is significantly diminished in comparison to the constitutively-active index 0 senders and the response from chain indices greater than 1 is negligible in all cases. Together, these data suggest that even with amplifier cells, a chained signaling cascade between minority populations is unlikely succeed past a single step.

tium sheds light on how amplifying cell-cell signaling can benefit consensus in an environment of spatially heterogeneous chemical signals. In this scenario, the senders stand in for components of a deployed consortium that alert the community to environmental events to which they are uniquely sensitive. Receivers respond to the simultaneous presence of signals from both sender cells, which can also be amplified by the amplifier strain. Simulations are performed that apply varied pulses of activity from the sender populations. The success of the communication strategy is evaluated on the basis of simulated fluorescence from the receiver population. A conceptual summary of this consortium and results provided in Figure 4.8a). The results show an unexpected pattern of receiver response in which the strongest activity arose from simulations with sender pulse timings at time 2 and 5 hours with a pulse width of 4 hours. This implies an overlap in both pulses of 1 hour, relatively short compared to other simulations that produced overlaps of 2 or 3 hours. While increasing the overlap time naturally results in more receiver activity, it appears that separating the initiating times of the two sender populations has a greater benefit. The reasons behind this are not clear and further investigation is required before arriving at conclusions. It is clear that the time-varying nutrient availability and spatial patterning is a key determinant of both the effectiveness of cell-cell communication within consortia executing complex programming and the benefit provided by amplifier strains.

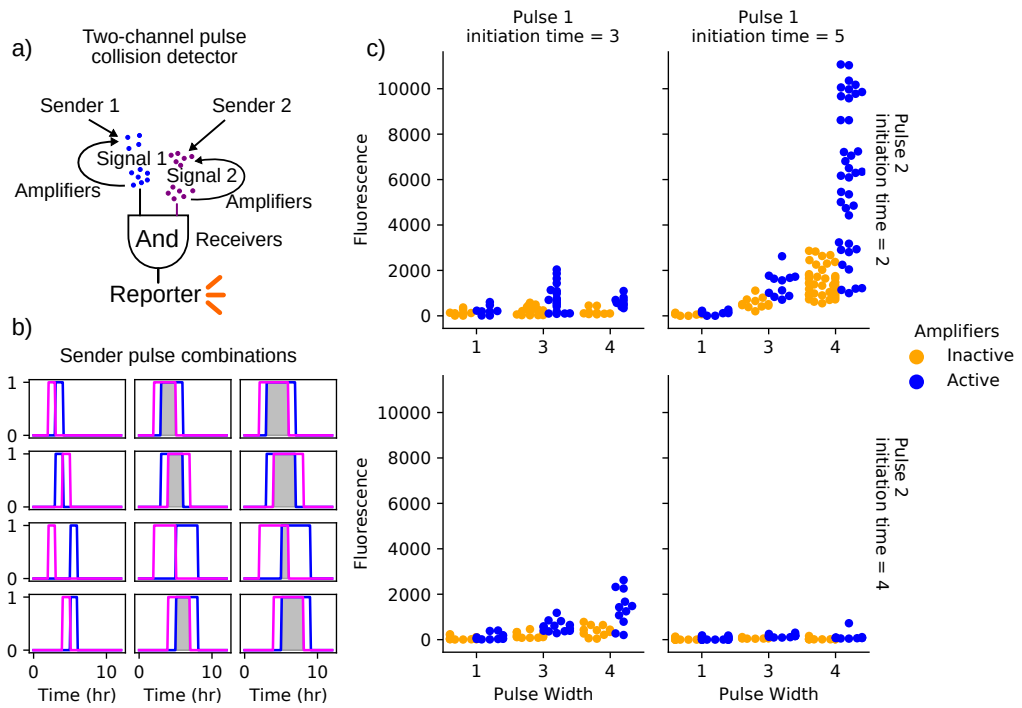

Figure 4.8: Outline and results of the coincidence detector consortium. (a) depicts a schematic representing the cell-cell signaling relationships represented in these simulations. The two sender populations initiate signaling through the release of diffusible signaling chemicals. Amplifiers, when active, are triggered by the sender populations to contribute to the local concentrations of these chemicals in a pulsatile fashion. The receiver population produces a fluorescent output in the presence of both signaling molecules. Simulations vary the timing and duration of synthase expression within the sender populations as shown in (b). Simulations span 12 hours. Gray-shaded regions indicate overlapping synthase expression within the sender populations. Plots in (c) show the maximum pixel values within the receiver population obtained over three simulations of each sender-emission pattern. Each of these three iterations employs a different random spatial patterning of the founding population, which is composed of a single founder of each of the sender and receiver strains and 100 amplifier/bystander cells. These random initial conditions are repeated for consortia with active and inactive amplifier cells such that active/inactive cases can be compared directly. In all assayed conditions, utilizing parameters governing gene networks of sender, amplifiers, and receivers as well as cell growth in B.3 and use the same arena and consortia geometry as in 4.5. The simulated receiver fluorescence indicates the degree to which the population receives input from both sender cells. All conditions show a greater average response in the receiver population when amplifiers are active. However, only the simulations in which one pulse initiates at 2 hours and the other at 5 hours results in significant expression. This is likely a result of the time-varying nutrient availability and suggests the relationship between event timing and response is non-trivial and not intuitive. Further exploring the spatiotemporal dynamics of nutrient availability in simulation requires more experimental data capturing these effects in order to be predictive.

Appendix A

BAYESIAN PARAMETER INFERENCE OF THE LIQUID CULTURE MODEL USING MARKOV CHAIN MONTE CARLO

A.1 Bayesian parameter inference

Bayesian parameter inference using Markov Chain Monte Carlo is an approach to curve-fitting that enables approximation of the probability distribution over model parameters given the observed experimental data. This section describes the selection of likelihood and prior functions used to estimate the posterior (Hogg, Bovy, and Lang (n.d.)).

The data available presents many options for defining the likelihood function. Many approaches to parameter inference for dynamical systems apply generative models to produce simulated data. At each parameter set sampled by the Markov process, the difference between the simulated and experimental values, or residual, is used to estimate the probability of the data given the sampled parameter set. The difficulty in taking this approach when dealing with a nonlinear dynamical system is that the distribution of residual values can be hard to determine or approximate. This is in part due to the fact that the residuals are not independent of one another. Consider, for instance, a generative model of the data that involves the numerical solution of ordinary differential equations that depends on the sampled parameters. Due to the continuity of the differential equations and, by assumption, the physical process that generated the experimental data set, large residuals are likely surrounded by large residuals. Instead of generating simulated data according to a sampled parameter set and the model equations, the approach we take here is to compute the residual of a self-consistency relationship derived from the model equations.

Defining the residual function

These self-consistency functions are derived from the differential equations described in Chapter 4. The equations governing protein expression include variables that represent intracellular concentrations. By rewriting these as equations relating the total quantities of each species, the equations then define an expected consistency between observable data, namely the derivatives of

the fluorescence and optical density traces, and Hill functions. A residual calculation is performed by treating the Hill functions as the generative model describing the observed data.

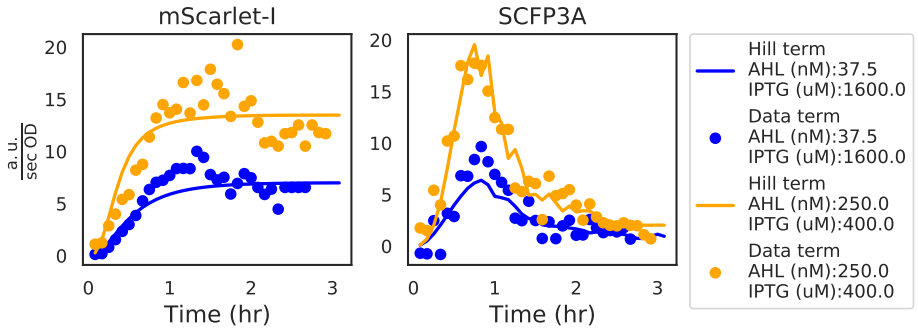

Figure A.1: The two plots show the Hill and Data terms based on data from two experimental wells and the maximum-posterior estimate for the liquid culture model parameters. The inducer concentrations from each well are included in the figure legend. While the Hill term values corresponding to the mScarlet-I channel are functions of simulated AHL transport and parameter values, those of the SCFP3A channel depend on observed density-normalized mScarlet-I fluorescence. Due to noise in measuring density-normalized fluorescence, the Hill term values of the SCFP3A channel appear less smooth than in the mScarlet-I channel.

Consider the differential equation that governs mScarlet-I expression, according to the liquid culture model:

$$\frac{\mathrm{d}R}{\mathrm{d}t} = x_R\,\mathcal{H}(\mathrm{A}, \lambda_R, K_R) - (\rho_R + r_c(1 - \frac{c}{C_m}))R.$$

Substitute the growth-rate dependent dilution term with the definition for the derivative of cell growth and simplify the notation by letting a time-varying function $H(t)$ stand in for the production term:

$$\frac{\mathrm{d}R}{\mathrm{d}t} = H(t) - \rho_R R - \frac{\mathrm{d}c}{\mathrm{d}t}\frac{R}{c}. \tag{A.1}$$

Now consider the relationship between the intracellular concentration of mScarlet-I, R, and the total amount $\hat{R} = cR$. Using the chain rule, we compute the derivative of the total mScarlet-I quantity:

$$\frac{\mathrm{d}\hat{R}}{\mathrm{d}t} = \frac{\mathrm{d}c}{\mathrm{d}t}R + c\frac{\mathrm{d}R}{\mathrm{d}t}. \tag{A.2}$$

Using this relationship, we can substitute the terms for intracellular concentration R with those of the total quantity $\hat{R}$ in Equation (A.1):

$$
\begin{aligned}
0 &= H(t) - \rho_R R - \frac{\mathrm{d}c}{\mathrm{d}t}\frac{R}{c} - \frac{\mathrm{d}R}{\mathrm{d}t}, \\
0 &= H(t) - \rho_R R - \frac{1}{c}(\frac{\mathrm{d}c}{\mathrm{d}t}R - c\frac{\mathrm{d}R}{\mathrm{d}t}), \\
0 &= H(t) - \rho_R R - \frac{1}{c}(\frac{\mathrm{d}\hat{R}}{\mathrm{d}t}), \\
0 &= H(t) - \frac{1}{c}(\rho_R R + \frac{\mathrm{d}\hat{R}}{\mathrm{d}t}).
\end{aligned}
$$

Just as $H(t)$ is a time-varying function, so to is the term $D(t) = \frac{1}{c}(\rho_R R + \frac{\mathrm{d}\hat{R}}{\mathrm{d}t})$. These are called the Hill and Data terms, respectively, and examples are shown in Figure A.1. The residual is the difference between the two terms

$$r = H(t) - D(t).$$

As a result, the residual distribution can be approximated by the distribution of the Data terms. This same derivation is used to determine the definition of the residual for SCFP3A data, where the Hill term is a product of the three Hill functions describing IPTG activation, AHL activation, and transcriptional repression via Laci, and the maximal expression rate parameter x_S.

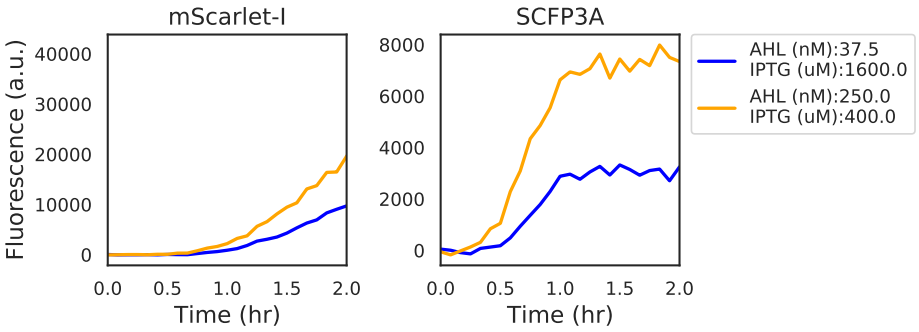

Figure A.2: Fluorescence data from two experimental wells with high expression activity show a marked delay in the initial accumulation of mScarlet-I fluorescence in comparison to SCFP3A fluorescence. Derivative estimates verify that there is no measured change in mScarlet-I fluorescence in the first half hour of data acquisition.

In the case of bicistronic LacI and mScarlet-I expression, we expect that the Hill functions vary in time only as the AHL concentration equilibrates between the

cell volume and the growth media. This occurs rapidly (Kaplan and Greenberg (1985)). However, the data in Figure A.2 show a significant delay before the expression rate gradually ramps up to a peak and, finally, descending. The ramp up may be due to the long maturation time of mScarlet-I (roughly 66 minutes to 90% maturation) and the descent is likely due to slowing cell growth and protein production. This delay is less apparent in the SCFP3A fluorescence channel, perhaps due to its more rapid mautration (roughly 24 minutes to 90% maturation) (Balleza, J. M. Kim, and Cluzel (2018)).

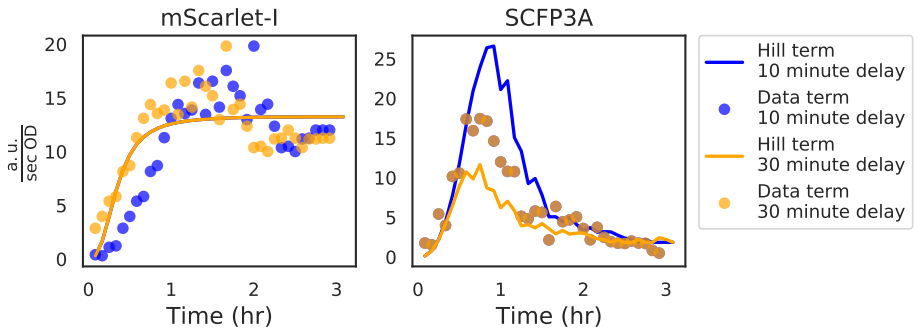

Figure A.3: Data from a single experimental well is shown here with two choices of mScarlet-I channel delay. The corresponding Hill and Data term values illustrate the effect of the delay parameter on how the residual is calculated. In the case of the mScarlet-I terms, the Hill term is left unchanged while the Data terms are moved backwards in time by the delay value. Data that fall into negative time are removed and replaced by resampling the same number of points from the end of the time series. The SCFP3A Data terms are unchanged while the Hill terms respond to the time-shifted density-normalized mScarlet-I fluorescence.

As the model does not consider maturation time or nutrient effects, an additional parameter is included while performing MCMC. This parameter represents impact of mScarlet-I's maturation time on the observed fluorescence data and is demonstrated in Figure A.3. For a delay time t_d, data from the mScarlet-I channel measured at time $t_m < t_d$ is removed and the observation times of subsequent data are shifted backwards in time by $\text{Max}(t_m : t_m < t_d)$. To avoid favoring large delay times by removing data from the likelihood calculation, observations are added by resampling. When n observations are removed, n are resampled from the terminal n points of the time series.

Approximating the residual distribution

The residual distribution is difficult to approximate whenever its dependencies on model parameters, observation values, and other residuals are unknown (Hogg, Bovy, and Lang (n.d.)). One of the benefits of calculating the residual using the self-consistency relationships instead of simulated data is that each residual calculation represents an independent sampling of the residual distribution. When comparing data to a simulated ODE time series, on the other hand, the residuals at each point are not independent of one another. This section describes how the residual distribution of the self-consistency relationships were approximated.

The mScarlet-I channel Hill term is a function of AHL concentration, maximal expression rate x_R, and Hill function parameters λ_R and K_R while the Data term depends both on the observed fluorescence and a single model parameter, d_r. In determining the residual distribution, the contribution of uncertainty from the Hill term is ignored. It is assumed that understanding the distribution of Data term values is sufficient to describe the residual distributions. As the uncertainty in measuring Data term values results from the measurement process and sample-to-sample variability, it is natural to assume that both sources of uncertainty will result in a monotonic increasing relationship between the expected value and width of the distribution of Data terms. The task then becomes determining this relationship.

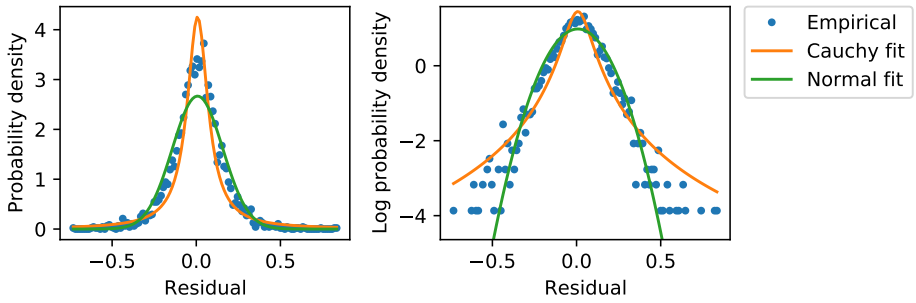

Figure A.4: Distribution of residual values from wells with Hill term values equal to 0. Empirical distribution values were determined through binning the n Data term values into $2\sqrt{n}$ bins and calculating the normalized bin occupancy. The fit to a Normal distribution was perfomed by calculating the sample mean and standard deviation. The fit to a Cauchy distribution used the same sample mean. The Cauchy scale parameter was taken to be half of the sample standard deviation. The sample standard deviation here is 0.24.

This is accomplished by first characterizing the residual distribution when the Hill term is equal to zero, which corresponds to no mScarlet-I expression. In the case that both the Data and Hill terms are expected to be zero, the distribution of Data term values represents the uncertainty in measurement and sample preparation. Sample variability can only introduce uncertainty, therefore the distribution when the Hill term is zero corresponds to the minimum uncertainty in the residual distribution. Figure A.4 depicts the distribution of mScarlet-I channel Data term values recovered from experimental wells without any added AHL.

The distribution of residuals at higher Data term values was approximated by sampling the highest 4 Data term values from experimental wells grouped by AHL concentration. Data terms depend on a single model parameter that at this point is unknown. For the purpose of estimating the residual distribution, the *a priori* expected value of $2 \cdot 10^{-4}$/second was selected. Data term values from the mScarlet-I channel reach a plateau after a few hours (see Figure A.1). By sampling values from the plateaus corresponding to different experimental wells, pooled according to inducer condition, the resulting distribution reflects both measurement residual and sample-to-sample variability. The sample means and standard deviations were computed for each pooled group of Data term values and plotted to determine the relationship between distribution width and expected value. Figure A.5 depicts the observed width - center pairs along with the scaling law implemented in the likelihood calculation.

To verify that the distribution scaling law appropriately described data from all ranges of Data term values, the data within each pooled group was rescaled according to the sample means and standard deviation as predicted by the scaling law. Following this, the data were collated and its empirical density function computed. The result, as shown in Figure A.6, appeared to be well-represented by a Normal distribution.

A similar approach was employed to determine the appriopriate scaling law for the SCFP3A residual distrubtion. Figures A.7 through A.9 show residual distributions and scaling law.

MCMC was performed using the emcee python package, making use of 1000 walkers and the kombine move selector (Foreman-Mackey et al. (2012), B. Farr and W. Farr (2015)). The chain was run for 10000 steps and the walker occupancy distrubtion appeared unimodal. Figure A.11 depicts the positions

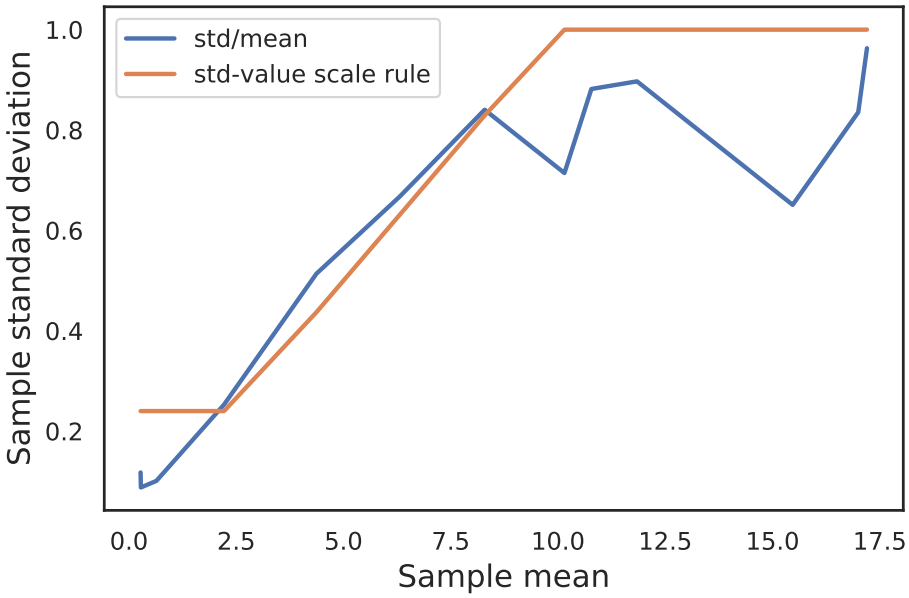

Figure A.5: Investigating the residual distribution in the mScarlet-I channel at high expression from liquid culture experiments. This plot of rescaled Hill and Data terms from high-expression samples suggests that the scaling law appropriately describes the residual distribution as mScarlet-I fluorescence increases. Data term values were rescaled according to the scaling law and an empirical density function was calculated. The empirical probability density function is plotted against fitted Normal and Cauchy distributions. This data set is much smaller than that of Figure A.4, so only $\sqrt{n}/2$ equal-sized bins were used in calculating the empirical density function. The by-eye match between the empirical density values and the Normal distribution was the basis of selecting a Normal distribution using the scaling law in the likelihood calculation.

of the walkers at each iteration. The acceptance rate is well below the optimal for the dimensionality of the posterior distribution. This typically indicates that the walkers have discovered multiple local minima such that the move selector cannot produce a likely proposal. However, due to the proximity of these apparent local minima, the large number of walkers used, and the appearance of the marginalized posterior distributions, the chain is accepted as is. Furthermore, the authors of the emcee package note that many problems suffer from a vanishing acceptance rate, but that this is no cause for concern outside of applications in high-accuracy machining or involving life-or-death decision-making (Foreman-Mackey et al. (2012)). Outside of these cases, a chain with low-acceptance that surpasses 10 autocorrelation times can be

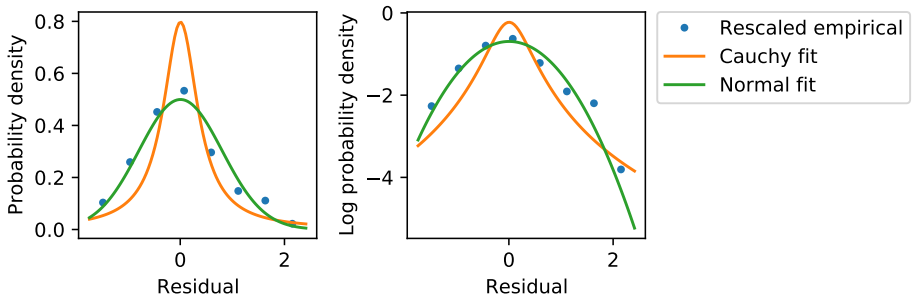

Figure A.6: Investigating the residual distribution in the mScarlet-I channel at high expression from liquid culture experiments. This plot of rescaled Hill and Data terms from high-expression samples suggests that the scaling law appropriately describes the residual distribution as mScarlet-I fluorescence increases. Data term values were rescaled according to the scaling law and an empirical density function was calculated. The empirical probability density function is plotted against fitted Normal and Cauchy distributions. This data set is much smaller than that of Figure A.4, so only $\sqrt{n}/2$ equal-sized bins were used in calculating the empirical density function. The by-eye match between the empirical density values and the Normal distribution was the basis of selecting a Normal distribution using the scaling law in the likelihood calculation.

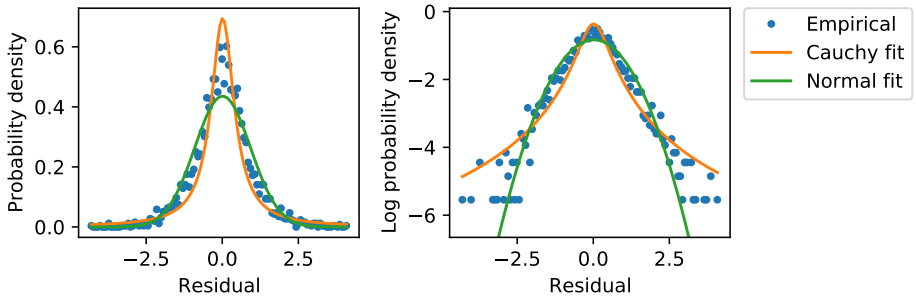

Figure A.7: Investigating the residual distribution in the SCFP3A channel at low expression from liquid culture experiments. Distribution of residual values from wells with Hill term values equal to 0. Empirical distribution values were determined through binning the n Data term values into $2\sqrt{n}$ bins and calculating the normalized bin occupancy. The fit to a Normal distribution was performed by calculating the sample mean and standard deviation. The fit to a Cauchy distribution used the same sample mean. The Cauchy scale parameter was taken to be half of the sample standard deviation. The sample standard deviation here is roughly 1.

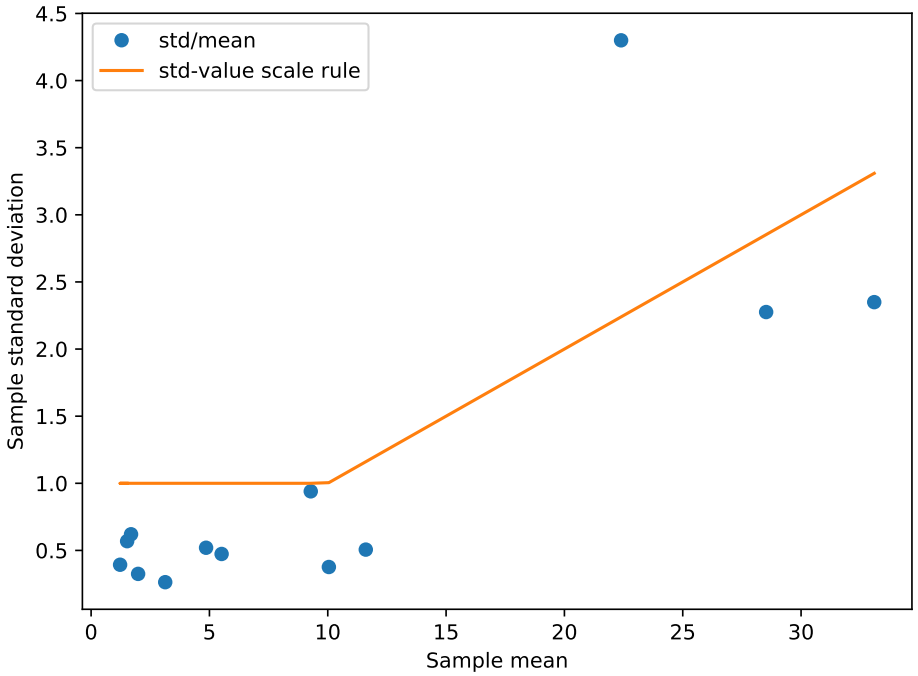

Figure A.8: Characterizing the scaling law for residual distributions in SCFP3A fluorescecne data from liquid culture experiments. This plot depicts in blue the sample means and standard deviations calculated of the Data term values, grouped according to AHL and IPTG concentration. The orange line depicts the scaling law used to define the residual distribution's standard deviation.

considered a reliable approximation of the posterior distribution. Empirical cumulative density functions for each parameter are depicted in Figure A.10.

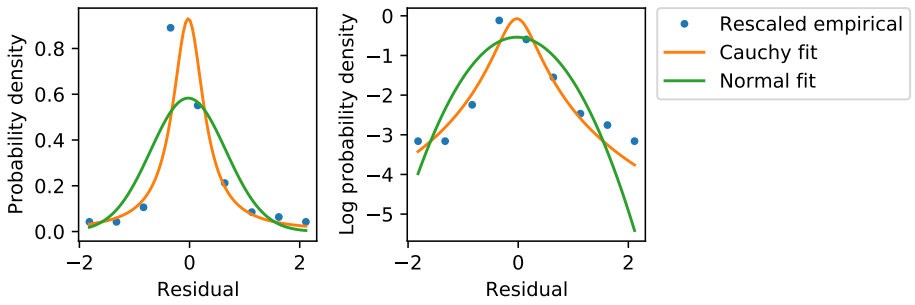

Figure A.9: Investigating the residual distribution in the SCFP3A channel at high expression from liquid culture experiments. This plot suggests that the scaling law appropriately describes the residual distribution when the Hill and Data terms are larger than zero. Data term values were rescaled according to the scaling law and an empirical density function was calculated. The empirical probability density function is plotted against fitted Normal and Cauchy distributions. Again, only $\sqrt(n)/2$ equal-sized bins were used in calculating the empirical density function. The by-eye match between the empirical density values and the Normal distribution was the basis of selecting a Normal distribution using the scaling law in the likelihood calculation.

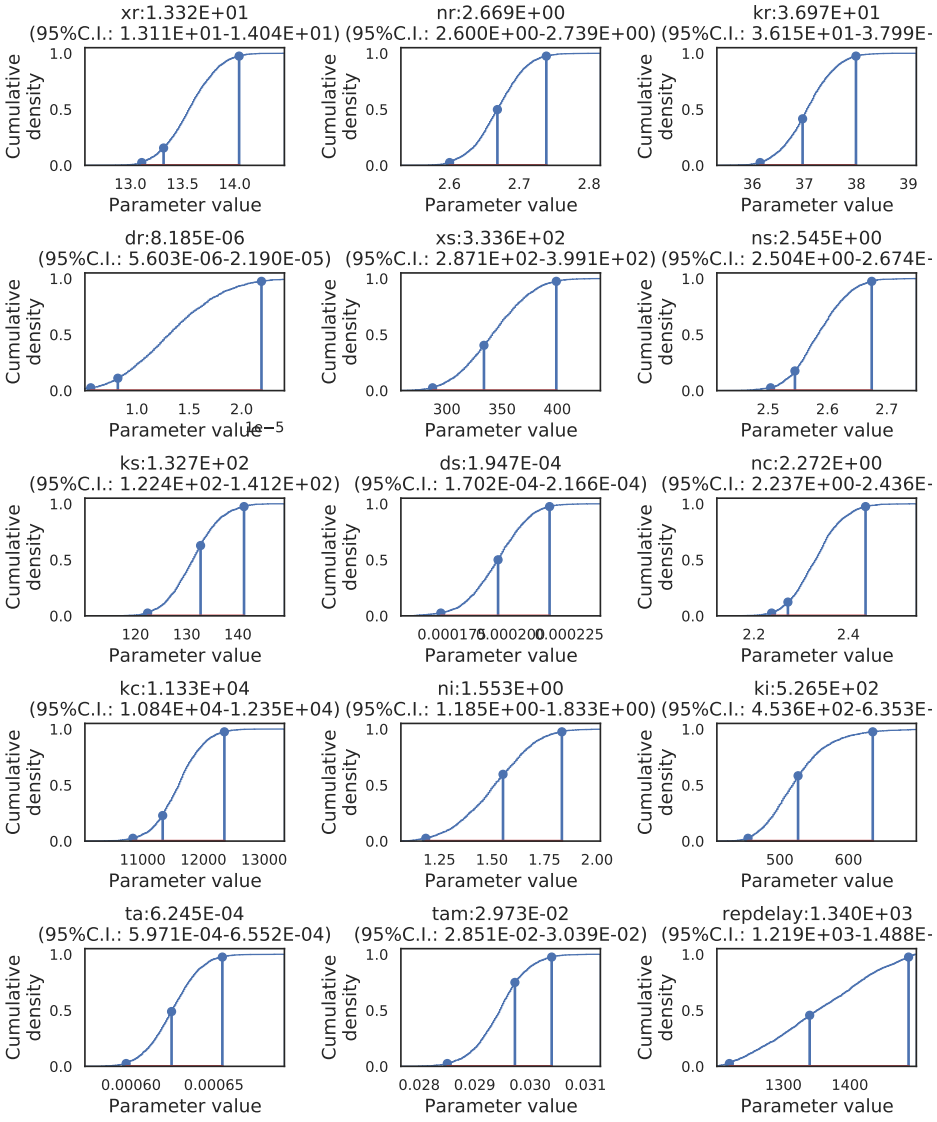

Figure A.10: Empirical cumulative density derived from the MCMC chain. Stem plots bracket the middle 95 centiles and mark the maximum estimated posterior parameter set between them.

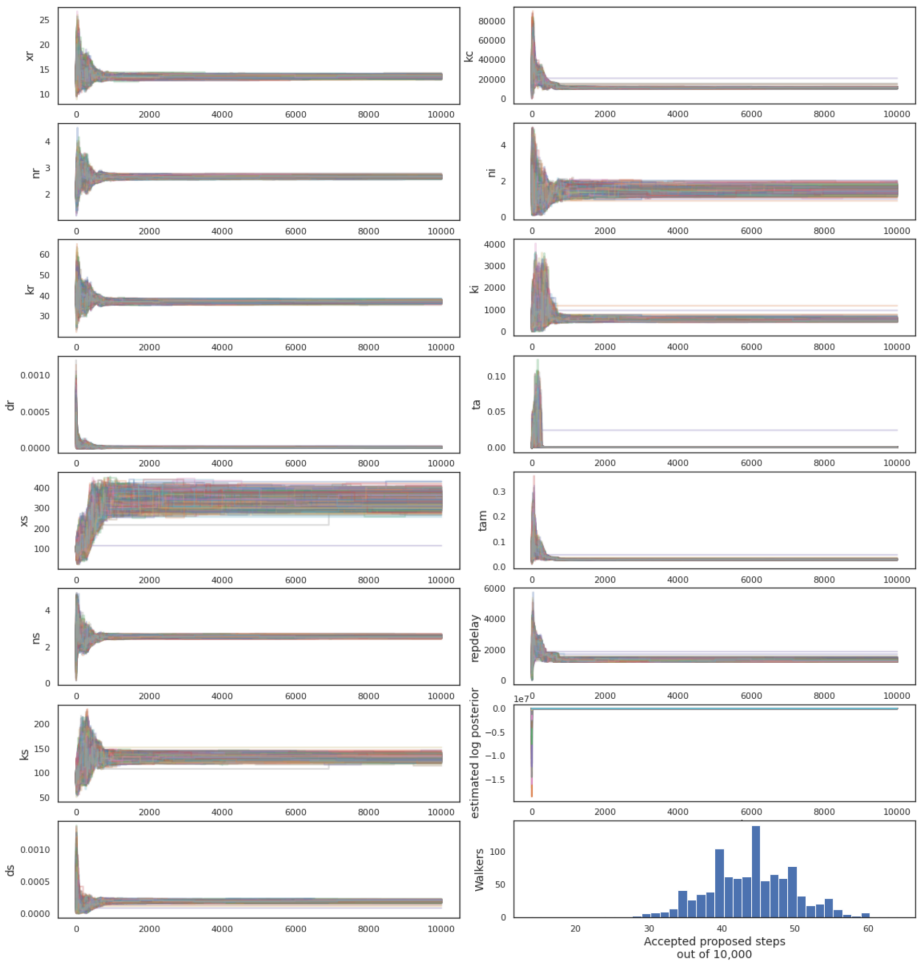

Figure A.11: Walker positions in parameter space over all iterations. Estimated log posterior values were calculated by the emcee ensemble sampler instance. The initial positions of the walker were selected from the *a priori* best fit, but all walkers quickly moved to a more likely region of parameter space.

A p p e n d i x B

TABLES OF MODEL SPECIES AND PARAMETERS

B.1 Liquid culture model

Table B.1: Liquid culture model parameters [1]

Parameter	Notation	Value	Units
Cell growth rate	r_c	2.6×10^{-4}	second^{-1}
Density carrying capacity	C_m	1	OD700
Synthase autodegradation	ρ_s	2×10^{-4}	second^{-1}
Repressor autodegradation	ρ_r	8×10^{-6}	second^{-1}
Synthase maximum expression	x_S	389.2	$\frac{a.u.}{\text{second OD700}}$
Synthase AHL-activated Hill coefficient	λ_S	2.3	
Synthase AHL-activated IC50	K_S	126.4	nM
Synthase IPTG-activated Hill coefficient	λ_I	1.4	
Synthase IPTG-activated IC50	K_I	558	uM
Repressor maximum expression	x_R	13.6	$\frac{a.u.}{\text{second OD700}}$
Repressor AHL-activated Hill coefficient	λ_R	2.65	
Repressor AHL-activated IC50	K_R	36.6	nM
Repressor-mediated inhibition Hill coefficient	λ_C	2.3	
Repressor-mediated inhibition IC50	K_C	11280	a.u. OD700^{-1}
AHL autodegradation	ρ_A	3×10^{-5}	second^{-1}

Table B.2: Liquid culture model species

Species	notation	Units
Synthase	S	$\frac{\text{a.u.}}{\text{OD700}}$
Repressor	R	a.u. OD700^{-1}
Cell density	c	OD700
Intracellular AHL concentration	A	nM
Extracellular AHL concentration	A_E	nM
IPTG concentration	I	uM

B.2 Reaction-diffusion model

Table B.3: Model terms unique to the reaction-diffusion model [2]

Parameter	Notation	Value	Units
Nutrient consumption rate	ρ_n	1.6×10^{-2}	second^{-1}
Nutrient Hill concentration	K_n	68	
Nutrient Hill coefficient	λ_n	4	
Sender cell synthase expression rate[3]	x_O	300	$\frac{a.u.}{\text{second OD700}}$
Sender cell synthase expression rate[4]	x_O	150	$\frac{a.u.}{\text{second OD700}}$
Amplifier cell synthase expression rate[5]	x_s	250	$\frac{a.u.}{\text{second OD700}}$
Synthase activity rate [6]	x_a	2×10^{-6}	$\frac{nM}{\text{second a.u.}}$
Synthase activity rate [7]	x_a	1×10^{-7}	$\frac{nM}{\text{second a.u.}}$
Nutrient diffusion rate	D_n	1.14×10^{-6}	$\frac{\text{mm}^2}{\text{second}}$
AHL diffusion rate	D_a	2×10^{-4}	$\frac{\text{mm}^2}{\text{second}}$

[1] Appendix A describes the parameter inference approach for the gene circuit parameters.

[2] Appendix C describes the parameter inference approach for the nutrient-based growth parameters.

[3]Out-of-plane diffusion model

[4]2D model

[5]Calculated by applying agarose IPTG concentration, 1.5mM, to the IPTG-dependent Hill function from the liquid culture model

[6]Out-of-plane diffusion model

[7]2D model

A p p e n d i x C

INFERRING PARAMETERS GOVERNING CELL GROWTH IN SEMI-SOLID MEDIA

C.1 Introduction

The growth process of cells in semi-solid media is different from growth in liquid culture and therefore requires different treatment within a model. In a diffusive environment, such as an agarose pad, cells deplete local nutrient concentrations while dividing exponentially. Following this depletion, microcolonies grow at a reduced rate, consuming nutrients as they diffuse away from unoccupied or more sparesly occupied by cells. Furthermore, cells divide more slowly as a result of contact inhibition (Payne et al. (2013), Cao et al. (2016)).

Disentangling the various mechanisms affecting division rates in semi-solid media is an active area of research that is beyond the scope of this work. The goal here is to apply a simple model describing cell growth that accurately reflects the empirical growth curves observed during time-lapse microscopy experiments. To do so, a model of cell growth in a diffusive environment is selected that can, under different parameter choices, reflect the range of expected growth behaviors. Then, parameters are selected to minimize the difference between experimental and simulated data.

C.2 Growth model

The model of cell growth used in the finite-differences approximation simulations was based on the work of C. Liu et al. (2011). In this model, the division rate at a position is related to the value of a nutrient species by a Hill function of the form

$$\frac{\partial_t c}{c} = r_c \frac{n^{\lambda_n}}{n^{\lambda_n} + K_n^{\lambda_n}} \tag{C.1}$$

As noted in Chapter 2, the shorthand

$$\mathcal{H}(a, n, K) = \frac{a^n}{a^n + K^n} \tag{C.2}$$

is used for Hill functions.

At the beginning of a simulation, the nutrient species is equal to 100, a nondimensional quantity. This is simply for convenience in notation; a value of 100 indicates 100% nutrient availability. Nutrients diffuse freely within the no-flux boundaries of the simulation. The rate of cell growth is proportional to the rate at which nutrient is consumed. Nutrient quantities are therefore diminished at a rate proportional to the Hill function as well. The model parameter ρ_n defines this proportionality. The two differential equations comprising this nutrient-based model are

$$\begin{aligned} \partial_t n &= D_n \nabla^2 - \rho_n \mathcal{H}(n, \lambda_n, K_n) \\ \partial_t c &= r_c \mathcal{H}(n, \lambda, K_n). \end{aligned} \tag{C.3}$$

C.3 Dataset

A set of time-lapse microscopy experiments were performed to capture the growth behavior of *E. coli* cells on agarose pads. Competent cells of the strain JS006 were transformed with a plasmid conferring constitutive expression of the mScarlet-I fluorescent protein. These cells were deposited onto agarose pads using an acoustic liquid handler. A variety of growth contexts were implemented that varied in both the seeding population density and shape, the position of the colonies on the agarose pad, and size of the agarose pad. Images depicting these contexts are shown in Figure C.1.

Seeding shapes were square patterns of Echo-deposited droplets of 1, 2, or 3 positions wide. Likewise, the square shapes of the agarose pads were multiples of 4.5mm to a side, ranging from 9mm$\times 9mmto 24.5mm\times$24.5mm. The positions of the colonies were selected near the center and near edges of the agarose pads to create different diffusive environments for nutrient transport. Furthermore, some conditions including two separated colonies were repeated with one of the pair missing. If the growing colonies could impact each others' growth even when separated by millimeters, a comparison of these matched agarose pads would illuminate that interaction.

The constitutive expression of fluorescent protein allowed for easy segmentation of the cells within each image. Average cell density was calculated by applying a threshold to each fluorescence micrograph and downsampling the resulting boolean arrays. Downsampling was performed using local means and a scale factor of 64. In this way, planar cell density values were calculated over square

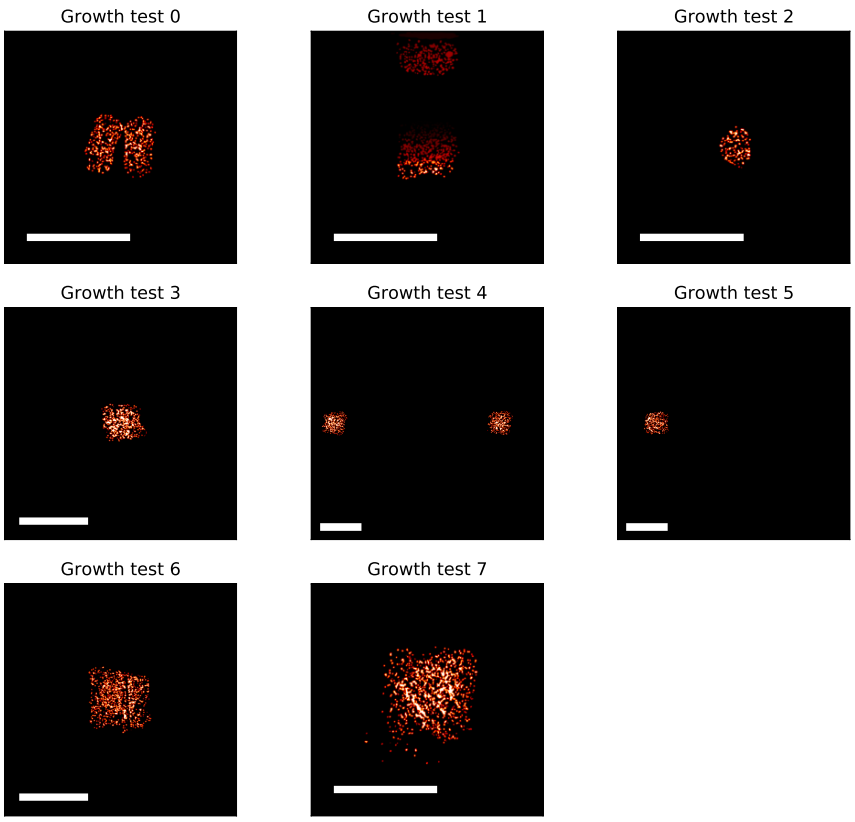

Figure C.1: The images in each plot show the fluorescence micrographs of the different agarose pad growth contexts. The images show the size and position of the cells on the square agarose pads of varying sizes. While the images were collected with a spatial resolution of 0.89 μm, these images were downsampled by a factor of 64 in order to match the spatial resolution of the simulated data. Scale bars in each image are 4mm.

regions 56.8 μm to a side.

C.4 Simulation and parameter fitting

Simulations were performed with the same spatial resolution and applied the first downsampled, stitched microscopy image from each agarose pad as the initial condition of the simulated cell density. Images were drawn from the ongoing simulations at timepoints equal to those of the experimental data as well. Examples of these simulated agarose pads are shown in Figure C.2.

Parameter fitting requires a generative model and a measure of goodness-of-fit between the simulated and experimental datasets. While the correspondence in spatiotemporal resolution of the two datasets encourage a measure that is a

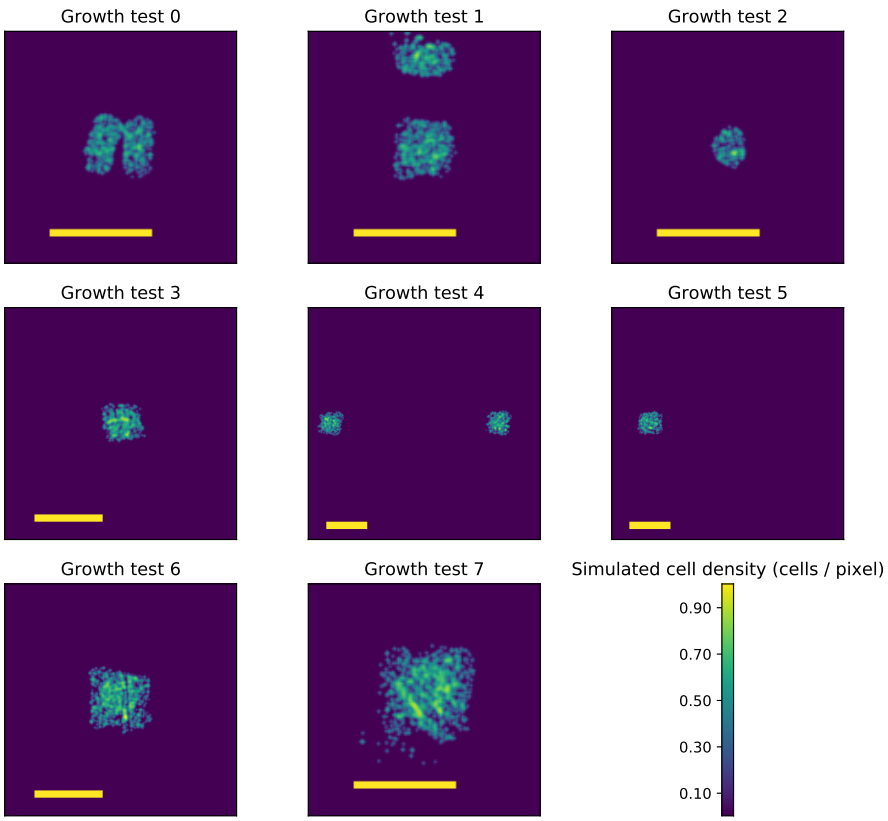

Figure C.2: The images in each plot show the simulated planar cell densities designed to match the experimental conditions shown in Figure C.1. Yellow scale bars in each image are 4mm.

function of the residuals at each position and time, this was not the approach taken. The reason is that cell crowding and displacement are not features of the model. As the founding cells grow from single cells into microcolonies, they form expanding disks. Some of these disks remain within the bins used when performing downsampling while others do not. Increasing the downsampling scale factor reduces the impact of these bin-crossing microcolonies at the expense of spatial resolution. The scale factor of 64 was selected to reduce the impact cells passing into neighboring bins. However, there is another issue. Cells lose fluorescence as a result of loss-of-function mutations. Without the burden of expressing the reporter, mutant cells grow more rapidly than their neighbors and form large, expanding gaps in the fluorescence micrographs. This results both in vanishing cell density estimates within the mutant-occupied regions and crowding of fluorescent cells around the edges of these regions. If goodness-of-fit were calculated on a sum of residuals, it is likely that parameter

inference routine would fit to these outlier data rather than overall distribution of growth rates.

To reduce the impact of loss-of-fluroescence mutants on parameter inference, the two-sample Kolmogorov-Smirnov goodness-of-fit function (K-S function) was used to compare the empirical distributions of the cell growth rates from the simulated and experimental data. The empirical cumulative distribution functions for the experimental and simulated datasets were calculated using 20 equal-width bins between growth rate values of 0 and 2 per hour. The K-S goodness-of-fit function f_{KS} used here to compare empirical cumulative density functions F_1 and F_2 describing samples of size p and q is

$$f_{KS}(F_1, F_2) = \frac{p\,q}{q+p}\sup_x |F_1(x) - F_2(x)|^2 \tag{C.4}$$

Cell growth rates were determined at each position and time by forward differences in cell densities. The parameter fitting routine used SciPy's "optimize" module to find the parameter vector that minimized the K-S values, summed over each frame and growth context. While doing so eliminates the spatial information from the comparison, spatial variability in nutrient availability leads directly to shift of the empirical distribution of the growth rate towards slower growth. The K-S function therefore allows for parameter fitting in spite of obstacles presented by loss-of-function mutations and cell crowding.

A comparison of the simulated behavior at best-fit parameters to experimental data are shown in Figure C.3. The clearest discrepancies are the wider distributions in growth rate seen in the simulated data as mean growth rate declines. However, the mean growth rates are very similar between datasets at all timepoints in all growth contexts.

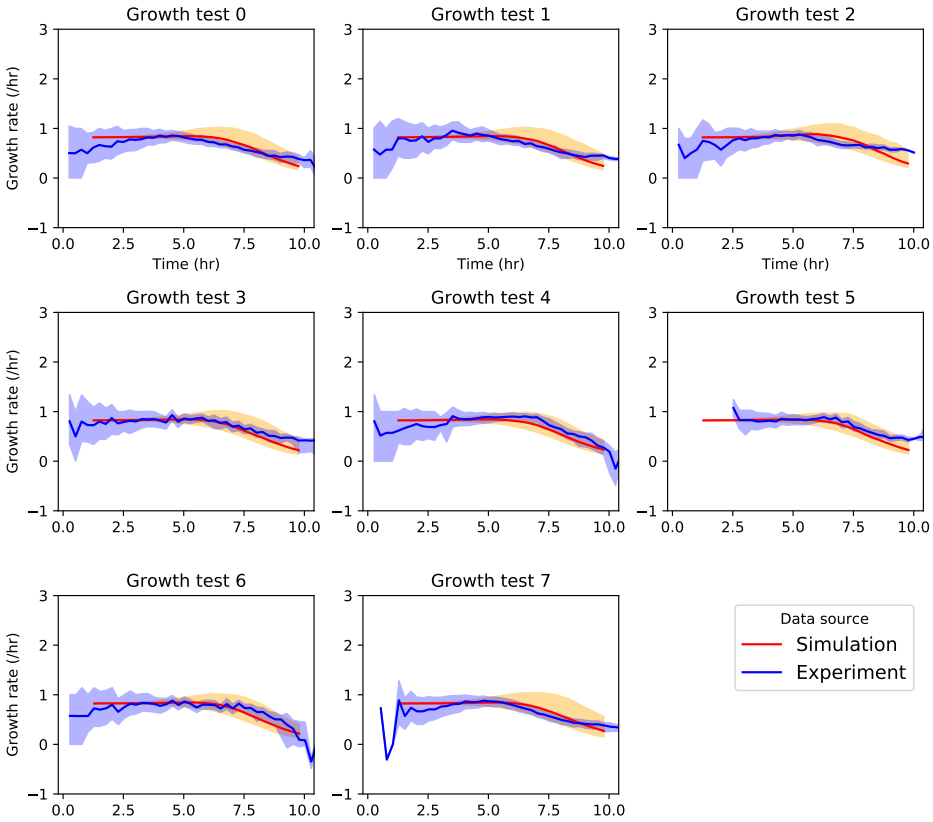

Figure C.3: Growth rate of cells in each growth context calculated by forward difference approximations. Simulation data was generated used the optimized parameter set. Solid lines represent mean growth rate, excepting empty positions. Shaded regions extend from the 25th to the 75th centile at each timepoint.

A p p e n d i x D

TABLES OF PLASMIDS AND STRAINS

D.1 Plasmid and strain tables

Table D.1: Plasmid names and descriptions

Code	Description
ha3_2	pSC101 carb pCinLacO_m B0032 CinI BCD2 sfCFP
har_2	p15a chlor pCinTetO B0032 LacI B0034 mScarlet-I
i1f_8	p15a chlor J23106 B0034 mScarlet-I
i38_13	pSC101 carb pRhlLacO_m B0032 RhlI-ssrA BCD2 sfCFP
i39_1	p15a chlor pRhl B0032 LacI-ssrA B0034 mScarlet-I-ssrA
j78_7	pSC101 carb pCinLacO_m B0032 truncated-CinI-ssrA BCD2 sfCFP
j7h_1	p15a chlor pCinTetO truncated-LacI mScarlet-I
jag_d3	pSC101 carb pRpa B0034 CinI BCD2 sfYFP T J23103 B0032 RpaR

Table D.2: Strain names and descriptions

Description	Code	Plasmids	*E. coli* strain
Cin Amplifier strain	i59_2	ha3_2;har_2	CY026
Weak sender strain	jao_1	jag_d3;har_2	CY026
Strong sender strain	k25_1	jag_d3;j78_7	CY026
Reporter strain	k25_3	j7h_1;j78_7	CY026
Constitutive mScarlet-I	i1f_8	i1f_8	JS006
Rhl positive feedback strain	i38_13	i38_13	CY026

Table D.3: E. coli strain genotypes

Description	Genotype	Source
CY026	BW25113 ΔlacI ΔaraC ΔsdiA Ptrc* -cinR Ptrc* -rhlR	Ye Chen, Addgene Bacterial strain #72340

Description	Genotype	Source
JS006	BW25113 ΔaraC ΔlacI	Jesse Stricker, Stricker et al 2008

D.2 Plasmids maps

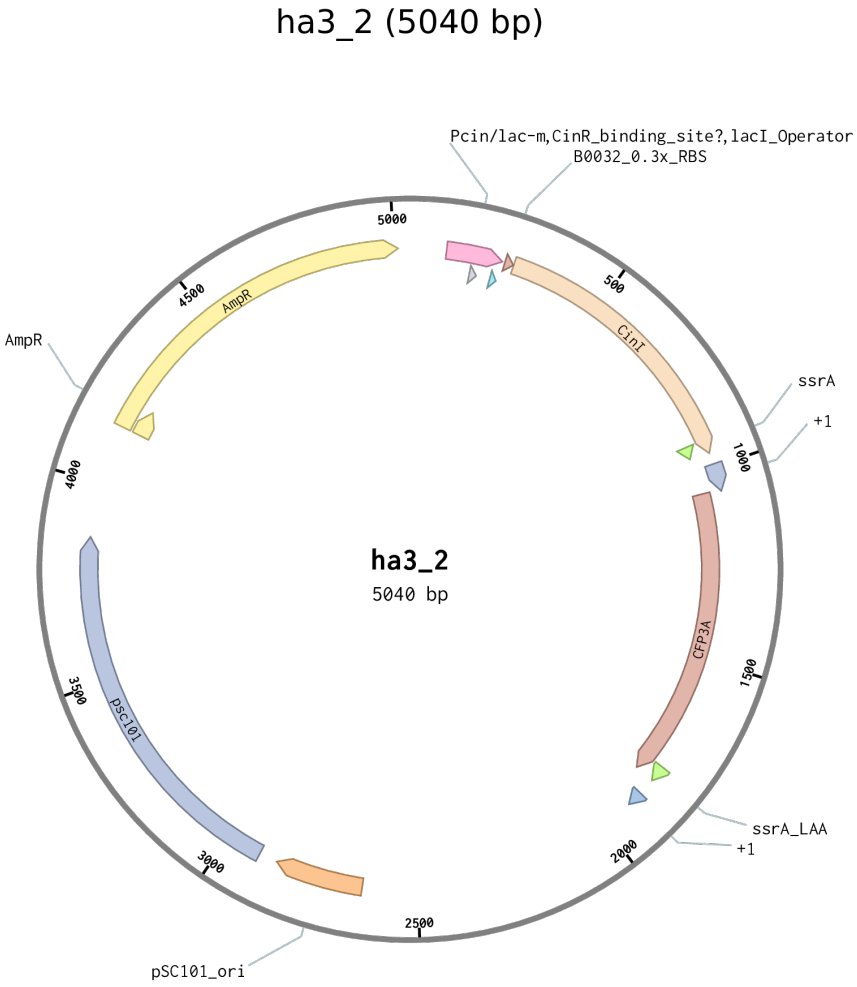

Figure D.1: Plasmid map for Cin positive feedback plasmid ha3_2.

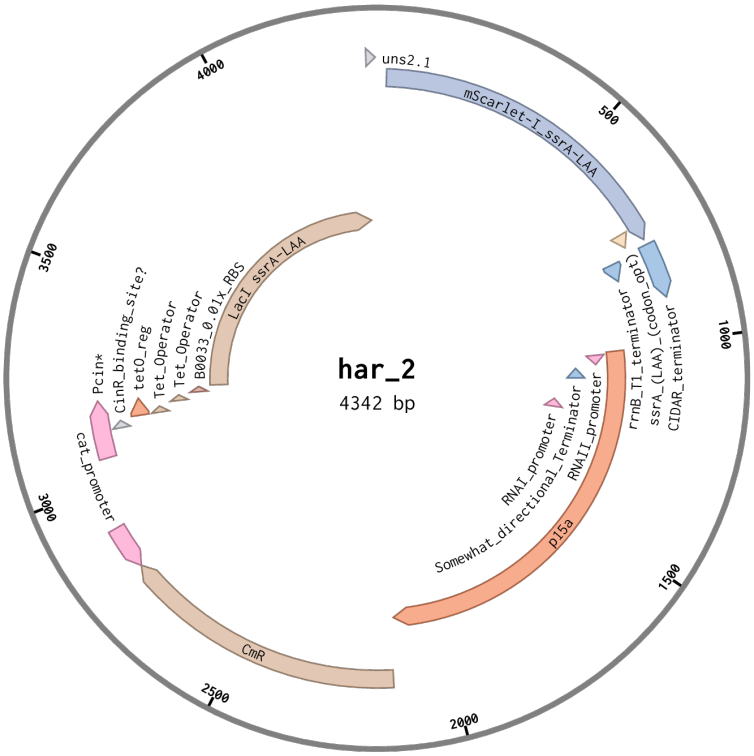

Figure D.2: Plasmid map for Cin negative feedforward plasmid har_2.

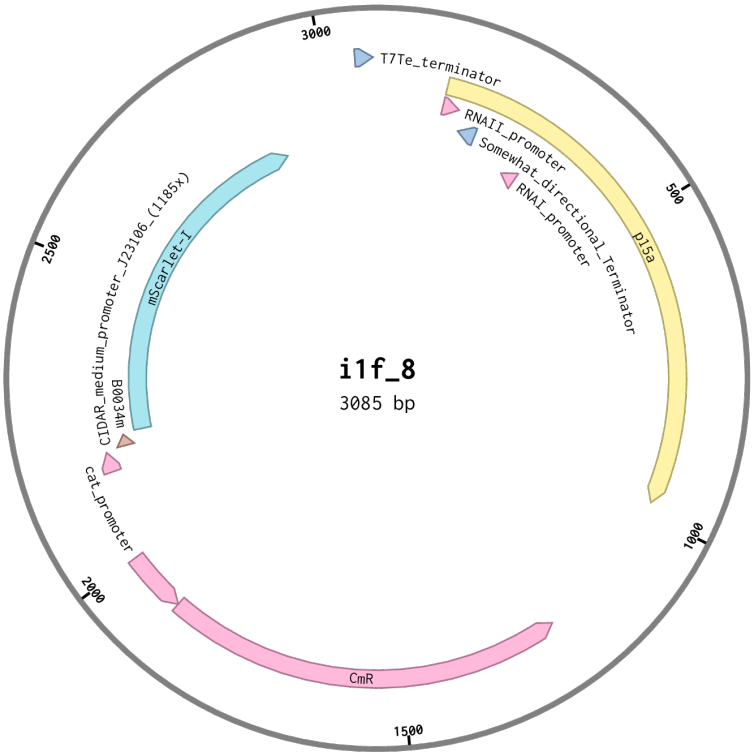

Figure D.3: Plasmid map for constitutive mScarlet-I source i1f_8.

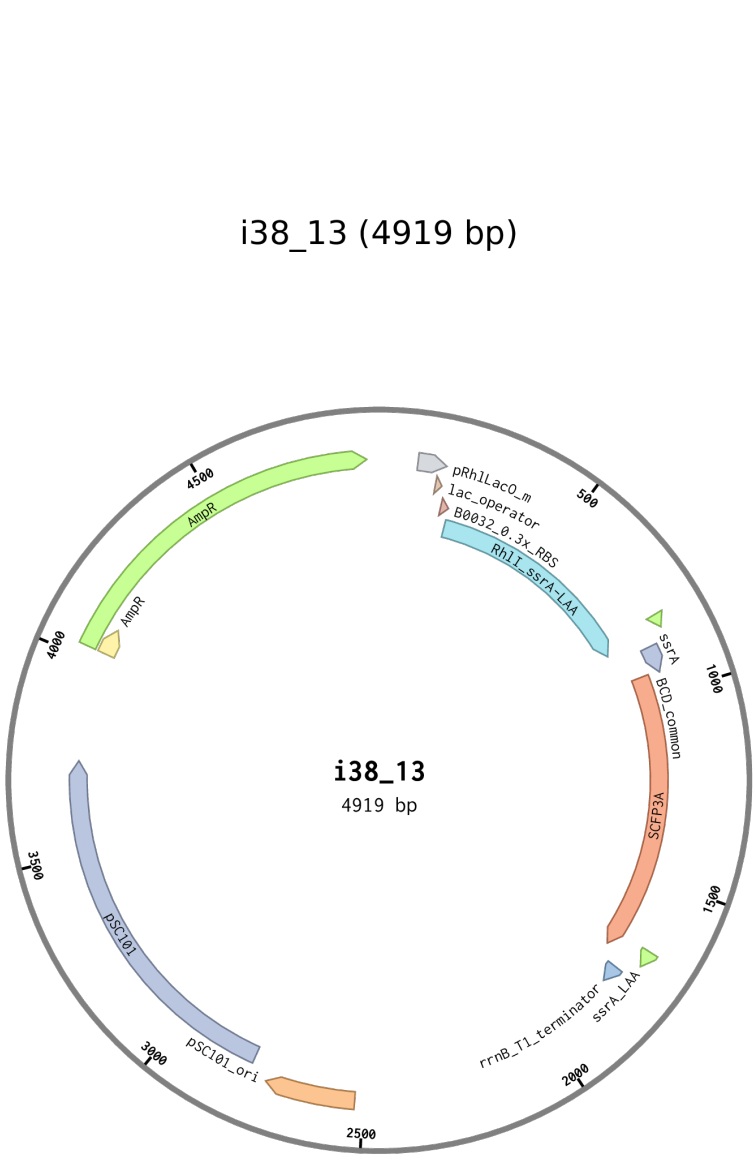

Figure D.4: Plasmid map for Rhl positive feedback plasmid i38_13.

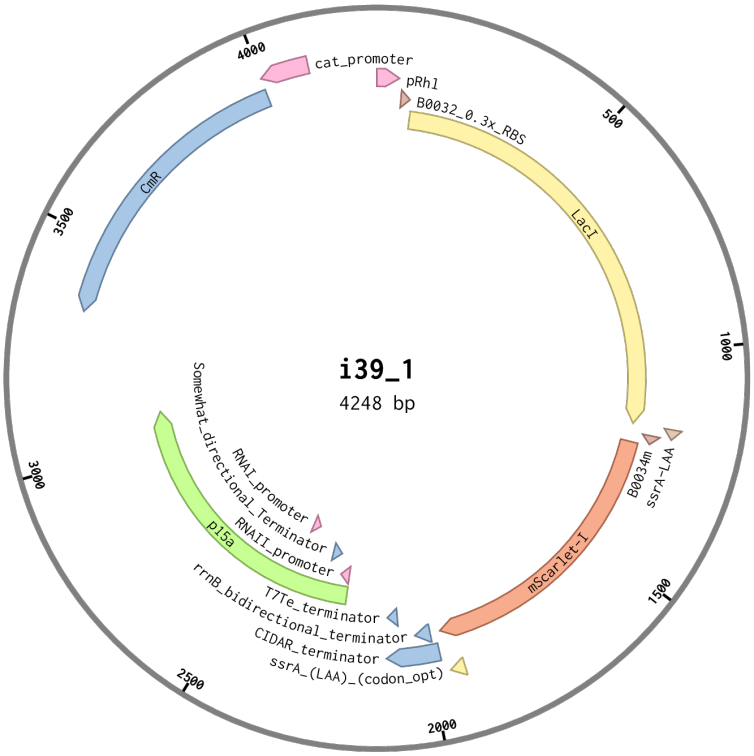

Figure D.5: Plasmid map for Rhl negative feedforward plasmid i39_1.

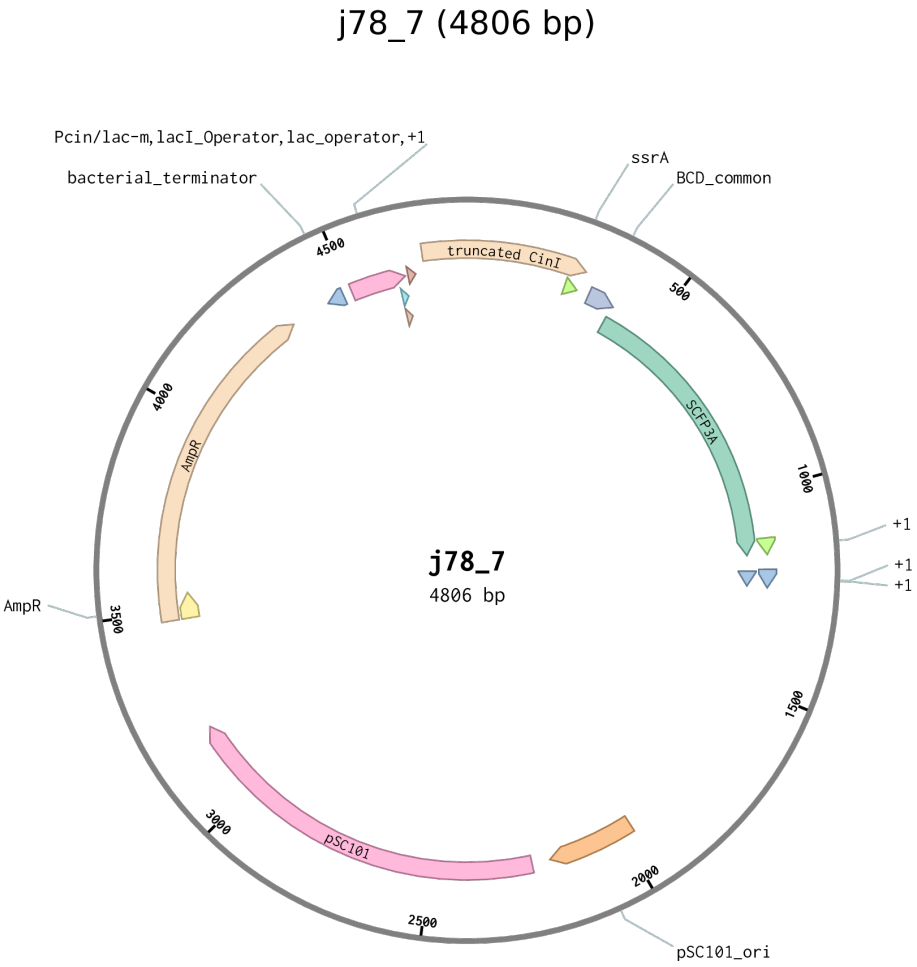

Figure D.6: Plasmid map for truncated CinI Cin reporter plasmid j78_7.

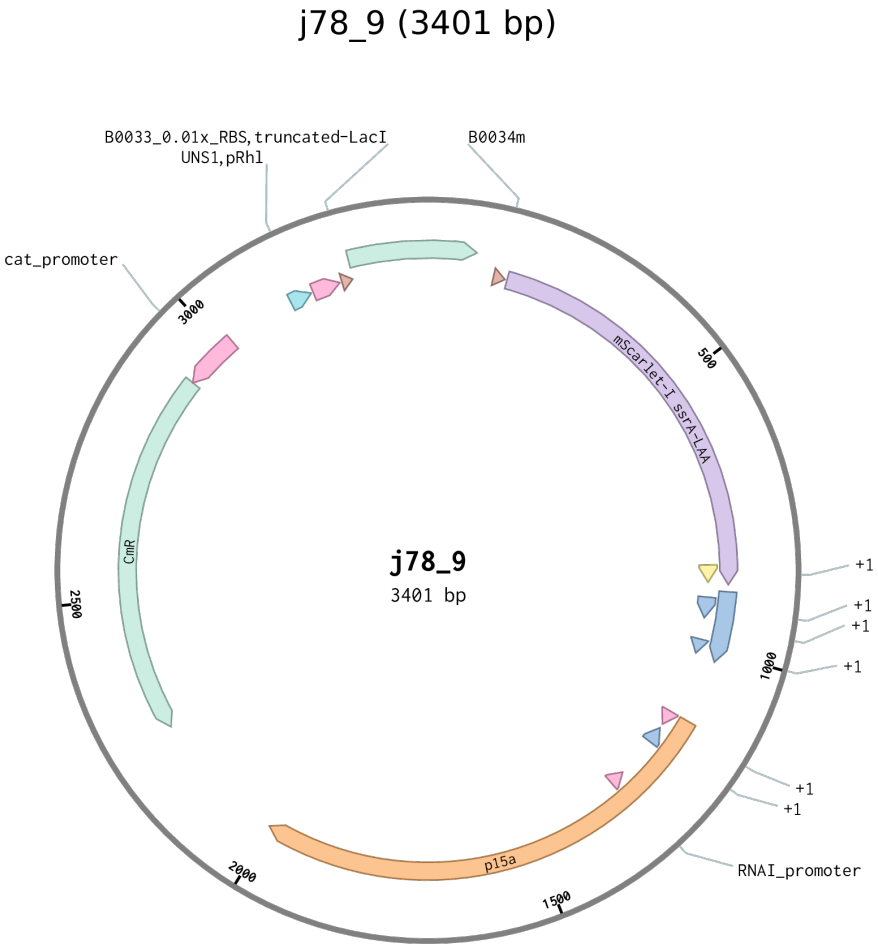

Figure D.7: Plasmid map for truncated LacI Cin reporter plasmid j78_9.

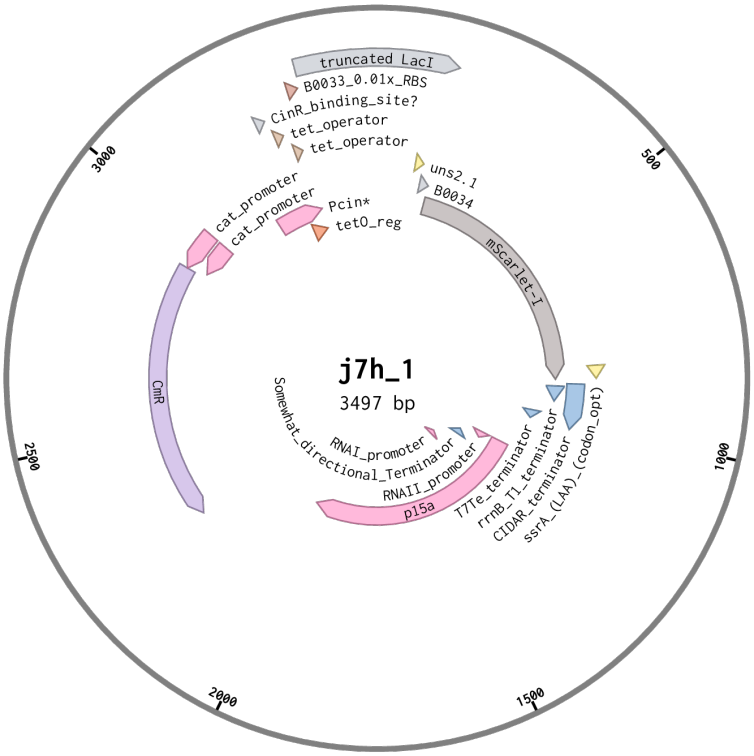

Figure D.8: Plasmid map for truncated LacI Rhl reporter plasmid j7h_1.

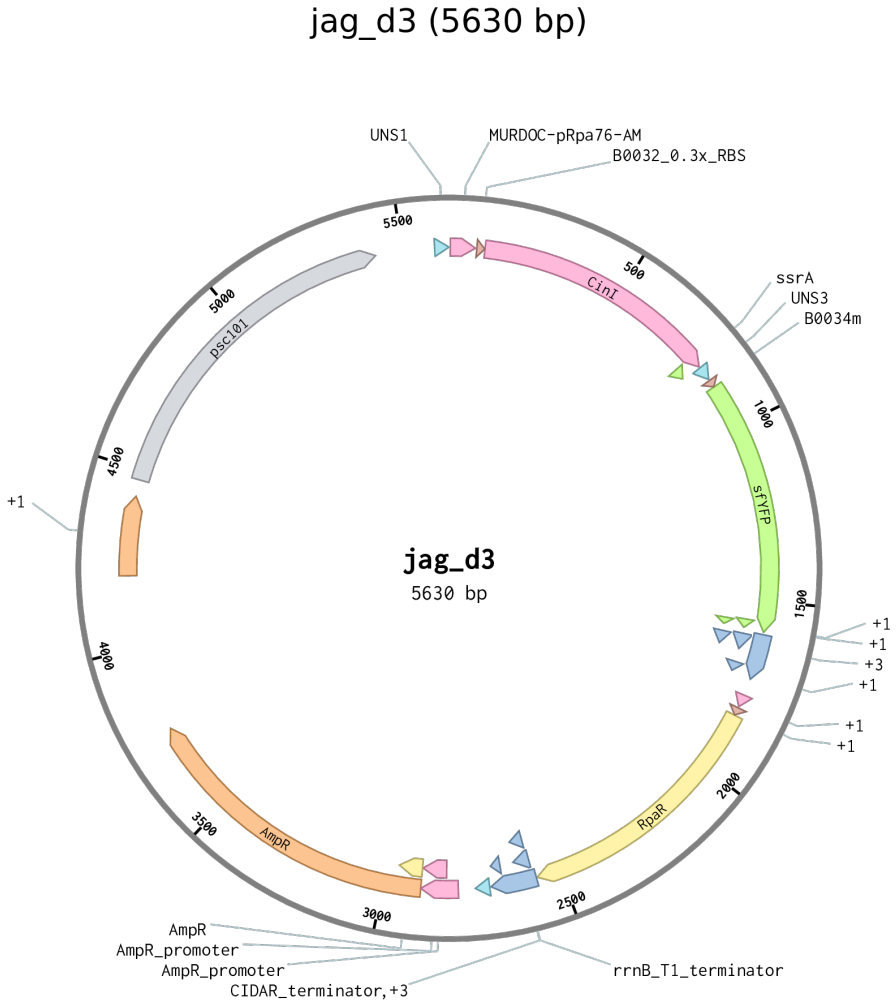

Figure D.9: Plasmid map for Cin sender plasmid jag_d3.

BIBLIOGRAPHY

Alon, Uri (2006). *An introduction to systems biology: Design principles of biological circuits*. ISBN: 9781420011432.

Avis, Tyler J. et al. (July 2008). "Multifaceted beneficial effects of rhizosphere microorganisms on plant health and productivity". In: *Soil Biology and Biochemistry* 40.7, pp. 1733–1740. ISSN: 00380717. DOI: 10.1016/j.soilbio.2008.02.013.

Balagaddé, Frederick K. et al. (2008). "A synthetic Escherichia coli predator-prey ecosystem". In: *Molecular Systems Biology* 4.187, p. 187. ISSN: 17444292. DOI: 10.1038/msb.2008.24.

Balleza, Enrique, J. Mark Kim, and Philippe Cluzel (2018). "Systematic characterization of maturation time of fluorescent proteins in living cells". In: *Nature Methods* 15.1, pp. 47–51. ISSN: 15487105. DOI: 10.1038/nmeth.4509.

Basu, Subhayu et al. (Apr. 2005). "A synthetic multicellular system for programmed pattern formation". In: *Nature* 434.7037, pp. 1130–1134. ISSN: 0028-0836. DOI: 10.1038/nature03461.

Belousov, B. P. (1959). "A periodic reaction and its mechanism." In: *Collection of short papers on radiation medicine*.

Berlec, Aleš et al. (2015). "In vivo imaging of Lactococcus lactis, Lactobacillus plantarum and Escherichia coli expressing infrared fluorescent protein in mice". In: *Microbial Cell Factories*. ISSN: 14752859. DOI: 10.1186/s12934-015-0376-4.

Bulgarelli, Davide et al. (Aug. 2012). "Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota". In: *Nature* 488.7409, pp. 91–95. ISSN: 0028-0836. DOI: 10.1038/nature11336. URL: http://www.nature.com/articles/nature11336.

Cao, Yangxiaolu et al. (Apr. 2016). "Collective Space-Sensing Coordinates Pattern Scaling in Engineered Bacteria". In: *Cell* 165.3, pp. 620–630. ISSN: 10974172. DOI: 10.1016/j.cell.2016.03.006.

Chen, Ye et al. (Aug. 2015). "Emergent genetic oscillations in a synthetic microbial consortium". In: *Science* 349.6251, pp. 986–989. ISSN: 10959203. DOI: 10.1126/science.aaa3794. arXiv: arXiv:1011.1669v3.

Cheng, Xianrui and James E. Ferrell (Aug. 2018). "Apoptosis propagates through the cytoplasm as trigger waves". In: *Science* 361.6402, pp. 607–612. ISSN: 0036-8075. DOI: 10.1126/science.aah4065. URL: https://www.sciencemag.org/lookup/doi/10.1126/science.aah4065.

Christian, David A. et al. (Oct. 2014). "Use of transgenic parasites and host reporters to dissect events that promote interleukin-12 production during toxoplasmosis". In: *Infection and Immunity* 82.10, pp. 4056–4067. ISSN: 10985522. DOI: 10.1128/IAI.01643-14.

Daeffler, Kristina N-M et al. (2017). "Engineering bacterial thiosulfate and tetrathionate sensors for detecting gut inflammation". In: *Molecular Systems Biology* 13.4, p. 923. ISSN: 1744-4292. DOI: 10.15252/msb.20167416. arXiv: 1701.00732.

Dalchau, Neil et al. (Dec. 2018). "Computing with biological switches and clocks". In: *Natural Computing* 17.4, pp. 761–779. ISSN: 15729796. DOI: 10.1007/s11047-018-9686-x.

Danino, Tal et al. (2010). "A synchronized quorum of genetic clocks". In: *Nature* 463.7279, pp. 326–330. ISSN: 00280836. DOI: 10.1038/nature08753. arXiv: 15334406.

Dieterle, Paul B et al. (2020). "Dynamics of diffusive cell signaling relays". In: *eLife* 9. ISSN: 2050-084X. DOI: 10.7554/elife.61771.

Dilanji, Gabriel E. et al. (2012). "Quorum activation at a distance: Spatiotemporal patterns of gene regulation from diffusion of an autoinducer signal". In: *Journal of the American Chemical Society* 134.12, pp. 5618–5626. ISSN: 00027863. DOI: 10.1021/ja211593q.

Dockery, Jack D. and James P. Keener (2001). "A mathematical model for quorum sensing in Pseudomonas aeruginosa". In: *Bulletin of Mathematical Biology* 63.1, pp. 95–116. ISSN: 00928240. DOI: 10.1006/bulm.2000.0205.

Doelman, Arjen (2019). *Pattern formation in reaction-diffusion systems- an explicit approach*. Tech. rep., pp. 129–182. DOI: 10.1142/9789813239609_0004.

Donaldson, Gregory P., S. Melanie Lee, and Sarkis K. Mazmanian (2015). "Gut biogeography of the bacterial microbiota". In: *Nature Reviews Microbiology* 14.1, pp. 20–32. ISSN: 17401534. DOI: 10.1038/nrmicro3552. arXiv: arXiv: 1011.1669v3.

Doong, Joy, James Parkin, and Richard M. Murray (2017). "Length and time scales of cell-cell signaling circuits in agar". In: *bioRxiv*, pp. 1–19. DOI: 10.1101/220244.

Farr, B. and W.M. Farr (2015). "kombine: a kernel-density-based, embarrassingly parallel ensemble sampler". In: *in preparation*.

Foreman-Mackey, Daniel et al. (Feb. 2012). "emcee: The MCMC Hammer". In: *Publications of the Astronomical Society of the Pacific* 125.925, pp. 306–312. DOI: 10.1086/670067. arXiv: 1202.3665.

Gantner, Stephan et al. (May 2006). "In situ quantitation of the spatial scale of calling distances and population density-independent N-acylhomoserine lactone-mediated communication by rhizobacteria colonized on plant roots". In: *FEMS Microbiology Ecology* 56.2, pp. 188–194. ISSN: 01686496. DOI: 10.1111/j.1574-6941.2005.00037.x.

Garcia-Ojalvo, Jordi, Michael B. Elowitz, and Steven H. Strogatz (July 2004). "Modeling a synthetic multicellular clock: Repressilators coupled by quorum sensing". In: *Proceedings of the National Academy of Sciences of the United States of America* 101.30, pp. 10955–10960. ISSN: 00278424. DOI: 10.1073/pnas.0307095101.

Gelens, Lendert, Graham A. Anderson, and James E. Ferrell (2014). "Spatial trigger waves: Positive feedback gets you a long way". In: *Molecular Biology of the Cell* 25.22, pp. 3486–3493. ISSN: 19394586. DOI: 10.1091/mbc.E14-08-1306.

Gines, G et al. (2017). "Microscopic agents programmed by DNA circuits". In: *NATURE NANOTECHNOLOGY* / 12. DOI: 10.1038/NNANO.2016.299.

Goentoro, Lea et al. (2009). "The Incoherent Feedforward Loop Can Provide Fold-Change Detection in Gene Regulation". In: *Molecular Cell* 36.5, pp. 894–899. ISSN: 10972765. DOI: 10.1016/j.molcel.2009.11.018. arXiv: NIHMS150003.

Goldbeter, Albert (2006). "Oscillations and waves of cyclic AMP in Dictyostelium: A prototype for spatio-temporal organization and pulsatile intercellular communication". In: *Bulletin of Mathematical Biology* 68.5, pp. 1095–1109. ISSN: 15229602. DOI: 10.1007/s11538-006-9090-z.

Gupta, Apoorv et al. (2017). "Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit". In: *Nature Biotechnology* 35.3, pp. 273–279. ISSN: 15461696. DOI: 10.1038/nbt.3796. arXiv: 15334406.

Gupta, Sonali et al. (Dec. 2020). "Investigating the dynamics of microbial consortia in spatially structured environments". In: *Nature Communications* 11.1. ISSN: 20411723. DOI: 10.1038/s41467-020-16200-0.

Halleran, Andrew D., Emanuel Flores-Bautista, and Richard M. Murray (Mar. 2019). "Quantitative characterization of random partitioning in the evolution of plasmid-encoded traits". In: p. 594879. DOI: 10.1101/594879. URL: https://doi.org/10.1101/594879.

Hogg, David W, Jo Bovy, and Dustin Lang (n.d.). *Data analysis recipes: Fitting a model to data **. Tech. rep. arXiv: 1008.4686v1.

Holzer, Matt, Arjen Doelman, and Tasso J. Kaper (2013). *Existence and stability of traveling pulses in a reaction-diffusion- mechanics system*. Tech. rep. 1, pp. 129–177. DOI: 10.1007/s00332-012-9147-0.

Hsiao, Victoria et al. (2016). "A population-based temporal logic gate for timing and recording chemical events". In: *Molecular Systems Biology* 12.5, p. 869. ISSN: 1744-4292. DOI: 10.15252/msb.20156663.

Jain, A. et al. (2012). "Microbial consortium-mediated reprogramming of defence network in pea to enhance tolerance against Sclerotinia sclerotiorum". In: *Journal of Applied Microbiology*. ISSN: 13645072. DOI: 10.1111/j.1365-2672.2011.05220.x.

Jin, Zi Jing et al. (2020). "Identification of a Strong Quorum Sensing- And Thermo-Regulated Promoter for the Biosynthesis of a New Metabolite Pesticide Phenazine-1-carboxamide in Pseudomonas strain PA1201". In: *ACS Synthetic Biology* 9.7, pp. 1802–1812. ISSN: 21615063. DOI: 10.1021/acssynbio.0c00161.

Johns, Nathan I. et al. (2016). "Principles for designing synthetic microbial communities". In: *Current Opinion in Microbiology* 31, pp. 146–153. ISSN: 18790364. DOI: 10.1016/j.mib.2016.03.010.

Kaplan, H. B. and E. P. Greenberg (1985). "Diffusion of autoinducer is involved in regulation of the Vibrio fischeri luminescence system". In: *Journal of Bacteriology* 163.3, pp. 1210–1214. ISSN: 00219193. DOI: 10.1128/jb.163.3.1210-1214.1985.

Kim, Hyun Jung, Wenbin Du, and Rustem F. Ismagilov (2011). "Complex function by design using spatially pre-structured synthetic microbial communities: Degradation of pentachlorophenol in the presence of Hg(ii)". In: *Integrative Biology* 3.2, pp. 126–133. ISSN: 17579694. DOI: 10.1039/c0ib00019a.

Kim, Jae Kyoung et al. (2019). "Long-range temporal coordination of gene expression in synthetic microbial consortia". In: *Nature Chemical Biology*. ISSN: 15524469. DOI: 10.1038/s41589-019-0372-9.

Kong, Wentao et al. (2018). "Designing microbial consortia with defined social interactions". In: *Nature Chemical Biology* 14.8, pp. 821–829. ISSN: 15524469. DOI: 10.1038/s41589-018-0091-7.

Langebrake, Jessica B. et al. (2014). "Traveling waves in response to a diffusing quorum sensing signal in spatially-extended bacterial colonies". In: *Journal of Theoretical Biology* 363, pp. 53–61. ISSN: 10958541. DOI: 10.1016/j.jtbi.2014.07.033.

Larkin, Joseph W. et al. (Aug. 2018). "Signal Percolation within a Bacterial Community". In: *Cell Systems* 7.2, 137–145.e3. ISSN: 24054720. DOI: 10.1016/j.cels.2018.06.005.

Lithgow, James K. et al. (2000). "The regulatory locus cinRI in Rhizobium leguminosarum controls a network of quorum-sensing loci". In: *Molecular Microbiology* 37.1, pp. 81–97. ISSN: 0950382X. DOI: 10.1046/j.1365-2958.2000.01960.x.

Liu, Chenli et al. (2011). "Sequential establishment of stripe patterns in an expanding cell population". In: *Science* 334.6053, pp. 238–241. ISSN: 10959203. DOI: 10.1126/science.1209042. arXiv: 20.

Macia, Javier et al. (2016). "Implementation of Complex Biological Logic Circuits Using Spatially Distributed Multicellular Consortia". In: *PLoS Computational Biology* 12.2. ISSN: 15537358. DOI: 10.1371/journal.pcbi.1004685.

Mimee, Mark et al. (July 2015). "Programming a Human Commensal Bacterium, Bacteroides thetaiotaomicron, to Sense and Respond to Stimuli in the Murine Gut Microbiota". In: *Cell Systems* 1.1, pp. 62–71. ISSN: 24054712. DOI: 10.1016/j.cels.2015.06.001. arXiv: 15334406.

Noorbakhsh, Javad et al. (2015). "Modeling oscillations and spiral waves in Dictyostelium populations". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 91.6, p. 62711. ISSN: 15502376. DOI: 10.1103/PhysRevE.91.062711.

Nunan, Naoise et al. (Jan. 2006). "Quantification of the in situ distribution of soil bacteria by large-scale imaging of thin sections of undisturbed soil". In: *FEMS Microbiology Ecology* 37.1, pp. 67–77. DOI: 10.1111/j.1574-6941.2001.tb00854.x.

Oleinik, O.A., A.N. Kolmogorov, and N.S. Piskunov (2019). *Studies of the Diffusion with the Increasing Quantity of the Substance; Its Application to a Biological Problem **. Tech. rep., pp. 106–132. DOI: 10.1201/9780367810504-7.

Payne, Stephen et al. (2013). "Temporal control of self-organized pattern formation without morphogen gradients in bacteria". In: *Molecular Systems Biology* 9.1, p. 697. ISSN: 17444292. DOI: 10.1038/msb.2013.55.

Pearson, James P., Christian Van Delden, and Barbara H. Iglewski (1999). "Active efflux and diffusion are involved in transport of Pseudomonas aeruginosa cell-to-cell signals". In: *Journal of Bacteriology* 181.4, pp. 1203–1210. ISSN: 00219193. DOI: 10.1128/jb.181.4.1203-1210.1999.

Potvin-Trottier, Laurent et al. (2016). "Synchronous long-term oscillations in a synthetic gene circuit". In: *Nature* 538.7626, pp. 514–517. ISSN: 14764687. DOI: 10.1038/nature19841.

Prindle, Arthur, Jintao Liu, et al. (2015). "Ion channels enable electrical communication in bacterial communities". In: *Nature* 527.7576, pp. 59–63. ISSN: 14764687. DOI: 10.1038/nature15709. arXiv: 15334406.

Prindle, Arthur, Phillip Samayoa, et al. (Jan. 2012). "A sensing array of radically coupled genetic 'biopixels"'. In: *Nature* 481.7379, pp. 39–44. ISSN: 00280836. DOI: 10.1038/nature10722.

Reátegui, Eduardo et al. (2017). "Microscale arrays for the profiling of start and stop signals coordinating human-neutrophil swarming". In: *Nature Biomedical Engineering* 1.7. ISSN: 2157846X. DOI: 10.1038/s41551-017-0094.

Redfield, Rosemary J. (Aug. 2002). "Is quorum sensing a side effect of diffusion sensing?" In: *Trends in Microbiology* 10.8, pp. 365–370. ISSN: 0966842X. DOI: 10.1016/S0966-842X(02)02400-9.

Saini, Mukesh et al. (2016). "Production of biobutanol from cellulose hydrolysate by the Escherichia coli co-culture system". In: *FEMS Microbiology Letters* 363.4. ISSN: 15746968. DOI: 10.1093/femsle/fnw008.

Sarma, Birinchi Kumar et al. (Aug. 2015). *Microbial consortium-mediated plant defense against phytopathogens: Readdressing for enhancing efficacy*. DOI: 10.1016/j.soilbio.2015.04.001.

Schaerli, Yolanda et al. (2014). "A unified design space of synthetic stripe-forming networks". In: *Nature Communications* 5, p. 4905. ISSN: 20411723. DOI: 10.1038/ncomms5905. arXiv: arXiv:1011.1669v3.

Scott, Spencer R. et al. (Aug. 2017). "A stabilized microbial ecosystem of self-limiting bacteria using synthetic quorum-regulated lysis". In: *Nature Microbiology* 2.8, p. 17083. ISSN: 20585276. DOI: 10.1038/nmicrobiol.2017.83.

Steidler, Lothar, Pieter Rottiers, and Bernard Coulie (2009). "Actobiotics™ as a novel method for cytokine delivery: The interleukin-10 case". In: *Annals of the New York Academy of Sciences*. Vol. 1182, pp. 135–145. ISBN: 9781573317832. DOI: 10.1111/j.1749-6632.2009.05067.x.

Taillefumier, Thibaud et al. (2017). "Microbial consortia at steady supply". In: *eLife* 6, pp. 1–65. ISSN: 2050084X. DOI: 10.7554/eLife.22644. arXiv: 1604.02733.

Tayar, Alexandra M. et al. (Sept. 2015). "Propagating gene expression fronts in a one-dimensional coupled system of artificial cells". In: *Nature Physics* 11.12, pp. 1037–1041. ISSN: 17452481. DOI: 10.1038/nphys3469.

Tecon, Robin et al. (Sept. 2018). "Cell-to-cell bacterial interactions promoted by drier conditions on soil surfaces". In: *Proceedings of the National Academy of Sciences of the United States of America* 115.39, pp. 9791–9796. ISSN: 10916490. DOI: 10.1073/pnas.1808274115.

Temme, Karsten (2019). "2019-Pivot-Bio-Performance-Report". In: *blog.pivotbio.com*.

Tsoi, Ryan et al. (Mar. 2018). "Metabolic division of labor in microbial systems". In: *Proceedings of the National Academy of Sciences of the United States of America* 115.10, pp. 2526–2531. ISSN: 10916490. DOI: 10.1073/pnas.1716888115.

Voigt, Christopher A. (Dec. 2020). "Synthetic biology 2020–2030: six commercially-available products that are changing our world". In: *Nature Communications* 11.1, p. 6379. ISSN: 2041-1723. DOI: 10.1038/s41467-020-20122-2.

Watts, Duncan J. and Steven H. Strogatz (1998). "Collective dynamics of 'small-world9 networks". In: *Nature* 393.6684, pp. 440–442. ISSN: 00280836. DOI: 10.1038/30918. arXiv: 0803.0939v1.

Willits, Tracy (Feb. 2020). *Pivot Bio PROVEN™ Creates Sustainably Self-Fertilizing Corn.*

Wolpert, L. (1969). "Positional information and the spatial pattern of cellular differentiation". In: *Journal of Theoretical Biology*. ISSN: 10958541. DOI: 10.1016/S0022-5193(69)80016-0.

Xue, Xiaoru, Chuan Xue, and Min Tang (June 2018). "The role of intracellular signaling in the stripe formation in engineered Escherichia coli populations". In: *PLoS Computational Biology* 14.6. Ed. by Kevin Painter, e1006178. ISSN: 15537358. DOI: 10.1371/journal.pcbi.1006178.

Yates, Edwin A. et al. (Oct. 2002). "N-acylhomoserine lactones undergo lactonolysis in a pH-, temperature-, and acyl chain length-dependent manner during growth of Yersinia pseudotuberculosis and Pseudomonas aeruginosa". In: *Infection and Immunity* 70.10, pp. 5635–5646. ISSN: 00199567. DOI: 10.1128/IAI.70.10.5635-5646.2002.

Youk, Hyun and Wendell A. Lim (2014). "Secreting and sensing the same molecule allows cells to achieve versatile social behaviors". In: *Science* 343.6171, pp. 1242782–1242782. ISSN: 10959203. DOI: 10.1126/science.1242782.

Young, Jonathan W. et al. (2012). "Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy". In: *Nature Protocols* 7.1, pp. 80–88. ISSN: 17542189. DOI: 10.1038/nprot.2011.432.

Zhang, Haoran and Gregory Stephanopoulos (2016). "Co-culture engineering for microbial biosynthesis of 3-amino-benzoic acid in Escherichia coli". In: *Biotechnology Journal* 11.7, pp. 981–987. ISSN: 18607314. DOI: 10.1002/biot.201600013.

Zomorrodi, Ali R. and Daniel Segrè (2016). "Synthetic Ecology of Microbes: Mathematical Models and Applications". In: *Journal of Molecular Biology* 428.5, pp. 837–861. ISSN: 10898638. DOI: 10.1016/j.jmb.2015.10.019.