Biological patterning in networks of interacting cells

Mindy Perkins

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2020-51
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2020/EECS-2020-51.html

May 18, 2020

Biological patterning in networks of interacting cells

by

Melinda Liu Perkins

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Murat Arcak, Chair
Professor Claire Tomlin
Assistant Professor Hernan G. Garcia

Spring 2020

The dissertation of Melinda Liu Perkins, titled Biological patterning in networks of interacting cells, is approved:

Chair _____ Date _____

_____ Date _____

_____ Date _____

University of California, Berkeley

Biological patterning in networks of interacting cells

Abstract

Biological patterning in networks of interacting cells

by

Melinda Liu Perkins

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Murat Arcak, Chair

Biological organisms rely on spatial variation in cell activity to coordinate diverse activities, such as microbial colonization and embryonic development. Interactions among neighboring cells play a crucial role in generating spatial patterns spontaneously from stochastic initial conditions or by refining simple spatially varying inputs, such as chemical concentration gradients, into complex outputs, such as stripes in gene expression. The ability to synthetically engineer multicellular patterning will facilitate advances in designing microbial communities, creating synthetic biomaterials, and programming tissue and organ growth, among other applications. Researchers face numerous theoretical and practical challenges to implementing patterning in synthetic biological setups, in part because living systems are extremely high dimensional and nonlinear. Thus, a handful of theories have guided most experimental efforts to implement multicellular patterning, and most successes have relied primarily on trial and error or numerical simulation. Future progress in synthetic multicellular patterning will benefit from interpretable mathematical theory for patterning phenomena, coupled with experimental platforms for validating these theories in practice.

In this thesis, I develop methods for understanding and implementing multicellular patterning through the lens of networked dynamical systems, in which individual cells are modeled as dynamical systems and their interactions are modeled as networks. With this mathematical representation, cell behavior and network structure can be partially decoupled, facilitating both analysis and intuition building. In Chapter 2, I introduce a spatial filtering approach to biological pattern refinement, in which a network of interacting cells is treated as a "filter" for interpreting a prepattern into a different output pattern. This approach separates the exact form of the prepattern from the resulting readout, enabling a spatial frequency-based interpretation of pattern refinement that is conducive to analysis with extant intuition from signal processing. In the remaining two chapters, I apply theory for contrasting patterning to two unique synthetic biological setups. Chapter 3 focuses on a platform in which colonies of bacteria physically connected by channels interact through the exchange of diffusible

molecules. Spatial parameters such as the distance between colonies and their placement is exploited to modulate patterning features, including contrast level and stability. Chapter 4 describes a setup in which chemical interaction among yeast cells is substituted with computer-controlled inputs that are adjusted in real time based on the measured gene expression levels. Theory predicted spontaneous contrasting patterning with qualitative and quantitative accuracy, demonstrating the potential benefits of this experimental platform for future avenues in multicellular patterning research.

A central theme throughout this thesis is how exploiting a dynamical systems framework can provide new insights into (possible) patterning behaviors by considering them with a different perspective from the usual. For example, I show in Chapter 2 that random, constant-in-time spatial variation can generate Turing-like patterns when system dynamics are stable; conventional wisdom requires an instability to observe such patterns. Similarly, in Chapter 4, I show how "lateral inhibition" can be achieved without cells physically communicating or even physically neighboring each other—an impossible paradigm in customary biological setups. Moving forward, I hope this work will stand as just one example of how an interdisciplinary approach can shed light on the structure and function of living systems, evolved and engineered alike.

To my father's mother, Carolyn G. Perkins,
and my mother's father, Dr. Chien Liu,
who saw me into graduate school
and are with me in spirit as I leave.

I would be honored to follow in your footsteps.

Contents

List of Figures

List of Tables

Acknowledgments

When I was interviewing for Ph.D. programs, a wise young assistant professor told me that "the end product of grad school is the grad student". His advice has stuck with me throughout these past five years of intense personal and intellectual growth. If I am indeed a better end product than when I arrived at Berkeley, I owe it to the people who have defined my time here in the Bay Area and elsewhere—whose kindness and support, guidance and understanding, have changed time and time again the course of my life. What follows can only ever be a partial list; know I am grateful for you whether you find your name here or not.

My research career (such as it is) began during undergraduate at Stanford. As early as freshman year I was lucky to meet Marvin Diogenes and Prof. Thomas S. Grey, who endorsed my first funded project, and Prof. Ellen Porzig, whose passion, respect, and advice inspired me to chart a course for my learning. I thank my research advisors Prof. John Pauly, for helping me carry my studies in signal processing to an interdisciplinary application, and Prof. Elizabeth Hadly, for welcoming an engineer to her biology lab with earnest energy and incisive feedback. I continue to be thankful for then-Ph.D.-candidate, now-professor Hannah K. Frank, who took me under her wing for no better reason than that we both like bats, and who guided me through my first steps into research with enthusiasm, consideration, and realism.

Every graduate student has an advisor, but not every advisor is a mentor. I was fortunate to have both in one. Murat Arcak was the first professor to meet with me during EECS Visit Days, and perhaps my single best decision as a Ph.D. student was to join his lab the following year. His gentle encouragement, paired with sound judgment in matters theoretical and practical, gave me both the space and the nourishment to grow. He made time for me no matter how busy he was or how little I had to communicate except my own consternation—which he would doubtless address with a reasonable suggestion or calm intervention. From him, I learned how to be curious and patient when pursuing new research directions, as well as how to balance my energy, prioritize my responsibilities, and value the most important aspects of any endeavor. I remember Skyping him one summer in a panic that the experiments he'd sent me overseas to run weren't working; I was terrified I wouldn't have results by the time I returned to Berkeley. He told me that as far as he was concerned, the project was already a success, because I was *learning* from the experience. He was, as always, looking out for me as much as for my career. I may never attain his level-headedness, but I can at least aspire to the deep thinking, careful attention, and humility that have characterized his work as a researcher, teacher, and mentor. I hope that in some small way I have made him proud.

In Murat's lab, I thank Marcella Gomez for offering mentorship early on and acting as a role model for unpretentious dedication and work-life balance. I was fortunate to have a neighbor in Eric Kim, whose quiet advice was invaluable to a floundering second- and third-year, and whose unique way of thinking never failed to cast ideas in a new light. Although we no longer sit in the same office space, I'm glad to have met Mikhail Burov,

Alex Devonport, Pierre-Jean Meyer, and Stan Smith, who along with Eric provided good company at occasional lab lunches on Euclid; as well as Matthias Hirche, for his amiability and boundless patience with my fumbling German over afternoon tea. Warmest thanks also to my collaborators Prof. Adam Arkin, Justin Hsia, and Mika Tei, who gave me my first glimpses into synthetic biology and the art of applying theory to experiments, and to Prof. Hernan G. Garcia and Jake Zhao, who have been instrumental in shaping my academic progression toward developmental biology.

The EECS Department at Berkeley has proved to be a welcoming harbor in an otherwise intimidatingly large institution. In my first two years I benefited especially from the Women in Computer Science and Electrical Engineering, which hosted lunches where I made many friends, and trusted me to act as social co-chair for a year. I've had several memorable encounters with EECS faculty, though I give special thanks to Prof. Miki Lustig, an academic cousin of sorts, for serving as my temporary advisor and sitting on my quals committee. I remain beholden to Prof. John DeNero (and all my students in CS 195) for giving me the incredible opportunity to lecture and develop course content from scratch for a subject outside my research area that nevertheless matters a good deal to me. And no acknowledgments would be complete without recognizing Shirley Salanio for cheerfully and skillfully managing my interactions with the university bureaucracy.

Among my most transformative months at Berkeley actually took place an ocean away from campus, in Basel, Switzerland, at the ETH Zürich Department of Biosystems Science and Engineering. Prof. Dr. Mustafa Khammash generously hosted me for six months in his lab, where I am indebted to him and his group for kindling my curiosity with their open-mindedness, intellectual engagement, and genuine friendship. I am particularly grateful to my collaborators Dirk Benzinger and Marc Rul·lan, without whose patience and insight I would never have started (much less finished) any experiments. For their good company in the office and out, I thank Dirk for his sardonic yet good-hearted humor; Sara Dionisi, whose focus, dedication, internal strength, and good taste were intensely inspiring; Timothy Frei, who furnished many a thoughtful discussion about a range of subjects with admirable nuance; Joaquín Gutiérrez, who shared his dry wit, philosophical musings, and recommendations for music and literature over fondue or hot chocolate; Linh Ho, who invited me to Art Basel, took me for a hike, spent seven hours making dumplings with me, and somehow still concluded I was worthy of her generosity; Serguei Ovinnikov, whose upbeat attitude and atmospheric music lightened my hours at the bench; and Maaike Welling, who introduced me to sledging in the Alps and provided an invaluable perspective on science as a profession. A final shoutout to Steph Aoki, Ankit Gupta, Sant Kumar, Jan Mikelsson, Prof. Guy-Bart Stan, Manu Upadhyaya, and everyone else I encountered in the lab, who all at some point or another made a fantastic half year even more special. I can't wait for another swim together in the Rhine.

Throughout graduate school, I have been gifted with the company of dear friends. Heidi Kleven has been a bastion of sunniness, propriety, and bravery since the beginning; together we've carved pumpkins and baked muffins, crossed countries and climbed mountains, laughed and cried in the span of a symphony. For her influence in shaping my path here, I'm equally

obliged to Neha Wadia, among the most dignified, well-versed, self-possessed, aware, and understanding of human beings, whose ability to tackle challenges with honesty and aplomb knows no parallel in my acquaintance.

The indomitable Regina Eckert I first met at our hotel stay during Visit Days, and thereafter I cannot express my gratitude enough for our jaunts: the walks and hikes, the birthday parties, department events and Christmas fairs; the planning and executing, complaining and commiserating, and of course, the Stanford Alumni volunteer event which she attended as a guest wearing Berkeley gear. Without her listening ear, her unwavering support, maturity, and daring, I do not doubt my journey here would have been dark and tortuous indeed—though I do hope our future paths are tortuous enough to cross many a time again.

To my once mentee, at-least-twice mentor, countless-times friend Sarah Sterman, I am also thankful beyond words. Anyone who loves dragons, plays viola, and went to Stanford for undergraduate must be a kindred spirit of mine, even if I've nothing akin to her resourcefulness, persistence, storytelling savvy, and backcountry skills. She kept me and my ideas fed and (in)sane during backpacking trips, B-grade fantasy films, home-cooked dinners, workouts, plotting and replotting, and lazy days on the couch. Although visiting her may never again be as easy as ascending one flight of stairs, I can think of a few peaks we may yet summit together, if I have my way.

I've benefitted too from the energetic company of Justin Yim, who readily shares his encyclopedic knowledge and excitement about robotics, animals, Matlab, and Wikipedia with as much passion as compassion, earnestness as sensitivity. Others I'm lucky to know include Emma Alexander, Carol Baumbauer, Carolyn Chen, non-Berkeley-but-SF-native Laurel Estes, Daria Genzel-Wehrfritz, Grace Kuo, Kristina Monakhova, Alyssa Morrow, Sandra Muroy, Emily Naviasky, and Fanny Yang, who have joined me at various times in bike riding, bouldering, crafting, concertgoing, cooking, dim sum, hiking, road tripping, running, tea, train riding, and Zooming, among other outings (or innings). I thank everyone who sightread chamber music with me for keeping a part of my soul alive; the Writer's Coven for tackling fantastical journeys with limitless enthusiasm; and my housemates for making me feel at home, especially during the final weeks of graduate school when pandemic locked us up together. And of course, I am grateful for Jaro Ferenc, with whom (despite the distance) I've embarked this past year upon an adventure that surprises me continually with its gifts. If indeed *hic sunt dracones*, I look forward to discovering them together.

To Berkeley I may owe my Ph.D., but my education I owe to my parents, who showed me new worlds: my mother in books and at the museum, my father in mathematics and the mountains. From my parents I learned what it means to work hard, take responsibility, seek truth, and strive to be someone I could be proud of. To my sister, I owe everything else. She has acted as my compass in matters moral, emotional, and geographical, modeling for me the culture of character—of conscientiousness, of commitment, of discipline, of depth—that most admire and few achieve. Thank you for never letting others tell me who I was, and for showing me all that I could aspire to be. I'm blessed to call you my family.

Chapter 1

Introduction

"[L]ife is infinitely stranger than anything which the mind of man could invent."

– Sherlock Holmes, in [1]

In some sense, life is a narrowing of possibilities: the process of selecting a particular set of outcomes from an infinite space of physically permissible arrangements of matter. In biology, the process operates at nested levels from the replication of chemical materials through the transformation of fertilized eggs into adult organisms, all the way up to physiological adaptation, evolution, and arguably cultural and sociological phenomena as well. The activities and interactions at each level both enable and constrain activities and interactions at other levels, generating what many of us would intuitively refer to as order or structure in the appearance, behavior, or configuration of the elements at each level. We might then apply the term "pattern" to *an observable, coherent set of such configurations, sharing certain discernible features in common, that arises reliably from a specified process.* Setting aside the philosophical question of whether the selection process itself generates order, or whether we merely define order with respect to the outcomes we observe to be favored by the selection process, we might ask: Given a particular set of mechanisms operating at a particular set of levels—a "system"—how do we predict, understand, or develop intuition for the emergence of patterns? What does pattern emergence tell us about the living creatures we see around us, and the biological possibilities that have not (yet) been explored?

This thesis concerns itself with a microscopic facet of life, situated between the levels of biochemical reactions in semi-isolated compartments and interactions among these compartments and with the environment. Practically, these compartments comprise one or multiple biological cells producing proteins, and the patterns that will be considered are differences in the concentration of protein products across compartments—specifically those resulting from differential levels of *gene expression* among cells. Since the compartments are understood to occupy separate physical locations, the differing expression levels can be visualized as recognizable spatial patterns including stripes, spots, or checkerboards. Such *multicellular gene expression patterning* is fundamental to organizing diverse biological activities, including cooperation among colonial microorganisms and body planning in developing embryos.

Encouraged by natural examples, genetic engineers have successfully modified living organisms to achieve multicellular patterning with aspirational applications in chemical production, biomaterials, and regenerative medicine. Yet theory often lags behind technological and experimental progress; many successes in multicellular patterning [181, 182, 266] have relied primarily upon numerical simulation of whole systems (e.g., [149,163,165]) rather than a holistic theory or methodological, modular characterization, and a few have not employed models at all (e.g., [180]). In contrast, conventional engineering approaches rely on the predictive power of theory both to design complex systems from simpler constituent parts and to build the intuition necessary to envision new capabilities. Future progress in synthetic multicellular patterning will benefit from a firm understanding of the underlying theoretical principles, as well as scalable, efficient methods for implementing—and validating—these principles in practice.

Here, I develop approaches to understanding multicellular gene expression patterning using techniques from signal processing and networked dynamical systems theory—tools traditionally employed in electrical and mechanical engineering. I then apply theoretical insights to help design and implement new experimental platforms for attaining multicellular patterning with genetically engineered microorganisms. My work demonstrates how taking a new theoretical perspective can reveal behaviors in existing systems that were not previously considered possible, as well as suggest new classes of systems capable of achieving identified behaviors.

The remainder of this section situates the thesis in the context of extant research in biological patterning. It begins with an introduction to theoretical paradigms for conceptualizing patterning from a century of work in developmental biology; progresses through a brief history of multicellular pattern formation in synthetic biology; justifies more generally the application of engineering approaches to biological research; and presents conceptual and mathematical conventions that will be used in the remainder of the work. This section concludes with an overview of the thesis organization and the contributions specific to each chapter.

1.1 Gene Expression Patterning in Natural Systems

There are many senses in which biological systems can be understood to generate patterns. On the cellular level, there are roughly two main classes of patterns: those that describe "shapes", or the placement of cells in space—e.g., the structure of a bacterial colony or the digits of a human hand—and those patterns that describe spatial variation in the activity or anatomy of individual cells—e.g., the color patterns on an animal's fur are produced by variation in the types and quantities of pigments produced by the cells that grow each hair. This thesis focuses on the latter type of pattern, specifically on cell-to-cell variation in gene expression.

Gene expression is the process by which a gene coded in DNA is *transcribed* into mRNA molecules that are then *translated* into protein molecules (Figure 1.1). These proteins deter-

mine many important elements of a cell's behavior and may generate effects that persist over long timescales. For example, the genes expressed by cells during embryonic development will determine the ultimate "identity" of the cell (e.g., a nerve or skin cell) in the adult organism. Environmental and biological cues, physical or chemical, contribute to determining whether, when, and how much a gene is expressed.

One particularly important avenue for modulating gene expression levels involves classes of proteins that interact locally with DNA so as to increase (promote) or decrease (inhibit or repress) the production rate for mRNA corresponding to a particular gene. In this case, the DNA-interacting protein is called a *transcription factor* because it directly influences whether mRNA for a particular gene (or set of genes) is transcribed. Because transcription factors are proteins coded by genes, their production is ultimately subject to regulation by other transcription factors. Collections of such interactions form *genetic circuits* that can be exceedingly complex in living organisms. It is not uncommon for the products of genetic circuits to cause cells to signal other cells in a way that changes those cells' gene expression levels, such that even relatively simple genetic circuits can generate sophisticated behaviors at the multicellular level. Indeed, cell-to-cell signaling is an essential ingredient in most (if not all) natural patterning systems, and will form the basis for the work presented in this thesis.

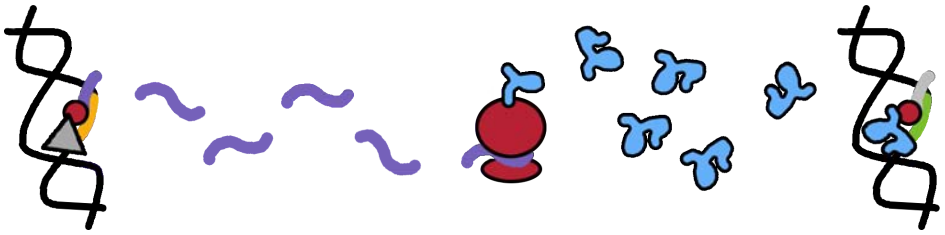

Figure 1.1: A schematic of gene expression. An input (gray triangle) modifies the *transcription rate* of mRNA (purple) from a gene (gold). The mRNA is then *translated* into protein (blue). The proteins may in turn modify the transcription rate of some other target gene (green). If the target gene is the same as the expressed gene, then the protein is said to be autoregulatory. Chemical, mechanical, or electrical signals from neighboring cells may also influence transcription rates (not pictured). Cell-to-cell variability in production rates that "persists" in time (i.e., is not due to the intrinsically stochastic nature of chemical interactions) may arise, for example, from variation in the concentrations of intercellular machinery (red circles) responsible for transcription and translation.

The remainder of this section provides a brief history of gene expression patterning in developmental biology, where a great deal of theoretical and experimental research has taken place. In keeping with the spirit of this thesis, emphasis is placed on conceptual frameworks with mathematical underpinnings rather than specific molecules or model organisms.

Reaction-diffusion and Positional Information

Biological patterns are easily found, and at least some aspects of their formation are readily observed, but *how* a particular pattern emerges is not straightforward to probe

or understand. Computer scientist Alan M. Turing is credited with the earliest formal mathematical hypothesis for spontaneous biochemical pattern formation, published in 1952. He envisioned a tissue with fixed geometrical form inside of which chemicals react with each other and with the tissue, promoting or inhibiting the production of other chemicals. Although "a perfectly homogeneous and possibly irregularly shaped mass of tissue [...] will continue indefinitely be homogeneous", statistical fluctuations in real systems can trigger pattern formation provided that "the system has an appropriate kind of instability"—that is, a dynamical system description indicates the homogeneous steady state is unstable. Turing demonstrated that such an instability could arise from diffusion of the involved chemicals, resulting in periodic, oscillatory, or traveling-wave patterns of chemical concentrations [2].

It is difficult to overstate the influence of Turing's lucid theoretical formalism on subsequent research in multicellular patterning [3–5]. Although decades would pass before mathematicians and experimental biologists recognized the importance of his ideas [6], his insights have since accumulated experimental evidence in such varied systems as feather patterning [7], hair follice spacing [8], fish skin striping [9], and mammalian palate ridges [10]. To date, most attempts to identify instances of Turing patterning in living systems have focused on the two-chemical "substrate depletion" or "activator-inhibitor" models introduced in 1972 by Alfred Gierer and Hans Meinhardt [11]. More recent work has expanded the scope of *reaction-diffusion* patterning to systems exhibiting localized or "solitary" (but not global) patterning [12,13], systems with space-dependent parameters [14–16], intrinsically stochastic chemical networks [17,18], systems capable of pattern "inception" regardless of whether the result is stable [19], generalized systems with local and nonlocal interactions [20], and locally mass-conserving systems that conceptually relate Turing instabilities to other patterning mechanisms [21,22].

Turing introduced the reaction-diffusion paradigm within a more general discussion of embryonic development, in which he postulated that chemical "morphogens" could induce growth and formation of distinct structures by "[catalyzing] the production of other morphogens" until eventually some substance was produced "whose duties [were] not purely catalytic" [2]. Thus, morphogens would guide initially indistinguishable, unspecified cells to adopt particular *fates* in an adult organism, with the identity and activity level of the morphogen determining the fate, e.g., a skin cell or a nerve cell. The exact characteristics of a "morphogen" have been a subject of some contention [23–25]—Turing himself had not intended the term to have "any very exact meaning" [2]—but his ideas foreshadowed modern understanding of genetic networks as chains of proteins modulating each other's production levels in context- and concentration-dependent fashion [25,26].

The notion of some "informative" substance specifying cell fate through signaling, thresholding, or some combination thereof had received attention in the early 1900s from scientists including Hilde Mangold, Hans Spemann, Leopold von Ubisch, and Hildegard F. Stumpf. These ideas were synthesized into their most famous form by Lewis Wolpert in a series of influential papers published in the late 1960s and early 1970s [27]. Wolpert distinguished between "mosaic" development, in which a single cell divides into many cells whose fates are specified by local interactions during division, and "regulative" development, in which cell

fates are specified by global interactions across many extant cells. Focusing on regulative development, he articulated the notion of *positional information*, which encodes a cell's relative location within a tissue and instructs further gene expression so that the cell adopts the fate appropriate to that location. His prototypical example was the "French flag model", in which cells arranged in a line must produce blue, white, or red pigment based on their position in the line. He discussed a few conceptual "algorithms" to accomplish this patterning, the most famous of which involves a preestablished gradient of some substance that decreases monotonically from one end of the line to the other. The absolute concentration of substance would then be sufficient to specify cell fate at the appropriate position in the line [28].

Following Wolpert's conceptualization—and Francis Crick's suggestion that morphogen gradients could be established through diffusion [29]—a spate of studies sought to identify just such molecules that "[spread] out from a localized source to form [...] a monotonic gradient [...] that determines, point by point, the responses of all cells in the field" [23]. Within thirty years researchers had identified putative morphogen gradients involved in chick limb development [30, 31], axolotl limb regeneration [32], fruit fly [33, 34] and zebrafish [35, 36] axis orientation, frog mesoderm formation [37], and vertebrate neurogenesis [38, 39], among other candidates with varying levels of experimental evidence (and agreement among scientists) [23, 24, 40]. Subsequent work expanded upon Wolpert's original conception to provide more detailed mathematical underpinnings, especially with regard to how concentration gradients were established and maintained [27, 41–46].

Scientists before Turing and Wolpert had postulated the existence of "formative substances", "organizers" [47], and gradients of signals that would guide the *differentiation* of cells into distinct fates [23, 27]. What reaction-diffusion and positional information contributed were compelling theoretical frameworks that linked previously proposed high-level concepts together with clearly hypothesized molecular means to particular patterning ends. These frameworks have remained appealing in part because they are sufficiently general to describe many systems, sufficiently specific to provide explanatory power, and sufficiently simple to permit intuitive understanding. Decades of mathematical and empirical studies alike have supported the basic tenets of reaction-diffusion and positional information, even as new observations have complicated the theories from their initial conceptions.

On the Shoulders of Giants

By the turn of the century enough experimental evidence had been accumulated to convince most developmental biologists that reaction-diffusion and morphogen gradient interpretation were reasonable explanations for observed patterning phenomena. Through their efforts to identify the involved biochemicals, researchers also made conceptual and empirical observations that drove the field to adopt a new concern. "[R]eliability," wrote Arthur D. Lander in 2007, "[or] as engineers call it, robustness [in] patterning, has become something of an obsession among experimentalists and theoreticians alike" [25].

The "obsession" arose in studies of both reaction-diffusion systems and morphogen gradient systems, but for somewhat different reasons. While the end of the 1900s saw several molecular candidates for gradient-forming morphogens, researchers had more difficulty identifying biochemical players in reaction-diffusion networks [48]. Overly "literal" interpretation of theoretical models debatably exacerbated the situation: Gierer and Meinhardt's two-species reaction-diffusion models [11] place stringent constraints on the kinetic rate parameters [49] that permit spontaneous patterning and typically require a large difference in diffusivity between the two species [50]—a requirement that is difficult to satisfy with biological molecules. Furthermore, even systems capable of patterning are not necessarily robust or reproducible, in that changes to boundary conditions [51] or initial conditions [52, 53] can alter or even eliminate patterning behaviors [54, 55]. This so-called "robustness problem" formed the basis for a number of works exploring networks of more than two species that are capable of spontaneous patterning in or near the Turing sense, with potentially more biologically plausible parameter values [56–66], or how mechanisms such as tissue growth might (de)stabilize reaction-diffusion patterning [67, 68].

Concurrently, researchers realized that the simplest model of morphogen gradient interpretation could not account for experimental observations. The original conception of a morphogen gradient suggested a purely feedforward network in which cells had only to read out their position relative to the morphogen concentration. But individual animals of the same species maintain their proportions despite variability in size, and evidence indicated that morphogen gradients tend to be produced by localized production, diffusion, and global degradation, which—left unregulated—would produce exponential gradients that do *not* scale with tissue size, and would therefore lead to errors in positional decoding by concentration alone. Theoreticians were obliged to suggest additional system constraints or regulatory mechanisms between morphogen and target gene that would scale gradients to match experimental observations [69–72]. Building on one of Wolpert's original suggestions [28], researchers also explored how positional decoding could be improved by processing two opposing gradients of morphogens [73–75].

Related challenges continued to emerge to the idea of target genes directly reading morphogen concentrations. Theoretical work identified limits to the positional information that could be decoded from gradients owing to intrinsic stochastic influences, and suggested strategies to buffer disruption from noise and cell-to-cell variability [76–84]. Meanwhile, experimental observations indicated that features other than morphogen concentration—such as the slope or temporal evolution of the gradient—might be key to the transfer of positional information [85–91]. These developments have led researchers to the next logical (and still open) question, namely, how the interplay between morphogen gradients and genetic networks actually *generates* an output pattern with appropriate features in the appropriate places [92–98].

Turing's and Wolpert's frameworks have historically been placed in opposition to each other [6], but there is no reason this must be so, as pointed out by Jeremy B. A. Green and James Sharpe in their 2015 paper illustrating how genetic networks might combine information both from a reaction-diffusion network and a morphogen gradient either sequentially

or in parallel [5]. Their argument highlights a growing need for integrative theoretical approaches to untangle the perplexing complexities behind natural patterning systems, most of which appear to resist classification into simple, siloed schemes. For now, it seems enough to predict that as researchers persist in pushing theoretical and experimental boundaries, the "two big ideas in developmental biology" [5] will continue to expand and adapt to the ever-evolving landscape of our scientific understanding.

From the Perches of Dwarfs

While discussions of gene expression patterning may have been dominated by reaction-diffusion and positional information, researchers have not restricted their studies or their mathematical models to systems that do not quite fit either scheme. For example, contemporaneously with Wolpert's investigations, Jonathan Cooke and Elizabeth C. Zeeman proposed a "clock-and-wavefront" model to describe the emergence of periodic, regularly spaced structures in embryonic vertebrates. According to this model, biochemical concentrations (forming a "segmentation clock") oscillate within cells at a rate faster than the anterior-to-posterior passage of a wavefront that "freezes" cells at some point in the oscillation [99]. Supported by experimental evidence for involved molecules, a moving morphogen gradient was proposed to generate the differentiation wavefront [100], and mathematical models were produced for the oscillators specifically [101, 102] as well as for the process as a whole [103–105]. Later models examined scaling [106, 107] and robustness to noise [108–110], and also challenged the initial premises by suggesting alternatives without the segmentation clock [111] or (separately) the gradient [105, 112, 113].

The clock-and-wavefront model addresses the appearance of periodic patterns along one dimension. An alternative mechanism is hypothesized to explain the emergence of checkerboard-like periodic patterns in two dimensions. As early as 1940, Sir Vincent B. Wigglesworth, in his study of the cuticle of kissing bugs, discovered that plaques and bristles would arise with regular spacing based not on absolute distance but on the number of cells separating them. His experimental observations led him to postulate that plaque- and bristle-destined cells inhibited other cells in the vicinity from becoming plaques and bristles [114]. Half a century later, Joanne Collier and colleagues revisited this idea in a different model system, with a mathematical description in which the more strongly a cell expresses a gene, the more strongly it inhibits its neighbors from expressing the same gene. Such *lateral inhibition* sets up a competition that can spontaneously generate "salt-and-pepper" or "checkerboard" patterns of expression across the whole field of cells [115]. Their model of lateral inhibition was based on a contact-based signaling scheme with known molecular components [116, 117], which establishes communication more locally than a diffusion-based signaling scheme such as the "short-range activation, long-range inhibition" invoked by Gierer and Meinhardt [11]. Subsequent research has generalized and elaborated upon lateral inhibition models in diverse contexts [118–124], often in conjunction with other patterning methods [125–131].

In addition to these more well-known schemes, specific study systems have also inspired unique but potentially generalizable models. For example, a "wave-pinning" mechanism

introduced to explain cell polarization (asymmetry) in response to a chemical signal [132,133] was subsequently extended to describe germ layer establishment in frog embryos [134]. As for many reaction-diffusion models, the crux of the wave-pinning model involves a disparity in diffusivity between two chemicals, and indeed theoretical work has subsequently unified more traditional Turing patterning with wave-pinning through a general analysis of mass-conserving chemical systems [21]. Studies have also examined alternatives to traditional morphogen gradient readouts for establishing gene expression boundaries, such as a recently proposed "speed regulation" strategy for segmentation in developing insects. According to this strategy, genes are always expressed in a particular sequence. Either high concentrations of or long exposures to a molecular factor are required to progress through the sequence, such that a gradient or a wavefront of the factor can generate experimentally observed stripes of gene expression [135, 136]. Researchers have also hypothesized a molecular implementation for speed regulation termed the "gradual enhancer switching model" [137].

In summary, the past seven decades of research into gene expression patterning during embryonic development have revealed—perhaps surprisingly—that a rich variety of outcomes may result from elaborations upon a few core ideas. Conceptual frameworks such as reaction-diffusion and positional information have inspired experimentalists to gather supporting evidence and theoreticians to develop explanations for evidence to the contrary, leading to a proliferation of models inspired by data and experiments inspired by models. Concurrently, researchers have begun to appreciate that the complexity and intricacy of living systems demands new approaches, including mathematical and quantitative analyses, to explain and predict the dynamics of patterning as well as its robustness, reliability, and reproducibility. In this regard, interdisciplinary work especially has the potential to inform our perspectives on patterning systems with an eye toward the underlying principles, which no doubt still hold surprises in store.

Further Reading

I have attempted here to cover in general terms the conceptual frameworks that have guided theoretical research into gene expression patterning in developmental biology, as appropriate to establish context for this thesis. As such, I have necessarily omitted much of the fascinating details (and particularly the experimental intricacies) that have facilitated theoretical advances. For related reviews in the literature, I recommend [4] (interplay of theory and experiment in pattern formation research), [138] (computational approaches to patterning in developmental biology), and [139] (self-organization in diverse experimental systems). For more focused overviews, I suggest [22] (recent developments in reaction-diffusion research) as well as [25] and [27] (morphogens and challenges related to their study). I will end by noting that new strands of research into gene expression patterning are increasingly considering the interaction between molecular concentrations and mechanical forces during embryonic development. Although these discoveries are exciting and relevant, they are also (sadly) beyond the scope of this thesis. I refer the interested reader to the latter half of [139] (as a survey) and to [140] (for more detailed treatment).

1.2 Multicellular Patterning in Synthetic Biology

The field of synthetic biology has yet to adopt a standard definition. For the purposes of this discussion, I will adapt a version from [141]: synthetic biology encompasses efforts to *create biological systems in order to establish control over cellular behaviors*. This framing excludes genetic engineering for measurement or perturbation of extant systems (e.g., "molecular recording" [142], applying fluorescent labels to endogenous components, adding optogenetic domains to existing proteins), as well as studies employing biomolecules outside the context of a living organism (e.g., [143,144]). A good deal of work has gone into developing synthetic microbes for technological purposes, including chemical production, medicine, and materials science. In contrast, most of the multicellular patterning systems to be discussed here have yet to find direct application. Tools from synthetic biology have also been used to reconstruct or recapitulate the function of known genetic circuits in model organisms, rather than to "control" the behavior of systems *per se*; this approach will be addressed in the next section during our discussion of "building to understand" natural systems.

Although the exact conception of synthetic biology is difficult to pinpoint, its birth might be recorded as January 20, 2000, when a pair of papers published in the same issue of the same journal reported the first synthetic biological circuits: a chemically inducible *toggle switch* comprising two mutually repressive transcription factors [145], and a network termed the "repressilator" comprising three transcriptional repressors with temporally oscillating concentrations [146]. A few months later, researchers published on a genetic network with negative feedback to reduce stochastic fluctuations in biochemical concentrations [147]. By the following year, genetic engineers had designed the first synthetic cell-to-cell communication [148]. Already in its first two years, then, the nascent field had demonstrated the potential to regulate three key ingredients for patterning: gene expression, robustness to stochasticity, and intercellular signaling.

The first pattern-forming circuit was of the Wolpert flavor, involving a chemical gradient that was interpreted by a microbial lawn into a stripe pattern. "Receiver" strains of bacteria expressed green fluorescent protein (GFP) for a specific range of concentrations of a diffusible chemical, which was produced by "sender" strains of bacteria seeded in localized regions. Diffusion from sender strains generated a chemical gradient, such that only receivers at a narrow range of distances from the sender expressed GFP. In this way, a green "ring" centered around a sender would emerge. The researchers called their implementation a "'band-detect' gene network"[1] [149]. A later version of the circuit allowed the position of the band to be tuned through externally administered signals in the growth medium of the bacteria [150]. Subsequent work has explored different implementations of the band-detect as a strategy for stripe formation [151–154].

Synthetic biologists have had more difficulty implementing "true" Turing patterning (with a diffusion-driven instability), for much the same reason developmental biologists have

[1]The terminology has since morphed in many places into "band-pass"—a most unfortunate development so far as anyone who has studied signal processing is concerned.

had difficulty identifying it in nature—namely, restrictive parameter ranges that biological parts do not tend to satisfy. Theoreticians have proposed alternatives, including a scheme for producing Turing patterns with a single diffuser [155,156], and have employed computational modeling to suggest practical requirements for achieving more standard reaction-diffusion patterning [157]. Experimental demonstrations have been shown in cell-free biological setups (e.g., [158]), but the only successful multicellular demonstrations thus far are of stochastic Turing patterning in engineered bacteria [159] and "solitary" reaction-diffusion patterning in mammalian cells [160]. The latter differs from conventional Turing patterning in that the "short-range activator" will only activate its own production once a certain threshold is reached. Thus, a perturbation to the homogeneous steady state that locally increases the activator concentration past the threshold will initiate patterning, but the faster-diffusing inhibitor will at some radius from the perturbation prevent enough activator from being produced to cross the threshold, thereby preventing propagation of the pattern past that radius [12]. As a result, patterns may strongly depend on initial conditions in a way they would not be expected to in Turing's original setup [160].

Concurrently, researchers have explored other patterning schemes based on models of naturally evolved systems. Genetic oscillations have been employed experimentally in conjunction with cell growth to generate ring patterns in microbial colonies [161], reminiscent of the clock-and-wavefront mechanism [162]. Lateral inhibition has been engineered into mammalian cells to produce mixed populations with a predictable ratio of cells expressing one gene vs. those expressing another, although the system did not produce salt-and-pepper patterns *per se* since the cells growing in a colony were not organized in a regular lattice [163]. Inspired by proposed mechanisms for morphogen gradient scaling in developing embryos, synthetic biologists engineered stripe-forming systems in circular bacterial colonies that scale the location [164] and width [165] of a ring of gene expression with the size of the colony, based on collective sensing of the size of the domain in which the colony is growing.

In addition to borrowing ideas from developmental patterning, synthetic biologists have also investigated schemes that are not as yet known to occur in "wild-type" organisms. Experimentalists have engineered bacteria to produce pigments in response to light, and have used colonies of these bacteria to identify outlines in black and white images (edge detection) [166] or reproduce RGB color images [167]. (Image processing schemes using reaction-diffusion of nucleic acids outside of living cells have also been proposed [144,168,169].) Theoreticians have uncovered dynamic patterning schemes in two-dimensional systems, including a "tug-of-war" mechanism that produces transient but long-lived Turing-like patterns [170], and have performed computational screens to identify two-node networks capable of producing spontaneous dynamic patterning in discrete cellular tissues [171].

Synthetic biology as a whole faces unique hurdles relative to conventional engineering fields, ranging from the practical difficulty of integrating and maintaining networks in living organisms to a lack of available methods to screen components or prototypes for desired functionality [172, 173]. Synthetic multicellular patterning additionally suffers from complications associated with cell-to-cell signaling, such as ensuring synthetic receptors respond to cues without activating native pathways, or that multiple signaling "channels" (including

engineered and endogenous) do not interfere with each other [174, 175]. On the theoretical end, challenges tend to arise from the high dimensionality of patterning systems. Genetic networks within individual cells are often already intricate and involve many components; coupling these networks with coordinated population-level behaviors can produce extraordinarily high-dimensional models, which may not be easy to analyze or interpret. How to reduce these models without losing important details—how to make them "as simple as possible but not simpler"—is an ongoing area of research and philosophical exploration within both the synthetic and developmental biological communities.

Further Reading

As before, I have grossly neglected examples of self-organization related to mechanical properties or physical arrangements of cells, including patterns arising from variable population densities [176–178] and cell-to-cell adhesion [179, 180]. These systems are reviewed alongside other synthetic multicellular patterning systems in [181] and [182]. For a history of synthetic biology in microbes, see [183]. For extensive reviews of practical challenges facing synthetic biologists, see [172] and [173].

1.3 Why Engineering?

> *"[E]nlightenment can come in different forms—not just in the elegant simplicity of a physicist's theory, but also in the more utilitarian guise of an engineer's analysis."*
>
> *– the (anonymous) editorial board of* Nature*, in [184]*

Denis Noble, one of the first scientists to apply computational models in physiology, has argued that researchers have yet to succeed in establishing a theoretical approach to the life sciences, having thus far relied rather on borrowed principles from chemistry and physics [185, 186]. His remarks motivate the mathematical aims of systems biology, or *the study of biological functions or pathways that consist of or comprise complex, integrated networks with dynamic interactions* (definition adapted from [187]). This somewhat nebulous field arose primarily in the 20th century and gained momentum into the early 2000s, concurrently with synthetic biology. Given the context, it is not entirely surprising that calls to apply engineering techniques to systems biology also emerged around this time [184, 188, 189].

I think it is relatively clear why engineering theory has strongly influenced synthetic biology, which, as an endeavor focused on "building things" for various applications, shares many of the same concerns as more traditional engineering disciplines (but see [190] as to whether it *should*). What requires more explaining is why an engineering perspective ought to inform our understanding of naturally evolved systems. This is of course a matter of some nuance and opinion, but at a philosophical level, engineering analysis is commonly justified for application to physiology and molecular biology on the basis of the following parallels (observed or proposed):

1. Living and engineered systems are distinguished from purely physical or chemical systems by their purpose, function [188], or performance objectives [191].

2. Living and engineered systems are modular, redundant, and hierarchical in structure [188, 189, 192, 193].

3. Both living and engineered systems "operat[e] in a world that is messy and unpredictable" [193], and therefore must regulate aspects of themselves or their environment so as to reject or adapt to disturbances [25, 189, 192, 194–196].

4. Both living and engineered systems are the result of optimization processes, the former through natural selection and the latter through deliberate design [197–199].

In considering these parallels, researchers must balance the desire to unify concepts between engineering and biology with the aims and details that are unique to each discipline. Put another way, researchers must recognize not just the analogies, but where they break down. For example, molecular signaling networks—in comparison to human-built automated systems—"do not have the luxury of employing specially-designed, dedicated components whose purpose is to sense or control biochemical signals" and thus "are comprised of elements that must serve both as the transmitted signals and their own controllers" [200]. The biochemical components employed by organisms impose constraints on system design that warrant careful consideration within—or in comparison to—existing frameworks [201] (e.g., [202–204]), especially with regard to intrinsic stochasticity (e.g., [205–207]).

Similarly, optimization is an entire field of research in engineering that is often formalized with respect to precise mathematical representations of the objectives, costs, and constraints relevant to a particular problem. For many biological systems, it is unclear what these objectives and costs would be, or how they ought to be expressed mathematically. There are also objections to this framing on philosophical grounds: Does optimization for one context necessarily entail suboptimality in other contexts? Do organisms even need to be optimal, or merely sufficient? Can sufficiency itself be encoded in, or emerge as a consequence of, objectives, costs, and constraints? Such challenges urge caution in applying principles and approaches established in one context to another context [25, 199, 201, 208–210]. At the same time, the unique characteristics of living matter open new avenues for theoretical, mathematical, and engineering research, by demonstrating principles of function and structure yet unimagined by human minds.

Theory-based engineering approaches to life sciences tend to fall into three camps. The first, which I will call "reverse engineering design", focuses on applying engineering techniques or analysis to uncover principles behind biological "designs"—meant here in the sense invoked by Daniel Dennett, as emergent structures resulting from environmental pressures, natural selection, and other "free-floating reasons", rather than from conscious intent [211]. The second, forward-engineering approach is sometimes referred to as "building to understand" [212,213], in the spirit of Richard Feynman's famous blackboard quip, "What I cannot

create, I do not understand." The central aim is to synthetically reconstitute existing systems, in whole or in part, to test current understanding of the mechanisms guiding their behavior. I would add to this approach the practice of implementing new or hypothesized mechanisms that have yet to be observed in order to evaluate their plausibility as biological "solutions". The third approach, which I tentatively term "behavior certification", is of a more practical bent. The goal is to develop tractable mathematical tests for whether a system is capable of exhibiting a particular behavior or class of behaviors. The test criteria that are necessary and/or sufficient to guarantee a behavior of interest are justified on analytical grounds, even if the test itself is administered computationally. Behavior certification may overlap with the other two approaches, particularly if the test criteria have physically meaningful interpretations.

Approach 1: Reverse Engineering Design

"[T]here is good reason for believing that there are a set of general and universal principles involved in the translation of genetic information into pattern and form."

– L. Wolpert, in [28]

For the purposes of this discussion, "design principles" can be thought of as *guidelines for selecting structural features that enable a system to fulfill its performance objective(s).* Defining a design principle necessitates (1) identifying the system's performance objectives and (2) identifying the structures responsible for achieving those objectives. Both steps require explicit consideration of the tradeoffs and constraints imposed by evolutionary history, biochemical and physical limitations, conditions of operation, and context (i.e., relation to other systems).

The hunt for design principles in systems biology has been predicated primarily on the notion of *modularity*, which maintains that living systems are organized into a series of interconnected *modules*, each of which carries out a function separate from that of other modules. Individual components (e.g., molecular species) may "belong to multiple modules at different times" and may be "quantitatively regulated, or switched between qualitatively different functions, by chemical signals from other modules" [188]. Many different combinations of molecules and their interactions may execute the same function, such that a given module can be implemented in myriad ways. The *implementation independence* of modules means a module-level analysis should apply to multiple similar systems operating in different contexts (e.g., in different organisms). Hence, with a modular representation, one ought to be able to consider "design principles" that apply to living systems in general, and not merely to specific examples.

Modularity received a critical vote of support with the reported discovery of "genetic network motifs" in bacteria [214]. A motif is a particular pattern of interactions among a small number of molecular species [215]. For example, stripes of gene expression can be produced by feedforward loops [149, 153, 216], in which one species activates or inhibits another species directly and also by way of an intermediary [217]. Interlaced cross-repressive

motifs ("toggle switches") are also well known for their role in positioning gene expression boundaries (e.g., [94, 218]). Similarly, a combination toggle switch/repressilator motif[2] was proposed based on observations of stripe formation in vertebrate neural tube development [219, 220] and was subsequently identified in fruit fly patterning as well [98]. Although genetic network motifs were originally envisioned to assume a specific function based solely on *topology* (the pattern of interactions) [221], the behavior of most motifs does in fact depend on parameter values to a significant degree [222]. Motifs that can perform the same function despite significant variation in parameter values are said to exhibit "parameter robustness" or to have "structural" behaviors [196, 223], a property that can also apply to modules more generally [224].

Whether a system is evaluated at the level of motifs or of other modules (e.g., [225]), it is the *interconnection* of these abstract elements that is responsible for producing full patterns of spatial and temporal gene expression observed in developing organisms [26, 226]. Engineers are familiar with the complications that can arise from hooking modules together: impedance matching, for example, is essential for designing electrical circuits in which each module in the full circuit behaves similarly to its behavior in isolation, implying that modules can be individually characterized before assembly of the whole device. Extending this idea to biochemical networks, Domitilla Del Vecchio and colleagues introduced the notion of "retroactivity" to capture the effect on a module of attaching its output to the input of another module. Accounting for retroactivity requires additional elements relative to models of more conventional engineered systems. Analysis based on retroactivity has identified certain biological mechanisms that enhance the "separation of function" among modules [227], as well as examples where genetic network motifs exhibit counterintuitive behaviors due to contextual factors [228]. Although this approach has yet to be widely applied outside synthetic biology, it serves as a welcome example of how certain high-level phenomena—in this case, changes in module behavior in isolation vs. in context—may be shared by both living and engineered systems, but nevertheless require mathematical formalisms tailored to the specific realities of each field.

Studies of genetic network motifs and interconnections thereof typically rely on *dynamical systems models*, which describe systems that change in time. These models originated in physics and have generated substantial interest among mathematicians; however, many practical methods for analyzing them were pioneered by engineers beginning in the late 19th century, who were concerned with applications related to electrical power, communications, manufacturing, and military technology [229–232]. In particular, the field of *control theory* has contributed a wealth of analysis and design techniques for regulating and controlling the behavior of automated systems. The notion of "control" implicitly frames the problem in terms of a performance objective (i.e., what manipulations or means can be employed to achieve predetermined ends), which distinguishes it from traditional approaches in the hard sciences. Control-theoretic representations of systems are modular by nature, with the "controller" characterized separately from the "plant" that it controls.

[2]The authors dubbed this motif the "AC-DC circuit" [219], confoundingly to this engineer.

The application of control theory to biological networks has resulted in some of the most successful attempts to explore mathematical principles behind evolutionary designs. A number of biological processes involve *homeostasis*, or maintaining certain quantities within certain ranges despite environmental perturbations; for example, a human's internal body temperature generally should not depart too drastically from ca. 37 C, regardless of the weather outside. Similarly, in a process sometimes referred to as "desensitization", many systems will respond to a stimulus initially but return to basal activity levels after a short time, thereby ensuring sensitivity to future stimuli. Both homeostasis and desensitization are examples of *adaptative* regulatory processes, many of which (e.g., bacterial chemotaxis [233–235], ion channel responsiveness [236]) have been analyzed from the standpoint of *feedback control systems* [237]. Feedback implies that measurements of a system's output modify the system's behavior so that the output achieves some desired response—in the case of adaptation, settling to basal levels after a transient [238]. In a sensory system, the transient is necessary because it encodes the response to a stimulus, while in a homeostatic system, the transient is rather a necessary evil, because it represents a departure from the constant *set point*. A given system may employ multiple "nested" feedbacks in serial or in parallel; modular decomposition of the heat shock pathway in bacteria has suggested functional roles for such feedbacks with respect to the performance objective of adaptation [239].

Glycolysis is an example of a biological process whose observed behavior is puzzling at first sight, but has a meaningful interpretation from an engineering perspective. This metabolic pathway transfers energy stored in glucose into energy stored in molecules of adenosine triphosphate (ATP). Although an organism would seem to benefit from maintaining a constant pool of ATP even as its energy demands vary, empirical observations in yeast [240, 241] and skeletal muscle cells [242] have shown that the concentrations of metabolites oscillate in time. A study of the glycolysis network from a control-theoretic perspective concluded that oscillations might arise not for any functional reason *per se*, but as a byproduct of constraints imposed by the simultaneous need for efficiency (low metabolic cost of manufacturing necessary network components) and low steady-state error (difference between actual and desired levels of ATP for a given energetic demand). Similar performance constraints were found to apply even when the purported molecular mechanisms in the model were replaced with an arbitrarily complex (linear) stabilizing controller [243].

This study well illustrates two of the advantages of engineering approaches. First, framing a problem with an eye toward design can make explicit the tradeoffs that influence system performance—in this case, between robustness (oscillations, steady-state error) and efficiency. Second, modular deconstruction—such as decomposition of a full system into controller and plant—facilitates analysis that is independent (to a degree) of molecular details, and can therefore allow generalizations beyond a particular model to a broader class of systems. In the aforementioned work on glycolysis, the ability to generalize was used to highlight fundamental constraints and tradeoffs (including in later work extending to autocatalytic networks beyond glycolysis [244]). More broadly, analytic techniques may also clarify the necessary and sufficient conditions for carrying out a particular aim, and therefore aid in the search for, or characterization of, systems that achieve that aim. For example,

robust perfect adaptation[3] of a biochemical system is equivalent to the control-theoretic requirement for robust steady-state tracking and disturbance rejection in deterministic linear systems, for which integral controllers are necessary and sufficient. Expanding this concept into the biological sphere, theoretical work has revealed an antithetic motif that attains integral control of stochastic unimolecular and bimolecular chemical reaction networks (satisfying reasonable constraints) [207]. Furthermore, it has been shown that any biomolecular controller achieving robust perfect adaptation for arbitrary stochastic reaction networks *must* include this motif [246]. This observation may be considered as a kind of "design principle" directly analogous to a concept in conventional engineered systems.

Control systems approaches are slowly finding their way into developmental biology as well. Theoretical research into how morphogen gradients scale with tissue size suggested an "expansion-repression" mechanism analogous to an integral controller [69], and a combination of modeling and experiments in fruit flies identified a control architecture with redundant negative feedback loops that are responsible for reducing developmental errors. Interestingly, the redundancy only appears to be necessary when cell metabolism is high and therefore development is fast; at slower metabolic rates, additional feedback loops have little effect on the dynamics of gene expression [247]. It remains to be seen whether the architecture of multiple feedback loops that are redundant in certain contexts and not in others constitutes a "design principle" that applies to other biological patterning systems, especially given that similar cascades of serially connected modules, with the output of the whole cascade fed back into each of the modules, have also been identified in intracellular signaling and metabolic pathways (e.g., [239, 248]) and are commonly employed in process engineering [237].

While the previous examples have considered control of molecular concentrations, Michael Levin has argued for further application of control theory to higher-level outcomes. He notes that salamanders may regenerate entire limbs with proper structure even from tails grafted onto the flanks of the animals, and suggests that this process may be interpreted as a homeostatic feedback system "work[ing] to reduce the error between the current morphogenetic state and an anatomical setpoint [...] that is a macrostate compatible with several molecular mechanisms that can implement it" [249]. Such a "top-down view" of biological systems is well supported by existing techniques in control theory [250], provided that states and set points can be defined in a meaningful way. More generally, the arguments of Levin and colleagues suggest that the design principles of patterning systems may be better understood at levels above genetic network motifs or other modules operating on similarly small scales.

Dynamical systems may be the most common representations invoked to study genetic networks mathematically, but other classes of models from engineering fields have also been employed to study biological pattern formation. For example, a Boolean network model—in which interacting state variables take on binary values—was invoked to conclude that the

[3] *Perfect* adaptation means that the system, after perturbation, returns to the set point without any steady-state error. *Robust* adaptation means adaptation to perturbing inputs occurs even despite (reasonable) perturbations to system parameters [245].

pattern of segment polarity genes in fruit fly embryos[4] results primarily from the topology of the network rather than exact kinetic rate parameters [251]. In this case, simplifying the model by treating genes as binary on/off switches facilitated mathematical analyses that would have been less tractable with other representations [187]. The results of this study support the hypothesized principle that naturally evolved systems are robust to parameter variation because they encode behaviors primarily into the structure of the underlying biochemical network.

Even more abstract representations of biological systems have emerged from applications of information theory, a branch of engineering that deals with communication over noisy channels. Techniques based on Claude Shannon's notion of information (as a reduction in uncertainty once an event is observed) have been employed to study transcriptional regulation [252] and thereafter to provide a literal interpretation of positional information [81,253]. Experimental studies in vertebrate neural tube [75] and early fruit fly embryos [254] indicate that genetic networks do appear to optimally decode a cell's relative position from the concentrations of underlying morphogen gradients or pre-established expression patterns of upstream genes. Taking optimal decoding as a performance objective for networks that implement sequential patterning, one might then ask what network features or design schemes facilitate optimal decoding. Unfortunately, the methodology employed in these studies cannot address this question[5], leading biologists Johannes Jaeger and Berta Verd to argue for a synergystic approach between information theory and dynamical systems in order to understand positional information [255].

In summary, researchers have exploited parallels between engineered and evolved systems to uncover common schemes underlying observed behaviors in diverse biological contexts. Motivated by a notion of living systems as modular entities, engineers have developed and adapted existing theories for control and information processing to describe the properties of genetic networks in terms of their ostensible function. These approaches have deepened our understanding of why genetic networks may exhibit nondesirable behaviors, or what properties may be required for a series of biochemical reactions to acheive a defined aim. Despite its intuitive appeal, an approach to biological design principles based on modularity still requires experimental and philosophical justification to address fundamental criticisms, including the apparent inability for most modules to be entirely separated from their context in a full system, the difficulty in defining a performance objective [256], and the underlying assumption that commonly observed modules were necessarily selected for functionality (the implication being that randomly generated networks could produce such modules with equal

[4]Propriety demands I mention that a sufficiently early fruit fly embryo is actually a *synctytium*, or a single cell with multiple nuclei. In the spirit of mathematical equivalence, theoretical simplification, and/or engineering practicality, I will continue to use it in the context of discussion about multicellular systems with the understanding that patterning in the synctytium takes place among nuclei rather than cells, and that any signaling must be diffusion- and not contact-based.

[5]As theoretical physicist Boris Shraiman has remarked, "The problem with [Shannon] information is that it's just one number and there's not much information in it" (overheard at the 2019 UCSB/KITP summer course on Morphogenesis: Form, Fate, and Function).

frequency) [257]. Fortunately, mathematical models are hypotheses whose predictions can be (and are being) tested experimentally. An explosion in methods and technologies, particularly in synthetic biology, is opening a new frontier for empirical verification of mathematical theories, founded on the notion that our understanding of a system should be sufficient for us to reconstitute it (as far as we can) from scratch.

Approach 2: Building to Understand

"The properties of a system can be explained by the properties of its components. They cannot be deduced from them."

– F. Jacob, in [208] (a lecture delivered at UC Berkeley in 1977)

In his excellent review entitled "Using synthetic biology to explore principles of development", Jamie Davies writes that "if a complex system is believed to achieve its action according to a simple principle, then constructing a new system based on that principle and assessing whether it performs the required action provides a powerful verification" [258]. This idea has motivated synthetic biology since that same fateful year of its "birth" in 2000, by which time researchers had already used a synthetic gene network to experimentally verify the theory that negative autoregulatory feedback loops in a genetic circuit might reduce the range over which chemical concentrations fluctuate [147]. Similarly, many of the examples of synthetic patterning systems presented earlier demonstrate "proof of principle" simply by virtue of their successful functioning. Some of them—constructed without the intent to exactly mimic a known genetic network or validate an existing theory—have even suggested new possible mechanisms for naturally observed phenomena. Rather than retread the territory discussed in Section 1.2, the following discussion will focus on two examples of "building to understand" gene expression patterning in developmental biology that seem conducive to further consideration with engineering approaches.

The pathways that sense morphogen gradients can be implemented in multiple ways [89], but why a particular system might employ a particular architecture is not fully understood. In 2018, Pulin Li and colleagues explored the design principles of morphogen gradient formation and signaling using a synthetic setup that allowed them to build and test various network structures in mammalian cells grown on a plate. Sender cells seeded in a point or column produced morphogen, which diffused through a field of receiver cells that could sequester (uptake) the morphogen by way of a ligand. The receiver cells were engineered with one of three network structures based on a naturally occurring pathway, and signaling levels (response to morphogen) were visualized. In an open-loop system in which the amount of sensed morphogen did *not* feed back on the sequestration rate, the spatial pattern of signaling was more sensitive to variation in morphogen production rate than the system with feedback. Other modifications to the architecture elucidated forms of feedback that could buffer changes in gradient length scale or amplitude to changes in morphogen or ligand production rate, backed by mathematical modeling [259].

Li and colleagues' observations have the potential to expand beyond the exact models that were studied, in a manner particularly ripe for further investigation with tools from control theory. Is there a more abstract framework in which to consider the role of the multiple feedback loops in gradient formation and maintenance, along the lines of the controls analysis performed on the heat shock system [239]? Could this framework account for the robustness of the various architectures to variation in different parameters, or for the behavior of other known morphogenetic systems in which positive rather than negative feedback loops are employed? Could theory suggest alternative architectures or implementations to achieve the results observed by the authors? Could these alternatives be constructed to validate the generality of the framework? These and related questions illustrate (1) how engineering abstractions could help to distinguish characteristics common to systems more broadly from characteristics specific to a particular experimental implementation; and (2) how the design principles identified through reverse engineering might be evaluated through systematic forward engineering, ideally across multiple model systems or implementations.

One of the benefits of using synthetic circuits to test design principles is that they can operate outside the context of a complex surrounding genetic network, and are therefore more amenable to characterization in isolation than are endogenous modules. For example, to explore how regulation of a single gene by multiple proteins might influence downstream patterning, Justin Crocker and colleagues engineered two synthetic transcription factors into a fruit fly embryo: an activator expressed in a gradient along the length, and a repressor expressed in a stripe through the center. A reporter was placed under the control of both transcription factors, causing it to be expressed in two stripes on either side of the anterior half of the embryo. The authors found that the number and arrangement of binding sites for the transcription factors modulated the sharpness of the stripes, with overlapping binding sites creating more well-defined boundaries, perhaps due to competition between the transcription factors at the enhancer (where the binding sites are located). This observation may help explain the prevalence of overlapping binding sites in developmental enhancers [260]. A similar system in its native context would have been more challenging to investigate in this manner, since the researchers would not have had as much control over the layout of binding sites or over the upstream and downstream elements possibly interacting with the gene of interest. This study also highlights how "microscale" properties such as the layout of genes and binding sites on DNA can affect macroscale patterning behaviors—a trait that control-theoretic engineering approaches, with their focus on network topologies and implementation-independent modules, have yet to account for in a systematic way.

The idea of "building to understand" pattern formation has gained traction within a larger context of growing interest in *synthetic (tissue) development* or *synthetic morphogenesis*, which aims to construct and control the processes by which single cells or homogeneous clusters of cells self-organize into complex, specialized structures [261–268]. In my experience the members of this nascent community are eager for multidisciplinary perspectives to tackle the exciting challenges that lie ahead [213, 269, 270]. Theoretical approaches from engineering have particular potential in this field, for engineering demands not simply a model of a system, but a model intuitive enough that a human can use it to design some-

thing new. Whether such designs would (or should) recapitulate natural ones remains to be seen, but the very process of creating them ought to shed light upon the "allowances and constraints" [141] of working with biological material, and the plausible principles that arise therefrom.

Approach 3: Behavior Certification

"Now is the time to incorporate computational modeling [...] and not merely to manipulate language as done here."

– A. Trewavas, in [271]

To understand the context for "behavior certification"-type approaches to biological systems, I think it is necessary to first take a step back and consider the interplay between analytical theory and computation. Thus far, I have tended to use "theory" (in the mathematical sense) to apply to *abstract representations with explanatory power*. This usage distinguishes theory from modeling, which need not be abstract (generalizable) nor have explanatory power (which derives primarily from the interpretability of the representation). Models in biology are most often studied in simulation through the use of computational approaches, that is, numerical methods that produce results or "conclusions", rather than through analytical means, since many such models do not have closed-form solutions.[6]

I have sometimes found that biologists question the value of mathematical theory, which seems to require vast oversimplifications that render it useless for modeling a particular system with any accuracy. In this view, theory is something like the austere and recalcitrant older sibling of affable, talkative simulation, which—with the appropriate guidance—can approximately solve some of the most repugnant of equations in milliseconds. In the spirit of bridging gaps, I submit that theory and computation in fact work closely to account for each other's deficits, and it is really the combination of approaches that most successfully imparts insight.

I will attempt to illustrate with an example. In 2009, researchers carried out a computational screen to identify network topologies capable of near-perfect adaptation, and found that only two "core topologies"—negative feedback and incoherent feed-forward—achieved this goal across a large range of kinetic parameter values [272]. Previous work had tended to focus on the possible behaviors of known motifs rather than on the possible motifs that could demonstrate known behaviors, thus this study represented a welcome new approach to modularity in biochemical networks [238, 273]. However, due to computational constraints, the authors restricted their study to systems with three species, each of which exists in either an active or inactive state, and they were only able to scan about 4% of parameter space. These limitations meant that their analysis failed to identify two more motifs capable of adaptation:

[6]It is worth reiterating that mathematical is not the same as quantitative. Quantitative ("numerical") measurements can be made without any underlying model to explain where they came from, and similarly mathematical theory can be conducted symbolically with only the barest of constraints upon the involved numbers.

a state-dependent inactivation mechanism (in which nodes can be "off" in addition to being active or inactive) [236] and the antithetic motif [207,238]. Even with greater computational power, research conducted with similar methods could not feasibly conclude that *any* system achieving adaptation is *required* to contain at least one of these motifs, in the way that the antithetic motif was found to be required for robust perfect adaptation *for any* stochastic reaction network. Thus, in this situation, analytical theory was able to address a class of questions impractical (or impossible) to answer with computational screens.

Nevertheless, in many cases computational techniques can be more accurate and more practical than "equivalent" analytical proofs. Biological systems are constrained to operate with physical parts, usually within certain environments, and can also tolerate occasional anomalous failures. Hence, metrics "empirically" assessing simulation results under ordinary circumstances may be adequate to identify "sufficient" or approximate solutions to satisfy a performance objective—and substantially easier to pursue than corresponding analytical avenues (which may be impossible). Furthermore, it is possible that many of the nuances in system performance arise from complexities that would not anyway be included in a simplified but mathematically tractable model. Of course, the appropriate choice of model might also confer benefits in the computational sphere; for example, a *hybrid systems* model incorporating both continuous and discrete dynamics was used to model lateral inhibition, and found to be as accurate as a full nonlinear model, but more interpretable and less computationally expensive [274]. (That accuracy is maintained with respect to the behavior of interest is an essential ingredient of a useful model, and is often best tested through collaboration with experimentalists.) Overall, these tradeoffs suggest a synergy between analytical theory and computation, with the former primarily concerned with formal justification and generalization, and the latter primarily focused on evaluating utility or plausibility in more complex or specific settings.

This synergy is exemplified by the *behavior certification* approach, in which theory serves not necessarily to explain the behavior of a system, but to justify in rigorous fashion that certain outcomes are inevitable or impossible if that system (model) meets a given set of requirements. In this way, models that evade exact analytical solutions may still be found numerically to satisfy criteria that are analytically guaranteed to result in or preclude particular classes of behaviors. These criteria are loosely analogous to the "necessary" or "sufficient" conditions that biologists traditionally discuss (though see [275] for arguments that researchers should move beyond this classical paradigm), with the difference being that such conditions need not admit a physical interpretation or provide intuitive insight into system behavior (though ideally they will). Control theory has developed a number of such methods for certification based on technological demands for controllability, reliability, safety and the like, some of which are applicable to biological systems.

Identifying mathematical conditions to guarantee the (non)existence of oscillatory solutions or the (in)stability of particular steady states in biochemical networks has a history at least as long as that of positional information. One of the earliest examples arises from mathematical treatments of enzymatic pathways, specifically a chain of reactions in which the end product modulates the rate of the first reaction [276, 277]. For pathways in which

the feedback reaction is inhibitory and all others are "facilitatory", initial approaches based on linearization identified a condition on the collective decay rates and interaction strengths sufficient to guarantee stability of the steady state [278–280]. Subsequently, researchers in control theory harnessed methods for analyzing networks of input-output systems to generalize the condition [281] to fully nonlinear models [282] and to interconnection structures other than the negative feedback loop [283].

Similar control-theoretic methods have been applied to other intracellular metabolic and genetic networks to probe the existence and stability of steady-state solutions (partially reviewed in [284]). Fewer such approaches have been applied specifically to networks capable of patterning. A handful of theoretical works have examined the stability of spatially uniform (homogeneous) solutions in reaction-diffusion settings [285–288], and a test for contrasting patterning through lateral inhibition [122,289] will form the basis for Chapters 3 and 4 in this thesis. A few extant procedures for verifying synchrony among coupled oscillators (e.g., [290–293]) might eventually be applied to clock-and-wavefront or other proposed mechanisms for translating temporal oscillations into spatial patterns, though to the best of my knowledge this work has yet to be done.

Despite the existence of many behavior certification methods for biochemical networks, they do not appear to have found wide use in the experimental community, especially outside synthetic biology. Many factors might contribute to the gap, including lack of awareness, a difference in timescale between theory and experiment, or a mismatch in concerns between fields (existence and uniqueness are not the customary purview of biologists). In my opinion, however, the most significant barrier is accessibility. Leaving aside the fact that most of these tests are published in engineering journals that biologists are unlikely to peruse, traditional education in life sciences does not equip biologists with the vocabulary or the training to make sense of control theory papers. With a few exceptions (e.g., [294]) those works framed for biological audiences tend to focus on synthetic biology, which has a relatively larger cadre of researchers trained as engineers than do other branches of life sciences, and therefore requires only a partial "repackaging" of the work for accessibility. That theoretical methods are not more widely adopted cannot be attributed solely to lack of effort; until education in biology incorporates more mathematical background, researchers reframing mathematical work in more intuitive terms will continue to run the risk of losing nuance in translation, which could exacerbate issues—perceived or actual—with the validity of theory. In the meantime, I believe both engineering and biology would benefit from more concerted efforts to communicate across the divide and unlock the full potential for interdisciplinary exploration and discovery.

Further Reading

My thoughts on the interface between engineering and biology have been shaped by a number of opinions with varying levels of philosophical and technical concern. For approachable yet insightful reads, I recommend Yuri Lazebnik's humorous "Can a biologist fix a radio?" (which argues convincingly for a more formal, quantitative language in sys-

tems biology) [295], Michael Elowitz and Wendell Lim's "Build life to understand it" [296], and Uri Alon's "Biological networks: the tinkerer as an engineer" [189] (a response in part to François Jacob's largely philosophical but fascinating lecture on "Evolution and tinkering" [208]). Arthur D. Lander has argued clearly and compellingly for considering performance objectives when analyzing living systems [191], while Andreas Wagner and William Rosen have proposed a philosophical framework for uniting "technological and biological innovation" [198] (see also [199] for an example of where these varieties of "evolution" might diverge). Further examples of control theory in genetic networks can be found in [201], with a more detailed discussion geared toward synthetic biology available in [297]. For those intrigued by synthetic morphogenesis and related topics, I recommend [258] and [266].

I have tried to focus my discussion on the principles behind gene expression patterning, with excursions into solely intracellular networks only insomuch as they are necessary to understand the context for, or have come to contribute to, patterning analysis. As such, I have excluded a great deal of related work without explicit relation to engineering theory, as well as a few engineering-inspired tools (e.g., sensitivity analysis in biochemical networks [298], noise filtering during transcription [299]) without immediate application to pattern-forming systems (though a creative mind would doubtless identify connections). Finally, the reader will forgive me for eschewing any discussion of network identification or parameter fitting, which has received extensive attention from electrical engineers and computer scientists, but bears little relation to the content of this thesis.

1.4 Preliminaries

"This model will be a simplification and an idealization, and consequently a falsification."

– A. M. Turing, in [2]

As with any work predicated on modeling, there are multiple ways of representing the same physical system mathematically. Particularly relevant to multicellular pattern formation is the distinction between *continuous* models, in which chemicals are produced at every point in a domain in continuous space, and *discrete* models, in which cells are considered as separate entities and therefore chemicals are produced at a countable number of points in space. Examples of continuous models include most reaction-diffusion networks (although Turing himself in his seminal paper started by analyzing a discrete ring of cells [2]), most clock-and-wavefront models, and many discussions of morphogen gradients and their subsequent interpretation. Discrete models are more commonly evoked for systems with (very) small numbers of cells or heterogeneous populations of cells, many of which are primarily examined in simulation. Systems in which cell-to-cell signaling requires cells to to be in physical contact with their neighbors also tend to be treated with discrete representations. While some behaviors can qualitatively be exhibited by both continuous and discrete models, there are often underyling subtleties between the two. For example, traveling wavefronts emerging from bistable dynamics can propagate in continuous media with any (nonzero)

amount of diffusion, but will stall in discrete media unless a diffusion (coupling) threshold is reached [300, 301] (see also [302] for another form of "propagation failure" in discrete lattices).

The work in this thesis is built off a discrete representation of cells as *networked dynamical systems*. In this view, each individual cell is an *input-output system* that can receive or read inputs from the environment or from other cells, and also produces one or more outputs that can act as inputs to other cells. Each cell has certain properties that evolve in time and are influenced by the inputs the cell receives. In this thesis, cellular properties will generally be chemical concentrations with dynamics described by sets of ordinary differential equations, or ODEs. Connecting the outputs of some cells to the inputs of others produces a *network* of dynamical systems. The fact that the network is explicitly considered—that is, represented separately from the behaviors of the cells—enables certain mathematical tools to be applied in ways that would be difficult to manage with other representations of multicellular systems.

In the following subsections I establish conventions for mathematical notation as well as introduce the basic models of gene expression that are employed throughout the remainder of the thesis.

Network Structure and Cell-to-Cell Signaling

In this thesis, discrete models of multicellular systems will be represented as *graphs*, which are mathematical structures consisting of *nodes* connected by *edges*. Nodes will correspond to cells and edges to an unspecified means of signaling between them (e.g., diffusion, contact). Specifically, we will consider *undirected graphs*, in which signaling is bidirectional; that is, if one cell is capable of signaling a second cell, the second cell is also capable of signaling the first.

Suppose we have a graph $\mathcal{G}$ with N nodes indexed 1 through N. The *adjacency matrix* of $\mathcal{G}$ is an $N \times N$ matrix representing the edges between nodes, with the (i, j)th entry equal to 1 if nodes i and j are connected and 0 otherwise. We will primarily work with conceptual variants of the adjacency matrix that are *weighted* (the nonzero entries are not necessarily equal to 1) and may include negative diagonal terms to quantify the "cost" to a cell of signaling its neighbors. These matrices explicitly model network structure (which cells signal to which) and may also capture aspects of cell-to-cell signaling. For example, in Chapter 2, the interconnectivity matrix M indicates whether cells communicate through diffusion- or contact-based mechanisms.

Input-Output Dynamical Systems

Cells are modeled as input-output dynamical systems. Let i index the cells 1 through N. I will generally consider models where the dynamics of the ith cell are given by

$$\begin{cases} \frac{d}{dt}x_i(t) = f_i(x_i(t), u_i(t)) \\ y_i(t) = g_i(x_i(t)) \end{cases} \tag{1.1}$$

where $x_i(t)$ is the vector of *states* (chemical concentrations) in cell i, $u_i(t)$ is an input (usually scalar), and $y_i(t)$ is an output (also usually scalar). Often $g_i(x_i(t))$ will simply select a single entry of $x_i(t)$ to act as a readout. Frequently, all entries of $x_i(t)$ will represent chemical concentrations and will therefore be nonnegative at all times t; similarly for $u_i(t)$ and $y_i(t)$. I will customarily use the dot notation to indicate time derivatives, i.e., $\dot{x}(t) = \frac{d}{dt}x(t)$. The vertical concatenations

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_N(t) \end{bmatrix}, \; u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_N(t) \end{bmatrix}, \; y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_N(t) \end{bmatrix} \tag{1.2}$$

will contain all of the states, inputs, and outputs respectively in the full system of N cells. Thus, a full networked system of interacting cells will be given by (1.1) in conjunction with

$$u = My \tag{1.3}$$

for some matrix M of appropriate dimension that captures the structure of the network (which cells signal to which). This matrix is more specifically defined by chapter for specific application.

Gene Expression

One of the great challenges to studying biological systems from a mathematical standpoint is their nonlinearities. These are present by necessity, because cells and organisms are of limited size and duration, and cannot contain or produce negative or infinite quantities of chemicals. Practically, this means that $f_i(x_i(t), u_i(t))$ as given in (1.1) will be a nonlinear function of $x_i(t)$ and $u_i(t)$, though it may be linearized in the course of analysis. (Note $g_i(x_i(t))$ need not be nonlinear, as it may simply serve to identify which of the states $x_i(t)$ is the output.)

A simple model of gene expression is given by

$$\begin{cases} \dot{m}(t) = \alpha_m h(u(t)) - \gamma_m m(t) \\ \dot{p}(t) = \alpha_p m(t) - \gamma_p p(t) \end{cases} \tag{1.4}$$

where $m(t) \in \mathbb{R}_+$ is the concentration of mRNA, $p(t) \in \mathbb{R}_+$ is the concentration of protein, α_m, α_p are transcription or translation rates respectively, and γ_m, γ_p are degradation (decay) rates. (All four of these constants are scalar and nonnegative.) At steady state for a constant-in-time input $u(t) = u^* \in \mathbb{R}_+$ the protein concentration is a linear multiple of steady-state mRNA concentration. For this reason, models that are not concerned with transient system responses will occasionally lump the dynamics of mRNA and protein.

In biological modeling, $h(\cdot)$ in (1.4) is typically a sigmoidal function called a Hill function, after Archibald V. Hill, who postulated their form in his studies of hæmoglobin dissociation

curves in the presence of aggregates [303]. Typically $u(t)$ is a transcription factor that binds the DNA near the gene so as to influence the transcription rate. For activation (the input $u(t)$ increases gene expression),

$$h(u) = \frac{\left(\frac{u}{K}\right)^n}{1 + \left(\frac{u}{K}\right)^n}, \tag{1.5}$$

while for inhibition ($u(t)$ decreases transcription),

$$h(u) = \frac{1}{1 + \left(\frac{u}{K}\right)^n}. \tag{1.6}$$

If multiple transcription factors regulate the expression of a gene, then depending upon the application one may multiply or add many Hill functions together, or adopt alternative forms such as

$$h(u_1, u_2) = \frac{a_1 u_1^{n_1} + a_2 u_2^{n_2}}{1 + b_1 u_1^{n_1} + b_2 u_2^{n_2} + b_3 u_1^{n_1} u_2^{n_2}}. \tag{1.7}$$

The activating Hill function (1.5) can be derived from a chemical model of a system in which molecular *receptors* are bound by other molecules called *ligands*. If there are two reactions in the system—n ligands bind or unbind a receptor simultaneously—then in a population of many molecules the fraction of bound to unbound receptors will be given by (1.5), where K is the ratio of unbinding to binding rate. In practice this model is not often physically realistic [304], and $h(\cdot)$ is rather used as an ansatz, with the "dissocation constant" K and Hill coefficient or "cooperativity" n obtained by fitting to data.[7] Indeed, many researchers (myself not excluded) find it convenient to use Hill functions for any sigmoidal response curve, regardless of the (lack of) physical interpretation.

Finally, I will stress that there are many more ways for gene expression to be regulated than simple binding of a transcription factor near a gene of interest, and many more ways to model such processes than through deterministic sigmoidal functions. These fascinating details have filled many a thesis, and are unfortunately out of scope for the introdution to this one.

1.5 Thesis Organization and Contributions

This thesis is organized into three main chapters. The first chapter introduces a general theoretical framework for examining prepattern processing (or pattern refinement), with analyses of examples in developmental and synthetic patterning. The second and third chapters each focus on one experimental implementation of spontaneous contrasting patterning in synthetic biology. Although both experimental systems rely on the same behavior

[7] Fittingly, Hill himself admitted to having guessed the form of his function, because otherwise "the calculation of the constants [...] is very tedious". His "object was rather to see whether an equation *of this type* can satisfy all the observations, than to base any direct physical meaning on n and K" [303].

certification-type theory, its realization is quite distinct between the two, based on the affordances and constraints of the respective setups. The design principle of bistability for large-scale patterning, however, remains intact. Where relevant, I have emphasized the process of translating theory into practice, as well as areas where empirical observations have inspired theoretical directions. A brief conclusion with outlook is also provided.

Chapter 2: Spatial Filtering Approach

We propose a discrete filtering approach to predict how networks of interacting cells modulate spatially varying input signals to produce more complicated or precise output signals. The interconnections between cells determine the set of spatial modes that are amplified or suppressed based on the coupling and internal dynamics of each cell, analogously to the way a traditional digital filter modifies the frequency components of a discrete signal. We apply the framework to two systems in developmental biology and to one synthetic biological system that acts as an edge detector. In the course of our analysis, we identify experiments to differentiate between similar models and find that Turing-like patterns may occur even in the absence of instabilities. Results also indicate patterning systems may be inherently robust to both correlated and uncorrelated noise sources. Our work shows that a spatial frequency-based interpretation simplifies the process of predicting patterning in living organisms when both environmental influences and intercellular interactions are present, and may facilitate designs for synthetic multicellular systems that execute image processing tasks *in vivo*.

Most of the material in this chapter first appeared in [305] and [306], co-authored with Murat Arcak.

Chapter 3: Contrasting Patterning through Diffusion-mediated Lateral Inhibition

Pattern formation and specific cell-to-cell interactions are important for microbial consortia to divide labor and perform complex functions. To obtain further insight into such interactions, we experimentally demonstrate control over gene expression patterning in a diffusion-mediated lateral inhibition circuit whose behavior is well predicted by theory. We highlight the importance of spatial arrangement as a control knob for modulating system behavior. Our systematic approach provides a foundation for future applications that require understanding and engineering of multistrain microbial communities for sophisticated, synergistic functions.

This work was published in [307] with co-first author Mika Tei, who conducted the experiments, and co-authors Justin Hsia, Murat Arcak, and Adam Arkin.

Chapter 4: Cell-in-the-loop Patterning with Optogenetically Emulated Cell-to-Cell Signaling

Here, we introduce a "cell-in-the-loop" approach where living cells interact through *in silico* signaling, establishing a new testbed to interrogate theoretical principles when internal cell dynamics are incorporated rather than modeled. We present a theory that offers an easy-to-use test to predict the emergence of contrasting patterns in gene expression among laterally inhibiting cells. Guided by the theory, we experimentally demonstrate spontaneous checkerboard patterning in an optogenetic setup where cell-to-cell signaling was emulated with light inputs calculated *in silico* from real-time gene expression measurements. The scheme successfully produces spontaneous, persistent checkerboard patterns for systems of sixteen patches, in quantitative agreement with theoretical predictions. Our research highlights how tools from dynamical systems theory may inform our understanding of patterning, and illustrates the potential of cell-in-the-loop for engineering synthetic multicellular systems.

This work, co-authored with Dirk Benzinger, Murat Arcak, and Mustafa Khammash, appears in [308].

Chapter 2

Spatial Filtering Approach

"Most of an organism, most of the time, is developing from one pattern into another, rather than from homogeneity into a pattern."

– A. M. Turing, in [2]

The main contribution of this work is a filtering perspective for analyzing prepattern processing in multicellular biological systems. While research into spontaneous patterning generally involves searching for conditions that destabilize a homogeneous solution, this framework considers patterning when that homogeneous solution is stable. The prepattern is then treated as a spatial perturbation about this point. A central component of the approach is spatial mode decomposition, a common tool in distributed systems analysis (e.g., [309]) that has previously been applied to detect instabilities in cellular networks lacking external inputs [310]. The spatial modes are determined from the network structure and the mode of communication among cells. The genetic network in each cell acts to attenuate or amplify ("filter") the spatial modes relative to their weights in the prepattern, in a manner analogous to image filtering in traditional signal processing. The primary benefit of this framework for biological analysis is that it decouples the activity of the genetic network from the exact shape of the prepattern, enabling one to build intuition without reliance on computational simulations of the entire system. Examples from developmental and synthetic biology highlight the insights gained from use of the framework.

2.1 Introduction

In multicellular patterning systems, prepattern processing describes the process by which an extant pattern drives a downstream genetic network [26] to output a different pattern [2]. Prepatterns may arise directly from environmental influences that differ by cell, from internal processes such as those that establish molecular concentration gradients, or from consistent parameter variation across space. In synthetic biology, prepatterns comprising chemicals [153] or light [166] may be manually established and subjected to post-processing by "lawns" of microbes. Despite its potential generality, mathematical analysis of prepattern

processing—where it exists—has been largely limited to data-driven modeling [92], [311], numerical simulations [312], [126], or information-theoretic approaches for specifying cell position [254].

We propose a discrete filtering approach to analyze how networks of interacting cells respond to prepatterns. The framework elucidates which components of spatial structure are amplified and which ones attenuated by the system to produce an output from any given input. The insights gained from this perspective challenge the conventional notion that instability is necessary for complex patterning; for example, our approach reveals that Turing-like stripes can emerge from a stable system lacking diffusion-driven instability, and furthermore that external noise reinforces rather than combats this behavior (Section 2.6).

To set up our framework, we combine the internal dynamics of cell behavior with interaction between cells by modeling each cell as an input-output module coupled to other modules. We examine the steady-state gains for constant-in-time, spatially varying inputs (prepatterns) and show that the system behaves as a discrete spatial filter, where the interconnectivity between cells determines the spatial modes, while the coupling and input-output dynamics dictate how each mode is scaled to generate a readout pattern. We also examine the system response to temporal and spatially varying noise inputs, measured with the $\mathcal{H}_2$ norm, to determine which spatial modes are sensitive to stochastic influence.

We next present a brief tutorial of spatial modes corresponding to various network structures for readers unfamiliar with the material. In the following section, we show how to apply the filtering approach by considering a simple model of gene expression that exemplifies two of the most common classes of filter behaviors—highpass or lowpass—depending on the choice of parameters. Finally, we demonstrate the utility of the filtering perspective by examining two developmental biological case studies, the Notch-Delta system in fruit fly wings (Section 2.5) and the Sox9/Wnt/Bmp network in vertebrate digit formation (Section 2.6), as well as a synthetic biological circuit for edge detection in binary images (Section 2.7). We conclude with a brief summary and areas for future research.

Throughout this chapter, we use the following notational conventions (illustrated in Fig. 2.1):

- Cells are indexed by i in vector form and spatial modes are indexed by k in vector form or (m, n) in an array, unless noted otherwise.
- Inputs except white noise in the context of the $\mathcal{H}_2$ norm are assumed constant in time.
- Vectors containing strictly constant-in-space entries are designated with an underline. The entries corresponding to any fixed point in space are additionally labeled with an overbar, e.g., $\underline{u} = \bar{u}\mathbb{1}_N$ where $\bar{u} \in \mathbb{R}$ and $\mathbb{1}_N$ is the length N vector of all ones.
- Steady-state values for time-dependent variables are designated with superscript asterisks. Constant-in-space steady-state (i.e., homogeneous) solutions to nonlinear systems are designated with both an asterisk and an underline, e.g., $\underline{y}^* = \bar{y}^*\mathbb{1}_N$.

- "Actual" values in the standard basis are unadorned. Perturbations from constant-in-space values are designated with a tilde; time-dependent perturbations are understood to be linear approximations of "actual" nonlinear solutions, e.g., $\tilde{x}_i(t) \approx x_i(t) - \underline{\bar{x}}^*$. Perturbed variables in the basis T are designated with a hat, e.g., $\hat{x}^* = T^{-1}\tilde{x}^*$.

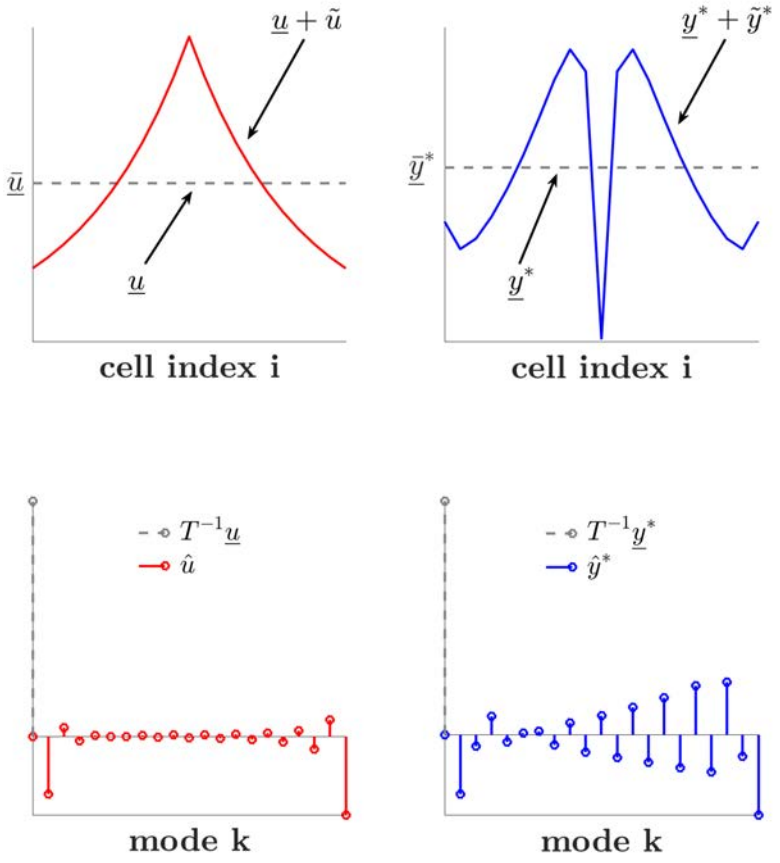

Figure 2.1: Schematic illustrating the variables utilized in the filter coefficient derivation. Left, red corresponds to input-related variables; right, blue corresponds to readout-related variables. Top, plots in the standard basis; the solid line for the input is the exact input to the full system, while the solid blue line is the approximated readout based on the filter coefficient analysis. Bottom, plots in the basis of the spatial modes. Dashed gray lines indicate values associated with the linearization; in this example, these constant-in-space inputs produce impulses at constant frequency (i.e., contribution to all other modes is zero—these values are not plotted). This example uses periodic boundary conditions and sinusoidal modes given by 2 with k (from 0 to $N-1$) indexing increasing frequency toward the middle of the x-axis, i.e., modes k and $N-k$ have the same frequency. The filter coefficients are symmetric about the midpoint but the mode representations of input and readout are not. This asymmetry captures the "location" of the standard-basis input and readout relative to the cell indices i, since modes k and $N-k$ have opposite phase (sign).

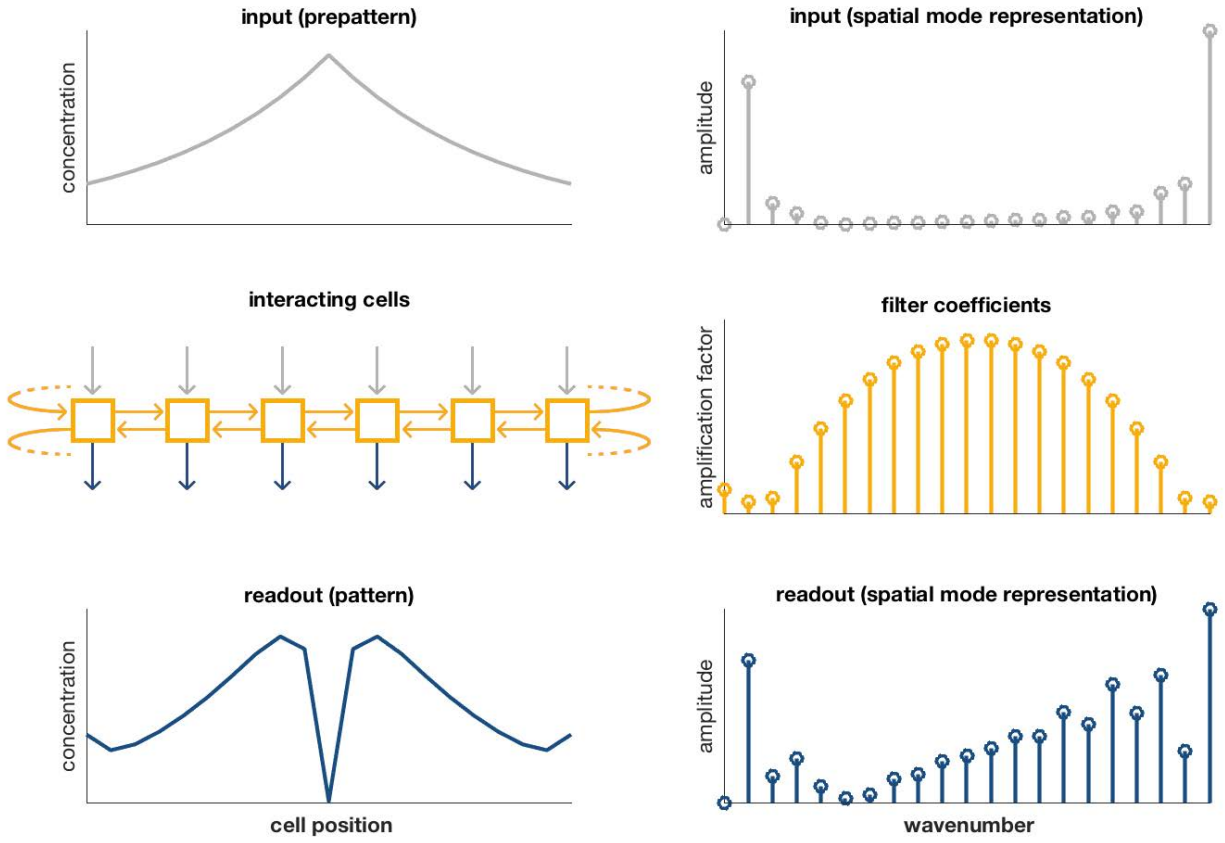

Figure 2.2: Weights of spatial modes in a prepattern are multiplied by filter coefficients, determined by the internal dynamics and interconnectivity among cells, to produce patterns. Each cell acts as an input/output module. The collective activity of cells produces the filtering behavior. Here, a single line of cells with periodic boundary conditions communicates through contact-based lateral inhibition, resulting in a highpass filter (see Section 2.5; note that wavenumber increases toward the center of the axis). The spatial modes differ from the standard Fourier basis (Observation 2), hence the asymmetry in the spatial mode representation of the input and readout (Figure 2.1).

2.2 Main Results

2.2.1 Filter Coefficients

We consider a generalized system of N identical cells with the state variables of the ith cell at time t given by $x_i(t) \in \mathbb{R}^n$, readout $y_i(t) \in \mathbb{R}$, and constant-in-time input $u_i \in \mathbb{R}$, which may represent an environmental stimulus or intrinsic parameter variation. Coupling occurs via $v_i(t) \in \mathbb{R}^q$ and output $w_i(t) \in \mathbb{R}^q$ where $q \leq n$. Let the vectors for the full system be the vertical concatenation $x(t)^T := [x_0(t)^T x_1(t)^T ... x_{N-1}(t)^T]$ and similarly for u, $y(t)$, $w(t)$, and $v(t)$. The dynamics of the ith cell and the full linear coupling between the N cells

are given by

$$\begin{cases} \dot{x}_i(t) = f(x_i(t), v_i(t), u_i) \\ w_i(t) = g(x_i(t)) \\ y_i(t) = h(x_i(t)) \\ v(t) = (M \otimes I_q)\, w(t) \end{cases} \tag{2.1}$$

where $\otimes$ is the Kronecker product, I_q is the $q \times q$ identity matrix, and $M \in \mathbb{R}^{N \times N}$.

The system (2.1) accommodates a wide range of specific deterministic models. Intercellular processes such as gene expression and protein decay are encapsulated by appropriate definition of the evolution function f for chemical concentrations x_i, including the effect of environmental stimuli or parameter values u_i as well as signals from neighbors v_i. The output w_i is the subset of elements in x_i that transmit signals to neighbors, with the method of transmission (e.g., diffusion, cell-to-cell contact) and the neighboring cells specified by the interconnection matrix M. The readout y_i isolates a quantity of interest to the user, which may be experimentally measurable (e.g., fluorescence) or simply relevant to a particular model (see examples in Sections 2.5, 2.6, and 2.7).

The vectors indexed by i describe patterns by the concentration of chemicals within individual cells at individual points in space. A full pattern is reconstructed from N elements, each of which represents the concentration of a chemical in a single cell. However, patterns can also be thought of as combinations of spatially varying components that span multiple cells, e.g., stripes of varying thickness (frequency). When weighted and summed, these spatial modes can represent arbitrary patterns of interest. We use the term "filtering" to refer to the process by which a network of interacting cells alters the weighting of the spatial modes of the input, thereby producing a readout that is built from the same components as, but differs in appearance from, the input. A key approximation to facilitate the analysis is that coupling between spatial modes is negligible, such that the readout can be expressed as a linear sum of the same set of spatial modes used to represent the input. In analogy to traditional signal processing, the network of cells plays the role of a linear time-invariant system (filter) that modifies the frequency components of a (spatially) varying signal. The following proposition formalizes this concept mathematically.

Proposition 1. *If the system described by* (2.1) *satisfies*

1. $M\mathbb{1}_N = \mu\mathbb{1}_N$ *and* M *is diagonalized by* T *(*$M = T\Lambda T^{-1}$*),*
2. *given* $\underline{\bar{u}} \in \mathbb{R}$, $\exists\ \underline{\bar{x}}^* \in \mathbb{R}^n$ *such that* $f\left(\underline{\bar{x}}^*, \mu g\left(\underline{\bar{x}}^*\right), \underline{\bar{u}}\right) = 0$ *and* $\underline{x}^* := \mathbb{1}_N \otimes \underline{\bar{x}}^*$, $\underline{u} := \underline{\bar{u}}\mathbb{1}_N$,
3. *the homogeneous steady state* $(\underline{x}^*, \underline{u})$ *is stable,*

then the system may be linearized about $(\underline{x}^*, \underline{u})$ *with linearization matrices*

$$A := \left.\frac{\partial f}{\partial x_i}\right|_{(\underline{\bar{x}}^*, \underline{\bar{v}}^*, \underline{\bar{u}})},\ B_v := \left.\frac{\partial f}{\partial v_i}\right|_{(\underline{\bar{x}}^*, \underline{\bar{v}}^*, \underline{\bar{u}})},\ B_u := \left.\frac{\partial f}{\partial u_i}\right|_{(\underline{\bar{x}}^*, \underline{\bar{v}}^*, \underline{\bar{u}})},\ C := \left.\frac{dh}{dy_i}\right|_{\underline{\bar{x}}^*}, G := \left.\frac{dg}{dx_i}\right|_{\underline{\bar{x}}^*}. \tag{2.2}$$

A constant-in-time, varying-in-space input u to (2.1) equates to a perturbing input

$$\hat{u} = T^{-1}(u - \underline{u}) \tag{2.3}$$

to the linearized system in the coordinate system T. Then the steady-state perturbed readout $\tilde{y}^*$ in the basis T is $\hat{y}^* := T^{-1}\tilde{y}^* = S\hat{u}$ where S is a diagonal matrix and

$$[S]_{kk} = -C(A + \lambda_k(M)B_vG)^{-1}B_u \tag{2.4}$$

is the steady-state gain of the kth eigenvector of T.

Derivation. Let $\underline{u} := \bar{u}\mathbb{1}_N$ be a spatially homogeneous input and assume $\exists\ \bar{x}^* \in \mathbb{R}^n$ such that $f(\bar{x}^*, \mu g(\bar{x}^*), \bar{u}) = 0$. Then $\underline{x}^* := \mathbb{1}_N \otimes \bar{x}^*$ is a homogeneous steady state. The remaining steady-state quantities are similarly designated $\underline{y}^* = \bar{y}^*\mathbb{1}_N$, $\underline{w}^* = g(\underline{x}^*)$, and $\underline{v}^* = M_q\underline{w}^* = \mathbb{1}_N \otimes \bar{v}^*$ (where $M_q := M \otimes I_q$). Let $\tilde{x}_i(t), \tilde{u}_i, \tilde{y}_i(t), \tilde{w}_i(t), \tilde{v}_i(t)$ denote perturbations about that steady state. The full system linearized about $(\underline{x}^*, \underline{u})$ yields perturbed dynamics

$$\dot{\tilde{x}}(t) = [(I_N \otimes A) + (I_N \otimes B_v)\, M_q\, (I_N \otimes G)]\, \tilde{x}(t) + (I_N \otimes B_u)\, \tilde{u} \tag{2.5}$$

where $A := \frac{\partial f}{\partial x_i}\big|_{(\bar{x}^*, \bar{v}^*, \bar{u})}$, $B_v := \frac{\partial f}{\partial v_i}\big|_{(\bar{x}^*, \bar{v}^*, \bar{u})}$, $B_u := \frac{\partial f}{\partial u_i}\big|_{(\bar{x}^*, \bar{v}^*, \bar{u})}$, $C := \frac{dh}{dy_i}\big|_{\bar{x}^*}$, and $G := \frac{dg}{dx_i}\big|_{\bar{x}^*}$ are the linearization matrices.

Assume that the interconnection matrix $M \in \mathbb{R}^{N\times N}$ is diagonalizable and let $M = T\Lambda T^{-1}$ be the diagonalization (so that M_q is diagonalized by $T \otimes I_q$). Define $\hat{x}(t) := (T^{-1} \otimes I_n)\, \tilde{x}(t)$, $\hat{u} := T^{-1}\tilde{u}$, and $\hat{y}(t) := T^{-1}\tilde{y}(t)$. Recasting (2.5) in the coordinate system T, we obtain the dynamical system

$$\dot{\hat{x}}(t) = [(I_N \otimes A) + \Lambda \otimes (B_vG)]\, \hat{x}(t) + (I_N \otimes B_u)\, \hat{u}. \tag{2.6}$$

In contrast to the conditions for spontaneous pattern formation, we will not require this system to be unstable; large amplification of spatial modes is possible even when the system is stable. The steady-state perturbed readout in basis T is

$$\hat{y}^* := -(I_N \otimes C)\, [(I_N \otimes A) + \Lambda \otimes (B_vG)]^{-1}\, (I_N \otimes B_u)\, \hat{u} \tag{2.7}$$

$$=: S\hat{u}, \tag{2.8}$$

where S is a diagonal matrix with entries

$$[S]_{kk} = -C(A + \lambda_k(M)B_vG)^{-1}B_u \tag{2.9}$$

for $k = 0, 1, ..., N-1$ and $[S]_{kk}$ collectively form the "filter coefficients" for the corresponding N spatial modes. The matrix S is thus analogous to a digital filter that processes the input $\tilde{u}$ into readout $\tilde{y}^*$ with respect to the eigenvectors, or spatial modes, of M as contained in T. ■

The conditions (1) through (3) ensure that the network, when given a constant-in-space input, will admit a stable, homogeneous steady-state solution, and that the expression pattern across cells can be represented in a complete orthonormal basis other than the standard; this basis T comprises the modes. The $[S]_{kk}$ collectively form the "filter coefficients", which dictate how the corresponding N spatial modes are multiplicatively scaled by the system when the input is no longer constant in space (Figure 2.2). In other words, the matrix S "filters" the perturbed input into a perturbed readout with respect to the eigenvectors, or spatial modes, of M as contained in T. In contrast to the conditions for spontaneous pattern formation, our approach does not require the perturbed system to be unstable; large amplification of spatial modes is possible even when the system is stable. Figure 2.8 shows eight examples of prototypical filter behaviors that vary with interaction type and cellular interconnectivity.

It is straightforward to extend the framework to multiple orthogonal signals sharing the same spatial modes. Let $M_0, M_1, ..., M_{q-1} \in \mathbb{R}^{N\times N}$ be the interconnection matrices for q orthogonal signals and assume they all commute (share the same basis). Let $\Delta_i \in \mathbb{R}^{q\times q}$ signify the matrix with (i,i)th entry one and all other entries zero, such that the full interconnectivity is

$$M = \sum_{i=0}^{q-1} \left(M_i \otimes \Delta_i\right). \tag{2.10}$$

If $\Lambda_i = T^{-1}M_iT$ for $i = 0, 1, ..., q-1$, then the basis $T \otimes I_q$ diagonalizes M such that the filter coefficients are given by

$$[S]_{kk} = -C\left(A + \sum_{i=0}^{q-1} \lambda_k\left(M_i\right) B_v \Delta_i G\right) B_u. \tag{2.11}$$

Many continuous pattern-forming and distributed dynamical systems exhibit spatial invariance of the dynamics with respect to linear transformations such as reflections, rotations, or translations [313], [309]. The discrete-space cellular network has a direct analog: If M is invariant under a linear transformation, then S is also invariant under the same transformation, since the system dynamics are identical within each cell and therefore the only spatial information contained within the system is contained in M. Formally:

Observation 1. *Under the assumptions listed in Proposition 1, if* $\Pi \in \mathbb{R}^{N\times N}$ *and* M *and* Π *commute, then* Π *and the filter coefficient matrix* S *also commute (i.e., the map from input to output is equivariant).*

Derivation. Let $\tilde{y}^* = S\tilde{u}$ and $\tilde{y}^*_\Pi = S\Pi\tilde{u}$. An equivalent statement to "S is equivariant" is then $\tilde{y}^*_\Pi = \Pi S\tilde{u} \implies \tilde{y}^*_\Pi = \Pi\tilde{y}^*$. To see this, take

$$\begin{aligned}
\tilde{y}^*_\Pi = S\Pi\tilde{u} &= -\left(I_N \otimes C\right)\left[\left(I_N \otimes A\right) + M \otimes \left(B_v G\right)\right]^{-1}\left(I_N \otimes B_u\right)\Pi\tilde{u} && (2.12)\\
&= -\left(I_N \otimes C\right)\left[\left(I_N \otimes A\right) + M \otimes \left(B_v G\right)\right]^{-1}\left(\Pi \otimes I_n\right)\left(I_N \otimes B_u\right)\tilde{u} && (2.13)\\
&= -\left(I_N \otimes C\right)\left[\left(\Pi^{-1} \otimes I_n\right)\left(\left(I_N \otimes A\right) + M \otimes \left(B_v G\right)\right)\right]^{-1}\left(I_N \otimes B_u\right)\tilde{u} && (2.14)\\
&= -\left(I_N \otimes C\right)\left[\left(\Pi^{-1} \otimes I_n\right)\left(I_N \otimes A\right) + \Pi^{-1}M \otimes \left(B_v G\right)\right]^{-1}\left(I_N \otimes B_u\right)\tilde{u}. && (2.15)
\end{aligned}$$

Since

$$\left(\Pi^{-1} \otimes I_n\right)\left(I_N \otimes A\right) = \left(I_N \otimes A\right)\left(\Pi^{-1} \otimes I_n\right), \tag{2.16}$$

then

$$M = \Pi M \Pi^{-1} \implies \Pi^{-1}M = M\Pi^{-1} \implies \Pi^{-1}M \otimes \left(B_v G\right) = M\Pi^{-1} \otimes \left(B_v G\right) \tag{2.17}$$

such that (2.15) becomes

$$\begin{aligned}
\tilde{y}^*_\Pi &= -\left(I_N \otimes C\right)\left[\left(\left(I_N \otimes A\right) + M \otimes \left(B_v G\right)\right)\left(\Pi^{-1} \otimes I_n\right)\right]^{-1}\left(I_N \otimes B_u\right)\tilde{u} && (2.18)\\
&= -\left(I_N \otimes C\right)\left(\Pi \otimes I_n\right)\left[\left(I_N \otimes A\right) + M \otimes \left(B_v G\right)\right]^{-1}\left(I_N \otimes B_u\right)\tilde{u} && (2.19)\\
&= -\Pi\left(I_N \otimes C\right)\left[\left(I_N \otimes A\right) + M \otimes \left(B_v G\right)\right]^{-1}\left(I_N \otimes B_u\right)\tilde{u} && (2.20)\\
&= \Pi\tilde{y}^*. && (2.21)
\end{aligned}$$

If $\Lambda = T^{-1}MT$, then since $M = \Pi M\Pi^{-1} \implies \Pi^{-1}M\Pi$, $\Lambda = T^{-1}\Pi^{-1}M\Pi T$ (i.e., T diagonalizes the permuted version of M with the same resultant eigenvalues). Note that the immutability of M under permutation Π confers immutability of Λ under permutation $T^{-1}\Pi T$, which is just the permutation in the basis of M; i.e.,

$$\Lambda = \left(T^{-1}\Pi T\right)\Lambda\left(T^{-1}\Pi T\right)^{-1}, \tag{2.22}$$

or equivalently, Λ and $(T^{-1}\Pi T)$ commute. ■

Thus, M's permutation $M_\Pi := \Pi M\Pi^{-1}$ shares the same eigenvectors T and corresponding eigenvalues Λ as M, which implies that the filter coefficients for a system with interconnection matrix M are the same as for that system with interconnection matrix M_Π.

2.2.2 Stochastic Influence on Patterning

The role of stochastic influences in biological patterning is a subject of ongoing theoretical and experimental interest (e.g., [17], [159]). Here, we concern ourselves with the response of spatial modes to time-varying white noise inputs, for which the $\mathcal{H}_2$ norm of the system quantifies the expected power of the perturbed readout. The $\mathcal{H}_2$ norm has previously been used to analyze energy amplification in channel flows [314], networks of cells [122], and reaction-diffusion systems [315], among others.

To begin our analysis we rewrite the linearized ordinary differential equations in the form of a nonlinear Langevin equation (Itô interpretation). Since the N modes are decoupled we can write the equation for the perturbed states in the kth mode as

$$d\hat{x}_k(t) = \left[(I_N \otimes A) + \Lambda \otimes (B_v G)\right] \hat{x}_k(t)dt + (I_N \otimes B_u)\, d\hat{u}_k(t). \tag{2.23}$$

Here $\hat{u}(t)$ is an n_u-dimensional independent standard Wiener process, also known as the standard Brownian motion process. Implicitly we assume that concentrations of reactants are high enough to permit us to neglect molecular-level fluctuations, which cannot accurately be described by the Langevin approach [316].

With slight abuse of notation, $d\hat{u}_k(t)$ is stationary, therefore the variance of the readout $y(t) = Cx(t)$ in mode k does not change in time. The variance is given by

$$\mathbb{E}\left[|\hat{y}_k|^2\right] = \mathbb{E}\left[Tr\left(\hat{y}_k\hat{y}_k^T\right)\right] = Tr\left(C\mathbb{E}\left[\hat{x}_k\hat{x}_k^T\right]C^T\right) = Tr\left(CQ_kC^T\right), \tag{2.24}$$

where $Q_k := \mathbb{E}\left[\hat{x}_k\hat{x}_k^T\right]$ is the covariance of the reactants in the kth mode.

Let $G_k(t)$ be the impulse response of (2.23) for readout $y(t)$. We could equivalently write

$$\mathbb{E}\left[Tr\left(\hat{y}_k\hat{y}_k^T\right)\right] = \int_0^\infty \mathbb{E}\left[Tr\left(G_k(t)d\hat{u}_k d\hat{u}_k^T G_k(t)^T\right)\right] dt \tag{2.25}$$

$$= \int_0^\infty Tr\left(G_k(t)\mathbb{E}\left[d\hat{u}_k d\hat{u}_k^T\right] G_k(t)^T\right) dt = \int_0^\infty Tr\left(G_k(t)G_k(t)^T\right) dt \tag{2.26}$$

$$=: ||G_k(t)||_{\mathcal{H}_2}^2, \tag{2.27}$$

from which we deduce that the $\mathcal{H}_2$ norm is equivalent to the variance of $\hat{y}_k$ and can be calculated as $Tr\left(CQ_kC^T\right)$ where Q_k is the positive semi-definite solution to the Lyapunov equation

$$(A + \lambda_k B_v G)\, Q_k + Q_k\, (A + \lambda_k B_v G)^T + B_u B_u^T = 0. \tag{2.28}$$

The unit variance of $d\hat{u}_k(t)$ allows us to interpret $||G_k(t)||_{\mathcal{H}_2}^2$ as the ratio of the variance of the readout to the variance of the input in mode k. Moreover, since $d\hat{u}_k(t)$ is zero mean, the squared $\mathcal{H}_2$ norm is also equivalent to the time integral of the expected power spectral density, or the factor by which the system amplifies the average power of the readout within mode k. Those modes with the highest $\mathcal{H}_2$ norms are most strongly amplified by the external noise source.

2.3 Spatial Modes and the Interconnection Matrix

In the remainder of this chapter we construct the interconnection matrix M for a particular signal as follows:

1. The length N vector of all ones $\mathbb{1}_N$ is an eigenvector of M, which implies that a homogeneous steady-state solution exists.

2. The ith, jth entry $[M]_{ij}$ for $i \neq j$ is 0 if cell i is not connected to cell j. Otherwise $0 < [M]_{ij}$, where the magnitude $[M]_{ij}$ captures the "strength" of the connection.

3. The diagonal entries $[M]_{ii}$ encapsulate the "signaling cost" associated with interaction. Negative values imply the cell loses signal to transmit to its neighbors, e.g., diffusion.

In many biological systems, cells can be approximated to have the same distance between them and the same communication strength with each of their neighbors. In such systems, the corresponding spatial modes are sinusoidal, giving rise to stripes or spots. Lower-frequency modes correspond to longer-wavelength spatial modes, while higher-frequency modes correspond to shorter-wavelength spatial modes. The relationship between patterning wavelength and spatial mode frequency enables these systems to be interpreted from the standpoint of how the weights of the frequency components in an input are scaled to produce the readout, analogous to filtering as it is understood in discrete signal processing. In this chapter we will consider basis vectors arising from a line or sheet of regularly spaced cells with periodic or no-flux boundary conditions. The modes then pertain to two common signal processing transforms: the discrete Fourier transform (DFT) for periodic boundaries or the second discrete cosine transform (DCT-2) for no-flux boundaries. The eigenvectors and eigenvalues for these transforms are well known (e.g., [317]); a review is offered in the following subsections, which the familiar reader may skip. We will assume modes are indexed in order of increasing frequency with increasing k toward $\frac{N}{2}$ for the DFT and N for the DCT-2.

2.3.1 Discrete Fourier Transform (DFT)

If the N cells form a ring indexed clockwise or counterclockwise, then M is circulant. The eigenvectors of a circulant matrix form the discrete Fourier basis such that the spatial modes of T correspond exactly to the frequencies of sinusoids.

We can choose T to be the discrete Fourier transform matrix (DFT) where the jth entry of the kth eigenvector, $j, k = 0, 1, ..., N-1$, is given by

$$[T]_{jk} = \frac{1}{\sqrt{N}} e^{-\frac{2\pi i j k}{N}} \tag{2.29}$$

with $i := \sqrt{-1}$. T is conjugate symmetric. If we let $m_0,\ m_1,\ ..., m_{N-1}$ denote the entries in the first row of M, then the eigenvalues of M are given by

$$\lambda_k(M) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} m_n e^{-\frac{2\pi i j k}{N}}, \tag{2.30}$$

which corresponds to the coefficients of the discrete Fourier transform (DFT) of the first row of M.

If M is symmetric in addition to circulant, then we can alternatively select the eigenvectors such that all entries are real.

Observation 2. *Let $M \in \mathbb{R}^{N\times N}$ be a symmetric circulant matrix where $m_0 \in \mathbb{R}^N$ is the first row and define m as the periodization of m_0. Let the matrix T have entries*

$$[T]_{jk} = \frac{1}{\sqrt{N}}\left(\cos\frac{2\pi jk}{N} + \sin\frac{2\pi jk}{N}\right). \tag{2.31}$$

Then T is a basis for M with eigenvalues

$$\lambda_k(M) = \sum_{n=0}^{N-1} m_n \cos\frac{2\pi nk}{N}, \tag{2.32}$$

$k = 0, 1, ..., N-1$.

Derivation. Let W be the unitary DFT matrix, i.e., the jth entry of the kth column is

$$W_k^j = \frac{1}{\sqrt{N}} e^{\frac{i2\pi jk}{N}}. \tag{2.33}$$

We can express T as

$$T = \frac{1}{2}\left[W + W^H + e^{\frac{-i\pi}{2}}W + \left(e^{\frac{-i\pi}{2}}W\right)^H\right]. \tag{2.34}$$

Then

$$T^{-1}MT = T^HMT = TMT \tag{2.35}$$

$$= \frac{1}{4}\left[W + W^H + e^{\frac{-i\pi}{2}}W + \left(e^{\frac{-i\pi}{2}}W\right)^H\right] M \left[W + W^H + e^{\frac{-i\pi}{2}}W + \left(e^{\frac{-i\pi}{2}}W\right)^H\right] \tag{2.36}$$

$$= \frac{1}{4}\left(W + W^H\right) M \left(W + W^H\right) + \frac{1}{4}\left(e^{\frac{-i\pi}{2}}W + e^{\frac{i\pi}{2}}W^H\right) M \left(e^{\frac{-i\pi}{2}}W + e^{\frac{i\pi}{2}}W^H\right) + \frac{1}{4}\left(W + W^H\right) M \left(e^{\frac{-i\pi}{2}}W + e^{\frac{i\pi}{2}}W^H\right) + \frac{1}{4}\left(e^{\frac{-i\pi}{2}}W + e^{\frac{i\pi}{2}}W^H\right) M \left(W + W^H\right) \tag{2.37}$$

$$= \frac{1}{2}W^HMW + \frac{1}{2}WMW^H + \frac{1}{2}\cos\frac{\pi}{2}\left[W^HMW + WMW^H\right] + \frac{1}{2}e^{\frac{i\pi}{2}}W^HMW^H + \frac{1}{2}e^{\frac{-i\pi}{2}}WMW \tag{2.38}$$

$$= \frac{1}{2}W^HMW + \frac{1}{2}WMW^H + \frac{1}{2}e^{\frac{i\pi}{2}}W^HMW^H + \frac{1}{2}e^{\frac{-i\pi}{2}}WMW. \tag{2.39}$$

Since M is real and even (symmetric), the DFT is also real. The symmetry of M together with the symmetry of W and the fact that diagonal matrices are symmetric also imply that

$W^H MW = (W^H MW)^T = W^T M^T W^{*T} = WMW^H$, so we can somewhat simplify (2.39) to

$$W^H MW + \frac{1}{2}\left[e^{\frac{i\pi}{2}} W^H MW^H + e^{\frac{-i\pi}{2}} WMW\right] \tag{2.40}$$

$$= W^H MW + \frac{1}{2}\left[e^{\frac{i\pi}{2}} \left(W^H MW\right) W^H W^H + e^{\frac{-i\pi}{2}} \left(WMW^H\right) WW\right] \tag{2.41}$$

$$= W^H MW + W^H MW \frac{1}{2}\left[e^{\frac{i\pi}{2}} W^H W^H + e^{\frac{-i\pi}{2}} WW\right]. \tag{2.42}$$

For $k = 1, 2, ..., \frac{N}{2}$ (N odd) or $k = 1, 2, ..., \frac{N-1}{2}$ (N even), the $(N-k)$th row or column of W is equal to the kth row or column of W^H. Therefore $[WW]_{k,N-k} = 1$. Because of the orthogonality of complex exponentials, the remaining entries are 0. By the same logic we deduce an identical structure for $W^H W^H$ such that $WW = W^H W^H$. Now (2.42) becomes

$$W^H MW + W^H MW \frac{1}{2}\left[e^{\frac{i\pi}{2}} WW + e^{\frac{-i\pi}{2}} WW\right] \tag{2.43}$$

$$= W^H MW + W^H MW\,(WW)\cos\frac{\pi}{2} \tag{2.44}$$

$$= W^H MW, \tag{2.45}$$

which is just M diagonalized by the complex exponential DFT matrices, as desired. From this we derive that the eigenvalues are the same as the DFT coefficients of h_0, which owing to symmetry may be calculated as

$$\lambda_m(M) = \sum_{n=0}^{N-1} m_n \cos\frac{2\pi nm}{N}. \tag{2.46}$$

■

Owing to the periodicity of cosine, the eigenvalues of symmetric circulant M (and hence the corresponding $[S]_{kk}$) are symmetric about the highest-frequency eigenvector associated with $k = \frac{N}{2}$.

A situation of particular interest occurs when v_i is a diffusible molecule and the connection strength is equal between cells. Then M is a scaled version of the circulant finite differences (Laplacian) matrix

$$M = \begin{bmatrix} -2 & 1 & 0 & \dots & 0 & 1 \\ 1 & -2 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \dots & 1 & -2 \end{bmatrix} \tag{2.47}$$

with eigenvalues

$$\lambda_k(M) = -2 + 2\cos\frac{2\pi k}{N}. \tag{2.48}$$

This form of M corresponds to the second differences matrix for a system with periodic boundary conditions [317].

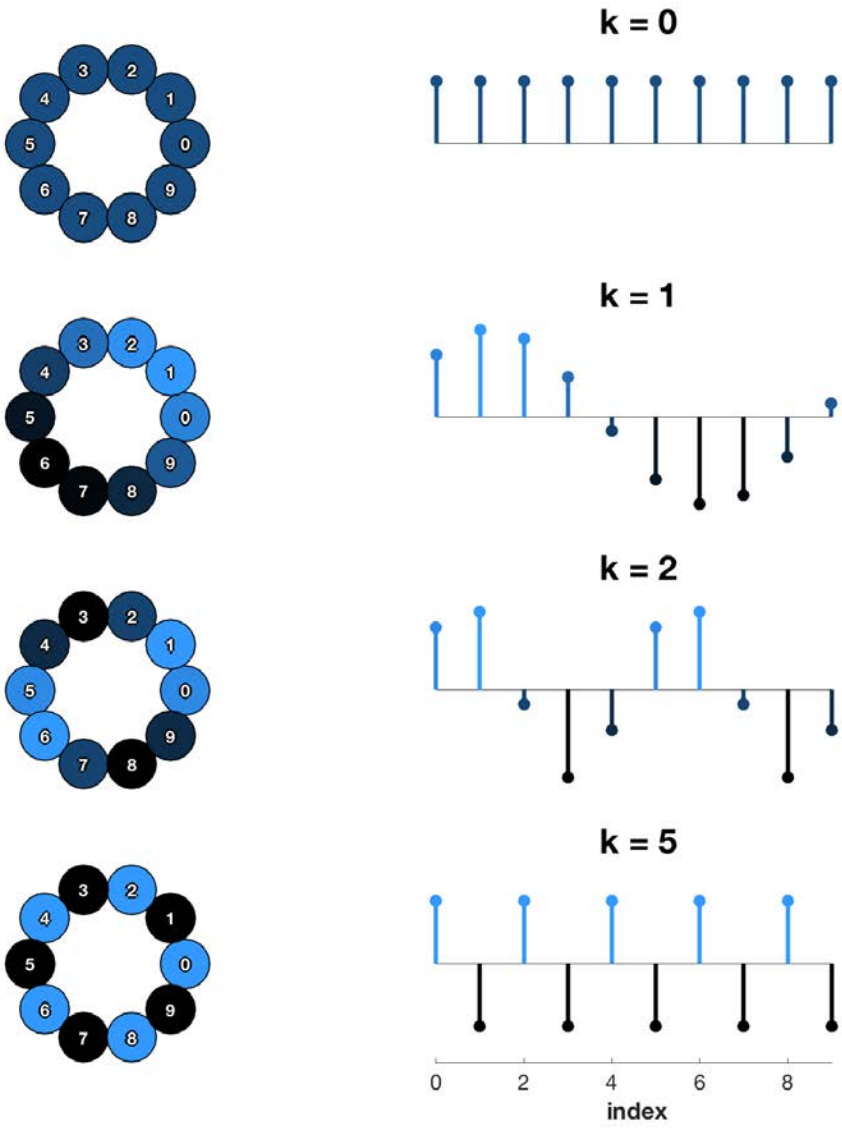

Figure 2.3: Sample spatial modes for a ring interconnectivity with $N = 10$ cells. Each cell is connected to each of its two neighbors with equal connection strength. M is circulant, so the eigenvectors form the discrete Fourier basis such that the spatial modes of T correspond exactly to the frequencies of sinusoids. The value k corresponds to the spatial frequency, or the number of complete periods present in a single cycle around the ring. Because the basis is the discrete Fourier transform (DFT) and N is even, the highest frequency is $\frac{N}{2} = 5$. Neighboring cells in this mode alternate between two values. Such a pattern is not possible in a ring configuration when N is odd.

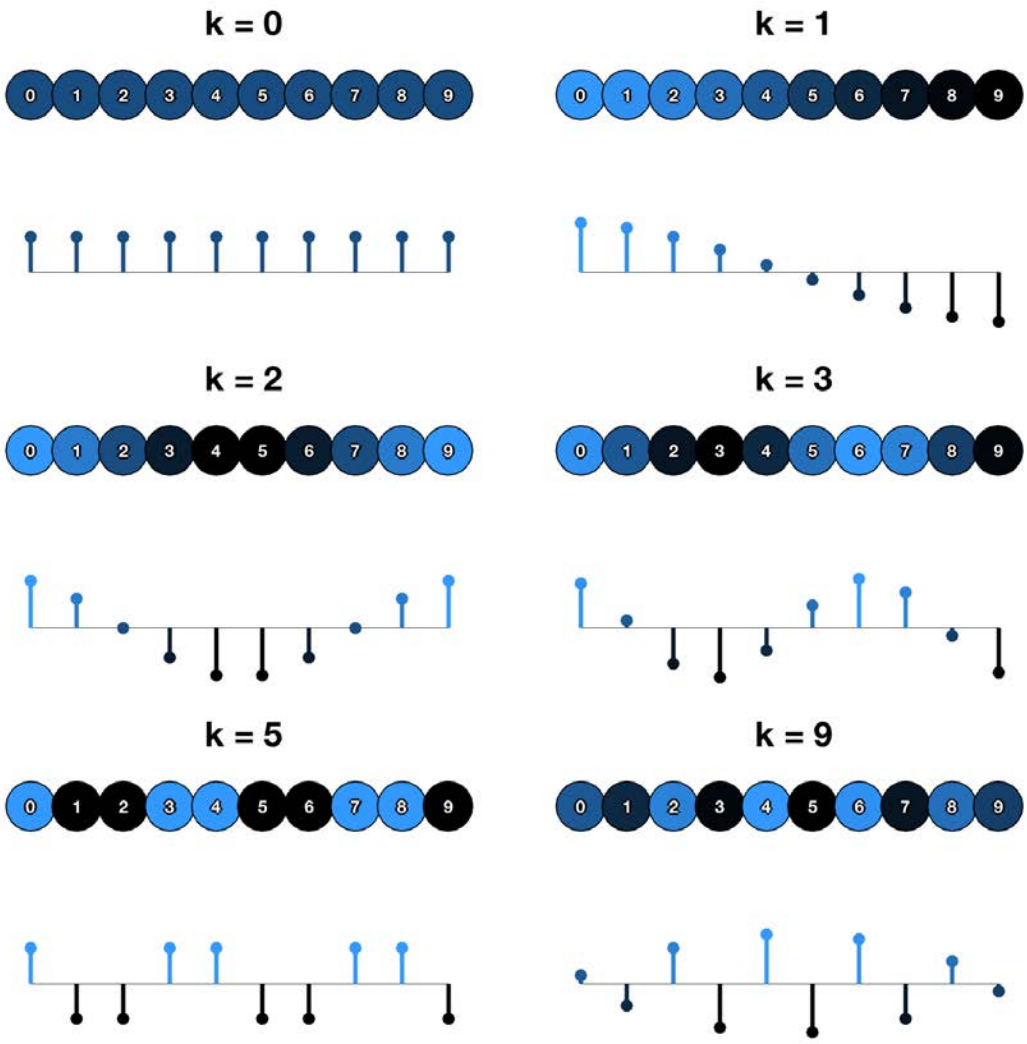

Figure 2.4: Sample spatial modes for a line interconnectivity with $N = 10$ cells. The two cells on the end each interact with only one neighbor such that the basis vectors are the DCT-2 vectors. The value k corresponds to twice the frequency of its corresponding mode, i.e., $\frac{k}{2}$ periods are represented in mode k. "Even" modes (k even) have symmetry about the midpoint between cells $\frac{N}{2} - 1$ and $\frac{N}{2} + 1$, while "odd" modes (k odd) are antisymmetric about this same point. If N were odd, the midpoint would instead be the $\left(\frac{N-1}{2}\right)$th cell.

2.3.2 Second Discrete Cosine Transform (DCT-2)

If the N cells are organized in a line, then the two cells on the end each communicate with only one neighbor. If the mode of communication is a diffusible molecule and the cells are indexed from one end of the line to the other, then the connectivity takes the form of a second differences matrix with Neumann boundary conditions centered at the midpoint:

$$M = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -2 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & -1 \end{bmatrix}. \tag{2.49}$$

The spatial modes of T form the basis for the second discrete cosine transform (DCT-2). The jth entry of the kth eigenvector, $j, k = 0, 1, ..., N-1$, is given by

$$[T]_{jk} = \sqrt{\frac{2}{N}} \cos\left[\left(j + \frac{1}{2}\right)\frac{k\pi}{N}\right] \tag{2.50}$$

(for $k = 0$, divide by additional factor of $\sqrt{2}$) with corresponding eigenvalue

$$\lambda_k(M) = -2 + 2\cos\frac{k\pi}{N}. \tag{2.51}$$

The highest frequency is $k = N - 1$ [317]. Unlike the case of circulant M, there are no guarantees of symmetry in the filter S and T itself is not conjugate symmetric.

2.3.3 2D Spatial Arrays

Consider a two-dimensional, rectangular $N_R \times N_C$ array of $N := N_C N_R$ cells indexed from 0 to $N_C N_R - 1$ starting in the upper corner from top to bottom and then left to right, i.e.,

0	N_R	$2N_R$	$\dots$	$(N_C-1)N_R$
1	N_R+1	$2N_R+1$	$\dots$	$(N_C-1)N_R+1$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
N_R-2	$2N_R-2$	$3R-2$	$\dots$	$N_C N_R-2$
N_R-1	$2N_R-1$	$3N_R-1$	$\dots$	$N_C N_R-1$

. (2.52)

It is known (e.g., [310]) that if any isolated row has interconnection matrix $M_R \in \mathbb{R}^{N_C \times N_C}$ and any isolated column has interconnection matrix $M_C \in \mathbb{R}^{N_R \times N_R}$, then the full matrix M for the interconnectivity of the entire array is

$$M := (M_R \otimes I_{N_R}) + (I_{N_C} \otimes M_C). \tag{2.53}$$

If T_R and T_C diagonalize M_R and M_C respectively then M is diagonalized by

$$T := (T_R \otimes I_{N_R})(I_{N_C} \otimes T_C) = T_R \otimes T_C, \tag{2.54}$$

giving $N_C N_R$ eigenvalues

$$\lambda_{m+nN_R}(M) = \lambda_m(M_C) + \lambda_n(M_R) \tag{2.55}$$

where $m = 0, 1, ...N_R - 1$, $n = 0, 1, ..., N_C - 1$. The (m, n)th spatial mode is given by $T_C^m T_R^{nT}$.

We can explicitly relate the spatial modes for a 2D array of cells to constituent modes in the horizontal and vertical directions by recasting the vector $\hat{y}$ in matrix form. Let U and Y be matrices arranged as in (2.52) where u_i is the input to compartment i. Vector form is recovered through the vectorization operation $vec(U) = u$. The readout matrix Y is defined similarly. If the matrices $\tilde{U}$ and $\tilde{Y}$ designate perturbations from steady state in the original basis and $\hat{Y}$, $\hat{U}$ designate perturbations in the basis for the spatial modes, then

$$\tilde{Y} = T_C \hat{Y} T_R^T = T_C \left(\Lambda_S \odot \hat{U}\right) T_R^T \tag{2.56}$$

where $\odot$ is the Hadamard product (element-by-element multiplication) and Λ_S has mth, nth entry

$$[\Lambda_S]_{mn} = -C\left[A + (\lambda_m(M_C) + \lambda_n(M_R))\, B_v G\right]^{-1} B_u. \tag{2.57}$$

From this it can be seen that the full system alters the input along the ith vertical spatial mode and the jth horizontal spatial mode defined by the vertical and horizontal connectivities.

Observation 3. *Consider an $N_R \times N_C$ array of cells. Let M_F describe the interconnectivity of $C_D := min(N_R, N_C)$ elements in the forward diagonal direction and M_B the interconnectivity of C_D elements in the backward diagonal direction. Let $P \in \mathbb{R}^{C_D \times C_D}$ be the permutation matrix with lower diagonal ones and the last entry of the first column also one, such that $P_F := P_B^T := diag\left(P^0, P^1, P^2, ..., P^{R_D}\right)$ where $R_D := max(N_R, N_C)$. In total, the interconnectivity of an array with horizontal, vertical, and diagonal components is described by*

$$M := (M_R \otimes I_{N_R}) + (I_C \otimes M_C) + P_F^T (M_F \otimes I_{R_D}) P_F + P_B^T (M_B \otimes I_{R_D}) P_B \tag{2.58}$$

where M_R, M_C, M_F, and M_B are circulant. Furthermore, if $N_R = N_C = N_0$ and the real or complex DFT basis T_0 diagonalizes each of M_R, M_C, M_F, and M_B individually, then $(T_0 \otimes T_0)$ diagonalizes M.

Derivation. Let T_0 diagonalize M_0 such that $(T_0 \otimes T_0)$ diagonalizes $(M_R \otimes I_{N_R}) + (I_C \otimes M_C)$. Since M_0 is circulant,

$$M_0 = \begin{bmatrix} m_0 & m_{N_R-1} & m_{N_R-2} & \dots & m_1 \\ m_1 & m_0 & m_{N_R-1} & \dots & m_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{N_R-1} & m_{N_R-2} & m_{N_R-3} & \dots & m_0 \end{bmatrix} = m_0 P^0 + m_1 P + m_2 P^2 + ... + m_{N_R-1} P^{N_R-1} \tag{2.59}$$

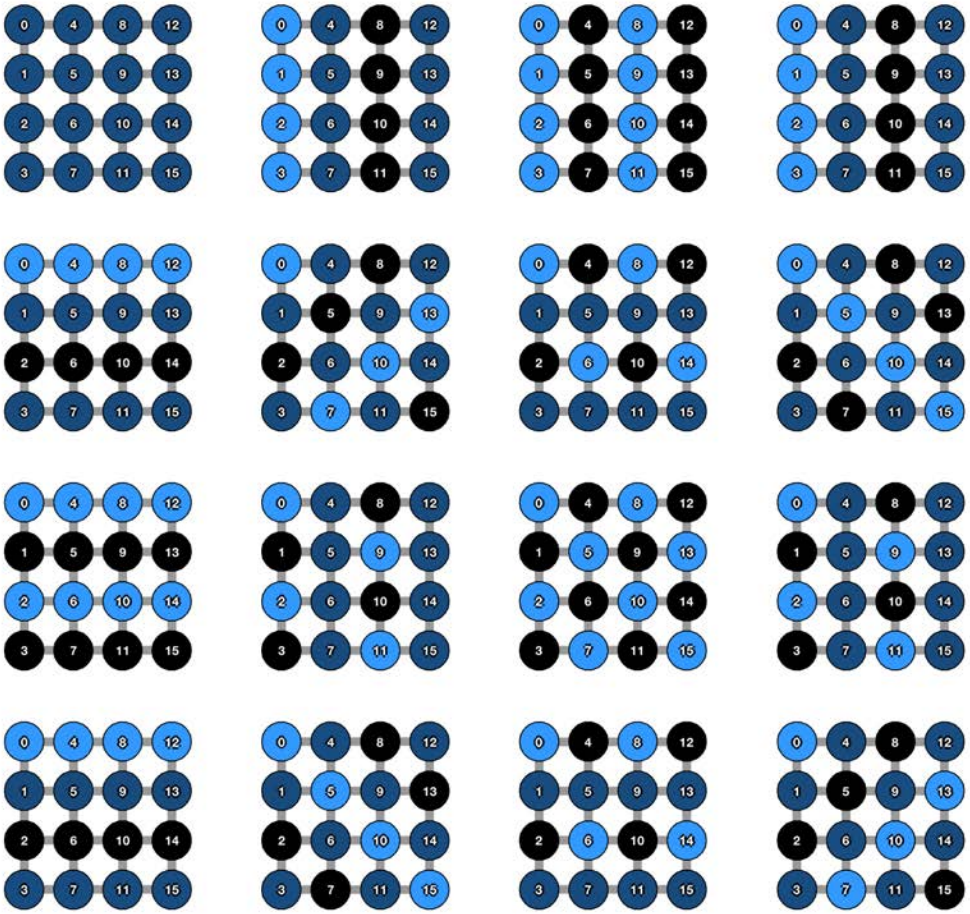

Figure 2.5: A complete set of spatial modes for a 2D periodic boundary interconnectivity (DFT basis) on a 4×4 rectangular array. Due to the symmetry in the eigenvectors, modes (m, n) and $(m, 4-n)$ are identical, as are (m, n) and $(4-m, n)$.

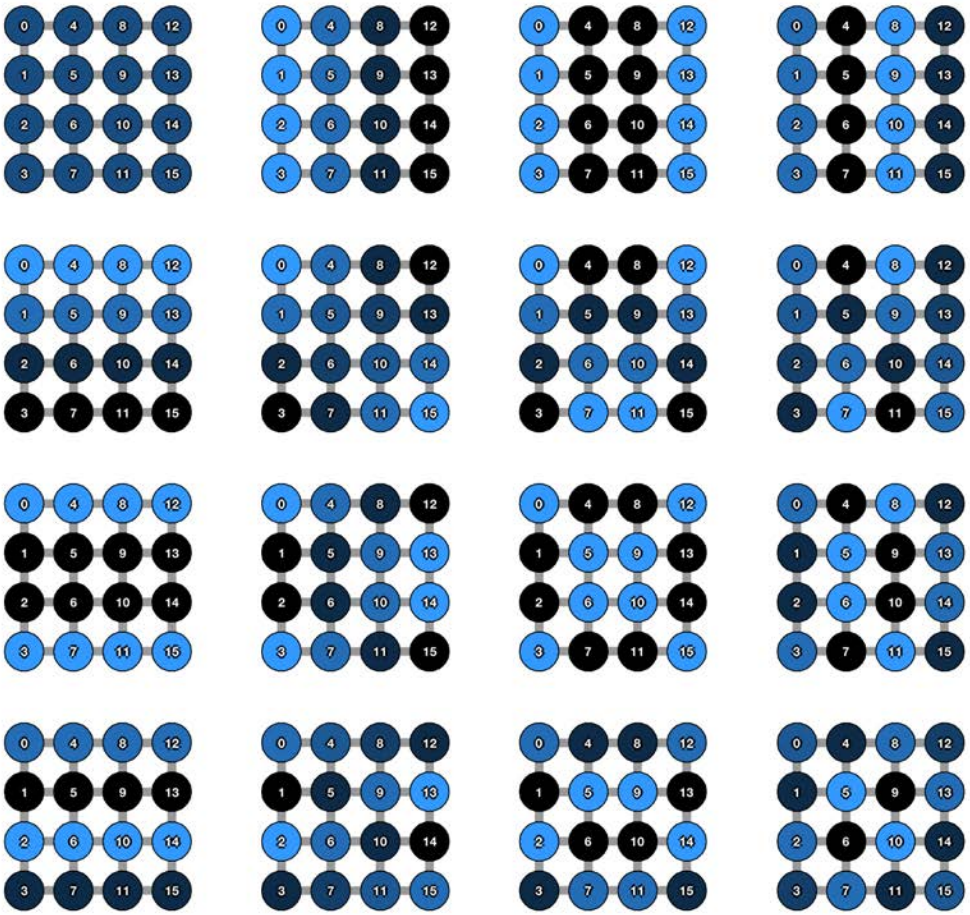

Figure 2.6: A complete set of spatial modes for a 2D interconnectivity with Neumann boundary conditions (DCT-2 basis) on a 4×4 rectangular array. Modes are indexed such that the (m, n)th mode has frequency $\frac{\pi m}{N_R}$ in the vertical direction (down rows) and $\frac{\pi n}{N_C}$ in the horizontal direction (across columns).

and therefore

$$P_F^{-1}\left(M_0 \otimes I_{N_R}\right) P_F = P_F^{-1}\left(\sum_{q=0}^{N_R-1} m_q P^q \otimes I_{N_R}\right) P_F \tag{2.60}$$

$$= \sum_{q=0}^{N_R-1} m_q P_F^{-1}\left(P^q \otimes I_{N_R}\right) P_F. \tag{2.61}$$

If we expand P_F and use the fact that $P^{-1} = P^{N_R-1} = P^T$, then the summation simplifies to

$$\sum_{q=0}^{N_R-1} m_q \left(P^q \otimes P^{-q}\right) \tag{2.62}$$

which is diagonalized by the complex exponential DFT vectors T_0^{-1} and T_0 as follows:

$$\left(T_0^{-1} \otimes T_0^{-1}\right)\left[\sum_{q=0}^{N_R-1} m_q \left(P^q \otimes P^{-q}\right)\right]\left(T_0 \otimes T_0\right) = \sum_{q=0}^{N_R-1} m_q \left(T_0^{-1} \otimes T_0^{-1}\right)\left(P^q \otimes P^{-q}\right)\left(T_0 \otimes T_0\right) \tag{2.63}$$

$$= \sum_{q=0}^{N_R-1} m_q \left(T_0^{-1} P^q T_0\right) \otimes \left(T_0^{-1} \otimes P^{-q} T_0\right), \tag{2.64}$$

which is a sum of diagonal matrices because permutation matrices are circulant and therefore diagonalized by the DFT matrices T_0, and the Kronecker product of two diagonal matrices is diagonal. Since $P_B = P_F^T = P_F^{-1}$, the same derivation for diagonal connectivity in the backward direction gives

$$P_B^{-1}\left(M_0 \otimes I_{N_R}\right) P_B = \sum_{q=0}^{N_R-1} m_q \left(P^q \otimes P^q\right), \tag{2.65}$$

which is also diagonalized by DFT matrices as

$$\sum_{q=0}^{N_R-1} m_q \left(T_0^{-1} P^q T_0\right) \otimes \left(T_0^{-1} \otimes P^q T_0\right) \tag{2.66}$$

where each summand is diagonal. This implies that the complex exponential form of the DFT is the basis for an array that is diagonally connected in either or both directions.

If M_0 is symmetric, then in addition to the complex exponential basis T_0 we might also choose the basis T_0' with the jth entry of kth eigenvector given by

$$T_0'(k) = \frac{1}{\sqrt{R}}\left(\cos\frac{2\pi jk}{R} + \sin\frac{2\pi jk}{R}\right), \tag{2.67}$$

in which case each term in the summation (2.64) is no longer diagonal because T_0' does not diagonalize the nonsymmetric permutation matrices. Hence for the real-valued basis T_0' we require that the array is diagonally connected in both forward and backward directions such that the full connectivity matrix is given by the sum of (2.65) and (2.62):

$$\sum_{q=0}^{N_R-1} m_q \left(P^q \otimes P^{-q}\right) + \sum_{q=0}^{N_R-1} m_q \left(P^q \otimes P^q\right) = \sum_{q=0}^{N_R-1} m_q P_q \otimes \left(P^{-q} + P^q\right). \tag{2.68}$$

The matrix $(P^{-q} + P^q)$ is circulant and symmetric and hence diagonalized by T_0'. To complete the argument we appeal to the symmetry of m_q. Specifically, if N_R is odd,

$$\sum_{q=0}^{N_R-1} m_q P^q \otimes \left(P^{-q} + P^q\right) \tag{2.69}$$

$$= 2m_0 I_{N_R} + \sum_{q=1}^{\frac{N_R-1}{2}} m_q P_q \otimes \left(P^{-q} + P^q\right) + \sum_{q=\frac{N_R-1}{2}+1}^{N_R-1} m_q P^q \otimes \left(P^{-q} + P^q\right) \tag{2.70}$$

$$= 2m_0 I_{N_R} + \sum_{q=1}^{\frac{N_R-1}{2}} m_q P^q \otimes \left(P^{-q} + P^q\right) + m_{R-q} P^{R-q} \otimes \left(P^{q-R} + P^{R-q}\right) \tag{2.71}$$

$$= 2m_0 I_{N_R} + \sum_{q=1}^{\frac{N_R-1}{2}} m_q P^q \otimes \left(P^{-q} + P^q\right) + m_q P^{-q} \otimes \left(P^q + P^{-q}\right) \tag{2.72}$$

$$= 2m_0 I_{N_R} + \sum_{q=1}^{\frac{N_R-1}{2}} m_q \left(P^q + P^{-q}\right) \otimes \left(P^{-q} + P^q\right), \tag{2.73}$$

where each individual term is diagonalized by T_0', and hence the whole summation is diagonal. If N_R is even, the summation (2.68) breaks into

$$2m_0 I_{N_R} + m_{\frac{N_R}{2}} P^{\frac{N_R}{2}} \otimes \left(P^{\frac{N_R}{2}} + P^{-\frac{N_R}{2}}\right) + \sum_{q=1}^{\frac{N_R-1}{2}} m_q \left(P^q + P^{-q}\right) \otimes \left(P^{-q} + P^q\right). \tag{2.74}$$

When N_R is even, $P^{R/2}$ alone is circulant symmetric, hence the additional term is also diagonalized by T_0'. Therefore when M_0 is circulant symmetric, the matrix

$$M = P_F^T \left(M_0 \otimes I_{N_R}\right) P_F + P_B^T \left(M_0 \otimes I_{N_R}\right) P_B \tag{2.75}$$

is diagonalized by $T' := T_0' \otimes T_0'$. Note that this implies

$$\left(M_0 \otimes I_{N_R}\right) + \left(I_{N_R} \otimes M_0\right) + P_F^T \left(M_0 \otimes I_{N_R}\right) P_F + P_B^T \left(M_0 \otimes I_{N_R}\right) P_B \tag{2.76}$$

is also diagonalized by T'. ■

Remark. For a system with diagonal connections only ($M_R = M_C = 0$), then if $N_R = N_C$ odd, the diagonal transformations are identical to a 2D DFT rotated 45°. For $N_R = N_C$ even, the array becomes divided into two separate classes that are transformed separately; i.e., the underlying network graph is no longer connected. This is because for $N_R = N_C$ odd, the array has a compartment at the center, while for $N_R = N_C$ even, the center would (in physical space) represent a crossing of intersections.

Observation 4. *Let $N_C = N_R = N_0$, M_0 circulant with complex exponential basis vectors T_0 and consider the full forward diagonal interconnection matrix $M = P_F^T (M_0 \otimes I_{N_0}) P_F$. The (m,n)th eigenvalue of M is*

$$\lambda_{m+nN_0}(M) = \begin{cases} m_0 + 2\sum_{q=0}^{\frac{N_0-1}{2}} m_q \cos\frac{2\pi q(n-m)}{N_0}, & N_0 \text{ odd}, \\ m_0 + m_{\frac{N_0}{2}}(-1)^{n-1} + 2\sum_{q=0}^{\frac{N_0}{2}-1} m_q \cos\frac{2\pi q(n-m)}{N_0}, & N_0 \text{ even}, \end{cases} \tag{2.77}$$

where m_k is the kth entry of the first row or column of M_0.

Derivation. For $N_C = N_R = N_0$ and M_F as defined above, the kth entry of the N_0-point DFT of the first row of P^q is $e^{\frac{2\pi ikq}{N_0}}$, $k = 0, 1, ..., N_0$ and the diagonalization $T_0^{-1} P^q T_0$ is the matrix with the DFT entries on the diagonal. This implies that we can write (2.64) as

$$\left(T_0^{-1} \otimes T_0^{-1}\right) P_F^T (I_{N_0} \otimes M_0) P_F (T_0 \otimes T_0) = \sum_{q=0}^{N_0-1} m_q \left(T_0^{-1} P^q T_0\right) \otimes \left(T_0^{-1} P^{-q} T_0\right) \tag{2.78}$$

$$= \sum_{q=0}^{N_0-1} m_q \begin{bmatrix} 1 & & & & \\ & e^{\frac{2\pi iq}{N_0}} & & & \\ & & e^{\frac{4\pi iq}{N_0}} & & \\ & & & \ddots & \\ & & & & e^{\frac{2\pi iq(N_0-1)}{N_0}} \end{bmatrix} \otimes \begin{bmatrix} 1 & & & & \\ & e^{\frac{-2\pi iq}{N_0}} & & & \\ & & e^{\frac{-4\pi iq}{N_0}} & & \\ & & & \ddots & \\ & & & & e^{\frac{-2\pi iq(N_0-1)}{N_0}} \end{bmatrix} \tag{2.79}$$

$$=: \sum_{q=0}^{N_0-1} m_q \begin{bmatrix} I_N & & & & \\ & E_1^q & & & \\ & & E_2^q & & \\ & & & \ddots & \\ & & & & E_{N_0-1}^q \end{bmatrix}, \tag{2.80}$$

where we have defined

$$E_k^q := \begin{bmatrix} e^{\frac{2\pi iqk}{N_0}} & & & & \\ & e^{\frac{2\pi iq(k-1)}{N_0}} & & & \\ & & e^{\frac{2\pi iq(k-2)}{N_0}} & & \\ & & & \ddots & \\ & & & & e^{\frac{2\pi iq(k-(N_0-1))}{N_0}} \end{bmatrix}. \tag{2.81}$$

For N_0 odd, we use the symmetry $m_q = m_{-q}$ to rewrite the summation (2.80) as

$$m_0 I_{N_0^2} + \sum_{q=1}^{\frac{N_0-1}{2}} m_q \begin{bmatrix} 2I_N & & & & \\ & E_1^q + E_1^{-q} & & & \\ & & E_2^q + E_2^{-q} & & \\ & & & \ddots & \\ & & & & E_{N_0-1}^q + E_{N_0-1}^{-q} \end{bmatrix}. \tag{2.82}$$

Conveniently,

$$E_k^q + E_k^{-q} = \begin{bmatrix} 2\cos\frac{2\pi iqk}{N_0} & & & & \\ & 2\cos\frac{2\pi iq(k-1)}{N_0} & & & \\ & & 2\cos\frac{2\pi iq(k-2)}{N_0} & & \\ & & & \ddots & \\ & & & & 2\cos\frac{2\pi iq(k-(N_0-1))}{N_0} \end{bmatrix}, \tag{2.83}$$

from which we infer

$$\lambda_{m+nN_0}\left(P_F^{-1}\left(M_0 \otimes I_{N_0}\right)P_F\right) = m_0 + 2\sum_{q=0}^{\frac{N_0-1}{2}} m_q \cos\frac{2\pi q\,(n-m)}{N_0} \tag{2.84}$$

is the eigenvalue for the (m,n)th spatial mode owing to diagonal connectivity, N_0 odd. If N_0 is even, we write (2.80) as

$$m_0 I_{N_0^2} + m_{\frac{N_0}{2}} \begin{bmatrix} I_N & & & \\ & E_1^{\frac{N_0}{2}} & & \\ & & \ddots & \\ & & & E_{N_0-1}^{\frac{N_0}{2}} \end{bmatrix} + \tag{2.85}$$

$$\sum_{q=1}^{\frac{N_0}{2}-1} m_q \begin{bmatrix} 2I_N & & & \\ & E_1^q + E_1^{-q} & & \\ & & \ddots & \\ & & & E_{N_0-1}^q + E_{N_0-1}^{-q} \end{bmatrix}, \tag{2.86}$$

(2.87)

and note that

$$E_k^{\frac{N_0}{2}} = \begin{bmatrix} (-1)^k & & & \\ & (-1)^{k-1} & & \\ & & \ddots & \\ & & & (-1)^{k-(N_0-1)} \end{bmatrix} \tag{2.88}$$

to write the (m,n)th eigenvalue as

$$\lambda_{m+nN_0}(M_F) = m_0 + m_{\frac{N_0}{2}}(-1)^{n-1} + 2\sum_{q=0}^{\frac{N_0}{2}-1} m_q \cos\frac{2\pi q\,(n-m)}{N_0}. \tag{2.89}$$

■

2.3.4 Spatial Modes on Planar Lattices

It is straightforward to generalize a frequency-based interpretation to cells arranged in periodic planar lattices, which are well described mathematically. Throughout the following discussion we will refer to coordinates in physical space as $\hat{e}$, the unit vector pointing "east," and $\hat{s}$, the unit vector pointing "south". This choice of vector orientations mimics the numbering scheme in an array, whereby indices increase horizontally left to right (with $\hat{e}$) and vertically top to bottom (with $\hat{s}$). We will assume a system of cells indexed in an $N_R \times N_C$ array with periodic boundary conditions such that M is diagonalized by $T_R \otimes T_C$ where both T_C and T_R are real or complex DFT bases of appropriate dimension.

Let cells in physical space be arranged in a planar lattice described by vectors a_R and a_C corresponding respectively to the rows and columns of the indexed array. Without loss of generality we orient a_R along $\hat{e}$ (such that $a_R \cdot \hat{e} = |a_R|$). We define the unit vectors $\hat{a}_R := \frac{a_R}{|a_R|} = \hat{e}$ and $\hat{a}_C := \frac{a_C}{|a_C|}$. Letting θ be the angle between a_R and a_C, we can write $a_R = |a_R|\hat{e}$ and $a_C = |a_C|\cos\theta\hat{e} + |a_C|\sin\theta\hat{s}$. Note that the eigenfunctions are periodic in n with period $N_C|a_R|$ along a_R and periodic in m with period $N_R|a_C|$ along a_C.

Observation 5. *For an $N_R \otimes N_C$ cellular lattice with lattice vectors a_R, a_C and periodic boundary conditions, the (m,n)th spatial mode corresponds to a plane wave of frequency*

$$\frac{n}{|a_R|N_C}\hat{a}_R + \frac{m}{|a_C|N_R}\hat{a}_C =: f_{a_R}\hat{a}_R + f_{a_C}\hat{a}_C = (f_{a_R} + f_{a_C}\cos\theta)\,\hat{e} + f_{a_C}\sin\theta\hat{s} \tag{2.90}$$

in physical space, with an "absolute" frequency of

$$f = \sqrt{|f_{a_R} + f_{a_C}\cos\theta|^2 + |f_{a_C}\sin\theta|^2} \tag{2.91}$$

pointing at an angle

$$\phi = \tan^{-1}\left(\frac{f_{a_C}\sin\theta}{f_{a_R} + f_{a_C}\cos\theta}\right) \tag{2.92}$$

from the $\hat{x}$ axis.

Remark. One may liken the translation from physical space into matrix space to "sampling" in space from an underlying pattern with "spatial sampling frequency" $\frac{1}{|a_R|}$ along a_R and $\frac{1}{|a_C|}$ along a_C. The translation into spatial modes, or the DFT, recovers normalized frequency components from the discrete samples. Maintaining a constant surface area but increasing the number of cells occupying that surface area (i.e., $N'_R = c_R N_R$, $N'_C = c_C N_C$, $a'_R = \frac{1}{C_R} a_R$, $a'_C = \frac{1}{c_C} a_C$) does not change the physical range of space over which the modes are described but does increase the "resolution" or "sampling rate" of the system by a factor of c_R along a_R and c_C along a_C, enabling the system to modify higher frequencies than before and therefore permitting finer filtering of a continuous-in-space input gradient.

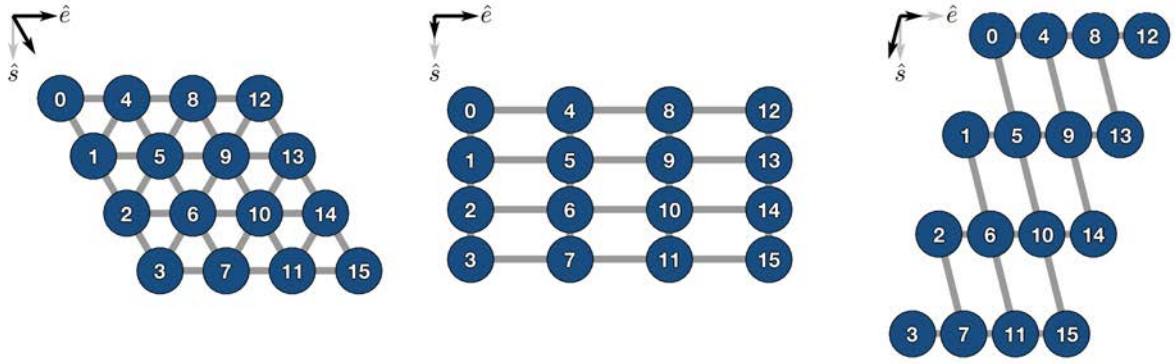

Figure 2.7: Examples of lattice configurations. Cells are uniformly spaced along the directions indicated by the black arrows. The gray arrow indicates a 90° angle from horizontal. Left, hexagonal ($\theta = 60°$) with nearest-neighbor (row/column/forward diagonal) interconnectivity; center, rectangular ($\theta = 90°$) with $|a_1| = 2|a_2|$ and row/column interconnectivity; right, an arbitrary lattice with $\theta = 105°$, $|a_2| = 2|a_1|$, and row/backward diagonal interconnectivity.

The hexagonal lattice is of particular importance in our later examples, as it is the tightest 2D packing arrangement for cells of fixed area, and is found in many natural systems such as the wing epithelial cells of *Drosophila* [318]. With the cells numbered as shown in Figure 2.7, the (m, n)th spatial mode corresponds to the mth mode horizontally and the nth mode on a line at a 60° angle from each row. Cells in the hexagonal lattice interact with each of their six nearest neighbors such that there are "diagonal interconnections" between rows of cells. Following notation in Observation 3, the diagonal interconnections are handled as follows: Define $C_D := min\,(N_R, N_C)$ and $R_D := max\,(N_R, N_C)$, and let $P \in \mathbb{R}^{C_D \times C_D}$ be the permutation matrix with lower diagonal ones and the last entry of the first column also one. Define $P_F := diag\left(P^0, P^1, ..., P^{R_D}\right)$. Then the interconnection matrix for a hexagonal lattice with periodic boundary conditions is given by

$$M := (M_R \otimes I_R) + (I_C \otimes M_C) + P_F^T \left(M_F \otimes I_{R_D}\right) P_F. \tag{2.93}$$

If we let $M_R = M_C = M_F = M_0 \in \mathbb{R}^{N_0 \times N_0}$ be the circulant diffusion matrix, then M has eigenvalues

$$[\Lambda]_{mn} = -6 + 2\left(\cos\frac{2\pi m}{N_0} + \cos\frac{2\pi n}{N_0} + \cos\frac{2\pi(m-n)}{N_0}\right). \tag{2.94}$$

From this we see that $\Lambda = \Lambda^T$ and $[\Lambda]_{mn} = [\Lambda]_{(N_0-m)(N_0-n)}$.

2.4 Minimal Model: Gene Expression with Autoregulation

The following example is a simple model of gene expression that is easy to solve analytically for the filter coefficients. We apply the filtering approach to the example, including an expansion of the matrix notation to emphasize the role of the filter coefficients as "weights" for the spatial modes. As this model focuses on biological and filtering concepts, intercellular interaction is described only in the most general terms, leaving exploration of the underlying mechanisms to later examples.

Here, we consider a simple model of an autoregulatory process in which each cell transcribes mRNA m that is translated into protein p that in turn modifies the production rate of m. The signaling molecule v, generated in exact proportion to p, also regulates p production in the self and neighbors by modulating the production rate of m. The system dynamics are

$$\begin{cases} \dot{m}_i = -\gamma_m m + \alpha_m f(v_i, u_i, p_i) \\ \dot{p}_i = -\gamma_p p + \alpha_p m \\ y = p \\ v = Mp \end{cases} \tag{2.95}$$

where γ_m, γ_p are the degradation (decay) rates of mRNA and protein respectively, and α_m, α_p are the corresponding transcription or translation rates. The function $f(v_i, u_i, p_i)$ captures the influence of the coupling, input, and protein on the production rate of the mRNA and therefore of the protein.

When linearized at steady state, the system becomes

$$\begin{cases} \dot{\tilde{m}}_i = -\gamma_m \tilde{m}_i + \alpha_m \left(F_v \tilde{v}_i + F_u \tilde{u}_i + F_p \tilde{p}_i \right) \\ \dot{\tilde{p}}_i = -\gamma_p \tilde{p}_i + \alpha_p \tilde{m}_i \\ \tilde{y} = \tilde{p} \\ \tilde{v} = M\tilde{p} \end{cases} \tag{2.96}$$

where $F_v := \frac{\partial f}{\partial v_i}\big|_{(\bar{m}^*, \bar{p}^*, \bar{v}^*, \bar{u}^*)}$, $F_u := \frac{\partial f}{\partial u_i}\big|_{(\bar{m}^*, \bar{p}^*, \bar{v}^*, \bar{u}^*)}$, and $F_p := \frac{\partial f}{\partial p_i}\big|_{(\bar{m}^*, \bar{p}^*, \bar{v}^*, \bar{u}^*)}$.

Define $\alpha := \alpha_m \alpha_p$ and $\gamma := \gamma_m \gamma_p$. Note that the steady-state protein concentration is a linear multiple of the steady-state mRNA concentration, such that mathematically a molecule produced at rate $\alpha f(v_i, u_i, p_i)$ and decayed at rate γ would have the same steady-state concentration as the protein in (2.95). Indeed, it is not uncommon for the dynamics of transcription and translation to be lumped together (usually by neglecting mRNA dynamics) in mathematical models such as those presented later in this chapter.

The steady-state solution to the perturbed system yields filter coefficients

$$[S]_{kk} = \frac{\frac{\alpha}{\gamma} F_u}{1 - \frac{\alpha}{\gamma}\left(F_p + F_v \lambda_k\left(M\right)\right)} \tag{2.97}$$

for $k = 0, 1, ..., N-1$. For the homogeneous steady state to be stable—and therefore for the filtering approach to be applicable—we require

$$\frac{\alpha}{\gamma}\left(F_v \lambda_k\left(M\right) + F_p\right) < 1. \tag{2.98}$$

We henceforth assume this condition is satisfied.

Recall that the spatial modes are given by t_k, the columns of the matrix T that diagonalizes M. The perturbing input can be written as

$$\tilde{u} = \left(u - \underline{u}\right) = T\hat{u} = \sum_{k=0}^{N-1} \hat{u}_k t_k. \tag{2.99}$$

The coefficients $\hat{u}_k$ (the entries of $\hat{u}$) are the weights assigned to each of the spatial modes t_k. The steady-state perturbed readout is given by

$$\tilde{y}^* = \sum_{k=0}^{N-1} [S]_{kk} \hat{u}_k t_k, \tag{2.100}$$

such that the readout in the ith cell is given by $\underline{\bar{y}} + \tilde{y}_i^*$.

The $\mathcal{H}_2$ norm for the kth spatial mode is analytically calculated to be $\frac{F_u}{2\alpha}[S]_{kk}$. This relationship indicates that the modes in the system respond identically to within a scaling factor to both persistent spatial disturbances and temporally varying white noise inputs.

Figure 2.8 exemplifies how the choice of interaction type and interconnectivity affects the filtering behavior of the system with no autoregulation. In particular, activation of neighbors tends to cause the system to amplify low spatial frequencies, while inhibition of neighbors introduces amplification at high spatial frequencies.

To investigate the effect of autoregulation, suppose we fix all parameters except F_p. As $F_p \to -\infty$ all filter coefficients approach 0. This attenuating behavior occurs because allowing a protein to effectively shut down its own production prevents the system from responding to signal.

For $F_p > 0$ (autoactivation), increasing F_p disproportionately increases the coefficients at spatial modes with low eigenvalues. For T corresponding to the DFT or DCT-2, the lower eigenvalues are associated with higher-frequency spatial modes. In the case of lateral inhibition ($F_v < 0$), the filter coefficients already amplify high-frequency spatial modes relative to intermediate ones (Figure 2.8), such that adding autoactivation enhances the filter's intrinsic highpass characteristics. Indeed, mechanisms involving lateral inhibition and autoactivation have been conjectured to increase the sharpness of boundary formation in systems of patterned cells responding to exponential input [319] [126].

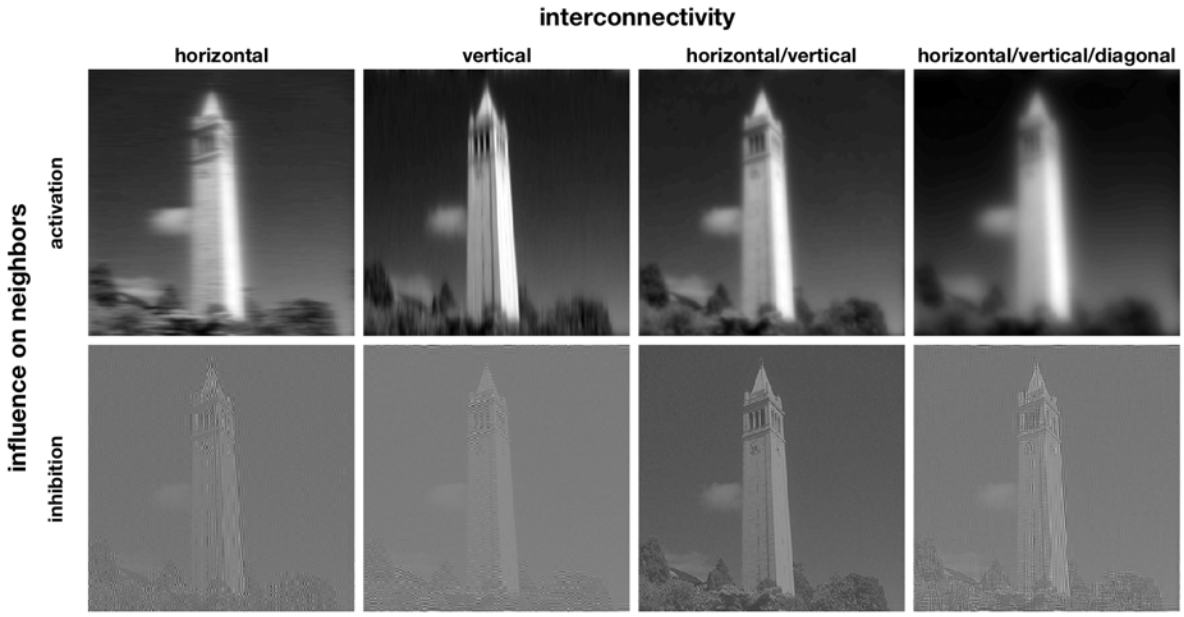

Figure 2.8: A minimal model of gene expression demonstrates that the same input prepattern produces different readouts depending on the interconnectivity and interaction type among cells. The ith cell has dynamics given by (2.95) for $\alpha = \gamma = 1$ with no autoregulation, i.e., $F_p = 0$. $F_v < 0$ corresponds to inhibition of neighbors while $F_v > 0$ implies activation of neighbors. The filter coefficients are thus $[S]_{kk} = (1 - F_v \lambda_k(M))^{-1}$. Pictured is the readout $\tilde{y}^*$ given the same perturbing input $\tilde{u}$ to $N = 62,500$ cells arranged in a 250×250 rectangular array, with one image pixel corresponding to each cell and the intensity of the pixel corresponding to the protein concentration. Interconnectivities vary by column; boundary conditions in all cases are periodic. Connection strengths are identical and assumed to incur no cost to the interacting cells (i.e., $[M]_{ii} = 0$). To best exemplify the effect of the interconnections, F_v was modified for each of the filtered images to give the highest magnitude of eigenvalues without destabilizing the underlying dynamical system. The readout in each cell is calculated according to (2.100). If cells activate their neighbors then the filter acts as a lowpass (attenuates short wavelengths) that blurs the underlying spatial input along the same dimension as the interconnections. Inhibition sharpens lines orthogonal to the interconnections by enhancing contrast parallel to the interconnections. The images are individually normalized.

2.5 Application: Fruit Fly Wing Veins

The Notch-Delta patterning mechanism is a lateral inhibition system that is responsible for diverse developmental phenomena including neural and epidermal fate determination in the fruit fly *Drosophila melanogaster*. Cells produce both Notch and Delta, which are proteins found in the cell membrane. Delta on the surface of one cell binds Notch on the surface of neighboring cells to inhibit those neighbors' Delta production, thereby relieving inhibition on the cell's own Delta production by decreasing the potential for the neighbors to bind its Notch. With the appropriate interaction strengths, such lateral inhibition will ultimately generate a checkerboard pattern in which cells expressing high Delta are adjacent to cells expressing low Delta. This has significant consequences for organismal development: Notch that is bound by Delta on an adjacent cell will cleave in two—preventing it from further signaling—and the portion left inside the cell will signal the cell to express target genes that influence the choice of cell identity [320]. A cell whose neighbors express more Delta is more likely to have bound Notch and therefore more likely to adopt a particular fate [321], [116].

Patterning in a Notch-Delta system may arise spontaneously [289], [115] or through modification of a prepattern. In the case of *Drosophila* wing development, the gene *veinless* is expressed in an exponential gradient decreasing in either direction from what will become the center of a vein. The level of *veinless* expression in a cell determines the Delta production rate at that cell. Notch activity occurs in two peaks, one on either side of the center, where further vein development is restricted to occur. One model of the Notch-Delta mechanism suggests that mutual inactivation (*cis*-inactivation), when Notch and Delta on the same cell inhibit each other's activity, enables sharper and more robust patterning than is achieved with lateral inhibition alone [126], [127]. The role of such *cis*-interactions is an ongoing area of research (e.g., [322], [323], [324]), and a particularly delicate one given the apparent context dependence of many Notch functions [320].

The authors of [126] considered a line of cells with periodic boundary conditions, corresponding to the interconnection matrix

$$M = \frac{1}{2}\begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}. \tag{2.101}$$

The diagonal entries are zero to reflect the fact that Notch (N) and Delta (D) interact via contact with neighbors rather than diffusion, while the factor of $\frac{1}{2}$ ensures that v_N is the average Notch from neighbors and v_D is the average Delta from neighbors. Because M is circulant, the spatial modes correspond to the eigenvectors of the DFT matrix.

We discretize the input gradient of Delta production rate $\beta_D(\cdot)$ into β_{D_i}, $i = 0, 1, ..., N-1$ and let $\bar{\beta}_D$ be the mean of the β_{D_i}. We then define $u_i := \beta_{D_i} - \bar{\beta}_D$, $x_i^T := [N_i, D_i, R_i]$, and $v_i := [v_{N_i}, v_{D_i}]$ where readout R is a reporter for Notch activity (i.e., is expressed from a target gene for Notch activity).

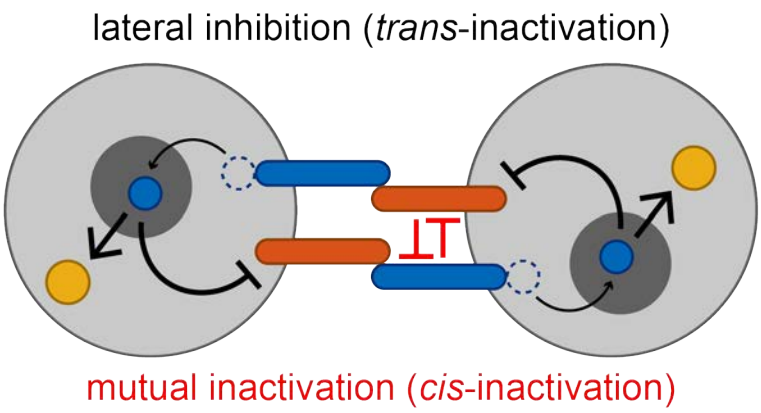

Figure 2.9: Schematic of lateral inhibition and mutual inactivation in the Notch-Delta system. In the lateral inhibition system, Notch (blue) and Delta (orange) are expressed on the surface of cells. Notch bound by Delta cleaves in two, and the domain within the cell travels to the nucleus to drive expression of target genes, such as reporter (yellow), and inhibit Delta production. Delta and Notch on the same cell may inhibit each other's activity, in a process termed "mutual" or "*cis*-inactivation" (see text).

The authors of [127] propose four models of the Notch-Delta patterning mechanism that involve mutual inactivation, lateral inhibition, or both. The linearization for these models is given in the following subsections.

Parameter	**Value**	**Description**	Source
α_N	10	"leakiness" of Notch expression (RFU/hr)	[127] Table S1 (Figure S4A)
β_{D0}	17.5	max. Delta production rate (RFU/hr)	[126] Table S3 (Figure 4C)
x_0	7	number of cell diameters	[126] Table S3 (Figure 4C)
β_{D_i}	$\beta_{D0}e^{-\lvert i\rvert/x_0}$	Delta production rate (RFU/hr) for cell i	[126] (Figure 4C)
$\bar{\beta}_D$	9.09	Delta production rate (RFU/hr) for linearization	$\frac{1}{N}\sum_{i=0}^{N-1}\beta_{D_i}$
β_N	10	Notch production rate (RFU/hr)	[126] Table S3 (Figure 4C)
β_R	150	reporter production rate (RFU/hr)	[126] Table S3 (Figure 4C)
γ	0.1	Notch, Delta decay rate (1/hr)	[126] Table S3 (Figure 4C)
γ_R	0.05	reporter decay rate (1/hr)	[126] Table S3 (Figure 4C)
k_c	0.25	inverse cis-interaction strength	[126] Table S3 (Figure 4C)
k_t	5	inverse trans-interaction strength	[126] Table S3 (Figure 4C)
n	2	Hill coefficient for Notch-Delta activation of reporter	-
m	2	Hill coefficient for reporter repression of Delta	-
k_{RS}	300,000	affinity of reporter induction	[127] Table S1 (Figure S4A)
k_{NS}	5×10^7	affinity of reporter induction	-

Table 2.1: Parameters used in the Notch-Delta model simulations, unless noted otherwise in the text (Figures 2.10 to 2.14).

2.5.1 Mutual Inactivation

The system equations for the MI model (Figure 2.9) are

$$\begin{cases} \dot{N}_i(t) = \beta_N - \gamma N_i(t) - \frac{N_i(t)v_{D_i}(t)}{k_t} - \frac{N_i(t)D_i(t)}{k_c} \\ \dot{D}_i(t) = \underline{\bar{\beta}}_D + u_i - \gamma D_i(t) - \frac{D_i(t)v_{N_i}(t)}{k_t} - \frac{N_i(t)D_i(t)}{k_c} \\ \dot{R}_i(t) = \beta_R \frac{\left(N_i(t)v_{D_i}(t)\right)^n}{k_{RS}+\left(N_i(t)v_{D_i}(t)\right)^n} - \gamma_R R_i(t) \\ y_i(t) = Cx_i(t) \\ w_i(t) = \begin{bmatrix} N_i(t) \\ D_i(t) \end{bmatrix} \\ v(t) = (M \otimes I_2)\, w(t) \end{cases} \tag{2.102}$$

where γ, γ_R are decay rates, k_t^{-1} is the rate at which Delta and Notch bind each other on neighboring cells, k_c^{-1} is the strength of mutual inactivation, and k_{RS}, n are parameters determining how strongly bound Notch promotes reporter expression. Note that mRNA is not explicitly incorporated into the model, such that the dynamics are effectively lumped into the production and degradation terms for the proteins.

Linearization about the steady state with all $u_i = 0$ yields

$$A = \begin{bmatrix} -\gamma - \frac{\bar{v}_D^*}{k_t} - \frac{\bar{D}^*}{k_c} & -\frac{\bar{N}^*}{k_c} & 0 \\ -\frac{\bar{D}^*}{k_c} & -\gamma - \frac{\bar{v}_N^*}{k_t} - \frac{\bar{N}^*}{k_c} & 0 \\ b_1 & 0 & -\gamma_R \end{bmatrix}, \tag{2.103}$$

$$B_v = \begin{bmatrix} 0 & -\frac{\bar{N}^*}{k_t} \\ -\frac{\bar{D}^*}{k_t} & 0 \\ 0 & b_2 \end{bmatrix}, \quad B_u = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \tag{2.104}$$

where we have defined

$$b_1 := \beta_R n k_{RS} \frac{\bar{v}_D^{*n} \bar{N}^{*n-1}}{\left(k_{RS} + \left(\bar{N}^* \bar{v}_D^*\right)^n\right)^2}, \; b_2 := \frac{\bar{N}^*}{\bar{v}_D^*} b_1 \tag{2.105}$$

and C is chosen depending on the readout.

For this system, we can show that the linearized dynamical system is stable for all nonnegative and thus biologically attainable parameter values, strengthening the argument that patterning may not require instability.

Observation 6. *The dynamical system* (2.102) *linearized about steady state for all* $u_i = 0$ *is stable when the parameter values are nonnegative.*

Derivation. We begin by explicitly calculating the steady-state values $\underline{\bar{N}}^*, \underline{\bar{D}}^*, \bar{\underline{v}}_N^*, \bar{\underline{v}}_D^*$ for $u = 0$. First we note that for our choice of M the homogeneous solution satisfies $\bar{\underline{v}}_N^* = \underline{\bar{N}}^*$ and $\bar{\underline{v}}_D^* = \underline{\bar{D}}^*$. After algebra, we find that $\underline{\bar{N}}^*$ is the positive root of a quadratic, $\underline{\bar{D}}^*$ is found in terms of $\underline{\bar{N}}^*$, and $\underline{\bar{R}}^*$ is expressed in terms of $\underline{\bar{N}}^*$ and $\underline{\bar{D}}^*$:

$$\begin{cases} -\frac{\gamma}{K}\underline{\bar{N}}^{*2} + \left(\frac{\beta_N}{K} - \gamma^2 - \frac{\bar{\beta}_D}{K}\right)\underline{\bar{N}}^* + \beta_N\gamma = 0 \\ \underline{\bar{D}}^* = \frac{\bar{\beta}_D}{\gamma + K\underline{\bar{N}}^*} \\ \underline{\bar{R}}^* = \frac{\beta_R}{\gamma_R}\frac{\left(\underline{\bar{N}}^*\underline{\bar{D}}^*\right)^n}{k_{RS} + \left(\underline{\bar{N}}^*\underline{\bar{D}}^*\right)^n} \end{cases} \tag{2.106}$$

where $K := \frac{k_c k_t}{k_c + k_t}$.

The filter coefficients $[S]_{kk}$ are given by $-C\left(A + \lambda_k(M)B_vG\right)^{-1}B_u$, $k = 0, 1, ..., N-1$. To find them we can exploit the structure of C and B_u. For the sake of demonstration we will take the readout to be the reporter protein such that $C = [0\ 0\ 1]$, although the procedure applies equally well to arbitrary choices of C.

First we notate

$$A + \lambda_k(M)B_vG = \begin{bmatrix} A_1 & 0 \\ [b_1,\ b_2\lambda_k(M)] & -\gamma_R \end{bmatrix} \tag{2.107}$$

and apply the matrix inversion lemma to obtain

$$\left(A + \lambda_k(M)B_vG\right)^{-1} = \begin{bmatrix} A_1^{-1} & 0 \\ \frac{1}{\gamma_R}[b_1,\ b_2\lambda_k(M)]A_1^{-1} & -\frac{1}{\gamma_R} \end{bmatrix}. \tag{2.108}$$

Observe that

$$A_1^{-1} = \frac{1}{\det A_1}\begin{bmatrix} -\gamma - \frac{\bar{\underline{N}}^*}{K} & \frac{\bar{\underline{N}}^*}{k_c} + \lambda_k(M_0)\frac{\bar{\underline{N}}^*}{k_t} \\ \frac{\bar{\underline{D}}^*}{k_c} + \lambda_k(M)\frac{\bar{\underline{D}}^*}{k_t} & -\gamma - \frac{\bar{\underline{D}}^*}{K} \end{bmatrix}. \tag{2.109}$$

Premultiplying (2.108) by C extracts the bottom row, while postmultiplying by B_u extracts the middle entry of that row, which is given by

$$\frac{1}{\gamma_R \det A_1}\left[b_1\left(\frac{\bar{\underline{N}}^*}{k_c} + \lambda_k(M)\frac{\bar{\underline{N}}^*}{k_t}\right) + b_2\lambda_k(M)\left(-\gamma - \frac{\bar{\underline{D}}^*}{K}\right)\right]. \tag{2.110}$$

Substituting b_1 and b_2, we simplify the expression to

$$[S]_{kk} = -\frac{\bar{\underline{N}}^* b_1}{\gamma_R k_c \det A_1}\left[1 - \lambda_k(M)\left(1 + \frac{\gamma k_c}{\bar{\underline{D}}^*}\right)\right] \tag{2.111}$$

where

$$\det A_1 = \left(-\gamma - \frac{\bar{\underline{N}}^*}{K}\right)\left(-\gamma - \frac{\bar{\underline{D}}^*}{K}\right) - \left(\frac{\bar{\underline{N}}^*}{k_c} + \lambda_k(M)\frac{\bar{\underline{N}}^*}{k_t}\right)\left(\frac{\bar{\underline{D}}^*}{k_c} + \lambda_k(M)\frac{\bar{\underline{D}}^*}{k_t}\right) \tag{2.112a}$$

$$= -\left[\frac{\bar{\underline{N}}^*\bar{\underline{D}}^*}{k_t^2}\lambda_k(M)^2 + 2\frac{\bar{\underline{N}}^*\bar{\underline{D}}^*}{k_t k_c}\lambda_k(M) + \left(\frac{\bar{\underline{N}}^*\bar{\underline{D}}^*}{k_c^2} - \gamma^2 - \frac{\gamma}{K}\left(\bar{\underline{N}}^* + \bar{\underline{D}}^*\right) - \frac{\bar{\underline{N}}^*\bar{\underline{D}}^*}{K^2}\right)\right]. \tag{2.112b}$$

The dynamical system corresponding to these filter coefficients is analytically stable for our chosen M when the parameters in 2.1 are positive. Since $A + \lambda_k(M_0) B_v G$ is a block triangular matrix, its eigenvalues are the eigenvalues of the diagonal blocks, i.e., $-\gamma_R$ along with the eigenvalues of A_1. Since $-\gamma_R$ is always negative, checking for stability amounts to checking the sign of the eigenvalues of A_1.

Using the fact that $\bar{v}_D^* = \underline{\bar{D}}^*$ and $\bar{v}_N^* = \underline{\bar{N}}^*$ for our choice of M yields

$$A_1 = \begin{bmatrix} -\gamma - \frac{\bar{D}^*}{K} & -\frac{\bar{N}^*}{k_c} - \lambda_k(M_0)\frac{\bar{N}^*}{k_t} \\ -\frac{\bar{D}^*}{k_c} - \lambda_k(M_0)\frac{\bar{D}^*}{k_t} & -\gamma - \frac{\bar{N}^*}{K} \end{bmatrix} \tag{2.113}$$

with K as and A_1 as defined earlier. The eigenvalues are given by the zeros of the characteristic polynomial, found by solving for s in

$$\left(-\gamma - \frac{\bar{D}^*}{K} - s\right)\left(-\gamma - \frac{\bar{N}^*}{K} - s\right) - \left(-\frac{\bar{N}^*}{k_c} - \lambda_k(M_0)\frac{\bar{N}^*}{k_t}\right)\left(-\frac{\bar{D}^*}{k_c} - \lambda_k(M_0)\frac{\bar{D}^*}{k_t}\right) = 0. \tag{2.114}$$

The first term multiplies out to

$$s^2 + \left[2\gamma + \frac{\bar{N}^* + \bar{D}^*}{K}\right]s + \left[\gamma^2 + \frac{\bar{N}^*\bar{D}^*}{K^2} + \frac{\gamma}{K}\left(\bar{N}^* + \bar{D}^*\right)\right] \tag{2.115}$$

and the second contributes the following terms, independent of s:

$$-\left(\frac{\bar{N}^*}{k_c} + \lambda_k(M_0)\frac{\bar{N}^*}{k_t}\right)\left(\frac{\bar{D}^*}{k_c} + \lambda_k(M_0)\frac{\bar{D}^*}{k_t}\right) = -\frac{\bar{N}^*\bar{D}^*}{k_c^2} - 2\lambda_k(M_0)\frac{\bar{N}^*\bar{D}^*}{k_c k_t} - \lambda_k(M_0)^2\frac{\bar{N}^*\bar{D}^*}{k_t^2}. \tag{2.116}$$

To be biologically attainable the parameters and steady-state values must all be positive, such that the quadratic in s has positive coefficients for the first- and second-order terms. If the roots are complex then assuming nonzero decay and nontrivial solutions, the real part is given by

$$\frac{-\left[2\gamma + \frac{\bar{N}^* + \bar{D}^*}{K}\right]}{2} < 0, \tag{2.117}$$

guaranteeing stability.

If the roots are real, then they will be negative if the zeroth-order term is positive. However, if the zeroth-order term is negative, then one root will be positive and the system will not be stable. Neglecting the expressions in γ, which by observation must be positive,

the contributions to the zeroth-order term are

$$\frac{\bar{\underline{N}}^* \bar{\underline{D}}^*}{K^2} - \frac{\bar{\underline{N}}^* \bar{\underline{D}}^*}{k_c^2} - 2\lambda_k(M_0)\frac{\bar{\underline{N}}^* \bar{\underline{D}}^*}{k_c k_t} - \lambda_k(M_0)^2 \frac{\bar{\underline{N}}^* \bar{\underline{D}}^*}{k_t^2} \tag{2.118}$$

$$= \frac{\bar{\underline{N}}^* \bar{\underline{D}}^* (k_c + k_t)^2 - k_t^2 \bar{\underline{N}}^* \bar{\underline{D}}^* - 2k_c k_t \lambda_k(M_0) \bar{\underline{N}}^* \bar{\underline{D}}^* - \lambda_k(M_0)^2 k_c^2 \bar{\underline{N}}^* \bar{\underline{D}}^*}{k_c^2 k_t^2} \tag{2.119}$$

$$= \frac{\bar{\underline{N}}^* \bar{\underline{D}}^* (k_c^2 + 2k_c k_t) - 2k_c k_t \lambda_k(M_0) \bar{\underline{N}}^* \bar{\underline{D}}^* - \lambda_k(M_0)^2 k_c^2 \bar{\underline{N}}^* \bar{\underline{D}}^*}{k_c^2 k_t^2} \tag{2.120}$$

$$= \frac{\bar{\underline{N}}^* \bar{\underline{D}}^*}{k_c^2 k_t^2} \left[k_c^2 \left(1 - \lambda_k(M_0)^2\right) + 2k_c k_t \left(1 - \lambda_k(M_0)\right) \right]. \tag{2.121}$$

By our choice of M_0, $\lambda_k(M_0) \in [-1, 1]$, therefore the bottom expression is minimized to 0 by $\lambda_k(M_0) = 1$. Since this expression contains all the possible negative contributions to the zeroth-order term, the overall zeroth-order term cannot be negative, and hence the roots of the characteristic polynomial must be negative, implying stability of the system with given M_0. ■

To examine the effects of identical (correlated) vs. separate (uncorrelated) white noise inputs to both Delta and Notch, we first modify 2.102 so that a single input appears in the equations for both $\dot{N}_i$ and $\dot{D}_i$ in the correlated case and two independent inputs appear in each of these equations for the uncorrelated case. Accordingly, we then calculate the $\mathcal{H}_2$ norm for

$$B_u^{corr} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \; B_u^{uncorr} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{2.122}$$

2.5.2 Lateral Inhibition with Mutual Inactivation (LIMI)

The system equations are the same as for the mutual inactivation model (2.102), except that now Delta production is repressed by reporter protein:

$$\begin{cases} \dot{N}_i(t) = \beta_N - \gamma N_i(t) - \frac{N_i(t) v_{D_i}(t)}{k_t} - \frac{N_i(t) D_i(t)}{k_c} \\ \dot{D}_i(t) = \left(\bar{\underline{\beta}}_D + u_i\right) \frac{1}{1 + R_i(t)^m} - \gamma D_i(t) - \frac{D_i(t) v_{N_i}(t)}{k_t} - \frac{N_i(t) D_i(t)}{k_c} \\ \dot{R}_i(t) = \beta_R \frac{\left(N_i(t) v_{D_i}(t)\right)^n}{k_{RS} + \left(N_i(t) v_{D_i}(t)\right)^n} - \gamma_R R_i(t) \\ y_i(t) = C x_i(t) \\ w_i(t) = \begin{bmatrix} N_i(t) \\ D_i(t) \end{bmatrix} \\ v(t) = (M \otimes I_2)\, w(t) \end{cases}. \tag{2.123}$$

When linearized at steady state, the relevant matrices are

$$A = \begin{bmatrix} -\gamma - \frac{\bar{v}_D^*}{k_t} - \frac{\bar{D}^*}{k_c} & -\frac{\bar{N}^*}{k_c} & 0 \\ -\frac{\bar{D}^*}{k_c} & -\gamma - \frac{\bar{v}_N^*}{k_t} - \frac{\bar{N}^*}{k_c} & -a \\ b_1 & 0 & -\gamma_R \end{bmatrix}, \tag{2.124a}$$

$$B_v = \begin{bmatrix} 0 & -\frac{\bar{N}^*}{k_t} \\ -\frac{\bar{D}^*}{k_t} & 0 \\ 0 & b_2 \end{bmatrix}, \quad B_u = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \tag{2.124b}$$

$$C = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{2.124c}$$

where b_1, b_2 are defined as before and

$$a := m \frac{\underline{\bar{R}}^{*m-1}}{\left(1 + \underline{\bar{R}}^{*m}\right)^2}. \tag{2.125}$$

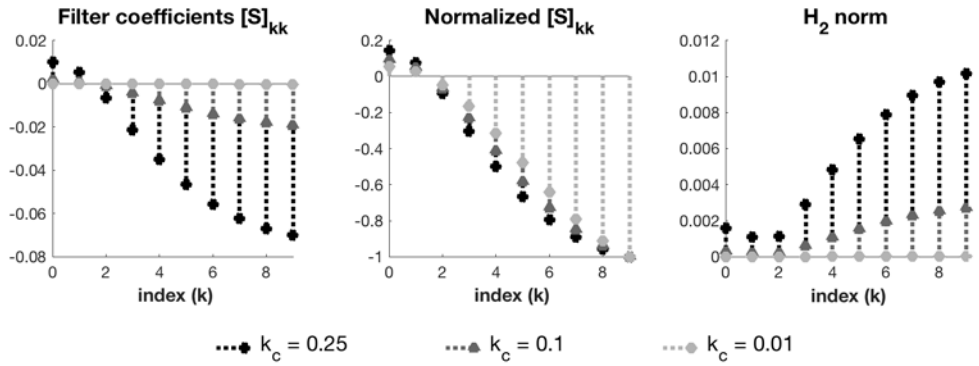

Figure 2.10: The filter coefficients and $\mathcal{H}_2$ norm for the MI model of Notch-Delta interaction reveal how changes to parameter values enhance high frequencies from an input gradient of Delta production to readout Notch activity. The magnitude of the coefficients decreases with greater mutual inactivation strength (lower k_c), indicating that greater inhibition reduces overall activity. The spatial modes correspond to the DFT basis and are indexed by k such that the kth mode has frequency $\frac{2\pi k}{N}$. The coefficients exhibit mirror-image symmetry about $k = \frac{N}{2}$; we plot only the first half of the coefficients to better visualize the filter's characteristic highpass shape. The greater the mutual inactivation, the greater the amplification of high frequencies relative to lower ones, as revealed by a plot of the coefficients individually normalized to the maximum in each set. The $\mathcal{H}_2$ norm is qualitatively similar to the highpass filter characteristic though smaller in magnitude, with more dramatic relative differences between values of k_c. Parameters are given in Table 2.1.

2.5.3 Simplest Lateral Inhibition by Mutual Inactivation (SLIMI)

The system equations are

$$\begin{cases} \dot{N}_i(t) = \alpha_N + \beta_N \frac{\left(N_i(t) v_{D_i}(t)\right)^n}{k_{NS} + \left(N_i(t) v_{D_i}(t)\right)^n} - \gamma N_i(t) - \frac{N_i(t) v_{D_i}(t)}{k_t} - \frac{N_i(t) D_i(t)}{k_c} \\ \dot{D}_i(t) = \underline{\bar{\beta}}_D + u_i - \gamma D_i - \frac{D_i(t) v_{N_i}(t)}{k_t} - \frac{N_i(t) D_i(t)}{k_c} \\ y_i(t) = C x_i(t) \\ w_i(t) = \begin{bmatrix} N_i(t) \\ D_i(t) \end{bmatrix} \end{cases} . \tag{2.126}$$

When linearized at steady state, the relevant matrices are

$$A = \begin{bmatrix} b_1 - \gamma - \frac{\bar{v}_D^*}{k_t} - \frac{\bar{D}^*}{k_c} & -\frac{\bar{N}^*}{k_c} \\ -\frac{\bar{D}^*}{k_c} & -\gamma - \frac{\bar{v}_N^*}{k_t} - \frac{\bar{N}^*}{k_c} \end{bmatrix}, \tag{2.127a}$$

$$B_v = \begin{bmatrix} 0 & b_2 - \frac{\bar{N}^*}{k_t} \\ -\frac{\bar{D}^*}{k_t} & 0 \end{bmatrix}, \quad B_u = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{2.127b}$$

$$G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{2.127c}$$

where now

$$b_1 := \beta_N n k_{NS} \frac{\bar{v}_D^{*n} \bar{N}^{*n-1}}{\left(k_{NS} + \left(\bar{N}^* \bar{v}_D^*\right)^n\right)^2}, \; b_2 := \frac{\bar{N}^*}{\bar{v}_D^*} b_1. \tag{2.128}$$

2.5.4 Comparison of Models

The MI, LIMI, and SLIMI models from reference [127] produce substantially similar readouts (Figure 2.11), filter characteristics, and $\mathcal{H}_2$ norms (Figures 2.11 and 2.12). Mutual inactivation (lower k_c) decreases the magnitude of the coefficients and therefore the final Notch and Delta concentrations, but exaggerates the intrinsic highpass characteristics of the filter, producing the sharper peaks in Notch activity predicted by [126].

Complementing the observation that mutual inactivation appears to enhance robustness to correlated gene expression perturbations [127], our analysis of the $\mathcal{H}_2$ norm reveals that regardless of readout, noise that is completely correlated between the production rates is almost uniformly rejected relative to uncorrelated noise. Furthermore, noise that is completely uncorrelated between Delta and Notch production rates is favored by the same frequencies as the system filter (Figures 2.10 and 2.12). Together, these observations suggest that time-varying stochastic inputs—unless they are of exceptionally large magnitude—do little to combat the intrinsic behavior of the filter, contributing to the robustness of the developmental program.

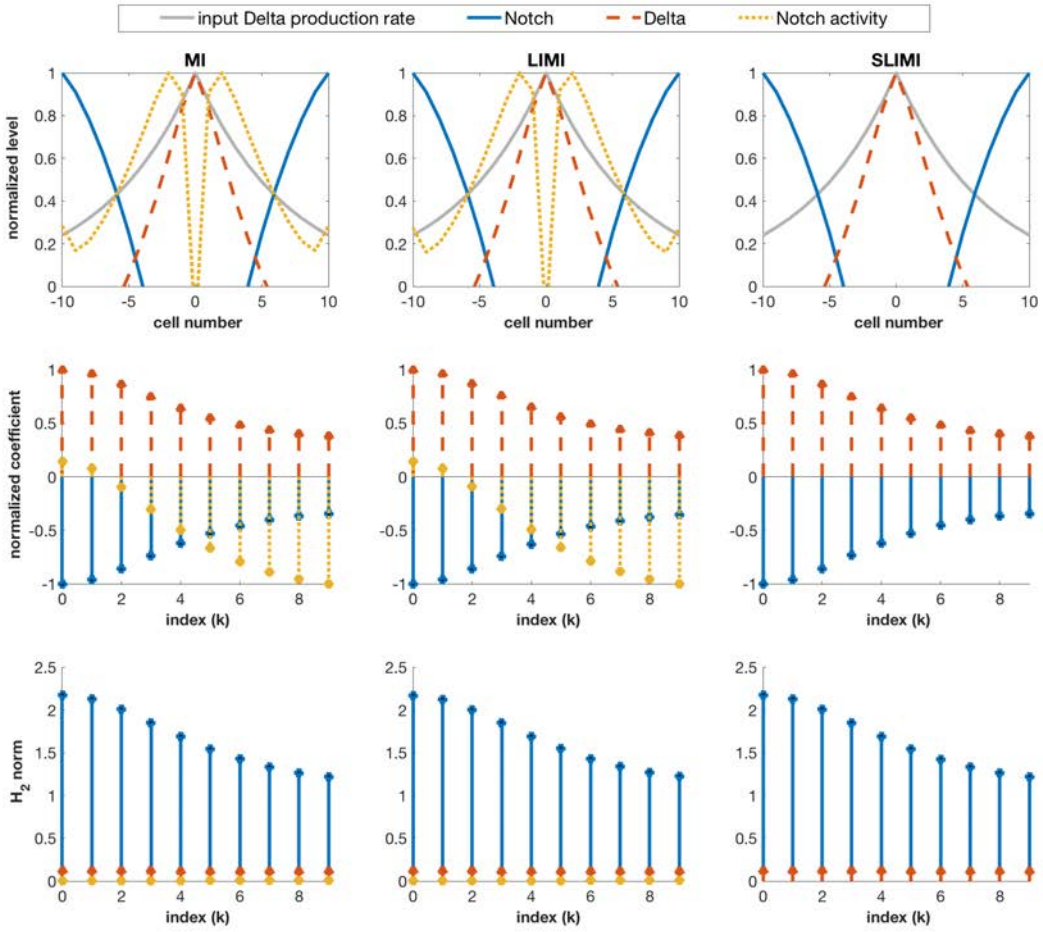

Figure 2.11: A linearized system qualitatively reproduces the double peaks in Notch activity predicted from full nonlinear simulations. The three different models for the Notch-Delta interaction produce qualitatively similar filter characteristics. Top row, a two-sided exponential input gradient of Delta production rate (solid light gray) results in two sharp bands of Notch activity (dotted yellow) that spatially segregates steady-state levels of Notch (solid blue) and Delta (dashed orange). Curves are normalized to their respective maxima. Note that the SLIMI model lacks a reporter protein and so does not have an output measure for Notch activity. These plots correspond to Figure 4C in [126]. Middle row, the magnitude of the filter coefficients for each possible output. Because the spatial modes correspond to the DFT basis, the coefficients exhibit mirror-image symmetry about $k = \frac{N}{2}$; we plot only the first half of the coefficients to better visualize the filter's characteristic highpass shape for output Notch activity, and lowpass shape for Delta and Notch (with a $\frac{pi}{2}$ phase shift in Notch expression). Each set of coefficients has been individually normalized to the maximum in each set. Bottom row, the $\mathcal{H}_2$ norm is qualitatively similar to the filter characteristic for the corresponding output. The coefficients here are not normalized in order to better visualize the large gain in Notch expression relative to Delta or activity levels. Parameters for all models are given in Table 2.1.

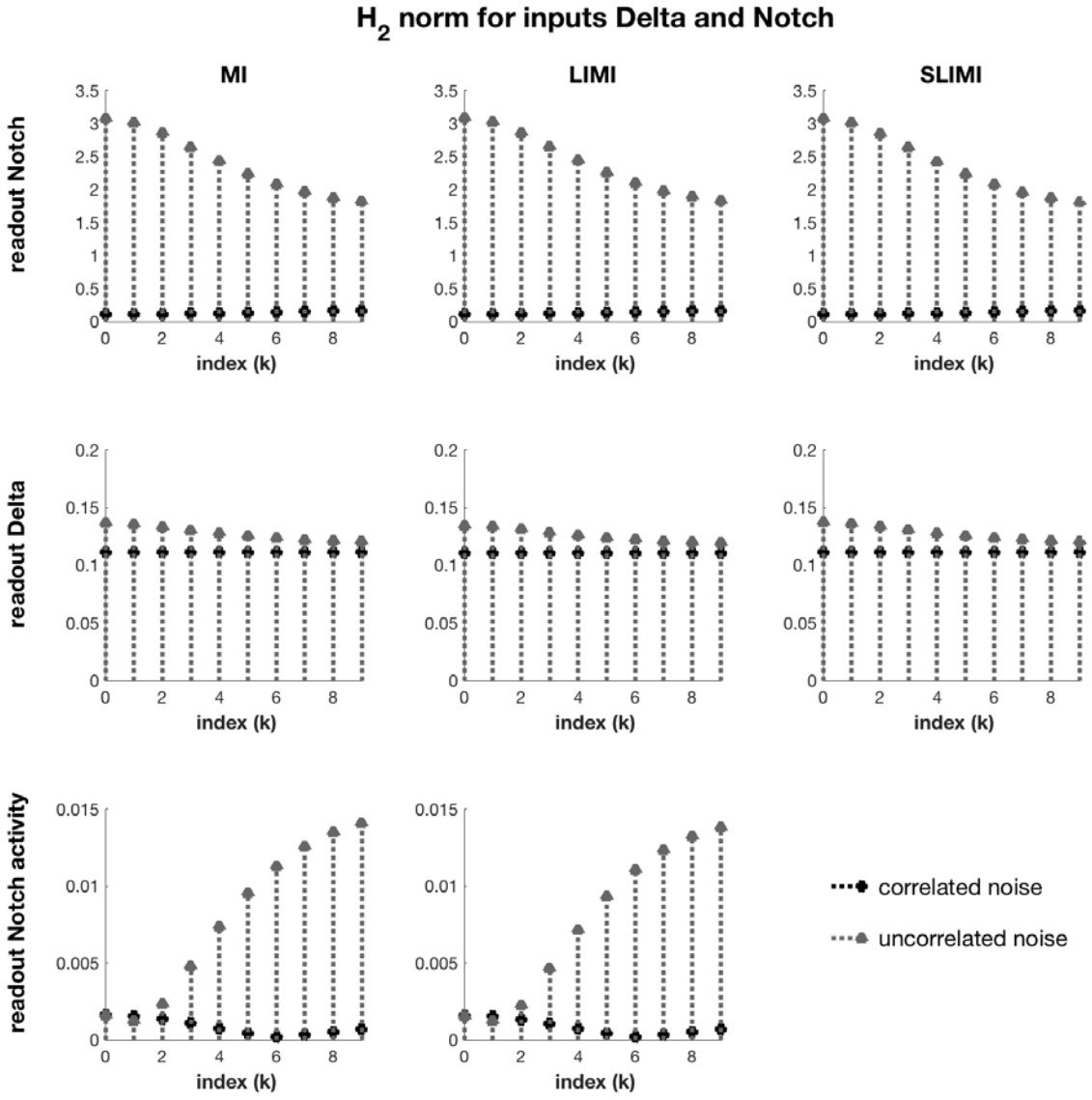

Figure 2.12: Filter behavior is robust with respect to noisy inputs to Notch or Delta, and system responses to correlated vs. uncorrelated noise is similar across models. Noise that is completely correlated between Notch and Delta is more strongly and uniformly rejected than completely uncorrelated noise. However, uncorrelated noisy inputs tend to emphasize the inherent highpass characteristics with respect to output Notch activity, suggesting that moderate levels of white noise do not compromise filter behavior. The most notable behavioral difference between models is that the LIMI model rejects uncorrelated noise slightly more strongly for readout Delta. Parameters for all models are given in Table 2.1.

While these three models exhibit near-identical filter coefficients for the model parameters in [127], increasing the mean input Delta production rate $\bar{\beta}_D$ causes the MI and SLIMI systems to more strongly attenuate high spatial frequencies in Notch and Delta concentration, without much changing the filtering behavior of the LIMI system (Figure 2.14). This observation implies the LIMI architecture is less sensitive to variations in absolute input relative to architectures relying upon mutual inactivation alone, which suggests that lateral inhibition and mutual inactivation may serve as complementary mechanisms to enhance patterning robustness: mutual inactivation to ensure sharpness of boundary formation [127], and lateral inhibition to buffer sensitivity to variation in overall Delta production rate or to the steepness of the *veinless* gradient. The analysis also identifies a set of experiments to distinguish which of the models is most reflective of the true biological mechanism. Namely, one could introduce perturbations that increase the mean Delta production rate—e.g., increase *veinless* expression or decrease diffusivity of the product—and compare the measured readouts to the model predictions.

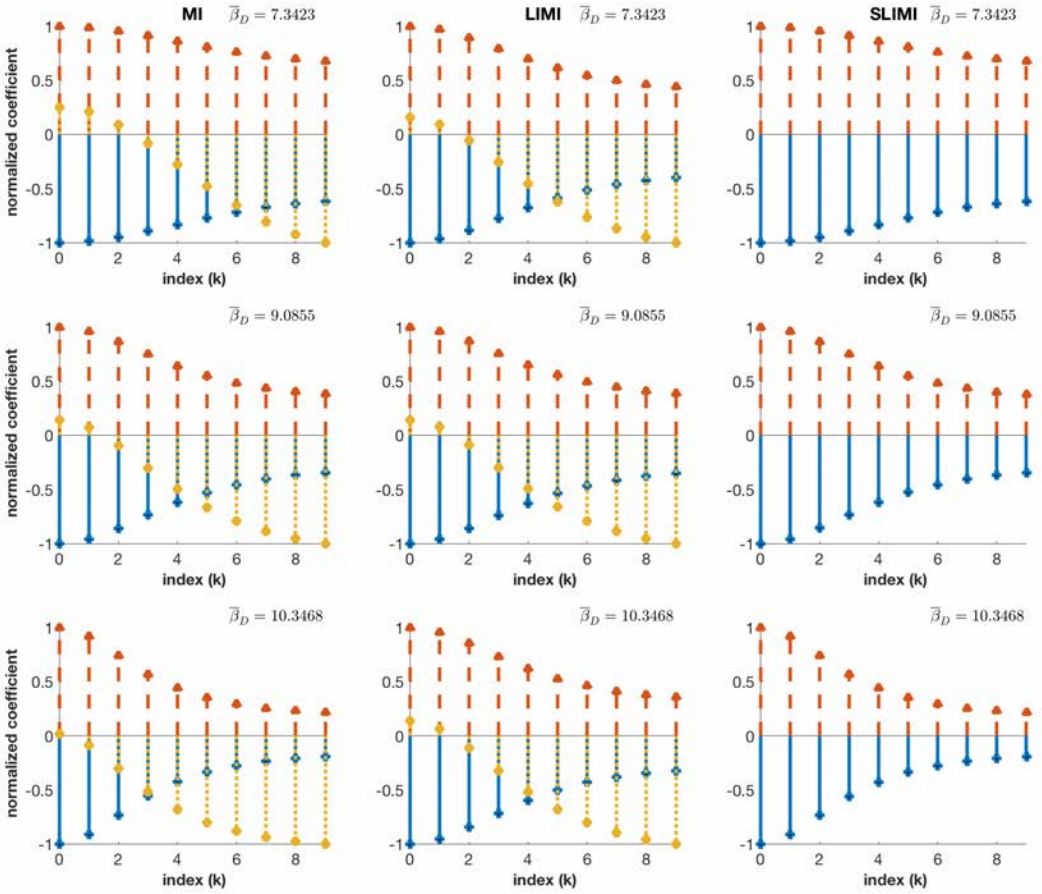

Figure 2.13: Filter coefficients depend on the mean value of the prepattern ($\bar{u}$), but some system architectures are more robust to changes in this mean value. Here, we show the filter coefficients for the MI, LIMI, and SLIMI models with varying choices of $\bar{\beta}_D$ used for the linearization (corresponding to $x_0 = 5, 7, 9$ respectively). The MI and SLIMI architectures show increasing attenuation of high frequencies with respect to output Notch and Delta. The MI model additionally shows alterations to how low frequencies are amplified in Notch activity levels. The LIMI model, in comparison, appears to be relatively buffered against such changes. See also Figure 2.14.

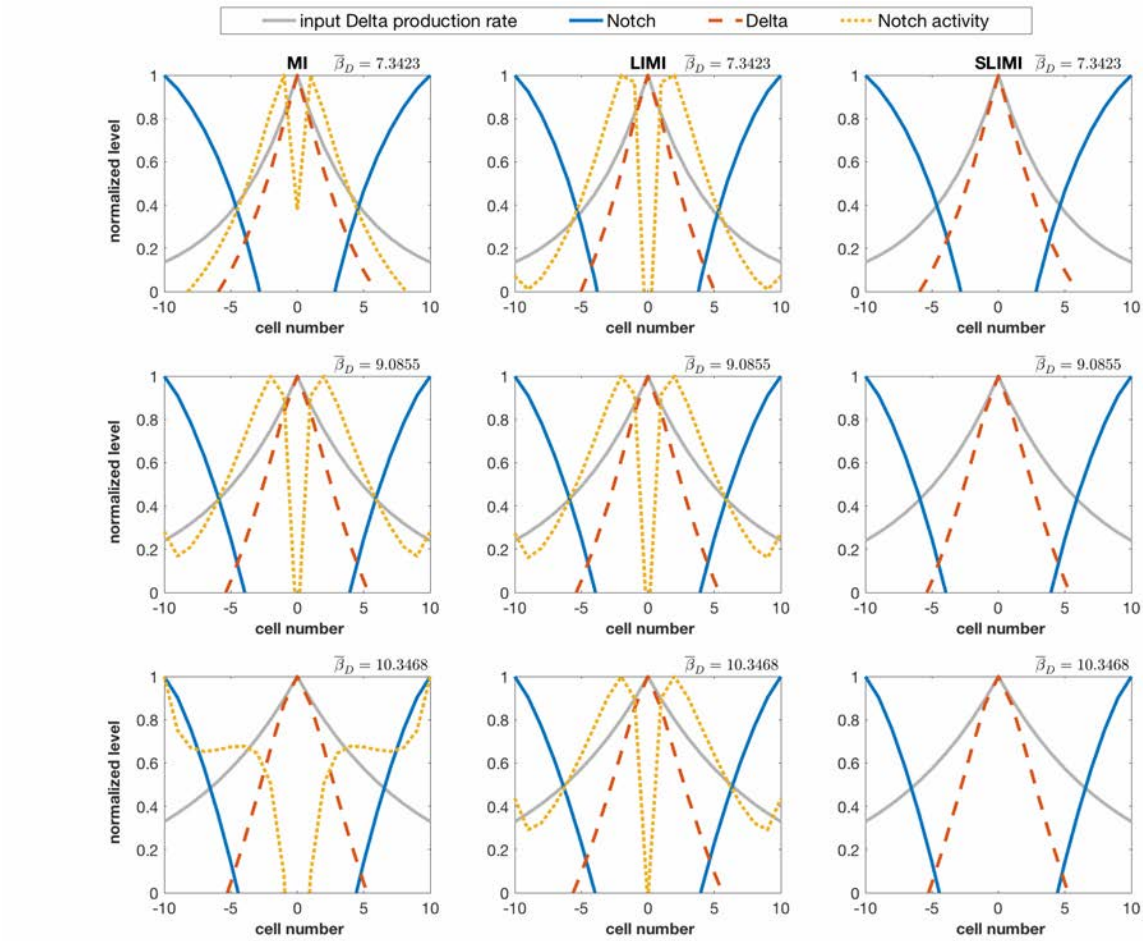

Figure 2.14: The LIMI architecture, while more complicated than MI or SLIMI, exhibits less sensitivity to changes in gradient steepness than the other models. Gradients correspond to $x_0 = 5, 7, 9$ from top to bottom, with corresponding $\bar{\beta}_D$ values identified on the upper left. Differences in readout owe chiefly to the differences in the filter coefficients with changes to mean input value, as seen in Figure 2.13.

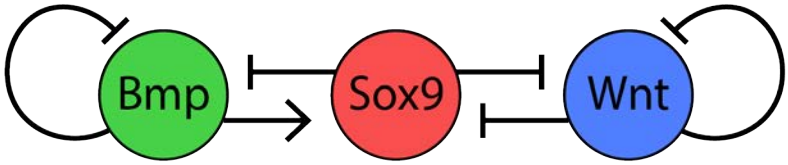

Figure 2.15: Schematic of the proposed Turing architecture for early gene expression patterning during digit formation in vertebrate limbs (see text).

2.6 Application: Digit Formation

Digits in developing vertebrate embryos originate from a flat paddle-shaped layer of cells that form the limb bud. A crucial step in digit patterning involves specifying which cells in the paddle will become digits and which will die to create the space between digits [325], [326]. This periodic pattern of digit with interdigit has been proposed to originate with spatially periodic expression of the gene *sox9*, which produces a protein that regulates transcription of the genes *wnt* and *bmp*. In turn, these genes code transcription factors Wnt and Bmp that regulate Sox9 production [327].

Cell cultures from developing embryos grown on plates show Turing-like patterns where Sox9 is out of phase with Wnt and Bmp. Turing patterns typically arise in chemical reaction systems with at least two types of diffusible molecules produced at every point in space, where the activation/inhibition relationship between the types is such that the homogeneous solution to the resulting dynamical system is unstable owing to the difference in diffusion rates between the two molecules. Such a reaction-diffusion model has been proposed to generate the observed Sox9/Wnt/Bmp pattern from stochastic initial conditions within a particular parameter range [327]. Our discretization of the model suggests that such a pattern might be observed even if the parameters do not satisfy the conditions for diffusion-driven instability.

Consider the Sox9/Bmp/Wnt network with distance l between cell centers (Figure 2.15). Let s, b, and z represent the concentrations of Sox9, Bmp, and Wnt respectively, such that $x_i = [s_i,\ b_i,\ z_i]$ and $v = [v_b,\ v_z]$. Let the input be random cell-to-cell variation in background protein production rate, i.e., the production rate of protein in the absence of promotion or inhibition, as from cell-to-cell variability in transcription or translation rates (see Figure 1.1). The dynamics within cell i and the coupling are given by

$$\begin{cases} \dot{s}_i(t) = \alpha_s + u_i + k_2 b_i(t) - k_3 z_i(t) - (s_i(t) - s_0)^3 \\ \dot{b}(t) = \alpha_{bmp} + u_i - k_4 s_i(t) - k_5 b_i(t) + \frac{d_b}{l^2} v_{b_i}(t) \\ \dot{z}_i(t) = \alpha_{wnt} + u_i - k_7 s_i(t) - k_9 z_i(t) + \frac{d_z}{l^2} v_{z_i}(t) \\ w_i(t) = \begin{bmatrix} b_i(t) \\ z_i(t) \end{bmatrix} \\ v(t) = (M \otimes I_2) w(t) \end{cases} \tag{2.129}$$

Parameter	Value	Description
α_{sox9}	0	constitutive Sox9 production rate
α_{bmp}	16.9	constitutive Bmp production rate
α_{wnt}	13.7	constitutive Wnt production rate
k_2	1	Bmp promotion of *sox9* expression
k_3	1	Wnt repression of *sox9* expression
k_4	1.59	Sox9 repression of *bmp* expression
k_5	0.1	Bmp decay rate
k_7	1.27	Sox9 repression of *wnt* expression
k_9	0.1	Wnt decay rate
d_b	2.5	diffusion coefficient for Bmp
d_w	1	diffusion coefficient for Wnt

Table 2.2: Parameters used in the simulations of digit formation, unless noted otherwise in the text (Figures 2.18, 2.16, 2.17). Values are from Table ST4 and text of [327].

where α are background production rates, k are interaction rates, and d are diffusivities.

Linearization about the homogeneous steady state yields

$$A = \begin{bmatrix} -3(\underline{s}^* - s_0)^2 & k_2 & -k_3 \\ -k_4 & -k_5 & 0 \\ -k_7 & 0 & -k_9 \end{bmatrix}, \tag{2.130}$$

$$B_v = \begin{bmatrix} 0 & 0 \\ \frac{d_b}{l^2} & 0 \\ 0 & \frac{d_z}{l^2} \end{bmatrix}, \quad B_u = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad G = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.131}$$

In this example, we consider a 2D rectangular array of evenly spaced cells interacting only with their immediate neighbors (i.e., no diagonal interconnections). For the remainder of this example, we will assume Neumann boundary conditions such that the spatial modes for the rows and columns of M correspond to the DCT-2 (Figure 2.6).

2.6.1 Analysis

We pick C to monitor Sox9, Bmp, or Wnt concentration and choose s_0 such that the Turing instability conditions are not satisfied, i.e., the eigenvalues of $(I_N \otimes A) + M \otimes (B_v G)$ are all negative. Nevertheless, the readout still replicates the spatially periodic patterns predicted by [327] for a range of intercellular distances (Figure 2.18) owing to the bandpass behavior of the filter (Figure 2.16). [Sox9] is out of phase with both [Bmp] and [Wnt], as indicated by the opposing signs of the coefficients in the passband.

The $\mathcal{H}_2$ norm measurements for the readouts qualitatively emphasize the same frequency bands as their respective filters $[S]_{kk}$ (Figure 2.17). For correlated or uncorrelated noise sources, readouts [Bmp] and [Wnt] experience much greater magnification than does [Sox9],

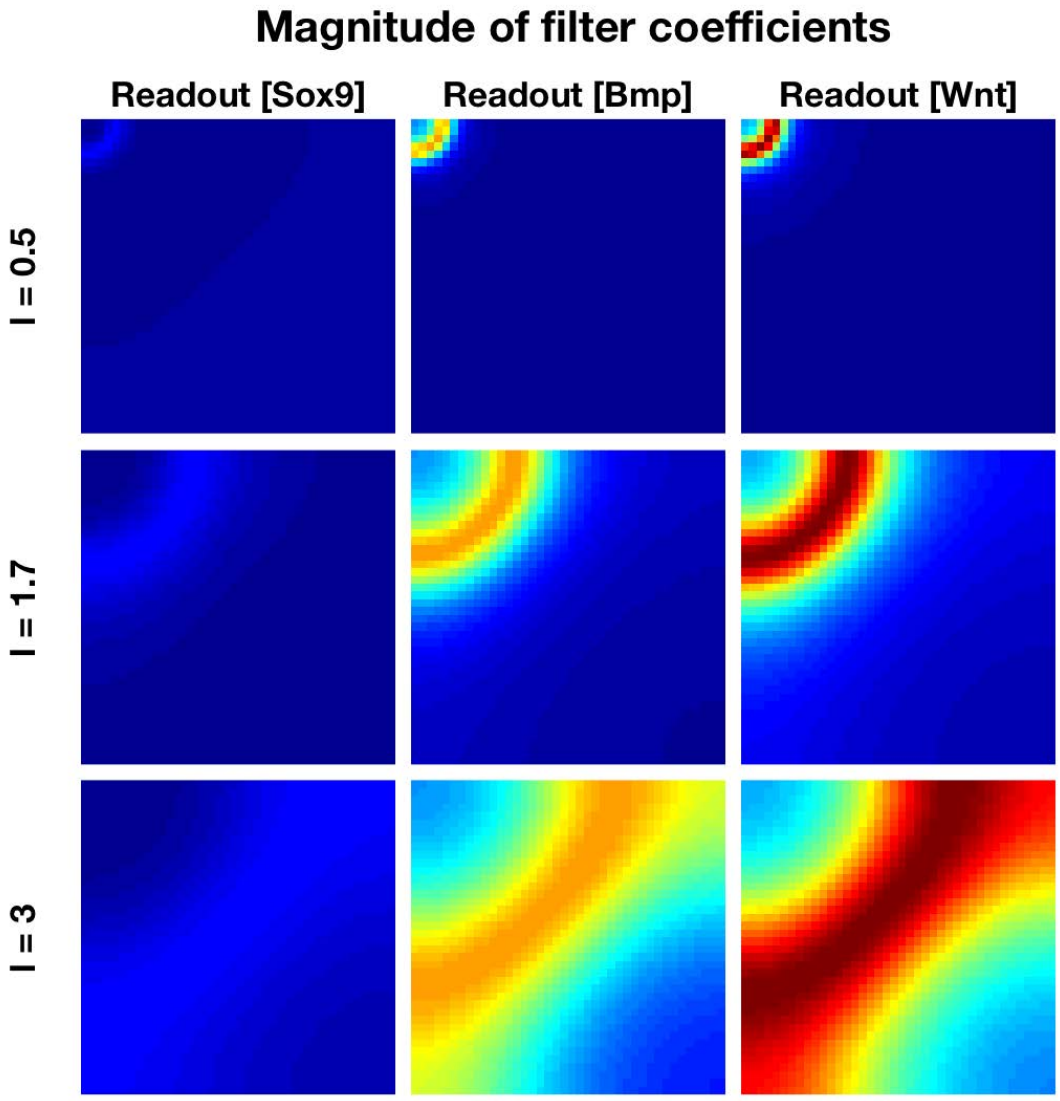

Figure 2.16: The stable, linearized reaction-diffusion system behaves as a bandpass filter for Sox9 (left), Bmp (center), and Wnt (right), resulting in a spatially periodic output (see Figure 2.18). Pictured are heat maps of the magnitude of the filter coefficients for the three readouts assuming Neumann boundary conditions (DCT-2 basis) in both dimensions, such that vertical frequency increases down rows (higher m) and horizontal frequency increases across columns (higher n). Increasing the distance between cells (l) increases the frequency of the passband but decreases the sharpness of the dropoff. The readout concentration of Sox9 is out of phase from the Bmp and Wnt concentrations due to the fact that the coefficients of S have an extra multiplicative factor of $-1 = e^{i\pi}$, or a phase shift of π, relative to the coefficients when the readout is [Bmp] or [Wnt]. Parameters are as given in Table ST4 of [327] with $s_0 = 11$ instead of 10, i.e., $s_0 \neq \underline{\bar{s}}^*$ and therefore $[A]_{00} \neq 0$ (see also Table 2.2). This choice of s_0 stabilizes the dynamical system with diffusion, thereby violating Turing conditions. Here, $N_R = N_C = 40$ for a total of $N = 1600$ cells. Images are normalized to the same scale (min. 0, max. 26.6).

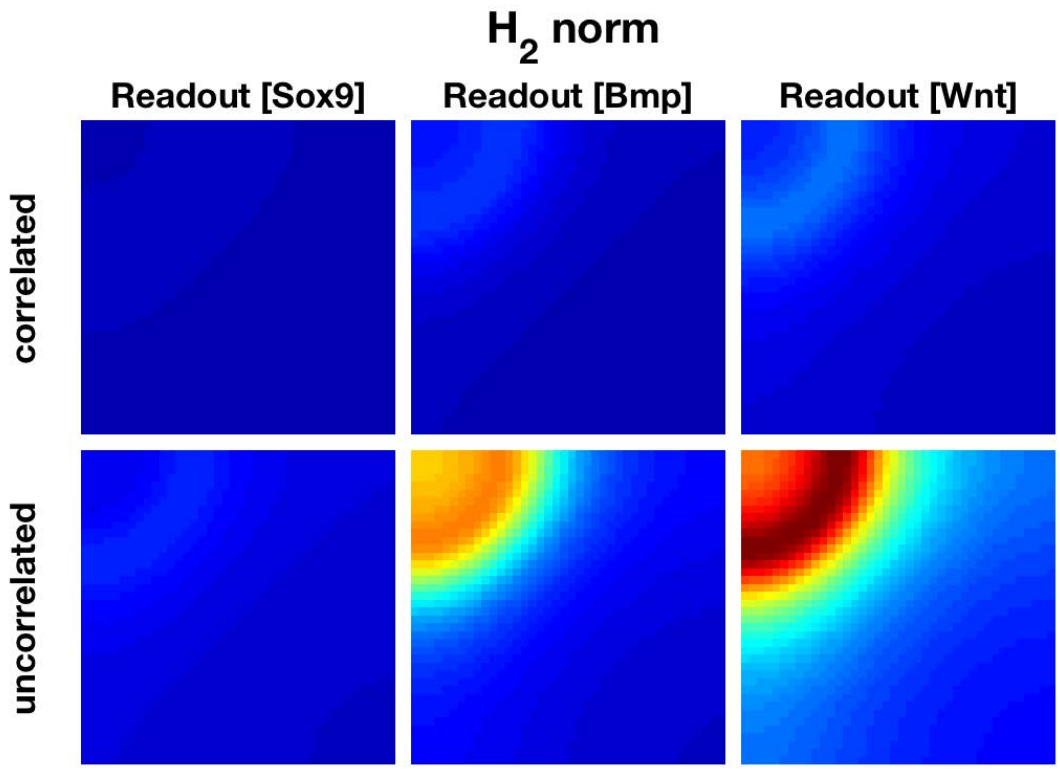

Figure 2.17: The stabilized Sox9-Bmp-Wnt network emphasizes noise in the same frequency bands as those favored by the filter. The $\mathcal{H}_2$ norm for correlated noise when the readout is [Sox9] is less than 1 in magnitude, indicating noise rejection at all frequencies, while uncorrelated noise is amplified at all frequencies for readout [Wnt]. Uncorrelated noise, despite being highly amplified for readouts [Bmp] and [Wnt], is rejected at frequencies higher than the upper end of the filter passband for [Sox9] and only weakly amplified at lower frequencies, perhaps as a result of the opposing influences of Bmp and Wnt on *sox9* expression. Parameters are as in 2.16 with $l = 1.7$ (see also Table 2.2). Images are normalized to the same scale (min. 0, max. 111).

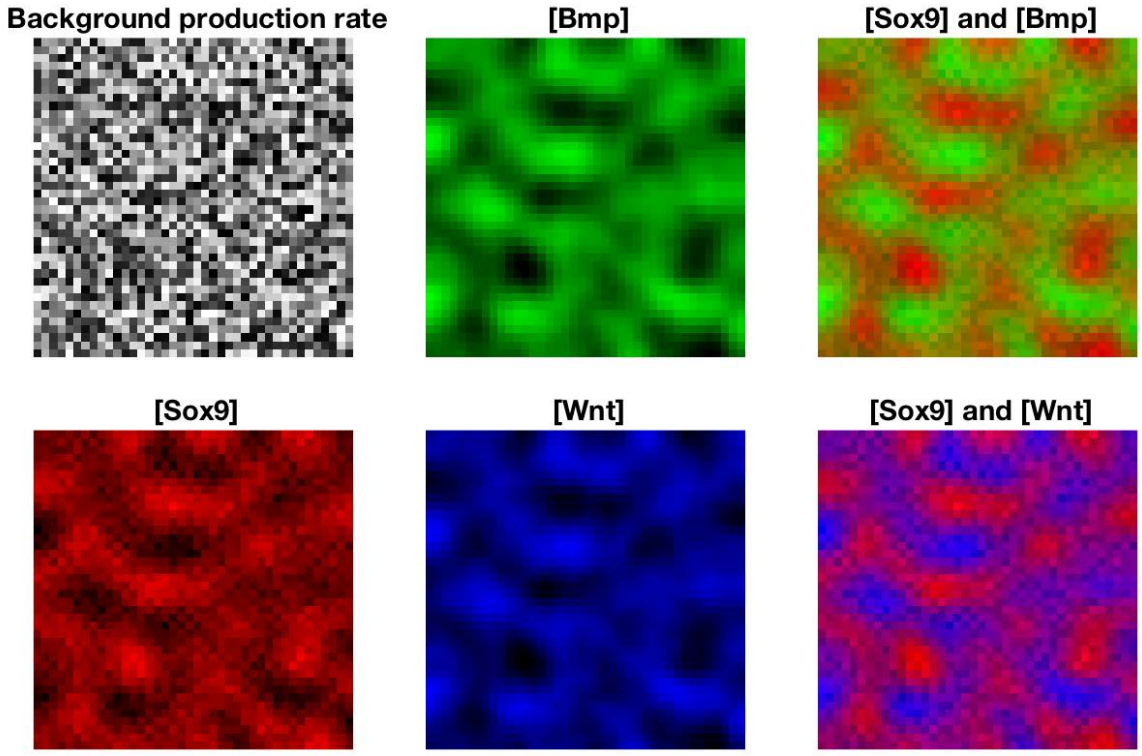

Figure 2.18: A linearized, discrete 2D system with random but constant-in-time variation between points replicates predictions from a Turing reaction-diffusion model even when the parameters do not satisfy the conditions for Turing-driven instability. The color-coded visualization reproduces images of micromass cultures from Supplementary Figure S3 in [327], showing similar periodic striped patterns. Higher intensity corresponds to higher concentration level and intensity is normalized individually by protein species, even in the overlaid images. The expression boundaries depicted here are not as sharp as the original Turing model due to the stability of the linearized system. The concentration of Sox9 is out of phase with Bmp and Wnt concentrations, as seen from the overlaid images. Parameters are as in 2.16 and 2.17 with $l = 1.7$ (see also Table 2.2), with a constant-in-time input background production rate input that is shared by all reactants (as in (2.129)). For outputs [Sox9], the filter coefficient of greatest magnitude occurs at $(14, 6)$ (and symmetrically also $(6, 14)$; see Figure 2.16), corresponding to a spatial mode comprising a sum of two cosines, the higher of which has period 25% the length of one side of the array. As predicted therefrom, the output pattern has approximately four complete periods at an angle about 67° from horizontal.

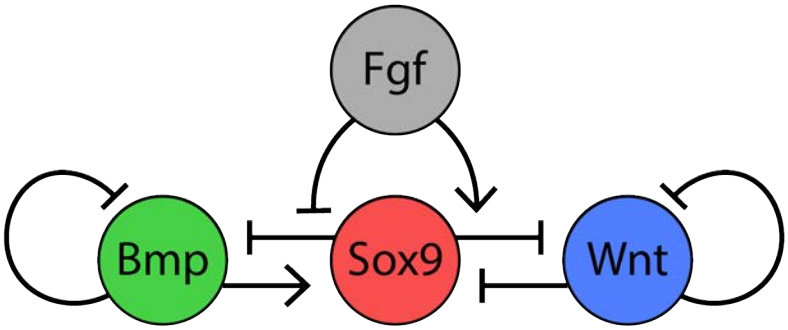

Figure 2.19: Schematic of the proposed Turing architecture for early gene expression patterning during digit formation in catfish shark fins (see text).

suggesting that the opposing effects of Bmp and Wnt on *sox9* expression may mostly cancel each other out at the level of Sox9 concentration. Relative noise amplification in the same modes favored by the $[S]_{kk}$ may ensure that stochastic influences do not counteract filter behavior, at the same time that attenuation and evenness in the response to other modes might reduce the relative influence of temporally varying inputs on the readout. The latter especially may be useful to maintain consistent behavior in a process such as digit formation that takes place over a long timespan.

While our simulations do not refute the hypothesis that a diffusion-driven instability constitutes the biological basis for digit formation, the fact that we can produce a similar pattern with an externally perturbed stable system suggests that not all apparent Turing patterns need arise from an instability. This observation could significantly ease the search for molecules and proteins that contribute to "spontaneous" stripe and spot patterning, as the parameter restrictions required for true Turing instabilities may not be biologically plausible.

2.6.2 Digit Formation with a Morphogen Gradient

Expanding on the work of [327], reference [328] demonstrated that changes to the parameters in the proposed Turing network for digit formation in mice can produce *sox9* expression patterns matching those found in embryonic catshark fins, suggesting that the mechanism has been evolutionarily conserved. The authors augmented the model with an exponential gradient of fibroblast growth factor (Fgf), a morphogen originating at the fin edge that has been experimentally demonstrated to facilitate normal digit arrangement in mice. In their model, Fgf represses Sox9 repression of *bmp* expression (k_4) and promotes Sox9 repression of *wnt* expression (k_7). In simulation, the authors observed that increasing the ratio of Wnt production to Bmp production or decreasing Bmp promotion of *sox*9 expression caused the Turing pattern to transition from stripes to spots.

We implemented the model from [328] using the following evolution equations (Figure

2.19):

$$\begin{cases} \dot{F}_i(t) = \alpha_F + u_{1i} - \mu_F F_i(t) + \frac{d_f}{l^2} v_{F_i}(t) \\ \dot{s}_i(t) = \alpha_s + u_{2i} + k_2 b_i(t) - k_3 z_i(t) - s_i(t)^3 \\ \dot{b}(t) = \alpha_{bmp} + u_{2i} - k_4 \left(1 - k_f F_i(t)\right) s_i(t) - k_5 b_i(t) + \frac{d_b}{l^2} v_{b_i}(t) \\ \dot{z}_i(t) = \alpha_{wnt} + u_{2i} - k_7 k_f F_i(t) s_i(t) - k_9 z_i(t) + \frac{d_z}{l^2} v_{z_i}(t) \\ \dot{w}_i(t) = \begin{bmatrix} F_i(t) \\ b_i(t) \\ z_i(t) \end{bmatrix} \\ v(t) = (M \otimes I_3)\, w(t) \end{cases} \tag{2.132}$$

where α_F is the Fgf production rate, u_1 represents the source of Fgf, and u_2 is random constant-in-time spatial variation in background production rate. Unlike [328], we did not normalize F to $[0, 1]$, but we chose $u_1, \underline{u}_1$ such that $0 \leq \tilde{F}_i^{\,*} + \underline{\bar{F}}^* \leq 1$ and $0 \leq u_{1_i} - \underline{\bar{u}}_1$.

As compared to (2.129), the equations in (2.132) are rendered as perturbations to prior steady-state protein concentrations, therefore "negative" steady-state values should be interpreted as reductions in concentration relative to preexisting levels.

To handle both background production rate and localized Fgf production we use the generalization to L inputs

$$\hat{y}^* := -\left(I_N \otimes C\right) \left[\left(I_N \otimes A\right) + \Lambda \otimes \left(B_v G\right)\right]^{-1} \left[\sum_{k=1}^{L} \left(I_N \otimes B_{u_k}\right) \hat{u}_k\right] \tag{2.133}$$

where $u_k \in \mathbb{R}^N$ is the kth input vector and

$$B_{u_k} := \frac{\partial f}{\partial u_k}\Big|_{(\underline{\bar{x}}^*, \underline{\bar{v}}^*, \underline{\bar{u}}_1, \underline{\bar{u}}_2, \ldots, \underline{\bar{u}}_L)} \tag{2.134}$$

is the linearization matrix for one subsystem with respect to the kth input when all inputs are held constant in time and space. To avoid ambiguity, the "filter" interpretation is defined with respect to one input, i.e., as one term in the summation (2.133).

The matrices for the linearization are

$$A = \begin{bmatrix} -\mu_F & 0 & 0 & 0 \\ 0 & -3\underline{\bar{s}}^{*2} & k_2 & -k_3 \\ k_4 k_f \underline{\bar{s}}^* & -k_4\left(1 - k_f \underline{\bar{F}}^*\right) & -k_5 & 0 \\ -k_7 k_f \underline{\bar{s}}^* & -k_7 k_f \underline{\bar{F}}^* & 0 & -k_9 \end{bmatrix}, \tag{2.135}$$

$$B_v = \begin{bmatrix} \frac{d_f}{l^2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \frac{d_b}{l^2} & 0 \\ 0 & 0 & \frac{d_z}{l^2} \end{bmatrix}, \quad B_{u_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad B_{u_2} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.136}$$

where the steady-state concentration of Fgf is $\underline{\bar{F}}^* = \frac{\alpha_F + \underline{\bar{u}}_1}{\mu_F}$ independent of the other variables. We stabilized the homogeneous steady-state solution by setting $\alpha_F = 0$ and using a small value of $\underline{\bar{u}}_1$ with the remaining parameters taken from Figures 4 and 5 in [328]. This choice of α_F completely localizes the source of Fgf to the input u_2.

As in the original Sox9-Bmp-Wnt Turing model in [327], the filter coefficients in the Fgf-augmented model form a bandpass at mid-range frequencies, resulting in the roughly periodic output patterning that alternates between Sox9 and Wnt (Figure 2.20). Increasing α_{wnt} decreases the magnitude of the bandpass, while decreasing k_2 concentrates amplification at a small range of frequencies inside the bandpass. Either of these parameter changes tends to shrink contiguous regions of high [Sox9], consistent with the transition from stripes to spots observed in [328]. Parameters yielding more spotlike patterns also tend to suppress the influence of both correlated and uncorrelated noise for readout [Sox9], though the effect on readout [Wnt] is negligible (Figures 2.21, 2.22, 2.23, and 2.24). The distal edge where the source of Fgf is localized exhibits relatively higher Wnt than Sox9 expression, as observed *in vivo* [328]; in our normalized images, the effect is most visible at higher values of α_{wnt}.

If we assume cells are approximately 12 to 15 μm in diameter [329], then for $\alpha_{wnt} = 1.2, k_2 = 1$ the filter and $\mathcal{H}_2$ norm analysis indicate that wavelengths of about 84 to 105 μm will be most strongly amplified in the result. The prediction is in decent agreement with the experimental images in Figure 2 of [328], which exhibit periodicity on the order of 80 to 100 μm. Some of the error may be accounted for by the difference in domain shape between filter simulations and actual limb paddles (rhomboidal vs. elliptical) as well as the presence of growth in the living animal. Nevertheless, this observation suggests that the framework correctly identifies the range of spatial modes that will be most influential in forming the "actual" biological pattern.

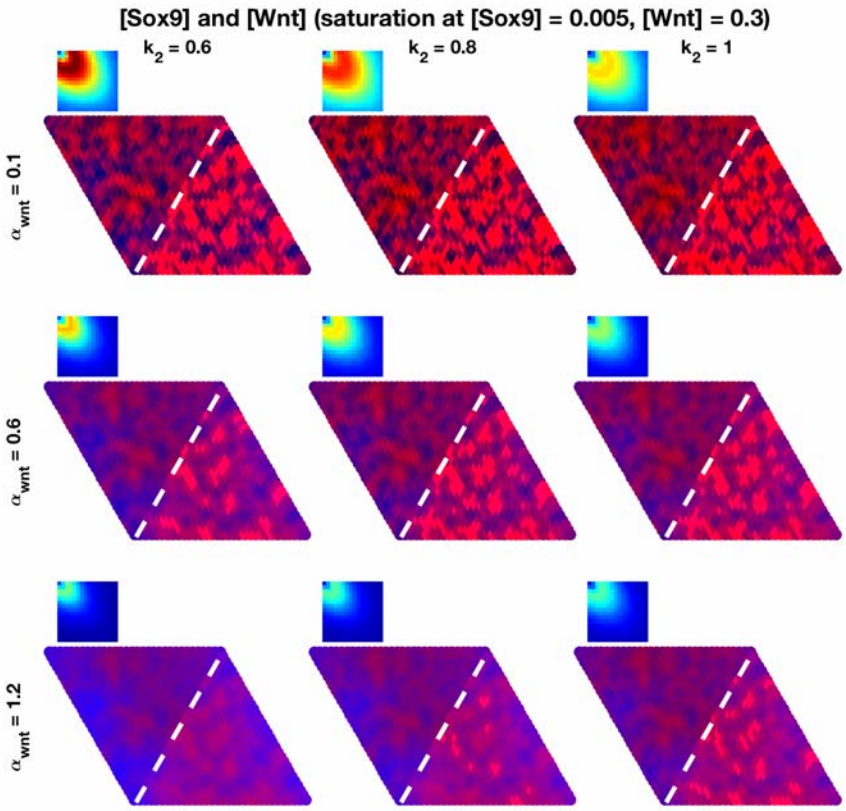

Figure 2.20: Either increasing Wnt production or decreasing Bmp promotion of *sox9* expression shrinks the size of contiguous high-[Sox9] regions, though filter analysis shows the two methods act through different mechanisms. Simulations are performed for nine (α_{wnt}, k_2) pairs on a hexagonal lattice ($N_R = N_C = 32$) with two external inputs: a random background production rate for Sox9, Wnt, and Bmp; and an Fgf source localized to 5 columns of cells on the left. The effect of the Fgf is visible as an increase in [Wnt] (blue) relative to [Sox9] (red) at both the left and right boundaries owing to the periodic boundary conditions. The "actual" readouts, normalized across all images independently by channel, are pictured above the dotted white line; readout values below the dotted line have been post-processed to saturate at a threshold (0.005 for [Sox9], 0.3 for Wnt) and are normalized in the same fashion as (but separately from) the "actual" readouts. For visual emphasis, saturated [Sox9] values are displayed at 10× the threshold intensity. Inset heat maps display the magnitude of the filter coefficients around the bandpass (from $k_C = 0$ to $k_C = \frac{N_R}{2}$ and $k_R = 0$ to $k_R = \frac{N_C}{2}$) from input background production to readout [Sox9], each normalized to the same range (min. 0, max. 3.36). Readouts that saturate above a certain threshold show more spotlike patterns for higher α_{wnt} or lower k_2, as observed in [328]. Increasing α_{wnt} decreases the overall amplitude of the filter and thereby shrinks the width of the passband, which suggests that spots rather than stripes may emerge when fewer cells express above a threshold. In contrast, decreasing k_2 increases the maximum magnitude of the bandpass in a concentrated region, suggesting that spots may also be obtained by exaggerating differences in amplification between frequencies. Parameters unless noted otherwise are as given in [328], Figures 4 and 5.

Parameter	Value	Description
α_{sox9}	0	constitutive Sox9 production rate
α_{bmp}	0.1	constitutive Bmp production rate
α_{wnt}	1.2	constitutive Wnt production rate
μ_F	0.1	Fgf decay rate
k_2	1	Bmp promotion of *sox9* expression
k_3	3	Wnt repression of *sox9* expression
k_4	6	Sox9 repression of *bmp* expression
k_5	0.1	Bmp decay rate
k_7	2.4	Sox9 repression of *wnt* expression
k_9	0.1	Wnt decay rate
k_f	$\frac{2}{3}$	strength of Fgf influence on k_4, k_7
d_b	160	diffusion coefficient for Bmp
d_w	25	diffusion coefficient for Wnt
d_F	600	diffusion coefficient for Fgf
l	4	distance between cells

Table 2.3: Parameters used in the simulations of digit formation with a morphogen gradient, unless noted otherwise in the text (Figures 2.20 to 2.24). Values are from Methods in [328] (Figures 4 and 5).

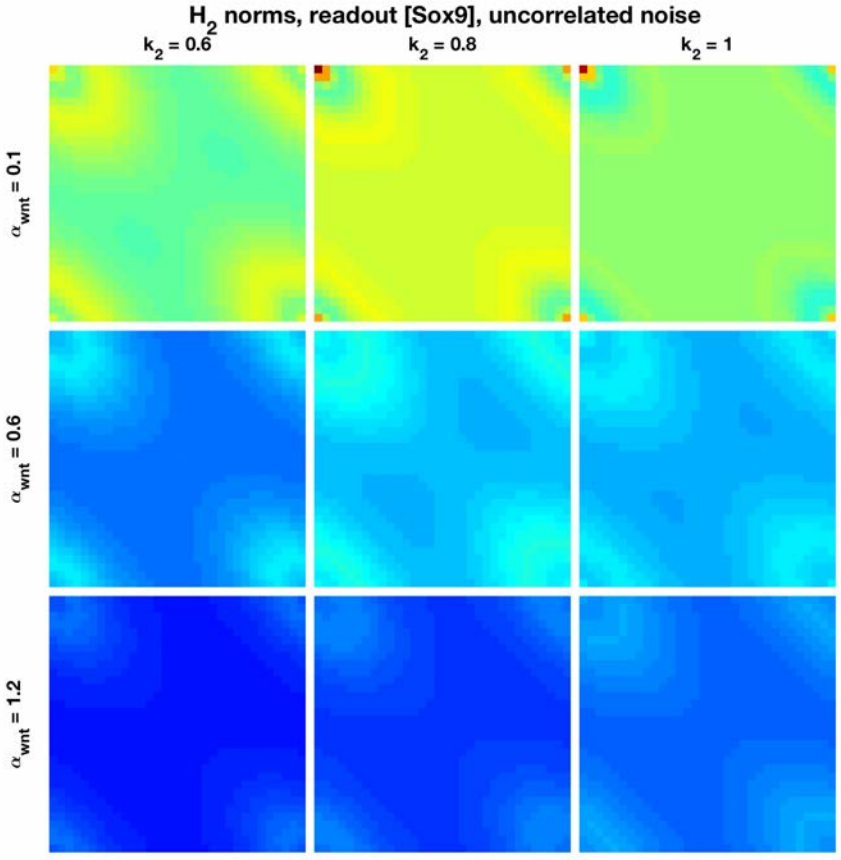

Figure 2.21: $\mathcal{H}_2$ norms with readout [Sox9] for 100% uncorrelated white noise inputs to *sox9*, *bmp*, and *wnt*. Images are normalized to the same scale as Figure 2.22 (min. 0, max. 3.43). Other parameters are from Table 2.3.

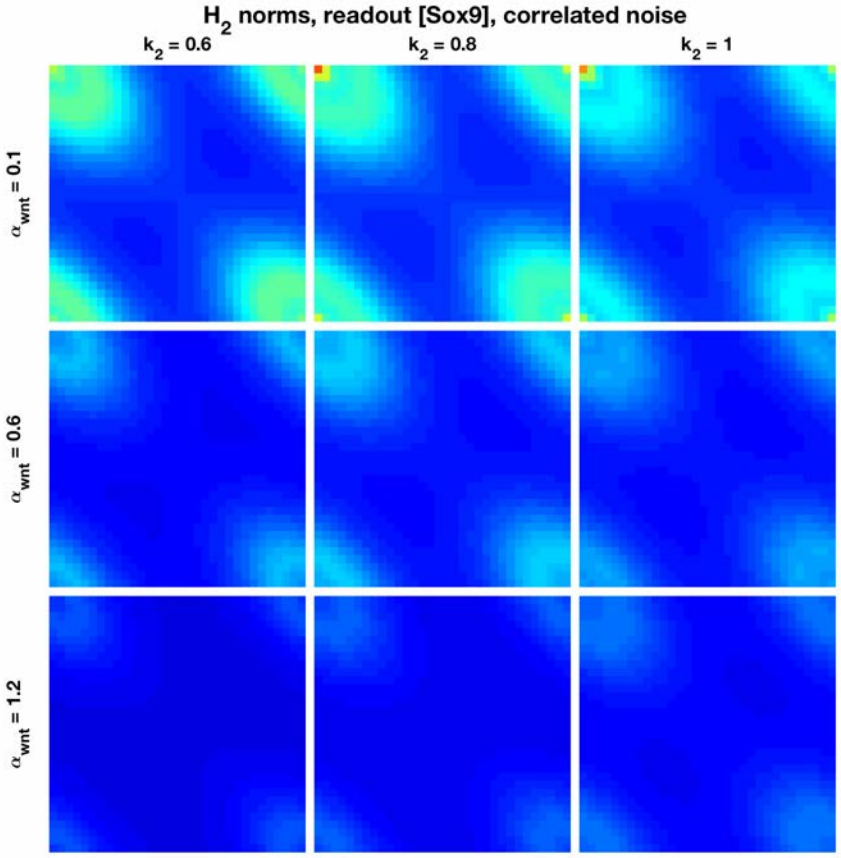

Figure 2.22: $\mathcal{H}_2$ norms with readout [Sox9] for 100% correlated white noise inputs to *sox9*, *bmp*, and *wnt*. Images are normalized to the same scale as Figure 2.21 (min. 0, max. 3.43). Other parameters are from Table 2.3.

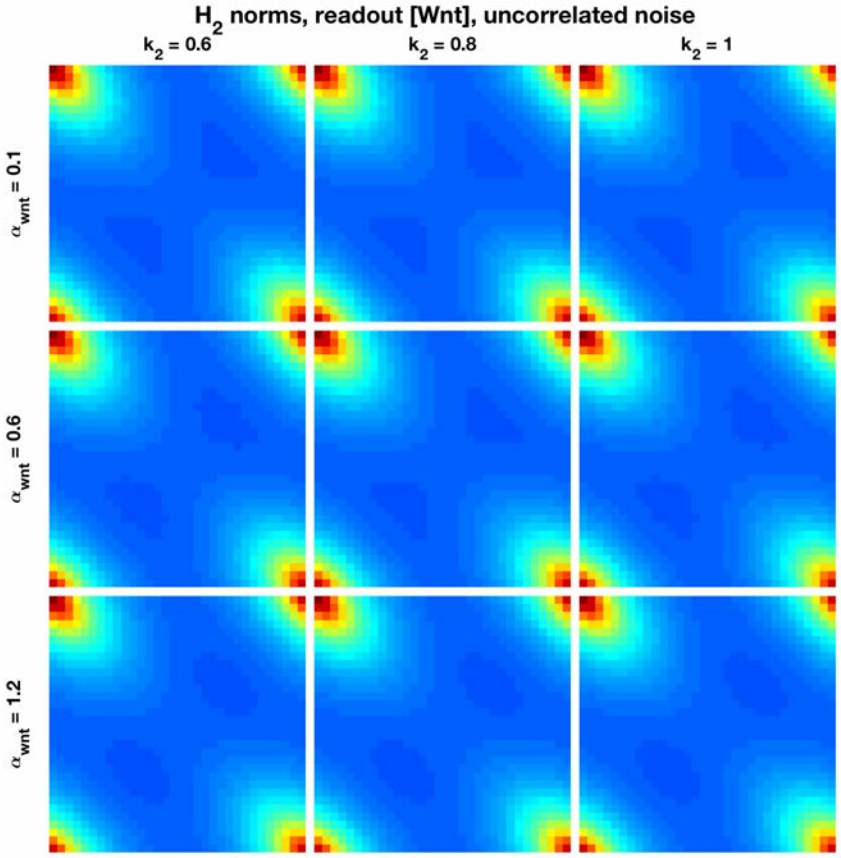

Figure 2.23: $\mathcal{H}_2$ norms with readout [Wnt] for 100% uncorrelated white noise inputs to *sox9*, *bmp*, and *wnt*. Images are normalized to the same scale as Figure 2.24 (min. 0, max. 2.28). Other parameters are from Table 2.3.

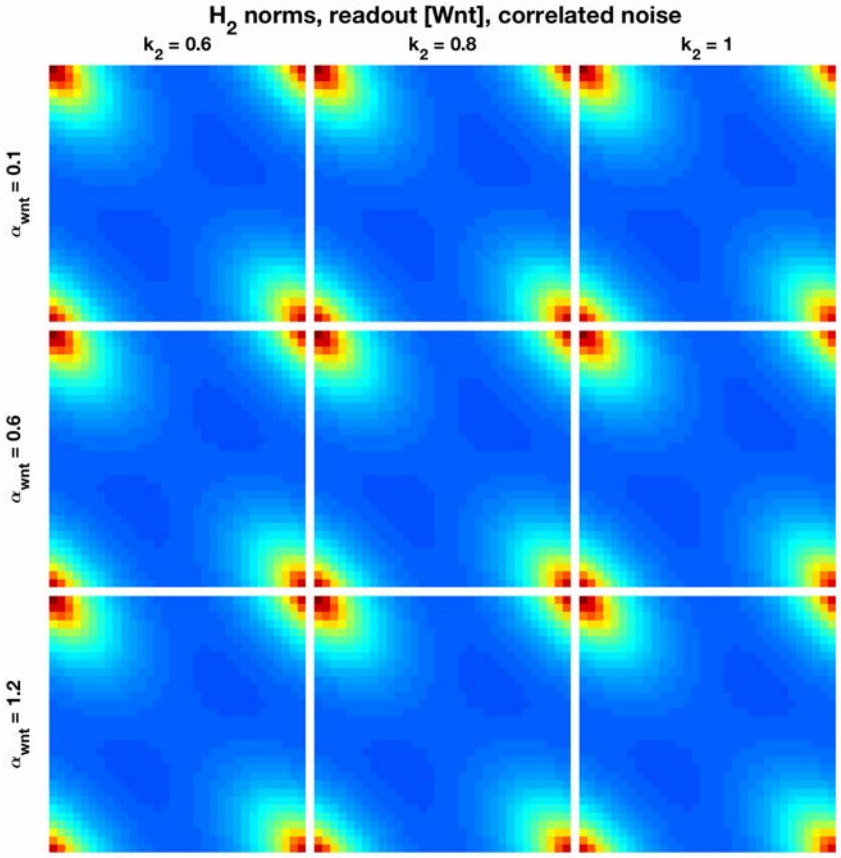

Figure 2.24: $\mathcal{H}_2$ norms with readout [Wnt] for 100% correlated white noise inputs to *sox9*, *bmp*, and *wnt*. Images are normalized to the same scale as Figure 2.23 (min. 0, max. 2.28). Other parameters are from Table 2.3.

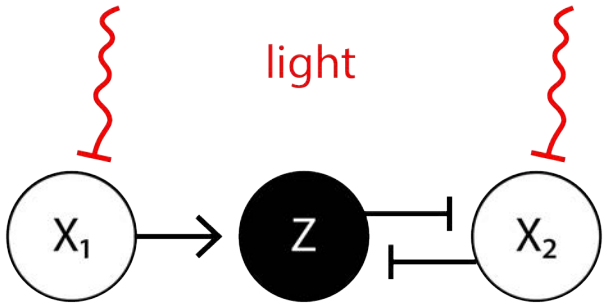

Figure 2.25: Schematic of the synthetic edge detector implemented in [166] (see text).

2.7 Application: Synthetic Edge Detector

Spatial patterning of gene expression is an area of ongoing interest to synthetic biologists for applications in microbial consortia [330] and biocomputation, among others (see [181] and [182] for recent reviews). A few efforts in particular have focused on image reproduction [167], [331] and processing [166] in systems of light-responsive cells, where "lawns" of engineered bacteria are illuminated with prepatterns of spatially varying intensity and/or color. We apply the filtering approach to analyze the synthetic edge detector implemented in [166], in which bacteria in the dark produce a diffusible chemical, and bacteria that sense both light and the chemical produce a pigment (Figure 2.25).

Cells in the dark produce an enzyme (LuxR) that facilitates production of X_1, the diffusible communication signal (AHL), which activates production of a black pigment Z that acts as the readout. The protein X_2 (Cl), also produced in the dark (inhibited by red light), inhibits Z production. Both enzyme and X_2 are inhibited by light, such that the steady-state transfer function (in Miller units) of X_1 and X_2 as a function of light input L is given by

$$f_{light}(L) = \frac{K}{K+L}\left(\beta_{max} - \beta_{min}\right) + \beta_{min}. \tag{2.137}$$

Both X_1 and X_2 bind the $P_{lux-\lambda}$ promoter controlling Z production with steady-state transfer function

$$f_{logic}\left(x_1, x_2\right) = \frac{c_0 + c_1 f_{\mathrm{Lux}}(x_1)}{1 + c_0 + c_1 f_{\mathrm{Lux}}(x_1) + c_2 f_{\mathrm{Cl}}^n(x_2) + c_1 c_2 f_{\mathrm{Lux}}(x_1) f_{\mathrm{Cl}}^n(x_2)} \tag{2.138}$$

where

$$f_{\mathrm{Lux}}(x_1) = \frac{1}{2}\left[\left(\mathrm{LuxR}_{tot} + \frac{K_A^2 K_D^{\mathrm{LuxR}}}{4x_1^2}\right) - \sqrt{\left(\mathrm{LuxR}_{tot} + \frac{K_A^2 K_D^{\mathrm{LuxR}}}{4x_1^2}\right)^2 - \mathrm{LuxR}_{tot}^2}\right] \tag{2.139}$$

$$f_{\mathrm{Cl}}(x_2) = \frac{x_2}{2} + \frac{1}{8K_D^{Cl}}\left(1 - \sqrt{1 + 8K_d^{Cl} x_2}\right). \tag{2.140}$$

The authors of [166] assume AHL diffusion is the slowest process in order to describe the system in just three dimensionless equations for AHL and Cl production and AHL diffusion. Because we ultimately only care about steady-state solutions under constant spatial perturbation, we can ignore the relative degradation rate of X_2 to X_1 and simply choose dynamics for a system that will produce the correct transfer functions as steady-state solutions. The simplest choice gives the following dynamical equations for cell i:

$$\begin{cases} \dot{x}_{1i} = \kappa_1 f_{light}(\bar{\underline{L}}) - \kappa_2 x_{1i} + D v_i \\ \dot{x}_{2i} = \kappa_3 f_{light}(\bar{\underline{L}}) - x_{2i} \\ z_i = \kappa_4 f_{logic}(x_{1i}, x_{2i}) \\ v = M x_1 \end{cases} \tag{2.141}$$

where D is the diffusivity of AHL and M is Laplacian, such that the homogeneous steady-state solutions under constant-in-space light input $\underline{L}$ are

$$\begin{cases} \bar{\underline{x}}_1^* = \frac{\kappa_1}{\kappa_2} f_{light}(\bar{\underline{L}}) \\ \bar{\underline{x}}_2^* = \kappa_3 f_{light}(\bar{\underline{L}}) \\ \bar{\underline{z}}^* = \kappa_4 f_{logic}(\bar{\underline{x}}_1^*, \bar{\underline{x}}_2^*) \end{cases} . \tag{2.142}$$

Then the linearization matrices are

$$A = \begin{bmatrix} -\kappa_2 & 0 \\ 0 & -1 \end{bmatrix}, \; B_v = \begin{bmatrix} \frac{D}{l^2} \\ 0 \end{bmatrix}, \quad B_u = -\frac{K\left(\beta_{max} - \beta_{min}\right)}{\left(K + \bar{\underline{L}}\right)^2} \begin{bmatrix} \kappa_1 \\ \kappa_3 \end{bmatrix}, \tag{2.143}$$

$$G = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad C = \kappa_4 \begin{bmatrix} a_1 & a_2 \end{bmatrix} \tag{2.144}$$

where

$$a_1 = \left.\frac{\partial f_{logic}}{\partial x_1}\right|_{\bar{\underline{x}}^*}, \; a_2 = \left.\frac{\partial f_{logic}}{\partial x_2}\right|_{\bar{\underline{x}}^*}. \tag{2.145}$$

In the absence of further information about system dynamics, we assume the doubling time of bacteria is faster than this system reaches steady state (reasonable given comparison to the AHL half-life—see Table 2.4). Thus, we take l to be a distance reflecting discretization in space, rather than a distance between individual cells. Then the filter coefficients are given by

$$[S]_{kk} = \frac{K\left(\beta_{max} - \beta_{min}\right)}{\left(K + \bar{\underline{L}}\right)^2} \left(\frac{a_1 \kappa_1}{\frac{D}{l^2}\lambda_k\left(M\right) - \kappa_2} - a_2 \kappa_3 \right). \tag{2.146}$$

For the remainder of this example, we use a hexagonal array of evenly spaced cells with periodic boundary conditions.

2.7.1 Analysis

The synthetic edge detector acts as a highpass filter. Although the magnitude of the filter coefficients decreases with increasing average input $\bar{\underline{L}}$, the shape of the filter is almost

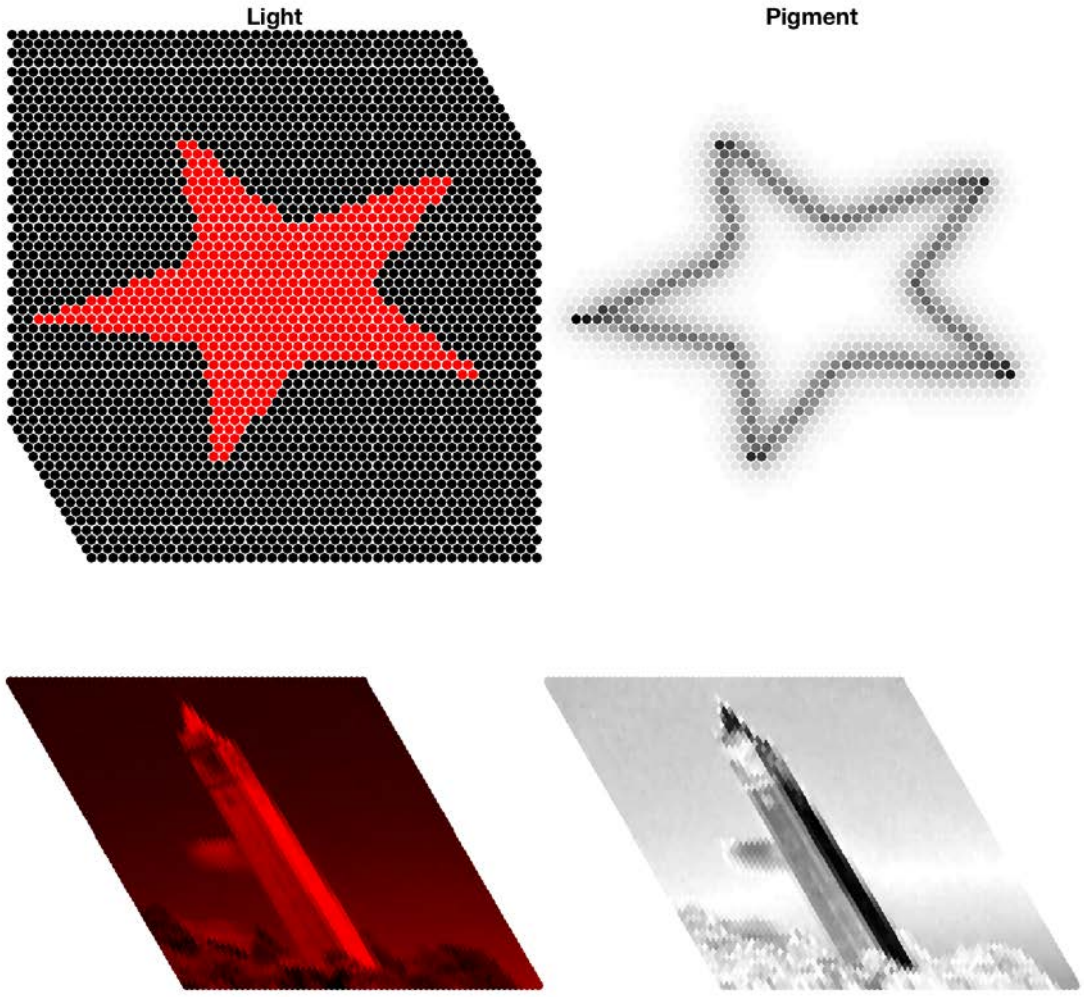

Figure 2.26: The synthetic edge detector acts as a highpass that is sufficiently sharp to detect edges in binary images, but too broad to distinguish edges in grayscale. Left, the light input seen by individual cells; right, the resulting output pigment pattern. Full lattice is hexagonal with 64×64 units as in Figure 2.7, with periodic boundary conditions. Top plots are zoomed in for clarity. Parameters are as given in [166] and can be found in Table 2.4.

Parameter	Value	Description
β_{max}	298	Miller
β_{min}	125	Miller
K	0.0017	Miller
D	$1.67 \times 10^{-7} \times 60^2$	AHL diffusivity (cm^2/hr)
c_0	0.04	
c_1	0.05	1/nM
c_2	0.011	$1/nM^n$
n	1.5	
κ_1	0.03	nM/hr
κ_2	0.012	AHL half-life (hr^-1)
κ_3	0.8	nM/Miller
κ_4	289	Miller
$K_A^2 K_D^{\mathrm{LuxR}}$	270,000	nM^3
LuxR_{tot}	2000	nM
K_D^{Cl}	5	nM
l	$1e-1$	distance between discretized units (cm)

Table 2.4: Values are from text and experimental procedures of [166].

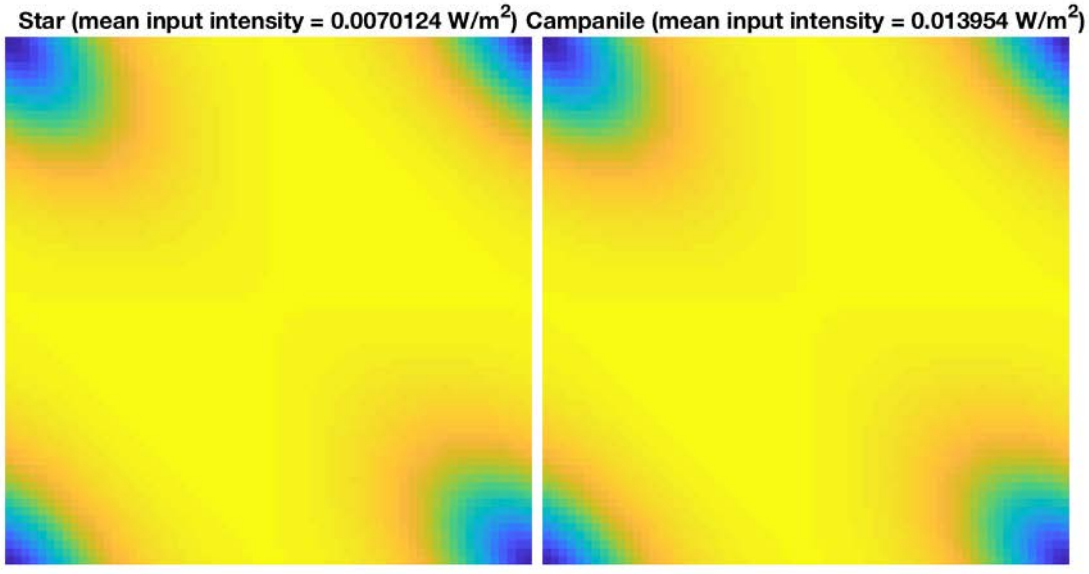

Figure 2.27: Filter coefficients for the synthetic edge detector, normalized separately with a range of 1.7×10^{-5} (left) and 8.0×10^{-6} (right). Note the loosely highpass behavior. Skew owes to the choice of a hexagonal lattice. The general shape appears invariant to the mean input intensity.

completely invariant to $\underline{\bar{L}}$ (Figure 2.27), which is a point in favor of the system's ability to properly detect edges in images differing in mean intensity. Owing to the diffusion of the communication signal, however, the filter is not a particularly sharp highpass, such that edges are not clearly identified in "low-contrast" images with areas of middling intensity. This is an understandable limitation given the biological components available to construct the circuit, and may partially reflect the binary logic (IF/AND/NOT) employed by the designers.

Partial differential equation simulations, such as those used by [166] to predict patterning in the synthetic edge detector, may be more accurate than linearization-based approaches such as ours, but can also be time consuming and memory intensive especially as the system increases in dimension. In comparison, to employ our approach the filter coefficients need only be calculated once per parameter choice, and a single matrix multiplication then suffices to predict readouts for a wide variety of inputs. Thus, the filtering approach may provide a foundation for efficiently testing and evaluating putative biological image processing circuits, particularly those with numerous components. In addition, a filter-based approach could permit synthetic biologists to tackle more complicated image processing functions that require finely tuned spatial frequency responses.

Lastly, filters can be serially applied (or "cascaded") to carry out more sophisticated image processing functionalities. In these cases, the readout pattern of one filter acts as the prepattern input to the next layer, and the filter coefficients of the full cascade are given by the componentwise product of the coefficients for each stage. Hence, synthetic biologists could design separate filter modules to execute specific functions and then connect them serially to carry out a more sophisticated overall program.

2.8 Conclusions

In this chapter we have presented a framework to analyze how networks of interacting cells modify spatially varying inputs, either from environmental factors or intrinsic parameter variation, to produce patterned outputs. Three biologically relevant examples indicate that qualitatively similar patterns may arise from different physical implementations (Section 2.5), from both stable and unstable fixed points (Section 2.6), as well as from variable filter behaviors when certain postprocessing steps are applied (Section 2.6.2). Furthermore, these biological models appear robust to correlated and uncorrelated space-and-time-varying white noise inputs—a critical feature for maintaining consistency during embryonic development.

We have demonstrated in a theoretical context how a filtering approach can offer insight into system behavior at an intermediate level between the exact physical implementation and the measured result. In an experimental context, evaluating systems at the filter level may clarify when alterations to the input are capable of distinguishing between alternative explanations for an observed behavior. The perturbation experiments suggested at the end of Section 2.5 are one example. Other systems may exhibit near-identical filter coefficients despite drastic changes to mean input levels, suggesting that pure input-output probing is

unlikely to illuminate the underlying mechanism. Moreover, model systems with similar coefficients for some range of spatial frequencies and disaparate values for another range will vary in their response to inputs with frequencies in both ranges, such that experiments in which inputs to the real system can be finely controlled may suffice to differentiate more accurate models from less accurate ones. Of interest in all cases is the extent to which a particular system may impose structure upon a readout pattern as compared to how much structure must be present in the prepattern. The inverse problem of how to infer filter coefficients from a given prepattern and readout pattern is also a subject for further investigation, as such a method would ease the requirement for an extant dynamical model in order to apply the filtering interpretation.

A critical assumption in our development of the filtering framework is that linearization about a homogeneous steady state is sufficient to capture relevant system behavior. Future work should focus on incorporating nonlinear dynamics as well as investigating the influence of external inputs on spatially distributed, networked systems in the vicinity of unstable or nonhomogeneous steady states. Additional areas for further research include system response to non-white noise inputs and patterning in time-varying or perturbed networks (an especially relevant consideration given that some chemical gradients appear to evolve on the same timescale as the gene regulatory networks responding to them [25]). Lastly, we believe the filtering framework could be especially informative for designing synthetic "image processing" circuits such as the edge detector examined in this text. Of note, the spatial frequency-based interpretation opens the door for synthetic biologists to leverage extant signal processing literature based around filter designs. We encourage further application of our tool to genetic circuits in engineered multicellular systems.

Overall, we believe a filtering approach simplifies the process of predicting how intermolecular and intercellular interactions affect patterning mechanisms in living organisms or communities thereof. It is our hope that the viewpoint developed here will help us to elucidate—and elaborate upon—nature's designs.

Acknowledgments

This work first appeared in [305] and [306], in which the authors thank Andy Packard for providing incisive feedback. This research was funded by the United States Air Force Office of Scientific Research award FA9550-14-1-0089.

Chapter 3

Contrasting Patterning through Diffusion-mediated Lateral Inhibition

The main contribution of this work to extant literature in synthetic biology is an experimental platform for facilitating diffusion-based communication among bacterial colonies. Guided by theory, we harnessed the platform to demonstrate the stability of contrasting gene expression states between neighboring colonies linked by channels etched in polydimethylsiloxane (PDMS) mold, and further showed that spatial parameters such as diffusion rates and the distance between colonies could act as a control knob for the stability and contrast level of the resulting patterns. Building on the original theory from [332], this chapter provides proofs for contrasting pattern emergence in the experimental setup *if and only if* the full system can be reduced to a two-compartment equivalent. The importance of spatial layout and geometry for pattern emergence are emphasized.

3.1 Introduction

Over the last two decades, synthetic biologists have sought to engineer microbes and their consortia to execute ever more complex tasks. These range from the relatively straightforward production of valuable chemicals [333] to the computation of complex logics [334–336] that allow the microbes to make sophisticated decisions to optimize such production [337–339] or release therapeutics *in situ* [340, 341]. There has been sustained interest in both expanding the intrinsic size of these "circuits" for implementation of even more ambitious functions and spreading these circuits among different members of a population that distributes the production load [342, 343], allows reuse of components [344], or better utilizes space [149, 166, 167, 345, 346] and growth [164, 347, 348]. Certainly, in nature, microbial populations and consortia spatially arrange themselves to form specialized structures that have mechanical, developmental, and chemical advantages over homogeneous distributions [349–353].

Ideally, spatial patterning arises in a self-organized fashion from control among individual

cells, but shaping communities of cells remains a challenge to engineer [354]. In natural systems, this organization often relies on highly specific communication among cells and/or among highly spatially constrained signaling such as cell contact-mediated mechanisms [180]. In bacteria, intercellular signaling systems like acyl homoserine lactones (AHL) used in quorum-sensing (QS) tend to be fairly non-specific and long range, while contact-mediated systems like Cdi-A/B [355, 356] are not easily programmable. However, it is possible to experimentally explore the principles of spatial organization among microbial populations by imposing external constraints on communication.

In this work, we designed a compartmentalized culturing platform to house multiple spatially separated bacterial colonies harboring different genetic circuits constrained to communicate through specified channels on the device. To demonstrate the utility of the setup, we adapted the theory developed in [289], [122], and [332] for spontaneous contrasting patterning through lateral inhibition to the specific case of quorum sensing between colonies of synthetically engineered bacterial cells. We found that it was experimentally difficult to match kinetic rate parameters across strains, leading us to harness the geometric arrangement of colonies on the culturing platform as a control parameter for attaining contrasting patterns.

In the following sections, we introduce the biological models for our experimental implementation alongside the theory predicting their behavior. We begin by developing intuition with a two-compartment system, which we will show to be mathematically equivalent to a larger system with a certain network structure for the purposes of assessing the existence and stability of contrasting patterns. Once the basic theoretical results are established, we present computational results to analyze the system followed by experimental results that validate the theory.

3.1.1 Two-compartment system

To build intuition, we first consider a two-compartment system of mutual inhibition comprising two colonies of bacterial cells of Strains A and B. Suppose each cell type produces a product along with an associated signaling molecule that inhibits the other cell type's production of signaling molecule. Depending on factors such as the strength of inhibition and the proximity of the colonies to each other, over time one colony may be "dominated" by the other, which produces vastly more product. In this application, we are interested in the situation where either colony could theoretically become the dominant colony and where a near-homogeneous solution (in which both colonies produce intermediate levels of product) is not stable. Such a system is considered bistable, with one stable point corresponding to strain A high, B low and the other corresponding to strain A low, B high.

We assume the colonies communicate by way of diffusible molecules and that signals from one cell type inhibit cells of the other type but not cells of the same type; i.e., the cells employ orthogonal quorum-sensing (QS) systems. Ultrasensitivity is necessary for bistability, but most QS systems alone are not ultrasensitive, therefore we introduce a repressor such as TetR that binds a promoter with an ultrasensitive response. A channel between the two

Variable	Description
X_A	HSL (A) in A (produced by A)
X_B	HSL (A) in B (diffused from A)
Y_A	HSL (B) in A (diffused from B)
Y_B	HSL (B) in B (produced by B)
m_{R_A}	repressor mRNA in A
m_{R_B}	repressor mRNA in B
R_A	repressor in A
R_B	repressor in B
m_{I_A}	AHL synthase mRNA in A
m_{I_B}	AHL synthase mRNA in B
I_A	AHL (A) in A
I_B	AHL (B) in B

Table 3.1: Chemical species (state space variables) whose behavior is represented by system ODEs. Brackets denote concentration.

compartments allows diffusion of AHL between the colonies. To prevent AHL buildup due to its slow decay rate, we allow the possibility of adding an extra "drain" channel of the same length as the other channels that connects each compartment to a reservoir with zero concentration AHL. We incorporate this drainage into the overall decay rate for AHL (Table 3.2). We assume colonies are in the exponential phase of growth and incorporate dilution directly into the decay rate of mRNA and protein within cells (Table 3.2).

The ordinary differential equations governing system behavior are divided into three boxes (modules) per cell, each describing a particular function. The state space variables are listed in Table 3.1. The parameters are defined and values provided in Table 3.2.

The production box $\boldsymbol{H_A}$ describes AHL (A) synthase production by A for a given input concentration of repressor. Repressor inhibition of synthase expression is modeled by a repressive Hill function:

$$\boldsymbol{H_A} \; : \; \begin{cases} \dot{m}_{I_A} = V_{I_A} N_{I_A} C \left(\frac{1}{1+\left(\frac{R_A}{K_{R_A}}\right)^{n_{R_A}}} + l_{I_A} \right) - \gamma_{m_{I_A}} m_{I_A} \\ \dot{I}_A = \epsilon_{I_A} m_{I_A} - \gamma_{I_A} I_A \end{cases} . \tag{3.1}$$

The steady-state concentration of synthase for a constant input R_A^* is then

$$I_A^* = \frac{\epsilon_{I_A} V_{I_A} N_{I_A} C}{\gamma_{I_A} \gamma_{m_{I_A}}} \left(\frac{1}{1+\left(\frac{R_A^*}{K_{R_A}}\right)^{n_{R_A}}} + l_{I_A} \right) . \tag{3.2}$$

The transceiver box $\boldsymbol{tx_{A \to B}}$ encapsulates the catalysis of AHL (A) by AHL (A) synthase and the diffusion of AHL (A) to B. It accepts an input I_A, the concentration of AHL (A)

produced by cell A, and outputs a vector $X = [X_A,\ X_B]^T$, where the first entry is the concentration of AHL (A) in A and the second is the concentration of AHL (A) in B after diffusion. The equations governing the behavior of this box are

$$\boldsymbol{tx_{A\to B}}\ :\ \begin{cases}\dot{X}_A = d(X_B - X_A) - \gamma_X X_A + \nu_X I_A \\ \dot{X}_B = d(X_A - X_B) - \gamma_X X_B\end{cases}. \tag{3.3}$$

where $d(X_B - X_A)$ represents the diffusion balancing of AHL (A) between A and B (the term $d := \frac{D}{l^2}$ includes the diffusion coefficient D and the channel length l). The steady-state concentration of X_B in response to constant input I_A^* is then

$$X_B^* = \frac{d\nu_Y I_A^*}{\gamma_X(\gamma_X + 2d)}. \tag{3.4}$$

The receiver box $\boldsymbol{rx_B}$ captures repressor production in B as a function of X_A. The upregulation of repressor mRNA transcription by AHL (A) binding to AHL (B) synthase is modeled by a Hill function:

$$\boldsymbol{rx_B}\ :\ \begin{cases}\dot{m}_{R_B} = V_{R_B}N_{R_B}C\left(\frac{\left(\frac{X_B}{K_X}\right)^{n_X}}{1+\left(\frac{X_B}{K_X}\right)^{n_X}} + l_{R_B}\right) - \gamma_{m_{R_B}} m_{R_B} \\ \dot{R}_B = \epsilon_{R_B} m_{R_B} - \gamma_{R_B} R_B\end{cases}. \tag{3.5}$$

The complementary boxes are identical in structure to those above, and are included here only for completeness. The production box $\boldsymbol{H_B}$ describes AHL (B) synthase production by B:

$$\boldsymbol{H_B}\ :\ \begin{cases}\dot{m}_{I_B} = V_{I_B}N_{I_B}C\left(\frac{1}{1+\left(\frac{R_B}{K_{R_B}}\right)^{n_{R_B}}} + l_{I_B}\right) - \gamma_{m_{I_B}} m_{I_B} \\ \dot{I}_B = \epsilon_{I_B} m_{I_B} - \gamma_{I_B} I_B\end{cases}. \tag{3.6}$$

Transceiver box $\boldsymbol{tx_{B\to A}}$ covers catalysis of AHL (B) by AHL (B) synthase and the diffusion of AHL (B) to A. The output vector $Y = [Y_A,\ Y_B]^T$ stores the concentrations of AHL (B) in A and B respectively:

$$\boldsymbol{tx_{B\to A}}\ :\ \begin{cases}\dot{Y}_A = d(Y_B - Y_A) - \gamma_Y Y_A \\ \dot{Y}_B = d(Y_A - Y_B) - \gamma_Y Y_B + \nu_Y I_B\end{cases} \tag{3.7}$$

The complementary receiver box $\boldsymbol{rx_A}$ is nearly identical to $\boldsymbol{rx_B}$ except that it describes repressor production in A as a function of Y_A, so $V_{R_B}N_{R_B}$ is replaced by $V_{R_A}N_{R_A}$:

$$\boldsymbol{rx_A}\ :\ \begin{cases}\dot{m}_{R_A} = V_{R_A}N_{R_A}C\left(\frac{\left(\frac{Y_A}{K_Y}\right)^{n_Y}}{1+\left(\frac{Y_A}{K_Y}\right)^{n_Y}} + l_{R_A}\right) - \gamma_{m_{R_A}} m_{R_A} \\ \dot{R}_A = \epsilon_{R_A} m_{R_A} - \gamma_{R_A} R_A\end{cases}. \tag{3.8}$$

The steady-state concentration of repressor for a constant input Y_A^* is given by

$$R_A^* = \frac{\epsilon_{R_A} V_{R_A} N_{R_A} C}{\gamma_{R_A} \gamma_{m_{R_A}}} \left(\frac{\left(\frac{Y_A^*}{K_Y}\right)^{n_Y}}{1 + \left(\frac{Y_A^*}{K_Y}\right)^{n_Y}} + l_{R_A} \right). \tag{3.9}$$

The steady-state solutions differ from their complements only in exact parameter values. For our experimental implementation, we selected LuxR/LuxI and LasR/LasI as the orthogonal QS systems. The parameters are as defined in Table 3.2.

Parameter	**Description**	**Value**	**Source**
D	diffusivity of HSL	4.9×10^{-6} cm^2/s	
l	distance between channels	4 mm	design parameter
d	edge weight constant	-	$:= \frac{D}{l^2}$
λ	doubling time (bacterial growth)	27.5 min	-
μ	dilution due to growth	0.00042 s^{-1}	$:= \frac{\ln 2}{\lambda}$
γ_X	decay rate of C6-HSL + drainage	0.3371×10^{-4} s^{-1}	
γ_Y	decay rate of C12-HSL + drainage	0.3371×10^{-4} s^{-1}	
ν_X	production (catalysis) rate of C12-HSL from LasI	1×10^{-4} s^{-1}	
ν_Y	production (catalysis) rate of C6-HSL from LuxI	5×10^{-4} s^{-1}	
V_{R_A}	transcription rate of LuxI mRNA from *plux* in A	0.02 s^{-1}	
V_{R_B}	transcription rate of LasI mRNA from *plas* in B	$0.771 V_{R_A}$ s^{-1}	experimental fit
N_{R_A}	copy number of *plux* in A	1	design parameter
N_{R_B}	copy number of *plas* in B	1	design parameter
C	concentration constant	1	-
K_X	C12-HSL dissociation constant to LasI in B	8.335	experimental fit
K_Y	C6-HSL dissociation constant to LuxR in A	47.514	experimental fit
n_X	Hill coefficient (cooperativity) for C12-HSL to LasR in B	1.907	experimental fit
n_Y	Hill coefficient (cooperativity) for C6-HSL to LuxR in A	1.831	experimental fit
l_{R_A}	rate of leaky transcription from *plux* in A	0 s^{-1}	

l_{R_B}	rate of leaky transcription from *plas* in B	$0\ s^{-1}$	
$\gamma_{m_{R_A}}$	decay rate of TetR mRNA in A (dilution + degradation)	$\mu + 0.0042\ s^{-1}$	
$\gamma_{m_{R_B}}$	decay rate of TetR mRNA in B (dilution + degradation)	$\mu + 0.0042\ s^{-1}$	
ϵ_{R_A}	translation rate of TetR in A	$0.035\ s^{-1}$	
ϵ_{R_B}	translation rate of TetR in B	$0.035\ s^{-1}$	
γ_{R_A}	decay rate of TetR in A (dilution + degradation)	$\mu + 1.2836 \times 10^{-4}\ s^{-1}$	
γ_{R_B}	decay rate of TetR in B (dilution + degradation)	$\mu + 1.2836 \times 10^{-4}\ s^{-1}$	
V_{I_A}	transcription rate of mRNA from *ptet* in A	$0.02\ s^{-1}$	
V_{I_B}	transcription rate of mRNA from *ptet* in B	$0.02\ s^{-1}$	
N_{I_A}	copy number of *ptet* in A	1	design parameter
N_{I_B}	copy number of *ptet* in B	1	design parameter
K_{R_A}	dissociation constant of TetR to *ptet* in A	50	
K_{R_B}	dissociation constant of TetR to *ptet* in B	50	
n_{R_A}	Hill coefficient (cooperativity) for TetR to *ptet* in A	4	
n_{R_B}	Hill coefficient (cooperativity) for TetR to *ptet* in B	4	
l_{I_A}	rate of leaky transcription from *ptet* in A	0	
l_{I_B}	rate of leaky transcription from *ptet* in B	0	
$\gamma_{m_{I_A}}$	decay rate of LasI mRNA in A (dilution + degradation)	$\mu + 0.0042\ s^{-1}$	
$\gamma_{m_{I_B}}$	decay rate of LuxI mRNA in B (dilution + degradation)	$\mu + 0.0042\ s^{-1}$	
ϵ_{I_A}	translation rate of LasI in A	$0.07\ s^{-1}$	
ϵ_{I_B}	translation rate of LuxI in B	$0.07\ s^{-1}$	
γ_{I_A}	decay rate of LasI in A (no degradation)	μ	
γ_{I_B}	decay rate of LuxI in B (no degradation)	μ	

Table 3.2: Parameter definitions and values used in the ODEs for an experimental implementation using the orthogonal QS systems LuxI/LuxR and LasI/LasR.

Depending on the exact parameter values, the same system of ordinary differential equations may describe a monostable or bistable system. For the system of mutual inhibition that we have described, one way to determine whether a certain parameter set yields a bistable system is to count the number of steady-state points: If there is one point and it is stable, then the system is monostable; if there are three points total, one unstable and two stable (one each corresponding to the dominance of A or B), then the system is bistable. While steady-state points may be located using a convenient graphical method, it can be more difficult to analytically determine whether they are stable. Fortunately, if the system satisfies a set of technical conditions that classify it as *monotone*, then the graphical method will also reveal whether the equilibria are stable. We now show that our system is monotone and therefore we may use the graphical method to evaluate bistability.

In vector form, we represent the entire system with state space

$$Z := \begin{bmatrix} m_{I_A} \\ I_A \\ X \\ m_{R_B} \\ R_B \\ m_{I_B} \\ I_B \\ Y \\ m_{R_A} \\ R_A \end{bmatrix} \tag{3.10}$$

as $\dot{Z} = f(Z)$. Let $J(\cdot)$ be the Jacobian of $f(\cdot)$. For the system to be monotone, evolution of the system with time must preserve ordering in the state space, i.e., for solutions $\phi(\cdot,\cdot)$ to the differential equation, $x_1(0) \preceq_K x_2(0) \implies \phi(t, x_1(0)) \preceq_K \phi(t, x_2(0))$ for all $t \geq 0$. The ordering is defined with respect to some positivity cone K in a Euclidean space. Although physical concentrations cannot be negative, due to the presence of inhibition our system is not monotone with respect to the positive orthant, but rather to a combination of positive and negative orthants. If we redefined the state space as $\tilde{Z} := [-m_{I_A},\ -I_A,\ -X^T,\ -m_{R_B},\ -R_B,\ m_{I_B},\ I_B,\ Y^T,\ m_{R_A},\ R_A]^T$, then $\tilde{Z}$ would be monotone with respect to $\mathbb{R}^{12}_+$. Hence Z is monotone with respect to the cone K spanned by all vectors $\tilde{Z}$ corresponding to feasible Z (i.e., $Z \in \mathbb{R}^{12}_+$).

To see that the system is monotone, we apply a graphical method for systems with state, input, and output spaces defined by orthants [357]. We proceed by constructing an incidence graph (signed digraph) where each node is a species and each edge describes the relationship between two distinct species: no edge if no direct interaction; + if one promotes the other; and - if one inhibits the other (Figure 3.1). By Z_j promotes Z_i we mean $J_{ij} := \frac{\partial f_i}{\partial Z_j} = \frac{\partial^2 Z_i}{\partial t \partial Z_j} \geq 0$ for all $Z_i, Z_j \in \mathbb{R}_+$ whereas by Z_j inhibits Z_i we mean $J_{ij} \leq 0$ for all $Z_i, Z_j \in \mathbb{R}_+$.

Because each chemical species in the system directly affects only one other species in the system in a sequential fashion, the Jacobian for this system is sparse. We consider the entries corresponding to $\boldsymbol{H_A}$, $\boldsymbol{tx_{B\to A}}$, and $\boldsymbol{rx_A}$; because of the symmetry of the system, we know the entries corresponding to $\boldsymbol{H_B}$, $\boldsymbol{tx_{A\to B}}$, and $\boldsymbol{rx_B}$ will differ only in exact parameter values. Then the nonzero entries in the Jacobian (except those corresponding to one species' influence on its own concentration, which do not appear in the incidence graph) are

$$\boldsymbol{H_A} \;:\; \begin{cases} \frac{\partial \dot{m}_{I_A}}{\partial R_A} = V_{I_A} N_{I_A} C \frac{-n_{R_A}\left(\frac{R_A}{K_{R_A}}\right)^{n_{R_A}-1}}{K_{R_A}\left(1+\left(\frac{R_A}{K_{R_A}}\right)^{n_{R_A}}\right)^2} & \leq 0 \\ \frac{\partial \dot{I}_A}{\partial m_{I_A}} = \epsilon_{I_A} & \geq 0 \end{cases} \tag{3.11}$$

$$\boldsymbol{tx_{B\to A}} \;:\; \begin{cases} \frac{\partial \dot{Y}_B}{\partial I_B} = \nu_Y & \geq 0 \\ \frac{\partial \dot{Y}_A}{\partial Y_B} = d & \geq 0 \\ \frac{\partial \dot{Y}_B}{\partial Y_A} = d & \geq 0 \end{cases} \tag{3.12}$$

$$\boldsymbol{rx_A} \;:\; \begin{cases} \frac{\partial \dot{m}_{R_A}}{\partial Y_A} = V_{R_A} N_{R_A} C \frac{n_Y\left(\frac{Y_A}{K_Y}\right)^{n_Y-1}}{K_Y\left(1+\left(\frac{Y_A}{K_Y}\right)^{n_Y}\right)^2} & \geq 0 \\ \frac{\partial \dot{R}_A}{\partial m_{R_A}} = \epsilon_{R_A} & \geq 0 \end{cases} \tag{3.13}$$

along with the corresponding entries for the complementary boxes.

From a conceptual standpoint, monotonicity means that regardless of context, an element always has the same qualitative effect on itself after its influence is propagated through the network; i.e., the influence of the element on itself is "consistent". From the graph for this system (shown in Figure 3.1) we see that there is only one cycle and it is positive in parity; that is, the product of the signs of each edge traversed to complete one cycle is positive, regardless of the direction of travel around the cycle. This implies that the graph is consistent, and hence the system it describes is closed-loop monotone.

By Theorem 3 in [357] and [294], the equilibria of a closed-loop monotone system (system with feedback) can be found by examining an open-loop (input-output, or I/O system) monotone system formed by "breaking" the feedback of the original system. For fixed points to exist the I/O system must have a static input-output characteristic.

In our case, we can define $\boldsymbol{T_A}$ as the cascade of the three boxes $\boldsymbol{H_A}$, $\boldsymbol{tx_{A\to B}}$, and $\boldsymbol{rx_B}$ having input R_A and output R_B. With $\boldsymbol{T_B}$ defined similarly for the other three boxes, then the entire system "broken" at R_A is the cascade of $\boldsymbol{T_A}$ and $\boldsymbol{T_B}$. This new I/O system accepts an input $u \in \mathbb{R}_+$ and produces an output $y \in \mathbb{R}_+$. If the I/O system admits a static input-output characteristic, then the points where $u = y$ are the steady-state solutions to the closed-loop system.

Since $\boldsymbol{H_A}$ and the combined cascade of $\boldsymbol{tx_{A\to B}}$ and $\boldsymbol{rx_B}$ each have a unique steady-state solution that is a global and asymptotically stable hyperbolic equilibrium [332], then the cascade of the three systems, i.e., $\boldsymbol{T_A}$, also has a static input-output characteristic. $\boldsymbol{T_B}$ is similarly endowed. Therefore the cascade of $\boldsymbol{T_A}$ and $\boldsymbol{T_B}$ also has a static input-output

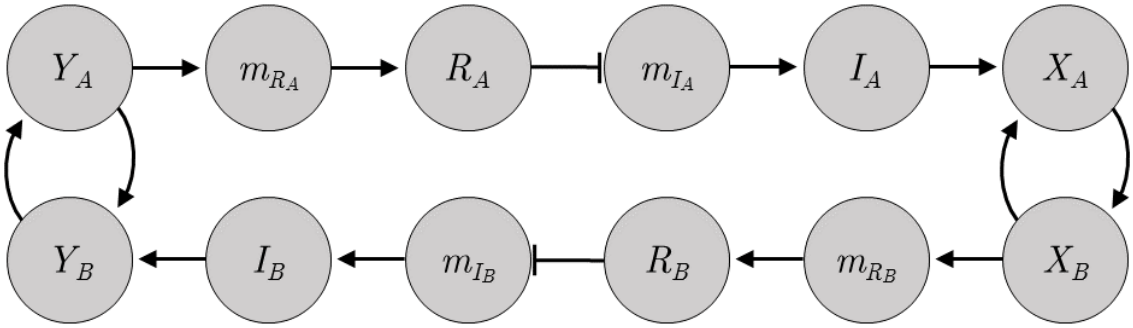

Figure 3.1: Schematic of the graph representing molecular interactions in the two-cell, closed-loop system. Sharp arrows indicate promotion, or positive parity; flat arrows indicate inhibition/repression, or negative parity. A cycle is any undirected sequence of edges and nodes beginning at one element and ending at the same element (i.e., ignoring direction of arrows). The parity of a cycle is the parity of the product of the signs of all edges traversed to complete the cycle. Since every cycle in the graph is positive in parity, the closed-loop system is monotone. To ensure that the corresponding open-loop system is strongly monotone we require a directed path (i.e., following the arrows) to exist between the input node and every other node, and between every node and the output node. The input and output nodes are determined by where the feedback loop is broken. In this system every element is reachable from every other element, so these conditions will be satisfied regardless of where the cycle is broken. Hence the open-loop (input-output) system is also monotone.

characteristic, and because the I/O system is also monotone, then its fixed points and the equilibrium points of the closed-loop system correspond.

Define $T_A(\cdot) : \mathbb{R} \to \mathbb{R}$ to be the static I/O characteristic of $\boldsymbol{T_A}$, i.e., for constant input $R^*_{A_i}$, $\boldsymbol{T_A}$ produces constant output $R^*_{B_o} = T_A(R^*_{A_i})$, and define $T_B(\cdot) : \mathbb{R} \to \mathbb{R}$ to be the static I/O characteristic of $\boldsymbol{T_B}$ for constant input $R^*_{B_i}$ and output $R^*_{A_o}$. Then the static I/O characteristic for the cascade of $\boldsymbol{T_A}$ and $\boldsymbol{T_B}$ is $T_B(T_A(\cdot))$, which maps constant input $R^*_{A_i}$ to output $R^*_{A_o}$. The function $T_B(T_A(\cdot))$ is nonnegative and sigmoidal, meaning that there must be exactly one or three intersection points between $y = u$ and $y = T_B(T_A(u))$ (the fixed-point solutions to the I/O system where $u = y$ or $R^*_{A_i} = R^*_{A_o}$). If there is only one intersection then the I/O characteristic at the intersection must have a slope less than unity ($T'_B(T_A(y^*))T'_A(y^*) < 1$ where y^* is the intersection point), implying that the corresponding equilibrium is stable. If there are three intersections, then the middle intersection must have a slope greater than one while the higher and lower intersections must have slopes less than one, implying that the middle equilibrium is unstable and the other two equilibria are stable [357]. Hence determining bistability amounts to graphically counting the intersections between $y = u$ and $y = T_B(T_A(u))$. (We could carry out the above analysis with equivalent results for $T_A(T_B(\cdot))$ mapping $R^*_{B_i}$ to $R^*_{B_o}$.)

3.1.2 Multicompartmental systems with symmetry

Up to this point we have considered a system of only two compartments, one each of Strains A and B. The formation of an interesting pattern, however, requires more than two elements. To that end we will now consider a class of systems with multiple compartments of each type, where each compartment of Strain A is connected to the same number of compartments of Strain B, and vice versa. In other words, each compartment of Strain A or B is essentially indistinguishable from any other. The symmetry present in such a system will allow us to apply the graphical method of analyzing bistability to sets of compartments arranged in particular geometries.

The interior mechanics of all compartments of the same type are the same, so the ODEs governing the behavior of the receiver boxes $\boldsymbol{rx}$ and production boxes $\boldsymbol{H}$ are the same for compartments of the same type. Adding more compartments does, however, change the concentration of diffusible signaling molecules that reach the compartments, and therefore changes the behavior of the transceiver box $\boldsymbol{tx}$.

We begin by noting that if we define the 2×2 matrix

$$L_2 := d \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \tag{3.14}$$

then we can reformulate (3.7) as

$$\dot{Y} = L_2 Y - \gamma_Y Y + \begin{bmatrix} 0 \\ \nu_Y I_B \end{bmatrix} \tag{3.15}$$

with steady-state solution

$$Y^* = (-L_2 + \gamma_Y I)^{-1} \begin{bmatrix} 0 \\ \nu_Y I_B^* \end{bmatrix}. \tag{3.16}$$

The matrix L_2 contains information on the diffusion of HSL between compartments. We now generalize to systems with N compartments and the corresponding $N \times N$ matrix L_N (which we will henceforth designate as simply L). If all compartments in the system are numbered from 1 to N, then the element $[L]_{ij}$ represents the connection strength between compartments i and j. Conceptually, the connection strength is the diffusion into compartment j from compartment i (or vice versa) if $i \neq j$, and the total diffusion out of a given compartment to all other compartments if $i = j$. Each element $[L]_{ij}$ is directly proportional to the diffusivity D of HSL and inversely proportional to the square of the distance between i and j. In the special case where there are only two compartments, the connection strength between them is identical, hence the elements in L_2 are all of magnitude d.

Mathematically, we can represent the multicompartmental system as an undirected graph where each vertex is a compartment and each edge is a channel. Let d_{ij} be the edge weight between vertices i and j. L is the Laplacian of this graph:

$$[L]_{ij} = \begin{cases} -\sum_{j=1}^{N} d_{ij} & i = j \\ d_{ij} & i \neq j \end{cases}. \tag{3.17}$$

Let N_A be the number of compartments of Strain A and N_B the number of compartments of Strain B such that $N_A + N_B = N$. Assume the diffusivity of HSL is constant, all channels have the same length, and compartments of the same type are not connected to each other. Then $d_{ij} = 0$ if i, j are of the same type and $d_{ij} = d$ between connected compartments i, j of opposite strains. Assume each compartment of Strain A is connected to q_B compartments of Strain B and each compartment of Strain B is connected to q_A compartments of Strain A.

Let the first N_A entries of a row or column of $L \in \mathbb{R}^{N\times N}$ designate compartments of Strain A and the last N_B entries designate compartments of Strain B. Then L has the form

$$L = d\begin{bmatrix} -q_B I_{N_A\times N_A} & F \\ F^T & -q_A I_{N_B\times N_B} \end{bmatrix} \tag{3.18}$$

where F is an $N_A \times N_B$ matrix for which $[F]_{ij} = 0$ indicates that the ith compartment of Strain A and the jth compartment of Strain B are not connected by a channel, and $[F]_{ij} = 1$ indicates that they are.

As shown in the following sections, the assumed structure of the system allows us to reduce our N-dimensional system to a two-dimensional one, which greatly simplifies the calculations for a steady-state contrasting pattern and enables us to use the graphical method introduced in the previous section to determine when the overall system is bistable.

Define $M \in \mathbb{R}^{N\times 2}$ as

$$M := \begin{bmatrix} \mathbb{1}_{N_A} & 0_{N_A} \\ 0_{N_B} & \mathbb{1}_{N_B} \end{bmatrix} \tag{3.19}$$

where $\mathbb{1}_n$ designates a length-n vector of all ones and 0_n designates a length-n vector of all zeros. Then because all compartments of the same type have the same number of connections to compartments of the opposite type, there exists some $\bar{L} \in \mathbb{R}^{2\times 2}$ such that

$$LM = M\bar{L}. \tag{3.20}$$

Let $d_A := dq_B$ be the (nonnegative) total outgoing edge weight for a compartment of Strain A and $d_B := dq_A$ be the (nonnegative) total outgoing edge weight for a compartment of Strain B. Because we have assumed no connections between compartments of the same type, $\bar{L}$ has the form

$$\bar{L} = \begin{bmatrix} -d_A & d_A \\ d_B & -d_B \end{bmatrix}. \tag{3.21}$$

We would like to solve for the steady-state value of $X = [X_A,\ X_B]^T \in \mathbb{R}^N$ in $\boldsymbol{tx_{A\to B}}$:

$$\left(-L + \gamma_X I_{N\times N}\right) X = \begin{bmatrix} \nu_X I_A \mathbb{1}_{N_A} \\ 0_{N_B} \end{bmatrix}. \tag{3.22}$$

We restrict our search to a subset of solutions for which the variables of interest are identical among compartments of the same type, i.e.,

$$X = M\begin{bmatrix} x_A \\ x_B \end{bmatrix} \tag{3.23}$$

for $x_A, x_B \in \mathbb{R}$. Let $x := [x_A,\ x_B]^T$. We can then rewrite (3.22) as

$$(-L + \gamma_X I_{N\times N})\, Mx = M \begin{bmatrix} \nu_X I_A \\ 0 \end{bmatrix} \tag{3.24}$$

$$\implies M\left(-\bar{L} + \gamma_X I_{2\times 2}\right) x = M \begin{bmatrix} \nu_X I_A \\ 0 \end{bmatrix} \tag{3.25}$$

which implies that solutions x^* to

$$\left(-\bar{L} + \gamma_X I\right) x^* = \begin{bmatrix} \nu_X I_A^* \\ 0 \end{bmatrix} \tag{3.26}$$

$$\implies x^* = \left(-\bar{L} + \gamma_X I\right)^{-1} \begin{bmatrix} \nu_X I_A^* \\ 0 \end{bmatrix} \tag{3.27}$$

provide solutions to (3.22) by way of (3.23). In other words, x_A^* is the steady-state concentration of X in any compartment of Strain A and x_B^* is the steady-state concentration of X in any compartment of Strain B for constant input I_A^*. The derivation for $\boldsymbol{tx_{B\to A}}$ proceeds similarly.

Now recall from (3.21) that the matrix $\bar{L}$ has form

$$\bar{L} = \begin{bmatrix} -d_A & d_A \\ d_B & -d_B \end{bmatrix}. \tag{3.28}$$

Then the inverse matrix in (3.27) can be directly evaluated, yielding

$$\left(-\bar{L} + \gamma_X I\right)^{-1} = \frac{1}{(d_A+\gamma_X)(d_B+\gamma_X) - d_A d_B} \begin{bmatrix} d_B+\gamma_X & d_A \\ d_B & d_A+\gamma_X \end{bmatrix} \tag{3.29}$$

$$= \frac{1}{\gamma_X(\gamma_X + d_A + d_B)} \begin{bmatrix} d_B+\gamma_X & d_A \\ d_B & d_A+\gamma_X \end{bmatrix}. \tag{3.30}$$

Since $d_A,\ d_B > 0$, the matrix is always invertible provided that $\gamma_X \neq 0$.

In essence, the new multicompartmental system is identical to the two-compartmental system with a revision to the transceiver boxes:

$$\boldsymbol{tx_{B\to A}}\ :\ \begin{cases} \dot{Y}_A = d_A(Y_B - Y_A) - \gamma_Y Y_A \\ \dot{Y}_B = d_B(Y_A - Y_B) - \gamma_Y Y_B + \nu_Y I_B \end{cases} \tag{3.31}$$

and similarly for $\boldsymbol{tx_{A\to B}}$. Since $d_A, d_B > 0$ the Jacobian equations from (3.12) maintain their parity and the monotonicity of the system is preserved.

3.1.3 Contrasting patterns may emerge if and only if the full system can be reduced to a two-compartment equivalent

We have just seen that if compartments of the same type have the same number of neighbors, then the system can be reduced to a two-compartment case for the purposes of

identifying contrasting steady states, i.e., those in which all compartments of the same type have the same steady-state concentrations. In other words, contrasting steady states emerge if the system can be reduced to the two-compartment case. We now show that contrasting steady states in a full system can *only* emerge if that system can be reduced to a two-compartment case. This allows us to conclude that stable contrasting steady states may emerge *if and only if* the system can be reduced to an equivalent two-compartment system that is bistable.

The first claim shows that if a steady state exists for which (a) all compartments of the same type have the same concentrations and (b) every compartment of Strain A is connected to at least one compartment of Strain B and vice versa, then the underlying edge weight matrix L can be reduced to a two-compartment system. We do not need to assume that compartments of the same type must have neighbors *only* of the opposite type, although a "checkerboard" of alternating high/low immediately adjacent neighbors will require this arrangement. We do not even require that all channels be of the same length. We only require that for a fixed compartment type, the total incoming edge weight from other compartments of the same type as well as the incoming edge weight from other compartments of the opposite type is the same regardless of the choice of compartment. The second claim shows that the (in)stability of the reduced system steady states implies (in)stability of the corresponding full system steady states.

For both claims, by "full system" we refer to a system of $N = N_A + N_B$ compartments with N_A of strain A each independently governed by equations (3.1), (3.3), (3.5) and N_B of strain B each independently governed by equations (3.6), (3.7), (3.8), with interconnection matrix L.

Claim 1. *Assume every compartment of type A is connected to at least one compartment of type B and vice versa. If there exists a steady-state solution to the full system of the form*

$$X^* = \begin{bmatrix} X_A^* \\ X_B^* \end{bmatrix} = \begin{bmatrix} \mathbb{1}_{N_A} & 0_{N_A} \\ 0_{N_B} & \mathbb{1}_{N_B} \end{bmatrix} \begin{bmatrix} x_A^* \\ x_B^* \end{bmatrix} =: Mx^* \tag{3.32}$$

where $x_A, x_B \in \mathbb{R}$ as in (3.23)*, then there exists a unique $\bar{L} \in \mathbb{R}^{2\times 2}$ such that $LM = M\bar{L}$.*

Proof. X^* a steady-state solution implies that

$$(-L + \gamma_X I_N)\, X^* = \begin{bmatrix} \nu_X I_A^* \mathbb{1}_{N_A} \\ 0_{N_B} \end{bmatrix} \tag{3.33}$$

$$\implies (-L + \gamma_X I_N)\, Mx^* = M \begin{bmatrix} \nu_X I_A^* \\ 0 \end{bmatrix} \tag{3.34}$$

since there cannot be multiple values of I_A^* that result in the same X_A^*. Now define

$$L = \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix} \tag{3.35}$$

where $L_{11} \in \mathbb{R}^{N_A \times N_A}$, $L_{12} = L_{21}^T \in \mathbb{R}^{N_A \times N_B}$, and $L_{22} \in \mathbb{R}^{N_B \times N_B}$, and observe that

$$(-L + \gamma_X I_N) M = \begin{bmatrix} -\sum_{j=1}^{N_A} [L_{11}]_j + \gamma_X \mathbb{1}_{N_A} & -\sum_{j=1}^{N_B} [L_{12}]_j \\ -\sum_{j=1}^{N_A} [L_{21}]_j & -\sum_{j=1}^{N_B} [L_{22}]_j + \gamma_X \mathbb{1}_{N_B} \end{bmatrix} \tag{3.36}$$

where $[L_{nm}]_j$ denotes the jth column of L_{nm}, $n, m = 1, 2$. Now define

$$Q_{nm}^i = \begin{cases} -\sum_{j=1}^{N_A} [L_{nm}]_{ij}, & m = 1 \\ -\sum_{j=1}^{N_B} [L_{nm}]_{ij}, & m = 2 \end{cases}. \tag{3.37}$$

Combining (3.36) with (3.34) yields

$$\begin{cases} (Q_{11}^a + \gamma_X) x_A^* + (Q_{12}^a) x_B^* = \nu_X I_A^*, & a = 1, 2, ..., N_A \\ \left(Q_{21}^b\right) x_A^* + \left(Q_{22}^b + \gamma_X\right) x_B^* = 0, & b = N_A + 1, ..., N_A + N_B \end{cases} \tag{3.38}$$

and similarly for Y^*,

$$\begin{cases} (Q_{11}^a + \gamma_Y) y_A^* + (Q_{12}^a) y_B^* = 0, & a = 1, 2, ..., N_A \\ \left(Q_{21}^b\right) y_A^* + \left(Q_{22}^b + \gamma_Y\right) y_B^* = \nu_Y I_B^*, & b = N_A + 1, ..., N_A + N_B \end{cases}. \tag{3.39}$$

Suppose we fix a and b and construct a system of four equations from (3.38) and (3.39). In these four equations there are four "unknowns" $(Q_{11}^a, Q_{12}^a, Q_{21}^b, Q_{22}^b)$, suggesting that there is a unique solution to

$$\begin{bmatrix} x_A^* & x_B^* & 0 & 0 \\ 0 & 0 & x_A^* & x_B^* \\ y_A^* & y_B^* & 0 & 0 \\ 0 & 0 & y_A^* & y_B^* \end{bmatrix} \begin{bmatrix} Q_{11}^a \\ Q_{12}^a \\ Q_{21}^b \\ Q_{21}^b \end{bmatrix} = \begin{bmatrix} \nu_X I_A^* - \gamma_X x_A^* \\ -\gamma_X x_B^* \\ -\gamma_Y y_A^* \\ \nu_Y I_B^* - \gamma_Y y_B^* \end{bmatrix} \tag{3.40}$$

provided that the LHS matrix is full rank, i.e., $\frac{x_A^*}{x_B^*} \neq \frac{y_A^*}{y_B^*}$. This unique solution is the same regardless of the choice of a and b, which implies that Q_{11}^a, Q_{12}^a must be the same for all a and Q_{21}^b, Q_{22}^b must be the same for all b (since the solution is unique). This, in turn, implies that we can write

$$LM = M \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} =: M\bar{L}. \tag{3.41}$$

The only fact left to show, then, is that our system cannot admit a solution for which $\frac{x_A^*}{x_B^*} = \frac{y_A^*}{y_B^*}$, which will guarantee a unique solution to (3.40). Our proof leverages the lateral inhibition structure of the system.

Suppose $\frac{x_A^*}{x_B^*} = \frac{y_A^*}{y_B^*}$ and without loss of generality, define $r := \frac{x_A^*}{y_A^*} = \frac{x_B^*}{y_B^*}$. Substituting $x_A^* = r y_A^*$ and $x_B^* = r y_B^*$ into (3.38) and (3.39) for fixed a, b (equivalently (3.40)), we obtain

$$\begin{cases} (Q_{11}^a + \gamma_X) x_A^* + Q_{12}^a x_B^* = \nu_X I_A^* \\ (Q_{11}^a + \gamma_Y) x_A^* + Q_{12}^a x_B^* = 0 \\ Q_{21}^b x_A^* + \left(Q_{22}^b + \gamma_X\right) x_B^* = 0 \\ Q_{21}^b x_A^* + \left(Q_{22}^b + \gamma_Y\right) x_B^* = r\nu_Y I_B^* \end{cases} \implies \begin{cases} x_A^* = \frac{-\left(Q_{22}^b + \gamma_X\right) x_B^*}{Q_{21}^b} = \frac{r\nu_Y I_B^* - \left(Q_{22}^b + \gamma_Y\right) x_B^*}{Q_{21}^b} \\ x_B^* = \frac{\nu_X I_A^* - \left(Q_{11}^a + \gamma_X\right) x_A^*}{Q_{12}^a} = \frac{-\left(Q_{11}^a + \gamma_Y\right) x_A^*}{Q_{12}^a} \end{cases}. \tag{3.42}$$

The assumption that every compartment of Strain A is connected to at least one cell of Strain B and vice versa guarantees that $Q^a_{12} \neq 0, Q^b_{21} \neq 0$. Rearranging terms, we find

$$\begin{cases} (\gamma_Y - \gamma_X)\, x^*_B = r\nu_Y I^*_B \\ (\gamma_X - \gamma_Y)\, x^*_A = \nu_Y I^*_A \end{cases} . \tag{3.43}$$

Now we know $\gamma_X, \gamma_Y, \nu_X, \nu_Y > 0$ and since $x^*_A, x^*_B, I^*_A, I^*_B \geq 0$ (we cannot have negative concentrations) we also know $r \geq 0$. Because of the mutual inhibition relationship we further know that $I^*_A =: h_1(rx^*_A)$ is a bounded, nonnegative, nonincreasing function of x^*_A and $I^*_B =: h_2(x^*_B)$ is a bounded, nonnegative, nonincreasing function of x^*_B. If we are concerned with a nontrivial system we must have $r \neq 0$ because otherwise

$$r = 0 \implies x^*_A = 0 \implies I^*_A = 0 \implies h_1(0) = 0 \implies h_1(z \geq 0) = 0 \ \forall z, \tag{3.44}$$

i.e., cells of Strain A are insensitive to inputs. Similarly, we must have r finite because we could as easily have set up the system in terms of y^*_A and y^*_B with $\frac{1}{r}$ as the ratio of interest, and following the same logic we would have (with slight abuse of notation) $\frac{1}{r} \neq 0$.

Since $r > 0$, then from (3.43), nonnegative x^*_B, I^*_B requires $\gamma_Y \geq \gamma_X$ while nonnegative x^*_A, I^*_A requires $\gamma_X \geq \gamma_Y$. These two conditions can only be satisfied if $\gamma_X = \gamma_Y$, which implies that $I^*_A = I^*_B = 0$. But this is impossible from the definition of h_1, h_2 in (3.2) (except in the limit as $R^*_A, R^*_B \to \infty$, which is anyway unattainable because R^*_A, R^*_B are bounded). Hence we cannot have a steady-state contrasting solution for which $\frac{x^*_A}{x^*_B} = \frac{y^*_A}{y^*_B}$, implying that the matrix in (3.40) will have a unique solution. This completes our claim. ■

We now show that the local stability or instability of a steady state in the reduced system implies local stability or instability of the state in the full system, and therefore we can continue to use the graphical intersection method to determine when the system admits a solution where all cells of the same type are identically (and reversibly) high or low.

Claim 2. *Let J represent the Jacobian at a steady state in the full system and $\bar{J}$ represent the Jacobian for the corresponding steady state in the reduced system. If $\bar{J}$ is contractive, then J is also contractive, thus the steady state is locally stable in both reduced and full systems. Otherwise, the reduced steady state is unstable, and the corresponding state in the full system is also unstable.*

Proof. Let $Z \in \mathbb{R}^{6N}_+$ be the vector of states for the full system where $\dot{Z} = f(Z)$. Let Z^* be a steady state where all cells of the same type have identical states, i.e., a steady state

identified in the reduced system. Assume the states are ordered such that

$$Z^* := \begin{bmatrix} Y^* \\ m^*_{R_A}\mathbb{1}_{N_A} \\ R^*_A\mathbb{1}_{N_A} \\ m^*_{I_A}\mathbb{1}_{N_A} \\ I^*_A\mathbb{1}_{N_A} \\ X^* \\ m^*_{R_B}\mathbb{1}_{N_B} \\ R^*_B\mathbb{1}_{N_B} \\ m^*_{I_B}\mathbb{1}_{N_B} \\ I^*_B\mathbb{1}_{N_B} \end{bmatrix}. \tag{3.45}$$

Define

$$K_{A1} := \left.\frac{\partial \dot{m}_{R_A}}{\partial Y_A}\right|_{Y^*_A} = V_{R_A} N_{R_A} C \frac{n_Y \left(\frac{Y^*_A}{K_Y}\right)^{n_Y - 1}}{K_Y \left(1 + \left(\frac{Y^*_A}{K_Y}\right)^{n_Y}\right)^2} \tag{3.46}$$

$$K_{A2} := -\left.\frac{\partial \dot{m}_{I_A}}{\partial R_A}\right|_{R^*_A} = V_{I_A} N_{I_A} C \frac{n_{R_A} \left(\frac{R^*_A}{K_{R_A}}\right)^{n_{R_A} - 1}}{K_{R_A} \left(1 + \left(\frac{R^*_A}{K_{R_A}}\right)^{n_{R_A}}\right)^2}. \tag{3.47}$$

K_{B1} and K_{B2} are defined analogously with the appropriate subscripts for cell B. Then the Jacobian of the full system evaluated at the steady state Z^* is given by

$$J = \begin{bmatrix} S_A & P_Y \\ P_X & S_B \end{bmatrix} \tag{3.48}$$

where $J \in \mathbb{R}^{6N \times 6N}$, $S_A \in \mathbb{R}^{N+4N_A \times N+4N_A}$, $S_B \in \mathbb{R}^{N+4N_B \times N+4N_B}$, $P_Y \in \mathbb{R}^{N+4N_A \times N+4N_B}$, and

$P_X \in \mathbb{R}^{N+4N_B \times N+4N_A}$, defined as

$$S_A := \begin{bmatrix} L-\gamma_Y I_N & 0 & 0 & 0 & 0 \\ \begin{bmatrix} K_{A1} I_{N_A} & 0 \end{bmatrix} & -\gamma_{m_{R_A}} I_{N_A} & 0 & 0 & 0 \\ 0 & \epsilon_{R_A} I_{N_A} & -\gamma_{R_A} I_{N_A} & 0 & 0 \\ 0 & 0 & -K_{A2} I_{N_A} & -\gamma_{m_{I_A}} I_{N_A} & 0 \\ 0 & 0 & 0 & \epsilon_{I_A} I_{N_A} & -\gamma_{I_A} I_{N_A} \end{bmatrix} \tag{3.49}$$

$$P_Y := \begin{bmatrix} 0_{N \times N+3N_B} & \begin{bmatrix} 0 \\ \nu_Y I_{N_B} \end{bmatrix} \\ 0_{4N_A \times N+3N_B} & 0_{4N_A \times N_B} \end{bmatrix} \tag{3.50}$$

$$P_X := \begin{bmatrix} 0_{N \times N+3N_A} & \begin{bmatrix} \nu_X I_{N_A} \\ 0 \end{bmatrix} \\ 0_{4N_B \times N+3N_A} & 0_{3N_A \times N_A} \end{bmatrix} \tag{3.51}$$

$$S_B := \begin{bmatrix} L-\gamma_X I_N & 0 & 0 & 0 & 0 \\ \begin{bmatrix} 0 & K_{B1} I_{N_B} \end{bmatrix} & -\gamma_{m_{R_B}} I_{N_B} & 0 & 0 & 0 \\ 0 & \epsilon_{R_B} I_{N_B} & -\gamma_{R_B} I_{N_B} & 0 & 0 \\ 0 & 0 & -K_{B2} I_{N_B} & -\gamma_{m_{I_B}} I_{N_B} & 0 \\ 0 & 0 & 0 & \epsilon_{I_B} I_{N_B} & -\gamma_{I_B} I_{N_B} \end{bmatrix} \tag{3.52}$$

The matrix measure of a matrix M with respect to the one-norm is defined as

$$\mu(M) := \max_j \left\{ M_{jj} + \sum_{i \neq j} |M_{ij}| \right\}. \tag{3.53}$$

If there exists an invertible diagonal matrix D such that

$$\mu(DJD^{-1}) < 0, \tag{3.54}$$

then the mapping described by J is contractive. This implies that the eigenvalues are negative, which for our nonlinear system means that the steady state around which J is linearized is locally stable.

Consider the reduced version of the full system, which has a Jacobian $\bar{J}$ identical in form to (3.48) for $N_A = 1$, $N_B = 1$, and

$$L = \bar{L} := \begin{bmatrix} -d_A & d_A \\ d_B & -d_B \end{bmatrix} \tag{3.55}$$

where d_A is the total (nonnegative) outgoing edge weight for a cell of Strain A and d_B is the total (nonnegative) outgoing edge weight for a cell of Strain B.

Let $\bar{D}$ take the form

$$\bar{D} = diag\left(l_{1A}, l_{1B}, a_1, a_2, a_3, a_4, l_{2A}, l_{2B}, b_1, b_2, b_3, b_4\right) \tag{3.56}$$

where $l_{1A}, l_{1B}, a_1, a_2, a_3, a_4, l_{2A}, l_{2B}, b_1, b_2, b_3, b_4 \in \mathbb{R}$ are arbitrary constants. For $\bar{D}$ to satisfy (3.54), we require

$$-\gamma_Y + \frac{a_1}{l_{1A}} K_{A1} - d_A + \frac{l_{1B}}{l_{1A}} d_A < 0 \tag{3.57}$$

$$-\gamma_Y - d_B + \frac{l_{1A}}{l_{1B}} d_B < 0 \tag{3.58}$$

$$-\gamma_{m_{R_A}} + \frac{a_2}{a_1} \epsilon_{R_A} < 0 \tag{3.59}$$

$$-\gamma_{R_A} + \frac{a_3}{a_2} K_{A2} < 0 \tag{3.60}$$

$$-\gamma_{m_{I_A}} + \frac{a_4}{a_3} \epsilon_{I_A} < 0 \tag{3.61}$$

$$-\gamma_{I_A} + \frac{l_{2A}}{a_4} \nu_X < 0 \tag{3.62}$$

and analogously for columns corresponding to compartments B.

We can combine the inequalities to obtain a single expression if we rearrange them to reflect the relative sizes of certain constants. Then we obtain

$$\frac{a_1 K_{A1} + l_{1B} d_A}{\gamma_Y + d_A} < l_{1A} \tag{3.63}$$

$$l_{1A} < \frac{\gamma_Y + d_B}{d_B} l_{1B} \tag{3.64}$$

$$a_2 < \frac{\gamma_{m_{R_A}}}{\epsilon_{R_A}} a_1 \tag{3.65}$$

$$a_3 < \frac{\gamma_{R_A}}{K_{A2}} a_2 \tag{3.66}$$

$$a_4 < \frac{\gamma_{m_{I_A}}}{\epsilon_{I_A}} a_3 \tag{3.67}$$

$$l_{2A} < \frac{\gamma_{I_A}}{\nu_X} a_4. \tag{3.68}$$

We can rearrange (3.63) and (3.64) as

$$a_1 < \frac{\gamma_Y(\gamma_Y + d_B + d_A)}{K_{A1} d_B} \tag{3.69}$$

whereupon the combination of inequalities yields

$$l_{2A} < \frac{\gamma_{I_A} \gamma_{m_{I_A}} \gamma_{R_A} \gamma_{m_{R_A}} \gamma_Y (\gamma_Y + d_B + d_A)}{\nu_X \epsilon_{I_A} K_{A2} \epsilon_{R_A} K_{A1} d_B} l_{1B} =: \frac{1}{C_1} l_{1B}. \tag{3.70}$$

Similarly, for compartments B,

$$l_{1B} < \frac{\gamma_{I_B} \gamma_{m_{I_B}} \gamma_{R_B} \gamma_{m_{R_B}} \gamma_X (\gamma_X + d_B + d_A)}{\nu_Y \epsilon_{I_B} K_{B2} \epsilon_{R_B} K_{B1} d_A} l_{1A} =: \frac{1}{C_2} l_{2A} \tag{3.71}$$

such that

$$l_{2A} < \frac{1}{C_1 C_2} l_{2A} \implies C_1 C_2 < 1. \tag{3.72}$$

Incidentally,

$$C_1 C_2 = \left(\frac{\epsilon_{I_A} K_{A2}}{\gamma_{I_A}\gamma_{m_{I_A}}}\right)\left(\frac{\epsilon_{R_B} K_{B1}}{\gamma_{R_B}\gamma_{m_{R_B}}}\right)\left(\frac{d_B \nu_X}{\gamma_X(\gamma_X + d_A + d_B)}\right) \times \tag{3.73}$$

$$\left(\frac{\epsilon_{I_B} K_{B2}}{\gamma_{I_B}\gamma_{m_{I_B}}}\right)\left(\frac{\epsilon_{R_A} K_{A1}}{\gamma_{R_A}\gamma_{m_{R_A}}}\right)\left(\frac{d_A \nu_Y}{\gamma_Y(\gamma_Y + d_A + d_B)}\right) \tag{3.74}$$

can be written as

$$\left(\frac{dI_A^*}{dR_{A_i}^*}\bigg|_{R_A^*}\frac{dX_B^*}{dI_A^*}\bigg|_{I_A^*}\frac{dR_B^*}{dX_B^*}\bigg|_{X_B^*}\right)\left(\frac{dI_B^*}{dR_B^*}\bigg|_{R_B^*}\frac{dY_A^*}{dI_B^*}\bigg|_{I_B^*}\frac{dR_{A_o}^*}{dY_A^*}\bigg|_{Y_A^*}\right) = \frac{dR_{A_o}^*}{dR_{A_i}^*}\bigg|_{R_A^*}, \tag{3.75}$$

where $R_{A_i}^*$ is a constant input and $R_{A_o}^*$ is the associated output. Hence (3.75) is the slope of the I/O system evaluated at the steady state $R_{A_i}^* = R_A^*$, which matches the graphical stability test.

Now consider the full system with Jacobian (3.48) corresponding to the reduced system for which $\bar{D}$ satisfies (3.54). We construct a D for the full system that has diagonal entries

$$[l_{1A}I_{N_A}, l_{1B}I_{N_A}, a_1 I_{N_A}, a_2 I_{N_A}, a_3 I_{N_A}, a_4 I_{N_A}, l_{2A}I_{N_B}, l_{2B}I_{N_B}, b_1 I_{N_B}, b_2 I_{N_B}, b_3 I_{N_B}, b_4 I_{N_B}] \tag{3.76}$$

where the constants are the same as for $\bar{D}$. DJD^{-1} is then effectively organized into "blocks" that correspond to columns of $\bar{D}\bar{J}\bar{D}^{-1}$. Let

$$m_j(M) := M_{jj} + \sum_{i \neq j} |M_{ij}|, \tag{3.77}$$

i.e., $\mu(M) = \max_j m_j(M)$. Because J is linearized about a reduced-system steady state, it is

straightforward to see that

$$m_j(J) = \begin{cases} m_1(\bar{J}), & j = 1, 2, ..., N_A \\ m_2(\bar{J}), & j = N_A + 1, N_A + 2, ..., N \\ m_3(\bar{J}), & j = N + 1, N + 2, ..., N + N_A \\ m_4(\bar{J}), & j = N + N_A + 1, ..., N + 2N_A \\ m_5(\bar{J}), & j = N + 2N_A + 1, ..., N + 3N_A \\ m_6(\bar{J}), & j = N + 3N_A + 1, ..., N + 4N_A \\ m_7(\bar{J}), & j = N + 4N_A + 1, ..., N + 5N_A \\ m_8(\bar{J}), & j = N + 5N_A + 1, ..., 2N + 4N_A \\ m_9(\bar{J}), & j = 2N + 4N_A + 1, ..., 3N + 3N_A \\ m_{10}(\bar{J}), & j = 3N + 3N_A + 1, ..., 4N + 2N_A \\ m_{11}(\bar{J}), & j = 4N + 2N_A + 1, ..., 5N + N_A \\ m_{12}(\bar{J}), & j = 5N + N_A, ..., 6N \end{cases} . \tag{3.78}$$

(Note that the equivalence of $m_j(J)$ to $m_1(\bar{J})$, $m_2(\bar{J})$, $m_7(\bar{J})$, and $m_8(\bar{J})$ for appropriate j arises out of the form of L, which has diagonal entries $-d_A$ and $-d_B$ repeated N_A and N_B times respectively, with non-diagonal entries summing to d_A for the first N_A columns and d_B for the last N_B columns.)

Because $m_j(\bar{J}) < 0$ for $j = 1, 2, ..., 12$, then $m_j(J) < 0$ for $j = 1, 2, ..., 6N$. Therefore for J linearized about a given reduced-system steady state, the conditions required for J to be contractive—and therefore for the full system to be locally stable at the steady state—is the same as the graphical condition for the reduced system to be stable or unstable at that steady state, namely, that the slope of the I/O system must be less than 1. Conversely, if (3.75) is greater than 1, the steady state in the reduced-system subspace is unstable and so the full system (which contains the reduced-system subspace) cannot be stable. Therefore the graphical test for the stability or instability of steady states in the reduced system is sufficient to determine the local stability or instability of corresponding steady states in the full system. ■

Note that when J is a contractive map, the steady state is locally exponentially stable because the real parts of all eigenvalues are negative, therefore sufficiently small perturbations to system behavior will not destroy the stability of the steady state.

We note this proof only suffices to show local stability of the steady state in the full system. Other solutions representing different spatial patterns may exist and be stable in the full system even though the system reduction as performed herein would not identify them. Equivalently, the fact that the bistable steady state is global in the reduced system but only local in the full system exemplifies how introducing more compartments enlarges the space of possible solutions beyond those that exist in lower-dimensional regimes.

3.1.4 Finite differences within channels

Thus far, implicit in our definition of L is that we approximate diffusion between compartments using the method of finite differences for a step size of l, the channel length. Specifically, we have taken Fisk's diffusion equation in one dimension

$$\frac{dX}{dt} = D\frac{d^2X}{dr^2} \tag{3.79}$$

where X is the concentration of some species and r is distance. Under the finite differences approximation, we discretize space along r and approximate the change in concentration at each point as

$$\frac{dX}{dt} \approx D\frac{(X(r+\Delta) - X(r)) + (X(r-\Delta) - X(r))}{\Delta^2}. \tag{3.80}$$

In the process of discretizing r we pick two boundary points. To model a single channel connecting a cell of type A to one of type B we simply pick one end, say $r = 0$, to correspond to A and the other end, $r = l$, to correspond to B. We assume no diffusion outside the channel. Since $X(0)$ has no neighbors $r < 0$ and $X(l)$ has no neighboring points $r > l$, the approximation at the boundaries is performed using only one difference. At the $r = 0$ boundary we have

$$\frac{dX(0)}{dt} \approx D\frac{X(\Delta) - X(0)}{\Delta^2} \tag{3.81}$$

and at the $r = l$ boundary

$$\frac{dX(l)}{dt} \approx D\frac{X(l-\Delta) - X(l)}{\Delta^2}. \tag{3.82}$$

If we let $\Delta = l$, $X(0) = X_A$, and $X(l) = X_B$, we recover the familiar

$$\begin{cases} \frac{dX_A}{dt} = \frac{D}{l^2}(X_B - X_A) \\ \frac{dX_B}{dt} = \frac{D}{l^2}(X_A - X_B) \end{cases}. \tag{3.83}$$

If l is sufficiently small, this approximation is appropriate. In practice, however, if the channel is too long, then setting $\Delta = l$ produces an extremely coarse approximation that may not be an accurate description of the physical process. Suppose now we discretize r such that there are N_L points between $r = 0$ and $r = l$, with a step size of $\Delta = \frac{l}{N_L+1}$. The equation to describe diffusion is written in matrix form as

$$\begin{bmatrix} \dot{X}(0) \\ \dot{X}(\Delta) \\ \dot{X}(2\Delta) \\ \vdots \\ \dot{X}(l-\Delta) \\ \dot{X}(l) \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & 0 & ... & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & ... & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & ... & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & ... & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & ... & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} X(0) \\ X(\Delta) \\ X(2\Delta) \\ \vdots \\ X(l-\Delta) \\ X(l) \end{bmatrix}. \tag{3.84}$$

Note that this describes only the diffusion, not the reactions that occur within $X(0)$ and $X(l)$.

We can accommodate multiple cells with multiple channels by adding more terms and associated concentration variables. Imagine that we number the discretized points within a channel from 1 to N_L where 1 is the point nearest to a cell of type A and N_L is the point nearest to a cell of type B. Then we assume all channels with the same number form a similar class; i.e., if we have N_C channels we define $M \in \mathbb{R}^{N+N_C N_L \times N_C+2}$ as

$$M := \begin{bmatrix} \mathbb{1}_{N_A} & 0_{N_L} & 0_{N_L} & ... & 0_{N_L} \\ 0_{N_L} & \mathbb{1}_{N_L} & 0_{N_L} & ... & 0_{N_L} \\ 0_{N_L} & 0_{N_L} & \mathbb{1}_{N_L} & ... & 0_{N_L} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_{N_L} & 0_{N_L} & 0_{N_L} & ... & \mathbb{1}_{N_B} \end{bmatrix} \tag{3.85}$$

such that

$$\bar{L} := \frac{D}{\Delta^2} \begin{bmatrix} -q_B & q_B & 0 & 0 & ... & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & ... & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & ... & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & ... & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & ... & 0 & -q_A & q_A \end{bmatrix} \tag{3.86}$$

where as before, q_B and q_A are the number of compartments to which Strains A and B respectively are connected. We then rewrite (3.22) as

$$\left(-\bar{L} + \gamma_X I\right) X = \begin{bmatrix} \nu_X I_A \\ 0_{N_L N_C+1} \end{bmatrix}. \tag{3.87}$$

The equation takes a similar form for Y.

In the above discussion we have assumed that production takes place only in single compartments, which shrink in size with increasing N_L. Let w be the width of a cell colony. To accurately reckon with arbitrary l and N_C we would have to either (a) dilute the concentrations of intercellular components ($l > w$); or (b) designate multiple compartments as production compartments, which would rapidly increase the size of the ODE system ($l < w$). To avoid these situations we set $\Delta = w$ such that $l = kw$ where $k \geq 0$ is an integer. This limitation in the considered range of l is justified by the small range of l that we implement expermientally as well as the minimal effect of l on the theoretically evaluated bistability of the system (for systems with and without multiple diffusive compartments).

3.2 Computational Results

Contrasting patterns result from disparity in the steady-state target gene expression between cross-repressive strains in a DLI system. We assessed the effect of changing the pa-

rameters in Table 3.2 on patterning in the system for a given connectivity of compartments—hereafter referred to as a geometry—satisfying the conditions for reduction to an equivalent two-compartment model. The types of changes leading to bifurcations in the two-strain circuits are summarized in Figure 3.3.

Ultrasensitivity (cooperativity) in $T_A(\cdot)$ and $T_B(\cdot)$ is necessary for bistable contrast. Proper kinetic rate matching ensures that ultrasensitivity is preserved in the feedback loop [358]. This equates to a condition on dynamic range matching between two bistable switches: the minimum output from one switch should be low enough to induce close-to-highest expression levels of the second switch while the maximum output should be high enough to induce close-to-lowest expression levels, and vice versa. The curves in Figure 3.2 show input/output curves (nullclines) for the first switch with input X_B and output Y_A (boxes $\mathbf{rx_B} \rightarrow \mathbf{H_B} \rightarrow \mathbf{tx_{B\rightarrow A}}$) and the second switch with input Y_A and output X_B (boxes $\mathbf{rx_A} \rightarrow \mathbf{H_A} \rightarrow \mathbf{tx_{A\rightarrow B}}$).

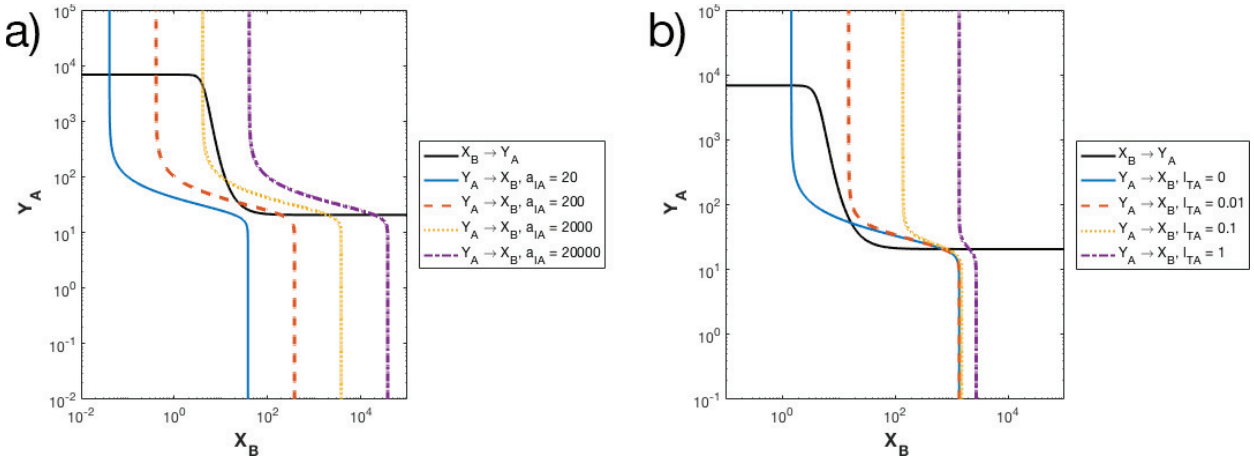

Figure 3.2: Kinetic rates determine whether the cross-repressive switches operate in the appropriate regime to turn each other on and off. Changes in relative output ranges between the two on/off switches for different values of (a) $a_{I_A} := \frac{V_{R_A} N C \epsilon_{I_A}}{\gamma_I \gamma_m}$ and (b) the leakiness of *plux* (l_{RA}). Remaining parameters are as given in Table 3.3. The intersections between two transfer functions $X_B \rightarrow Y_A$ and $Y_A \rightarrow X_B$ indicate the steady states of the full system. As a_{IA} scales, so does the maximum and minimum output of LasI, which geometrically translates the composite transfer function $Y_A \rightarrow X_B$. Increasing the leakiness increases the minimum output of LasI, which decreases the dynamic range of the composite transfer function $Y_A \rightarrow X_B$. Leakiness reduces ultrasensitivity of $Y_A \rightarrow X_B$, which leads to loss of bistability. The translational movement of $Y_A \rightarrow X_B$ causes mismatched tuning between $X_B \rightarrow Y_A$ and $Y_A \rightarrow X_B$ also leads to loss of bistability.

In addition to ultrasensitivity, DLI systems must have sufficiently similar inhibition strength between strains to be bistable. When the system loses bistability due to unbalanced inhibition strengths, monostable contrast emerges, with the extent of the contrast depending on the degree of imbalance). Changing spatial configuration triggers a bifurcation by modifying the effective inhibition strength between strains (Figure 3.3).

For the parameters as given in Table 3.2, the system is bistable across a range of geometries including checkerboards and asymmetric lattice arrangements. Of particular theoretical

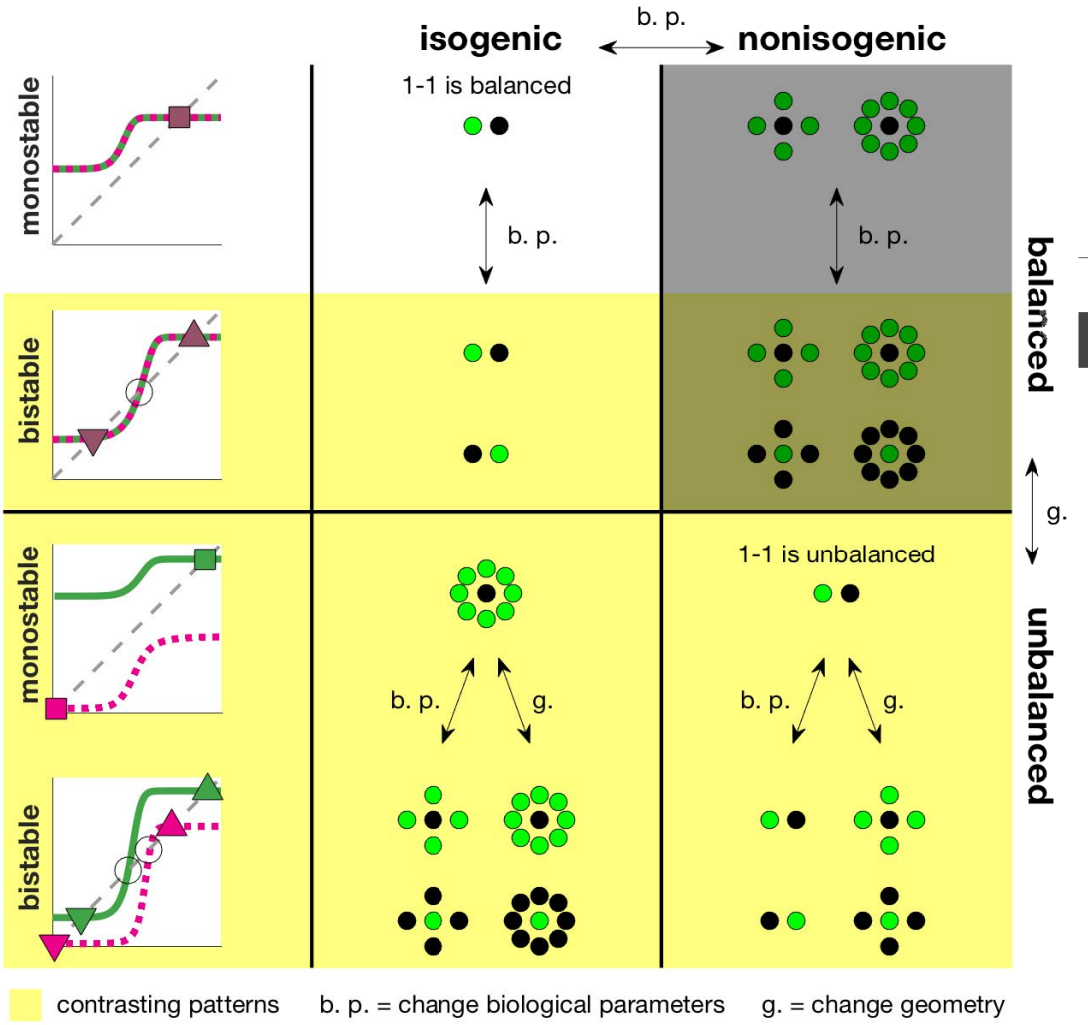

Figure 3.3: Schematic qualitatively identifying parameter changes that lead to bifurcation in two-strain circuits in isogenic (parameters identical between strains) and nonisogenic (parameters differ between strains) cases. Yellow background identifies systems that form contrasting patterns. In a balanced system, the cross-representative compartments exhibit identical transfer functions ($\boldsymbol{T_A}(\cdot) = \boldsymbol{T_B}(\cdot)$), while in an unbalanced system, the transfer functions differ. Monostable isogenic populations are homogeneous while nonisogenic populations exhibit contrast in the one-to-one geometry. Geometry may introduce imbalance or offset the biochemical difference, depending on the arrangement. Contrast in balanced systems is only attainable through bistability.

interest are "star" geometries in which one center colony is surrounded by a ring of colonies of the opposite type such that the overall spatial arrangement exhibits radial symmetry. Investigating stars of different sizes gives insight into how changes to the ratio of A compartments to B compartments affect the system's location on the stability diagram. Our results indicate that the growth rates of the cells and the length of the channels have little effect on the bistability of the system.

Theory predicts that the system is bistable for rings from 1 to at least 8 surrounding colonies for both central A/surrounding B and central B/surrounding A. Stability diagrams reveal that changing the number of colonies in the ring preserves the overall shape of the bistable region (plotted on a log scale) but shifts the system location relative to the edge of the bistable region (Figure 3.4).

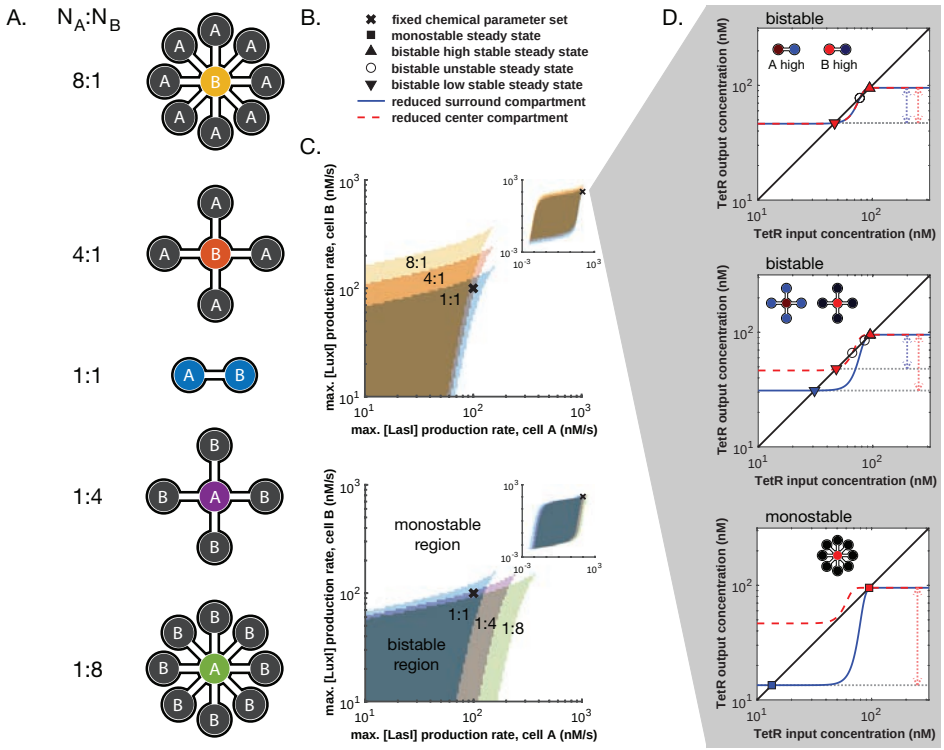

Figure 3.4: Patterning mechanisms and contrast levels determined by biochemical parameters and geometries in simulation. (A) Star geometries with varying numbers of surround compartments. Throughout the figure, ratios of colony numbers are given as $N_A : N_B$. (B) Character legend for plots in (C) and (D). (C) Overlaid stability plots show biochemical parameter ranges for which the system is monostable (white) and bistable (shaded, colors corresponding to center compartment of appropriate geometry in (A). Parameters on the axes are maximum steady-state production rates for LasI (x-axis) and LuxI (y-axis). Remaining parameters are as given in Table 3.2 where corresponding biochemical parameter values are equal between strains. As the number of points in the star changes, the shape of the bistable region remains the same (relative to log-scale axes) but shifts relative to the exact biochemical parameter values (insets show the full shape of the bistable region). $\times$ indicates an arbitrary set of fixed biochemical parameters that is bistable in the 1:1 and 4:1 cases but monostable for the 8:1 case. (D) For the biochemical parameters indicated by $\times$ in (C), a graphical test reveals that contrast may arise from a bistable system (1:1 and 4:1) or from a monostable system with imbalance (8:1) between the input/output characteristics of strains in the reduced systems. Steady states are indicated by ▲ (high expression) and ▼ (low expression) for the bistable case or ■ for the monostable case. In the bistable case with imbalance (4:1), the contrast level ($\leftrightarrow$) is greater when expression in the center compartment (dashed red) is high than when expression in the surrounding compartments (solid blue) is high. Small insets show corresponding configurations and possible steady-state solutions.

Parameter	Description	Value
D	diffusivity of HSL	5×10^{-4}
l	distance between channels	5×10^{-5}
d	edge weight constant	$:= \frac{D}{l^2}$
μ	dilution due to growth	0.1
γ	decay rate of HSL + drainage	0.01
ν	production (catalysis) rate of HSL	1
V	transcription rate	1
N	copy number	1
C	concentration constant	1
K	HSL dissociation constant	10
n	Hill coefficient (cooperativity) for HSL	2
l_R	rate of leaky transcription from *plux/plas*	0
γ_{m_R}	decay rate of TetR mRNA (dilution + degradation)	0.105
ϵ_R	translation rate of TetR	1
γ_R	decay rate of TetR	μ
K_R	dissociation constant of TetR to *ptet*	10
n_R	Hill coefficient (cooperativity) for TetR to *ptet*	4
l_I	rate of leaky transcription from *ptet*	0
γ_{m_I}	decay rate of LasI/LuxI mRNA	0.01
ϵ_I	translation rate of LasI/LuxI	1
γ_I	decay rate of LasI/LuxI (no degradation)	μ

Table 3.3: Parameter definitions and values used in the toy system with even parameters, i.e., corresponding parameters between Strains A and B are equal.

3.3 Experimental Results

Detailed experimental methods can be found in Appendix A.

3.3.1 DLI network design and implementation require diffusible cross-repression and a geometric culturing platform

Two *Escherichia coli* strains, A and B, were constructed using a pair of orthogonal QS systems [359] and a highly cooperative repressor, *tetR* [360] (Figure 3.5A). In both strains,

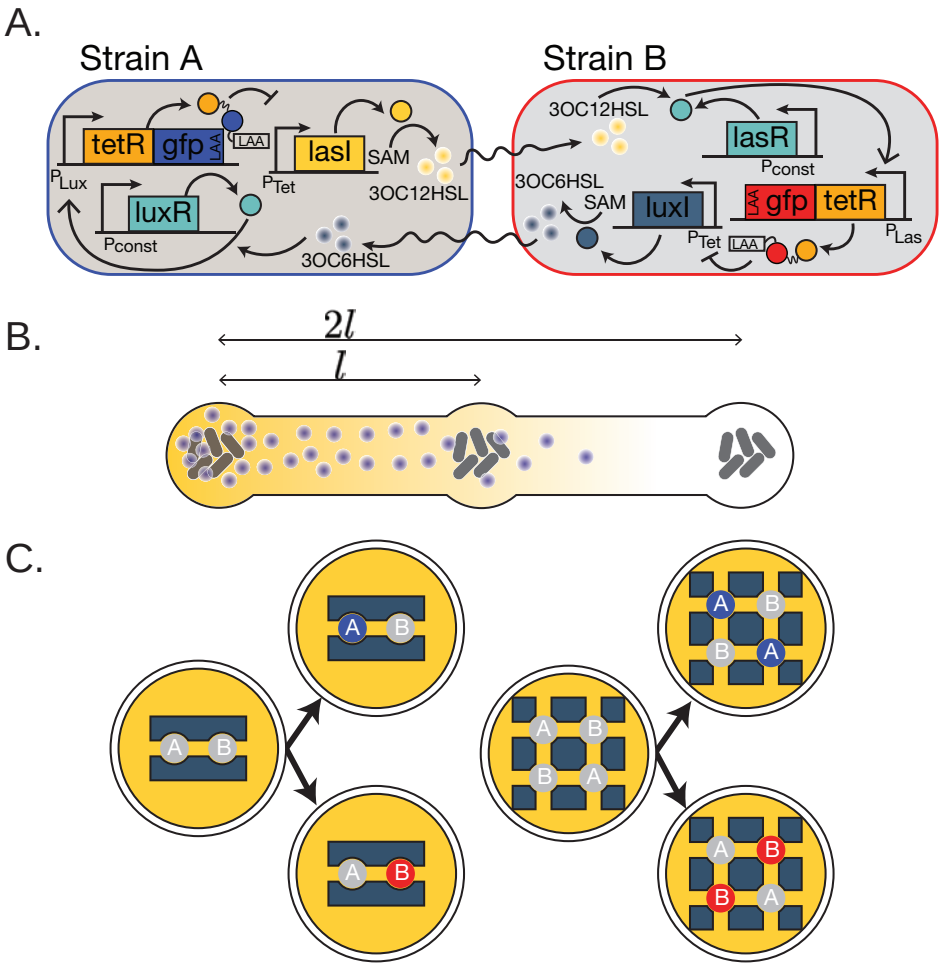

Figure 3.5: Schematic designs of the DLI system. Arrow-headed lines indicate activation and bar-headed lines indicate inhibition. (A) Genetic circuit diagram of cross-repressive strains. (B) Channel length l is chosen such that AHL diffusion establishes communication between adjacent compartments, but not between non-adjacent compartments with distance $\geq 2l$. (C) Each compartment of the DLI device is inoculated with one strain type. PDMS mold (indigo) is placed on a tissue culture plate to shape solid medium (yellow) into compartments and channels. Contrasting patterns emerge when two strains have different sfGFP-tagged TetR levels, either high (represented by green or magenta colored colonies) or low (represented by gray colonies).

tetR is translationally fused to the green fluorescent protein reporter, $sfGFP$, with LAA *ssrA* degradation tag in the C-terminus to allow dynamic tracking of the cell state [361].

The length of the channel (l) between compartments determines the AHL concentration in the neighboring compartments as well as the inter-compartmental communication lag time (Figure 3.5B). A partial differential equation (PDE) model of AHL production, degradation, and diffusion is used to optimize l for sufficient diffusion of AHL to the immediate neighbors while preventing communication between nonadjacent compartments (Appendix A). Since AHLs can be stable with a half-life of 6 hours up to days [362], an efflux channel is added to each compartment to match the dilution rate of AHL to the degradation rates of other proteins in the DLI circuit (Figure §5B, *Table* S3). We use polydimethylsiloxane (PDMS) as the mold to shape solid medium into compartments and channels in specific geometries (Figure 3.5C).

3.3.2 Bistable and contrasting gene expression was observed for DLI circuit in liquid coculture

Before testing pattern formation in compartmental structures, we verified the predicted bistability of the constructed circuits in liquid coculture. Numerical parameters required for biochemical modeling were determined by experimental measurements of individual modules of AHL reception/activation, transcriptional repression, and AHL synthesis and diffusion in the DLI circuit (Figures S7-9).

Single-cell reporter gene expression was measured using flow cytometry. To examine the existence of two stable steady states, the cocultures were biased with varying external concentrations of 3OC6HSL or 3OC12HSL. While the external AHL inductions in monocultures of Strains A (Figure 3.6A diamonds and the solid line for the Hill equation fit) and B (Figure 3.6B diamonds and the solid line for the Hill equation fit) resulted in gently sloped sigmoid responses, the coculture showed a sharp transition in steady state at 10 nM 3OC6HSL (Figure 3.6 squares), which is a characteristic for a bistable feedback loop [145].

Another characteristic of bistable systems is hysteresis. To investigate whether our system can reach two heterogeneous steady states in the same culture condition depending on initial conditions, we pre-conditioned Strains A and B monocultures with the saturated concentration (1 μM) of either 3OC6HSL or 3OC12HSL prior to mixing them into a coculture. The cocultures maintained the distinct gene expression states determined by pre-conditions over time, while the similarly pre-induced monocultures lost the pre-conditioned state and exhibited sigmoidal induction curves when the cultures were transferred to the fresh media (Figure 3.6 asterisks).

The appearance of both the sharp transition and hysteresis confirms that cross-repression between Strains A and B produces an effective intercellular bistable switch.

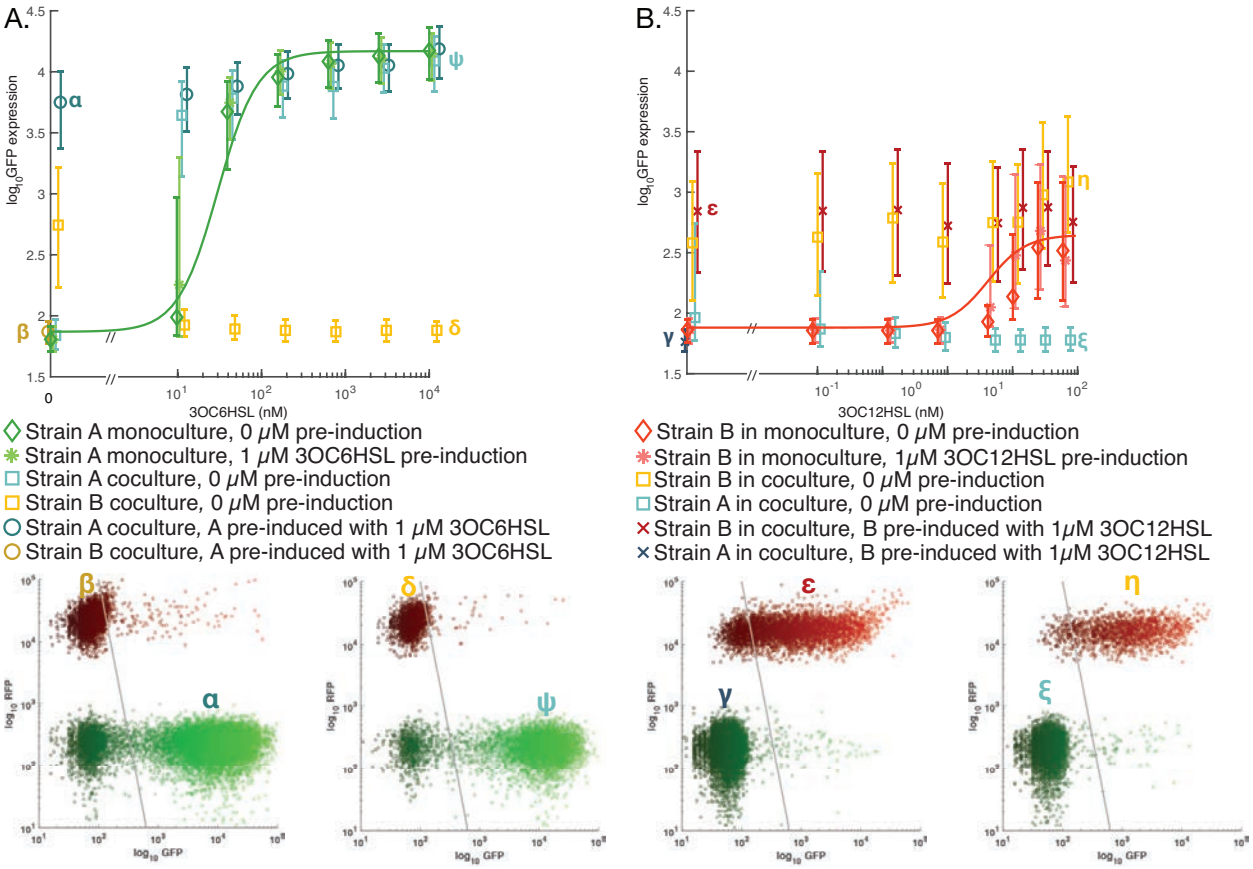

Figure 3.6: Steady states of the cross-repressive circuit characterized using flow cytometry measurements in liquid cultures. Strain B was identified using constitutively expressed *mRFP1*. Varying concentrations of (A) 3OC6HSL or (B) 3OC12HSL were externally added to the liquid medium (x-axis) and the medians of sfGFP fluorescence after 8 hours of growth were recorded (y-axis). Error bars represent 1st and 3rd quartiles of sfGFP fluorescence. All of the multi-strain cocultures, each indicated by $\square, \circ, \times$, exhibited contrasting expression profiles between Strains A and B. While the monocultures of strains A and B showed gently sloped responses to external AHL with Hill function fits of $K_d \simeq 50\ nM$ and $K_d \simeq 20\ nM$ (solid lines), the two-strain coculture showed a switch-like response at threshold [3OC6HSL] $= 10\ nM$. Hysteresis was tested by pre-inducing one of the strains with appropriate AHL prior to washing and mixing the strains into a coculture with fresh medium. The two-strain cocultures maintained pre-induced states after 8 hours of growth even without external AHL whereas the monocultures lost their pre-induction states. Bottom scatter plots show similarity in gene expression patterns between Strain-A-biased coculture (ψ, δ) and Strain-A-pre-induced coculture at 8 hours of growth after removing external AHL (α, β), and Strain-B-biased coculture (ζ, η) and Strain-B-pre-induced coculture at 8 hours of growth after removing external AHL (γ, ϵ).

3.3.3 Bifurcation is observed for cells grown on the geometric culturing platform

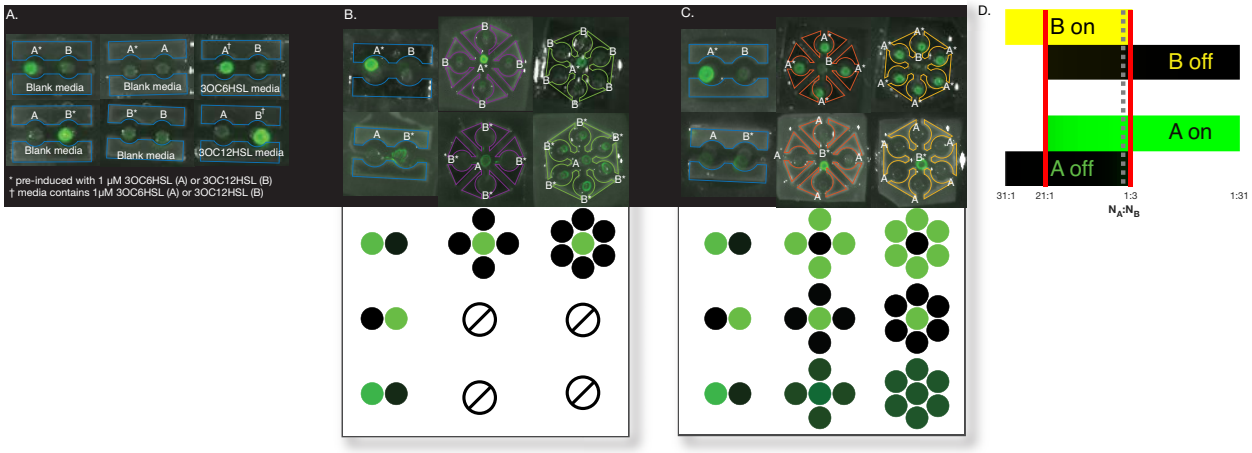

Figure 3.7: Contrasting pattern formation in various DLI devices. The fluorimeter images were taken after 12 hours of growth in room temperature. ∗indicates pre-induced strains with 1 μM AHL and † indicates strains that were biased to be fluorescent by externally added AHL in medium. (A) 1:1 spatial configuration seeded with cells that had different initial conditions and strain combinations. Devices were seeded with a pair of complementary strains (left), negative controls consisting a single strain (middle), and positive controls of complementary strains where either 1 μM 3OC6HSL or 1 μM 3OC12HSL was mixed in solid medium (right). (*B,C*) 1:1 (left), 1:4 (middle), and 1:6 (right) spatial configurations seeded with Strain A at the center surrounded by Strain B (B) or Strain B surrounded by Strain A (C). Top panel shows the fluorimeter images and the bottom panel shows predicted steady-state pattern from computational simulations with the parameter values given in *Table* S3. When multiple equilibria exist, the predicted patterns are plotted in the order of "A high", "B high", and "unstable". (*D*) Simulated one-dimensional bifurcation diagram in which the ratio of compartments of Strain A:B is used as the bifurcation parameter. The remaining parameters are given in *Table* S3. Brighter color indicates higher steady-state [sfGFP].

Theoretically, the DLI system with two compartments with proper channel length should behave similarly to the liquid coculture. To investigate bistability and hysteresis in the geometric platform, we plated Strains A and B pre-conditioned with either 1 μM 3OC6HSL or 1 μM 3OC12HSL on solid medium after washing the pre-conditioning media and observed their gene expression over time using the plate fluorimeter. The two-compartments geometry plated with complementary strains maintained the pre-conditioned states over 12 hours whereas the single-strain systems on the same setup quickly lost the pre-induced gene expression (Figure 3.7A). When 3OC6HSL pre-induced Strain A was seeded adjacent to Strain B, Strain A showed high reporter expression while Strain B showed basal expression comparable to the single-strain control, and this persisted in time. Similarly, 3OC12HSL pre-induced Strain B seeded adjacent to Strain A showed high reporter expression in Strain B and low reporter expression in Strain A, indicating that contrasting patterns depended on initial conditions.

Next, we evaluated DLI systems in star geometries (Figure 3.7B, C). We experimentally observed that the system that was bistable in the two-compartments geometry became monostable as the number of outer compartments of Strain B surrounding Strain A increased to four or more (Figure 3.7B). The monostable contrasting pattern exhibited high reporter fluorescence in the center Strain A regardless of the initial cell states. Strain B, on the other hand, did not shift from monostable to bistable in our experimentally tested geometries, but rather augmented the bistable expression as the number of surrounding A increased. Although the experimental setup cannot physically accommodate > 6 compartments, the mathematical simulation suggests that further increasing the number of surrounding compartments would shift the system from bistable to monostable contrast with Strain B expressing high reporter (Figure 3.7D).

Our results show that we can easily control the geometry of the DLI system to affect the circuit behavior and trigger a bifurcation. Geometry may also offset imbalance in biochemical parameters and improve the stability of bistable steady states (Figure 3.7C).

3.4 Discussion

Recent advancement in high-throughput sequencing has revealed that an astonishing range of microbial biodiversity may exist in a single ecosystem [363]. For the analyzed DLI system, we explored the role of geometry in system behavior. Although the implementation was done in multi-strain bacterial colonies, our theory can be applied to isogenic populations as well. Here, the units of interest are individual cells rather than colonies and communication must be contact-mediated since diffusion-based signaling would form self-loops. Dimensionality reduction still applies when cells can be categorized into two separate "classes" by virtue of spatial configuration. Replacing the Laplacian matrix with the adjacency matrix simulates cell-to-cell contact rather than diffusion. The remainder of the analysis then proceeds as before. Extensive and detailed research has been performed to accurately model the developmental processes in metazoans [289, 364, 365], and a handful of recent studies have highlighted that spatially relevant parameters such as the number of neighbors or the contact area between them can influence patterning activity even in genetically isogenic cell populations [123, 163, 366]. Our work offers a unified interpretation of these results with respect to the imbalance in transfer functions between pairs of representative cells. With sufficient imbalance, the system becomes monostable, essentially guaranteeing the fate of the involved cells, and in fact only spatial control knobs can introduce monostable contrast in isogenic populations, since changes to biochemical parameters affect all cells equally. Furthering our understanding of micro-scale pattern formation would require experimental implementation of controllable contact-based systems. In bacteria, several contact-dependent inhibition systems have been discovered [355, 356] in which potential harnessing strategies have been discussed [367].

Genetic circuit design and implementation are hampered by context-dependent gene expression [368]. Spatial control has advantages over biochemical parameter modification in

that it can linearly modulate the effective interaction strength (Figure 3.7D) via the number of connected channels, and the modulation is robust to intracellular conditions. Furthermore, physical separation of the composite strains reduces resource competition among different strains [352] to stabilize the intercellular network. However, spatial control is constrained by structural limitations, such as the maximum number of compartments that fit on the mold or whether the layout of the desired communication network can be laid out without channels intersecting each other. Thus, synthetic biologists should exploit both biochemical and spatial control knobs for precise design of microbial consortia engineering.

3.5 Acknowledgments

This work appears in [307], in which the authors thank the NIH National Institute of General Medical Sciences grant 1R01GM109460-01 and Air Force Office of Scientific Research grant FA9550-14-1-0089 for support, as well as Prof. Michel Maharbiz for comments and discussions, and Prof. Dave Savage for use of a plate fluorimeter. The DH10B variant TR117 was provided by Dr. Thomas L. Ruegg and the backbone image processing scripts were provided by Lynn D. Kong.

Chapter 4

Cell-in-the-loop Patterning with Optogenetically Emulated Cell-to-Cell Signaling

In Chapter 3, we tested theory for spontaneous contrasting pattern formation in lateral inhibition systems using a purely biochemical system of bacterial colonies interacting through diffusible molecules. In the course of our experiments, we found that it was challenging to match biochemical parameters between the two engineered strains, and we were ultimately unable to show spontaneous patterning without preinduction. Furthermore, due to spatial constraints, we were restricted to systems with seven colonies or fewer. Inspired in part by the challenges presented by this setup, we sought to validate the same original theory from [289] and [122] using a "cell-in-the-loop" approach, in which physical or chemical cell-to-cell signaling is substituted with inputs calculated and controlled *in silico*. Cells were engineered to increase gene expression in response to light. Then, a computer-controlled system automatically measured the gene expression levels of individual living cells and administered new light inputs to the cells based on those measurements. Certain parameters such as the strength of inhibition could be easily tuned and tested by programming them directly into the computer without going through the effort and expense to implement them fully *in vivo*. Using this new setup, we showed spontaneous emergence of contrasting gene expression patterns in systems of sixteen interacting components. We showed that our theory qualitatively predicted patterning outcomes and quantitatively predicted overall brightness (for non-patterning) or contrast level (for patterning) on average across experiments. In addition to testing theories, the cell-in-the-loop approach has potential practical applications as a prototyping tool or as a control unit to drive real-world results.

4.1 Introduction

There are a number of challenges associated with engineering spontaneous gene expression patterning into biochemical systems, including how to facilitate interaction among cells [369] and achieve spatial precision in the resulting patterns [150, 163, 307]. Even when successful, these implementations are still constrained by time, expense, and the availability of biological parts satisfying parameter requirements [172, 370]. Moreover, it may be difficult to measure or monitor particular system components in real time, which can hinder "debugging" and slow down the design-build-test cycle [173].

While numerical simulation is an important method for efficient prototyping, simulations are only as valid as the models underlying them, and simplifications or faulty assumptions can limit the experimental applicability of simulation results. We propose that future efforts in synthetic patterning would benefit from an intermediate step between pure simulation and full biochemical implementation, which could be used to validate theories or incrementally test synthetic designs before they are fully incorporated into the organism. Inspired by "human-in-the-loop" approaches for engineering systems that must interact with complex, living individuals [371], we propose a "cell-in-the-loop" approach in which physical signaling among cells is substituted with computer-controlled inputs calculated *in silico* from real-time measurements of gene expression. Cell-in-the-loop, by incorporating live cells into the "simulation", eliminates the need to make assumptions about individual cell behavior during dynamic evolution, while retaining flexibility in testing parameters that remain under computational control. These benefits are particularly essential for patterning systems, in which the large number of interacting cells can make detailed simulations prohibitive or impossible.

We implemented cell-in-the-loop using optogenetics, which have been shown to afford excellent spatiotemporal precision in applications including feedback control [372–375], and which were previously used to emulate cell-to-cell signaling for oscillatory synchronization [376]. We engineered *Saccharomyces cerevisiae* to respond to blue light [377] by increasing gene expression as measured by a fast-acting fluorescent reporter [378]. We used an optogenetic platform capable of targeting individual cells independently of each other [375], such that the light input to any given cell could be calculated based on the gene expression levels of other cells that were interacting with the target cell. Both the network architecture (which cells interacted with which) as well as the exact form of interaction were programmed into the computer, allowing us to precisely modulate system parameters related to cell-to-cell signaling.

We adapted a general theory for pattern emergence in large-scale lateral inhibition systems [122, 289] to inform our designs and predict steady-state outcomes. Lateral inhibition regulated by the Notch-Delta signaling pathway is responsible for patterning in a range of developmental contexts, including proneural stripe formation [130] and subsequent neural precursor selection [116] in fruit flies, as well as patterning in the central nervous system [120], inner ear [118, 119], and intestine [121] of vetebrates [320]. Inspired by these systems, we programmed a computational *signaling relation* to emulate mutual inhibition among groups

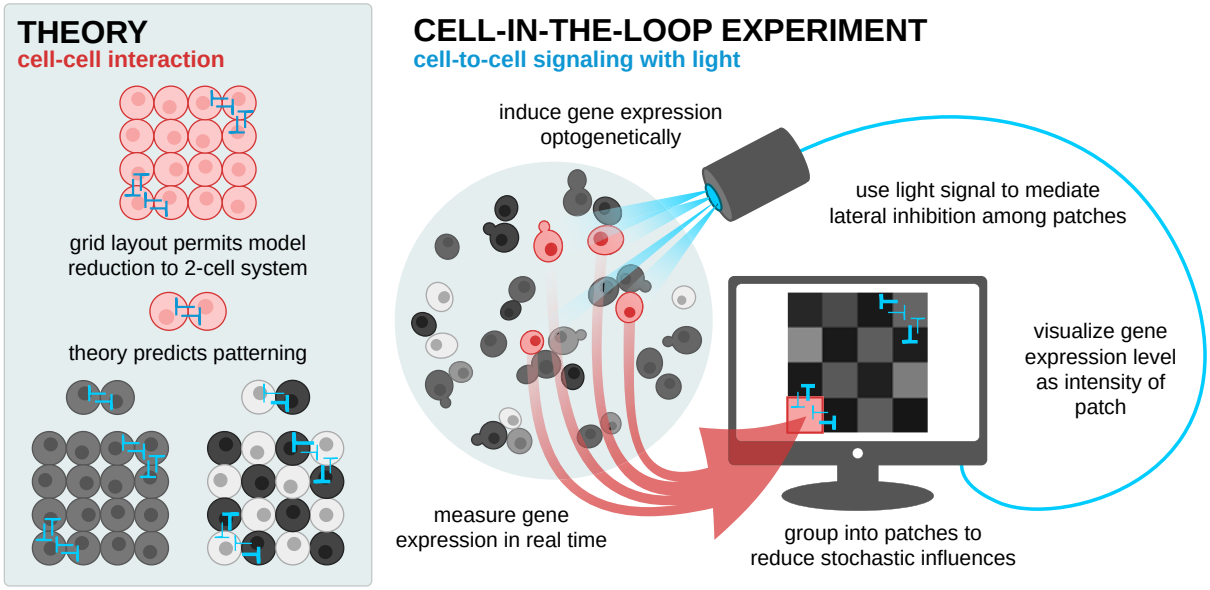

Figure 4.1: Spontaneous checkerboard patterning with optogenetically emulated cell-to-cell signaling. Optogenetically responsive cells signal to each other through computer-controlled light inputs that vary in intensity based on the gene expression levels of other cells. We enact lateral inhibition according to the theory in Section 4.2.1 that predicts when cells will spontaneously separate into two classes of high and low gene expression. In all figures, red denotes *in vivo* and blue denotes *in silico* components.

of cells and varied the strength of the inhibition by tuning a single digital bifurcation parameter. Once the network architecture and signaling relation were defined, inputs to cells were calculated solely based on measurements of those cells without any further external control, creating a self-contained dynamical system. Using this setup, we visualized gene expression levels of real cells by the brightness of square patches on a virtual grid (Fig. 4.1). We showed spontaneous emergence of contrasting "checkerboard" patterns in which neighboring patches alternated between expressing high and low levels of gene. The theory accurately predicted which values of the bifurcation parameter would produce patterns, and on average across experiments the theory also quantitatively predicted contrast levels and overall patch brightness. Our results demonstrate the utility of a cell-in-the-loop approach for designing and evaluating systems of interacting cells, as well as probing the limits of deterministic theory in the face of stochastic influence.

4.2 Results

4.2.1 Theory predicts patterning using a test for bistability

We developed theory to predict the emergence of stable contrasting patterns in deterministic systems of laterally inhibiting cells [122, 289]. Here, we adapt the theory to the present

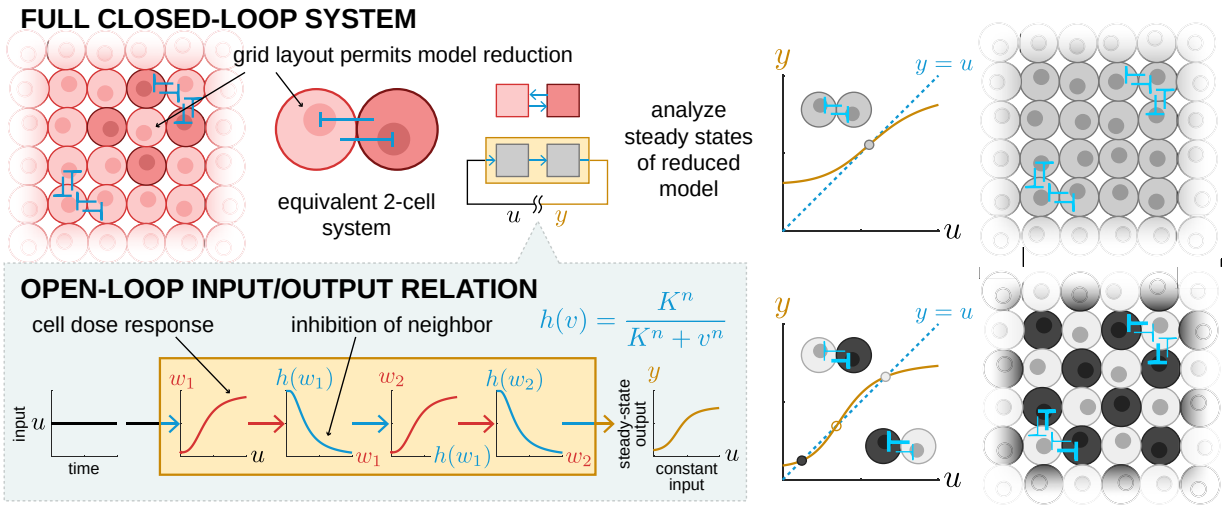

Figure 4.2: Spontaneous contrasting patterning in a large lateral inhibition system can be predicted from bistability in a 2-cell system. The bistability of a 2-cell system can be used to evaluate contrasting patterning in a full network of isogenic, mutually inhibiting cells arranged in a grid with periodic boundary conditions. Bistability of the corresponding 2-cell system implies contrasting patterns exist in the full system (see box entitled "Summary of theory"). Bistability can be assessed through an analytical test in which the closed-loop dynamical system is "opened" into an input/output system by breaking the feedback loop. If a constant-in-time input to the open-loop system produces a steady-state output value equal to the input value, then that value is a steady state for the corresponding closed-loop system.

optogenetic implementation. We emphasize how our system was decomposed into *in vivo* and *in silico* components, each of which corresponds to a particular element in the theory, and how this correspondence enables empirical measurement and experimental design.

Consider a system of N isogenic cells signaling to each other. Suppose we measure for each cell a scalar output such as fluorescence that correlates positively with gene expression level and is designated by w_i for the ith cell. The input u_i to a cell affects output levels with an empirically characterizable dose response, which describes the steady-state level of w_i for a constant-in-time input. In our setup, the input u_i is light, and increasing input intensity increases gene expression. This portion of the theory represents the *in vivo* component of the system.

To synthesize u_i, we first average the measured gene expression, w_j, over all cells j signaling to cell i, and denote this average as v_i. We then set the input to the ith cell to $u_i = h(v_i)$, where $h(\cdot)$ is the signaling relation programmed into the computer. To enact mutual inhibition, increasing gene expression in one cell must decrease gene expression in neighboring cells. Therefore, since higher-intensity light induces higher gene expression, we select $h(\cdot)$ to be decreasing.

We chose a grid layout with periodic boundary conditions in which each cell signals four other cells reciprocally. This layout satisfies all assumptions discussed in the box entitled "Summary of theory", therefore we can predict contrasting patterning in a full system of N

cells based on the bistability of an equivalent 2-cell system. If the 2-cell system is monostable, then both cells express the same level of gene, and the N-cell system also has a stable state in which all cells express the same level of gene. Inversely, if the 2-cell system is bistable, then one of the stable states corresponds to one cell expressing high levels of gene and the other, low, and the other stable state corresponds to the opposite situation. In this case, two stable, contrasting steady-state patterns also exist for the N-cell system; that is, one subset of the N cells expresses identically high levels of gene, and the remaining cells express identically low levels of gene (or vice versa). Contrasting steady states can be visualized as checkerboards in which neighbors alternate between high and low. The 2-cell system can be assessed for bistability using a standard technique illustrated in Figure 4.2 and described in Section 4.2.1.

To apply the theory, we must know (1) the dose response, in our case *in vivo* gene expression levels under varying intensities of light; and (2) the form of the signaling relation, here programmed *in silico*. Thus, to carry out lateral inhibition experiments, we needed to measure an empirical dose response of cells to light, and define the computational signaling relation controlling light inputs such that intensity was inversely related to the responsiveness of cells interacting with the target.

Summary of theory

Consider a system of N identical cells modeled as single-input, single-output dynamical systems. Biochemical concentrations $\mathbf{x}_i(t) \in \mathbb{R}_+^n$ in the ith cell evolve according to

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}_i(t) = f(\mathbf{x}_i(t), \mathbf{u}_i(t)). \tag{4.1}$$

Each cell has output $w_i(t) = g(\mathbf{x}_i(t)) \in \mathbb{R}_+$ and input $u_i(t) \in \mathbb{R}_+$. Let the vector $\mathbf{x}(t) \in \mathbb{R}_+^{Nn}$ be the vertical concatenation of the vectors $\mathbf{x}_i(t)$ for all N cells, and similarly for $\mathbf{w}(t) \in \mathbb{R}_+^N$ and $\mathbf{u}(t) \in \mathbb{R}_+^N$. We assume each cell has a *static input-output characteristic* $T(\cdot)$, that is, if a cell is given constant-in-time input $u_i(t) = u_i^\dagger$, it will reach a globally asymptotically stable hyperbolic equilibrium $x_i^\dagger$ solving $0 = f(\mathbf{x}_i^\dagger, u_i^\dagger)$ with output $w_i^\dagger = T(u_i^\dagger) = g(\mathbf{x}_i^\dagger)$. We assume $T(\cdot)$ is bounded and increasing, meaning that increasing the input increases the output. In our setup, the static input-output characteristic corresponds to the empirically measured dose response.

Suppose the outputs of cells are connected to the inputs of other cells, forming a network. We capture information about which cells signal to which by way of the interconnectivity matrix $M \in \mathbb{R}_+^{N\times N}$ with entries $[M]_{ij} = 0$ if cell j does not signal to cell i and $[M]_{ij} > 0$ otherwise, with the value $[M]_{ij}$ indicating the strength of signaling. We require that the sum over all entries in a row equal the same constant, $\mu \in \mathbb{R}_+$, regardless of the row, i.e., $\sum_j [M]_{ij} = \mu$ for all i. In our setup each cell receives signals from four other cells with equal weights $\frac{1}{4}$, therefore $\mu = 1$. Defining $\mathbf{v}(t) = M\mathbf{w}(t)$, we model lateral inhibition by letting the input to cell i be given by $u_i(t) = h(v_i(t))$, where $h(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ is bounded and decreasing.

***Model reduction theorem**: Let $\mathbb{1}_m$ represent the length-m column vector of all ones, and similarly for 0_m. If there exists a matrix $\bar{M} \in \mathbb{R}_+^{2\times 2}$ such that*

$$ML = L\bar{M} \text{ where } L = \begin{bmatrix} \mathbb{1}_m & 0_m \\ 0_{N-m} & \mathbb{1}_{N-m} \end{bmatrix} \tag{4.2}$$

for some indexing of cells, then $\bar{M}$ is an interconnectivity matrix for an equivalent 2-cell system whose steady-state solutions correspond to steady states of the N-cell system with interconnectivity matrix M. In other words, if $\bar{\mathbf{w}}^* \in \mathbb{R}_+^2$ is the output corresponding to a steady-state solution to the 2-cell system, then $\mathbf{w}^* = L\bar{\mathbf{w}}^*$ is a steady-state output to the N-cell system. □

Note that the cells indexed 1 through m take on steady-state output values $\bar{w}_1^*$ while those indexed $m+1$ through N take on steady-state output values $\bar{w}_2^*$. Condition (4.2) is satisfied when cells can be grouped into two subsets within which nodes are interchangeable; that is, reindexing nodes within a subset will not change M [122].

When

$$\bar{M} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{4.3}$$

the steady states of the 2-cell system are determined graphically from the fixed points of $h(T(h(T(\cdot))))$, as shown in Figure 4.2 and explained in Section 4.2.1. For the reduced 2-cell system the graphical test also ensures stability of the points corresponding to the lower/upper intersections in Figure 4.2, and instability of the point corresponding to the middle intersection, when the cellular dynamics are *monotone* in the input/output sense [357]. The stability properties established graphically for the 2-cell system are preserved in the full N-cell system when additional assumptions hold. Our setup satisfies one such assumption from [289], which stipulates cells within a subset not signal to each other. Thus, if $\bar{\mathbf{w}}^*$ is the output corresponding to a stable state in a bistable 2-cell system, then in the N-cell system, cells in one subset have higher output than cells in the other subset. If the cells belonging to different subsets are spatially interlaced or alternating, then the high/low dichotomy produces a spatially contrasting pattern such as a checkerboard.

Graphical test for bistability

Consider a 2-cell system of mutual inhibition with dynamics (and notation) as given in Section 4.2.1, in which $u_2(t) = h(v_2(t)) = h(w_1(t))$ and similarly $u_1(t) = h(v_1(t)) = h(w_2(t))$. Now break the loop such that $u_1(t)$ becomes the input and $y(t) = h(w_2(t))$ the output of a the *open-loop system*. The static input-output characteristic for the cascade, given by $y^\dagger = h(T(h(T(u_1^\dagger))))$, is increasing. The points where $u_1^\dagger = h(T(h(T(u_1^\dagger))))$ (i.e., intersections of the static input-output characteristic with a line of slope 1) are the steady states of the corresponding *closed-loop system* in which $u_1(t) = y(t)$ (and we still have $u_2(t) = h(w_1(t))$). We will designate values at such intersections by superscript asterisks (e.g., u_1^*). When cell dynamics are monotone, then the steady state corresponding to u_1^* is stable if the slope of

$h(T(h(T(u_1^*))))$ is greater than 1 and unstable if it is less than one. In particular for our setup, this implies that if there is one intersection, the closed-loop system is monostable, and if there are three intersections, the system is bistable, with the homogeneous solution ($u_1^* = u_2^*$) being unstable and the other two stable points corresponding to w_1^* high, w_2^* low, and vice versa [294].

4.2.2 Empirical characterization informs computational parameters

We combined the blue light-inducible VP-EL222 expression system [377,379] with a fast-acting nuclear translocation reporter (dPSTR) [378] to control and measure gene expression in *Saccharomyces cerevisiae*. In the dark, constitutively expressed red fluorescent protein (RFP) fused to the synthetic bZip domain SZ2 [380] is equally distributed between nucleus and cytoplasm due to passive diffusion through the nuclear membrane. Under exposure to blue light, VP-EL222 molecules dimerize and bind the cognate promoter to activate expression of a protein comprising two nuclear localization signals (NLS) and SZ1 [380]. This protein then forms a heterodimer with the RFP reporter, thereby localizing fluorescence in the nucleus. We quantitated the degree of nuclear localization (nuclear localization score) as the difference between mean cytoplasmic and mean nuclear fluorescence normalized to the mean fluorescence across the entire cell. In principle, the score is 0 if cells are not at all responding (there is no nuclear localization) and positive otherwise (Figure 4.3a,b).

We characterized the dose response of individual cells to constant, targeted blue light exposure (Figure 4.3b,c). On average, cells exhibited a graded response to light intensity well described by a Hill function (Figure 4.3c). Variability from cell to cell was greater than for individual cells across time, perhaps owing to variation in cell cycle state [381]. As the theory is deterministic, for patterning experiments we ultimately substituted single cells with computationally defined patches of 4 or 6 cells, with the patch response determined as the average response of the constituent cells. Generating score distributions for such patches by bootstrapping from the single-cell dose response data shows reduced temporal and patch-patch variability as well as reduced difference between temporal and patch-patch variability relative to the single-cell case (Figure 4.3d, Figure 4.4).

Based on the range of cellular response scores, we defined the signaling relation $h(\cdot)$ for use in patterning experiments, which determined the light input administered to a patch as a function of the average scores of neighboring patches at each time step. We chose an inhibiting Hill function with fixed Hill coefficient $n = 2$ and a single free parameter K with smaller values corresponding to sharper inhibition. We combined the empirical dose response with the computational signaling relation to generate theoretical predictions for the mono- or bistability of a 2-cell lateral inhibition system as K was varied between 0 and 1, corresponding to non-patterning or patterning outcomes in a full system.

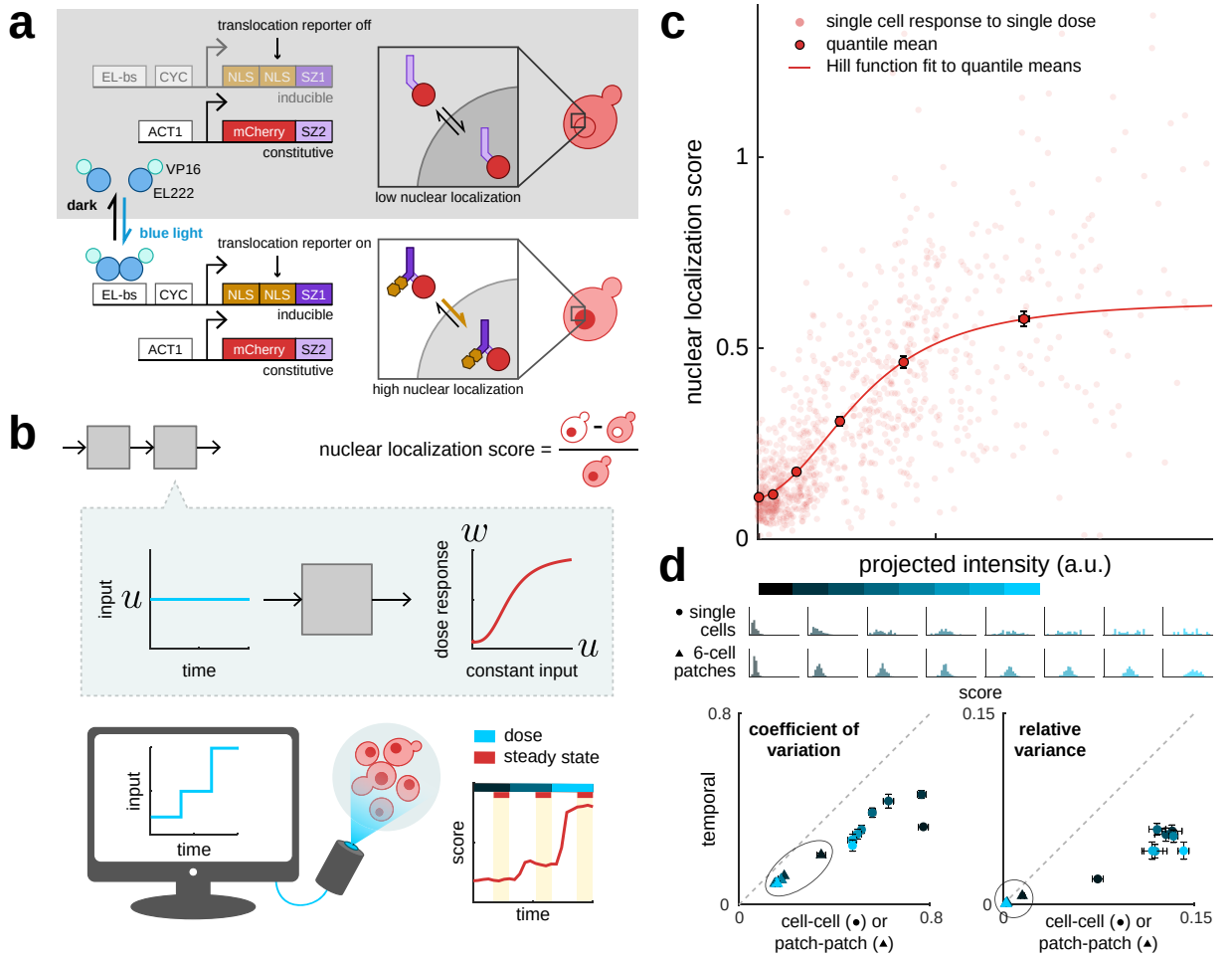

Figure 4.3: Experimental characterization of cellular dose response curves shows population-level gradedness. (a) **Blue light induces gene expression in optogenetically responsive *Saccharomyces cerevisiae* via the VP-EL222 system [375], as measured by a nuclear translocation reporter (dPSTR) [378].** Cells constitutively express RFP fused to SZ2. In the dark, RFP is distributed equally between nucleus and cytoplasm due to diffusion. Blue light induces VP-EL222 dimerization, which activates expression of an NLS fused to SZ1. SZ1 binds SZ2, such that the NLS promotes localization of RFP in the nucleus, increasing nuclear fluorescence relative to cytoplasmic fluorescence. (b) For a single dose, individual cells were illuminated for 80 min with constant-intensity light. Single-cell responses to a single intensity were calculated as the average score over the last 40 min. Pictured in the schematic are average input intensities and responses for approx. 700 cells for 3 consecutive doses. (c) **Although individual cell responses vary, on the population level nuclear localization is graded with respect to the intensity of the light input.** Responses of single cells to single doses were pooled across 3 experiments (no outline). A Hill function was fit to the quantile means (180 or 181 samples per quantile) to generate the final dose response (solid outline). (d) **Grouping cells into patches reduces both cell-cell and temporal variability.** Scores were binned by projected intensity (blue shaded bar) to generate histograms of score distributions for given input levels (samples per bin left to right: 344, 183, 69, 108, 102, 57, 35, 26). Bin cutoff is twice the maximum projected intensity used in final patterning experiments. Histograms for 6-cell patches were generated by bootstrapping from individual cell scores within a bin for a total of 200 bootstrapped patches per intensity level. For an ergodic process, cell-cell and temporal variation would be equal (gray dotted line); here, responses appear to be more variable from cell to cell than for single cells across time. Grouping cells into patches of 6 across which response scores are averaged (triangles) reduces the magnitude of difference between cell-cell or patch-patch and temporal variability. Error bars are s.e.m.

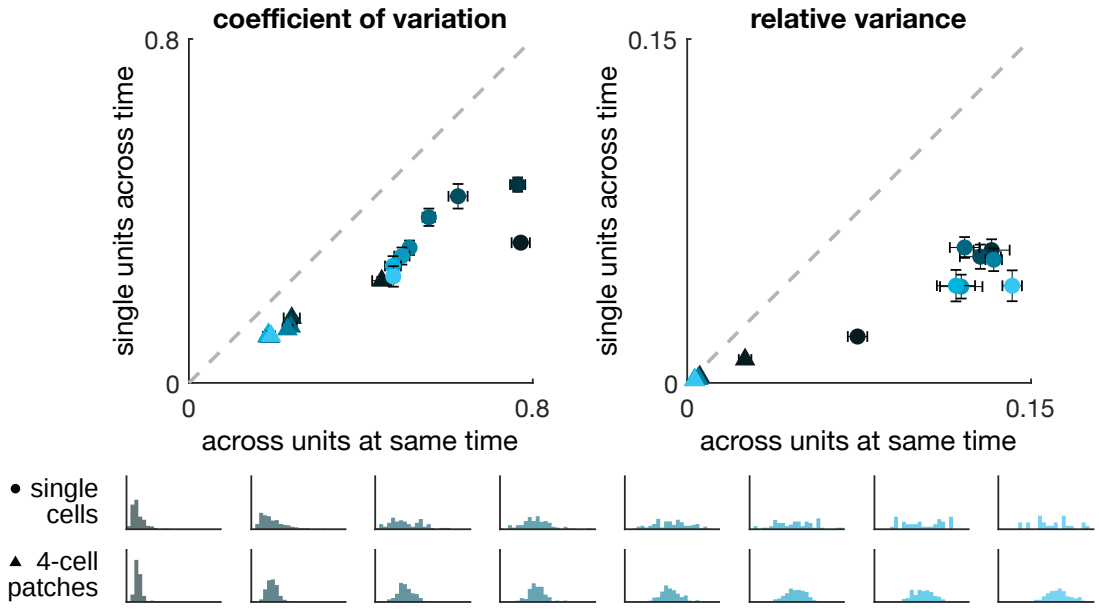

Figure 4.4: Cell-cell and patch-patch variability in cell response score for 4-cell patches at different illumination intensities. Plots correspond to those in Figure 4.3d, but for 4-cell instead of 6-cell patches. Scores were binned by projected intensity (increasing from black to light blue) to generate histograms of score distributions for given input levels (samples per bin left to right: 344, 183, 69, 108, 102, 57, 35, 26). Bin cutoff is twice the maximum projected intensity used in final patterning experiments. Histograms for 4-cell patches were generated for each bin by bootstrapping from individual cell scores within that bin for a total of 200 bootstrapped patches per intensity level. Circles indicate cell-cell variability; triangles indicate patch-patch variability. Error bars are s.e.m.

4.2.3 Cell-in-the-loop generates spontaneous checkerboard patterns

We ran a series of patterning experiments emulating lateral inhibition. Cells were randomly assigned to patches such that cells belonging to the same patches were not necessarily neighbors in physical space, thereby reducing spurious correlations that might arise from spatially dependent factors other than the targeted light input. Once assigned, cells remained in the same patch throughout the duration of an experiment. Patches were arranged to neighbor each other in "virtual space" as visualized on a checkerboard (Figure 4.5).

During patterning experiments, we took measurements and calculated new light inputs every 10 min, about the same rate as the estimated time constant for cell response to a switch in illumination intensity (Section B.3.1). This choice allowed us to avoid adjusting inputs more frequently than responses to the previous input could be detected. We tested systems of 16 patches with 6 cells per patch for four values of K between 0.1 and 1. Spontaneous patterning was always achieved in the $K = 0.1$ case and never in the $K = 1$ case, with

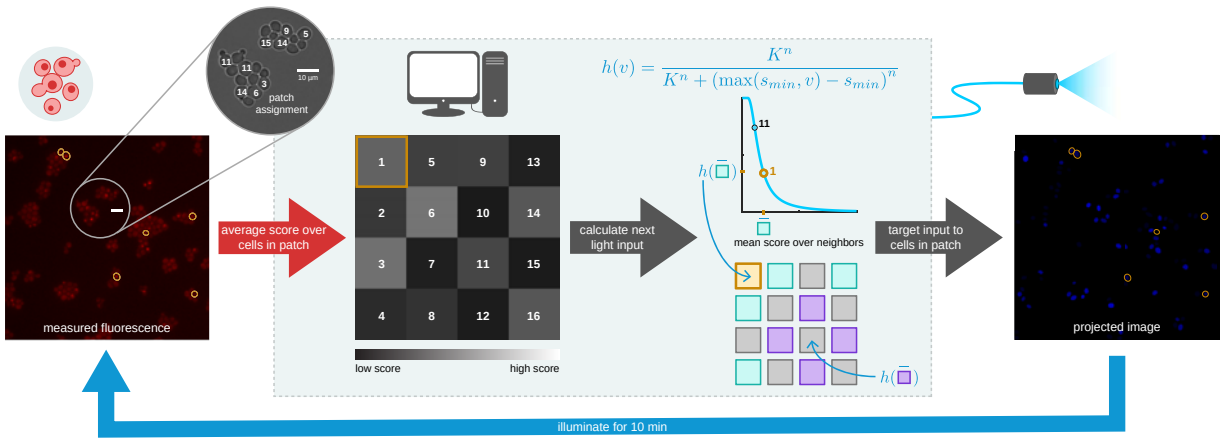

Figure 4.5: Schematic of automated workflow for patterning experiments. Cells growing in a monolayer were placed under the microscope, which was connected to a computer system that controlled the camera and blue light projection system. Cells were segmented and tracked in brightfield images and scored in fluorescence images (here, false-colored red). After the first image was acquired, cells were randomly assigned to patches to which they belonged for the remainder of the experiment. The grid visualizes scores at a single time step, with each square representing one patch of cells. The brightness of a square corresponds to the average score of the constituent cells in the corresponding patch. The signaling relation, defined before experiment start, determined the input light intensity administered to patches for the next time step (10 min) based on the scores of neighboring patches to the north, south, east, and west, with periodic boundary conditions. The relation $h(\cdot)$ was chosen to be an inhibiting Hill function with a minor correction to better utilize the range of available illumination intensities (i.e., a patch received maximum illumination intensity if the average score over its neighbors was at or below $s_{min} = 0.05$). The Hill coefficient was fixed at $n = 2$ for all experiments. The bifurcation parameter K was fixed for each experiment at a value between 0.1 and 1. Cells in the same patch received the same input intensity targeted individually to each cell, as shown in the projected image. Scale bar in fluorescence and brightfield images is 10 µm.

mixed results for $K = 0.2$, 0.3, near one of the theoretically predicted critical points (Figure 4.6a). Sample time traces at $K = 0.1$ and $K = 1$ show, respectively, the gradual deviation in score between sets of alternating patches that characterizes a contrasting pattern, or a rapid adoption of a non-patterning state. Visualizing the checkerboard at individual time points or averaged over the last hour clearly depicts the distinction between the two cases (Figure 4.7a). In the bistable case, which of the two possible checkerboards emerged in a given experiment depended on the stochastic initial conditions and continued noisy influences during system evolution (Figure 4.6).

When averaged over the last hour and across experiments, the contrast level (mean scores of sets of alternating patches) was quantitatively well predicted by theory in the bistable region and the overall brightness (mean score across all patches) was well predicted in the monostable region (Figure 4.7b). When considering individual experiments, the variability in overall brightness increased with increasing K. In the predicted monostable cases, stochasticity also introduced a difference between the means over alternating patches, though statistical analyses confirm that the difference was indistinguishable from random (Table 4.1).

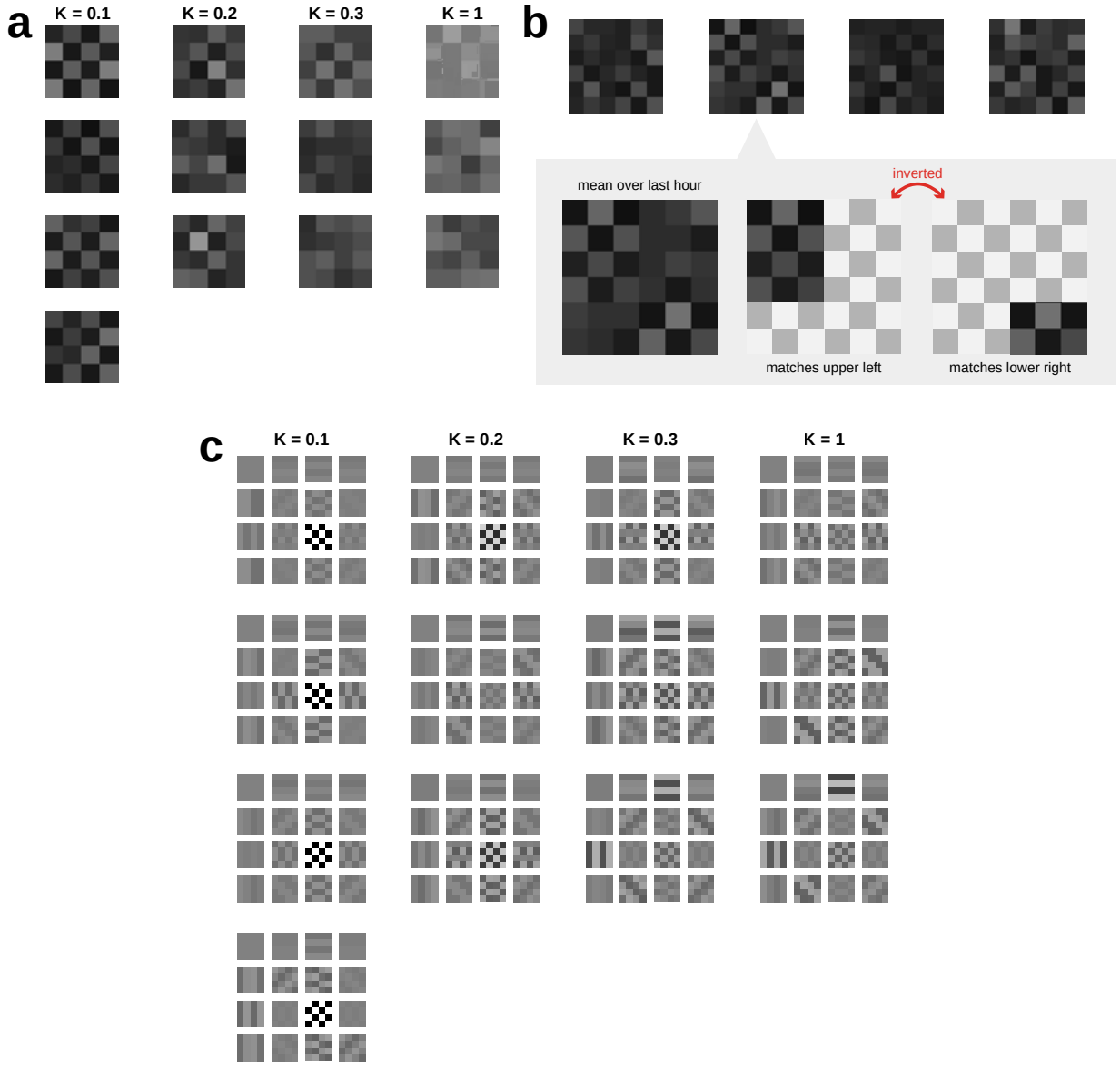

Figure 4.6: Checkerboards observed in all experiments. Visualizations are rendered from scores averaged over the last hour for (a) 16-patch and (b) 36-patch ($K = 0.1$) patterning experiments. Although the 36-patch experiments did not achieve global checkerboard patterning, one experiment showed persistent local patterning in which two small checkerboards appeared in opposing corners. The local patterns were inverted relative to each other and did not resolve before the end of the experiment. (c) 2D Fourier transforms conducted on the checkerboard averages from (1) reveal that the greatest weight is given to the highest-contrast spatial mode in all 4 experiments with $K = 0.1$, 2 out of 3 experiments with $K = 0.2$, 1 out of 3 experiments with $K = 0.3$, and none of the 3 experiments with $K = 1$. Pictured are the spectral components (horizontal frequency increasing top to bottom, vertical frequency increasing left to right) that contribute to last-hour checkerboards for each experiment. The intensity of each component indicates its weighting relative to other components in the same experiment. Note also the relatively higher weighting for lower-frequency spatial modes in the $K = 0.3$ and $K = 1$ cases relative to the $K = 0.1$ and $K = 0.2$ cases.

Taken together, these results suggest that the deterministic theory calibrated to population averages is an excellent quantitative predictor for mean system behavior across time, patches (cells), and experiments, while at the same time even small amounts of cell-cell variability and temporal stochasticity may cause a given experiment to deviate considerably from quantitative forecasts.

K	trial1	trial2	trial3
0.1	0	0.0002	0
0.1	0	0	0
0.1	0	0	0
0.1	0	0	0
0.2	0.0006	0.0008	0.0008
0.2	0.4422	0.4348	0.4356
0.2	0.0234	0.0226	0.017
0.3	0.0158	0.0156	0.0164
0.3	0.1262	0.1282	0.1322
0.3	0.146	0.1348	0.1388
1	0.6516	0.6444	0.645
1	0.4108	0.4116	0.4018
1	0.1966	0.1942	0.2144

Table 4.1: Differences in mean score between alternating sets of patches exceeded differences between random sets of patches. Two-tailed permutation tests for difference in mean score averaged across the last hour between sets of alternating patches, conducted against empirical distributions constructed from N = 5000 relabelings for each experiment individually. The empirical distributions of difference in means between relabeled sets of patches were randomly drawn three times per experiment and the resulting p-value calculated for the alternating patch allocation. Tests were performed against the null hypothesis that scores were drawn from the same distribution for all patches. For $K = 0.1$, the null hypothesis was rejected at a p-value of 0.0002 in all four experiments. For $K = 1$, the null hypothesis failed to be rejected at any reasonable p-value. Results were less clear for values close to the critical point, for which one experiment with $K = 0.2$ (theoretically bistable) failed to reject at any reasonable p-value, and one experiment with $K = 0.3$ (theoretically monostable) rejected at a p-value of 0.02.

Because our setup allowed us to monitor both gene expression levels and cell signaling levels, we were able to assess convergence to (quasi-)steady state by comparing instantaneous input-output curves (patch score vs. administered intensity) to the steady-state dose response (Figure 4.7c). Specifically, since the time for cells to converge to steady state under exposure to light of constant intensity (40 min) was longer than the time between changes to input intensity during experiments (10 min), the instantaneous input-output curve during a patterning experiment would only match the empirical dose response curve if the administered intensity remained relatively constant for several frames before a given time point—i.e., if there was little temporal variability for at least 40 min preceding the frame. Directly plotting the temporal variability in administered input to individual patches does indeed reveal

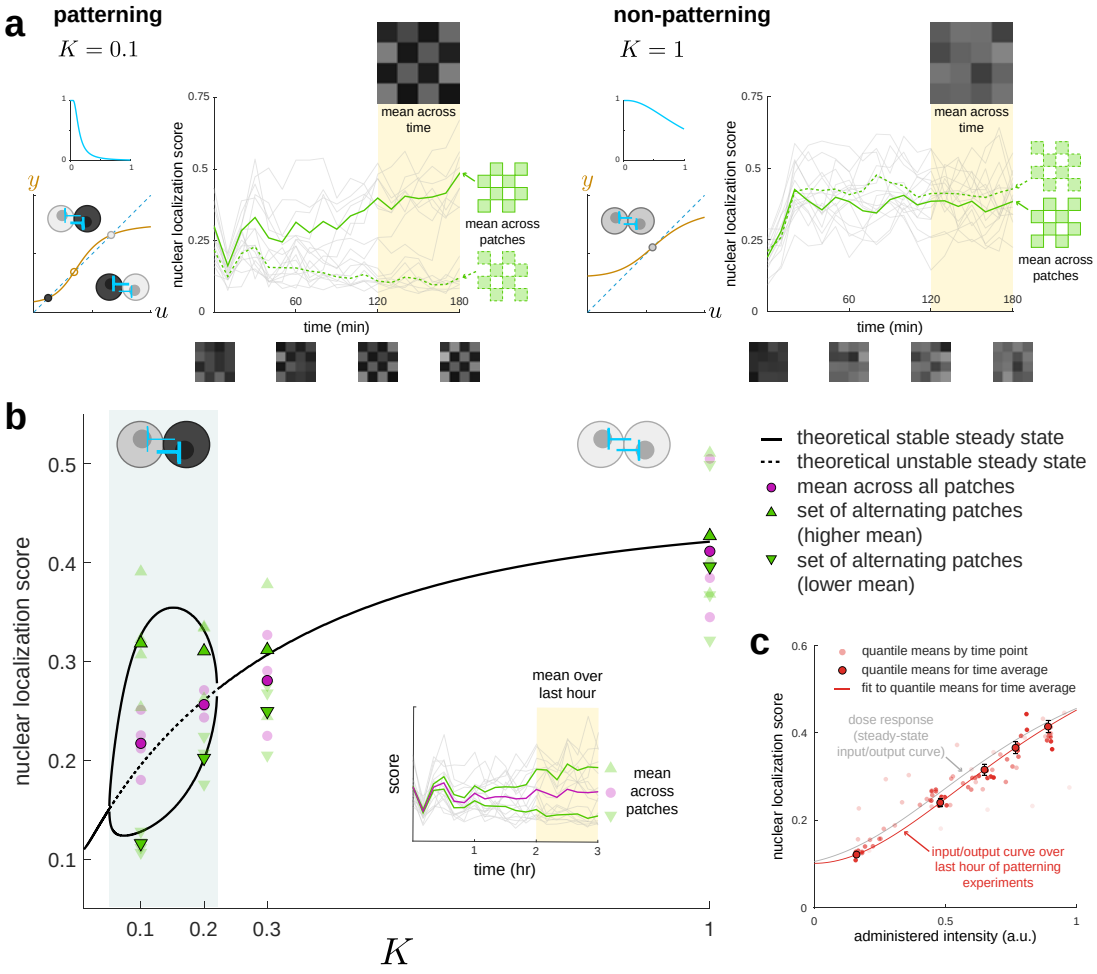

Figure 4.7: Theory quantitatively predicts spontaneous patterning and patch intensity during cell-in-the-loop experiments. (a) **Sample time traces show emergence of a contrasting pattern ($K = 0.1$) or convergence to a non-patterning state ($K = 1$).** Gray lines correspond to score traces for individual patches; green lines indicate the mean scores of sets of alternating patches. Checkerboards visualize scores at single time points (bottom) or averaged over the last hour (top). (b) **Theory quantitatively recapitulates experimental results for mean patch response score.** Black lines denote theoretical steady-state points as a function of the bifurcation parameter K, with solid indicating stable and dashed, unstable. All points are averaged over the last hour. Faded points correspond to individual experiments; solid outlines are averages across experiments ($N = 4$ for $K = 0.1$, $N = 3$ for $K = 0.2, 0.3, 1$). Magenta circular points are averages over all patches. Green points are averages over sets of alternating patches, with upward- and downward-facing triangles denoting the higher and lower of the two means respectively. Stochasticity in the experimental system introduced contrast between the average means of sets even in regions that were deterministically suggested to be monostable. As predicted, the contrast level (difference between means of sets of alternating patches) was higher for lower K. Experiment-to-experiment variation in overall brightness (score averaged across all patches) was greater for higher K, an effect that cannot be accounted for in a purely deterministic theory. (c) **The 16-patch system with 6 cells per patch converges to a steady state by about 2 hr into patterning experiments.** Solid outline, score values averaged over the last hour for individual patches and split into quantiles by administered intensities show decent agreement with the empirical steady-state dose response curve. Error bars are s.e.m. Circles without outlines are quantiles at individual time points pooled from $N = 13$ experiments (darker red at later times). For comparison, the dose response curve fit to empirical data is plotted in gray.

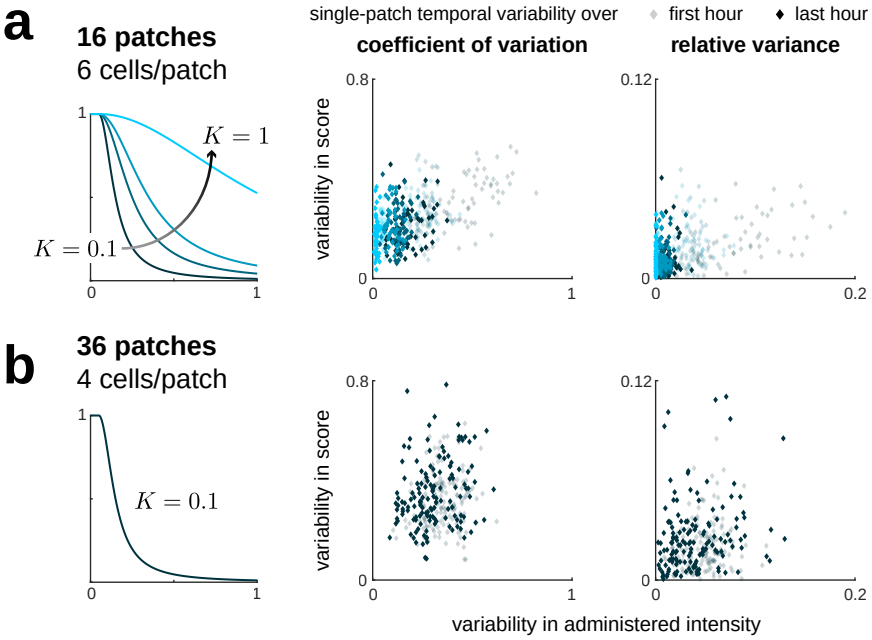

Figure 4.8: Temporal variability in administered intensity during first or last experimental hour. (a) Temporal variability in score and administered intensity to individual patches decreased from the first to the last experimental hour for all 16-patch patterning experiments, while temporal variability in score remained relatively constant. The result is not surprising if cells were comparably variable regardless of input intensity, in which case convergence of the mean score need not imply reduction in temporal variability. The administered intensity may also appear less variable since it was calculated based on an average over the scores of 4 neighbors, which at a given time step would have a variance $\frac{1}{16}$ that of an individual patch. (b) In four experiments with 36 patches and 4 cells per patch, temporal variability in administered intensity and score for individual patches remained equally high throughout the experiment duration (3 hr), suggesting steady state was not reached in that time.

a decrease from the first to the last experimental hour regardless of K value (Figure 4.8a).

Lastly, we examined the effect of patch number on patterning outcomes through four experiments with 36 patches, 4 cells per patch, and $K = 0.1$. None of the experiments spontaneously achieved a checkerboard pattern across the whole board in 3 hr, although a control experiment preinduced with the pattern showed that it was indeed persistent (Figure 4.9). The input/output curve did not approach the empirical dose response (Figure 4.10) and temporal variability in administered intensity was the same during the first and third experimental hour (Figure 4.8b), further supporting the conclusion that the system never reached steady state. Interestingly, one experiment produced two checkerboards in opposite corners that persisted throughout the last experimental hour, but were inverted relative to each other and did not resolve before the end of the experiment (Figure 4.6b). Other experiments also exhibited transient local patterning, although to a lesser degree. The local patterning and the increased convergence time are consequences of the fact that a 36-

patch system admits a much larger space of possible configurations than a 16-patch system. Although variability in 4-cell patches was only modestly larger than in 6-cell patches (Figure 4.4), the stochasticity may also have contributed to a longer convergence time. These and related challenges will require further investigation in future efforts to synthetically generate gene expression patterns with single-cell granularity.

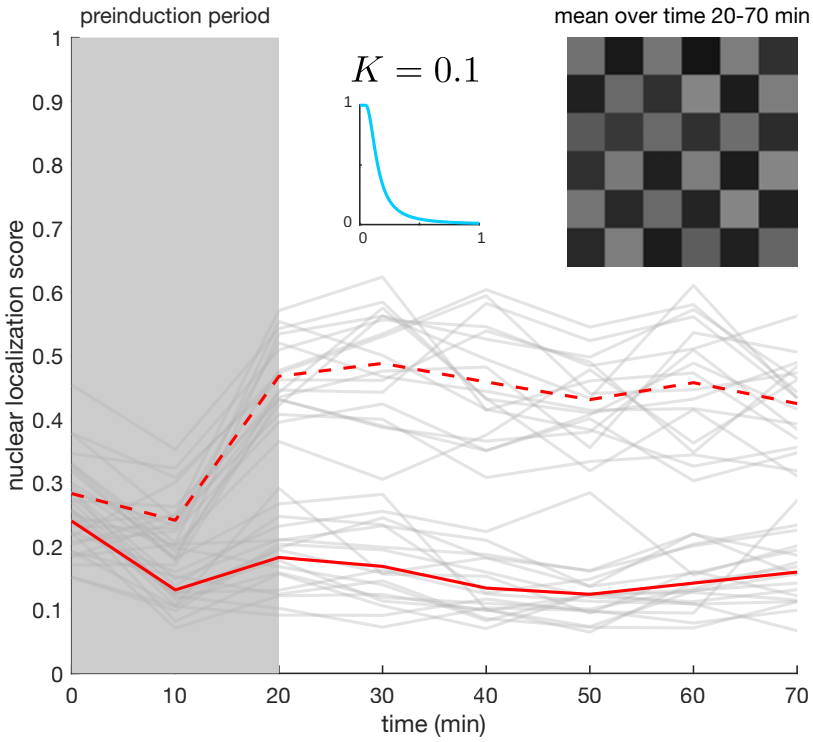

Figure 4.9: Control experiment verifying persistence of checkerboard pattern in full lateral inhibition system with 36 patchs, 4 cells/patch. Patches were preinduced for 20 min to display a checkerboard pattern. From 20 to 70 min, the system was run in closed-loop with lateral inhibition signaling relation $K = 0.1$. Pictured is the time course for individual patch scores (gray) with averages over sets of alternating patches in red (as in Figure 4.5). Board is visualization for individual patch scores averaged over the last 50 min.

4.3 Discussion

In this work, we employed cell-in-the-loop, a closed-loop, hybrid *in vivo*/*in silico* approach, to validate a theory for spontaneous gene expression patterning among laterally inhibiting cells. We engineered *S. cerevisiae* to respond optogenetically to light inputs, then emulated cell-to-cell signaling in real time by modulating the intensity of light inputs to cells based on real-time measurements of gene expression. The theory made accurate quantitative predictions for average steady-state patterning outcomes across a range of parameters.

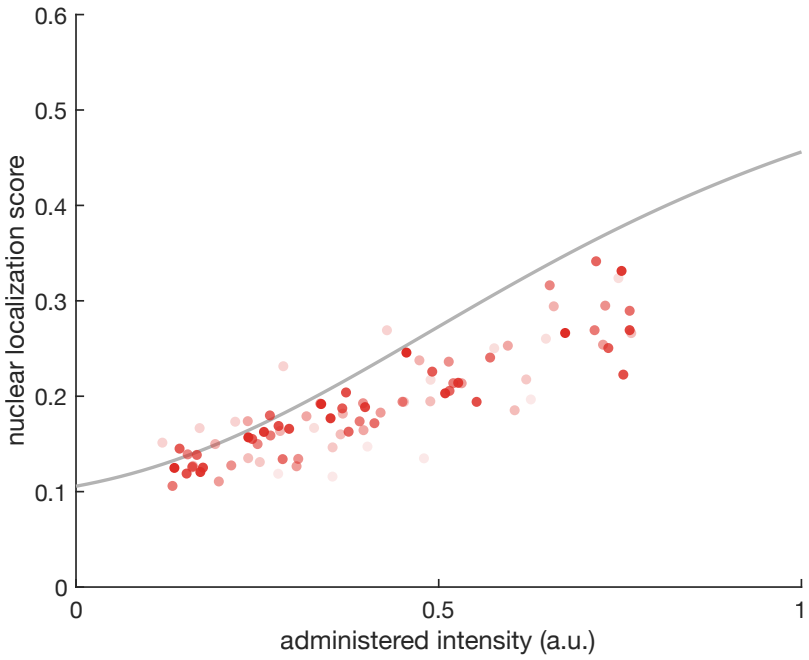

Figure 4.10: Administered input vs. measured patch score over time in the 36-patch experiments. The scatterplot does not show convergence to the steady-state dose response curve in 3 hr, unlike in the 16-patch case (Figure 4.7). Points are quantiles across $N = 4$ experiments with $K = 0.1$ for single time points, with more opaque (darker) points corresponding to later times.

Increasing system size—by increasing the space of possible dynamic behaviors—diminished the probability of achieving global patterning on short timescales in the absence of initial or external bias. Further theroetical research should explicitly incorporate cell-cell variability and temporal stochasticity in order to improve our understanding of variation in individual experimental outcomes, patterning robustness, and the link between individual-level and population-level behavior.

Prior work using optogenetics to generate persistent spatial patterns in living cells has focused on reproducing [167, 177, 331] or processing [166] pre-existing images projected by the light input. In comparison to these studies, light in our system does not *a priori* encode a pattern to which the cells conform; rather, light acts as a virtual signal transmitted from cell to cell. The input intensity is determined by cellular responses that are in turn influenced by the received intensity, establishing a closed-loop relationship independent of external control. That similar patterns are ultimately observed in both the cellular responses and the optogenetic inputs arises as a consequence of their mutual dependence.

Depending on the application, cell-in-the-loop offers benefits over purely biochemical or purely computational approaches. First, it reduces the number of components that must be engineered into cells. We integrated a single optogenetically induced promoter and a single reporter, and were able to modulate patterning outcomes simply by reprogramming the com-

puter. In this way, we circumvented issues associated with synthetic cell-to-cell signaling, including parameter matching and crosstalk [175, 307], and alleviated complications such as burden [382, 383] that arise from integrating complex networks into cells [172, 173]. We were also able to achieve spatiotemporal control over the whole population of interacting individuals and probe stochasticity at a finer level than would be attainable with a conventional biochemical implementation.

Compared to a computer simulation, cell-in-the-loop makes no assumptions about cell behavior or the form of biological noise, since the cells themselves are incorporated into the system. Although we used this setup to test the validity of a theoretical principle, one could also envision testing the accuracy of a full model for cell-to-cell signaling by simulating a proposed physical mechanism of interaction, then comparing the outcome of such a system to the outcome of a purely physical system. Cell-in-the-loop also allows one to track system components that might otherwise be inaccessible or difficult to measure. For example, we were able to monitor the levels of both gene expression and virtual signal simultaneously, which could be difficult to achieve in a solely biochemical setup.

Once established, a cell-in-the-loop system could couple with more complex cellular processes to achieve real-world results. One could also envision using cell-in-the-loop as a rapid prototyping platform or "stepping stone" to a fully biochemical implementation. In this paradigm, one would begin with minimally engineered cells and then sequentially replace *in silico* components with biochemical ones, testing at each stage whether the remaining portions of the network ought to be modified in structure or value before the next component is incorporated into the cell. As an illustration, suppose one were to implement lateral inhibition with the basic transcriptional repressor circuit from [163], in which cleaved intracellular Notch domain activates a repressor for expression of Delta, the target molecule of Notch recognition domain. One might employ the synNotch toolbox for mammalian cells, which includes as design choices the recognition domain/target molecule pair and the transcription factor constituting the intracellular synNotch domain [369]. One might begin with a cell-in-the-loop approach by engineering just the repressor under optogenetic control, and define computational variables for the concentrations of recognition domain, transcriptional activator (cleaved synNotch intracellular domain), and target molecule (plus intervening pathway components) in each cell. Using calculated activator concentration to modulate the light input, one could simulate different parameters for the remaining components to decide upon an activator. One could then engineer a circuit with repressor controlled by the chosen activator and place the activator under optogenetic control. With another round of cell-in-the-loop, one could further restrict the range of parameters for the combination of target molecule and recognition domain that would produce the desired effect *in vivo*, and use these numbers to guide construction of the final biochemical circuit in live cells.

Although the above example has relatively few components and therefore few design choices, the benefits of cell-in-the-loop should scale with circuit complexity as the parameter space—and therefore the potential number of full circuits to test—increases with each added component. The approach could also reveal shortcomings in proposed designs: for example, in our setup, 36-patch systems failed to produce spontaneous patterns on our experimental

timescale even with "perfect" deterministic signaling, suggesting that it could be challenging to achieve large-scale lateral inhibition patterning biochemically in a similar context. Thus, the addition of cell-in-the-loop to the biological engineering process could greatly decrease the time, expense, and effort required to develop synthetic multicellular systems for an increasingly rich and promising array of applications.

Acknowledgments

This work appears in [308], in which the authors thank Marc Rul·lan for his help with the setup and guidance throughout the project, as well as Brian M. Lang for his advice on statistics and general linear modeling for the dose response data. This project received funding from the United States Air Force Office of Scientific Research (AFOSR) grants FA9550-14-1-0089 and FA9550-18-1-0253, the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (CyberGenetics; grant agreement 743269), and Swiss National Foundation (SNF) grant 316030_177079.

Chapter 5

Conclusion: Multicellular Patterning, Evolved and Engineered

"The living organism is so complicated that we seldom have enough data to be able to work out exactly what is happening by means of the mathematics the engineer uses. Up to the present, the general ideas and terminology used by these engineers have been of more use to biologists than have the detailed application of their techniques."

– J. Z. Young, in [384]

In this thesis, I have explored multicellular gene expression patterning in biological systems using a combination of theoretical and experimental approaches informed by networked dynamical systems theory. First, I presented a general tool for analyzing prepattern processing from the perspective of spatial filtering. I demonstrated how this approach might be used to predict qualitative patterning outcomes, suggest experiments for discriminating among proposed genetic network architectures, and probe the robustness of outcomes to noisy biochemical components. I also showed that random, constant-in-time spatial variation can generate Turing-like patterns when system dynamics are stable; conventional wisdom requires an instability to observe such patterns.

Second, I adapted theory for spontaneous contrasting patterning to two distinct experimental setups in synthetic biology. Both setups verified aspects of the theory, lending credence to its validity and generality in practice. Both setups demonstrate how theory can analogize a known concept to classes of unconventional systems—in one case, lateral inhibition among spatially separated colonies instead of individual cells in physical contact, and in the other case, inhibition among cells that are not physically communicating or even physically neighboring each other. Finally, both setups permit further applications beyond the theory, and may thus be of some interest to the synthetic biological community.

The work presented in this thesis focuses on patterning at steady state. It is not clear, however, that biological systems in nature operate very often at steady state, and the dynamics of patterning networks in particular may have enormous consequences. In fact, both of the experimental studies in this thesis evidenced situations where practical constraints prevented theoretically predicted patterning behaviors from emerging on an experimental

timescale. The compartmental lateral inhibition system failed to spontaneously pattern before colonies outgrew their compartments, although patterns could be retained through preinduction. Similarly, the largest tested networks in the cell-in-the-loop setup could also be preinduced to retain a contrasting pattern, but did not spontaneously pattern because local defects failed to resolve into a globally consistent pattern before cells outgrew a monolayer. Future theoretical work should focus on incorporating an understanding of patterning dynamics at a level more rigorous than simulation. I expect such research would be of particular relevance to developmental biology with regard to the process of pattern refinement.

A central theme throughout this thesis is how engineering approaches can provide new insights into the structure and function of living systems—actual or potential—by considering them with a different perspective from the usual. It is my belief that continued cross-disciplinary dialog among engineers, biologists, mathematicians, physicists, and others will facilitate our efforts to uncover the conceptual and molecular mechanisms behind patterning in the natural world. This work is just one such attempt to bridge the gap between theory and practice, between evolved and engineered, that we may one day better comprehend the "endless forms most beautiful" [385] of life on Earth.

Bibliography

[1] Doyle, A. C. A case of identity. In *The Adventures of Sherlock Holmes* (1891).

[2] Turing, A. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London* **237**, 37–72 (1952).

[3] Murray, J. D. Discussion: Turing's theory of morphogenesis–its influence on modelling biological pattern and form. *Bulletin of Mathematical Biology* **52**, 117–152 (1990).

[4] Othmer, H. G., Painter, K., Umulis, D. & Xue, C. The intersection of theory and application in elucidating pattern formation in developmental biology. *Mathematical Modelling of Natural Phenomena* **4**, 3–82 (2009).

[5] Green, J. B. A. & Sharpe, J. Positional information and reaction-diffusion: two big ideas in developmental biology combine. *Development* **142**, 1203–1211 (2015).

[6] Harrison, L. G. What is the status of reaction-diffusion theory thirty-four years after Turing? *Journal of Theoretical Biology* **125**, 369–384 (1987).

[7] Jung, H.-S. *et al.* Local inhibitory action of BMPs and their relationships with activators in feather formation: Implications for periodic patterning. *Developmental Biology* **196**, 11–23 (1998).

[8] Sick, S., Reinker, S., Timmer, J. & Schlake, T. Wnt and dkk determine hair follicle spacing through a reaction-diffusion mechanism. *Science* **314**, 1447–1450 (2006).

[9] Watanabe, M. & Kondo, S. Is pigment patterning in fish skin determined by the turing mechanism? *Trends in Genetics* **31**, 88–96 (2015).

[10] Economou, A. D. *et al.* Periodic stripe formation by a Turing mechanism operating at growth zones in the mammalian palate. *Nature Genetics* **44**, 348–351 (2012).

[11] Gierer, A. & Meinhardt, H. A theory of biological pattern formation. *Kybernetik* **12**, 30–39 (1972).

[12] Koga, S. & Kuramoto, Y. Localized patterns in reaction-diffusion systems. *Progress of Theoretical Physics* **63**, 106–121 (1980).

[13] Vanag, V. K. & Epstein, I. R. Localized patterns in reaction-diffusion systems. *Chaos* **17**, 037110 (2007).

[14] Maini, P. K., Benson, D. L. & Sherratt, J. A. Pattern formation in reaction-diffusion mdoels with spatially inhomogeneous diffusion coefficients. *IMA Journal of Mathematics Applied in Medicine and Biology* **9**, 197–213 (1992).

[15] Page, K. M., Maini, P. K. & Monk, N. A. M. Complex pattern formation in reaction-diffusion systems with spatially varying parameters. *Physica D: Nonlinear Phenomena* **202**, 95–115 (2005).

[16] Kolokolnikov, T. & Wei, J. Pattern formation in a reaction-diffusion system with space-dependent feed rate. *SIAM Review* **60**, 626–645 (2018).

[17] Butler, T. & Goldenfeld, N. Fluctuation-driven Turing patterns. *Physical Review E* (2011).

[18] Biancalani, T., Jafarpour, F. & Goldenfeld, N. Giant amplification of noise in fluctuation-induced pattern formation. *Physical Review Letters* **118** (2017).

[19] Smith, S. & Dalchau, N. Beyond activator-inhibitor networks: the generalised turing mechanism. *arXiv* (2018).

[20] Cencetti, G., Battiston, F., Carletti, T. & Fanelli, D. Generalized patterns from local and non local reactions. *Chaos, Solitons & Fractals* **134**, 109707 (2020).

[21] Halatek, J. & Frey, E. Rethinking pattern formation in reaction-diffusion systems. *Nature Physics* **14**, 507–514 (2018).

[22] Landge, A. N., Jordan, B. M., Diego, X. & Müller, P. Pattern formation mechanisms of self-organizing reaction-diffusion systems. *Developmental Biology* (2020).

[23] Lawrence, P. A. Morphogens: how big is the big picture? *Nature Cell Biology* **3**, E151–E154 (2001).

[24] Gurdon, J. B. & Bourillot, P.-Y. Morphogen gradient interpretation. *Nature* **413**, 797–803 (2001).

[25] Lander, A. D. Morpheus unbound: Reimagining the morphogen gradient. *Cell* **128**, 245–256 (2007).

[26] Davidson, E. H. Emerging properties of animal gene regulatory networks. *Nature* **468**, 911–920 (2010).

[27] Rogers, K. W. & Schier, A. F. Morphogen gradients: from generation to interpretation. *Annual Review of Cell and Developmental Biology* **27**, 377–407 (2011).

[28] Wolpert, L. Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology* **25**, 1–47 (1969).

[29] Crick, F. Diffusion in embryogenesis. *Nature* **225**, 420–422 (1970).

[30] Tickle, C., Summerbell, D. & Wolpert, L. Positional signalling and specification of digits in chick limb morphogenesis. *Nature* **254**, 199–202 (1975).

[31] Eichele, G. & Thaller, C. Characterization of concentration gradients of a morphogenetically active retinoid in the chick limb bud. *The Journal of Cell Biology* **105**, 1917–1923 (1987).

[32] Maden, M. Vitamin a and pattern formation in the regenerating limb. *Nature* **295**, 672–675 (1982).

[33] Driever, W. & Nüsslein-Volhard, C. The *bicoid* protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell* **54**, 95–104 (1988).

[34] Ferguson, E. L. & Anderson, K. V. *decapentaplegic* acts as a morphogen to organize dorsal-ventral pattern in the Drosophila embryo. *Cell* **71**, 451–461 (1992).

[35] Sampath, K. *et al.* Induction of the zebrafish ventral brain and floorplate requires cyclops/nodal signalling. *Nature* **395**, 185–189 (1998).

[36] Chen, Y. & Schier, A. F. The zebrafish Nodal signal Squint functions as a morphogen. *Nature* **411**, 607–610 (2001).

[37] Gurdon, J. B., Harger, P., Mitchell, A. & Lemaire, P. Activin signalling and response to a morphogen gradient. *Nature* **371**, 487–492 (1994).

[38] Ericson, J., Briscoe, J., Rashbass, P., van Heyningen, V. & Jessell, T. M. Graded Sonic hedgehog signaling and the specification of cell fate in the ventral neural tube. *Cold Spring Harbor Symposia on Quantitative Biology* **62**, 451–466 (1997).

[39] Briscoe, J., Chen, Y., Jessell, T. M. & Struhl, G. A Hedgehog-insensitive form of Patched provides evidence for direct long-range morphogen activity of Sonic hedgehog in the neural tube. *Molecular Cell* **7**, 1279–1291 (2001).

[40] Neumann, C. & Cohen, S. Morphogens and pattern formation. *BioEssays* **19**, 721–729 (1997).

[41] Gregor, T., Wieschaus, E. F., McGregor, A. P., Bialek, W. & Tank, D. W. Stability and nuclear dynamics of the bicoid morphogen gradient. *Cell* **130**, 141–152 (2007).

[42] Wartlick, O., Kicheva, A. & González-Gaitán. Morphogen gradient formation. *Cold Spring Harbor Perspectives in Biology* **1** (2009).

[43] Müller, P., Rogers, K. W., Yu, S. R., Brand, M. & Schier, A. F. Morphogen transport. *Development* **140**, 1621–1638 (2013).

[44] Shilo, B.-Z., Haskel-Ittah, M., Ben-Zvi, D., Schejter, E. D. & Barkai, N. Creating gradients by morphogen shuttling. *Trends in Genetics* **29**, 339–347 (2013).

[45] Bryant, S. V. & Gardiner, D. M. The relationship between growth and pattern formation. *Regeneration* **3**, 103–122 (2016).

[46] Rogers, K. W. & Müller, P. Nodal and bmp dispersal during early zebrafish development. *Developmental Biology* **447**, 14–23 (2019).

[47] Spemann, H. & Mangold, H. über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Archiv für mikrokopische Anatomie und Entwicklungsmechanik* **100**, 599–638 (1924).

[48] Maini, P. K., Baker, R. E. & Chuong, C.-M. The Turing model comes of molecular age. *Science* **314**, 1397–1398 (2006).

[49] Murray, J. D. Parameter space for Turing instability in reaction diffusion mechanisms: A comparison of models. *Journal of Theoretical Biology* **98**, 143–163 (1982).

[50] Koch, A. J. & Meinhardt, H. Biological pattern formation: from basic mechanisms to complex structures. *Reviews of Modern Physics* **66**, 1481–1507 (1994).

[51] Dillon, R., Maini, P. K. & Othmer, H. G. Pattern formation in generalized Turing systems. *Journal of Mathematical Biology* **32**, 345–393 (1994).

[52] Bard, J. & Lauder, I. How well does Turing's theory of morphogenesis work? *Journal of Theoretical Biology* **25**, 501–531 (1974).

[53] Arcuri, P. & Murray, J. D. Pattern sensitivity to boundary and initial conditions in reaction-diffusion models. *Journal of Mathematical Biology* **24**, 141–165 (1986).

[54] Maini, P. K., Woolley, T. E., Baker, R. E., Gaffney, E. A. & Lee, S. S. Turing's model for biological pattern formation and the robustness problem. *Journal of the Royal Society Interface Focus* **2** (2012).

[55] Yi, F., Gaffney, E. A. & Seirin-Lee, S. The bifurcation analysis of Turing pattern formation induced by delay and diffusion in the Schnakenberg system. *Discrete and Continuous Dynamical Systems Series B* **22**, 647–668 (2017).

[56] Lengyel, I. & Epstein, I. R. A chemical approach to designing Turing patterns in reaction-diffusion systems. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 3977–3979 (1992).

[57] Pearson, J. E. & Bruno, W. J. Pattern formation in an n+q component reaction-diffusion system. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2** (1992).

[58] Satnoianu, R. A., Menzinger, M. & Maini, P. K. Turing instabilities in general systems. *Journal of Mathematical Biology* **41**, 493–512 (2000).

[59] Strier, D. E. & Dawson, S. P. Role of complexing agents in the appearance of Turing patterns. *Physical Review E* **69**, 066207 (2004).

[60] Korvasová, K., Gaffney, E. A., Maini, P. K., Ferreira, M. A. & Klika, V. Investigating the Turing conditions for diffusion-driven instability in the presence of a binding imnmobile substrate. *Journal of Theoretical Biology* **367**, 286–295 (2015).

[61] Marcon, L., Diego, X., Sharpe, J. & Müller, P. High-throughput mathematical analysis identifies Turing networks for patterning with equally diffusing signals. *eLife* **5**, e14022 (2016).

[62] Zheng, M. M., Shao, B. & Ouyang, Q. Identifying network topologies that can generate Turing pattern. *Journal of Theoretical Biology* **308**, 88–96 (2016).

[63] Diego, X., Marcon, L., Müller, P. & Sharpe, J. Key features of turing systems are determined purely by network topology. *Physical Review X* **8** (2018).

[64] Smith, S. & Dalchau, N. Model reduction enables turing instability analysis of large reaction-diffusion models. *Journal of the Royal Society Interface* **15** (2018).

[65] Scholes, N. S., Schnoerr, D., Isalan, M. & Stumpf, M. P. H. A comprehensive network atlas reveals that turing patterns are common but not robust. *Cell Systems* **9**, 243–257 (2019).

[66] Economou, A. D., Monk, N. A. M. & Green, J. B. A. Perturbation analysis of a multi-morphogen Turing reaction-diffusion stripe patterning system reveals key regulatory interactions. *bioRxiv* .

[67] Crampin, E. J., Gaffney, E. A. & Maini, P. K. Reaction and diffusion on growing domains: scenarios for robust pattern formation. *Bulletin of Mathematical Biology* **61**, 1093–1120 (1999).

[68] Woolley, T. E., Baker, R. E., Gaffney, E. A. & Maini, P. K. Influence of stochastic domain growth on pattern nucleation for diffusive systems with internal noise. *Physical Review E* **84**, 041905 (2011).

[69] Ben-Zvi, D. & Barkai, N. Scaling of morphogen gradients by an expansion-repression integral feedback control. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 6924–6929 (2010).

[70] Umulis, D. M. & Othmer, H. G. Mechanisms of scaling in pattern formation. *Development* **140**, 4830–4843 (2013).

[71] Averbukh, I., Ben-Zvi, D., Mishra, S. & Barkai, N. Scaling morphogen gradients during tissue growth by a cell division rule. *Development* **141**, 2150–2156 (2014).

[72] Čapek, D. & Müller, P. Positional information and tissue scaling during development and regeneration. *Development* **146**, dev177709 (2019).

[73] McHale, P., Rappel, W.-J. & Levine, H. Embryonic pattern scaling achieved by oppositely directed morphogen gradients. *Physical Biology* **3** (2006).

[74] Xu, P.-F., Houssin, N., Ferri-Lagneau, K. F., Thisse, B. & Thisse, C. Construction of a vertebrate embryo from two opposing morphogen gradients. *Science* **344**, 87–89 (2014).

[75] Zagorski, M. *et al.* Decoding of position in the developing neural tube from antiparallel morphogen gradients. *Science* **356**, 1379–1383 (2017).

[76] Eldar, A. *et al.* Robustness of the BMP morphogen gradient in *Drosophila* embryonic patterning. *Nature* **419**, 304–308 (2002).

[77] Tostevin, F., Rein ten Wolde, P. & Howard, M. Fundamental limits to position determination by concentration gradients. *PLoS Computational Biology* **3**, e78 (2007).

[78] Gregor, T., Tank, D. W., Wieschaus, E. F. & Bialek, W. Probing the limits to positional information. *Cell* **130**, 153–164 (2007).

[79] Erdmann, T., Howard, M. & Rein ten Wolde, P. Role of spatial averaging in the precision of gene expression patterns. *Physical Review Letters* **103**, 258101 (2009).

[80] Rodrigo, G. & Elena, S. F. Structural discrimination of robustness in transcriptional feedforward loops for pattern formation. *PLoS One* **6**, e16904 (2011).

[81] Dubuis, J. O., Tkačik, G., Wieschaus, E. F., Gregor, T. & Bialek, W. Positional information, in bits. *Proceedings of the National Academy of Sciences of the United States of America* **11**, 16301–16308 (2013).

[82] Lo, W.-C., Zhou, S., Wan, F. Y.-M., Lander, A. D. & Nie, Q. Robust and precise morphogen-mediated patterning: trade-offs, constraints and mechanisms. *Journal of the Royal Society Interface* **12**, 20141041 (2015).

[83] Perez-Carrasco, R., Guerrero, P., Briscoe, J. & Page, K. M. Intrinsic noise profoundly alters the dynamics and steady state of morphogen-controlled bistable genetic switches. *PLoS Computational Biology* **12**, e1005154 (2016).

[84] Akieda, Y. *et al.* Cell competition corrects noisy Wnt morphogen gradients to achieve robust patterning in the zebrafish embryo. *Nature Communications* **10**, 4710 (2019).

[85] Bergmann, S. *et al.* Pre-steady-state decoding of the Bicoid morphogen gradient. *PLoS Biology* **5**, e46 (2007).

[86] Saunders, T. & Howard, M. When it pays to rush: interpreting morphogen gradients prior to steady-state. *Physical Biology* **6**, 046020 (2009).

[87] Tamari, Z. & Barkai, N. Improved readout precision of the Bicoid morphogen gradient by early decoding. *Journal of Biological Physics* **38**, 317–329 (2012).

[88] Huang, A., Amourda, C., Zhang, S., Tolwinski, N. S. & Saunders, T. E. Decoding temporal interpretation of the morphogen Bicoid in the early *Drosophila* embryo. *eLife* **6**, e26258 (2017).

[89] Sagner, A. & Briscoe, J. Morphogen interpretation: concentration, time, competence, and signaling dynamics. *WIREs Developmental Biology* **6**, e271 (2017).

[90] Verd, B., Crombach, A. & Jaeger, J. Dynamic maternal gradients control timing and shift-rates for *Drosophila* gap gene expression. *PLoS Computational Biology* **13**, e1005285 (2017).

[91] Heemskerk, I. *et al.* Rapid changes in morphogen concentration control self-organized patterning in human embryonic stem cells. *eLife* **8**, e40526 (2019).

[92] Jaeger, J. *et al.* Dynamic control of positional information in the early *Drosophila* embryo. *Nature* **430**, 368–371 (2004).

[93] Cotterell, J. & Sharpe, J. An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. *Molecular Systems Biology* **6**, 425 (2010).

[94] Balaskas, N. *et al.* Gene regulatory logic for reading the Sonic hedgehog signaling gradient in the vertebrate neural tube. *Cell* **148**, 273–284 (2012).

[95] Sokolowski, T. R., Erdmann, T. & Rein ten WOlde, P. Mutual repression enhances the steepness and precision of gene expression boundaries. *PLoS Computational Biology* **8**, e1002654 (2012).

[96] Cohen, M., Page, K. M., Perez-Carrasco, R., Barnes, C. P. & Briscoe, J. A theoretical framework for the regulation of Shh morphogen-controlled gene expression. *Development* **141**, 3868–3878 (2014).

[97] Clark, E. Dynamic patterning by the *Drosophila* pair-rule network reconciles long-germ and short-germ segmentation. *PLoS Biology* **15**, e2002439 (2017).

[98] Verd, B., Monk, N. A. M. & Jaeger, J. Modularity, criticality,. and evolvability of a developmental gene regulatory network. *eLife* **8**, e42832 (2019).

[99] Cooke, J. & Zeeman, E. C. A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *Journal of Theoretical Biology* **58**, 455–476 (1976).

[100] Aulehla, A. & Herrmann, B. G. Segmentation in vertebrates: clock and gradient finally joined. *Genes & Development* **18**, 2060–2067 (2004).

[101] Lewis, J. Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. *Current Biology* **13**, 1398–1408 (2003).

[102] Monk, N. A. M. Oscillatory expression of Hes1, p53, and NF-κB driven by transcriptional time delays. *Current Biology* **13**, 1409–1413 (2003).

[103] Baker, R. E., Schnell, S. & Maini, P. K. A mathematical investigation of a *clock and wavefront* model for somitogenesis. *Journal of Mathematical Biology* **52**, 458–482 (2006).

[104] Baker, R. E., Schnell, S. & Maini, P. K. Mathematical models for somite formation. *Current Topics in Decvelopmental Biology* **81**, 183–203 (2008).

[105] Murray, P. J., Maini, P. K. & Baker, R. E. The clock and wavefront model revisited. *Journal of Theoretical Biology* **283**, 227–238 (2011).

[106] Lauschke, V. M., Tsiairis, C. D., François, P. & Aulehla, A. Scaling of embryonic patterning based on phase-gradient encoding. *Nature* **493**, 101–105 (2012).

[107] Ishimatsu, K. *et al.* Size-reduced embryos reveal a gradient scaling-based mechanism for zebrafish somite formation. *Development* **145**, dev161257 (2018).

[108] Horikawa, K., Ishimatsu, K., Yoshimoto, E., Kondo, S. & Takeda, H. Noise-resistant and syncrhonized oscillation of the segmentation clock. *Nature* **441**, 719–723 (2006).

[109] Riedel-Kruse, I. H., Müller, C. & Oates, A. C. Synchrony dynamics during initiation, failure, and rescue of the segmentation clock. *Science* **317**, 1911–1915 (2007).

[110] Naoki, H. *et al.* Noise-resistant developmental reproducibility in vertebrate somite formation. *PLoS Computational Biology* **15**, e1006579 (2019).

[111] Dias, A. S., de Almeida, I., Belmonte, J. M., Glazier, J. A. & Stern, C. D. Somites without a clock. *Science* **343**, 791–795 (2014).

[112] Cotterell, J., Robert-Moreno, A. & Sharpe, J. A local, self-organizing reaction-diffusion model can explain somite patterning in embryos. *Cell Systems* **1**, 257–269 (2015).

[113] Boareto, M., Tomka, T. & Iber, D. Positional information encoded in the dynamic differences between neighbouring oscillators during vertebrate segmentation. *bioRxiv* .

[114] Wigglesworth, V. B. Local and general factors in the development of "pattern" in *Rhodnius prolixus* (Hemiptera). *Journal of Experimental Biology* **17**, 180–201 (1940).

[115] Collier, J. R., Monk, N. A. M., Maini, P. K. & Lewis, J. H. Pattern formation by lateral inhibition with feedback: a mathematical model of Delta-Notch intercellular signalling. *Journal of Theoretical Biology* **183**, 429–446 (1996).

[116] Heitzler, P. & Simpson, P. The choice of cell fate in the epidermis of drosophila. *Cell* **64**, 1083–1092 (1991).

[117] Artavanis-Tsakonas, S., Matsuno, K. & Fortini, M. E. Notch signaling. *Science* **268**, 225–232 (1995).

[118] Lewis, J. Notch signalling and the control of cell fate choices in vertebrates. *Seminars in Cell and Developmental Biology* **9**, 583–589 (1998).

[119] Bryant, J., Goodyear, R. J. & Richardson, G. P. Sensory organ development in the inner ear: molecular and cellular mechanisms. *British Medical Bulletin* **63**, 39–57 (2002).

[120] Pierfelice, T., Alberi, L. & Gaiano, N. Notch in the vertebrate nervous system: An old dog with new tricks. *Neuron* **69**, 840–855 (2011).

[121] Vooijs, M., Liu, Z. & Kopan, R. Notch: Architect, landscaper, and guardian of the intestine. *Gastroenterology* **141**, 448–459 (2011).

[122] Ferreira, A. S. R. & Arcak, M. A graph partitioning approach to predicting patterns in lateral inhibition systems. *SIAM Journal on Applied Dynamical Systems* **12**, 2012–2031 (2013).

[123] Lee, S. S. Lateral inhibition-induced pattern formation controlled by the size and geometry of the cell. *Journal of Theoretical Biology* **404**, 51–65 (2016).

[124] Shaya, O. *et al.* Cell-cell contact area affects Notch signaling and Notch-dependent patterning. *Developmental Cell* **40**, 505–511.e6 (2017).

[125] Plahte, E. & Øyehaug, L. Pattern-generating travelling waves in a discrete multicellular system with lateral inhibition. *Physica D* **226**, 117–128 (2007).

[126] Sprinzak, D. *et al.* *Cis*-interactions between Notch and Delta generate mutually exclusive signalling states. *Nature* **465**, 86–90 (2010).

[127] Sprinzak, D., Lakhanpal, A., LeBon, L., Garcia-Ojalvo, J. & Elowitz, M. B. Mutual inactivation of Notch receptors and ligands facilitates developmental patterning. *PLoS Computational Biology* **7** (2011).

[128] Lubensky, D. K., Pennington, M. W., Shraiman, B. I. & Baker, N. E. A dynamical model of ommatidial crystal formation. *Proceedings of the National Academy of the United States of America* **108**, 11145–11150 (2011).

[129] Vasilopoulos, G. & Painter, K. J. Pattern formation in discrete cell tissues under long range filopodia-based direct cell to cell contact. *Mathematical Biosciences* **273**, 1–15 (2016).

[130] Corson, F., Couturier, L., Rouault, H., Mazouni, K. & Schweisguth, F. Self-organized Notch dynamics generate stereotyped sensory organ patterns in *Drosophila*. *Science* **356** (2017).

[131] Cohen, R. *et al.* Shear forces drive precise patterning of hair cells in the mammalian inner ear. *bioRxiv* .

[132] Mori, Y., Jilkine, A. & Edelstein-Keshet, L. Wave-pinning and cell polarity from a bistable reaction-diffusion system. *Biophysical Journal* **94**, 3684–3697 (2008).

[133] Mori, Y., Jilkine, A. & Edelstein-Keshet, L. Asymptotic and bifurcation analysis of wave-pinning in a reaction-diffusion model for cell polarization. *SIAM Journal on Applied Mathematics* **71**, 1401–1427 (2011).

[134] Middleton, A. M., King, J. R. & Loose, M. Wave pinning and spatial patterning in a mathematical model of Antivin/Lefty-Nodal signalling. *Journal of Mathematical Biology* **67**, 1393–1424 (2013).

[135] Zhu, X. *et al.* Speed regulation of genetic cascades allows for evolvability in the body plan specification of insects. *Proceedings of the National Academy of Sciences of the United States of America* E8646–E8655 (2017).

[136] Boos, A., Distler, J., Rudolf, H., Klingler, M. & El-Sherif, E. A re-inducible gap gene cascade patterns the anterior-posterior axis of insects in a threshold-free fashion. *eLife* **7**, e41208 (2018).

[137] Kuhlmann, L. & El-Sherif, E. Speed regulation and gradual enhancer switching models as flexible and evolvable patterning mechanisms. *bioRxiv* .

[138] Morelli, L. G., Uriu, K., Ares, S. & Oates, A. C. Computational approaches to developmental patterning. *Science* **336**, 187–191 (2012).

[139] Schweisguth, F. & Corson, F. Self-organization in pattern formation. *Developmental Cell* **49**, 659–677 (2019).

[140] Gilmour, D., Rembold, M. & Leptin, M. From morphogen to morphogenesis and back. *Nature* **541**, 311–320 (2017).

[141] Bashor, C. J. & Collins, J. J. Understanding biological regulation through synthetic biology. *Annual Review of Biophysics* **47**, 399–423 (2018).

[142] Shipman, S. L., Nivala, J., Macklis, J. D. & Church, G. M. Molecular recordings by directed CRISPR spacer acquisition. *Science* **353**, aaf1175 (2016).

[143] Church, G. M., Gao, Y. & Kosuri, S. Next-generation digital information storage in DNA. *Science* **337**, 1628 (2012).

[144] Chirieleison, S. M., Allen, P. B., Simpson, Z. B., Ellington, A. D. & Chen, X. Pattern transformation with DNA circuits. *Nature Chemistry* **5**, 1000–1005 (2013).

[145] Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).

[146] Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).

[147] Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593 (2000).

[148] Weiss, R. & Knight, T. F. Engineered communications for microbial robotics. In Condon, A. & Rozenberg, G. (eds.) *DNA Computing*, 1–16 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2001).

[149] Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130–1134 (2005).

[150] Sohka, T. *et al.* An externally tunable bacterial band-pass filter. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 10135–10140 (2009).

[151] Greber, D. & Fussenegger, M. An engineered mammalian band-pass network. *Nucleic Acids Research* **38**, e174 (2010).

[152] Kämpf, M. M. *et al.* Rewiring and dosing of systems modules as a design approach for synthetic mammalian signaling networks. *Molecular BioSystems* **8**, 1824–1832 (2012).

[153] Schaerli, Y. *et al.* A unified design space of synthetic stripe-forming networks. *Nature Communications* (2014).

[154] Kong, W., Blanchard, A. E., Liao, C. & Lu, T. Engineering robust and tunable spatial structures with synthetic gene circuits. *Nucleic Acids Research* **45**, 1005–1014 (2017).

[155] Hsia, J., Holtz, W. J., Huang, D. C., Arcak, M. & Maharbiz, M. M. A feedback quenched oscillator produces Turing patterning with one diffuser. *PLoS Computational Biology* **8**, e1002331 (2012).

[156] Miyazako, H., Hori, Y. & Hara, S. Turing instability in reaction-diffusion systems with a single diffuser: characterization based on root locus. In *52nd IEEE Conference on Decision and Control*, 2671–2676 (2013).

[157] Diambra, L., Senthivel, V. R., Menendez, D. B. & Isalan, M. Cooperativity to increase Turing pattern space for synthetic biology. *ACS Synthetic Biology* **4**, 177–186 (2015).

[158] Glock, P. *et al.* Stationay patterns in a two-protein reaction-diffusion system. *ACS Synthetic Biology* **8**, 148–157 (2019).

[159] Karig, D. *et al.* Stochastic Turing patterns in a synthetic bacterial population. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 6527–6577 (2018).

[160] Sekine, R., Shibata, T. & Ebisuya, M. Synthetic mammalian pattern formation driven by differential diffusivity of Nodal and Lefty. *Nature Communications* **9** (2018).

[161] Potvin-Trottier, L., Lord, N. D., Vinnicombe, G. & Paulsson, J. Synchronous long-term oscillations in a synthetic gene circuit. *Nature* **538**, 514–517 (2016).

[162] Wu, F. *et al.* A synthetic biology approach to sequential stripe patterning and somitogenesis. *bioRxiv* .

[163] Matsuda, M., Koga, M., Woltjen, K., Nishida, E. & Ebisuya, M. Synthetic lateral inhibition governs cell-type bifurcation with robust ratios. *Nature Communications* **6:6195** (2015).

[164] Payne, S. *et al.* Temporal control of self-organized pattern formation without morphogen gradients in bacteria. *Molecular Systems Biology* **9** (2013).

[165] Cao, Y. *et al.* Collective space-sensing coordinates pattern scaling in engineered bacteria. *Cell* **165**, 620–630 (2016).

[166] Tabor, J. J. *et al.* A synthetic genetic edge detection program. *Cell* **137**, 1272–1281 (2009).

[167] Fernandez-Rodriguez, J., Moser, F., Song, M. & Voigt, C. A. Engineering RGB color vision into Escherichia coli. *Nature Chemical Biology* **13**, 706–708 (2017).

[168] Scalise, D. & Schulman, R. Designing modular reaction-diffusion programs for complex pattern formation. *Technology* **2**, 55–66 (2014).

[169] Wang, S. S. & Ellington, A. D. Pattern generation with nucleic acid chemical reaction networks. *Chemical Reviews* **119**, 6370–6383 (2019).

[170] Gomez, M. M. & Arcak, M. A Tug-of-War Mechanism for Pattern Formation in a Genetic Network. *ACS Synthetic Biology* **6**, 2056–2066 (2017).

[171] Dang, Y., Grundel, D. & Youk, H. Cellular dialogues that enable self-organization of dynamic spatial patterns. *bioRxiv* .

[172] Cardinale, S. & Arkin, A. P. Contextualizing context for synthetic biology–identifying causes of failure of synthetic biological systems. *Biotechnology Journal* **7**, 856–866 (2012).

[173] Brophy, J. A. N. & Voigt, C. A. Principles of genetic circuit design. *Nature Methods* **11**, 508–520 (2014).

[174] Pai, A., Tanouchi, Y., Collins, C. H. & You, L. Engineering multicellular systems by cell-cell communication. *Current Opinion in Biotechnology* **20**, 461–470 (2009).

[175] Hennig, S., Rödel, G. & Ostermann, K. Artificial cell-cell communication as an emerging tool in synthetic biology applications. *Journal of Biological Engineering* **9** (2015).

[176] Liu, C. *et al.* Sequential establishment of stripe patterns in an expanding cell population. *Science* **334**, 238–231 (2011).

[177] Frangipane, G. *et al.* Dynamic density shaping of photokinetic *E. coli*. *eLife* (2018).

[178] Xue, X., Xue, C. & Tang, M. The role of intracellular signaling in the stripe formatino in engineered *Escherichia coli* populations. *PLoS Computational Biology* **14**, e1006178 (2018).

[179] Cachat, E. *et al.* 2- and 3-dimensional synthetic large-scale *de novo* patterning by mammalian cells trhough phase separation. *Scientific Reports* **6** (2016).

[180] Toda, S., Blauch, L. R., Tang, S. K. Y., Morsut, L. & Lim, W. A. Programming self-organizing multicellular structures with synthetic cell-cell signaling. *Science* **361**, 156–162 (2018).

[181] Santos-Moreno, J. & Schaerli, Y. Using synthetic biology to engineer spatial patterns. *Advanced Biosystems* **3**, 1800280 (2019).

[182] Luo, N., Wang, S. & You, L. Synthetic pattern formation. *Biochemistry* **58**, 1478–1483 (2019).

[183] Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nature Reviews Microbiology* **12**, 381–390 (2014).

[184] Can biological phenomena be understood by humans? *Nature* **403**, 345 (2000).

[185] Noble, D. The rise of computational biology. *Nature Reviews Molecular Cell Biology* **3**, 460–463 (2002).

[186] Noble, D. The future: putting Humpty-Dumpty together again. *Biochemical Society Transactions* **31**, 156–158 (2003).

[187] Tomlin, C. J. & Axelrod, J. D. Biology by numbers: mathematical modelling in developmental biology. *Nature Reviews Genetics* **8**, 331–340 (2007).

[188] Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. From molecular to modular cell biology. *Nature* **402**, C47–C52 (1999).

[189] Alon, U. Biological networks: the tinkerer as an engineer. *Science* **301**, 1866–1867 (2003).

[190] Davies, J. A. Real-world synthetic biology: is it founded on an engineering approach, and should it be? *Life* **9** (2019).

[191] Lander, A. D. Pattern, growth, and control. *Cell* **144**, 955–969 (2011).

[192] Csete, M. E. & Doyle, J. C. Reverse engineering of biological complexity. *Science* **295**, 1664–1669 (2002).

[193] Del Vecchio, D. *et al.* What have the principles of engineering taught us about biological systems? *Cell Systems* **2**, 5–7 (2016).

[194] Stelling, J., Sauer, U., Szallasi, Z., Doyle, F. J. & Doyle, J. C. Robustness of cellular functions. *Cell* **118**, 675–685 (2004).

[195] Khammash, M., Tomlin, C. J. & Vidyasagar, M. Guest editorial - special issue on systems biology. *IEEE Transactions on Automatic Control* **53**, 4–7 (2008).

[196] Khammash, M. An engineering viewpoint on biological robustness. *BMC Biology* **14**, 22 (2016).

[197] Arkin, A. & Doyle, J. Appreciation of the machinations of the blind watchmaker. *IEEE Transactions on Automatic Control* **53**, 8–9 (2008).

[198] Wagner, A. & Rosen, W. Spaces of the possible: universal Darwinism and the wall between technological and biological innovation. *Journal of the Royal Society Interface* **11**, 20131190 (2014).

[199] Yan, K.-K., Wang, D., Xiong, K. & Gerstein, M. Comparing technological development and biological evolution from a network perspective. *Cell Systems* **10**, 219–222 (2020).

[200] Araujo, R. P. & Liotta, L. A. The topological requirements for robust perfect adaptation in networks of any size. *Nature Communications* **9** (2018).

[201] Baetica, A.-A., Westbrook, A. & El-Samad, H. Control theoretical concepts for synthetic and systems biology. *Current Opinion in Systems Biology* **14**, 50–57 (2019).

[202] Xiao, F. & Doyle, J. C. Robust perfect adaptation in biomolecular reaction networks. In *2018 IEEE Conference on Decision and Control (CDC)*, 4345–4352 (2018).

[203] Olsman, N. *et al.* Hard limits and performance tradeoffs in a class of antithetic integral feedback networks. *Cell Systems* **9**, 49–63.e16 (2019).

[204] Chevalier, M., Gómez-Schiavon, M., Ng, A. H. & El-Samad, H. Design and analysis of a proportional-integral-derivative controller with biological molecules. *Cell Systems* **9**, 338–353.e10 (2019).

[205] Berg, O. G., Paulsson, J. & Ehrenberg, M. Fluctuations and quality of control in biological cells: zero-order ultrasensitivity reinvestigated. *Biophysical Journal* **79**, 1228–1236 (2000).

[206] Lestas, I., Vinnicombe, G. & Paulsson, J. Fundamental limits on the suppression of molecular fluctuations. *Nature* **467**, 174–178 (2010).

[207] Briat, C., Gupta, A. & Khammash, M. Antithetic integral feedback esnures robust perfect adaptation in noisy biomolecular networks. *Cell Systems* **2**, 15–26 (2016).

[208] Jacob, F. Evolution and tinkering. *Science* **196**, 1161–1166 (1977).

[209] Dudley, R. & Gans, C. A critique of symmorphosis and optimality models in physiology. *Physiological Zoology* **64**, 627–637 (1991).

[210] Sontag, E. D. Some new directions in control theory inspired by systems biology. *Systems Biology* **1**, 9–18 (2004).

[211] Dennett, D. C. *From Bacteria to Bach and Back: The Evolution of Minds* (W. W. Norton Company, 2017).

[212] Wang, L.-Z., Wu, F., Flores, K., Lai, Y.-C. & Wang, X. Build to understand: synthetic approaches to biology. *Integrative Biology* **8**, 394–408 (2016).

[213] Way, M. "What I cannot create, I do not understand". *Journal of Cell Science* **130**, 2941–2942 (2017).

[214] Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genetics* **31**, 64–68 (2002).

[215] Tyson, J. J. & Novák, B. Functional motifs in biochemical reaction networks. *Annual Review of Physical Chemistry* **61**, 219–240 (2010).

[216] Ishihara, S., Fujimoto, K. & Shibata, T. Cross talking of network motifs in gene regulation that generates temporal pulses and spatial stripes. *Genes to Cells* **10**, 1025–1038 (2005).

[217] Mangan, S. & Alon, U. Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 11980–11985 (2003).

[218] von Dassow, G. & Odell, G. M. Design and constraints of the drosophila segment polarity module: robust spatial patterning emerges from intertwined cell state switches. *Journal of Experimental Zoology* **294**, 179–215 (2002).

[219] Panovska-Griffiths, J., Page, K. M. & Briscoe, J. A gene regulatory motif that generates oscillatory or multiway switch outputs. *Journal of the Royal Society Interface* **10**, 20120826 (2013).

[220] Perez-Carrasco, R. *et al.* Combining a toggle switch and a repressilator within the AC-DC circuit generates distinct dynamical behaviors. *Cell Systems* **6**, 521–530.e3 (2018).

[221] Alon, U. Network motifs: theory and experimental approaches. *Nature Reviews Genetics* **8**, 450–461 (2007).

[222] Ingram, P. J., Stumpf, M. P. H. & Stark, J. Network motifs: structure does not determine function. *BMC Genomics* **7**, 108 (2006).

[223] Blanchini, F. & Franco, E. Structural analysis of biological networks. In Kulkarni, V. V., Stan, G.-B. & Raman, K. (eds.) *A Systems Theoretic Approach to Systems and Synthetic Biology I: Models and System Characterizations*, 47–71 (Springer, 2014).

[224] Ingolia, N. T. Topology and robustness in the *Drosophila* segment polarity network. *PLoS Biology* **2**, e123 (2004).

[225] Goelzer, A. & Fromion, V. Towards the modular decomposition of the metabolic network. In Kulkarni, V. V., Stan, G.-B. & Raman, K. (eds.) *A Systems Theoretic Approach to Systems and Synthetic Biology I: Models and System Characterizations*, 121–152 (Springer, 2014).

[226] Briscoe, J. Understanding pattern formation in embryos: Experiment, theory, and simulation. *Journal of Computational Biology* **26**, 1–7 (2019).

[227] Del Vecchio, D., Ninfa, A. J. & Sontag, E. D. Modular cell biology: retroactivity and insulation. *Molecular Systems Biology* **4**, 161 (2008).

[228] Gyorgy, A. & Del Vecchio, D. Modular composition of gene transcription networks. *PLoS Computational Biology* **10**, e1003486 (2014).

[229] Bennett, S. *A History of Control Engineering 1800-1930*. No. 8 in IET Control Engineering Series (The Institution of Engineering and Technology, London, 2008). Reprint of 1979 edition from Peter Peregrinus Ltd.

[230] Bennett, S. *A history of control engineering, 1930-1955*. No. 47 in IEE Control Engineering Series (Peter Peregrinus Ltd. on behalf of the Institution of Electrical Engineers, London, 1993).

[231] Strogatz, S. H. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering* (Perseus Books Publishing, LLC, 1994).

[232] Willems, J. C. In control, almost from the beginning until the day after tomorrow. *European Journal of Control* **13**, 71–81 (2007).

[233] Barkai, N. & Leibler, S. Robustness in simple biochemical networks. *Nature* **387**, 913–917 (1997).

[234] Alon, U., Surette, M. G., Barkai, N. & Leibler, S. Robustness in bacterial chemotaxis. *Nature* **397**, 168–171 (1999).

[235] Yi, T.-M., Huang, Y., Simon, M. I. & Doyle, J. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 4649–4653 (2000).

[236] Friedlander, T. & Brenner, N. Adaptive response by state-dependent inactivation. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 2258–22563 (2009).

[237] Cosentino, C. & Bates, D. *Feedback control in systems biology* (CRC Press, 2011).

[238] Ferrell, J., James E. Perfect and near-perfect adaptation in cell signaling. *Cell Systems* **2**, 62–67 (2016).

[239] El-Samad, H., Kurata, H. T., Gross, C. A. & Khammash, M. Surviving heat shock: control strategies for robustness and performance. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2736–2741 (2005).

[240] Ghosh, A. K. & Chance, B. Oscillations of glycolytic intermediates in yeast cells. *Biochemical and Biophysical Research Communications* **16**, 174–181 (1964).

[241] Richard, P. The rhythm of yeast. *FEMS Microbiology Reviews* **27**, 547–557 (2003).

[242] Tornheim, K. & Lowenstein, J. M. The purine nucleotide cycle. control of phosphofructokinase and glycolytic oscillations in muscle extracts. *The Journal of Biological Chemistry* **250**, 6304–6314 (1975).

[243] Chandra, F. A., Buzi, G. & Doyle, J. C. Glycolytic oscillations and limits on robust efficiency. *Science* **333**, 187–192 (2011).

[244] Siami, M. *et al.* Fundamental limits and tradeoffs in autocatalytic pathways. *IEEE Transactions on Automatic Control* **65**, 733–740 (2020).

[245] Ni, X. Y., Drengstig, T. & Ruoff, P. The control of the controller: molecular mechanisms for robust perfect adaptation and temperature compensation. *Biophysical Journal* **97**, 1244–1253 (2009).

[246] Aoki, S. K. *et al.* A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature* **570**, 533–537 (2019).

[247] Cassidy, J. J. *et al.* Repressive gene regulation synchronizes development with cellular metabolism. *Cell* **178**, 1–13 (2019).

[248] Bhartiya, S., Chaudhary, N., Venkatesh, K. V. & Doyle, F. J. Multiple feedback loop design in the tryptophan regulatory network of *Escherichia coli* suggests a paradigm for robust regulation of processes in series. *Journal of the Royal Society Interface* **3**, 383–391 (2006).

[249] Levin, M. The biophysics of regenerative repair suggests new perspectives on biological causation. *BioEssays* **42**, 1900146 (2020).

[250] Pezzulo, G. & Levin, M. Top-down models in biology: explanation and control of complex living systems above the molecular level. *Journal of the Royal Society Interface* **13**, 20160555 (2016).

[251] Albert, R. & Othmer, H. G. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *Journal of Theoretical Biology* **223**, 1–18 (2003).

[252] Tkačik, G., Callan, J., Curtis G. & Bialek, W. Information flow and optimization in transcriptional regulation. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 12265–12270 (2008).

[253] Tkačik, G., Dubuis, J. O., Petkova, M. D. & Gregor, T. Positional information, positional error, and readout precision in morphogenesis: A mathematical framework. *Genetics* **199**, 39–59 (2015).

[254] Petkova, M. D., Tkačik, G., Bialek, W., Wieschaus, E. F. & Gregor, T. Optimal decoding of cellular identities in a genetic network. *Cell* **176**, 1–12 (2019).

[255] Jaeger, J. & Verd, B. Dynamic positional information: patterning mechanism versus precision in gradient-driven systems. In *Current Topics in Developmental Biology*, vol. 137, chap. 7, 219–246 (Elsevier, Inc., 2020).

[256] Szallasi, Z., Periwal, V. & Stelling, J. On modules and modularity. In Szallasi, Z., Stelling, J. & Periwal, V. (eds.) *System modeling in cell biology: from concepts to nuts and bolts*, chap. 3, 41–50 (The MIT Press, 2006).

[257] Konagurthu, A. S. & Lesk, A. M. On the origin of distribution patterns of motifs in biological networks. *BMC Systems Biology* **2**, 73 (2008).

[258] Davies, J. Using synthetic biology to explore principles of development. *Development* **144**, 1146–1158 (2017).

[259] Li, P. *et al.* Morphogen gradient reconstitution reveals Hedgehog pathway design principles. *Science* **360**, 543–548 (2018).

[260] Crocker, J., Tsai, A. & Stern, D. L. A fully synthetic transcriptional platform for a multicellular eukaryote. *Cell Reports* **18**, 287–296 (2017).

[261] Davies, J. A. Synthetic morphology: prospects for engineered, self-constructing anatomies. *Journal of Anatomy* **212**, 707–719 (2008).

[262] Davies, J. A. & Cachat, E. Synthetic biology meets tissue engineering. *Biochemical Society Transactions* **44**, 696–701 (2016).

[263] Teague, B. P., Guye, P. & Weiss, R. Synthetic morphogenesis. *Cold Spring Harbor Perspectives in Biology* (2016).

[264] Johnson, M. B., March, A. R. & Morsut, L. Engineering multicellular systems: Using synthetic biology to control tissue self-organization. *Current Opinion in Biomedical Engineering* **4**, 163–173 (2017).

[265] Kamm, R. D. *et al.* Perspective: the promise of multi-cellular engineered living systems. *APL Bioengineering* **2**, 040901 (2018).

[266] Ebrahimkhani, M. R. & Ebisuya, M. Synthetic developmental biology: build and control multicellular systems. *Current Opinion in Chemical Biology* **52**, 9–15 (2019).

[267] Santorelli, M., Lam, C. & Morsut, L. Synthetic development: building mammalian multicellular structures with artificial genetic programs. *Current Opinion in Biotechnology* **59**, 130–140 (2019).

[268] Toda, S., Brunger, J. M. & Lim, W. A. Synthetic development: learning to program multicelllar self-organization. *Current Opinion in Systems Biology* **14**, 14–49 (2019).

[269] Kicheva, A. & Rivron, N. C. Creating to understand–developmental biology meets engineering in Paris. *Development* **144**, 733–736 (2017).

[270] Levin, M. & Arias, A. M. Reverse-engineering growth and form in Heidelberg. *Development* **146**, dev177261 (2019).

[271] Trewavas, A. A brief history of systems biology. *The Plant Cell* **18**, 2420–2430 (2006).

[272] Ma, W., Trusina, A., El-Samad, H., Lim, W. A. & Tang, C. Defining network topologies that can achieve biochemical adaptation. *Cell* **138**, 760–773 (2009).

[273] Artyukhin, A. B., Wu, L. F. & Altschuler, S. J. Only two ways to achieve perfection. *Cell* **138**, 619–621 (2009).

[274] Ghosh, R. & Tomlin, C. J. Lateral inhibition through Delta-Notch signaling: a piecewise affine hybrid model. In Di Benedetto, M. & Sangiovanni-Vincentelli, A. (eds.) *Hybrid Systems: Coputation and Control*, vol. 2034 of *Lecture Notes in Computer Science*, 232–246 (Springer, 2001).

[275] Bizzarri, M. *et al.* A call for a better understanding of causation in cell biology. *Nature Reviews Molecular Cell Biology* **20**, 261–262 (2019).

[276] Goodwin, B. C. *Temporal organization in cells: a dynamic theory of cellular control processes* (Academic Press, 1963).

[277] Othmer, H. G. The qualitative dynamics of a class of biochemical control circuits. *Journal of Mathematical Biology* **3**, 53–78 (1976).

[278] Tyson, J. J. & Othmer, H. G. The dynamics of feedback control circuits in biochemical pathways. In Rosen, R. & Snell, F. M. (eds.) *Progress in Theoretical Biology*, vol. 5, 1–62 (Academic Press, 1978).

[279] Thron, C. D. The secant condition for instability in biochemical feedback control–I. The role of cooperativity and saturability. *Bulletin of Mathematical Biology* **53**, 403–424 (1991).

[280] Thron, C. D. The secant condition for instability in biochemical feedback control–II. Models with upper Hessenberg Jacobian matrices. *Bulletin of Mathematical Biology* **53**, 403–424 (1991).

[281] Sontag, E. D. Passivity gains and the "secant condition" for stability. *Systems & Control Letters* **55**, 177–183 (2006).

[282] Arcak, M. & Sontag, E. Diagonal stability of a class of cyclic systems and its connection with the secant criterion. *Automatica* **42**, 1531–1537 (2006).

[283] Arcak, M. & Sontag, E. A passivity-based stability criterion for a class of biochemical reaction networks. *Mathematical Biosciences and Engineering* **5**, 1–19 (2008).

[284] Blanchini, F., El-Samad, H., Giordano, G. & Sontag, E. D. Control-theoretic methods for biological networks. In *IEEE Conference on Decision and Control*, 466–483 (2018).

[285] Arcak, M. Certifying spatially uniform behavior in reaction-diffusion PDE and compartmental ODE systems. *Automatica* **47**, 1219–1229 (2011).

[286] Aminzare, Z., Shafi, Y., Arcak, M. & Sontag, E. D. Guaranteeing spatial uniformity in reaction-diffusion systems using weighted L^2 norm contractions. In Kulkarni, V. V., Stan, G.-B. & Raman, K. (eds.) *A Systems Theoretic Approach to Systems and Synthetic Biology I: Models and System Characterizations*, 73–101 (Springer, 2014).

[287] Hori, Y., Miyazako, H., Kumagai, S. & Hara, S. Coordinated spatial pattern formation in biomolecular communication networks. *IEEE Transactions on Molecular, Biological and Multi-scale Communications* **1**, 111–121 (2015).

[288] Aminzare, Z. & Sontag, E. D. Some remarks on spatial uniformity of solutions of reaction-diffusion PDEs. *Nonlinear Analysis: Theory, Methods & Applications* **147**, 125–144 (2016).

[289] Arcak, M. Pattern formation by lateral inhibition in large-scale networks of cells. *IEEE Transactions on Automatic Control* **58**, 1250–1262 (2013).

[290] Wang, W. & Slotine, J.-J. E. On partial contraction analysis for coupled nonlinear oscillators. *Biological Cybernetics* **92**, 38–53 (2005).

[291] Stan, G.-B. & Sepulchre, R. Analysis of interconnected oscillators by dissipativity theory. *IEEE Transactions on Automatic Control* **52**, 256–270 (2007).

[292] Scardovi, L., Arcak, M. & Sontag, E. D. Synchronization of interconnected systems with applications to biochemical networks: An input-output approach. *IEEE Transactions on Automatic Control* **55**, 1367–1379 (2010).

[293] Aminzare, Z. & Sontag, E. D. Synchronization of diffusively-connected nonlinear systems: results based on contractions with respect to general norms. *IEEE Transactions on Network Science and Engineering* **1**, 91–106 (2014).

[294] Angeli, D., Ferrell, J., James E. & Sontag, E. D. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 1822–1827 (2004).

[295] Lazebnik, Y. Can a biologist fix a radio?—or, what i learned while studying apoptosis. *Cancer Cell* **2**, 179–182 (2002).

[296] Elowitz, M. & Lim, W. A. Build life to understand it. *Nature* **468**, 889–890 (2010).

[297] Del Vecchio, D., Qian, Y., Murray, R. M. & Sontag, E. D. Future systems and control research in synthetic biology. *Annual Reviews in Control* **45**, 5–17 (2018).

[298] Ingalls, B. P. A frequency domain approach to sensitivity analysis of biochemical networks. *The Journal of Physical Chemistry B* **108**, 1143–1152 (2004).

[299] Simpson, M. L., Cox, C. D. & Sayler, G. S. Frequency domain chemical Langevin analysis of stochasticity in gene transcriptional regulation. *Journal of Theoretical Biology* **229**, 383–394 (2004).

[300] Zinner, B. Existence of traveling wavefront solutions for the discrete nagumo equation. *Journal of Differential Equations* **96**, 1–27 (1992).

[301] Keener, J. P. & Sneyd, J. *Mathematical Physiology*, vol. 8 of *Interdisciplinary Applied Mathematics* (Springer, New York, NY, 2009), 2nd edn.

[302] Cahn, J. W., Mallet-Paret, J. & Van Vleck, E. S. Traveling wave solutions for systems of ODEs on a two-dimensional spatial lattice. *SIAM Journal on Applied Mathematics* **59**, 455–493 (1998).

[303] Hill, A. V. The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *Proceedings of the Physiological Society* (1910).

[304] Weiss, J. N. The Hill equation revisited: uses and misuses. *The Federation of American Societies for Experimental Biology Journal* **11** (1997).

[305] Perkins, M. L. & Arcak, M. Discrete spatial filtering by networks of cells facilitates biological pattern formation. In *Proceedings of the American Controls Conference* (IEEE, 2018).

[306] Perkins, M. L. & Arcak, M. A spatial filtering approach to biological patterning. *SIAM Journal on Applied Dynamical Systems* **18**, 1694–1721 (2019).

[307] Tei, M., Perkins, M. L., Hsia, J., Arcak, M. & Arkin, A. P. Designing spatially distributed gene regulatory networks to elicit contrasting patterns. *ACS Synthetic Biology* **8**, 119–126 (2019).

[308] Perkins, M. L., Benzinger, D., Arcak, M. & Khammash, M. Cell-in-the-loop pattern formation with optogenetically emulated cell-to-cell signaling. *Nature Communications* **11**, 1355 (2020).

[309] Bamieh, B., Paganini, F. & Dahleh, M. A. Distributed control of spatially invariant systems. *IEEE Transactions on Automatic Control* **47**, 1091–1107 (2002).

[310] Othmer, H. G. & Scriven, L. E. Instability and dynamic pattern in cellular networks. *Journal of Theoretical Biology* **32**, 507–537 (1971).

[311] Manu *et al.* Canalization of gene expression in the *Drosophila* blastoderm by gap gene cross regulation. *PLoS Biology* **7** (2009).

[312] Dillon, R. & Othmer, H. G. A mathematical model for outgrowth and spatial patterning of the vertebrate limb bud. *Journal of Theoretical Biology* **197**, 295–330 (1999).

[313] Cross, M. & Greenside, H. *Pattern Formation and Dynamics in Nonequilibrium Systems* (Cambridge University Press, Cambridge, UK, 2009).

[314] Jovanović, M. R. & Bamieh, B. Componentwise energy amplification in channel flows. *Journal of Fluid Mechanics* **534**, 145–183 (2005).

[315] Hori, Y. & Hara, S. Noise-induced spatial pattern formation in stochastic reaction-diffusion systems. In *51st IEEE Conference on Decision and Control (CDC)*, 1053–1058 (2012).

[316] van Kampen, N. G. *Stochastic processes in physics and chemistry* (Elsevier, Oxford, UK, 2007), 3 edn.

[317] Strang, G. The discrete cosine transform. *SIAM Review* **41**, 135–147 (1999).

[318] Classen, A.-K., Anderson, K. I., Marois, E. & Eaton, S. Hexagonal packing of *Drosophila* wing epithelial cells by the planar cell polarity pathway. *Developmental Cell* **9**, 805–817 (2005).

[319] Jaeger, J. The gap gene network. *Cellular and Molecular Life Sciences* **68**, 243–274 (2011).

[320] Henrique, D. & Schweisguth, F. Mechanisms of Notch signaling: a simple logic deployed in time and space. *Development* **146** (2019).

[321] Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch signaling: cell fate control and signal integration in development. *Science* **284**, 770–776 (1999).

[322] LeBon, L., Lee, T. V., Sprinzak, D., Jafar-Nejad, H. & Elowitz, M. B. Fringe proteins modulate Notch-ligand *cis* and *trans* interactions to specify signaling states. *eLife* **3** (2014).

[323] Langridge, P. D. & Struhl, G. Epsin-dependent ligand endocytosis activates Notch by force. *Cell* **171**, 1383–1396.e12 (2017).

[324] Nandagopal, N., Santat, L. A. & Elowitz, M. B. *Cis*-activation in the Notch signaling pathway. *eLife* **8** (2019).

[325] Tickle, C. Making digit patterns in the vertebrate limb. *Nature Reviews: Molecular Cell Biology* **7**, 45–53 (2006).

[326] Zeller, R., López-Ríos, J. & Zuniga, A. Vertebrate limb bud development: moving towards integrative analysis of organogenesis. *Nature Reviews Genetics* **10**, 845–858 (2009).

[327] Raspopovic, J., Marcon, L., Russo, L. & Sharpe, J. Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. *Science* **345**, 566–570 (2014).

[328] Onimaru, K., Marcon, U., Musy, M., Tanaka, M. & Sharpe, J. The fin-to-limb transition as the re-organization of a Turing pattern. *Nature Communications* **7** (2016).

[329] Milo, R., Jorgensen, P., Moran, U., Weber, G. & Springer, M. Bionumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Research* **38**, D750–D753 (2010).

[330] Brenner, K., You, L. & Arnold, F. H. Engineering microbial consortia: a new frontier in synthetic biology. *Trends in Biotechnology* **26**, 483–489 (2008).

[331] Levskaya, A. *et al.* Engineering *Escherichia coli* to see light. *Nature* **438**, 441–441 (2005).

[332] Ferreira, A. S. R., Hsia, J. & Arcak, M. A compartmental lateral inhibition system to generate contrasting patterns. *IEEE Life Sciences Letters* **1**, 7–10 (2015).

[333] Hsu, T. M. *et al.* Employing a biochemical protecting group for a sustainable indigo dyeing strategy. *Nature Chemical Biology* **14**, 256–261 (2018).

[334] Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C. & Voigt, C. A. Genetic programs constructed from layered logic gates in single cells. *Nature* **491**, 249–253 (2012).

[335] Nielsen, A. A. K. *et al.* Genetic circuit design automation. *Science* **352**, 53 (2016).

[336] Lee, Y. J., Kim, S.-J. & Moon, T. S. Multilevel regulation of bacterial gene expression with the combined STAR and antisense RNA system. *ACS Synthetic Biology* **7**, 853–865 (2018).

[337] Bokinsky, G. *et al.* Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 19949–54 (2011).

[338] Soma, Y., Tsuruno, K., Wada, M., Yokota, A. & Hanai, T. Metabolic flux redirection from a central metabolic pathway toward a synthetic pathway using a metabolic toggle switch. *Metabolic Engineering* **23**, 175–184 (2014).

[339] Venturelli, O. S. *et al.* Programming mRNA decay to modulate synthetic circuit resource allocation. *Nature Communications* **8**, 15128 (2017).

[340] Anderson, J. C., Clarke, E. J., Arkin, A. P. & Voigt, C. A. Environmentally controlled invasion of cancer cells by engineered bacteria. *Journal of Molecular Biology* **355**, 619–627 (2006).

[341] Anderson, J. C., Voigt, C. A. & Arkin, A. P. Environmental signal integration by a modular and gate. *Molecular Systems Biology* **3**, 133 (2007).

[342] Klavins, E. Lightening the load in synthetic biology. *Nature Biotechnology* **32**, 1198–1200 (2014).

[343] Tsoi, R. *et al.* Metabolic division of labor in microbial systems. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 201716888 (2018).

[344] Tamsir, A., Tabor, J. J. & Voigt, C. A. Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. *Nature* **469**, 212–215 (2011).

[345] Prindle, A. *et al.* A sensing array of radically coupled genetic 'biopixels'. *Nature* **481**, 39–44 (2012).

[346] Cao, Y. *et al.* Programmable assembly of pressure sensors using pattern-forming bacteria. *Nature Biotechnology* **35**, 1087–1093 (2017).

[347] Din, M. O. *et al.* Synchronized cycles of bacterial lysis for *in vivo* delivery. *Nature* **536**, 81–85 (2016).

[348] Kong, W., Meldgin, D. R., Collins, J. J. & Lu, T. Designing microbial consortia with defined social interactions. *Nature Chemical Biology* 1 (2018).

[349] Mobarry, B. K., Wagner, M., Urbain, V., Rittmann, B. E. & Stahl, D. A. Phylogenetic probes for analyzing abundance and spatial organization of nitrifying bacteria. *Applied Environmental Microbiology* **62**, 2156–2162 (1996).

[350] Fröstl, J. M. & Overmann, J. Physiology and tactic response of the phototrophic consortium "chlorochromatium aggregatum". *Archives of Microbiology* **169**, 129–135 (1998).

[351] Boetius, A. *et al.* A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* **407**, 623–626 (2000).

[352] Kim, H. J., Boedicker, J. Q., Choi, J. W. & Ismagilov, R. F. Defined spatial structure stabilizes a synthetic multispecies bacterial community. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 18188–18193 (2008).

[353] Mousa, W. K. *et al.* Root-hair endophyte stacking in finger millet creates a physicochemical barrier to trap the fungal pathogen fusarium graminearum. *Nature Microbiology* **1**, 1–12 (2016).

[354] Volke, D. C. & Nikel, P. I. Getting bacteria in shape: Synthetic morphology approaches for the design of efficient microbial cell factories. *Advanced Biosystems* **2** (2018).

[355] Morse, R. P. *et al.* Structural basis of toxicity and immunity in contact-dependent growth inhibition (cdi) systems. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 21480–21485 (2012).

[356] Garcia, E. C., Perault, A. I., Marlatt, S. A. & Cotter, P. A. Interbacterial signaling via Burkholderia contact-dependent growth inhibition system proteins. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 8296–8301 (2016).

[357] Angeli, D. & Sontag, E. D. Multi-stability in monotone input/output systems. *Systems and Control Letters* **51**, 185–202 (2004).

[358] Hooshangi, S., Thiberge, S. & Weiss, R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3581–3586 (2005).

[359] Grant, P. K. *et al.* Orthogonal intercellular signaling for programmed spatial behavior. *Molecular Systems Biology* **12**, 849–849 (2016).

[360] Stanton, B. C. *et al.* Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nature Chemical Biology* **10**, 99–105 (2014).

[361] Flynn, J. M. *et al.* Overlapping recognition determinants within the ssra degradation tag allow modulation of proteolysis. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10584–10589 (2001).

[362] Kaufmann, G. F. *et al.* Revisiting quorum sensing: Discovery of additional chemical and biological functions for 3-oxo-N-acylhomoserine lactones. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 309–314 (2005).

[363] Thompson, L. R. *et al.* A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457–463 (2017).

[364] Amonlirdviman, K. *et al.* Mathematical modeling of planar cell polarity to understand domineering nonautonomy. *Science* **307**, 423–426 (2005).

[365] Samoilov, M. S., Price, G. & Arkin, A. P. From fluctuations to phenotypes: the physiology of noise. *Science's Signal Transduction Knowledge Environment* **2006**, 1–10 (2006).

[366] Guisoni, N., Martinez-Corral, R., Garcia-Ojalvo, J. & de Navascués, J. Diversity of fate outcomes in cell pairs under lateral inhibition. *Development* **144**, 1177–1186 (2017).

[367] Chen, D. *Designing Genetic Circuits for Memory and Communication*. Ph.D. thesis, University of California, Berkeley (2014).

[368] Cardinale, S., Joachimiak, M. P. & Arkin, A. P. Effects of genetic variation on the *E. coli* host-circuit interface. *Cell Reports* **4**, 231–237 (2013).

[369] Morsut, L. *et al.* Engineering customized cell sensing and response behaviors using synthetic Notch receptors. *Cell* **164**, 780–791 (2016).

[370] Ang, J., Harris, E., Hussey, B. J., Kil, R. & McMillen, D. R. Tuning response curves for synthetic biology. *ACS Synthetic Biology* **2**, 547–567 (2013).

[371] Folds, D. J. Human in the loop simulation. In Loper, M. L. (ed.) *Modeling and Simulation in the Systems Engineering Life Cycle: Core Concepts and Accompanying Lectures*, Simulation Foundations, Methods, and Applications, chap. 15, 175–183 (Springer-Verlag, London, 2015).

[372] Toettcher, J. E., Gong, D., Lim, W. A. & Weiner, O. D. Light-based feedback for controlling intracellular signaling dynamics. *Nature Methods* **8**, 837–839 (2011).

[373] Milias-Argeitis, A. *et al. In silico* feedback for *in vivo* regulation of a gene expression circuit. *Nature Biotechnology* **29**, 1114–1116 (2011).

[374] Lugagne, J.-B. *et al.* Balancing a genetic toggle switch by real-time feedback control and periodic forcing. *Nature Communications* **8** (2017).

[375] Rullan, M., Benzinger, D., Schmidt, G. W., Milias-Argeitis, A. & Khammash, M. An optogenetic platform for real-time, single-cell interrogation of stochastic transcriptional regulation. *Molecular Cell* **70**, 745–756 (2018).

[376] Chait, R., Ruess, J., Bergmiller, T., Tkačik, G. & Guet, C. C. Shaping bacterial population behavior through computer-interfaced control of individual cells. *Nature Communications* **8:1535** (2017).

[377] Benzinger, D. & Khammash, M. Pulsatile inputs achieve tunable attenuation of gene expression variability and graded multi-gene regulation. *Nature Communications* **9** (2018).

[378] Aymoz, D., Wosika, V., Durandau, E. & Pelet, S. Real-time quantification of protein expression at the single-cell level via dynamic protein synthesis translocation reporters. *Nature Communications* **7** (2016).

[379] Motta-Mena, L. B. *et al.* An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nature Chemical Biology* 196–202 (2014).

[380] Thompson, K. E., Bashor, C. J., Lim, W. A. & Keating, A. E. SYNZIP protein interaction toolbox: *in vitro* and *in vivo* specifications of heterospecific coiled-coil interaction domains. *ACS Synthetic Biology* **1**, 118–120 (2012).

[381] Zopf, C. J., Quinn, K., Zeidman, J. & Maheshri, N. Cell-cycle dependence of trnascription dominates noise in gene expression. *PLoS Computational Biology* (2013).

[382] Borkowski, O., Ceroni, F., Stan, G.-B. & Ellis, T. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Current Opinion in Microbiology* **33**, 123–130 (2016).

[383] Wu, G. *et al.* Metabolic burden: Cornerstones in synthetic biology and metabolic engineering applications. *Trends in Biotechnology* **34**, 652–664 (2016).

[384] Young, J. Z. *Doubt and certainty in science: a biologist's reflections on the brain* (Oxford University Press, 1950). The B.B.C. Reith lectures.

[385] Darwin, C. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life* (John Murray, 1859).

[386] Lee, T. *et al.* BglBrick vectors and datasheets: A synthetic biology platform for gene expression. *Journal of Biological Engineering* **5**, 12 (2011).

[387] Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology* **27**, 946–950 (2009).

[388] Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nature Methods* **10**, 354–360 (2013).

[389] Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* (2008).

[390] Gietz, R. D. & Woods, R. A. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in Enzymology* **350**, 87–96 (2002).

[391] Lang, M., Rudolf, F. & Stelling, J. Use of YouScope to implement systematic microscopy protocols. *Current Protocols in Molecular Biology* **98**, 14.21.1–14.21.23 (2012).

[392] Dimopoulos, S., Mayer, C. E., Rudolf, F. & Stelling, J. Accurate cell segmentation in microscopy images using membrane patterns. *Bioinformatics* **30**, 2644–2651 (2014).

[393] Ricicova, M. *et al.* Dissecting genealogy and cell cycle as sources of cell-to-cell variability in MAPK signaling using high-throughput lineage tracking. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11403–11408 (2013).

Appendix A

Diffusion-mediated Lateral Inhibition: Materials and Methods

This appendix describes the detailed materials and methods for the experiments in Chapter 3, based on material published in [307].

Bacterial strains, plasmid construction, and growth conditions

Escherischia coli strain DH10B (NEB) was used for cloning. PCR amplifications were performed using Phusion High-Fidelity DNA Polymerase (Thermo) and oligonucleotides (IDT). BsaI (NEB) and T7 DNA ligase (NEB) were used to construct plasmids using parts obtained from the MIT Registry of Standard Biological Parts, JBEI registry [386], or synthesized gBlocks (IDT). RBS calculator [387] was used to generate balanced RBS strengths for $luxI$ and $lasI$. TR117 (gift of Thomas L. Ruegg) is a DH10B variant with genomically integrated mRFP1 driven by a constitutive promoter. MOPS EZ Rich Medium (Teknova) and MOPS with 1.5 % UltraPure agarose (Thermo) were used for liquid and solid medium. When appropriate, 50 $\mu g/mL$ Kanamycin or 20 $\mu g/mL$ Chloramphenicol were added to medium.

Plate reader assays

Overnight cultures of cells in MOPS were washed three times and diluted to fresh MOPS at OD600 of 0.3. After 8 hours in 30°C at 750 rpm, cells were washed three times and diluted to OD600 of 0.025 in a 96 well flat clear bottom black polystyrene microplate (Corning) containing 196 μL MOPS and appropriate concentrations of AHLs (Sigma) dissolved in 4 μL of dimethyl sulfoxide (DMSO) for Fig. A.1, or 192 μL MOPS and appropriate concentrations of AHLs in 4 μL DMSO and anhydrotetracycline (aTc) (Sigma) in 4 μL ethanol for Fig. A.2. Synergy 2 (Biotek Instruments) was used to measure cell density (OD600) and fluorescence of growing culture every 8 min for 12 hours at room temperature. The BioTek excitation

and emission wavelengths were 485 nm, 528 ± 20 nm for sfGFP and 560 nm, 620 ± 20 nm for mRFP1.

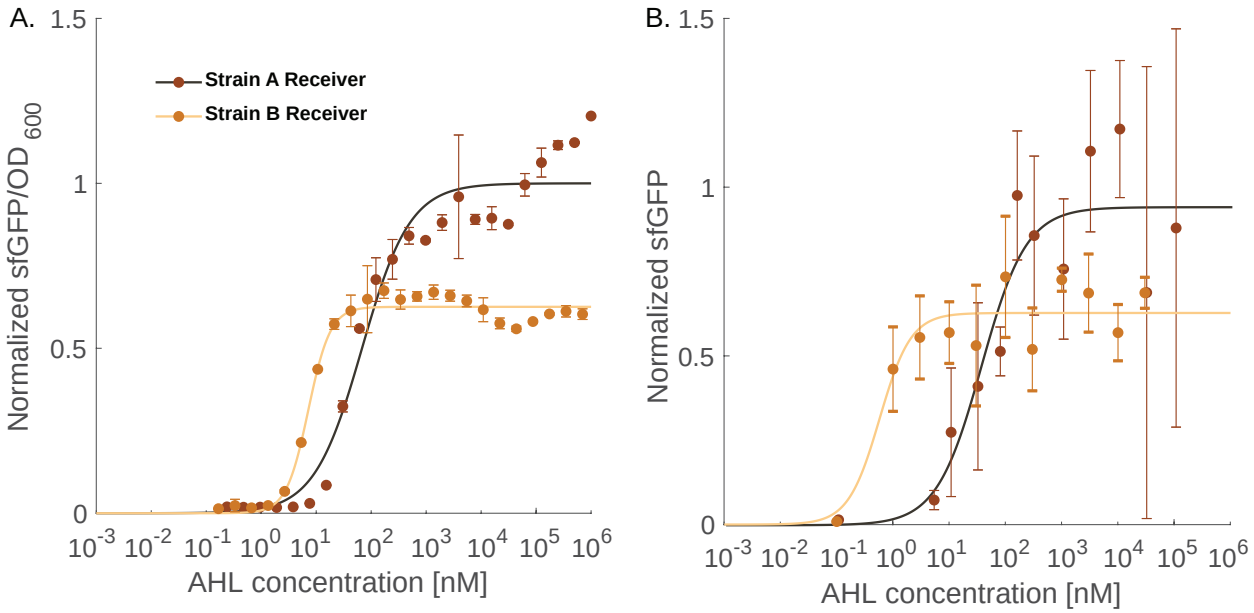

Figure A.1: Plate reader assays of output steady-state TetR-sfGFP fluorescence in response to input AHL concentration for the reception modules of Strains A and B in liquid (A) and solid (B) medium. The strains used for reception module characterization lack $ptet \rightarrow luxI$ or $lasI$. Steady-state sfGFP fluorescence (t = 10 h) divided by OD600 in liquid medium and steady-state sfGFP fluorescence ($t = 10\ h$) at the center of colonies on solid medium were each normalized to the maximum steady-state values across strains and input AHL conditions. Strain A receiver showed similar threshold AHL concentrations regardless of medium and Strain B receiver showed more AHL-sensitive response in solid than in liquid medium. In both conditions, the approximated K_d of Strain A reception module was higher than Strain B, suggesting higher sensitivity of *plas* in Strain B than *plux* in Strain A. Error bars show standard deviation and solid circles show the average of the measurement (n = 2). The solid curves show best-fit models when measurements were fit to activation Hill functions in the form of $a\frac{[AHL]}{([AHL]+K_d^n}$, where a denotes the maximal production of TetR-sfGFP, K_d denotes the apparent dissociation constant of AHL binding to the promoter and n denotes the apparent Hill coefficient. The following equation parameters were used for the best-fit models: $a = 1$ (Strain A) and 0.771 (Strain B), $K_d = 47.5\ nM$ (A) and 8.33 nM (B), $n = 1.91$ (A) and 1.83 (B) for liquid medium, $a =$ 0.9423 (A) and 0.6269 (B), $K_d = 34.71\ nM$ (A) and 0.577 $nM4(B)$,n=4 1.088 (A) and 1.556 (B) for solid medium.

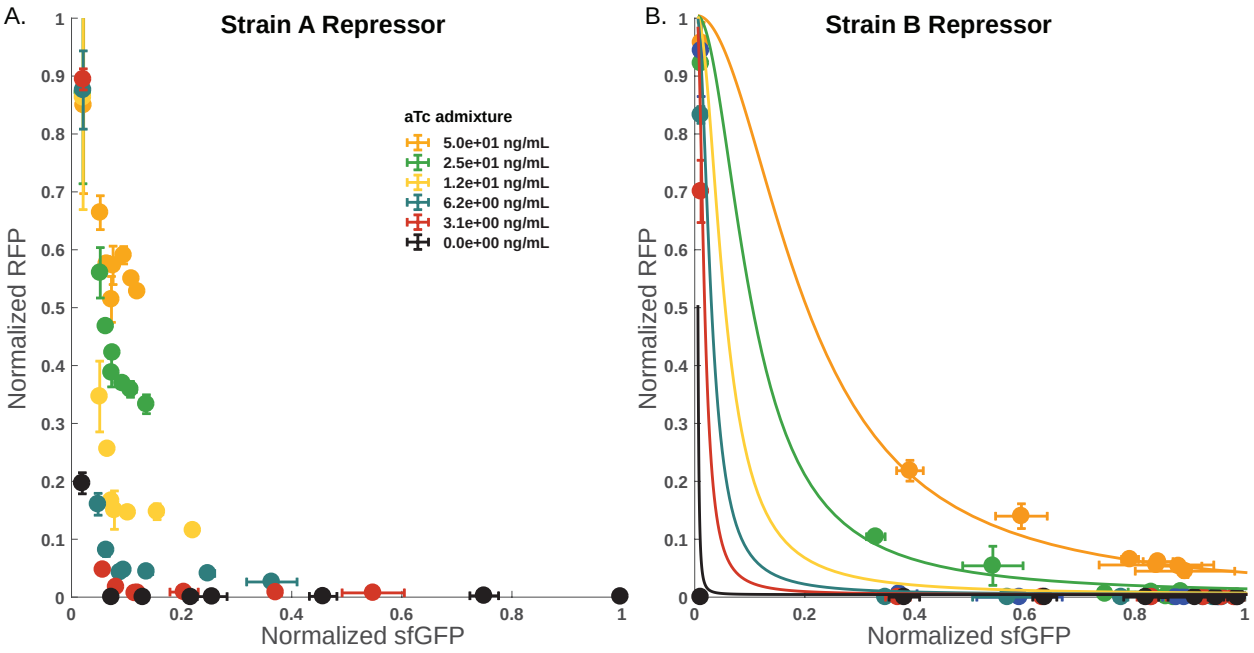

Figure A.2: Plate reader assays of output $ptet \rightarrow mRFP1$ fluorescence in response to input TetR-sfGFP fluorescence for the repression modules of Strains A (A) and B (B). The strains used for repression module characterization contain *ptet* driving mRFP1 instead of AHL synthases. Steady-state sfGFP fluorescence and steady-state mRFP1 fluorescence ($t = 10\ h$) were divided by OD600 and then each fluorescence was normalized to the maximum steady-state values across all conditions within the same strain. Different levels of input TetR-sfGFP were induced by AHL to repress mRFP1. As the leaky expression of $tetR - sfGFP$ caused significant repression of mRFP1 even in the absence of any AHL, a range of aTc was added to the medium to sequester basal level of TetR. Re-measurement of mRFP1 fluorescence showed a broad output range when at least 3.1 ng/mL aTc was added to medium. Unexpectedly, aTc also repressed TetR-sfGFP production in Strain A at high concentrations, which constrained viable aTc concentration to be less than 10 ng/mL. Solid circles show the average and error bars show standard deviation of the measurements (n = 2). The solid curve represents the best-fit model when the measurements were fit to repressive Hill function in the form of $\frac{1}{1+x_{max}\frac{[TetR]/K_t}{1+[aTc]/K_{aTc}}.^n}$, where $K_{aTc} = 0.098\ nM$, $K_{aTc} = 0.27\ nM$ as determined in Table 3.2. The resulting fit had a range of Hill coefficient between 1.7 to 2.8 and x_{max} ranged from 1100 nM to 1621 nM across different aTc concentrations.

Flow cytometry

Overnight cultures of Strains A and B in MOPS were washed three times and diluted to OD600 of 0.03 in fresh MOPS added with 5 ng/mL aTc and 0 or 1 μM AHL for pre-induction. After 6 hours of shaking at 750 rpm in 30°C, cells were washed three times and diluted to OD600 of 0.025 in a 96 well deep well plate (Green BioResearch) containing 196 μL MOPS, 5 ng/mL aTc, and appropriate concentrations of AHLs (Sigma). After 8 hours of shaking at 750 rpm in 30°C, cells were analyzed using BD LSRFortessa (BD Biosciences). Blue (488 nm) and green (561 nm) lasers were used in combination with 530/30 nm and 610/20 nm filters.

Construction of DLI device

The compartments and channels in the patterns were cut into 1/8 inch acrylic sheet (McMaster Carr) using a laser cutter (Universal Laser Systems) and then filled with SYLGARD 182 Silicone Elastomer (Dow Corning). PDMS molds were attached to the bottom of 6 well clear flat bottom cell culture plate (Falcon), and 3.4 mL of MOPS solid medium was poured into each mold to create DLI devices. aTc in ethanol was added to the final concentration of 5 ng/mL (Fig. A.3).

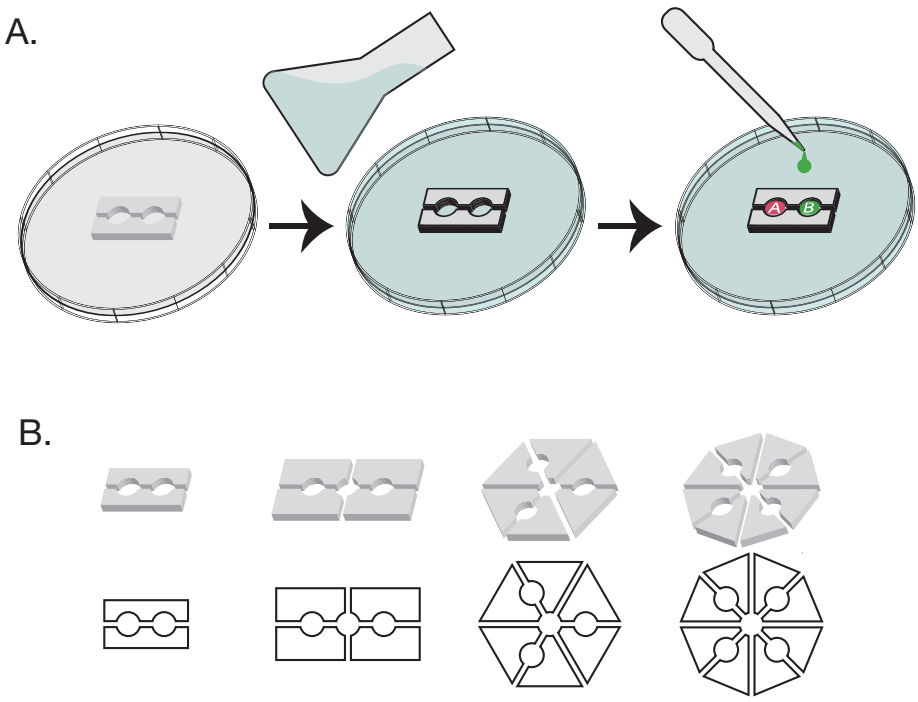

Figure A.3: Schematic of DLI device preparation. (A) PDMS molds are attached to the bottom of a culture plate and 3.4 mL of 1.5 % agarose-mixed medium is poured into the devices. After solidifying, 0.5 μL of 2.0 OD600 cells are pipetted at the center of each compartment and allowed to grow for 10-12 hours in room temperature to study pattern formation. (B) Various spatial configurations are prepared as PDMS molds. A central compartment with different number of neighbors can be manufactured.

DLI assays

Overnight cultures of Strains A and B in MOPS were washed three times and diluted to OD600 of 0.03 in fresh MOPS added with 5 ng/mL aTc and 0 or 1 μM AHL for pre-induction. After 8 hours of shaking at 750 rpm in 30°C, cells were washed three times and rediluted to OD600 of 2.0 in fresh MOPS. 0.5 μL of the culture was seeded onto each compartment of the DLI devices. Gel Doc XR+ System (Bio-rad) was used to image bacterial colonies every 30 min for 12 hours in room temperature. The blue epi illumination at 488 nm and 530/28 nm filter was used for sfGFP and the green epi illumination at 532 nm and 605/50 nm filters were used for mRFP1. Camera exposure time of 100 ms was used for all images.

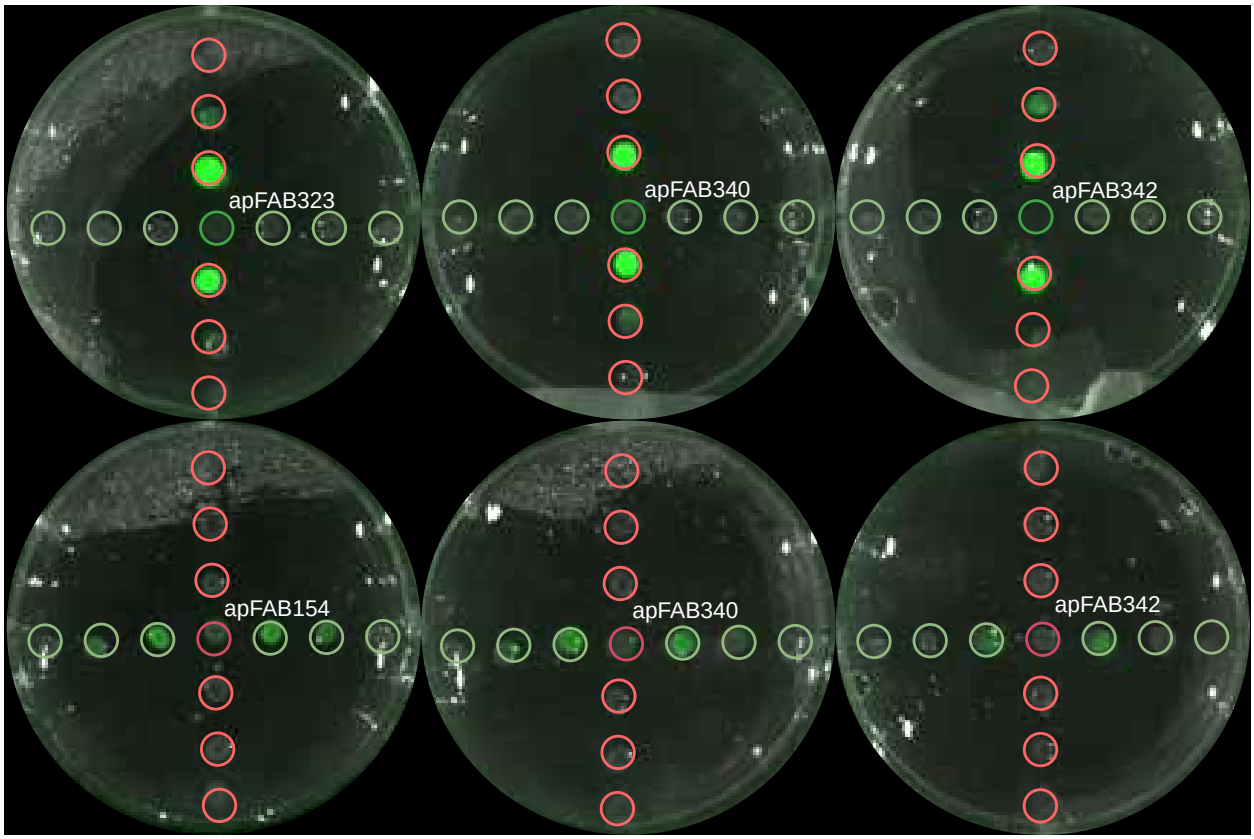

Figure A.4: Plate fluorimeter assay of AHL production and reception by Strains A and B. On a well of 3.4 mL of MOPS EZ Rich solid medium with 5 ng/mL aTc, the complete strains with different *ptet* promoter strength [388] were seeded at the center, surrounded by receiver strains (Strain A receiver in horizontal direction and Strain B receiver in vertical direction). The top row is seeded with Strain A variants at the center and the bottom row is seeded with Strain B variants at the center. The promoter strength of *ptet* seems to have little effect on AHL production, as receiver strains showed similar levels of activation regardless of promoter variants within the same strain type.

Computational modeling and simulation

We used custom code for computational modeling and data analysis in Matlab (Mathworks).

Channel length optimization

We optimized the channel length l to mimic close-range lateral inhibition with the diffusion-based system using a PDE model of AHL production, degradation, and diffusion. The goal was to allow sufficient diffusion of AHL to the immediate neighbors ([AHL] $\geq K_d$ at $\Delta x = l$) while restricting diffusion among nonadjacent compartments ([AHL] $< K_d$ at $\Delta x \geq 2l$).

Let $\theta(t)$ be AHL concentration over time at the center of an AHL-producing colony ($x = 0$). During exponential growth ($t \leq 10h$), the concentration of AHL can be expressed using parameters defined in Table 3.2 as

$$\dot{\theta}(t) = \nu I^* P_0 \exp(\mu t) - \left(\gamma + \frac{D}{l_{res}^2}\right)\theta(t) \tag{A.1}$$

$$\xrightarrow{\mathcal{L}} s\Theta(s) = \nu I^* P_0 \frac{1}{s-\mu} - \left(\gamma + \frac{D}{l_{res}^2}\right)\Theta(s) \tag{A.2}$$

$$\Leftrightarrow \Theta(s) = \nu I^* P_0 \frac{1}{(s-\mu)(s+\gamma+\frac{D}{l_{res}^2})} \tag{A.3}$$

$$\xrightarrow{\mathcal{L}^{-1}} \theta(t) = \nu I^* P_0 \frac{1}{\mu+\gamma+\frac{D}{l_{res}^2}}\left[\exp(\mu t) - \exp\left(-\left(\gamma + \frac{D}{l_{res}^2}\right)t\right)\right] \tag{A.4}$$

where P_0 is the initial population size of the colony and l_{res} is the length of channels that connect compartments to reservoir.

Now, consider diffusion of AHL across a channel. Define $[\text{AHL}](x,t) = \theta(x,t)$ in a one-dimensional, infinite-length channel, with the following boundary conditions:

$$\begin{cases} \theta(x,0) &= 0, \ \forall x, \\ \theta(0,t) &= \nu I^* P_0 \frac{1}{\mu+\gamma+\frac{D}{l_{res}^2}}\left[\exp(\mu t) - \exp\left(-\left(\gamma+\frac{D}{l_{res}^2}\right)t\right)\right], \ 0 \leq t \leq 10h, \\ \theta(\infty,t) &= 0, \ \forall t. \end{cases} \tag{A.5}$$

Assume the degradation of AHL in the channel is negligible. Then, from Fick's second law,

$$\frac{\partial\theta}{\partial t} = D\frac{\partial^2\theta}{\partial x^2} \tag{A.6}$$

$$\xrightarrow{\mathcal{L}} s\Theta - 0 = D\frac{d^2\Theta}{dx^2} \tag{A.7}$$

$$\Leftrightarrow \Theta(x,s) = A\exp\left(\sqrt{\frac{s}{D}}x\right) + B\exp\left(-\sqrt{\frac{s}{D}}x\right). \tag{A.8}$$

Using the boundary conditions A.5,

$$\Theta(x,s) = \nu I^* P_0 \frac{1}{(s-\mu)(s+\gamma+\frac{D}{l_{res}^2})} \exp\left(-\sqrt{\frac{s}{D}}x\right) \tag{A.9}$$

$$= \nu I^* P_0 \frac{1}{\mu+\gamma+\frac{D}{l_{res}^2}} \left(\frac{\exp(-\sqrt{\frac{s}{D}}x)}{s-\mu} - \frac{\exp(-\sqrt{\frac{s}{D}}x)}{s+\gamma+\frac{D}{l_{res}^2}}\right). \tag{A.10}$$

Since there is no simple solution to inversely transform A.10, for the optimization we instead replace the boundary conditions with an upper bound for AHL concentration.

Let $t \leq \tau_f$,

$$\theta(0,t) \leq \theta(0,\tau_f) = \nu I^* P_0 \frac{1}{\mu+\gamma+\frac{D}{l_{res}^2}} \left[\exp(\mu\tau_f) - \exp(-(\gamma+\frac{D}{l_{res}^2})\tau_f)\right] \tag{A.11}$$

$$\xrightarrow{\mathcal{L}} \Theta(0) \leq \frac{\nu I^* P_0 \frac{1}{\mu+\gamma+\frac{D}{l_{res}^2}} \left[\exp(\mu\tau_f) - \exp(-(\gamma+\frac{D}{l_{res}^2})\tau_f)\right]}{s}. \tag{A.12}$$

Then

$$\xrightarrow{\mathcal{L}^{-1}} \theta(x,t) \leq \nu I^* P_0 \frac{1}{\mu+\gamma+\frac{D}{l_{res}^2}} \left[\exp(\mu\tau_f) - \exp(-\left(\gamma+\frac{D}{l_{res}^2}\right)\tau_f)\right] erfc\left(\frac{x}{2\sqrt{Dt}}\right). \tag{A.13}$$

Using the upper-bound concentration, we determine the optimum channel length for short communication time $T_{1/2}$ (defined as the time for a compartment to reach half the concentration of the center) between the adjacent colonies and long the communication time between the non-adjacent compartments. $T_{1/2}$ at location x can be calculated as

$$\frac{x}{2\sqrt{DT_{1/2}}} = erfc^{-1}(0.5),$$

thus,

$$T_{1/2} = \frac{1}{4D}\left(\frac{x}{erfc^{-1}(0.5)}\right)^2. \tag{A.14}$$

We set the objective function and the constraints as follows:

$$\min_L T_{1/2}(x=l) + (\tau_f - T_{1/2}(x=2l)), \text{ such that} \tag{A.15}$$

$$\theta_{C_6}(L,\tau_f) \geq K_d^{C_6\to A}, \tag{A.16}$$

$$\theta_{C_6}(L,\tau_f) < K_d^{C_6\to B}, \tag{A.17}$$

$$\theta_{C_6}(2L,\tau_f) < K_d^{C_6\to A}, \tag{A.18}$$

$$\theta_{C_{12}}(L,\tau_f) \geq K_d^{C_{12}\to B}, \tag{A.19}$$

$$\theta_{C_{12}}(L,\tau_f) < K_d^{C_{12}\to A}, \tag{A.20}$$

$$\theta_{C_{12}}(2L,\tau_f) < K_d^{C_{12}\to B} \tag{A.21}$$

where

$$[\text{3OC6HSL}](t) = \theta_{C_6}(x,t), \tag{A.22}$$
$$[\text{3OC12HSL}](t) = \theta_{C_{12}}(x,t), \tag{A.23}$$
$$t = [0, \tau_f] \text{ is the time period when cells are in the exponential growth,} \tag{A.24}$$
$$K_d^{C_6 \to A} \text{ is the dissociation constant of 3OC6HSL on } plux, \tag{A.25}$$
$$K_d^{C_{12} \to A} \text{ is the cross-talk dissociation constant of 3OC12HSL on } plux, \tag{A.26}$$
$$K_d^{C_{12} \to B} \text{ is the dissociation constant of 3OC12HSL on } plas, \tag{A.27}$$
$$K_d^{C_6 \to B} \text{ is the cross-talk dissociation constant of 3OC6HSL on } plas. \tag{A.28}$$

With the parameter values in Table 3.2, the optimum channel length was determined to be 4.5 mm $\leq l \leq$ 9 mm (Figs. A.5, A.6).

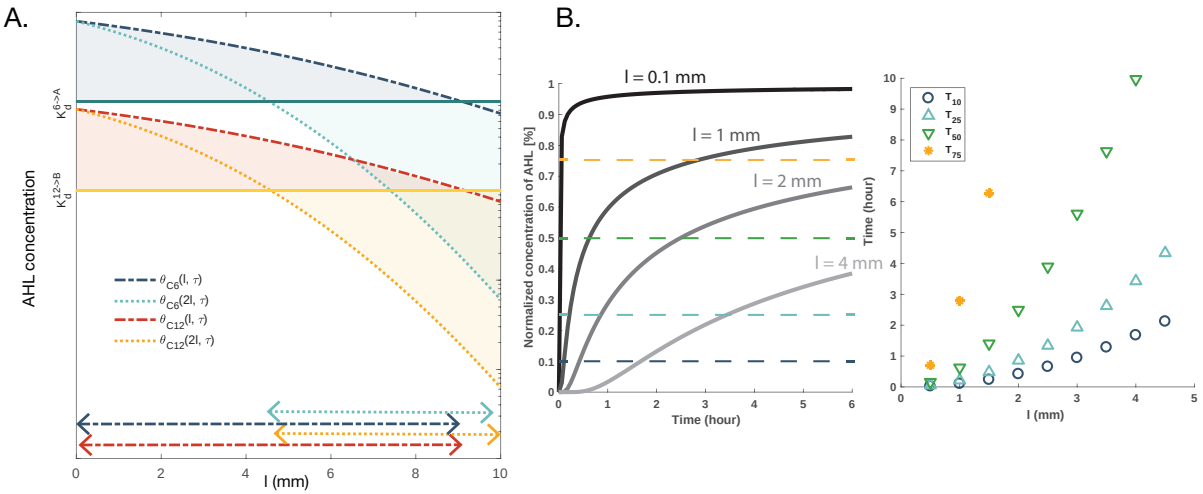

Figure A.5: PDE simulation for DLI device design to optimize the inter-compartmental channel length (l) and the channel length connecting a compartment and reservoir (l_{res}). (A) Graphical representation of constraints on l. The inter-compartmental channel length should allow sufficient AHL diffusion in directly adjacent compartments while keeping the AHL concentration in the second closest compartments (distance $2l$) to be below the threshold (K_d of AHL-inducible promoters). A constraint can be represented as an area under or above a function, and the l values where all areas overlap indicates appropriate channel length range. The smallest l value satisfying the criteria is optimized for shortest communication time, while largest l value is optimized for longest non-neighbor communication time. 4.5mm $< l <$ 9 mm satisfies the criteria with the experimentally evaluated parameters given in Table 3.2. (B) Characterization of l_{res} length and decay time. The channel length connecting a compartment and a reservoir adds extra decaying mechanism for AHL by allowing AHL efflux from every compartment to the bulk solid medium of $AHL \simeq 0$. The efflux diffusion rate from the compartment to medium depends on the length of the channel. The left plot shows the portion of AHL diffused out with respect to time, normalized to the AHL concentration inside the compartment at $t = 0$. The right plot shows the time constant for a portion of the AHL concentration inside the compartment to diffuse outside with respect to the channel length.

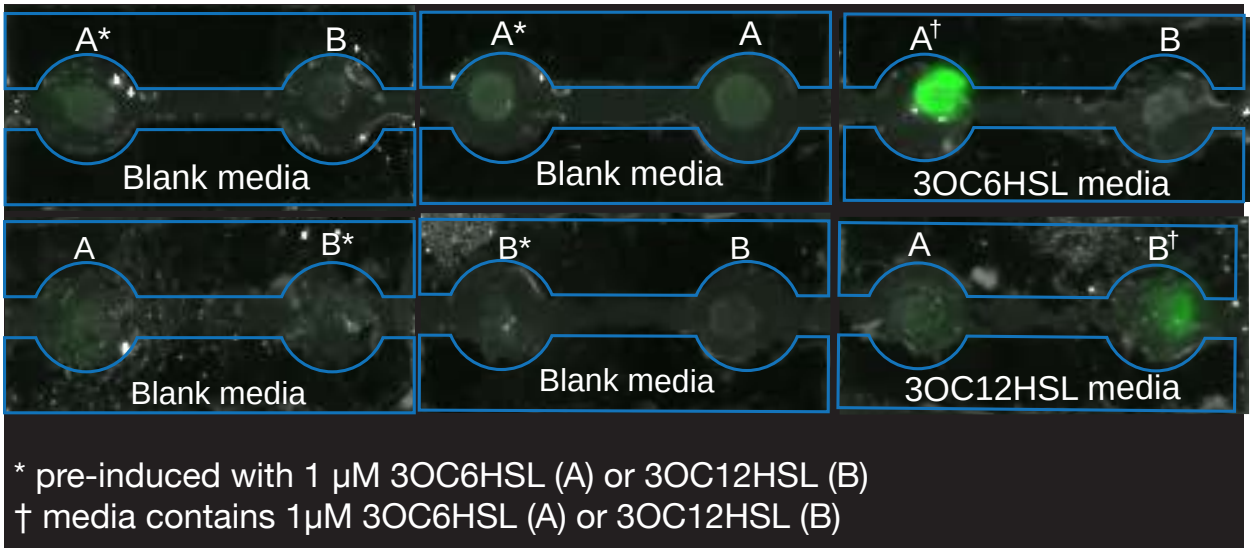

Figure A.6: A plate fluorimeter image of the 1:1 spatial configuration devices with channel length of 9 mm. The image was taken after 12 hours of growth at room temperature. ∗indicates pre-induced strains with 1 μM AHL and † indicates strains that were biased to be fluorescent by externally added AHL in medium. A 9-mm channel length was too long to establish communication between the adjacent compartments, deviating from predictions based on computational optimization. Positive controls were prepared by mixing either 1 μM 3OC6HSL or 3OC12HSL in solid medium on the right-most devices in order to bias the gene expression pattern to be either Strain A fluorescent or Strain B fluorescent. The deviation is possibly caused by the approximation to boundary conditions in the PDE solution using the upper bound of AHL concentration.

Appendix B

Cell-in-the-loop: Materials and Methods

This appendix describes the detailed materials and methods for the experiments in Chapter 4, based on material published in [CITE].

B.1 Plasmid and yeast strain construction

Escherichia coli TOP10 cells (Invitrogen) were used for plasmid cloning and propagation. The dPSTR reporter plasmid (pDB161) contains the coding sequences of UbiY-2xSV40NLS-SynZip1 [378], expressed from an EL222-responsive promoter (p5xBS-CYC180) [377], and mCherry-SynZip2 [378], expressed from the constitutive *ACT1* promoter. It was constructed by first replacing the promoter p*RPL24A* in the plasmid pDA183 [378] by p*ACT1* using SacI-XbaI cut sites and subsequently replacing the promoter p*STL1* by P5xBS-CYC180 using PCR and SapI-based Golden Gate cloning [389].

The yeast strain used in this study (DBY165) was constructed by transforming the PacI digested plasmid pDB161 into DBY41 [377], a strain with BY4741 background expressing VP-EL222 [379] from the *ACT1* promoter. The transformation was performed using the standard lithium acetate method [390].

B.2 Culture preparation

Cells were grown at 30°C in synthetic medium (SD) consisting of 2% glucose, low fluorescence yeast nitrogen base (Formedium), pH 5.8, 5 g/l ammonium sulfate, and complete supplement of amino acids and nucleotides. Cultures were started from plate, diluted, and maintained at $OD_{600} < 1.5$ between 24 and 32 hr before an experiment. For each experiment, between 3 and 5 mL of cell culture were centrifuged at 20, 3000 RCF for 6 min and enough supernatant was removed to achieve an approximate OD_{600} of 4 after resuspension. Cells were then immediately placed on agarose pads.

B.2.1 Agarose pad preparation

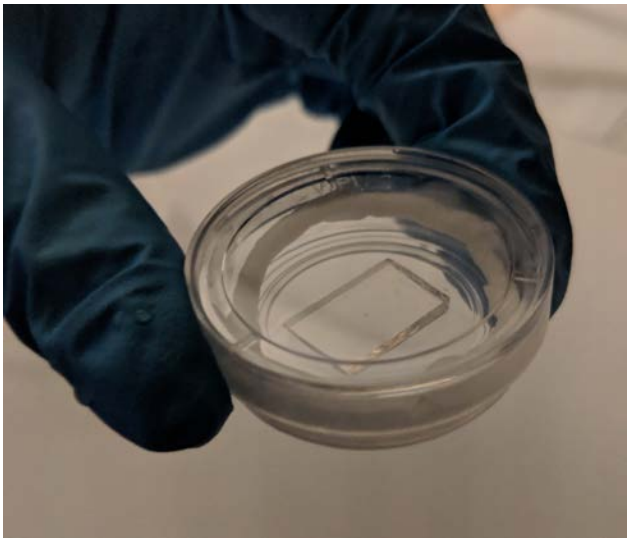

Figure B.1: Example agarose pad used in experiments.

Two 2-slide-tall stacks of microscope slides were placed 1 cm apart parallel to each other on top of a single microscope slide. 70 µL of 2% agarose (UltraPure™Agarose, Invitrogen) and SD medium solution were pipetted between the two stacks. A square 18 mm × 18 mm cover slip was gently placed on the top. The pad was solidified for 1 hr. Immediately before placement under the microscope, the stacks and cover slip were removed and the ends of the pad were sliced off with a scalpel such that the final pad was about 15 mm × 15 mm and level across the top. 3 µL of cell suspension were pipetted in increments of 1 µL onto three separate areas of the pad to ensure that at least one would have the correct density for use in the experiment. The pad was overturned into a circular tissue culture dish with cover glass bottom (35 mm FluoroDish™, World Precision Instruments) lined inside with a strip of damp paper towel to maintain humidity throughout the course of the experiment. The dish was closed and sealed with a strip of parafilm, then placed in the microscope's environmental box (Life Imaging Services, Switzerland). Cells were allowed to settle for 30 min before experiment start.

B.3 Imaging

Images were taken under a Nikon Ti-Eclipse inverted microscope (Nikon Instruments) with a 40x oil-immersion objective (MRH01401, Nikon AG, Egg, Switzerland), pE-100 bright-field light source (CoolLED, UK), and CMOS camera ORCA-Flash4.0 (Hamamatsu Photonic, Solothurn, Switzerland) water-cooled with a refrigerated bath circulator (A25 Refrigerated Circulator, Thermo Scientific). The temperature was maintained at 30°C by an

opaque environmental box (Life Imaging Services, Switzerland), and a dark cloth was additionally placed over the microscope to fully shield cell samples from external light. Experiments were conducted with a diffusor and a green interference filter placed in the bright-field light path, with the Nikon Perfect Focus System (+/- 5 AU) enabled. Fluorescence images were acquired using a Spectra X Light Engine fluorescence excitation light source (Lumencor, Beaverton, USA), filter cube with excitation filter 565/24 nm, emission filter 620/52 nm, and beam splitter HC BS 585 (AHF Analysetechnik AG, Tübingen, Germany). The final fluorescence images used for analysis were maximum projections across z-stacks of 5 images spanning 0.6 μm.

During experiments, the microscope was operated by the open-source software YouScope [391]. Cell segmentation and tracking were performed on brightfield images using software tools developed by [375] based on [392] and [393]. For each cell, the mean fluorescence in the nucleus, cytoplasm, and across the entire cell were automatically calculated using custom Matlab® (MathWorks) scripts.

B.3.1 Scoring

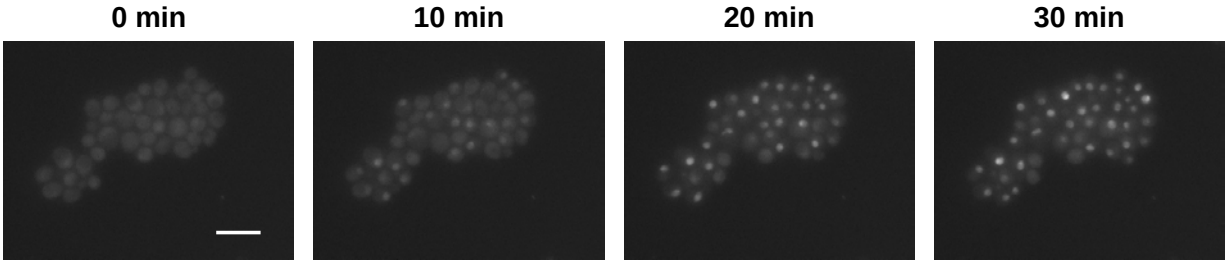

Figure B.2: Representative time lapse of reporter system. Pictured are maximum projection fluorescence images for cells under constant illumination, taken from one of four preliminary dose response experiments (see Section B.5.1). Scale bar is 10 µm.

We assessed induction of gene expression using a fast-acting, nuclear translocation-based reporting system [378], where higher responses corresponded to greater fluorescence in the cell nucleus. Accordingly, we defined the scoring scheme as

$$\text{score} = \frac{\text{mean fluorescence in nucleus} - \text{mean fluorescence in cytoplasm}}{\text{mean fluorescence across entire cell}} \tag{B.1}$$

such that a score of 0 indicates no difference between nuclear and cytoplasmic fluorescence.

Our experiments required us to calculate the response score automatically at each time step. The score for each automatically segmented cell was calculated for an approximated location of the nucleus, determined as follows: First, an image was formed by taking a box around the segmented cell in the maximum projection fluorescence image and converting the non-cell pixels to black. A black border 5 pixels wide was added around the image.

Then the image was blurred with a Gaussian filter of standard deviation 2 pixels and the brightest pixel in the blurred image was located. This point was considered to be the center of the nucleus. We noted by observation that the nucleus was almost always 5 pixels in radius, therefore the pixels falling within a circle of radius 5 pixels around the center were presumed to belong to the nucleus and were subsequently used in the calculations of the mean fluorescence. All remaining pixels belonging to the cell were considered to belong to the cytoplasm. The score was then calculated as indicated above.

We estimated the time constant for cell response as the time to reach half the maximum score, calculated from an average score over cells. For these experiments, we sampled cells every 2 min for a step input from zero to the maximum intensity delivered during patterning experiments, and obtained a time constant of 10.9 min ($N = 63$ cells). We also calculated the time constant based on an input drop from maximum intensity to zero, and obtained a similar time constant of 10.3 min ($N = 65$ cells).

B.3.2 Segmentation/tracking errors

The automated imaging pipeline occasionally failed to identify cells in particular frames. In the majority of cases the system was able to recover the cell within one or two frames. In dose response experiments, cells that were not tracked for the entirety of a dose were discarded. During patterning experiments, cells that were not tracked in a frame did not receive input for the following 10 min, and scores for their corresponding patches were calculated as averages over the scores of the remaining cells. In no experiments were all cells in a patch simultaneously dropped in the same frame.

B.4 Light-delivery system

Optogenetic inputs were delivered to cells using the setup developed in [375], in which images generated on the computer are projected by a digital mirror device through a system of lenses that focuses the light onto a microscope slide. Two neutral density filters (Thorlabs, 25 mm absorptive, optical densities 0.5 and 1.3) were placed serially to achieve a total density of 1.8.

To ensure light mapped properly from the projector to the cell, images were modified prior to projection in order to map pixels on the DMD to pixels in the camera images. The mapping was determined through the procedure outlined in [375], Figure S6B. The procedure was performed immediately before experiment start on an area of the agarose pad unoccupied by cells.

Custom Matlab® code was used for manually calibrating projector intensity before experiment start and for automatically carrying out experiments.

16-patch

K	Fraction Dropped Measurements
0.1	0.041666667
0.1	0.029605263
0.1	0.049342105
0.1	0.026864035
0.2	0.038377193
0.2	0.038377193
0.2	0.091557018
0.3	0.10252193
0.3	0.097039474
0.3	0.064144737
1	0.029057018
1	0.04879386
1	0.090460526

36-patch

K	Fraction Dropped Measurements
0.1	0.094298246
0.1	0.041666667
0.1	0.119152047
0.1	0.148391813

Table B.1: Fraction of all measurements that were dropped per experiment. One measurement is a single cell in a single frame.

B.4.1 Ensuring uniform illumination intensity

We observed that there was a sigmoidal relationship between the administered illumination intensity and the measured illumination intensity in the projection images, and also that images of uniform intensity did not evenly illuminate the sample plane. To compensate for these effects, we modified images before projection to linearize the administered-to-measured intensity relationship and also to reduce the intensity of overilluminated regions to match the level attained by underilluminated regions.

We calculated our intensity corrections based on the following model: If u is an $N \times N$ input image (normalized to $[0, 1]$) and u_r is that image magnified to size $M \times M$, then the measured intensity y (normalized to $[0, 1]$) is an $M \times M$-dimensional image given by

$$y = s(v(u_r)) \tag{B.2}$$

where $s(\cdot)$ is a sigmoidal function applied identically to each pixel, and $v(\cdot)$ is a function that varies by pixel to represent the uneven illumination intensity. This suggests that, to achieve a desired y_d, an input image should be calculated as

$$u_r = v^{-1}(s^{-1}(y_d)) \tag{B.3}$$

(which can be appropriately resized to obtain u).

$s^{-1}(\cdot)$ was already determined from previous experiments with the setup [375]. This correction was calculated once, as it did not appear to change between experiments. $v^{-1}(\cdot)$ was calculated before each experiment by sampling the measured intensity at a number of points spaced across the sample plane when the administered intensity was maximized. To do so, one hundred circles arranged in a grid were projected one at a time onto the slide at maximum administered intensity and the reflected images for each circle were measured. The mean measured intensity of all the pixels in a single circle was treated as a sample of the "true" illumination intensity at the point on the sample plane corresponding to the center of the circle, such that a 2D quadratic polynomial surface could be fit to all mean circle intensities in order to interpolate the illumination intensity at all points on the sample plane. These sample points were normalized to the maximum and a second 2D quadratic polynomial surface was fit to the inverse sigmoid of these normalized sample points. A target intensity value was chosen and the surface fit was renormalized relative to this intensity value to obtain a matrix V representing the factor by which bright areas were overilluminated relative to areas of underillumination, such that the matrix $\tilde{V}$ consisting of the element-by-element inverse of V (i.e., $\left[\tilde{V}\right]_{ij} = \frac{1}{[V]_{ij}}$) was normalized to $[0, 1]$. Subsequently, inputs u_r were calculated to achieve the desired output y_d as

$$u_r = \tilde{V} \odot s^{-1}(y_d) \tag{B.4}$$

where $\odot$ indicates element-by-element multiplication.

The sigmoidal intensity correction was used for all experiments. The flattening correction procedure was carried out before experiment start on an area of the plate unoccupied by cells, and was performed for all experiments except the dose response experiments.

B.5 Dose response

The theory relies on a deterministic dose response curve in which the expected steady-state response score of a cell increases as a function of constant input intensity. We performed a series of dose response experiments to verify that these conditions held for our yeast strain and then calculated a dose response curve from the average response of cells to varying measured projected intensities.

B.5.1 Preliminary dose response experiments verify gradedness and independence of dose history

In order for the theory to apply, we needed to verify (a) that the magnitude of cell response increased with received light intensity; and (b) that the dose response would be independent of dose history for the duration of the final patterning experiments. By "dose

history", we refer to the number, intensity, and order of doses administered to cells prior to a particular time of interest.

The following experiments were performed with an ND filter of optical density 2 (Thorlabs, 25 mm absorptive). Cells were illuminated for 10 min with uniform light, then left in the dark for 20 min before doses were administered. Cells were imaged every 10 min. Doses were administered for 40 min and the administered intensity of a dose varied for different collections of cells on the same plate simultaneously. Hereafter, we use "dose response experiment" to refer to a particular collection of cells on the same plate receiving the same series of administered doses. Two dose response experiments were conducted simultaneously per plate, with four or five doses per experiment. For a single time point, the measured projected intensity received by a cell was calculated as the average of the mean measured projected intensity across all pixels occupied by the cell.

A series of general linear models were fit to the data for each cell. The natural log of the score (response variable) was treated as a function of the agarose pad, frame, dose number (ordered by time of appearance during experiment), frames since dose start, current measured illumination intensity, and the measured illumination intensity for all frames up to the minimum experimental duration (16 frames) before the current frame. Intensities for time points before the start of the experiment were set to 0. An analysis of deviance for the full model (Table B.2) suggests that the illumination history past the current illumination contributes little to the current score.

Furthermore, we determined by observation that cells had reached a quasi-steady state response before 30 min, and found that fitting a general linear model to steady-state times only (30 min and 40 min into a dose) greatly diminishes the importance of plate, frame, dose number, and frames since dose start (Table B.3). Moreover, a reduced model for the steady-state score that includes only the current intensity has an AIC substantially similar to that of the full steady-state model (2675 for the reduced vs. 2622 for the full). Together, these analyses suggest that it is reasonable to consider steady-state dose response as a function of current illumination intensity alone.

B.5.2 Procedure for final dose response experiments

Cells tended to respond much more strongly and unpredictably to the first administered input than to later inputs regardless of the intensity of the first input. Therefore, before administering any doses, all cells on the dish were illuminated for 10 min with uniform, middling intensity light, then left in the dark for 20 min to allow the response to decay. Multiple doses were then administered in immediate succession to cells on the plate. For a single dose, cells were illuminated with individually targeted light with constant administered intensity per cell. Cells were imaged every 10 min. Cells that were not successfully segmented and tracked at all sampled time points in a dose were discarded. The steady-state response of a single cell was calculated as the average of the scores from 40 to 80 min under illumination.

The final dose response curve was fit to data aggregated from three experiments. Individual cell responses were binned by projected intensity into 6 quantiles. The final dose

Variable	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL	NA	NA	4117	3905.828561	NA
Im0	1	1051.353764	4116	2854.474798	0
Im1	1	148.2219005	4115	2706.252897	1.65E-53
Im2	1	0.229287964	4114	2706.023609	0.544725097
Im3	1	4.35808067	4113	2701.665528	0.008275942
Im4	1	1.320218071	4112	2700.34531	0.146119572
Im5	1	0.074542198	4111	2700.270768	0.729833304
Im6	1	0.240410136	4110	2700.030358	0.535126523
Im7	1	1.161222828	4109	2698.869135	0.17286582
Im8	1	0.187482186	4108	2698.681653	0.583904657
Im9	1	2.286447549	4107	2696.395205	0.055792838
Im10	1	0.249079692	4106	2696.146126	0.527855487
Im11	1	0.596308281	4105	2695.549817	0.328685196
Im12	1	0.142415919	4104	2695.407402	0.633115034
Im13	1	0.323652973	4103	2695.083749	0.471767645
Im14	1	0.836572299	4102	2694.247176	0.24730139
Im15	1	1.060493657	4101	2693.186683	0.192713497
Im16	1	3.170352656	4100	2690.01633	0.024309053
FramesSinceDoseStart	1	106.9288382	4099	2583.087492	4.29E-39
Pad	1	13.86651829	4098	2569.220973	2.47E-06
Frame	1	0.002190109	4097	2569.218783	0.952796489
DoseNumber	1	9.159326761	4096	2560.059457	0.00012912

Table B.2: General linear model of cell response score at all points in time. Model was calculated using data from four preliminary dose response experiments. ImX indicates the measured projected intensity preceding the current by X frames (0 is current frame). The AIC for this model is 9745. As expected, the time since dose start contributes a large drop in deviance, as cells did not instantly settle to a new steady state when the intensity was changed.

response curve in the form of a leaky activating Hill function

$$f(x) = a + b\frac{x^n}{c^n + x^n} \tag{B.5}$$

was fit to the mean score values within each quantile vs. the mean measured projected intensity within that quantile.

B.5.3 Patch-level dose response

The deterministic theory operates under the assumption that the dose response curve is identical for all cells and therefore all patches. To examine the similarities between the empirical fitted dose response curve and actual patch-level dose responses, we bootstrapped

Variable	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL	NA	NA	1185	1140.234564	NA
Im0	1	481.0998049	1184	659.1347596	1.03E-201
Im1	1	15.65834255	1183	643.476417	4.58E-08
Im2	1	0.206888356	1182	643.2695287	0.529734222
Im3	1	1.351828641	1181	641.9177	0.108199518
Im4	1	1.058581525	1180	640.8591185	0.155179456
Im5	1	0.658939118	1179	640.2001794	0.262075464
Im6	1	0.111574008	1178	640.0886054	0.644450512
Im7	1	0.009561578	1177	640.0790438	0.892536317
Im8	1	2.783572744	1176	637.2954711	0.021164596
Im9	1	1.374292916	1175	635.9211782	0.105311673
Im10	1	3.334521299	1174	632.5866569	0.011640495
Im11	1	1.633214394	1173	630.9534425	0.077459177
Im12	1	0.519193207	1172	630.4342493	0.319493668
Im13	1	0.000278015	1171	630.4339712	0.981621402
Im14	1	3.592110318	1170	626.8418609	0.008831956
Im15	1	2.201528614	1169	624.6403323	0.040371525
Im16	1	3.868935675	1168	620.7713966	0.006577405
FramesSinceDoseStart	1	4.202474492	1167	616.5689221	0.004622429
Pad	1	0.000989428	1166	616.5679327	0.965336552
Frame	1	6.225817085	1165	610.3421156	0.000566296
DoseNumber	0	0	1165	610.3421156	NA

Table B.3: General linear model of cell response score during steady state. Model was calculated using data from four preliminary dose response experiments for steady-state frames (30 and 40 min after dose start) only. ImX indicates the measured projected intensity preceding the current by X frames (0 is current frame). Note that the frames since dose start, pad, frame, and dose number are less significant in this model relative to the full model in B.2. The AIC for this model is 2622.

from a single dose response experiment for $N = 100$ patches with 4 or 6 cells per patch. Our final setup did not permit us to construct full dose response curves for individual cells or patches, as we could administer a maximum of four doses per experiment, which is not sufficient to fit a curve with four degrees of freedom. Therefore, we visually compared a linear interpolation of bootstrapped dose responses to the administered doses with the empirical fitted dose response curve (Figure B.3). The patch-level dose responses do exhibit some level of variation (higher for higher projected intensity), which probably manifests in the experiments as systematic variation in average patch intensity. This patch-patch variation, in turn, may be the most notable contributor to the experiment-to-experiment variation in exact contrast level and overall brightness observed in our patterning system (Figure 4.7, Figure 4.6).

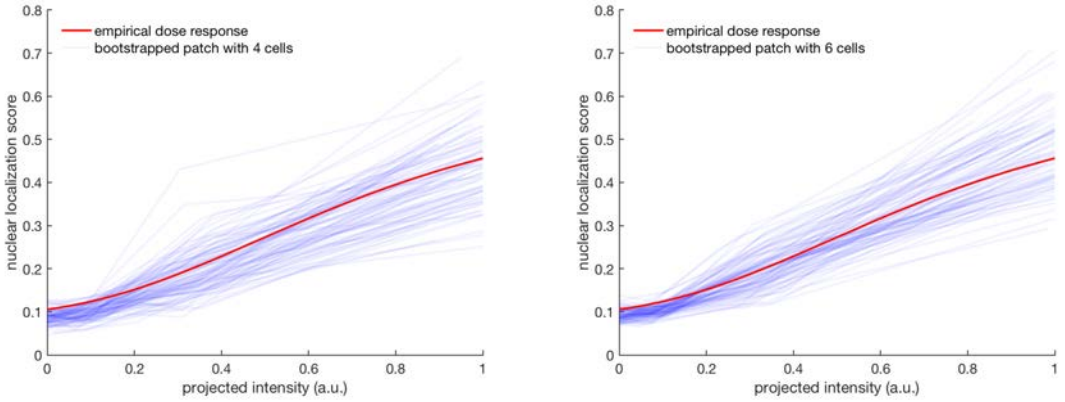

Figure B.3: Patch-level dose responses approximate empirically fitted dose response. Fitted dose response (red) vs. 100 bootstrapped dose responses (blue) for patches of 4 or 6 cells, drawn from a single dose response experiment. Note that the fitted response is a curve, while the bootstrapped responses are linear interpolations over four points.

B.6 Patterning experiments

Before experiment start, all cells on the plate were illuminated for 10 min with uniform, middling intensity light, then left in the dark for 10 min such that responses would not fully decay, allowing for some initial variation in nuclear localization score. Experiments lasted 3 hr, during which cells were imaged and their inputs adjusted every 10 min. Preliminary experiments confirmed that cells were alive and responsive up to 6 hr after placement under the microscope, although final experiments were constrained to 3 hr to ensure cells remained in a monolayer.

B.6.1 Patch construction

Cells were randomly assigned to groups with a fixed number of cells per group. Each group corresponded to a single computationally defined "patch". The score for a patch was calculated as the average of the scores of the cells comprising the patch. By "administering an input to a patch", we mean the cells in that patch were individually targeted with the same administered input. Most experiments were performed with 16 patches of 6 cells per patch. Higher-dimensional experiments were performed with 36 patches of 4 cells per patch.

Patches were arranged in "virtual space" into a square grid where each patch was connected to (interacted with) each of the patches to the north, south, east, and west, with periodic (wrap-around) boundary conditions such that each patch interacted with four other patches (its "neighbors"). Note that, because cells were randomly assigned to patches, cells to neighboring patches in virtual space were not necessarily adjacent in real space.

B.6.2 Input adjustment time

The frequency of input adjustment was lower bounded by the amount of time it took for image acquisition and analysis (minimum ∼2 min, somewhat longer the more cells were in the field). Adjustment time was also loosely upper bounded based on the practicality that cells would eventually grow out of a monolayer, which mandated that inputs be adjusted with enough frequency for a pattern to appear within 3 or 4 hours.

Within these constraints, we chose the frequency of input adjustment to be on the same order of magnitude as the response time of the cells. Specifically, we decided to adjust inputs faster than the ∼20 or 30 minutes for individual cells to reach steady state so that we could (a) ascertain when the *full system* (not just individual cells) had reached steady state by plotting the instantaneous input/output relation during the experiment, and (b) subject the theory to a more stringent test (i.e., make sure transient patch dynamics could still affect the dynamics of the full system without nullifying predictions made from single-cell steady-state behavior alone).

B.6.3 Signaling relation

Every imaging period (10 min), the input to each patch was adjusted according to the signaling relation, which was chosen to be a Hill function with a minor computational adjustment to better utilize the available range of illumination intensities. Specifically, the relation was given by

$$h(v) = \frac{K^n}{K^n + (\max(s_{\min}, v) - s_{\min})^n} \tag{B.6}$$

where v was the average score across all four neighbors of a patch, n was fixed at 2, and the minimum score $s_{\min} := 0.05$ was the cutoff to deliver maximum illumination intensity. The bifurcation parameter K was fixed within a given experiment at a value between 0.1 and 1, with higher K corresponding to weaker repression.

In particular, let x_k^i be the score of the ith patch in the kth frame, and u_k^i be the input to the patch between the kth and $(k+1)$th frames. Then u_k^i was calculated as

$$u_k^i = h\left(\sum_j x_k^j\right) \tag{B.7}$$

where the summation is taken over the neighbors of patch i and $L(\cdot)$ is as given in equation (B.6).