

Inferring bifurcations between phenotypes

Grisha Szep

A dissertation submitted in fulfillment
of the requirements for the degree of

Doctor of Philosophy

at

King's College London

Randall Division of Cell & Molecular Biophysics



funded by **Microsoft Research Cambridge**



October 29, 2021

Abstract

The gene-expression history of an organism and its environment determine the organism's phenotype. The phenotype is an inherently qualitative state, deduced by relative biochemical concentration measurements collected by methods such as flow cytometry or fluorescence microscopy. The biochemical threshold concentrations that distinguish different phenotypes can be modelled by applying bifurcation analysis to differential equation models and the search for these boundaries in experimental data can be done using dimensionality reduction and clustering techniques. This establishes a relationship between bifurcations, phenotypes and machine learning techniques that are the subject of this thesis.

The first chapter presents an interactive tool for exploring phenotypes in flow cytometry data. In particular we explore a multi-tissue, high-dimensional, immune cell dataset. The tool bridges machine learning methods and the popular FlowJo, used to annotate cells with gating strategies. An assortment of dimensionality reduction techniques are applied to create two dimensional embeddings and confusion matrices are used to quantify annotation agreement between immunologists. By leveraging the geospatial mapping library OpenLayers to render, annotate and analyze cells, immunologists can now efficiently navigate the phenotype space of Human Cell Atlas datasets.

The next chapter focuses on a model-driven approach for exploring and designing phenotypes, where we demonstrate how model-guided design of synthetic E. Coli can elucidate pattern formation mechanisms in

multicellular development. We infer the parameters of a biochemically motivated system of differential equations against time course fluorescence data acquired from plate reader experiments. Our design goals however were not in the temporal domain, rather we wanted to control the shape and size of a cusp bifurcation in the space of experimentally controlled input concentrations.

To address these limitations, I define a differentiable semi-supervised cost function that uses bifurcation locations as targets. Bifurcations are encouraged by an unsupervised term that extremises the curvature of the determinant of the Jacobian. By exploring the cost landscape for minimal models that span the space of saddle-nodes and pitchforks, I show that the parameter space basins define regions of qualitatively equivalent differential equations. The differentiability of the cost function enables efficient optimisation using libraries such as Flux.jl that leverage automatic differentiation. The impact of this work would enable experimentalists to efficiently navigate design spaces of differential equation models.

Acknowledgements

Acknowledgements to people I love

Contents

1	Introduction & Motivation	14
1.1	Bottom-up and Top-down Biology	15
1.2	Design, Build, Test, Learn	17
1.3	State of the Art	18
1.3.1	Inference of Differential Equations	18
1.3.2	Model Reduction and Classification	19
1.4	Design–Learn Pipeline	19
1.4.1	Experimental Design Loop	20
1.4.2	Hypothesis Scoring and Model Refinement Loop . .	21
1.4.3	System Design	22
2	Theoretical Background	23
2.1	Describing Phenotypes with Bifurcations	24
2.1.1	Differential Equation Models	27
2.1.2	Bifurcation Analysis	35
2.1.3	Spatially Extended Systems	39
2.2	Applications in Cell Biology	40
2.2.1	Genetic Switches & Phenotype Boundaries	40
2.2.2	Populations of Neurons	41
2.2.3	Self-organised Patterns & Development	41
2.3	Phenotype Inference with Machine Learning	41
2.3.1	Classification and Regression	42
2.3.2	Dimensionality Reduction Methods	42

<i>Contents</i>	6
2.3.3 Clustering Methods	42
2.4 Applications in Flow Cytometry	42
2.4.1 Immunophenotyping Panels	42
2.4.2 Inferring Bifurcations from Data	42
2.4.3 Multiomics Integration	48
3 Interpretation of Morphogen Gradients by a Bistable Circuit	49
3.1 Preface	49
3.1.1 Problem Statement & Context	50
3.1.2 Contributions	54
3.2 Abstract	55
3.3 Introduction	56
3.4 Results	56
3.4.1 Engineering mutual exclusivity	56
3.4.2 Mutual inhibition results in bistability	57
3.4.3 Hysteresis produces stable boundaries	58
3.4.4 A secondary gradient creates self-organised domains .	58
3.5 Discussion	59
3.6 Methods	60
3.6.1 Plasmid construction	60
3.6.2 Plate fluorometer assay	60
3.6.3 Flow-cytometric analysis of hysteresis	60
3.6.4 Microfluidics	61
3.6.5 Microfluidics microscopy	61
3.6.6 Solid culture assays	61
3.7 Afterword	63
4 Parameter Inference with Bifurcation Diagrams	64
4.1 Preface	64
4.1.1 Problem Statement & Context	64
4.1.2 Contributions	65

	<i>Contents</i>	7
4.2 Abstract	67	
4.3 Introduction	67	
4.3.1 Preliminaries	68	
4.4 Proposed Method	70	
4.4.1 Semi-supervised Cost Function	70	
4.4.2 Differentiating the semi-supervised cost function	71	
4.5 Experiments & Results	72	
4.5.1 Minimal Models	72	
4.5.2 Genetic Toggle Switch	72	
4.5.3 Complexity	73	
4.6 Conclusion & Broader Impact	75	
4.7 References	76	
4.8 Afterword	78	
5 Exploring bifurcations between phenotypes	79	
5.1 Flow Cytometry	79	
5.1.1 FlowJo	79	
5.1.2 Immunophenotyping	79	
5.1.3 Human Cell Atlas	79	
5.2 Data Exploration	79	
5.2.1 Dimensionality Reduction	79	
5.2.2 Dataset Integration	79	
5.2.3 Interactivity	79	
5.3 FlowAtlas.jl	79	
5.3.1 Julia & SciML	79	
5.3.2 OpenLayers & Data-driven Documents	79	
6 Conclusions	80	
6.1 Limitations	80	
6.2 Future work	80	
Appendices	81	

<i>Contents</i>	8
A Interpretation of Morphogen Gradients by a Bistable Circuit	81
A.1 Supplementary Figures	83
A.2 Supplementary Methods	99
A.2.1 Differential Equation Models & Parameter Inference .	100
A.2.2 Bistability Analysis	121
A.2.3 Boundary Experiments	123
A.2.4 Models of the Exclusive Receiver Relay Circuits . . .	131
B Parameter Inference with Bifurcation Diagrams	133
B.1 Bifurcation Diagrams as Tangent Fields	134
B.2 Bifurcation Measure Properties	135
B.3 Leibniz Rule for Space Curves	136
B.4 Application to the Double Exclusive Model	138
B.5 Extension for Hopf Bifurcations	139
Bibliography	140

List of Figures

1.1	<i>Enigma I</i> reverse engineered by Marian Adam Rejewski	15
1.2	Top-down and bottom-up modelling methods	17
1.3	Design-Build-Test-Learn cycle from synthetic biology	18
1.4	Workflow loop for optimal experimental design, without mechanistic model	20
1.5	Overview of hypothesis scoring pipeline and model refinement loop	21
1.6	Overview of system design pipeline	22
2.1	a) The bifurcations (shown in gold) at critical voltage p^* above argon concentration θ^* give rise to sudden jumps in brightness (right-hand panels). θ^* separates two bulb phenotypes and p^* separates the <i>on</i> and <i>off</i> states.	24
2.2	Waddington landscapes representing the set of available behaviours for two phenotypes in response to a controlled condition p . Phenotypes are related by changes in landscape topology via changes in θ	26
2.3	Illustration of a trajectory between $u = u(t)$ and $u' = u(t')$ over finite time $t - t'$ following the field lines of F_θ . In this example field lines either point towards steady state u^* or a limit cycle $\omega(t)$. There are two unstable fixed points marked by crosses: a saddle point separating the basins of attraction and an unstable focus enclosed by the limit cycle.	28

2.4	Illustration how the Jacobian matrix $\frac{\partial F_\theta}{\partial u}$ transforms the basis vectors $\hat{u}(t)$ defining a small patch of initial conditions for a short period of time Δt . The eigenvalues λ and eigenvectors \hat{v}_λ define the deformation of the patch	29
2.5	Illustration how the time ordering operator $\mathcal{T} \left\{ e^{\int_{t'}^t J(s) ds} \right\}$ propagates the separation $\delta u(t)$ between adjacent trajectories u and $u + \delta u$ by repeated application of the exponentiated Jacobian $e^{J(t)\Delta t}$ evaluated along unperturbed trajectory $u(t)$	32
2.6	Classification of stable and unstable fixed points for a general two dimensional Jacobian in terms of trace $\text{Tr}[J]$ and determinant $ J $	34
2.7	All bifurcations	37
2.8	Hopf, Cyclic-fold and Infinite-Period bifurcation diagram showing stable and unstable fixed points and limit cycles as a function of parameter $\sigma[r]$. Insets show nullclines [tbc] .	39
2.9	Bifurcation point extraction along the condition p via two possible routes: steady state inference (a,d→c) and fixed point inference (a,d→b,e→c). The fixed point inference route requires the geometry of the field $F(u)$ at different values of the condition p and yields additional information about unstable fixed points.	43
2.10	Protocols (a) for extracting limit points with respect to conditions p, p' . Steady state distributions (b). Similarity measure (c)	44
2.11	Gaussian process regressors estimating derivative of the trajectories \hat{f} from example trajectory datasets $\mathcal{D}_1 \dots \mathcal{D}_K$ with varying signal to noise ratios. Interpolation error E is shown as a heatmap; extrapolation fails	47
3.1	Geometry of opposing gradients experiment	49

3.2	Top: <i>Hox</i> gene expression patterns in body segments of drosophila. Bottom: Mutation where legs grow in-place of antenna	51
3.3	Expression patterns of pair-rule <i>Gap</i> genes in Drosophila embryo []	52
3.4	Pigment patterns hypothesised to be generated by Turing mechanism	52
3.5	Left: Diagram of double exclusive reporter showing wiring between promoters and complexes Right: Repression of <i>pTet</i> and <i>pLac</i> promoters via <i>TetR</i> and <i>LacI</i> respectively	53
3.6	Top Left: Optical density measurements obtained for one well. Top Right: Final fluorescence values at 16h across different C_6 , C_{12} concentrations. Bottom: Steady state predictions between data regions.	54
3.7	<i>Fig. 1</i> A synthetic gene circuit for morphogen interpretation.	56
3.8	<i>Fig. 2</i> Mutual inhibition produces bistability.	57
3.9	<i>Fig. 3</i> Formation of stable boundaries.	59
3.10	<i>Fig. 4</i> Addition of a Relay circuit creates self-organised domains of gene expression.	60
4.1	<i>Fig. 1</i> Illustration of bifurcation diagrams for minimal models of bifurcations. A. Saddle-node bifurcations arise for $F_\theta(u, p) = p + \theta_1 u + \theta_2 u^3$ when $\theta = (\frac{5}{2}, -1)$. B. Pitchfork bifurcations arise for $F_\theta(u, p) = \theta_1 + pu + \theta_2 u^3$ when $\theta = (\frac{1}{2}, -1)$. Targets are illustrated by light yellow vertical lines. Bifurcation curves are shown as solid blue and red lines, with lighter shades indicating the determinant crossing zero at locations $\mathcal{P}(\theta)$ giving rise to unstable solutions.	69

4.2 <i>Fig. 2</i> Bifurcation measure $\varphi_\theta(s)$ and determinant $\left \frac{\partial F_\theta}{\partial u} \right $ along the arclength s of two different bifurcation curves demonstrating how maximising the measure along the curve maintains the existing bifurcation marked by a circle, while encouraging new bifurcations marked by stars.	70
4.3 <i>Fig. 3</i> Saddle-node $F_\theta(u, p) = p + \theta_1 u + \theta_2 u^3$ and pitchfork $F_\theta(u, p) = \theta_1 + up + \theta_2 u^3$ optimised with respect to θ so that predicted bifurcations $\mathcal{P}(\theta)$ match targets \mathcal{D} in control condition p . The right panel shows bifurcations diagrams for the three optimal θ^* marked by stars on the left panel. The optimisation trajectories in white follow the gradient of the cost, approaching the black lines of global minima in the left panel	72
4.4 <i>Fig. 4</i> Bifurcation inference for the two-state model (11). A. Optimal parameter estimates θ^* for the targets $\mathcal{D} = \{4, 5\}$ reveal two clusters of qualitatively different regimes: mutual activation ($a_1 < 1$; cluster 1) and mutual inhibition ($a_1 > 1$; cluster 2). B. Example bifurcation diagrams indicate positively and negatively correlated dependencies between the two model states, as a function of the control condition.	73
4.5 <i>Fig. 5</i> A. Execution time (time to calculate cost gradient) with respect to states N . B. Convergence times (the time it takes to find and match a bifurcation to within 1% of a specified target) with respect to the number of parameters M , comparing against a gradient-free approach: Nelder-Mead. Calculations were performed on an Intel Core i7-6700HQ CPU @ 2.60GHz x 8 without GPU acceleration.	74
A.1 <i>Supplementary Figure 1</i> Circuit variants	83
A.2 <i>Supplementary Figure 2</i> Raw timecourse fluorescence traces .	85

A.3	<i>Supplementary Figure 13</i> Hysteresis flow cytometry experiments	96
A.4	<i>Supplementary Figure 25</i> Bifurcation curves for uniform and protected degradation models	122
A.5	<i>Supplementary Figure 26</i> bifurcation curve insensitivity specific growth rate γ_0	122
B.1	<i>Supplementary Figure 1</i> Two implicit surfaces $f_\theta(z) = 0$ and $g_\theta(z) = 0$ in \mathbb{R}^3 intersecting to form a space curve which is tangent to field $T_\theta(z)$ and perpendicular to gradients $\partial_z f_\theta$ and $\partial_z g_\theta$	134
B.2	<i>Supplementary Figure 2</i> Left/Right : Determinant $\left \frac{\partial F_\theta}{\partial u} \right $ and tangent field $T_\theta(z)$ for the saddle-node/pitchfork models for some set values of θ revealing that $\left \frac{\partial F_\theta}{\partial u} \right = 0$ defines bifurcations	135
B.3	<i>Supplementary Figure 3</i> Bifurcation inference for the <i>double exclusive reporter</i> . A. Optimal parameter estimates θ^* for the targets $\mathcal{D} = \{1,2\}$ (indicated by yellow lines in panel B) reveal four regions with two geometrically different regimes: mutual activation (region 1) and mutual inhibition (regions 2-4). B. Example bifurcation diagrams indicate that region 2 has swapped kinetics between L and T to region 3. Region 4 has models with non-zero imaginary parts to eigenvalues indicating damped oscillations (shown in light green).	138
B.4	<i>Supplementary Figure 4</i> Bifurcation measure $\varphi_\theta(s)$ and eigenvalues $\lambda(s)$ along the arclength s for two different bifurcation curves demonstrating how the measure detects non-zero imaginary parts $\Im[\lambda]$ (onset of damped oscillations marked by circle) and sign changes in real parts $\Re[\lambda]$ (Hopf bifurcations marked by stars)	139

Chapter 1

Introduction & Motivation



one thing that won't change with time is the memories of younger years

Minuet of Forest — Ocarina of Time

The advent of the modern digital computer, as formalised by Alan Turing,[1] ignited the field of computational physics, aided by pre-existing theoretical formulations of algorithms. Starting from the first experiments with Monte Carlo (MC) simulations in the 1930s by Fermi and the formulation of the Markov-Chain Monte Carlo (MCMC) technique by Ulam in the 1940s, von Neumann programmed the 18,000 vacuum-tube Electronic Numerical Integrator and Computer (ENIAC) computer to investigate neutron diffusion in fissionable materials.[2] This success paved the way for the integration of Newton's equations of motion to compute the time evolution of a many-body system.

Consider being Marian Adam Rejewski, a Polish mathematician and cryptologist in 1932, almost seven years before the beginning of World War II. It was known at the time that the Germans were using a machine called *Enigma I* to send secret messages, but no one knew what the machine looked like nor how it worked. Based on a handful of intercepted en-

cripted messages and pictures of the operating instructions obtained by Polish intelligence, he derived a set of permutation equations that allowed him the reverse-engineer the machine [3]. The cryptologic achievements of Rejewski and colleagues Jerzy Różycki and Henryk Zygalski are considered the some of the greatest feats in reverse-engineering of the 20th century.



Figure 1.1: *Enigma I* reverse engineered by Marian Adam Rejewski

Today we are getting ever closer to reverse-engineering the cell. The fields of synthetic and systems biology are beginning resemble engineering disciplines; genetic engineering is becoming more precise, high-throughput in vitro experiments are performed by robots and measurements of many desired observables can be obtained with high spatio-temporal resolution. Advances in micro-fabrication [4] and in-vitro reconstitutive methods [5] have allowed biologists isolate pathways and mechanisms to a level of mathematical and computational tractability [6]. This section outlines the scientific paradigms in these fields, their methods and limitations, and finally what this thesis will attempt to contribute.

1.1 Bottom-up and Top-down Biology

Fortunately for biologists copies of the target of their reverse-engineering attempts are available all around us. Less fortunate is the fact that most attempts at deconstructing the cell end in loss of function and destruction of the individual components. This restriction motivates system biologists

to manipulate environmental signals *in vivo* and build mechanistic models from correlations between signals and responses [7]. Models focus on relationships between macroscopic variables where the underlying mechanisms are not known.

In vitro reconstitutive methods aim to isolate minimal mechanisms from the complexity of the whole organism in order to unpick the relative importance of microscopic details [5]. In situations where purified proteins and crystal structures are available these methods can quite accurately characterise the kinetics of proteins. Relating these parameters to the *in vivo* context however may not be relevant, as too much of the complexity may have been stripped away.

From a theoretical modelling perspective, it is important to choose a time-scale and space-scale that is relevant to the problem. If one is interested in tissue dynamics, attempting to model DNA conformations within each cell will render the problem intractable. As George Booth aptly put “*most models are wrong but some are useful*” so the role of theoretical descriptions in these settings is not necessarily to describe the way reality *is* but serve as tools to bridge the non-intuitive gap between bottom-up and top-down approaches. Where intuition fails is where the *in silico* hypothesis testing playground becomes most valuable [6].

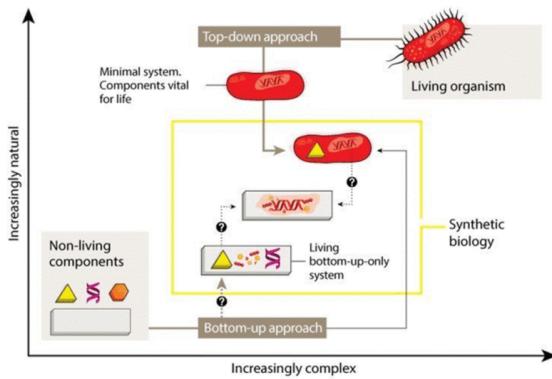


Figure 1.2: Top-down and bottom-up modelling methods

1.2 Design, Build, Test, Learn

Systems and synthetic biology have historically made progress through a process of brute force trial and error. This usually involves the interaction of many custom-made parts that are iteratively optimised by human intervention. A trend first observed in the 1980s known as *Eroom's law* reveals that discoveries in biotechnology are becoming slower and more expensive over time, despite improvements in technology [8]. This problem is compounded by the ongoing reproducibility crisis [9]. In the effort to transform methods used in academia and industry to become more systematic and predictable, a standard for the Design–Build–Test–Learn cycle has emerged – shown in figure 1.3.

This workflow has now been established as a paradigm with some aspects that have been automated by liquid handling robots, bioreactor environments and image processing pipelines. However, humans in the loop and custom moving parts still persist. The challenge in automating these processes lies in defining a programming language that has a sufficiently high level of abstraction for transparent implementation while allowing for low-level customisation [10].

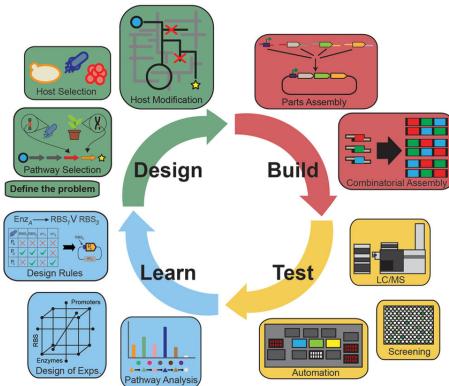


Figure 1.3: Design-Build-Test-Learn cycle from synthetic biology

This thesis will focus on modelling and inference using systems of differential and partial differential equations, which fits into the Design–Learn part of the cycle. Differential equations occupy a small subset of possible modelling tools, however they are amongst the most popular due to their ease of use formulation and simulation. This ease of access creates a zoo of models in literature making it difficult to identify the key ingredients that distinguish different models. Furthermore the relationship between multiple plausible *and* in-plausible hypotheses is rarely investigated. This motivates desire for an automated Design–Learn pipeline – outlined in Section 1.4 – which can generate and catalogue models in a transparent manner while producing insights for the Build stage. This pipeline is applied in an experiment-theory collaboration described in Chapter 3.

1.3 State of the Art

1.3.1 Inference of Differential Equations

Following the initial literature on smooth and match estimators [11] – which overcome the bottleneck having to integrate a proposed hypothesis for every parameter update – a plethora of methods for the inference of differential equations became available [12, 13]. The essence of these methods is to estimate the derivatives of the data rather than integrate the model and simultaneously estimate the qualitative and quantitative behaviours.

In most biological experiments batch-to-batch variations decrease the certainty with which it is possible to quantify a behaviour. Whether or not a parameter is identifiable, redundant or sloppy have become key questions in biology [14, 15, 16].

1.3.2 Model Reduction and Classification

Sloppiness and sensitivity analysis have been extensively used in the search for reduced models. Linear mappings between models that preserve stoichiometry and reactant structure were investigated [17, 18] and computational tools based on partition-refinement were released [19]. Structural similarity between reaction networks can be revealed by such mappings, elucidating the functional aspects of complex networks in terms of simpler networks. The aim of the Design–Learn pipeline is to extend this framework to nonlinear mappings with an emphasis on geometry rather than kinetics. Most inference techniques attempt to match geometry and kinetics simultaneously in an attempt to obtain a quantitative model. This thesis emphasises that geometry alone should be prioritised in order to obtain qualitatively equivalent models. Furthermore recent results in pattern formation theory [20] do not depend on kinetics at all, only geometry.

1.4 Design–Learn Pipeline

This section outlines the proposal for a design–learn pipeline for the purposes of model reduction and system design. Suppose experimental collaborators have provided us with time-course gene expression data \mathcal{D} , which could be taken via time-lapse microscopy of cells growing on microfluidic plates, optical density measurements from microtiter plate assays or temporal snapshots of flow cytometry measurements. On the other hand one may want to specify a behaviour \mathcal{H} which may have not yet been observed. This would be specified with top-down constraints, i.e. there exist oscillations of a fixed frequency or a region of bistability.

1.4.1 Experimental Design Loop

The experimentalists may want to know whether the data collected could result in a model of predictive power without mechanistic knowledge of the underlying biochemistry. Moreover it would be desirable gain insights in parameter regimes in the vicinity of observations, without having to wait for the theorists to produce a refined model. Such real-time insights would guide data collection protocols, optimising the amount of information gained while keeping the number of experiments performed to a minimum. Often data is noisy and at worst case contradictory; these issues must also be exposed.

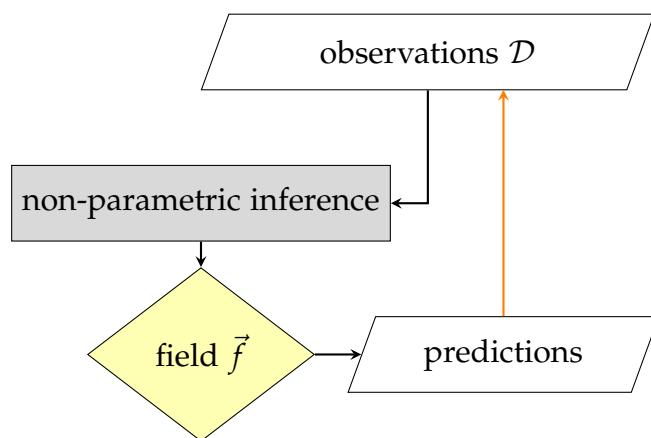


Figure 1.4: Workflow loop for optimal experimental design, without mechanistic model

Figure 1.4 outlines the workflow for optimal experimental design. The term *non-parametric* defines the procedures that have prioritised functional generality and flexibility over mechanistic insights gained from the values of the parameters and shapes of the mathematical forms. Neural networks and Gaussian process regressors are examples of non-parametric estimation procedures, which produce an estimate of the field \vec{f} that predicts gene expression rates at given expression levels. Based on these predictions, the experimentalist may proceed to collect data in the most informative parameter regions, which would in turn more accurately estimate \vec{f} .

1.4.2 Hypothesis Scoring and Model Refinement Loop

Often predictions from field \vec{f} are not enough. Models constructed with feasible biophysical assumptions $\vec{h}(\vec{\theta})$ have the potential to extrapolate predictions and give concrete biophysical meanings to each parameter $\vec{\theta}$. This way the experimentalist knows exactly which modification to the system they must make in order to achieve a desired behaviour. More often than not it is also unclear whether the model and its assumptions are reasonable, which brings us to the desire to score our hypotheses. For increased accuracy and efficiency [21] the mechanistic model $\vec{h}(\vec{\theta})$ is inferred against non-parametric estimate \vec{f} rather than the data \mathcal{D} directly. Furthermore as discussed in Section 1.3 the aim is to optimise geometry rather than kinetics. Alternatively one may construct $\vec{h}(\vec{\theta})$ to cover a whole class of models, such as those that satisfy mass-action. The expectation is that most of the parameters would be zero but some would be informative. One can obtain a distribution of optimal parameters $\rho(\vec{\theta})$ by running multiple optimisations. From this distribution one may construct alternative hypotheses and update $\vec{h}(\vec{\theta})$. By iterating this procedure one would identify the minimal model within the model class that explains the data. This process is known as model refinement or reduction.

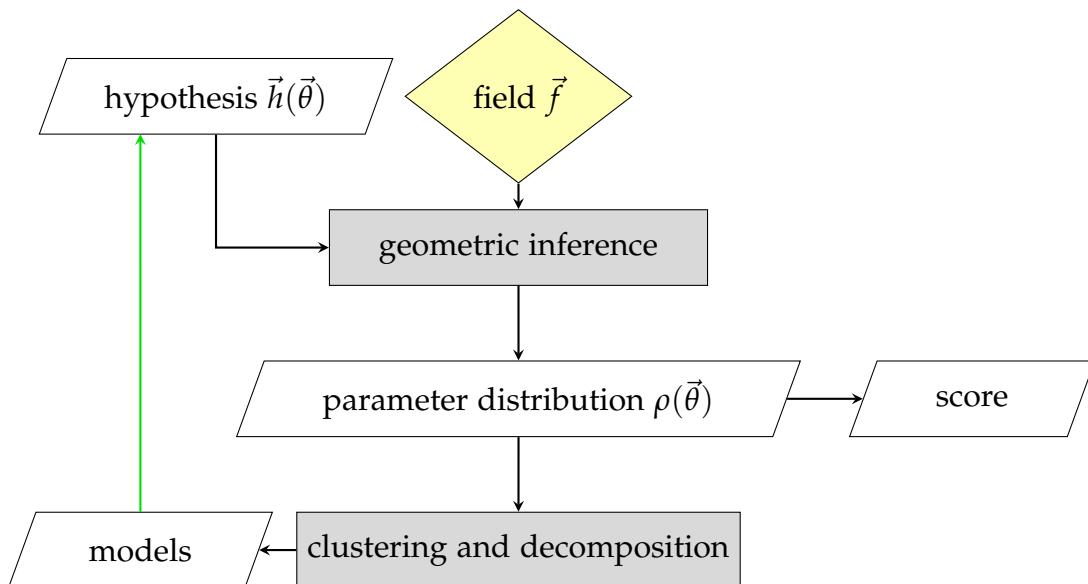


Figure 1.5: Overview of hypothesis scoring pipeline and model refinement loop

1.4.3 System Design

Suppose now that refined mechanistic models of the form $\vec{h}(\vec{\theta})$ have been obtained using the model refinement loop. These could be a library of known parts that have been individually characterised, but never combined to form a larger system. The experimentalists would like to create a system with a specified behaviour \mathcal{H} and would like to know which parts to combine and which modifications to make. The model refinement loop can be used with the design as an input.

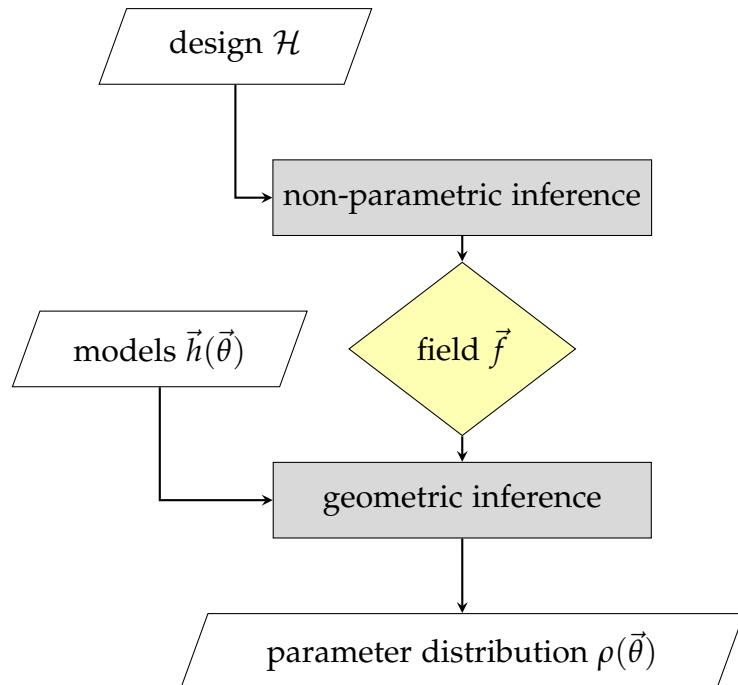


Figure 1.6: Overview of system design pipeline

Chapter 2

Theoretical Background



the passion of friendship will soon blossom into a righteous power

Bolero of Fire — Ocarina of Time

This Chapter lays out a background bifurcation theory and machine learning methods relevant to cell biology and flow cytometry. We establish the connection between the concept of phenotypes and bifurcations in Section 2.1 and lay out assumptions and definitions that are used throughout other chapters. Section 2.2 follows up with concrete applications of differential equations in cell biology and will prepare the reader for the incorporated synthetic biology publication in Chapter 3.

The problem of inferring phenotypes from data is defined in Section 2.3 with a survey relevant machine learning methods. This section discusses the pre-processing techniques to extract bifurcations from raw data, as has been done in Chapter 3, and used as a starting point in the incorporated machine learning publication in Chapter 4. Section 2.4 gives the reader a background in flow cytometry and how the machine learning methods discussed in previous sections are used by immunologists to identify immune cell phenotypes. This section lays the foundations for the interactive exploration tool presented in Chapter 5.

2.1 Describing Phenotypes with Bifurcations

A phenotype is a qualitative state or behaviour of an organism that can be described by several quantitative features. Before considering the additional complexity that comes with biology, let us first consider a simple everyday metaphor: light bulbs. Light bulbs come in various combinations of quantitative features θ which could be its shape, colour, materials and circuit design. The purpose of a light bulb is to fulfil a single function: increase in brightness as a function of voltage p , pushing the qualitative state of the bulb from *off* to *on*. Changes to bulb shape affect neither its function nor its states. Changes in color affect the quality of the *on* state but not the *off* state. Changes to circuit design and materials may change or even break the bulb's function. Fluorescent bulbs, for example, only have two possible brightness states in response to changes in voltage p , while incandescent bulbs have a continuous brightness response.

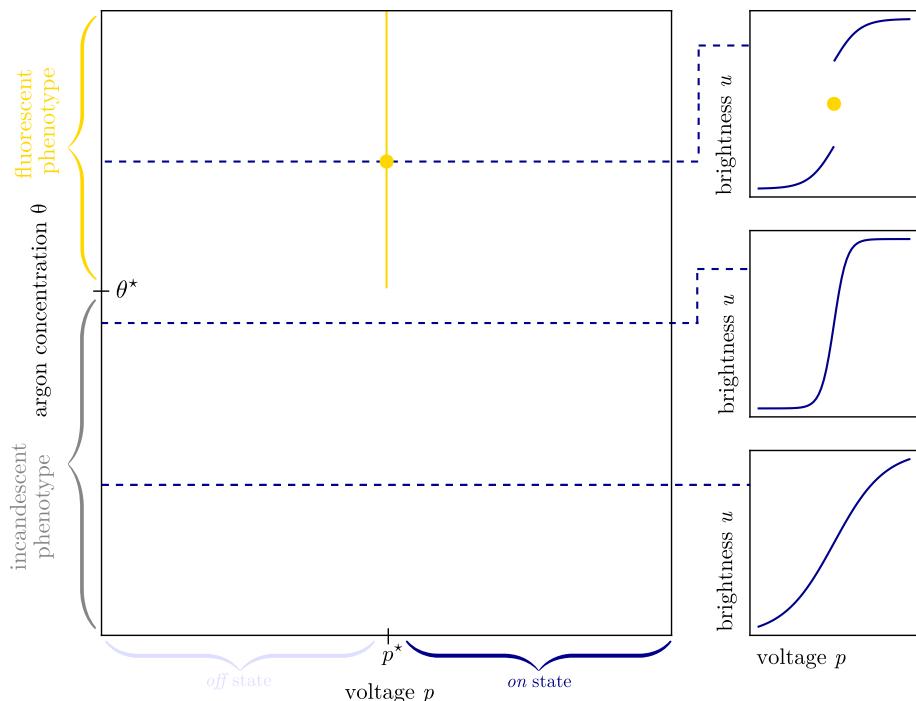


Figure 2.1: a) The bifurcations (shown in gold) at critical voltage p^* above argon concentration θ^* give rise to sudden jumps in brightness (right-hand panels). θ^* separates two bulb phenotypes and p^* separates the *on* and *off* states.

The fluorescent and incandescent bulbs can be considered as two different phenotypes distinguished by the quality of their response to voltage changes. Different colour bulbs can also be considered phenotypes, distinguished by the quality of their *on* state, rather than their response to voltage. Bifurcation theory allows us to describe the transitions between qualitative states and can be leveraged to distinguish and organise phenotypes. In this context a bifurcation becomes a punctuation that either *distinguishes between phenotypes* or *distinguishes between behaviours within a phenotype*.

Suppose we inserted a component into our light bulb that changed parameter θ in such a way so that we can change between the discrete response of the fluorescent bulb and the continuous response of the incandescent bulb. Perhaps θ could be the concentration of argon gas; the bulb would have to be wired to behave like an incandescent bulb at low concentrations. We could collect brightness u as a function of voltage p and gas concentration θ produce something similar to that shown in Figure 2.1. The bifurcations at critical voltage p^* above argon concentration θ^* give rise to sudden jumps in brightness. The bifurcations separate the *on* and *off* states of the bulb and are only present in the fluorescent phenotype. The two phenotypes lie either side of the onset of bifurcations at concentration θ^* . We can see from Figure 2.1 how knowing the locations of bifurcations allows us to organise qualitative behaviours and hence phenotypes of the bulb.

In the biological context, we can consider changes in θ as changes in the organism genotype that may or may not lead to a new phenotype. The idea of describing phenotypes in this way has been done before by Waddington []. His epigenetic landscape is a metaphor for how changes gene regulation, in our case represented by changes in θ , determines the fate of cells. He imagines the cell as a marble, rolling down a series of forking valleys representing a differentiation cascade, eventually settling

in its final phenotype. Let us adopt this metaphor for organism behaviours in response to a controlled condition (Figure 2.2). Changing the control condition places the marble at different points in the valley and each fork in the valley corresponds to different available behaviours to the organism. Changing the parameters θ changes the topology of the valleys potentially giving rise to new behaviours and therefore new phenotypes. Due to the robustness [] and fragility [] of organisms we expect most changes in θ to either kill the organism or do nothing at all. The emerging picture suggests that the route between phenotypes is a carefully created sequence of changes in θ .

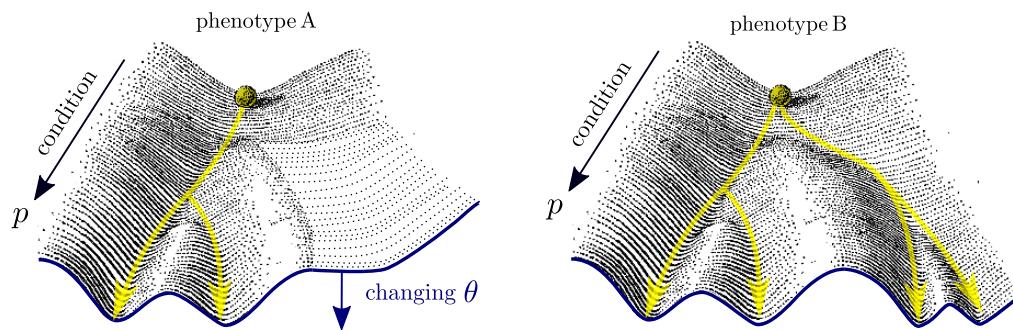


Figure 2.2: Waddington landscapes representing the set of available behaviours for two phenotypes in response to a controlled condition p . Phenotypes are related by changes in landscape topology via changes in θ

In order to enumerate the phenotype distribution in the the high-dimensional parameter space θ for a given organism we must construct a model. By parametrising such a model $F_\theta(u, p)$ by θ , we can relate the states of the organism u to experimentally controlled conditions p . Ideally a subset of u can be observed experimentally so that we may observe bifurcations in the data, as demonstrated in the right-hand panel of Figure 2.1. In the following section we shall go through a class of models that can leverage bifurcation theory.

2.1.1 Differential Equation Models

For the purposes of this thesis we will assume that the behaviour of the organism under study can be cast in terms of differential equations in N states u , M parameters θ and P control conditions p . For now we shall state the general class of models and follow up with concrete biological examples as we explore different types of bifurcations. Throughout the thesis we will consider models of the form

$$\frac{du}{dt} = F_\theta(u, p) \quad \text{where} \quad \begin{cases} F : \mathbb{R}^{N+P} \rightarrow \mathbb{R}^N \\ \theta \in \mathbb{R}^M, u \in \mathbb{R}^N, p \in \mathbb{R}^P \end{cases} \quad (2.1)$$

The total derivative $\frac{du}{dt}$ gives us the rate of change of the states with respect to time t and all variables have been vectorised with the appropriate dimension. In principle the right-hand-side F_θ can be arbitrarily complicated, containing spatial derivatives or even machine learning models such as neural networks. We shall see later when such models become relevant in biomedical modelling.

2.1.1.1 Trajectories & Field Geometry

In principle once equations (2.1) have been written down they can be integrated to obtain trajectories $u(t)$ for various initial conditions $u(t')$. We can write the solution down formally as

$$u(t) = u(t') + \int_{t'}^t F_\theta(u(s), p) ds \quad (2.2)$$

Here the integral reveals that in order to determine where the state is at time t we need to sum all the contributions of the function F_θ from the initial time t' all the way up to final time t . The function F_θ depends on the state u and must be updated with the integration variable s . This calculation can be interpreted, as shown in Figure 2.3, as choosing an initial point $u' = u(t')$ in a vector field F_θ and following the field lines until time t at which the final point $u = u(t)$ has been reached.

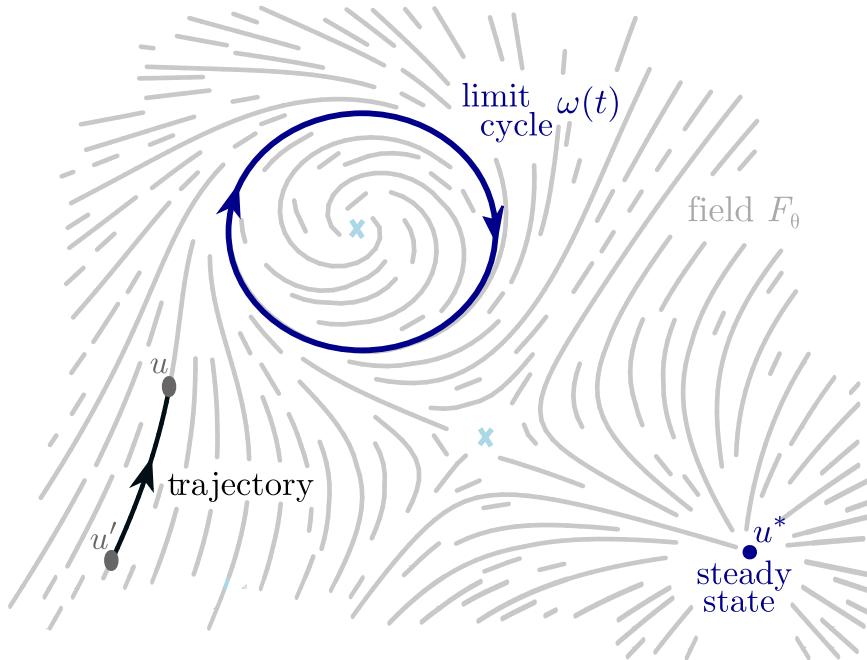


Figure 2.3: Illustration of a trajectory between $u = u(t)$ and $u' = u(t')$ over finite time $t - t'$ following the field lines of F_θ . In this example field lines either point towards steady state u^* or a limit cycle $\omega(t)$. There are two unstable fixed points marked by crosses: a saddle point separating the basins of attraction and an unstable focus enclosed by the limit cycle.

By considering the geometry of the field F_θ in state space we can determine the fate of sets of trajectories, which are ultimately pulled towards *dynamical attractors*. Such attractors can be static steady states u^* or dynamic like the limit cycle $\omega(t)$ illustrated in Figure 2.3. *Dynamical attractors* create basins of attraction defined as regions of state space in which trajectories are pulled towards the same stable structure. These basins must be separated by unstable structures, such as the saddle point marked by a cross in Figure 2.3, which define a boundary between the basins called the separatrix. Note that the direction of the field F_θ and not its magnitude $|F_\theta|$ determines the fate of a trajectories.

When equations (2.1) describe the behaviours of an organism, the attractors determine the set of qualitative behaviours available to the organism of genotype θ , whilst experiencing experimental conditions p . If changes to the genotype θ change the number, type or stability of the at-

tractors then we will observe new qualitative behaviours and hence a new phenotype. If changes to conditions p lead to changes in the state space geometry, this is interpreted as a different behavioural state available to the same phenotype. We can see therefore how casting a biomedical problem into the language of differential equations, allows us to characterise phenotypic traits with the geometry of attractors in state space.

2.1.1.2 The Jacobian Matrix

In order to quantify the geometry of a local patch of state space u we can imagine F_θ as a velocity field of water and place a tiny blob of ink surrounding the location u . The so called Jacobian matrix $\frac{\partial F_\theta}{\partial u}$ of partial derivatives at the location u transforms the basis vectors $\hat{u}(t)$ defining the blob over a short period of time Δt as depicted in Figure 2.4. The transformation is

$$\hat{u}(t + \Delta t) = \left(\mathbb{1} + \frac{\partial F_\theta}{\partial u} \Delta t \right) \hat{u}(t) \quad \text{where} \quad \hat{u} \in \mathbb{R}^N \quad \frac{\partial F_\theta}{\partial u} \in \mathbb{R}^{N \times N} \quad (2.3)$$

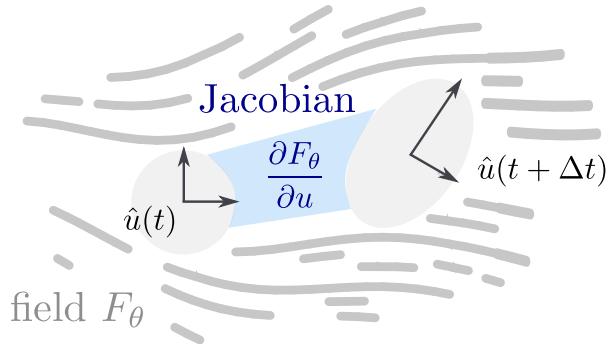


Figure 2.4: Illustration how the Jacobian matrix $\frac{\partial F_\theta}{\partial u}$ transforms the basis vectors $\hat{u}(t)$ defining a small patch of initial conditions for a short period of time Δt . The eigenvalues λ and eigenvectors \hat{v}_λ define the deformation of the patch

where $\mathbb{1}$ is the identity matrix. The eigenvalues λ and eigenvectors \hat{v}_λ of the Jacobian reveal to us the magnitudes and directions of the stretching or squeezing of the blob. We can diagonalise the Jacobian to obtain the eigenvalues and eigenvectors at any state space location u to quantify the local geometry of the field. The eigenvalue equation is written as

$$\lambda \hat{v}_\lambda = \frac{\partial F_\theta}{\partial u} \hat{v}_\lambda \quad \text{where} \quad \hat{v}_\lambda \in \mathbb{R}^N : |\hat{v}_\lambda| = 1 \quad (2.4)$$

where the eigenvectors are normalised to have unit magnitude. We can take the limit $\Delta t \rightarrow 0$ of equation (2.3) to obtain a first order matrix differential equation for the evolution of basis vectors $\hat{u}(t)$ in any patch u

$$\frac{d\hat{u}}{dt} = \frac{\partial F_\theta}{\partial u} \hat{u} \quad (2.5)$$

The coefficients given by the Jacobian matrix are time-independent and therefore the general solution can be written as a matrix exponential

$$\hat{u}(t) = e^{\frac{\partial F_\theta}{\partial u}(t-t')} \hat{u}(t') \quad (2.6)$$

where $\hat{u}(t')$ are the basis vectors that define the initial shape of the blob centred on location u . If we let the basis for the initial blob have axes parallel to the eigenvectors \hat{v}_λ of the Jacobian, then the evolution any vector $\sigma(t)$ within the blob expressed in its basis becomes

$$\sigma(t) = \sum_{\lambda} \hat{v}_{\lambda} \sigma_{\lambda}(t') e^{\lambda(t-t')} \quad (2.7)$$

where $\sigma_{\lambda}(t')$ are the initial components of the vector in the basis of the blob. This equation reveals explicitly how the sign of real parts to eigenvalues $\Re \lambda$ determine the exponential growth or shrinkage of the blob in the direction \hat{v}_λ . The imaginary parts $\Im \lambda$ determine the magnitude of rotations of the blob. Note that since the Jacobian matrix is real and eigenvalues and eigenvectors appear in conjugate pairs and therefore the overall evolution (2.7) yields real transformations of blob vectors $\sigma(t)$.

A field F_θ is called *degenerate* in regions of state space u where adjacent field lines do not change and hence do not cause shape changes to the blob in one or more directions. This means that one or more of the eigenvalues

of the Jacobian vanish and have associated eigenvectors \hat{T}_θ that obey

$$\frac{\partial F_\theta}{\partial u} \hat{T}_\theta(u) = 0 \quad (2.8)$$

This equation constructs vectors $\hat{T}_\theta(u)$ that span the degenerate subspace at field location u . Note that a degenerate subspace may not always exist. In linear algebra we will see that this equation is useful not for constructing tangent spaces,

Equations (2.17) suggest that small perturbations away from stable steady states u^* decay exponentially with respect to time. However if the state space flow F_θ at the steady state u^* is degenerate, then at least one of the eigenvalues of the Jacobian λ vanish. Expanding the exponentials for a subset of λ that are close to zero yields

$$\delta u^*(t) \approx \sum_{\lambda \approx 0} \alpha_\lambda \hat{v}_\lambda (1 + \lambda(t - t') + \mathcal{O}(\lambda^2)) + \sum_{\lambda' \neq \lambda} \alpha_{\lambda'} v_{\lambda'} e^{\lambda'(t-t')} \quad (2.9)$$

This gives rise to relaxation dynamics that are polynomial in the vicinity of the degenerate point along the degenerate directions defined by eigenvectors \hat{v}_λ . Perturbations along $v_{\lambda'}$ still decay exponentially and will vanish before the polynomial terms. By considering changes in the magnitude of perturbations $|\delta u^*(t)| - |\delta u^*(t')|$ and dropping higher order terms $\mathcal{O}(\lambda^2)$, and substituting an expression for the eigenvalues we arrive at

$$\frac{|\delta u^*(t)| - |\delta u^*(t')|}{t - t'} \sim \sum_{\lambda \approx 0} \hat{v}_\lambda \cdot \left. \frac{\partial F_\theta}{\partial u} \right|_{F_\theta(u,p)=0} \cdot \hat{v}_\lambda \quad \text{where} \quad \left. \frac{\partial F_\theta}{\partial u} \right. \cdot \hat{T}_\theta = 0 \quad (2.10)$$

2.1.1.3 Lyapunov Spectrum Analysis

We can determine the stability of attractors by looking at the geometry of field lines in its vicinity. Do they converge or diverge? Does the field rotate? We could answer these questions tracking the evolution of small

perturbations δu between adjacent trajectories. Let us begin by substituting the perturbed trajectory $u + \delta u$ into equations (2.1)

$$\frac{d}{dt}(u + \delta u) = F_\theta(u + \delta u, p) \quad (2.11)$$

After Taylor expanding the field F_θ and recognising that equations (2.1) lead to a cancellation in the terms involving the unperturbed trajectory u we arrive at

$$\frac{d}{dt}(\delta u) = \left. \frac{\partial F_\theta}{\partial u} \right|_{u=u(t)} \delta u + \mathcal{O}(\delta u^2) \quad (2.12)$$

where $\frac{\partial F_\theta}{\partial u}$ is the Jacobian of the field and additional terms $\mathcal{O}(\delta u^2)$ involve taking higher order derivatives of the field F_θ . Choosing a perturbation sufficiently small such that we can ignore higher order terms yields a first order homogenous ordinary matrix differential equation of the form $\dot{\delta u}(t) \approx J(t)\delta u(t)$ where $J(t)$ is a time-varying Jacobian. These equations can be solved formally with time-ordered matrix exponentiation []

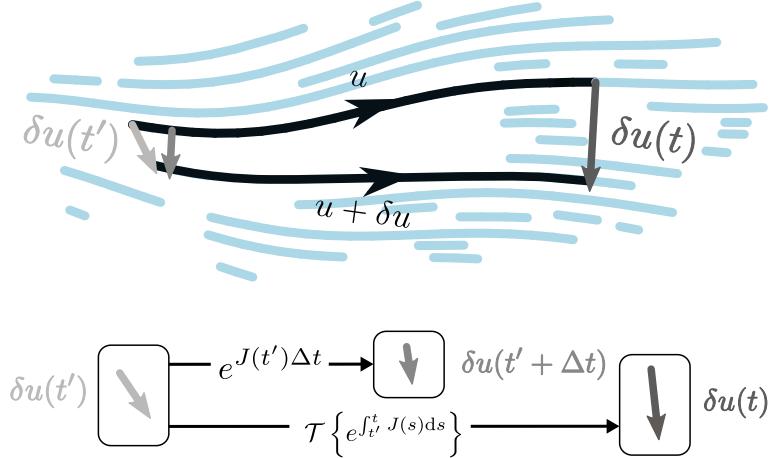


Figure 2.5: Illustration how the time ordering operator $\mathcal{T}\left\{e^{\int_{t'}^t J(s)ds}\right\}$ propagates the separation $\delta u(t)$ between adjacent trajectories u and $u + \delta u$ by repeated application of the exponentiated Jacobian $e^{J(t)\Delta t}$ evaluated along unperturbed trajectory $u(t)$

$$\delta u(t) \approx \mathcal{T} \left\{ e^{\int_{t'}^t J(s) ds} \right\} \delta u(t') \quad \text{where} \quad J(t) := \frac{\partial F_\theta}{\partial u} \Big|_{u=u(t)} \quad (2.13)$$

The time-ordering defines an ordered product of matrix exponentials

$$\mathcal{T} \left\{ e^{\int_{t'}^t J(s) ds} \right\} := \lim_{\Delta t \rightarrow 0} \left(e^{J(t)\Delta t} e^{J(t-\Delta t)\Delta t} \dots e^{J(t'+\Delta t)\Delta t} e^{J(t')\Delta t} \right) \quad (2.14)$$

which can be calculated by repeated exponentiation of the Jacobian $e^{J(t)\Delta t}$ evaluated at different times along the unperturbed trajectory $u(t)$. Although this expression may look complicated it is just another matrix whose eigenvalues reveal whether the trajectories diverge $|\delta u(t \rightarrow \infty)| \rightarrow \infty$ or converge $|\delta u(t \rightarrow \infty)| \rightarrow 0$. The rates of convergence or divergence are known as Lyapunov exponents, and can be defined in a local patch of state space t and finite time Δt as the eigenvalues of

$$\Lambda(u, \Delta t) := \frac{1}{\Delta t} \log \left(\mathcal{T} \left\{ e^{\int_t^{t+\Delta t} J(s) ds} \right\} \right) \quad (2.15)$$

Note that to calculate this matrix we need to choose an initial point in state space $u = u(t)$, obtain the trajectory up to $u(t + \Delta t)$ and evaluate the Jacobian along the trajectory in the time ordered integral.

2.1.1.4 Linear Stability Analysis

Thusfar we dealt with stability in arbitrary patches of state space u and asked questions about local flows. What happens when we look at flows in the vicinity of fixed points u^* that satisfy the equation

$$\frac{du}{dt} \Big|_{u=u^*} = F_\theta(u^*, p) = 0 \quad (2.16)$$

In this case, since there is no time-dependence in the equation (2.13), the time-ordered integral simplifies yielding a matrix exponential equation for

perturbations from fixed points $\delta u^*(t)$

$$\delta u^*(t) \approx e^{(t-t') \frac{\partial F_\theta}{\partial u} \Big|_{u=u^*}} \delta u^*(t') \quad (2.17)$$

The Jacobian evaluated at fixed points $\frac{\partial F_\theta}{\partial u} \Big|_{u=u^*}$ can be expressed in terms of its eigenvectors \hat{v}_λ and eigenvalues λ and therefore

$$\delta u^*(t) \approx \sum_{\lambda} \alpha_{\lambda} \hat{v}_{\lambda} e^{\lambda(t-t')} \quad \text{where} \quad \lambda \hat{v}_{\lambda} = \hat{v}_{\lambda} \cdot \frac{\partial F_\theta}{\partial u} \Big|_{u=u^*} \quad (2.18)$$

with α_{λ} given by initial condition $\delta u^*(t')$

We can see that the sign of the real parts to eigenvalues $\Re \lambda$ determine the exponential growth or shrinkage of fixed point perturbations $\delta u^*(t)$ in the direction parallel to the corresponding eigenvector \hat{v}_{λ} . The imaginary parts $\Im \lambda$ determine the frequency of oscillations around the fixed point. If at least one of the eigenvalues has a positive real part $\Re \lambda$ the field pushes trajectories away from the fixed point in direction \hat{v}_{λ} which means the fixed point is unstable.

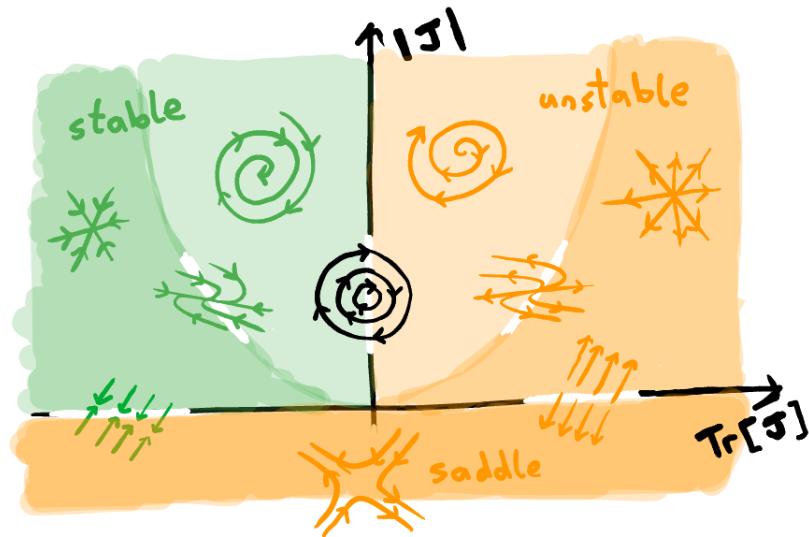


Figure 2.6: Classification of **stable** and **unstable** fixed points for a general two dimensional Jacobian in terms of trace $\text{Tr}[J]$ and determinant $|J|$

For a two dimensional system we can express the two eigenvalues of

the 2×2 Jacobian in terms of its trace $\text{Tr}[J]$ and determinant $|J|$. Figure 2.6 reveals the regions of stability and field flows for an arbitrary two dimensional system in terms of the Jacobian trace and determinant. We can see now how changing the model can lead to changes in sign in one of the eigenvalues of the Jacobian. This leads to a change in stability of a fixed point which is in fact type of bifurcation. This diagram gives us an intuition for the types of fixed points that are possible, but not all of them can be characterised in this way.

Degenerate fixed points may lead to vanishing of the Jacobian and hence require investigation of higher-order derivates of the field F_θ which were dropped in the approximation (2.13). In contrast to what is shown in Figure 2.6 it is not possible to know whether a limit cycle is present by looking at the eigenvalues of its corresponding fixed point alone.

In the vicinity of non-degenerate fixed points u^* the Lyapunov spectrum (2.15) is simply given by the Jacobian eigenvalues. Another scenario in which this simplification occurs is for small time intervals $\Delta t \rightarrow 0$ when evaluating the spectrum. Therefore

$$\lim_{\Delta t \rightarrow 0} \Lambda(u) = \Lambda(u^*) = \frac{\partial F_\theta}{\partial u} \quad (2.19)$$

While the rest of this thesis does not deal with chaotic systems, we would like make the reader aware of equation (2.15) which can be used as a substitute for the Jacobian for characterising chaotic attractors. Note that a suitably chosen timescale Δt to match the onset of chaos must be chosen in this case.

2.1.1.5 Limit Cycle Analysis

2.1.2 Bifurcation Analysis

2.1.2.1 Dynamics near Bifurcations

revealing that the speed of perturbation decay to stable steady states u^* scales as the eigenvalues do when the degenerate point is approached.

in a certain part of the state space. the relaxation dynamics become polynomial with respect

from equations One of the indicators for the onset of a bifurcation is the

2.1.2.2 Static Non-degenerate Bifurcations

Varying the continuos parameters $\sigma[r]$ moves the nullclines and may result in the creation or annihilation of fixed points of different classes. While an individual fixed point may change location and local phase space flow, it cannot change class without involving another fixed point.

These are called bifurcations and also fall into various categories. Figure 2.7 illustrates some of the possible one parameter supercritical bifurcations; subcritical cases are obtained by permuting stabilities of fixed points. Note the hysteresis loop in the saddle-node bifurcation.

Saddle-node Bifurcation

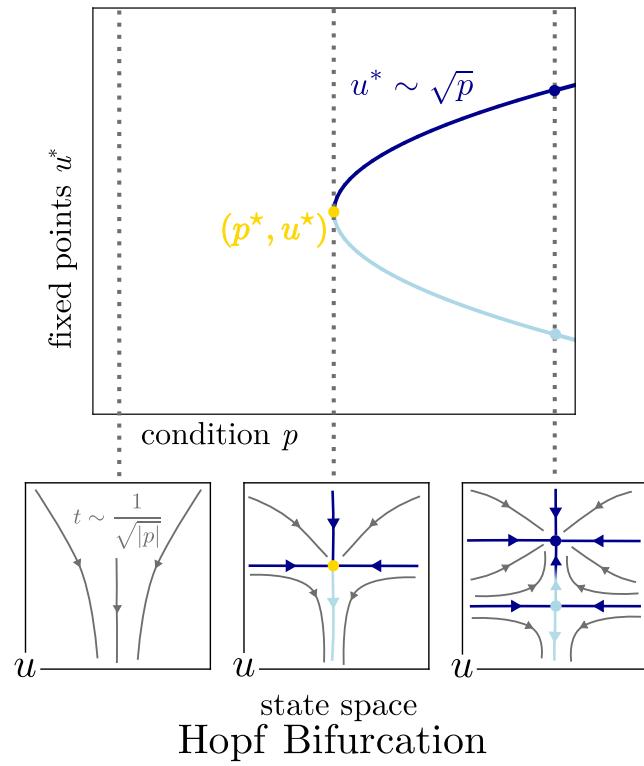
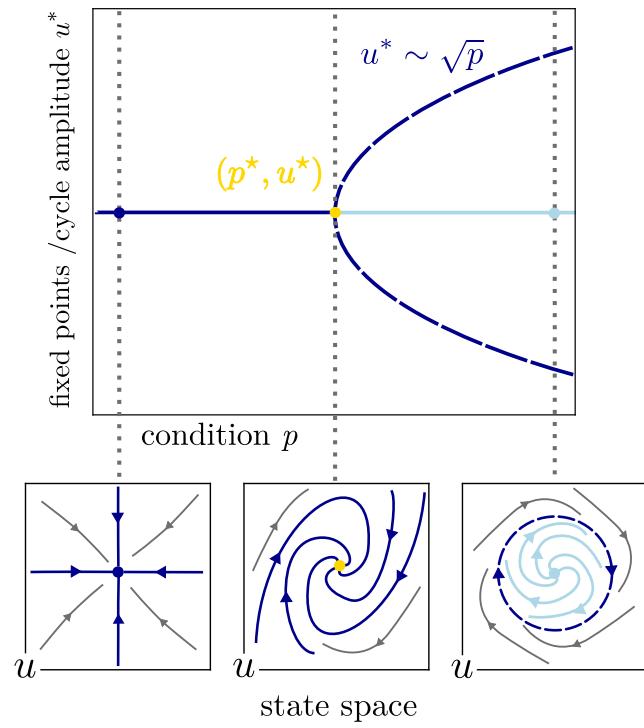
state space
Hopf Bifurcation

Figure 2.7: All bifurcations

2.1.2.3 Dynamic Bifurcations

Another category of bifurcations involves limit cycles, which emerge from fixed points where the linearised Jacobian eigenvalues have no real part. Limit cycles have circulating field flow as shown in Figure 2.6 along the $\text{Tr}[\mathbf{J}] = 0, |\mathbf{J}| > 0$ axis.

Note how oscillations emerge at small amplitudes in the Hopf bifurcation, whereas the large amplitude oscillations may instantly emerge in an infinite-period or cyclic-fold bifurcation. In Figure 2.8 the shaded regions represent the peaks and troughs of the oscillations.

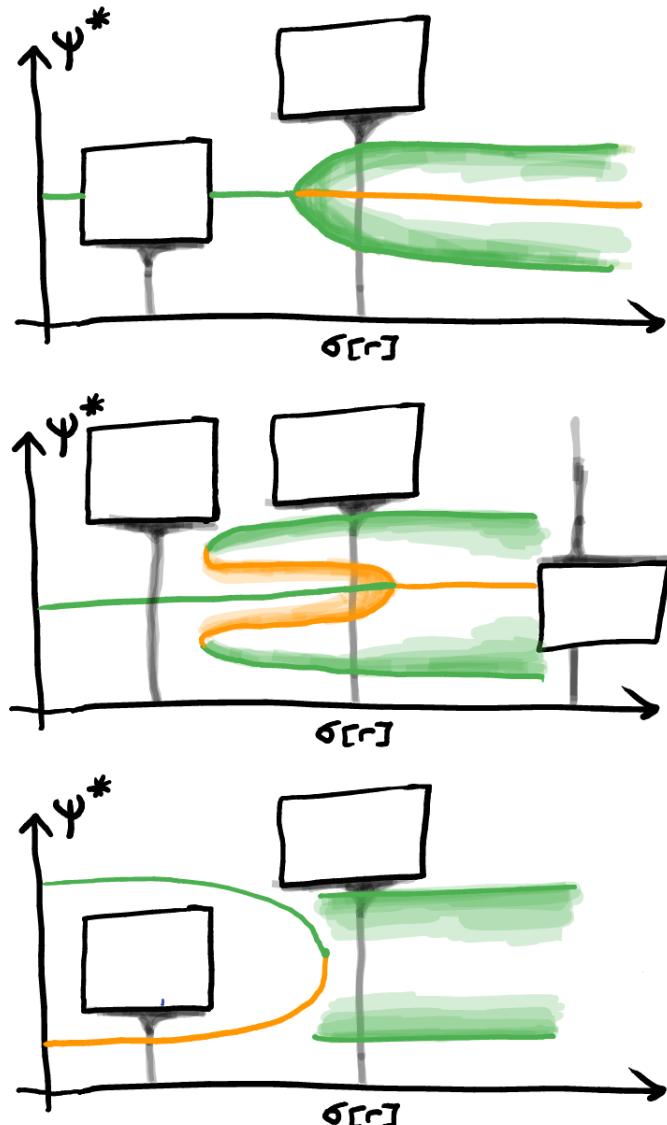


Figure 2.8: Hopf, Cyclic-fold and Infinite-Period bifurcation diagram showing **stable** and **unstable** fixed points and limit cycles as a function of parameter $\sigma[r]$. Insets show nullclines [tbc]

2.1.2.4 Parameter Continuation

2.1.3 Spatially Extended Systems

If F_θ contains spatial derivatives, additional conditions on u with respect to spatial variables $x \in \mathbb{R}^D$, where D is the dimension of the space considered, must be supplied to give a unique solution $u(x, t)$.

2.1.3.1 Partial Differential Equations

When studying spatio-temporal dynamics of organisms, a popular choice is to incorporate a diffusive term that models the exchange of matter between spatial locations

$$\partial_t \psi = \Gamma \omega(\psi | \Gamma) \quad (2.20)$$

A bifurcation diagram tracks changes in attractors such as fixed points u^* and limit cycles $\omega(t)$ in response to changes in control conditions p . In the next section we will see how fixed points and limit cycles are characterised with linear stability analysis

this can be done using In this thesis we extensively use the library `DifferentialEquations.jl` [22] for numerically solving differential equations and `BifurcationKit.jl` [23] for calculating bifurcation diagrams.

Partial observation can be modelled explicitly with latent variables [] or the explicit modelling of system and measurement device as is done in control theory []. For the scope of this thesis we will let the modeller have complete freedom on how to choose relevant state variables $u \in \mathbb{R}^N$ and write down their own set of differential equations, with unknown parameters $\theta \in \mathbb{R}^M$ of the form

2.2 Applications in Cell Biology

2.2.1 Genetic Switches & Phenotype Boundaries

These equations can originate from mean-field approximations of the chemical master equation [] representing biochemical reaction networks []. This typically yields models with a large number of parameters M and states N . The equations could have also undergone a series of quasi-steady state approximations [] yielding hill-functions. Various other model reduction methods exist [] in an attempt to reduce the number of state variables N and parameters M .

2.2.2 Populations of Neurons

2.2.2.1 Neural Differential Equations & Gaussian Processes

While principled derivation from microscopic rules exist, often the the number of variables N participating in the mechanism under study becomes intractable. When the desire for predictive performance out-weights the desire for mechanistic explanations, the states u are chosen to be the observable subset for which we have data, and the right-hand side F_θ can be a machine learning model such as a neural network or gaussian process []. In this case the number of parameters becomes much larger than the number of states $M \gg N$. A recent breakthrough in back-propagation through differential equation solvers [] has made these models tractable in optimisation routines, and popularised them in literature [].

2.2.2.2 Stochastic Differential Equations

2.2.3 Self-organised Patterns & Development

In this section we would explore the geometrisation of phase space, lyapunov exponents and universality. Perhaps a discussion on phase transitions and the relation to Landau-Ginzberg approaches is required. Maybe also periodic orbit theory? Depends how useful it is.

2.2.3.1 Reaction-Diffusion

Here we first introduce diffusion macroscopically by simply adding the laplacian to mean field equation (2.20). We introduce the turning bifurcation and show how linear stability analysis is insufficient to capture pattern formation and rich inhomogeneous steady states. A promising approach may be geometrisation of the moving local equilibria [20].

2.3 Phenotype Inference with Machine Learning

- Preface on SciML community and gray-box modelling

2.3.1 Classification and Regression

- Logistic regression
- Nearest neighbour methods
- Mixture Models

2.3.2 Dimensionality Reduction Methods

- UMAP and Tsne
- Sloppy parameters and identifyability

2.3.3 Clustering Methods

2.4 Applications in Flow Cytometry

2.4.1 Immunophenotyping Panels

2.4.2 Inferring Bifurcations from Data

In sections 4.2 – 4.6 we assume that the bifurcation data \mathcal{D} with respect to the experimental control condition $p \in \mathbb{R}$ is known or already extracted from raw data. We can break down the extraction process into steps: steady state inference, field inference and fixed-point inference; each discussed separately in its own subsection. These steps are sequenced together into two possible routes towards bifurcation data \mathcal{D} as depicted in Figure 2.9.

At best we could have state trajectory data like single cell trajectories extracted via segmentation and tracking in fluorescence microscopy movies (for example the data acquired using the CellASIC platform in Figure 3.8c). Having the state as a function of time $u(t)$, sampled at sufficiently broad initial conditions $u(0)$, enables the *non-parametric* estimation of field geometry $F(u)$. We can then extract important state-space structures, such as stable and unstable fixed points, their eigenvalues λ , from the field geometry $F(u)$ at different values of the control condition p , and hence identify possible bifurcations.

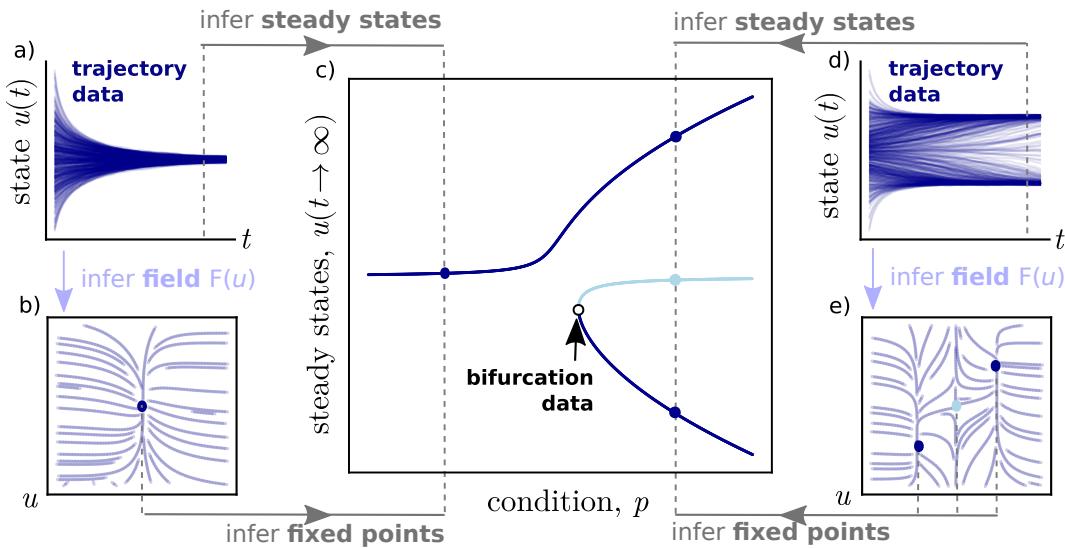


Figure 2.9: Bifurcation point extraction along the condition p via two possible routes: steady state inference (a,d \rightarrow c) and fixed point inference (a,d \rightarrow b,e \rightarrow c). The fixed point inference route requires the geometry of the field $F(u)$ at different values of the condition p and yields additional information about unstable fixed points.

Data on steady states $u(t \rightarrow \infty)$ are more abundant and typically found in flow cytometry measurements. It is possible to infer the steady state manifold directly from such data (Figure 2.9 a,d \rightarrow c) but we would not be able to extract information on unstable fixed points or eigenvalues λ of state-space structures. Since the criteria for bifurcations are not directly accessible, indirect measures (such as the population separation that can be seen in Figure A.3) must be used to extract bifurcations.

2.4.2.1 Bifurcations via Steady State Inference

Suppose we would like to detect a cusp bifurcation and limit points in a cell population with respect to two experimental control conditions $p, p' \in \mathbb{R}$. We can set up a serial dilution along the columns for p and along the rows for p' in two 96-well plates. Cells allowed to grow in exponential phase in a finite concentration of either $p \ll p'$ or $p \gg p'$. Let us call this the *priming* stage of the protocol as shown in Figure 2.10a, resulting in two cell populations: p -primed cells and p' -primed cells. The priming concentrations must be chosen sufficiently high so that the resultant population states lie

either side of the cusp. The two populations are transferred into separate 96-well plates containing the dilutions of p, p' ; we call this the *conditioning* stage in Figure 2.10a.

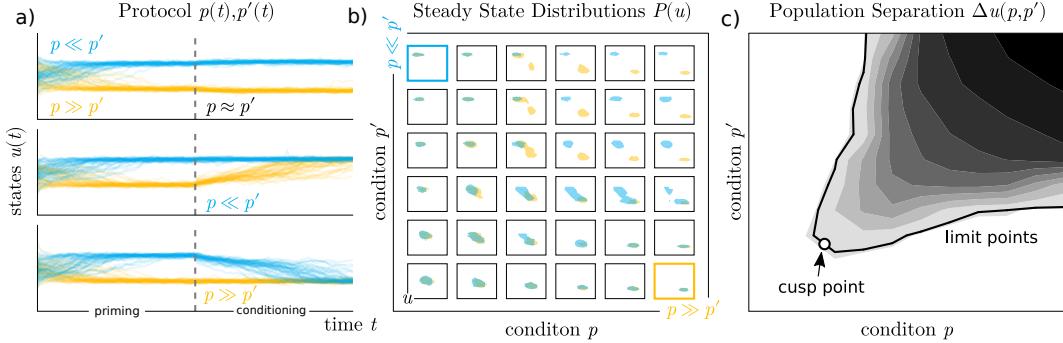


Figure 2.10: Protocols (a) for extracting limit points with respect to conditions p, p' . Steady state distributions (b). Similarity measure (c)

The cells are then transferred into a flow cytometer, gated for live singlets and processed with relevant compensation and auto-fluorescent normalisation, which would produce the population distributions $P(u)$ for the p -primed cells and $Q(u)$ for p' -primed cells in each well. By overlaying distributions $P(u), Q(u)$ for each well, a figure similar to Figure 2.10b (or Figure A.3) can be produced. Finally, the limit points can be defined by a level set of distribution similarity measure $D(P||Q)$ in the p, p' -plane as shown in Figure 2.10c. This similarity measure could be Kullback–Leibler divergence or something as simple as the distance between distribution medians. The level set must be some small positive amount ϵ above zero, picking out the onset of dissimilarity between steady state distributions $P(u)$ and $Q(u)$, and hence the onset hysteresis. We can define the limit points as

$$\mathcal{D} = \{(p, p') : D(P||Q) = \epsilon, \epsilon > 0\} \quad (2.21)$$

This approach may break down if multi-modal distributions exist in the data. This would be the case if something happened to prevent a sub-population of cells to switch from one state another other. Reasons for this

could include too much cell burden or not enough time given for cells to reach a steady state. In this case, quantifying the efficiency of switching from either side of the cusp could be a quantity of interest.

This approach would not work for extracting dynamic bifurcations such as *Hopf* since only steady state information is available. Furthermore, the accuracy of this method is subject to noise amplitude. This method works well for cases where changes in the number of stable steady states can be resolved in the measured steady state distributions.

- quantify efficiency of switching?

2.4.2.2 Non-parameteric Vector Field Inference

In this early days of this thesis, we investigated whether it was possible to transform the time-domain data into state-space. This approach, and related works, are discussed in this section and can in principle be used with the microfluidic fluorescence microscopy data for parameter inference.

Consider we are given K cell trajectories $\mathcal{D}_1, \mathcal{D}_2 \dots \mathcal{D}_K$, each containing N noisy observations of the state of the cell. Let the cell state be represented by state vector $u(t) \in \mathbb{R}^N$ which is hypothesized to obey a set of ordinary differential equations of the form (2.1). Instead of integrating the equations (2.1) we would find an estimate for the derivative of the trajectories \hat{f} .

This is known as the *smoothing* step [11] should be done using unsupervised methods, for example with Gaussian Process Regressors [24] as shown in Figure 2.11. This requires the inversion of an $K' \times K'$ data matrix where $K' := \sum_k |\mathcal{D}_k|$ is the total number of trajectory data points. This has a computational complexity K'^3 which is only tractable with sparse datasets.

Let the region $\partial\mathcal{D}$ be a boundary defined by the Delaunay tessellation of the input data. Let us define the estimate \hat{f} only within the region $\partial\mathcal{D}$ so that there are no extrapolation artefacts. For the Gaussian Process approach

the estimate would be

$$\hat{f}(u) \sim \mathcal{N}(\mu(u), \Sigma(u)) \quad \text{for } u \in \partial\mathcal{D} \quad (2.22)$$

where at any given state u the field estimate \hat{f} is generated by Gaussian distributions of mean vector μ and covariance matrix Σ . Solving for these requires a choice of matrix-valued kernel function $\mathbf{K}(u, v)$ which encodes our knowledge about the local structure of the field. Sophisticated kernels for learning vector fields exist [25] for decomposing fields in conservative and solenoidal components, which aid in localising fixed points and cycles.

The simplest choice of kernel assumes the components are independent and have a finite correlation length γ , such as Gaussian radial basis functions. Here \mathbf{I} is the identity matrix and the hyperparameter γ has to be optimised.

$$\mathbf{K}(\vec{u}, \vec{v}) = \mathbf{I} e^{-\gamma|\vec{u}-\vec{v}|^2} \quad (2.23)$$

The second step is called *matching* where the estimated field \hat{f} is used as an optimisation target against some parametrised function F_θ with unknown parameters θ .

In our setting we would like to match the geometry of the field but not its magnitude; in this sense we are focusing on the qualitative aspects of the dynamics of a set of differential equations, rather than the quantitative dynamics or kinetics. This could be achieved with the following objective function

$$\mathcal{L}(\theta | \mathcal{D}) := e^{-\frac{\hat{f} \cdot F_\theta}{|\hat{f}| |F_\theta|}} \quad (2.24)$$

where the cost is minimal when the data derivative \hat{f} and the parametrised model F_θ point in the same direction and maximal when they point in opposing directions.

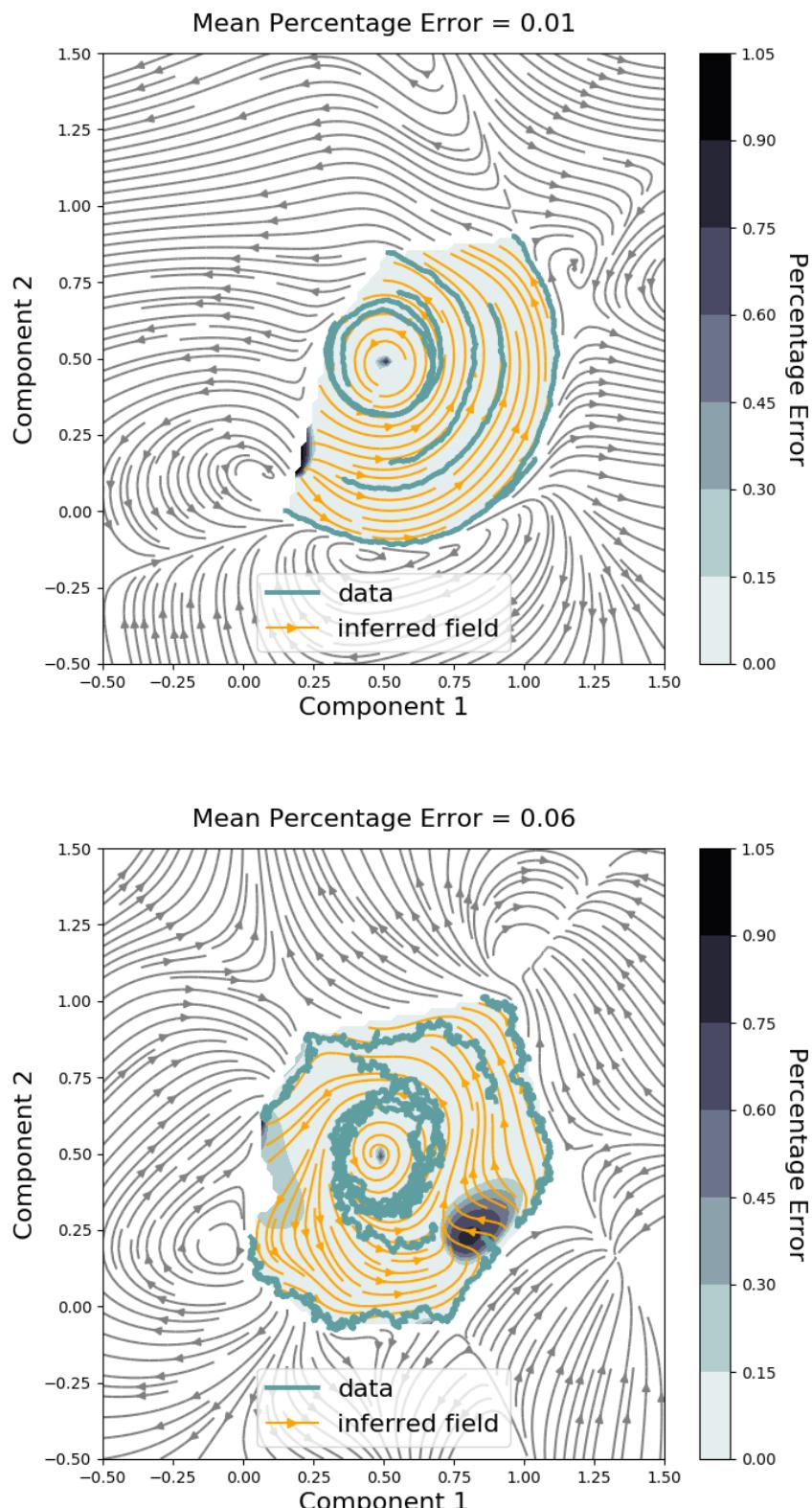


Figure 2.11: Gaussian process regressors estimating derivative of the trajectories \hat{f} from example trajectory datasets $\mathcal{D}_1 \dots \mathcal{D}_K$ with varying signal to noise ratios. Interpolation error E is shown as a heatmap; extrapolation fails

Although we are getting close to focusing on qualitative features of a model, this objective function is still sensitive to the locations and shapes of fixed point and limit cycles. What if we cared about even higher-level features such as the number of fixed points? Or perhaps whether a system oscillates or not? This is where the language of bifurcation theory described in Chapter 2 is optimally suited for this task, but first we need to discuss how to set up experiments to detect bifurcations from flow cytometry data.

- Vector field estimates too noisy to get bifurcation points?
- Divergence and the stability of fixed points
- Curl and limit cycles. How does this relate to hopf example 2.11
- Do we need a bistable example?

The accuracy of the cell trajectories is limited by cell segmentation and tracking algorithms. Initial investigations into this approach also suggested that trajectories need to be of sufficient temporal resolution and sampled from a wide variety of initial conditions. Such data is not widely available and ultimately we decided to focus on a method that could be used with a well-known workhorse in biomedical research: flow cytometry.

2.4.2.3 Fixed Point Inference

2.4.3 Multiomics Integration

Chapter 3

Interpretation of Morphogen Gradients by a Bistable Circuit

memory of younger days

Ocarina of Time

3.1 Preface

1. Keep focus on developmental biology
2. Revised supplement as this chapter

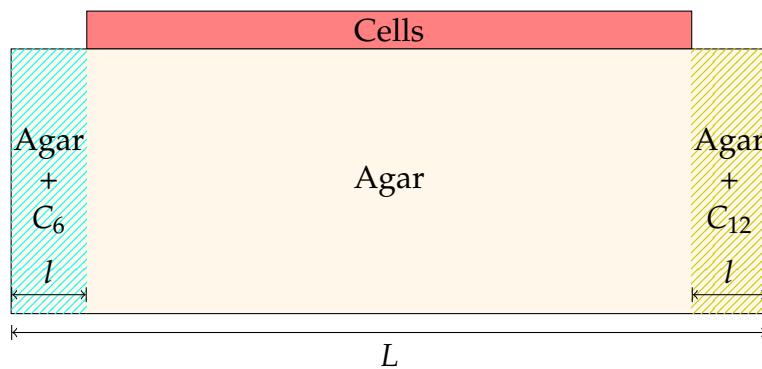


Figure 3.1: Geometry of opposing gradients experiment

This section outlines how the Design—Learn pipeline may help achieve a specific aim in a typical collaboration between theory, computation and experiment. The aim of this project is to reconstitute and con-

trol minimal self-organisation mechanisms which are believed play crucial roles in developmental biology. To this end *E. Coli* has been genetically engineered to produce orthogonal responses to two different input signals — henceforth this organism will be referred to as the *double exclusive reporter* circuit [26]. The colony of reporters serve as a reduced model for a multi-cellular organism during embryonic stages of development. While patterns with sharp boundaries have successfully been realised, producing Turing instabilities remains challenging as the system needs to be such that patterns develop before the colony reaches stationary phase. The role of theory and computation in this project is to help identify the parameter regimes that produce controllable and self-organised patterns.

3.1.1 Problem Statement & Context

During the development of any organism a hierarchy of self-organisation takes place that leads to the breaking of symmetry from a spherical cluster of undifferentiated cells to the formation of organ segments and limbs. This process is known as morphogenesis. figure 3.2 shows an example in which *Hox* gene expression patterns in the body segments of *Drosophila* can drastically affect its development.

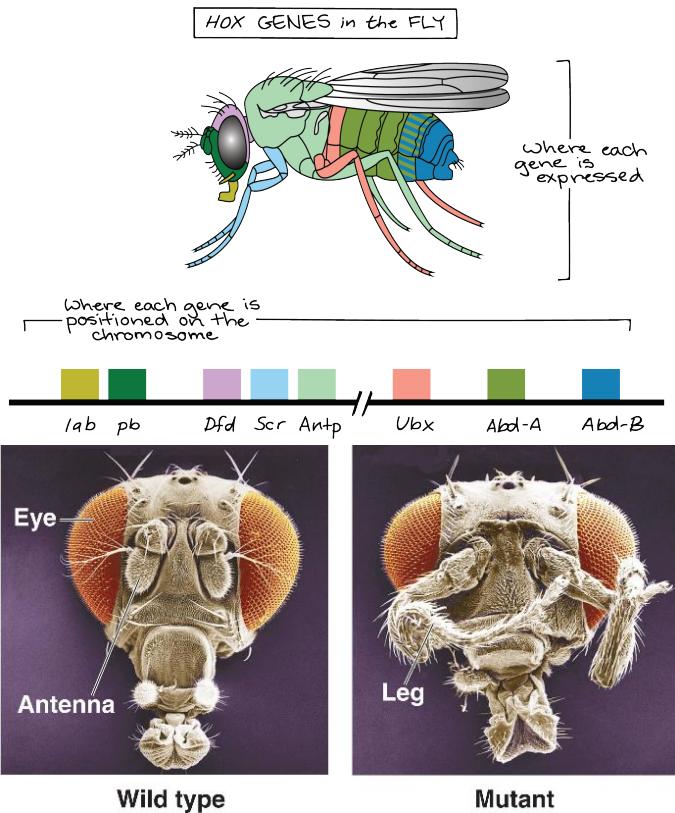


Figure 3.2: Top: *Hox* gene expression patterns in body segments of drosophila.
Bottom: Mutation where legs grow in-place of antenna

3.1.1.1 Morphogen-driven Patterns

One of the central questions in developmental biology is how positional information is sensed by a population of cells and how sharp gene expression boundaries between populations are maintained for robust organ and body segment development. The French Flag model [27] proposes that cells have a threshold response to external signalling molecules – henceforth referred to as morphogens – which pre-pattern the organism from anterior to posterior and laterally. Some examples of morphogens include Wingless, Decapentaplegic and Sonic Hedgehog. figure 3.3 show the *Gap* expression patterns that partition the Drosophila embryo into segments which are later differentiated by *Hox* genes.

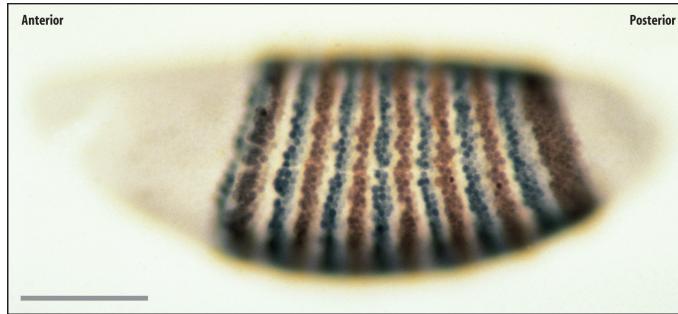


Figure 3.3: Expression patterns of pair-rule *Gap* genes in *Drosophila* embryo []

3.1.1.2 Self-organised Patterns

How are morphogen gradients set up and maintained? How can they be robust against changes in size and geometry? A canonical example of self-organisation in bacteria is the quorum sensing system [28]. Each cell secretes a signalling molecule resulting in the total concentration being proportional to the population density. This signal induces adaptive responses in metabolic and mobility in the whole colony. A long standing mathematical hypothesis that Turing patterns underlie self-organisation in cell populations. Recent literature suggests both morphogen-driven and Turing patterning mechanisms play a role in development [29]



Figure 3.4: Pigment patterns hypothesised to be generated by Turing mechanism

3.1.1.3 Learnings & Limitations

This section outlines the results on determining the bi-stability landscape for the *double-exclusive reporter*, and how its geometry helps determine the conditions for moving and stationary sharp boundary formation. The design of the device is summarised in figure 3.5. Each wiggly line represents a promoter. Incoming solid lines to a vertex represents the formation of a complex, which lead to the induction of a promoter. Outgoing solid lines represent the transcription-translation of genes downstream of a promoter. Arrows that end in t-junctions represent the repression of a pathway.

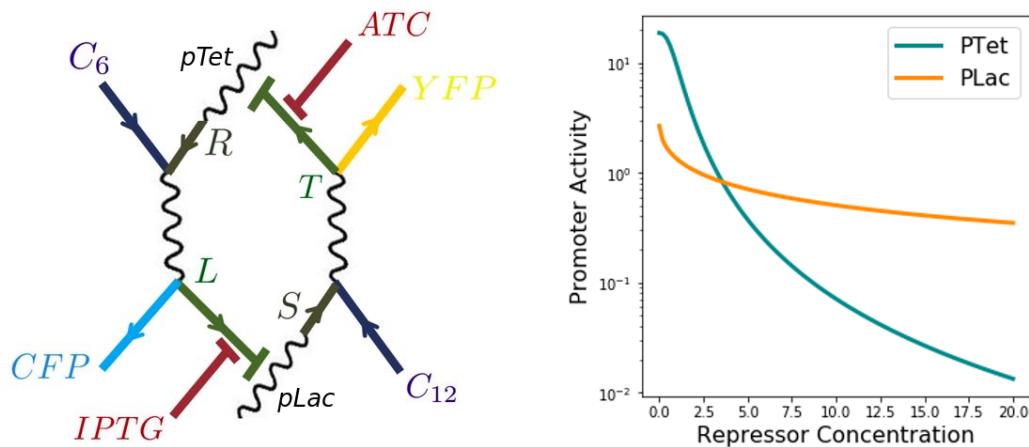


Figure 3.5: Left: Diagram of double exclusive reporter showing wiring between promoters and complexes Right: Repression of $pTet$ and $pLac$ promoters via $TetR$ and $LacI$ respectively

The observable chemicals, labelled YFP and CFP , are proportional to their upstream promoter activities. The experimentally controllable chemicals are quorum sensing molecules C_6 , C_{12} and de-repressors $ATC, IPTG$. All other chemicals cannot be directly observed. The double-negative feedback loop leads to bi-stability in the observable signals, which is key to producing sharp gene expression boundaries.

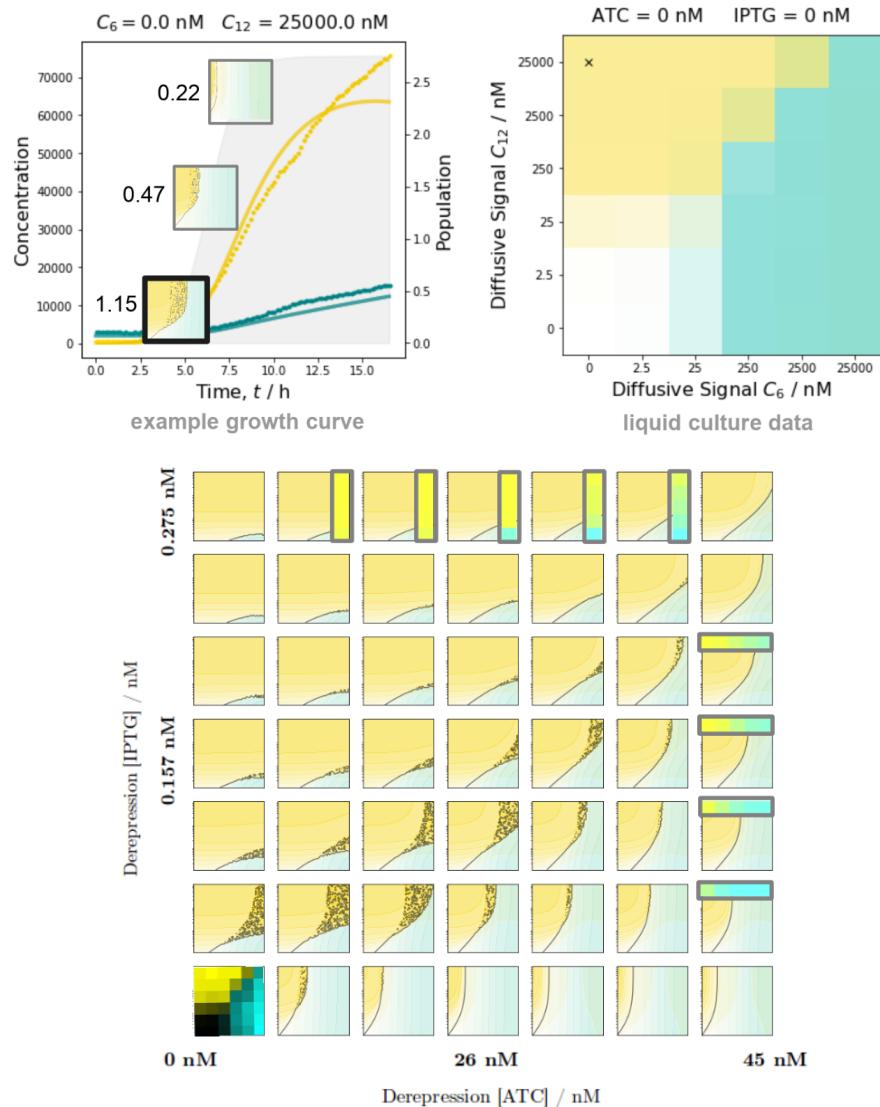
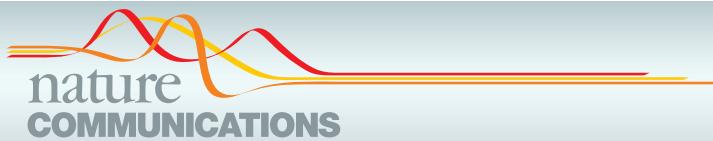


Figure 3.6: Top Left: Optical density measurements obtained for one well. Top Right: Final fluorescence values at 16h across different C_6 , C_{12} concentrations. Bottom: Steady state predictions between data regions.

3.1.2 Contributions

Contributions for this work are as follows



ARTICLE

<https://doi.org/10.1038/s41467-020-19098-w>

OPEN

Interpretation of morphogen gradients by a synthetic bistable circuit

Paul K. Grant¹✉, Gregory Szep^{1,2,9}, Om Patange¹ , Jacob Halatek¹ , Valerie Coppard¹, Attila Csikász-Nagy¹ , Jim Haseloff¹ , James C. W. Locke^{1,3,6}, Neil Dalchau¹ & Andrew Phillips¹

During development, cells gain positional information through the interpretation of dynamic morphogen gradients. A proposed mechanism for interpreting opposing morphogen gradients is mutual inhibition of downstream transcription factors, but isolating the role of this specific motif within a natural network remains a challenge. Here, we engineer a synthetic morphogen-induced mutual inhibition circuit in *E. coli* populations and show that mutual inhibition alone is sufficient to produce stable domains of gene expression in response to dynamic morphogen gradients, provided the spatial average of the morphogens falls within the region of bistability at the single cell level. When we add sender devices, the resulting patterning circuit produces theoretically predicted self-organised gene expression domains in response to a single gradient. We develop computational models of our synthetic circuits parameterised to timecourse fluorescence data, providing both a theoretical and experimental framework for engineering morphogen-induced spatial patterning in cell populations.

¹ Microsoft Research, 21 Station Road, Cambridge CB1 2FB, UK. ² Randall Centre for Cell and Molecular Biophysics, King's College London, London WC2R 2LS, UK. ³ Sainsbury Laboratory, University of Cambridge, Cambridge CB2 1LR, UK. ⁴ Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest 1083, Hungary. ⁵ Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK. ⁶ Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK. ⁷ Present address: Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA. ⁸ Present address: Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ⁹ These authors contributed equally: Gregory Szep, Om Patange. ✉email: grant.paul@microsoft.com; andrew.phillips@microsoft.com

ARTICLE

NATURE COMMUNICATIONS | <https://doi.org/10.1038/s41467-020-19098-w>

The positional information solution to the French flag problem, in which cells compute their spatial position by responding to the concentration of a morphogen in a gradient¹, provides crucial insight into how patterns of gene expression form in a developing organism. The simplest formulation of this model – concentration thresholds leading directly to gene expression states – requires a static morphogen gradient to produce a stable pattern of gene expression^{2,3}. However, quantitative measurements in developing embryos reveal that morphogen gradients are both dynamic and transient^{4,5}, and genetic perturbations reveal that pattern formation is robust to changes in morphogen concentration^{6–8}. A gene regulatory network topology of mutual inhibition downstream of antiparallel morphogen gradients^{9–12} (Fig. 1a) has been proposed to robustly interpret dynamic gradients (Fig. 1b). However, while certain features of this topology are common to a number of developmental contexts such as the early *Drosophila* embryo and the vertebrate neural tube (reviewed in³), demonstrating how this network functions and whether it is indeed sufficient remains a challenge, due to the complexities of the different biological contexts in which it operates. Recent work in synthetic biology has proven the utility of building multicellular patterning circuits both for understanding development and for learning engineering principles^{13–18}.

Here we show that the mutual inhibition motif¹⁹ is sufficient to produce stable domains of gene expression in response to dynamic and transient morphogen gradients. By taking a synthetic biology approach^{20–23} we have built a morphogen-induced

mutual inhibition circuit from scratch that acts in isolation in *E. coli* and used it to investigate the conditions under which patterning occurs. We have also added morphogen production to the core circuit to create a reaction-diffusion patterning system that responds to a single gradient by producing two domains of gene expression with a self-organized boundary. The experimental control and precise measurement afforded by a synthetic biology framework allowed us to understand the behaviour of these patterning mechanisms at a quantitative level in the context of a mathematical model parameterized against data, and to uncover general design principles for engineering multicellular systems.

Results

Engineering mutual exclusivity. To investigate whether a simple mutual inhibition network topology can interpret dynamic gradients, we built a synthetic Exclusive Receiver circuit (Fig. 1c), based on a previous Receiver circuit design (pR33S175²⁴) that responds to two homoserine lactone (HSL) input signals, 3O-C6-HSL (C6) and 3O-C12-HSL (C12) with fluorescent protein outputs. We engineered mutual inhibition by introducing genes encoding TetR, expressed bicistronically with eYFP, and LacI, expressed bicistronically with eCFP. In addition, the C12-binding receiver protein LasR was expressed under the control of a LacI-repressible promoter, while the C6-binding receiver protein LuxR was expressed under the control of a TetR-repressible promoter. The Exclusive Receiver therefore consists of two signalling pathways that mutually repress each other, such that LasR, eYFP and TetR are expressed in the

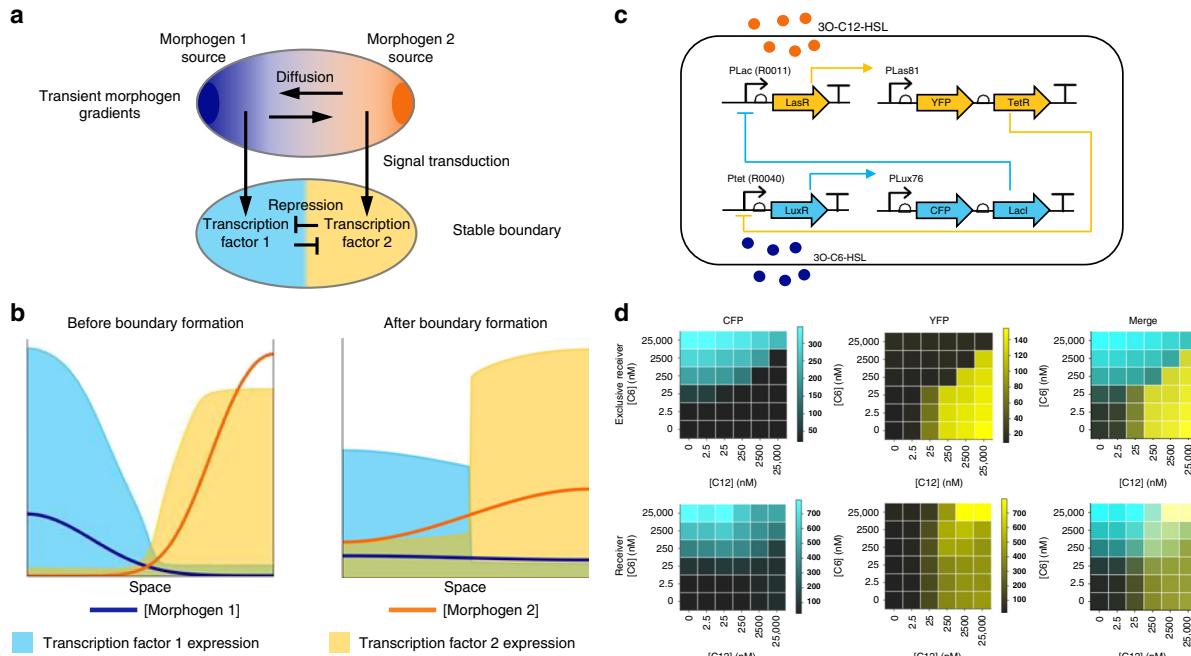


Fig. 1 A synthetic gene circuit for morphogen interpretation. **a** Schematic representation of a developing embryo. Mutual inhibition of transcription factors (cyan and yellow) downstream of antiparallel morphogen gradients (dark blue and orange) has been hypothesized to produce mutually exclusive domains of gene expression. **b** Morphogen gradients can be dynamic and transient, yet sharp, stable boundaries are observed between domains of gene expression. **c** A diagram of the Exclusive Receiver circuit. When 3O-C12-HSL (C12) levels are high, C12 binds to LasR, activating the expression of YFP and TetR, which represses the expression of LuxR, preventing expression of CFP and LacI. When 3O-C6-HSL (C6) levels are high, C6 binds to LuxR activating expression of CFP and LacI, which represses the expression of LasR, preventing expression of YFP and TetR. **d** Fluorescence output, measured in microplate fluorometer assays, of the Exclusive Receiver (top) and the Receiver (bottom) circuits represented as a ratio of CFP- (left) or YFP- (right) fluorescence to RFP fluorescence during exponential phase³⁸, cultured in the presence of the concentrations of C6 and C12 indicated. Data are representative of $n=3$ biological replicate experiments conducted on different days. See Supplementary Figs. 11 and 12 for replicates. Source data are provided as a Source Data file.

presence of the signal C12, while LuxR, eCFP and LacI are expressed in the presence of the signal C6. Unlike the Receiver circuit, which responds to the presence of both signals by producing both fluorescent proteins, the Exclusive Receiver was designed to respond exclusively to the two signals, by producing either eCFP or eYFP. Mutually exclusive reception will only occur when the repressors (TetR and LacI) are expressed such that the level of repression produced by high concentrations of one HSL is sufficient to repress detectable quantities of the other and low concentrations of HSL produce little repression. We used this mutually exclusive response as a design goal for the Exclusive Receiver circuit. We constructed a range of designs and chose a variant that exhibited a strong difference between CFP- and YFP-expressing states while maintaining a roughly equal sensitivity to the two HSLs (Supplementary Fig. 1).

To characterise the response of the Exclusive Receiver circuit to varying concentrations of C6 and C12, we performed timecourse plate fluorometry assays and calculated the promoter activity from the CFP and YFP channels using a ratiometric method²⁵. The circuit responded to high concentrations of C6 and C12 by producing CFP and YFP, respectively. As intended by the mutual inhibition design, mixtures of both signals resulted in only one fluorescent protein being produced, depending on the relative concentrations of the two signals (Fig. 1d; Supplementary Fig. 2). We confirmed these gene expression states at the single cell level using flow cytometry (Supplementary Fig. 3). In contrast, a Receiver circuit lacking mutual inhibition produced both fluorescent proteins simultaneously when both signals were present (Fig. 1d; Supplementary Figs. 2 and 3). Similar results were obtained when chemical inducers were used to suppress the repressors in the Exclusive Receiver circuit (Supplementary Fig. 4).

Mutual inhibition results in bistability. Central to the ability of mutual repression to produce a robust signal response is the

property of bistability, in which two stable steady states of gene expression are possible²⁰. We first explored this property mathematically by developing a dynamic (ordinary differential equation) model of the Exclusive Receiver circuit. The model is based on one derived for the Receiver circuit²⁴, but incorporates the repressor proteins, TetR and LacI, and their regulation of LuxR/LasR expression (see Supplementary Methods for a complete derivation). We identified parameter values that enabled the model to reproduce timecourse fluorescence data using a previously established inference methodology in which a sequence of parameter inference tasks are applied to models and data for circuits of increasing complexity²⁶ (Supplementary Methods). This enabled us to simplify the identification of parameter values of the Exclusive Receiver model by reusing values of the subset of parameters that also appear in the Receiver model. We then applied numerical continuation methods to our data-constrained model to locate saddle-node bifurcations (see Supplementary Methods), and thus the concentrations of C12 and C6 for which bistability was predicted (Fig. 2a,b, red lines).

To test whether the Exclusive Receiver circuit exhibited hysteresis, a hallmark of bistability, at the concentrations predicted by the model, we first conditioned cells in either C6 or C12 and then exposed them to varied concentrations of both signals. At concentrations that produce bistability, we expected C6-conditioned and C12-conditioned cells to remain in the CFP-expressing and YFP-expressing states, respectively. In contrast, at concentrations that produce monostability, gene expression states would be determined solely by the final concentrations. We measured CFP and YFP expression by flow cytometry (Supplementary Fig. 5). The C6-conditioned cells expressed CFP at a wider range of concentrations, while C12-conditioned cells expressed YFP at a wider range of concentrations (Fig. 1b, Fig. 2a, b). We interpret this history-dependent difference in gene expression to be due to hysteresis. Thus, the region in concentration space in which we observe this difference is the

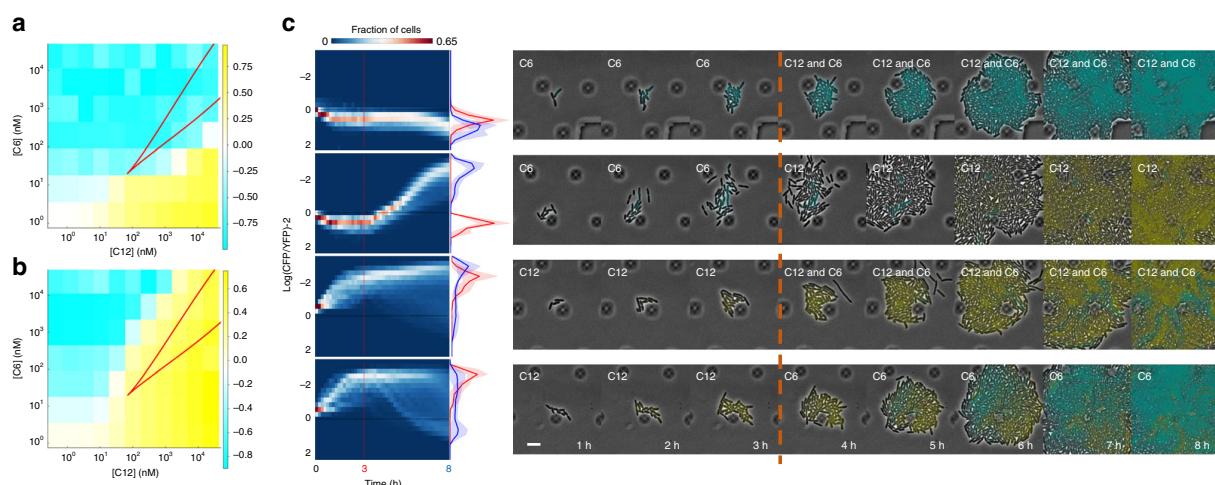


Fig. 2 Mutual inhibition produces bistability. Cells transformed with the Exclusive Receiver circuit were conditioned in either 500 nM C6 (a), or 500 nM C12 (b), and then exposed to the combinations of concentrations of C6 and C12 indicated. Cells were measured using flow cytometry and their normalized CFP minus YFP expressions were plotted. The region of bistability predicted by the parameterized model is the area within the red lines. See Supplementary Fig. 3 for gating strategy for all flow cytometry and Supplementary Fig. 13 for replicates. Source data are provided as a Source Data file. c Microfluidics cultures of cells transformed with Exclusive Receiver circuit in changing combinations of signals. Cells were grown for 3 h in the presence of either 37 nM C6 (rows 1 and 2) or 100 nM C12 (rows 3 and 4). Then media was changed to 100 nM C12 + 37 nM C6 (rows 1 and 3) or 100 nM C12 (row 2) or 37 nM C6 (row 4). Cells were imaged with a frame rate of (1 frame/10 min). Left panels are kymographs of the log-ratio of CFP expression per-cell to YFP expression per-cell, and fraction of cells as a heat map. Histograms represent the populations at 3 h (red) and 8 h (blue). Lines and shaded region represent the mean and standard deviation, respectively, over $n = 4$ biological replicates performed on 4 different days. Right panels are sample montages of cells switching state (rows 2 and 4) or exhibiting bistability (rows 1 and 3); phase contrast and fluorescence channel ranges chosen for display. Scalebar = 6 μ m.

ARTICLE

NATURE COMMUNICATIONS | <https://doi.org/10.1038/s41467-020-19098-w>

region of bistability. This region was slightly larger than that predicted by the model but was qualitatively similar, which suggested that our model captured the essential features of our circuit.

To determine whether individual cells were switching or retaining their gene expression states in response to signal combinations, as suggested by our flow cytometry data, we used microfluidic devices to observe the dynamics of single cells over time. We cultured cells with 37 nM C6 or 100 nM C12, concentrations expected to produce bistability in combination, and then switched to the opposite signal. We found that cells with high fluorescence expression could, indeed, switch to the alternative state (Fig. 2c rows 2 and 4, Supplementary Fig. 6), even when we used the much higher concentration of 1 μ M C6 (both to initially condition cells and to switch their state Supplementary Figs. 7 and 8). When cells conditioned with either C6 or C12 were exposed to both signals, the cells exhibited bistable behaviour, mostly maintaining their initial fluorescence states (Fig. 2c rows 1 and 3, Supplementary Figs. 6–8) although a comparatively small population of C12-conditioned cells do begin expressing CFP in response to mixed signals (Fig. 2c, row 3) indicating some heterogeneity in the population with respect to bistability. We hypothesize that cell to cell differences in gene expression result in differences in the HSL concentration regimes in which the cells exhibit bistability resulting in heterogeneity in switching behaviour. Together, these data indicate that, when signal concentrations gave rise to bistability, the final gene expression state was determined by the history of exposure to signals.

Hysteresis produces stable boundaries. To test how the Exclusive Receiver circuit interpreted dynamic morphogen gradients, we grew cells on filter paper printed with hydrophobic ink, such that growing colonies remained within the square in which they were inoculated but signals could diffuse through the agar beneath the filter^{24,27}. We performed these experiments in agar containing 10 μ M IPTG (1% of the standard induction concentration of 1 mM) to create a regime in which both CFP-dominating and YFP-dominating regions were observed. The addition of IPTG was required to compensate for differences in culture conditions between solid and liquid cultures (see Supplementary Fig. 34 for experiments without IPTG). While the precise mechanism is unknown, differences in culture conditions when switching to solid culture appeared to shift the bistability region such that even very low concentrations of C6 enabled bistability, preventing the appearance of a YFP-dominant regime. This was redressed through addition of IPTG, which partially derepresses LacI, shifting the region to coincide with the region in liquid culture. We cast either C6 or C12 into the agar at each end of the filter paper at varying concentrations and performed timelapse imaging of the fluorescence output of the circuit in response to the dynamic gradients produced by diffusion. When C6 and C12 were provided at concentrations that, if allowed to diffuse to homogeneity would result in 200 nM C6 and 2000 nM C12, (i.e., a 200 nM and 2000 nM spatial average, respectively), a sharp boundary was produced between mutually exclusive domains of CFP and YFP (Fig. 3a). Plotting the point in space at which CFP and YFP expression were equal (see Supplementary Methods) against time revealed that the boundary between domains did not move over time (Fig. 3b). In contrast, when spatial average concentrations of 20 nM C6 and 2000 nM C12 were used, there was an initial production of CFP near the source but it was quickly overwhelmed by YFP production and the point of equal expression moved toward the C6 source (Fig. 3a). These images measured bulk (rather than per-cell) fluorescent protein

expression so highly stable fluorescent proteins would remain detectable even after they stopped being expressed. It was therefore unsurprising that CFP remained detectable in cells close to the C6 source due to protein perdurance. Importantly, however, the CFP intensity did not increase, whereas the level of YFP did increase, indicating that the circuit had switched from CFP expression to YFP expression in these cells. At varying combinations of signal concentrations we classified boundaries (see Supplementary Methods) as static (S), moving (M), or not present (N) (Fig. 3c, Supplementary Fig. 9, Supplementary Movie 1) and observed that static boundaries occurred at signal concentrations similar to those that produced hysteresis (Fig. 2a and b).

To understand this behaviour we performed simulations and analysis of our mathematical model. We plotted the concentration of C6 and C12 experienced by physically separated cells and visualized the changes in concentration that they experienced due to diffusion (Fig. 3d, Supplementary Movie 2). We observed that cells at different points in space all converge to the spatial average concentration, but they do so by taking different paths, resulting in differences in CFP and YFP expression. Points closest to the C6 source begin in the monostable CFP region, whereas those closest to the C12 source begin in the monostable YFP region. This means that the cells are traversing the bistable region along different paths and will therefore exhibit hysteresis as they converge to the spatial average. If the spatial average concentrations lie within the bistable region, which is defined in concentration space, all cells will eventually experience concentrations within this region as the morphogens diffuse. Thus, over time, the region of cells in physical space that exhibit bistability expands to encompass the entire domain (Fig. 3e, S; Supplementary Movie 3). The result is that a cell's state will be determined by its history and cells with different histories that originated on different sides of the boundary will end up in different stable states. The cells close to the C6 source will behave like cells conditioned in C6 and express CFP while the cells close to the C12 source will behave like cells conditioned in C12 and express YFP (as in Fig. 2) and will maintain their states even after the two morphogens mix via diffusion. The result is the formation of two mutually exclusive domains of gene expression with a sharp boundary that is stable and stationary, even though the morphogen gradients that produced those domains were only present transiently. If, on the other hand, the spatial average concentrations lie outside the bistable region, a transient boundary will form and cells will switch fates as they leave the bistable region, taking on the fate determined by the morphogen of greater concentration (Fig. 3e, M; Supplementary Movies 4 and 5).

A secondary gradient creates self-organized domains. Given the ability of the mutual inhibition topology to produce stable domains of gene expression in response to antiparallel morphogen gradients, we hypothesized that it could function similarly in response to a single morphogen gradient with the addition of a secondary gradient produced by the cells themselves, which functions as a lateral inhibitor²⁸. This circuit mimics the sequential induction of organizing centres found in *C. elegans* vulval development²⁹, the rhombomeres of the vertebrate hindbrain^{30,31}, and the *Drosophila* wing disc³².

To explore this mechanism we added previously characterized Relay circuits²⁴, which produce one signal in response to the other, to the Exclusive Receiver circuit. This created an Exclusive Relay circuit that both produces and interprets morphogen gradients (Fig. 4a and Supplementary Methods). We created a transient gradient of C6 by replacing a cylinder of agar in the centre of a plate with agar containing 40 μ M C6, and plated cells

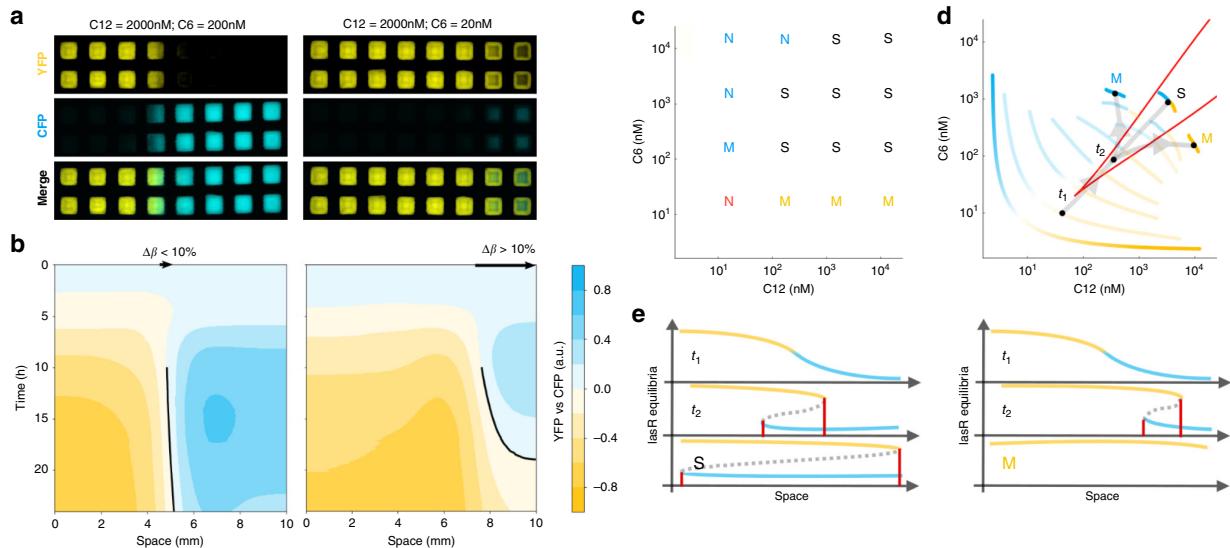


Fig. 3 Formation of stable boundaries. **a** Endpoint fluorescence microscopy of Exclusive Receiver cells grown in transient gradients of signals (C12 diffusing from the left, C6 diffusing from the right) at the spatial average concentrations indicated and in the context of 10 μ M IPTG throughout. Representative examples ($n=3$ biological replicates performed on 3 different days) of a static boundary (left) and a moving boundary (right). **b**, Corresponding kymographs of CFP and YFP fluorescence (intensity) over time (y-axes, hours) at different spatial positions (x-axes, mm). If the location of the boundary (location of equal normalized CFP and YFP fluorescences, black lines) at the end of the timelapse minus its location when it became detectable ($\Delta\beta$, arrows) was less than 10% of the domain size we considered the boundary stable. **c** Boundaries were evaluated as above at the signal concentrations indicated by letters. S indicates equilibrium concentrations at which static boundaries were observed. M indicates a moving boundary. "N" indicates no boundary. The colour of the letter indicates which FP was dominant and red indicates neither FP dominant. See supplementary Figs. 31-33 for replicates. **d** Schematic representation of the concentrations of C6 and C12 experienced by cells at different points in physical space (cyan and yellow curves) as gradients diffuse to homogeneity. Paler curves represent different timepoints. If the spatial average concentrations lie within the region of bistability, the boundary will be static (S), otherwise the boundary will move (M) and will eventually be abolished as cells adopt either CFP or YFP expression. t_1 and t_2 indicate timepoints considered in **e**. **e** Corresponding schematic representing LasR expression, coloured according to resultant fluorescent protein expression. Dashed line indicates the location of an unstable local equilibrium. Red lines indicate the spatial location in which cells are exhibiting bistability. In the case of a stationary boundary (S), the region of space containing cells exhibiting bistability expands to encompass all cells and their gene expression state is determined by their history. In the case of a moving boundary (M), the region exhibiting bistability moves rightward and disappears and the domain becomes dominated by a single monostable state.

transformed with the Exclusive Relay circuit onto gridded filters. The primary C6 gradient resulted in cells in the centre expressing CFP, LacI, and LasI (Fig. 4b and c). These cells produced C12 but were unable to sense it because they did not express LasR due to its repression by LacI. The gradients of C6 and C12 overlapped but the C12 gradient could extend further due to C12 being actively produced by a large region of cells (Fig. 4d). At a certain distance from the source of C6, the ratio of C12 to C6 favoured the C12 state such that the bistable switch 'flipped' and cells expressed YFP and TetR, repressing LuxR and the ability to sense C6. The result was two domains of mutually exclusive gene expression from a single primary morphogen gradient. As in the case of antiparallel gradients, the stability of the boundary between these domains of gene expression can be understood using the same framework: The spatial average concentration of the primary morphogen remained constant as the total density that was added at the beginning of the experiment was unchanging. The secondary morphogen, in contrast, was being produced by cells so the total density increased over time. The spatial average of both morphogens therefore moved along the axis of the secondary morphogen as cells that were sensing the primary morphogen produced the secondary (Fig. 4e, red arrows). The result is that a metastable boundary is produced that is present as long as the system remains within the region of bistability but will eventually be abolished as the secondary morphogen accumulates. Transforming the Exclusive Receiver

with a P81-LuxI Relay and creating a C12 primary gradient resulted in equivalent patterning with the physical location of the states reversed (Supplementary Fig. 10).

Discussion

By building a synthetic gene circuit composed of mutual inhibition downstream of diffusible morphogens, we have shown that this network topology in isolation is sufficient to recapitulate the behaviour, seen in multicellular developmental systems, of mutually exclusive domains of gene expression separated by a boundary that is sharp and stable despite transient and dynamic morphogen gradients. This topology also proves to be robust to differences in morphogen concentration, as variations in concentration spanning orders of magnitude result in only small changes in boundary location (Supplementary Fig. 9, Supplementary Movie 1). In addition, we have built a patterning circuit that creates a self-organised boundary between two gene expression domains in response to a primary morphogen by creating a secondary morphogen that functions as a lateral inhibitor. This circuit shares features of both a positional information model¹, in that it interprets a preexisting morphogen gradient to produce domains of gene expression, and a reaction-diffusion model^{33,34}, in that morphogen production is coincident with interpretation. These two models have previously been presented in opposition to each other, but it is likely that both mechanisms are at work in development³⁵. Our Exclusive Relay

ARTICLE

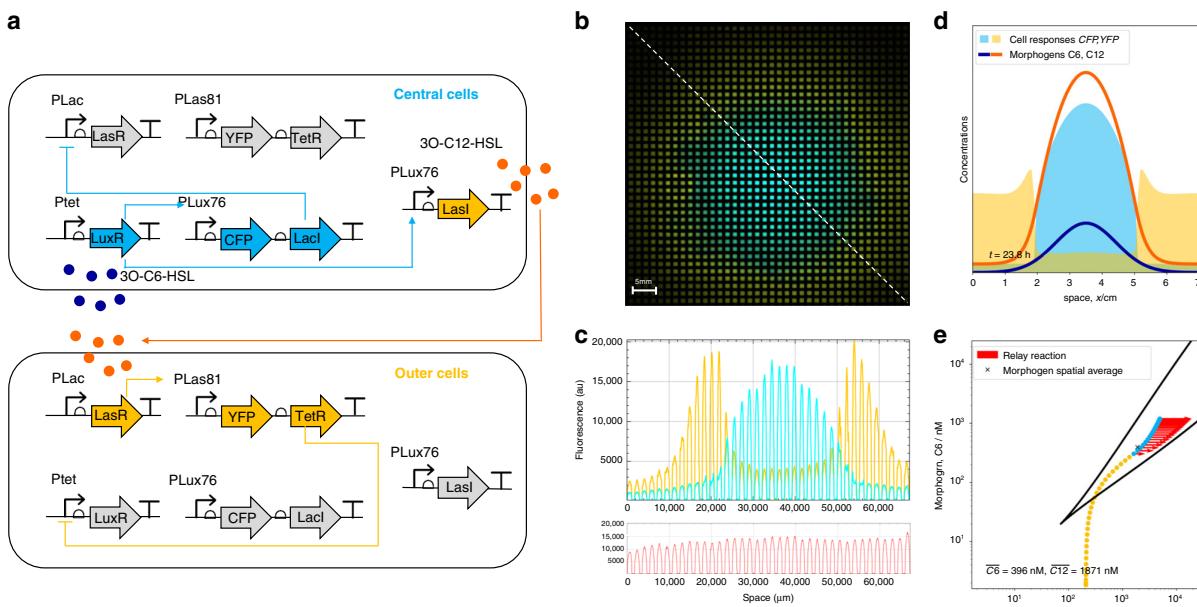
NATURE COMMUNICATIONS | <https://doi.org/10.1038/s41467-020-19098-w>

Fig. 4 Addition of a relay circuit creates self-organized domains of gene expression. **a** Circuit diagram of Exclusive Receiver cells co-transformed with a Relay circuit (P76-LasI) that responds to C6 by producing C12. **b** Isogenic cells transformed with the circuit shown in **a** and grown for 24 h in the presence of a gradient of C6 diffusing from the centre. Cells that experience high levels of C6 (central cells) will express CFP, LacI, and LasI, causing them to produce C12 but be unable to sense it. Neighbouring cells (outer cells) that do not experience C6 will sense C12 and express YFP and TetR, resulting in mutually exclusive domains of gene expression. Cells also constitutively express mRFP1 via a genomic transgene. Image is representative of 3 biological replicates performed on 3 different days. **c** Quantitation of fluorescence along the dotted line in **b**. Cyan, yellow, and red indicate CFP, YFP, and RFP expression, respectively. **d** Final timepoint of simulation shows a secondary gradient of C12 (orange) produced in response to the primary C6 gradient (dark blue). Cyan and yellow indicate simulated CFP and YFP expression, respectively. **e** Final time point of simulation in C6-C12 space labelling points in physical space by their CFP and YFP expression (cyan and yellow points), and showing the production of C12 as vectors (red arrows) that move the spatial average (x) toward increasing C12 (see supplementary video 6).

circuit forms patterns by using two in-phase morphogen gradients to produce a primary response whose spread is limited by the gradient of the secondary morphogen, due to its larger magnitude and diffusive radius. The result is concentration-dependent domains of gene expression produced in response to a morphogen gradient, just as in the French flag model. However, due to the hysteresis of the bistable core of the circuit, these domains of gene expression are metastable even though the primary morphogen gradient diffuses to homogeneity. It is worth noting that the boundary between the eYFP-expressing domain and the domain that expresses neither fluorescent protein is determined simply by the threshold of response to C12 and is therefore not a stable boundary. This boundary could be stabilised by recapitulating the mechanism we have described via the addition of a third morphogen (and mutual inhibition with C12) either diffusing from the opposite direction to the primary gradient (as in Fig. 3) or as a second relay mechanism (as in Fig. 4). The fact that genetic circuits optimized in different contexts can be directly composed to produce more complex patterns suggests that the synthesis of reaction-diffusion and positional information mechanisms may be readily obtainable through evolution, and therefore common in development. This also provides a blueprint for designing synthetic gene circuits that produce spatiotemporal patterns in cell populations, which could lay the groundwork for rationally designing self-organizing, self-repairing materials and tissues.

Methods

Plasmid construction. The exclusive receiver circuit and variants described in Supplementary Fig. 1 were cloned using Gibson Assembly³⁶ using pRS3S175²⁴ as a

starting point. Primers used for Gibson assembly can be found in Supplementary Table 1.

Plate fluorometer assays. The exclusive reporter construct was transformed into EC10G *E. coli* cells with a chromosomally integrated mRFP1 construct²⁴. Overnight cultures were grown from glycerol stocks in M9 media supplemented with 0.4% glucose, 0.2% casamino acids, and 50 $\mu\text{g}/\text{ml}$ kanamycin (supplemented M9) then diluted back 1:100, allowed to grow to an OD of 0.3 then diluted 1:1000. Cultures were aliquoted into black-walled, clear-bottom 96-well plates (Greiner μClear) in a volume of 200 μl per well and measurements taken every 10 min for ~1000 min in a BMG FLUOstar Omega plate fluorometer using BMG FluoStar Omega Reader Control Software 5.10R2. 3-oxohexanoyl-homoserine lactone, C6, (Cayman Chemicals) and 3-oxododecanoyl-homoserine lactone C12, Cayman Chemicals were dissolved to a concentration of 200 mM in DMSO then C6 was diluted in supplemented M9 to the concentrations described, while C12, due to its limited solubility in aqueous media, was first diluted 1:50 in ethanol then diluted in supplemented M9 medium to the concentrations described.

Flow-cytometric analysis of hysteresis. EC10G cells transformed with the exclusive receiver construct were grown overnight from glycerol stocks as described for plate fluorometer assays. Overnight culture was diluted 1:100 and incubated until OD 0.2. Cells were then resuspended 1:100 in supplemented M9 supplemented with either C6 or C12 at 500 nM each and conditioned for 2 h. Following conditioning, cells were washed three times in supplemented M9 by centrifugation at 3200 $\times g$ for 4 min. and seeded at 1:1000 into wells of a 96-well plate containing combinations of varying concentrations of C6 and C12 as indicated in Supplementary Fig. 5. The plate was incubated for 5 h with continuous OD monitoring. OD measurements at 5 h were consistently within the exponential growth range (0.3–0.8). Following 5 h incubation, cells were diluted 1:6 in PBS and analysed by flow cytometry on a BD FACSCelesta (BD Biosciences, San Jose, CA, USA) equipped with HTS and a standard optical setup. Data was collected using FACSDiva 8.01. CFP was excited with violet laser 405 nm and detected with 525/50 BP filter - 505 LP mirror combination. YFP was excited with blue laser 488 nm and detected with 530/30 BP filter - 505 LP mirror combination. RFP was excited with yellow-green laser 561 nm and detected with 610/20 BP filter - 600 LP mirror combination. Instrument quality control was performed prior to each experiment

using BD CS&T beads. Fluorescence compensation parameters were determined using induced and untreated exclusive reporter cells and 30,000 events were counted within RFP gate for each sample. Data analysis was performed with FCS Express v.7 software (DeNovo Software, Glendale, CA, USA). The gating strategy for all flow cytometry was the same and is shown in Supplementary Figs. 3c and 5c.

Microfluidics. Glycerol stocks of EC10G transformed with the Exclusive Receiver were streaked on LB agar plates. Colonies were picked into M9 and grown at 37 °C overnight, then diluted 1:1000 into M9 and grown for 4 h 45 min at 37 °C into exponential phase. The culture was typically diluted 1:100 in M9 before being loaded into the CellASIC ONIX B04A-03 microfluidic device using the manufacturer's protocol (EMD Millipore Corporation). Cells were supplied with media using a pressure of 5 psi in the device. The entire device, along with most of the microscope, was incubated at 37 °C during movie acquisition. Cell segmentation was done on the RFP channel using the published Schnitzcells software (release 1.1, 2005)³⁷. The mean YFP and CFP fluorescence normalized to cell area was then calculated by averaging the corresponding pixels in the respective channels. Movies with no cells or non-growing cells were excluded by keeping only those movies with greater than 20 and 50 cells at 3 and 6 h, respectively. Non-cell segmentation artefacts were excluded by area (<200 pixels) and Euler number (<1) computed with the regionprops function of MATLAB 2014a.

Microfluidics microscopy. Microfluidics devices were imaged using a widefield microscope with epifluorescence and phase contrast imaging modes (Nikon Ti-eclipse, Nikon, UK) equipped with the Nikon Perfect Focus (PFS) Unit. Illumination for the epifluorescence was provided by a white light LED source (SOLA SE Light Engine or Spectra X Light Engine, Lumencor, USA), transmitted by a liquid light guide (Lumencor, USA), through a fluorescence filter cube (YFP Channel: 49003; ET/Sputtered series ET-EYFP, exciter: ET500/20x, dichroic: T515LP, emitter: ET535/30m; CFP Channel: 49001; ET/Sputtered series ET-CFP, exciter: ET436/20x, dichroic: T455LP, emitter: ET480/40m; RFP Channel: 41027-Calculm Crimson, excitation: HQ580/20x, dichroic: Q595LP, emitter: HQ630 /60m, Chroma, USA), and a CFI Plan Apochromat 100x oil immersion objective (NA 1.45, Nikon). Phase contrast illumination was provided by a 100 W lamp via a condenser unit (Nikon). Images were acquired on CoolSNAP HQ2 camera (Photometrics, USA). The sample was held in motorized stages (Nikon). The sample was incubated along with much of the microscope body using a temperature controlled, heated chamber (Solent Scientific, UK). The microscope was controlled with MetaMorph software (version 7.8.10.0, Molecular Devices, USA). Fluorescent beads (TetraSpeck microspheres, 0.5 μm, Molecular Probes, USA) were imaged as a calibration standard.

Solid culture assays. Exponential phase cultures were grown to an OD of 0.3 and plated onto Iso-Grid membranes (Neogen) on supplemented M9 with 1.5% agar at a volume of 0.5 μl per square. Gradients were created by cutting holes in supplemented M9-agar (1.5%) plates (cast in OmniTray [Nunc]) containing 10 μM IPTG. Holes were cut on both ends of a domain to be inoculated at a size of 25% of the domain, each. Holes were then filled with liquid M9-agar to which either 3O-C6- or 3O-C12-HSL had been added at 4X concentration. After hardening, excess agar was cut away leaving each domain isolated. For relay circuit assays, circular holes were punched in the centre of plates using the back of a pipette tip and the holes were filled with ~200 μl of liquid agar containing 40 μM of the appropriate HSL. Plates were sealed with parafilm and imaged using a motorized Leica M205 FA fluorescence stereo microscope controlled using Leica LAS X software. Plates were incubated at 37 °C using a DigiTherm microscope temperature control air bath (Tritech Research). Illumination was an LED white light source (Lumencor) with excitation filters of 426–446 nm, 490–510 nm, and 555–589 nm, and emission filters of 460–500 nm, 520–550 nm, and 608–682 nm. Tiled images were taken every 10 min and were stitched using Leica LAS X software.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Source data are provided with this paper. Plate fluorometer and flow cytometry datasets can be found in supplementary file sourcedata.zip. Raw microscopy images (Figures 2c, 3a, 4b) are available on request to the authors. The exclusiver receiver plasmid is available from Addgene (Addgene ID 160376). All other relevant data are available from the authors upon reasonable request. Source data are provided with this paper.

Code availability

Code is available at GitHub repository <https://github.com/gszep/double-exclusive-reporter>. Source data are provided with this paper.

Received: 15 July 2020; Accepted: 23 September 2020;

Published online: 02 November 2020

References

- Wolpert, L. Positional information and the spatial pattern of cellular differentiation. *J. Theor. Biol.* **25**, 1–47 (1969).
- Jaeger, J. & Reinitz, J. On the dynamic nature of positional information. *BioEssays* **28**, 1102–1111 (2006).
- Briscoe, J. & Small, S. Morphogen rules: design principles of gradient-mediated embryo patterning. *Development* **142**, 3996–4009 (2015).
- Kicheva, A. et al. Kinetics of morphogen gradient formation. *Science* **315**, 521–5 (2007).
- Durrieu, L. et al. Bicoid gradient formation mechanism and dynamics revealed by protein lifetime analysis. *Mol. Syst. Biol.* **14**, e8355 (2018).
- Driever, W. & Nüsslein-Volhard, C. A gradient of bicoid protein in *Drosophila* embryos. *Cell* **54**, 83–93 (1988).
- Balaskas, N. et al. Gene regulatory logic for reading the sonic hedgehog signaling gradient in the vertebrate neural tube. *Cell* **148**, 273–284 (2012).
- Fried, P. & Iber, D. Read-out of dynamic morphogen gradients on growing domains. *PLOS ONE* **10**, e0143226 (2015).
- Clyde, D. E. et al. A self-organizing system of repressor gradients establishes segmental complexity in *Drosophila*. *Nature* **426**, 849–853 (2003).
- Sokolowski, T. R., Erdmann, T. & TenWolde, P. R. Mutual repression enhances the steepness and precision of gene expression boundaries. *PLoS Comput. Biol.* **8**, e1002654 (2012).
- Zagorski, M. et al. Decoding of position in the developing neural tube from antiparallel morphogen gradients. *Science* **356**, 1379–1383 (2017).
- Srinivasan, S. et al. A BMP-FGF morphogen toggle switch drives the ultrasensitive expression of multiple genes in the developing forebrain. *PLoS Comput. Biol.* **10**, 1–16 (2014).
- 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).
- Kong, W., Blanchard, A. E., Liao, C. & Lu, T. Engineering robust and tunable spatial structures with synthetic gene circuits. *Nucleic acids Res.* **45**, 1005–1014 (2017).
- Sekine, R., Shibata, T. & Ebisuya, M. Synthetic mammalian pattern formation driven by differential diffusivity of Nodal and Lefty. *Nat. Commun.* **9**, 5456 (2018).
- Schaerli, Y. et al. A unified design space of synthetic stripe-forming networks. *Nat. Commun.* **5**, 4905 (2014).
- Payne, S. et al. Temporal control of self-organized pattern formation without morphogen gradients in bacteria. *Mol. Syst. Biol.* **9**, 697 (2013).
- Cao, Y. et al. Programmable assembly of pressure sensors using pattern-forming bacteria. *Nat. Biotechnol.* **35**, 1087–1093 (2017).
- Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat. Genet.* **31**, 64–68 (2002).
- Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
- 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).
- Youk, H. & Lim, W.A. Secreting and sensing the same molecule allows cells to achieve versatile social behaviors. *Science* (New York, N.Y.) **343**, 1242782 (2014).
- Cao, Y. et al. Collective space-sensing coordinates pattern scaling in engineered bacteria. *Cell* **165**, 620–630 (2016).
- Grant, P. K. et al. Orthogonal intercellular signaling for programmed spatial behavior. *Mol. Syst. Biol.* **12**, 849 (2016).
- Rudge, T. J. et al. Characterization of Intrinsic Properties of Promoters. *ACS Synth. Biol.* **5**, 89–98 (2016).
- Dalchau, N. et al. Scalable dynamic characterization of synthetic gene circuits. Preprint at <https://www.biorxiv.org/content/10.1101/635672v1> (2019).
- Boehm, C. R., Grant, P. K. & Haseloff, J. Programmed hierarchical patterning of bacterial populations. *Nat. Commun.* **9**, 776 (2018).
- Meinhardt, H. & Gierer, A. Pattern formation by local self-activation and lateral inhibition. *BioEssays* **22**, 753–760 (2000).
- Simske, J. S. & Kirn, S. K. Sequential signalling during *caenorhabditis elegans* vulval induction. *Nature* **375**, 142–146 (1995).
- Hernandez, R. E., Rikhof, H. A., Bachmann, R. & Moens, C. B. vhnfl integrates global RA patterning and local FGF signals to direct posterior hindbrain development in zebrafish. *Development* **131**, 4511–4520 (2004).
- Schilling, T. F., Nie, Q. & Lander, A. D. Dynamics and precision in retinoic acid morphogen gradients. *Curr. Opin. Genet. Dev.* **6**, 562–569 (2012).
- Piddini, E. & Vincent, J. P. Interpretation of the Wingless Gradient Requires Signaling-Induced Self-Inhibition. *Cell* **136**, 296–307 (2009).
- Turing, A. M. The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond. Ser. B, Biol. Sci.* **237**, 37–72 (1952).
- Gierer, A. & Meinhardt, H. A theory of biological pattern formation. *Kybernetik* **12**, 30–9 (1972).
- Green, J. B. A. & Sharpe, J. Positional information and reaction-diffusion: two big ideas in developmental biology combine. *Development* **142**, 1203–1211 (2015).

ARTICLE

NATURE COMMUNICATIONS | <https://doi.org/10.1038/s41467-020-19098-w>

36. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
37. Young, J. W. et al. Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy. *Nat. Protoc.* **7**, 80–88 (2012).
38. Yordanov, B. et al. A Computational Method for Automated Characterization of Genetic Components. *ACS Synth. Biol.* **3**, 578–588 (2014).

Acknowledgements

We would like to thank Jeremy Green for helpful comments on the manuscript.

Author contributions

P.K.G., N.D., J.H., and A.P. conceived and designed the study. P.K.G. designed and built the genetic circuits. P.K.G., O.P., and V.C. performed the experiments. G.S., J.H., and N.D. conceived and implemented theory and modelling and wrote the supplementary information. All authors analysed and interpreted the data. P.K.G. and A.P. wrote the main text. All authors provided input into the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-19098-w>.

Correspondence and requests for materials should be addressed to P.K.G. or A.P.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

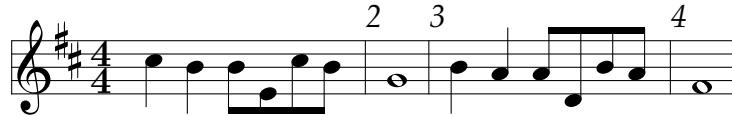
3.7 Afterword

The decision to focus on single cell trajectories and flow cytometry came from the limitations of using microplate data in Chapter 3. The model parameters θ were estimated using a hierarchical monte-carlo approach and time-course fluorescence microplate measurements (details of which can be found in Appendix A.2.1). The time-courses include information about dynamical transients and colony growth in liquid culture as shown in Figure 3.6. The desired cusp bifurcation, however, lives in state-space rather than the time-domain. The disconnect between the domain that the data lives in and the domain of the design goals poses the risk of overfitting the model on undesired information that exists in the data domain.

It is not possible to observe the cusp bifurcation in microplate data, due to the averaging of signals originating from heterogeneous cell populations. Instead, the cusp bifurcation can be observed in flow cytometry measurements of colonies in exponential phase (Supplementary Figure A.3) and microfluidic fluorescence microscopy data (Figure 3.8c) where computations on single-cell trajectories reveal the hysteresis loop which must necessarily accompany the cusp.

Chapter 4

Parameter Inference with Bifurcation Diagrams



the melody that will draw you into the infinite darkness

Nocturne of Shadow — Ocarina of Time

4.1 Preface

4.1.1 Problem Statement & Context

This chapter focuses on the problem of looking for parameter regimes for dynamical systems that result in bifurcations. This has been coined as *inverse bifurcation analysis* [30]. Formally we are looking for parameters θ for which the equation (2.1) has at least one fixed point for which one of the bifurcation criteria (see Section 2.1.2) is satisfied. Furthermore we would like to create a specified type and number of bifurcations to be placed along control condition p .

Inverse bifurcation analysis becomes relevant to biomedical researchers in settings where there is a design goal to engineer an organism with a distinct phenotype (as discussed in Chapter 1). Such a phenotype can be a the *double exclusive reporter* investigated in the interdisciplinary

collaboration in synthetic developmental biology presented in Chapter 3. This *E. Coli* phenotype exhibits a cusp bifurcation leading to a bistable response with respect to two input signals as shown in Figure 3.8. The engineering goal could be a phenotype with a specific cell cycle [31, 32] which involves the positioning of *Hopf bifurcations* that mark the onset of oscillations in protein concentrations from stable equilibria. Finally the design of self-organised patterns such as stripes and spots in mammalian coat patterns could be another engineering goal. According to a long standing mathematical hypotheses [33] that underpins developmental mechanisms this may involve the search for *Turing bifurcations*.

Before we get carried away and think that with the right tooling we could genetically control the length scale of spots or stripes on cats, we must remind ourselves that *in vivo* gene regulatory networks mostly consist of unknown and experimentally inaccessible parameters. Furthermore, we find that usually there exist multiple equally valid models that describe the observed behaviour, and so turn to *model selection* methods. Even if we had an accurate and unique model to describe an organism, generic tools for *inverse bifurcation analysis* are limited as we've explored in Chapter 2 and experienced in practice in Chapter 3. In an attempt to address these limitations, the incorporated publication (sections 4.2 – 4.6) focuses on the design of systems of ordinary differential equations with *pitchfork* and *saddle-node* bifurcation diagrams. Although the publication focuses on parameter synthesis for a subset of bifurcations, steps towards *Hopf* bifurcations are made (see Appendix B.5), and the approach lays foundations for differentiable optimisation methods that leverage bifurcation theory. A view towards how this approach can be used for the design of *Turing patterns* and model selection is discussed in concluding Chapter 6.

4.1.2 Contributions

Grisha Szep prepared the manuscript, designed the cost function, derived mathematical results, wrote and released the Julia package under the su-

pervision of **Neil Dalchau** and **Attila Czikasz-Nagy**.

Parameter Inference with Bifurcation Diagrams

Gregory Szep

King's College London
London, WC2R 2LS, UK
gregory.szep@kcl.ac.uk

Attila Csikász-Nagy

Pázmány Péter Catholic University
Budapest, 1083, Hungary
csikasznagy@gmail.com

Neil Dalchau

Microsoft Research Cambridge
Cambridge, CB1 2FB, UK
ndalchau@gmail.com

Abstract

Estimation of parameters in differential equation models can be achieved by applying learning algorithms to quantitative time-series data. However, sometimes it is only possible to measure qualitative changes of a system in response to a controlled condition. In dynamical systems theory, such change points are known as *bifurcations* and lie on a function of the controlled condition called the *bifurcation diagram*. In this work, we propose a gradient-based approach for inferring the parameters of differential equations that produce a user-specified bifurcation diagram. The cost function contains an error term that is minimal when the model bifurcations match the specified targets and a bifurcation measure which has gradients that push optimisers towards bifurcating parameter regimes. The gradients can be computed without the need to differentiate through the operations of the solver that was used to compute the diagram. We demonstrate parameter inference with minimal models which explore the space of saddle-node and pitchfork diagrams and the genetic toggle switch from synthetic biology. Furthermore, the cost landscape allows us to organise models in terms of topological and geometric equivalence.

1 Introduction

Inverse problems [1] arise in biology and engineering in settings when the model is not fully known and the desire is to match model behaviour to a given set of observations. This helps systematically guide both model and experimental design. While we would like to understand the quantitative details of a system, often only qualitative changes in response to varying experimental conditions can be robustly measured across independent studies [2, 3]. For example, several studies are likely to agree that the human immune system activates above a threshold concentration of a pathogen and deactivates at a lower threshold concentration, but may disagree on the exact quantities of the thresholds or the magnitudes of the immune response. Bifurcation theory provides us a framework for studying these transitions in a manner that is independent of quantitative details [4]. The emerging picture suggests that identification of the qualitative behaviour – the bifurcation diagram – should precede any attempt at inferring other properties of a system [5].

Inferring the parameters of a model directly from a bifurcation diagram is difficult because it is not obvious how multiple parameters in concert control the existence and position of a bifurcation. It could even be impossible for the model to bifurcate in the manner desired. For models with a sufficiently small number of parameters, finding specific bifurcation diagrams is typically done by

hand [6]. Several approaches exist to place bifurcations to desired locations once a manifold is present [7–9] yet typically resort to sampling techniques to search for them in the first place [10, 11]. It is always possible to design bespoke goodness of fit measures to find specific model behaviours, for example using the period and phase of limit cycle oscillations [12]. However, this approach does not generalise across a wider set of qualitative behaviours. Progress has been made in cases where model structure and stability conditions are used to refine the search space [13, 14] yet the resulting objectives are still not explicit in the bifurcation targets and also not differentiable. In the emerging field of scientific machine learning [15–17], parameters of structured mechanistic models are favoured over flexible models in larger parameter spaces. A scalable method for navigating the space of bifurcation diagrams would enable design of differential equations with high-level qualitative constraints. Furthermore one could begin organising models according to qualitatively distinct behaviours.

Back-propagation through differential equation solvers has been a breakthrough over the past couple of years [18, 19] that enabled scalable parameter inference for differential equations from trajectory data. Although one could use trajectory data to create the aforementioned qualitative constraints [20, 21] this would entail over-constraining information originating from the kinetics and dynamical transients of the model. Furthermore, such data usually does not contain sufficient information about dynamical transients in order to identify kinetic parameters. Techniques for back-propagating through implicit equation solvers have also been developed [22, 23] although to the best of the authors' knowledge have not been applied to bifurcation diagrams at the time of writing this paper.

The problem of inferring differential equation parameters against a user-specified bifurcation diagram decomposes into two parts: searching for bifurcating regimes and matching the locations of bifurcation points to desired values. Matching bifurcation locations is a supervised problem where the data are expressed as bifurcations points [8, 11]. Searching for bifurcations is an unsupervised problem because when bifurcations are not present, there is no distance defined between data and prediction [10]. Therefore only properties of the model can be used to start the search. We propose an approach for performing both tasks in an end-to-end fashion. The bifurcation diagram encodes high-level qualitative information defined by state space structures, rather than kinetics. We apply the strategy of implicit layers [22, 23] to calculate gradients. To compute the diagram we use a predictor-corrector method called deflated continuation [24, 25] developed for partial differential equations.

We find that the cost function landscape contains basins that not only allow us to synthesise models with a desired bifurcation diagram but also allow us to organise models in terms of topological and geometric equivalence. We discuss the relevance of this in model selection. In summary, our paper has the following main contributions:

- An end-to-end differentiable method for locating bifurcations in parameter space and then matching their dependency on a control condition to user-specified locations
- Implementation of the method as a Julia package `BifurcationInference.jl`
- Leveraging the cost landscape for a novel way of organising differential equation models in terms of geometric and topological equivalence

1.1 Preliminaries

Suppose we collected observations along a scalar control condition $p \in \mathbb{R}$ and conclude that there are specific values of p for which there are qualitative changes in system behaviour. Let \mathcal{D} be the set of those values and let us hypothesise that these transitions occur due to bifurcations in the dynamics that drive the underlying mechanism. Let us model the mechanism with a parametrised set of differential equations for states $u \in \mathbb{R}^N$ with a vector function F_θ in a parameter space $\theta \in \mathbb{R}^M$.

For the purposes of introducing this work, we will consider the simplest class of bifurcations known as *co-dimension one* bifurcations not including limit cycles. Therefore \mathcal{D} should contain conditions for which we hypothesise changes in multi-stable behaviour. Let the equations be

$$\frac{\partial u}{\partial t} = F_\theta(u, p) \quad \text{where} \quad F_\theta : \mathbb{R}^{N+1} \rightarrow \mathbb{R}^N \quad (1)$$

In the context of the differential equations, and not considering limit cycles for now, we show that a static non-degenerate bifurcation can be defined by a set of conditions on the determinant of the Jacobian $|\frac{\partial F_\theta}{\partial u}|$. The determinant of the Jacobian quantifies the rate at which trajectories in a local

patch of state-space $u \in \mathbb{R}^N$ converge or diverge. Let $s \in \mathbb{R}$ parametrise the curves that trace out the bifurcation diagram. Any location on the curve $u(s)$ and $p(s)$ must satisfy the steady-state of equations (1). Directional derivatives $\frac{d}{ds}$ along the diagram require the calculation of a vector that is tangent to the diagram (see Supplementary A). The determinant approaching zero along the diagram means that the dynamics of the system are slowing down, which is an important indicator for the onset of a transition between qualitative behaviours. Furthermore, the slowing down must necessarily be followed by a breakdown of stability; for this to be true it is sufficient *but not necessary* to require that the determinant cross zero with a finite slope, meaning that its directional derivative along the diagram $\frac{d}{ds} \left| \frac{\partial F_\theta}{\partial u} \right|$ is not zero. This is the non-degeneracy condition. The set of predicted values for the control condition $\mathcal{P}(\theta) \subset \mathbb{R}$ at which bifurcations occur are defined as

$$\mathcal{P}(\theta) := \left\{ p \mid \exists u : F_\theta(u, p) = 0, \left| \frac{\partial F_\theta}{\partial u} \right| = 0, \frac{d}{ds} \left| \frac{\partial F_\theta}{\partial u} \right| \neq 0 \right\} \quad (2)$$

A proof of how the conditions (2) are necessary and sufficient for static non-degenerate bifurcations is detailed in Supplementary B. The most common bifurcations between steady states, not including limit cycles, are saddle-nodes and pitchforks [26]. Saddle-node bifurcations, which often appear in pairs (Figure 1A) are defined by stable and unstable fixed points meeting and disappearing. Pitchfork bifurcations occur where a single steady state splits into two stable and one unstable steady state (Figure 1B shows an *imperfect* pitchfork; a *perfect* pitchfork arises when $\theta_1 = 0$). To illustrate these bifurcations, we define minimal models (Figure 1) that span the space of saddle-node and pitchforks, where indeed zero crossings in the determinant with a finite slope define the set of prediction $\mathcal{P}(\theta)$. The location of these crossings in general may not match the targets \mathcal{D} .

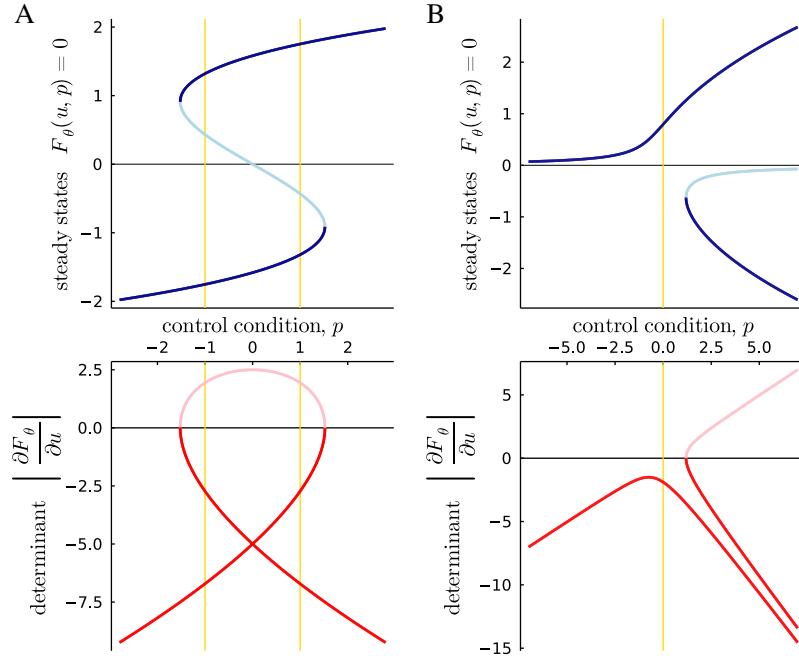


Figure 1: Illustration of bifurcation diagrams for minimal models of bifurcations. A. Saddle-node bifurcations arise for $F_\theta(u, p) = p + \theta_1 u + \theta_2 u^3$ when $\theta = (\frac{5}{2}, -1)$. B. Pitchfork bifurcations arise for $F_\theta(u, p) = \theta_1 + pu + \theta_2 u^3$ when $\theta = (\frac{1}{2}, -1)$. Targets are illustrated by light yellow vertical lines. Bifurcation curves are shown as solid blue and red lines, with lighter shades indicating the determinant crossing zero at locations $\mathcal{P}(\theta)$ giving rise to unstable solutions.

For a given set of parameters θ one could compute the set of predicted bifurcations $\mathcal{P}(\theta)$ using parameter continuation methods [25, 24]. Our goal is to find optimal parameters θ^* that match predictions $\mathcal{P}(\theta^*)$ to specified targets \mathcal{D} . We must design a suitable cost function L so that

$$\theta^* := \operatorname{argmin}_\theta L(\theta|\mathcal{D}) \quad (3)$$

The optimal θ^* is not expected to always be unique, but is in general a manifold representing the space of qualitatively equivalent models. Ideally, the cost function L should reward θ for which the number of predicted bifurcations is equal to the number of targets, $|\mathcal{P}(\theta)| = |\mathcal{D}|$. This is especially important in the case where there are no predictions $|\mathcal{P}(\theta)| = 0$.

2 Proposed Method

2.1 Cost Function

To identify parameter sets that give rise to bifurcation diagrams with specified bifurcation points, we propose a cost function that comprises two terms. The role of the error term is simply to reward predicted bifurcations to coincide with the specified target locations. This of course relies on such bifurcations existing. The role of the eigenvalue term is to encourage an optimiser to move towards parameter regimes that do exhibit bifurcations.

2.1.1 Error term: matching bifurcations to target locations

In order for predicted bifurcations $p(\theta) \in \mathcal{P}(\theta)$ to match targets $p' \in \mathcal{D}$ we need to evaluate an error term $|p(\theta) - p'|$. A naive approach might take an average over the norms for all prediction-target pairs. However this gives rise to unwanted cross-terms and the possibility of multiple predictions matching the same target without any penalty for unmatched targets. Therefore, we choose a geometric mean over the predictions and an arithmetic mean over targets:

$$E(\theta, \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{p' \in \mathcal{D}} \prod_{p(\theta) \in \mathcal{P}(\theta)} |p(\theta) - p'|^{\frac{1}{|\mathcal{P}|}} \quad (4)$$

The error term is only zero when each target is matched by at least one prediction and allows for cases where the number of predictions is greater than or equal to the number of targets $|\mathcal{P}| \geq |\mathcal{D}|$. An alternative approach, which undesirably introduces more hyper-parameters, would be to let each prediction $\mathcal{P}(\theta)$ represent the centroid of a mixture distribution and use expectation-maximisation to match the centroids to targets \mathcal{D} .

2.1.2 Eigenvalue term: encouraging bifurcations

We can see from Figure 1 and definition (2) that predictions $p(\theta)$ can be identified by looking for points along the curve where the determinant crosses zero $|\frac{\partial F_\theta}{\partial u}| = 0$ with a finite slope $\frac{d}{ds} |\frac{\partial F_\theta}{\partial u}| \neq 0$. Using these quantities we can define a positive semi-definite measure $\varphi_\theta(s)$ of zero crossings in the determinant along a curve parametrised by s which we define as

$$\varphi_\theta(s) := \left(1 + \left| \frac{\frac{d}{ds} |\frac{\partial F_\theta}{\partial u}|}{\frac{d}{ds} |\frac{\partial F_\theta}{\partial u}|} \right| \right)^{-1} \quad (5)$$

The bifurcation measure $\varphi_\theta(s)$ is maximal at bifurcations and has finite gradients in non-bifurcating regimes (Figure 2). More specifically, the measure $\varphi_\theta(s)$ is one at bifurcation points and goes to zero an odd number of times between bifurcations. This is because $|\frac{\partial F_\theta}{\partial u}|$ must eventually turn around in order to return back to zero, resulting in the directional derivative $\frac{d}{ds} |\frac{\partial F_\theta}{\partial u}|$ going to zero. Hence the measure $\varphi_\theta(s)$ goes to zero for each turning point (see Figure 2).

On the other hand, as the determinant $|\frac{\partial F_\theta}{\partial u}|$ diverges, we approach regimes far away from any bifurcations and hence $\varphi_\theta(s) \rightarrow 0$. Since we would still like to have non-zero gradients with respect to θ in these regimes we designed the measure to go to zero sufficiently slowly.

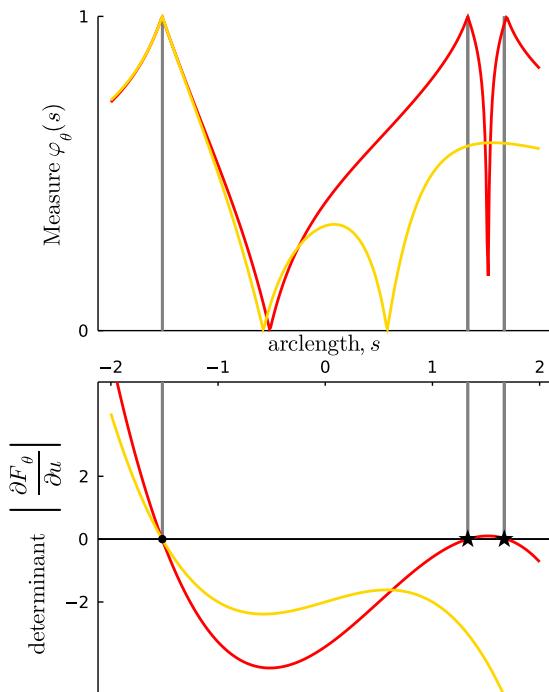


Figure 2: Bifurcation measure $\varphi_\theta(s)$ and determinant $|\frac{\partial F_\theta}{\partial u}|$ along the arclength s of two different bifurcation curves demonstrating how maximising the measure along the curve maintains the existing bifurcation marked by a circle, while encouraging new bifurcations marked by stars.

While the calculation of the determinant is straightforward, its directional derivative requires a tangent vector to the bifurcation curve. Fortunately the tangent vector $T_\theta(s)$ at the solution $u(s), p(s)$ anywhere along the curve s can be calculated as the nullspace of the rectangular $N \times (N + 1)$ Jacobian

$$\frac{\partial F_\theta}{\partial(u, p)} \Big|_{F_\theta(u(s), p(s))=0} \cdot T_\theta(s) = 0 \quad (6)$$

This equation guarantees that the tangent vector $T_\theta(s)$ is orthogonal to all hyper-planes defined by the components of F_θ . In this setting the dimension of the nullspace is always known, and therefore can reliably be calculated using QR factorisation methods [27].

Equipped with a measure that quantifies the appearance of bifurcations along a bifurcation arc we can define the total measure for a bifurcation diagram as

$$\Psi(\theta) := \frac{\int_{F_\theta(u, p)=0} \varphi_\theta(s) ds}{\int_{F_\theta(u, p)=0} ds}. \quad (7)$$

Here we denote $\int_{F_\theta(u, p)=0} ds$ as the sum of the line integrals in $(u, p) \in \mathbb{R}^{N+1}$ defined by the level set $F_\theta(u, p) = 0$ with s being an arbitrary parametrisation of the curves. The total measure $\Psi(\theta)$ is normalised such that $\Psi(\theta) \rightarrow 1$ in the regimes where the controlled condition region p is densely packed with bifurcations. The total measure $\Psi(\theta)$ is added to the error term as if it were a likelihood. This defines the cost function as

$$L(\theta|\mathcal{D}) := (|\mathcal{P}| - |\mathcal{D}|) \log \Psi(\theta) + E(\theta, \mathcal{D}), \quad (8)$$

The pre-factor $|\mathcal{D}| - |\mathcal{P}|$ in the eigenvalue term ensures that the gradients are always pushing optimisers towards a state where $|\mathcal{D}| = |\mathcal{P}|$. This can be seen as a step-wise annealing of the eigenvalue term until the desired state is reached.

2.2 Differentiating the cost function

To make use of gradient-based optimisers to locate desired bifurcation diagrams, we show here how to differentiate the cost function. First, we note that while individual bifurcations $p(\theta)$ depend smoothly on θ , the total number of predictions $|\mathcal{P}|$ does not have gradient contributions with respect to θ . Therefore, we can safely drop the dependency in the prediction counter and now proceed in taking gradients with respect to θ knowing that the only dependencies we need to track are for individual bifurcations $p(\theta)$ within the definition the error term (4) and the total measure (7). Therefore,

$$\frac{\partial L}{\partial \theta} = (|\mathcal{P}| - |\mathcal{D}|) \lambda \frac{\partial \Psi}{\partial \theta} \Psi(\theta)^{-1} + \frac{1}{|\mathcal{D}||\mathcal{P}|} \sum_{p'} \prod_{p(\theta)} |p(\theta) - p'|^{\frac{1}{|\mathcal{P}|}} \sum_{p(\theta)} \frac{\partial p}{\partial \theta} (p(\theta) - p')^{-1} \quad (9)$$

In a similar vein to back-propagation through neural differential equations [18] we would like to be able to calculate the gradient $\frac{\partial L}{\partial \theta}$ without having to differentiate through the operations of the solver that finds the bifurcation diagram $F_\theta(u, p) = 0$ and the bifurcation locations $p(\theta)$. To calculate the gradient of the measure $\frac{\partial \Psi}{\partial \theta}$ we need to differentiate line integrals that depend on θ . Fortunately this can be done by the application of the generalised Leibniz integral rule, details of which can be found in Supplementary C.

The gradient of the bifurcation points $\frac{\partial p}{\partial \theta}$ is found by application of the implicit function theorem to a vector function $G_\theta : \mathbb{R}^{N+1} \rightarrow \mathbb{R}^{N+1}$ whose components represent the two constraints $F_\theta(u, p) = 0$ and $|\frac{\partial F_\theta}{\partial u}| = 0$. By following a similar strategy to that used by implicit layers [22] we yield an $(N + 1) \times M$ Jacobian representing a deformation field [28] for each θ direction. The gradient we are looking for becomes

$$\frac{\partial p}{\partial \theta} = -\hat{p} \cdot \frac{\partial G_\theta}{\partial(u, p)}^{-1} \frac{\partial G_\theta}{\partial \theta} \Big|_{G_\theta(u, p)=0} \quad \text{where} \quad G_\theta(u, p) := \begin{bmatrix} F_\theta(u, p) \\ |\frac{\partial F_\theta}{\partial u}| \end{bmatrix} \quad (10)$$

Here \hat{p} is a unit vector in $(u, p) \in \mathbb{R}^{N+1}$ that picks out the deformations along the p -direction. If we wanted to place the bifurcation at target steady state u' as well as target control condition p' we would use the full $(N + 1) \times M$ deformation matrix. Calculation of this matrix involves inverting an

$(N + 1) \times (N + 1)$ Jacobian $\frac{\partial G_\theta}{\partial(u,p)}$. Instead of explicitly inverting the Jacobian the corresponding system of linear equations is solved. The determinant of this Jacobian goes to zero in the degenerate case where $\frac{d}{ds} |\frac{\partial F_\theta}{\partial u}| = 0$, further justifying our choice of measure $\Psi(\theta)$ which discourages the degenerate case.

The cost function is piece-wise smooth and differentiable with undefined gradients only in parameter contours where the number of predictions $|\mathcal{P}|$ changes; this is when $\Psi(\theta)$ is undefined and the inverse of $\frac{\partial G_\theta}{\partial(u,p)}$ does not exist. Given a set of solutions to $F_\theta(u, p) = 0$ and locations $p(\theta)$ the gradient $\frac{\partial L}{\partial \theta}$ can be evaluated using automatic differentiation methods [29–31] without needing to back-propagate through the solver that obtained the level set $F_\theta(u, p) = 0$ in the forward pass.

3 Experiments & Results

In this section, we apply the method first to minimal examples that can produce saddle-node and pitchfork bifurcations (both $N = 1, M = 2$), and then a slightly more complex model ($N = 2, M = 5$) that has multiple parametric regimes producing saddle-node bifurcations. We also demonstrate our method on a model of greater complexity, to convince the reader that the method can be used on more realistic examples with practical significance. In Supplementary D we demonstrate the identification of saddle-node bifurcations and damped oscillations in a model ($N = 4, M = 21$) of a synthetic gene circuit in *E. coli* [3].

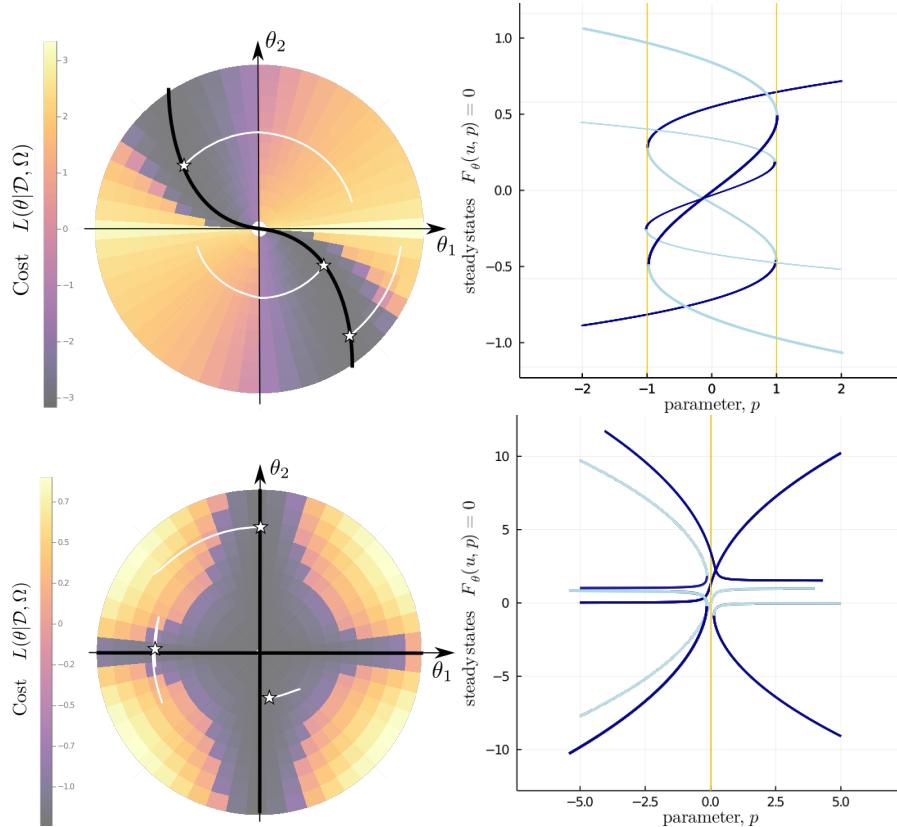


Figure 3: Saddle-node $F_\theta(u, p) = p + \theta_1 u + \theta_2 u^3$ and pitchfork $F_\theta(u, p) = \theta_1 + up + \theta_2 u^3$ optimised with respect to θ so that predicted bifurcations $\mathcal{P}(\theta)$ match targets \mathcal{D} in control condition p . The right panel shows bifurcations diagrams for the three optimal θ^* marked by stars on the left panel. The optimisation trajectories in white follow the gradient of the cost, approaching the black lines of global minima in the left panel

3.1 Minimal Models

Optimisations of two parameters (θ_1, θ_2) using simple gradient descent from Flux.jl with learning rate $\eta = 0.01$ for the minimal saddle-node and pitchfork models (Figure 1) yield trajectories approach-

ing lines of global minima in the cost function (Figures 3) which represent a set of geometrically equivalent models. Two bifurcation diagrams are geometrically equivalent if the number, type and locations of bifurcations match the specified targets \mathcal{D} .

We can see that the geometrically equivalent lines are contained within larger basins where the correct number and type of bifurcations are present but do not match the locations of targets \mathcal{D} . All models within this basin are in some sense topologically equivalent. This hierarchical classification allows us to identify the set of models that satisfy observed qualitative behaviour [5] before any attempt at inferring kinetic parameters, which is done by choosing a model along the line of geometrically equivalent models.

Optimisation trajectories for the two minimal models appear mostly circumferential. This is because the models were set up such that the radial direction from the origin in θ space mostly scale kinetics whereas the circumferential direction changes the bifurcation topology. This suggests that the gradients of our cost function seek to change model geometry over kinetics.

3.2 Genetic Toggle Switch

In this section we optimise a model where the states share a Hill function relationship with co-operatively $n = 2$; these models often emerge from mass action kinetics with quasi-steady state approximations and are used to model species concentrations. After re-scaling the equations governing the dynamics of concentrations, the simplified equations for state u_1 and u_2 become

$$\partial_t u_1 = \frac{a_1 + (pu_2)^2}{1 + (pu_2)^2} - \mu_1 u_1 \quad \partial_t u_2 = \frac{a_2 + (ku_1)^2}{1 + (ku_1)^2} - \mu_2 u_2 \quad (11)$$

where a_k is the baseline production rate for species k in the absence of the other species. Each species has a finite degradation rate μ_k . Finally we have two sensitivity constants p and k , one of which is chosen as our control condition. A baseline production rate $a_k > 1$ recovers an inhibitor type hill function for species k and is an activator otherwise. The sensitivities are proportional to the slope of the hill productions. Solving for the steady states, substituting the equation for u_1 into u_2 and rearranging gives rise to the relationship

$$\frac{k}{\mu_1} = \frac{(1 + (\frac{p}{\mu_2}u')^2)\sqrt{a_2 - u'}}{(a_1 + (\frac{p}{\mu_2}u')^2)\sqrt{u' - 1}} \quad \text{where } u' := u_2\mu_2 \quad (12)$$

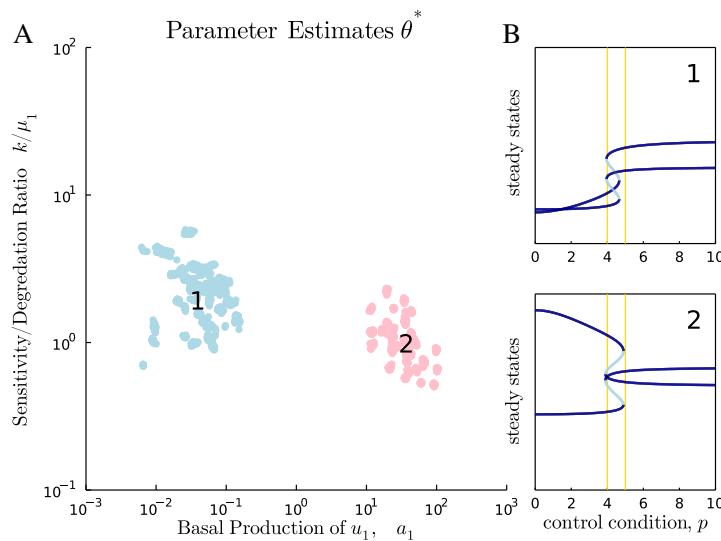


Figure 4: Bifurcation inference for the two-state model (11). A. Optimal parameter estimates θ^* for the targets $\mathcal{D} = \{4, 5\}$ reveal two clusters of qualitatively different regimes: mutual activation ($a_1 < 1$; cluster 1) and mutual inhibition ($a_1 > 1$; cluster 2). B. Example bifurcation diagrams indicate positively and negatively correlated dependencies between the two model states, as a function of the control condition.

which reveals that only a_1, a_2 and the ratio between the sensitivity and degradation parameters, $\frac{k}{\mu_1}$, affect the solutions to this equation, and hence the locations of the bifurcations (Figure 4A). In 98% of 800 runs, optimisation using the ADAM optimiser [32] from Flux.jl with learning rate $\eta = 0.1$ converged to one of two clusters: mutual activation ($a_1 < 1, a_2 < 1$; cluster 1) and mutual inhibition ($a_1 > 1, a_2 > 1$; cluster 2) regimes. Example bifurcation diagrams illustrate how the bifurcation curves of each species are positively correlated in mutual activation and negatively correlated for mutual inhibition (Figure 4B).

In order to maintain biological interpretability, optimisation was restricted to the positive parameter regime by transforming the parameters to log-space $\theta \rightarrow 10^\theta$. At the beginning of each optimisation run an initial θ was chosen in the log-space by sampling from a multivariate normal distribution with mean zero and standard deviation one.

3.3 Complexity

Performing one iteration of the optimisation requires the computation of the gradient of the cost (9), requiring a computation of the bifurcation diagram with parameter continuation methods, which includes the evaluation of matrix inversions (10). Instead of evaluating the inversions directly, we solve a system of linear equations, applying the same strategy as implicit layers [22, 23]. This leaves us with the computational bottleneck of calculating the determinant of the state space Jacobian, required in both the bifurcation measure (5) and gradient (10). This calculation scales like N^2 where N is the number of state space variables (Figure 5A).

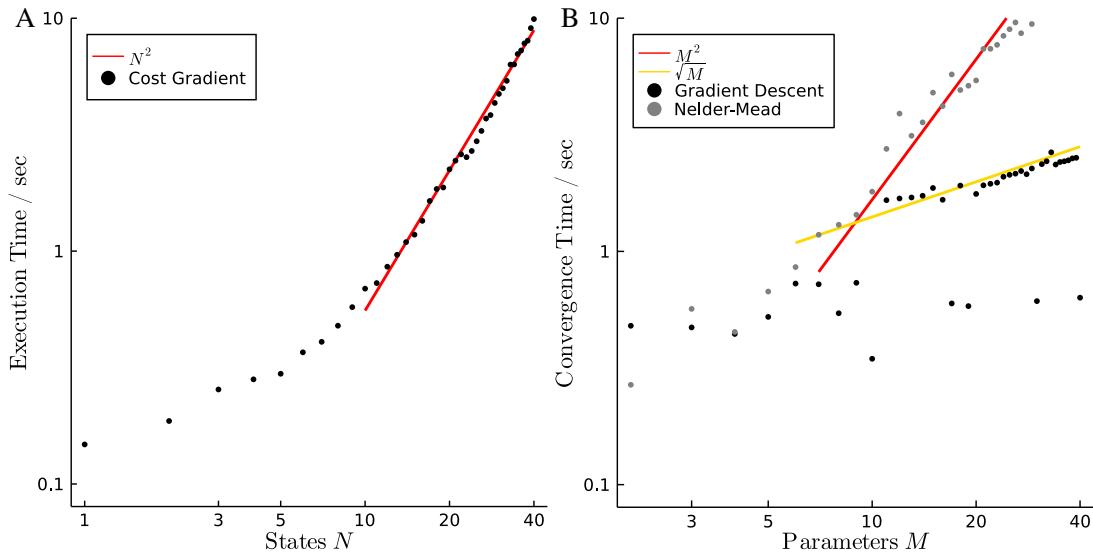


Figure 5: A. Execution time (time to calculate cost gradient) with respect to states N . B. Convergence times (the time it takes to find and match a bifurcation to within 1% of a specified target) with respect to the number of parameters M , comparing against a gradient-free approach: Nelder-Mead. Calculations were performed on an Intel Core i7-6700HQ CPU @ 2.60GHz x 8 without GPU acceleration.

For the complexity study, a model was designed so that it is extensible both in the number of parameters M and the number of states N . There are many choices for this; we opted for a model of the form

$$\begin{cases} \partial_t u_1 = \sin^2 p - (\theta_1 \sin^2 p + 1)u_1 \\ \partial_t u_n = u_{n-1} - (\mu_n^2 + 1)u_n \end{cases} \quad 2 \leq n \leq N \quad (13)$$

In this model only the first state u_1 defines the shape of the bifurcation diagram, while the remaining states are merely linearly proportional to the first. The parameters μ_n contain sums of θ_m allowing us a flexible choice on the number of parameters while maintaining stable solutions for the bifurcation diagram.

While still tractable on laptop computers for states $N < 100$ our implementation currently does not scale well for partial differential equations where a large the number of states N arises from discretisation of the spatial variables. The only reason we need this determinant is because it

is an indicator of bifurcations. We can address the computational bottleneck by finding a more computationally efficient way of calculating this indicator. One approach would be to take the product of a finite subset of eigenvalues of the system. Note that any more efficient calculation must still permit back-propagation through it.

To demonstrate the benefits of the gradient-based aspect of our method we compare convergence times of gradient descent against a gradient-free approach. We use the Nelder-Mead method from `Optim.jl` [33] and obtain convergence times as the number of parameters M is increased (Figure 5B). We observe that for our method convergence times scale like \sqrt{M} compared to M^2 for the gradient-free approach.

4 Conclusion & Broader Impact

We proposed a gradient-based approach for inferring the parameters of differential equations that produce a user-specified bifurcation diagram. By applying implicit layers [22, 23] and the generalised Leibniz rule [34] to the geometry of the implicitly defined steady states [35] it is possible to use automatic differentiation methods to efficiently calculate gradients. We defined a bifurcation measure that uses the determinant of the state-space Jacobian as an indicator for bifurcating parameter regimes in the eigenvalue term of the cost function. The gradients of the cost can be efficiently computed using automatic differentiation methods. The computational bottleneck is the evaluation of the state-space Jacobian determinant which limits the implementation to ordinary differential equations.

We demonstrated our approach on models with one bifurcation parameter that can give rise to pitchforks and saddle-nodes. The estimated parameters form distinct clusters, allowing us to organise models in terms of topological and geometric equivalence (Figure 3). In the case of the genetic toggle switch (Figure 4) and a more complex model [3] (Figure D.1) we recovered mutual activation and inhibition regimes. In the more complex model we found a damped oscillatory regime that was not known about in the original paper.

Although we did not consider limit cycles, the bifurcation measure can be extended to detect Poincaré-Andronov-Hopf bifurcations alongside changes in stability of fixed points (see Supplementary E for details). This measure enables detection of the onset of damped oscillations and/or the emergence of limit cycles (Figure E.1). Used together with a steady state solver that detects periodic solutions and gradient-based optimisation, we can specify regions of damped oscillation and limit cycles. Our approach generalises naturally to bifurcation manifolds such as limit point curves or surfaces. This is because the normal components of implicit derivatives can still be calculated for under-determined systems of equations [28, 36, 37]. In the case of manifolds it would be more appropriate to use isosurface extraction algorithms rather than continuation to obtain the steady-state manifold. Our approach does not depend on the details of the steady-state solver and therefore can still be applied.

In dynamical systems theory the geometry of state-space determines all of the qualitative behaviours of a system. Our work makes progress towards designing models directly in state-space, rather than the spatial or temporal domain. This is valuable to experimentalists who only have qualitative observations available to them and wish to navigate the space of qualitative behaviours of their system. Our work lies within a trend of progress in the scientific machine learning community, where structured domain-informed models are favoured over flexible models that live in large parameter spaces.

5 Acknowledgements

We would like to acknowledge Kieran Cooney for the fruitful conversations that helped guide the derivations and computational approach. A special thanks go to Romain Veltz and the Julia community for helpful pointers on package development and discussions over Slack. This work was supported by Microsoft Research through its PhD Scholarship Programme and the EPSRC Centre for Doctoral Training in Cross-Disciplinary Approaches to Non-Equilibrium Systems (CANES, EP/L015854/1).

References

- [1] U. G. Abdulla, R. Poteau, A. Binder, H. W. Engl, C. Flamm, P. K. Å. Ugler, J. Lu, S. M. Å. Uller, and P. Schuster, “Inverse problems in systems biology,” *Inverse Problems*, vol. 25, p. 51, 2009.
- [2] J. J. Tyson, K. Chen, and B. Novak, “Network dynamics and cell physiology,” *Nature reviews Molecular cell biology*, vol. 2, no. 12, pp. 908–916, 2001.
- [3] P. Grant, G. Szep, O. Patange, J. Halatek, V. Coppard, A. Csikász-Nagy, J. Haseloff, J. Locke, N. Dalchau, and A. Phillips, “Interpretation of morphogen gradients by a synthetic bistable circuit,” *Nature Communications*, vol. 11, no. 1, 2020.
- [4] Y. A. Kuznetsov, “Topological Equivalence, Bifurcations, and Structural Stability of Dynamical Systems,” in *Elements of Applied Bifurcation Theory*, pp. 39–76, Springer New York, 2004.
- [5] M. P. H. Stumpf and E. Roesch, “Parameter inference in dynamical systems with co-dimension 1 bifurcations,” *Royal Society*, vol. 6, no. 10, 2019.
- [6] A. Csikász-Nagy, D. Battogtokh, K. C. Chen, B. Novák, and J. J. Tyson, “Analysis of a Generic Model of Eukaryotic Cell-Cycle Regulation,” *Biophysical Journal*, vol. 90, pp. 4361–4379, 6 2006.
- [7] K. Iwasaki and Y. Kamimura, “An inverse bifurcation problem and an integral equation of the Abel type,” *Inverse Problems*, vol. 13, pp. 1015–1031, 1997.
- [8] J. Lu, H. W. Engl, and P. Schuster, “Inverse bifurcation analysis: Application to simple gene systems,” *Algorithms for Molecular Biology*, vol. 1, pp. 1–16, 7 2006.
- [9] I. Dobson, “Distance to Bifurcation in Multidimensional Parameter Space: Margin Sensitivity and Closest Bifurcations,” in *Bifurcation Control: Theory and Applications*, pp. 49–66, Springer, Berlin, Heidelberg, 4 2004.
- [10] V. Chickarmane, S. R. Paladugu, F. Bergmann, and H. M. Sauro, “Bifurcation discovery tool,” *BIOINFORMATICS APPLICATIONS NOTE*, vol. 21, no. 18, pp. 3688–3690, 2005.
- [11] E. D. Conrad, J. Tyson, R. Laubenbacher, J. Phillips, and M. Renardy, “Bifurcation Analysis and Qualitative Optimization of Models in Molecular Cell Biology with Applications to the Circadian Clock,” *Virginia Tech*, 4 2006.
- [12] J. C. W. Locke, A. J. Millar, and M. S. Turner, “Modelling genetic networks with noisy and varied experimental data: the circadian clock in *Arabidopsis thaliana*,” *Journal of theoretical biology*, vol. 234, no. 3, pp. 383–393, 2005.
- [13] I. Otero-Muras and J. R. Banga, “Optimization-based prediction of fold bifurcations in nonlinear ODE models,” *IFAC-PapersOnLine*, vol. 51, pp. 485–490, 1 2018.
- [14] I. Otero-Muras, P. Yordanov, and J. Stelling, “A method for inverse bifurcation of biochemical switches: inferring parameters from dose response curves,” *BMC Systems Biology*, vol. 8, p. 114, 2014.
- [15] C. Rackauckas and Q. Nie, “Differentialequations.jl - a performant and feature-rich ecosystem for solving differential equations in julia,” *Journal of Open Research Software*, vol. 5, no. 1, 2017.
- [16] C. Rackauckas, Y. Ma, V. Dixit, X. Guo, M. Innes, J. Revels, J. Nyberg, and V. Ivaturi, “A comparison of automatic differentiation and continuous sensitivity analysis for derivatives of differential equation solutions,” *arXiv preprint arXiv:1812.01892*, 2018.
- [17] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, and A. Ramadhan, “Universal Differential Equations for Scientific Machine Learning,” *arXiv*, 2020.
- [18] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, “Neural Ordinary Differential Equations,” *NIPS*, vol. 109, pp. 31–60, 6 2018.

- [19] C. Rackauckas, M. Innes, Y. Ma, J. Bettencourt, L. White, and V. Dixit, “DiffEqFlux.jl - A Julia Library for Neural Differential Equations,” *ArXiv*, 2019.
- [20] S. Ranciati, C. Viroli, and E. Wit, “Bayesian Smooth-and-Match strategy for ordinary differential equations models that are linear in the parameters,” *ArXiv*, 2017.
- [21] F. Khadivar, I. Lauzana, and A. Billard, “Learning dynamical systems with bifurcations,” *Robotics and Autonomous Systems*, vol. 136, p. 103700, 2 2021.
- [22] A. Look, S. Doneva, M. Kandemir, R. Gemulla, and J. Peters, “Differentiable Implicit Layers,” *ArXiv*, 10 2020.
- [23] S. Bai, J. Z. Kolter, and V. Koltun, “Deep Equilibrium Models,” *arXiv*, 9 2019.
- [24] P. E. Farrell, C. H. L. Beentjes, and Å. Birkisson, “The computation of disconnected bifurcation diagrams,” *arXiv*, 3 2016.
- [25] R. Veltz, “BifurcationKit.jl,” tech. rep., Inria Sophia-Antipolis, 7 2020.
- [26] M. Haragus and G. Iooss, *Local Bifurcations, Center Manifolds, and Normal Forms in Infinite-Dimensional Dynamical Systems*. London: Springer London, 2011.
- [27] Z. Drmač and Z. Bujanović, “On the Failure of Rank-Revealing QR Factorization Software – A Case Study,” *ACM Transactions on Mathematical Software*, vol. 35, 7 2008.
- [28] S. Jos and R. Schmidt, “On the velocity of an implicit surface,” *ACM Transactions on Graphics*, vol. 30, pp. 1–7, 5 2011.
- [29] J. Revels, M. Lubin, and T. Papamarkou, “Forward-Mode Automatic Differentiation in Julia,” *ArXiv*, 2016.
- [30] M. Innes, E. Saba, K. Fischer, D. Gandhi, M. C. Rudilosso, N. M. Joy, T. Karmali, A. Pal, and V. Shah, “Fashionable Modelling with Flux,” *CoRR*, vol. abs/1811.01457, 2018.
- [31] M. Innes, “Flux: Elegant Machine Learning with Julia,” *Journal of Open Source Software*, 2018.
- [32] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *arXiv*, 12 2014.
- [33] P. K. Mogensen and A. N. Riseth, “Optim: A mathematical optimization package for Julia,” *Journal of Open Source Software*, vol. 3, p. 615, 4 2018.
- [34] H. Flanders, “Differentiation Under the Integral Sign,” *The American Mathematical Monthly*, vol. 80, p. 615, 6 1973.
- [35] R. Goldman, “Curvature formulas for implicit curves and surfaces,” in *Computer Aided Geometric Design*, vol. 22, pp. 632–658, Elsevier, 10 2005.
- [36] M. Tao, J. Solomon, and A. Butscher, “Near-Isometric Level Set Tracking,” *Computer Graphics Forum*, vol. 35, pp. 65–77, 8 2016.
- [37] M. Fujisawa, Y. Mandachi, and K. T. Miura, “Calculation of Velocity on an Implicit Surface by Curvature Invariance,” *Information and Media Technologies*, vol. 8, no. 4, pp. 674–680, 2013.
- [38] M. Kratochvíl, O. Hunewald, L. Heirendt, V. Veríssimo, J. Vondrášek, V. P. Satagopam, R. Schneider, C. Trefois, and M. Ollert, “GigaSOM.jl: High-performance clustering and visualization of huge cytometry datasets,” *GigaScience*, vol. 9, pp. 1–8, 11 2020.

4.8 Afterword

Chapter 5

Exploring bifurcations between phenotypes

Who knows what might happen to those who are consumed by greed

Ocarina of Time

5.1 Flow Cytometry

5.1.1 FlowJo

5.1.2 Immunophenotyping

5.1.3 Human Cell Atlas

5.2 Data Exploration

5.2.1 Dimensionality Reduction

5.2.2 Dataset Integration

5.2.3 Interactivity

5.3 FlowAtlas.jl

5.3.1 Julia & SciML

5.3.2 OpenLayers & Data-driven Documents

Chapter 6

Conclusions

*A childish mind will turn
to noble ambition*

Ocarina of Time

6.1 Limitations

6.2 Future work

1. How can thesis results be applied to biological computing [34]

Appendix A

Interpretation of Morphogen Gradients by a Bistable Circuit

Supplementary Information

Interpretation of morphogen gradients by a synthetic bistable circuit

Paul K. Grant¹, Gregory Szep^{1,2,9}, Om Patange^{3,7,8,9}, Jacob Halatek¹, Valerie Coppard¹, Attila Csikász-Nagy^{2,4}, Jim Haseloff⁵, James C. W. Locke^{1,3,6}, Neil Dalchau¹, and Andrew Phillips¹

¹Microsoft Research, 21 Station Road, Cambridge CB1 2FB, UK

²Randall Centre for Cell and Molecular Biophysics, King's College London WC2R 2LS, London, UK

³Sainsbury Laboratory, University of Cambridge, Cambridge CB2 1LR, UK

⁴Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, 1083 Budapest, Hungary

⁵Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK

⁶Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK

⁷Present address: Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA

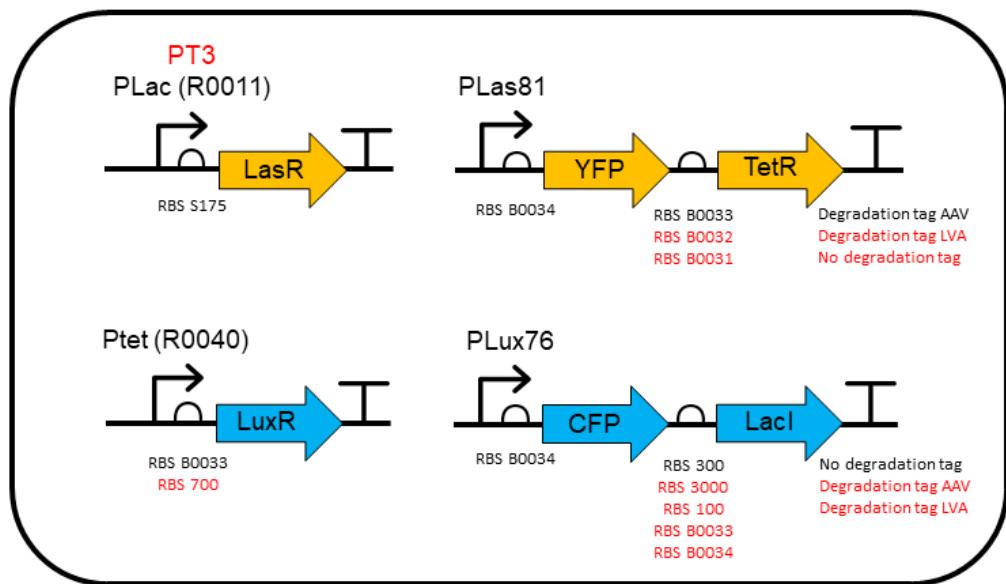
⁸Present address: Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

⁹These authors contributed equally: Gregory Szep, Om Patange

Contents

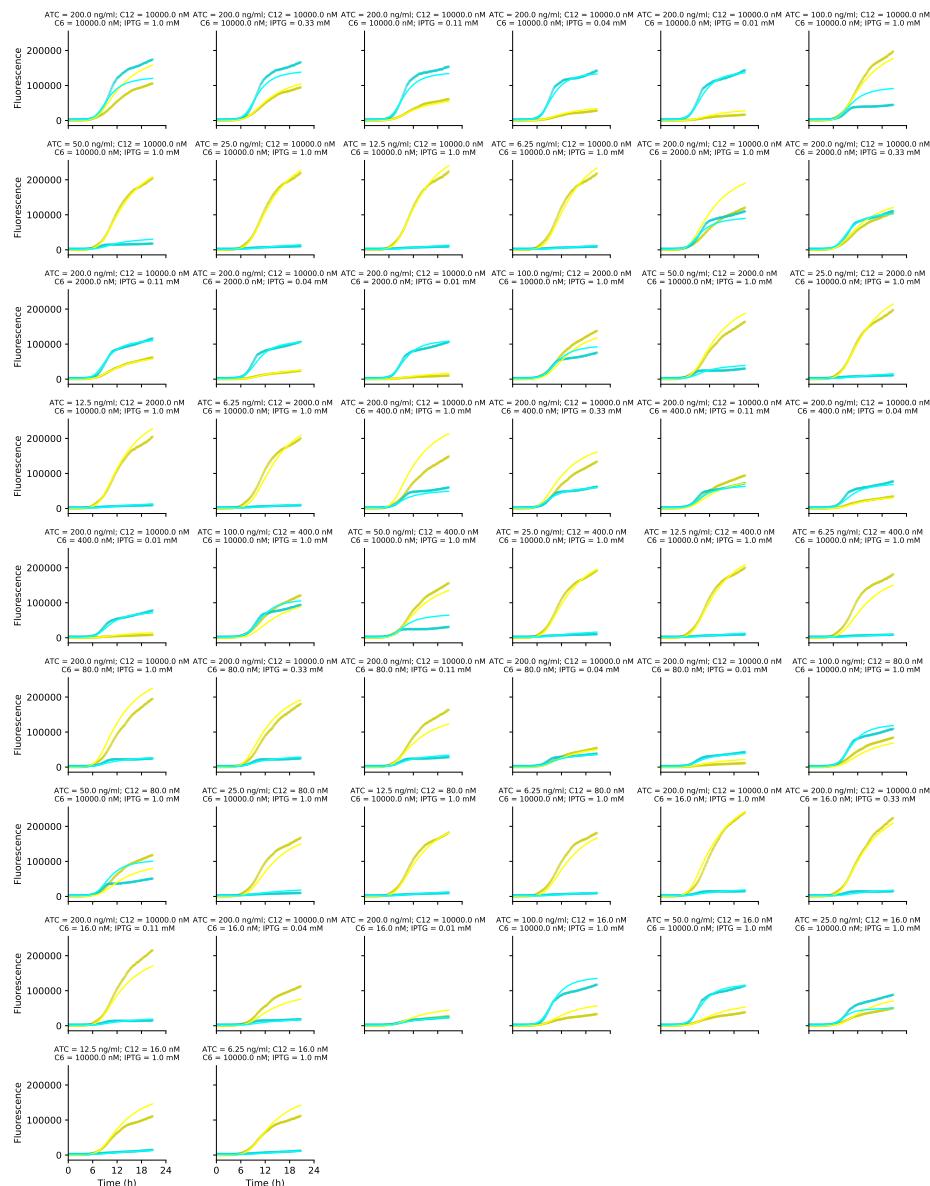
1 Supplementary Figures	2
2 Supplementary Methods	18
2.1 Differential Equation Models and Parameter Inference	19
2.2 Bistability Analysis	40
2.3 Boundary Experiments	42
2.4 Models of the Exclusive Receiver Relay Circuits	50

1 Supplementary Figures

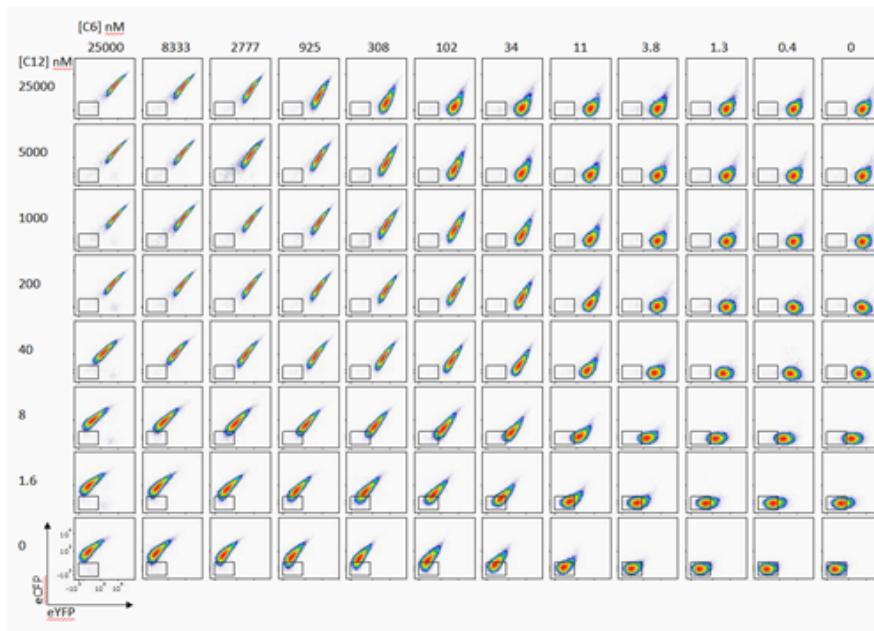
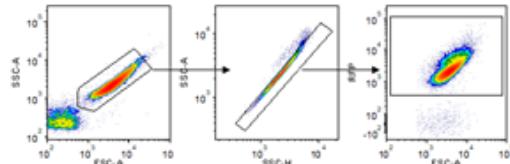


Supplementary Figure 1: Circuit variants. A circuit diagram of the exclusive receiver circuit with genetic parts labelled in black that were included in the final circuit and in red that were evaluated but discarded. Not all combinations of parts were tested.

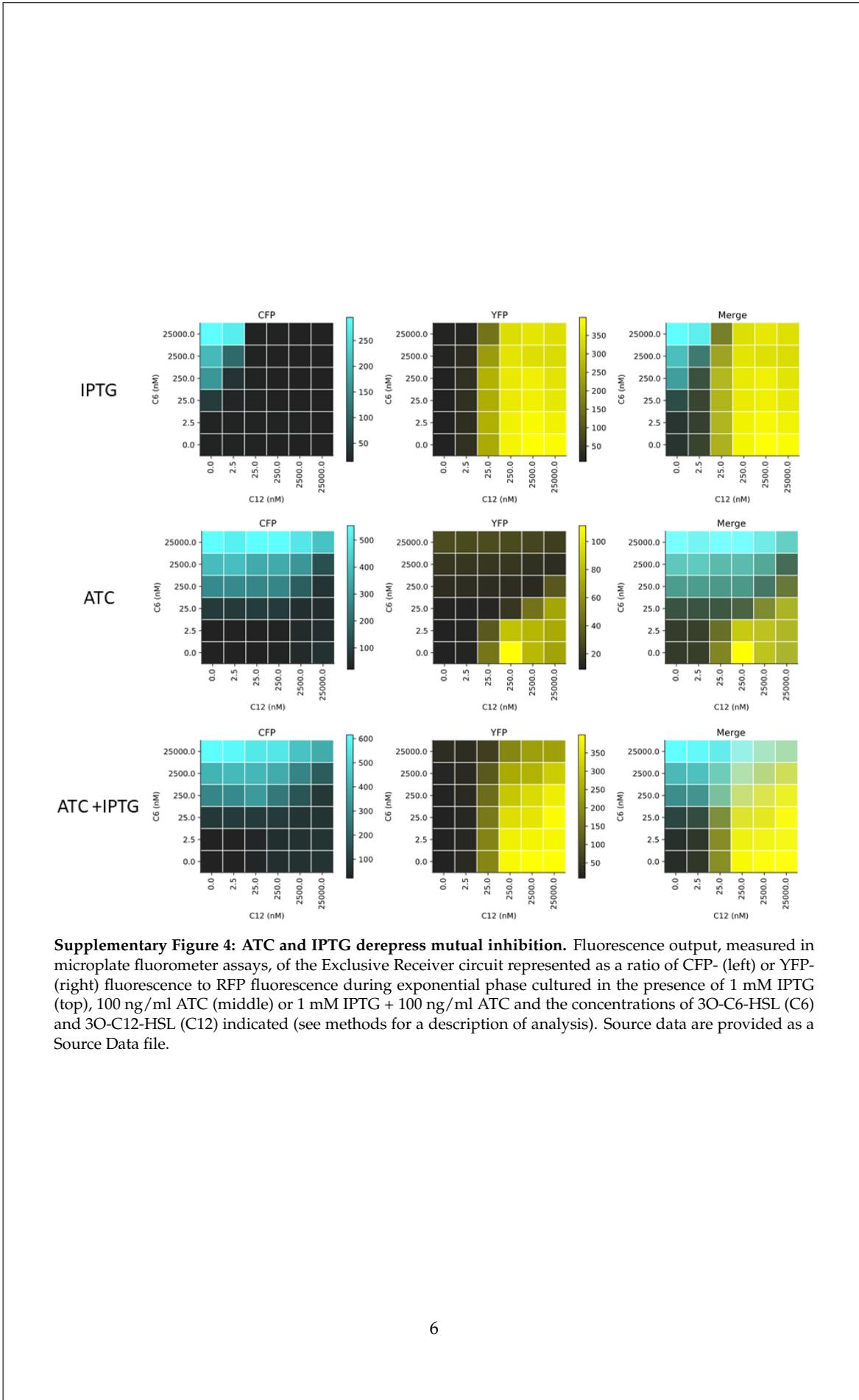


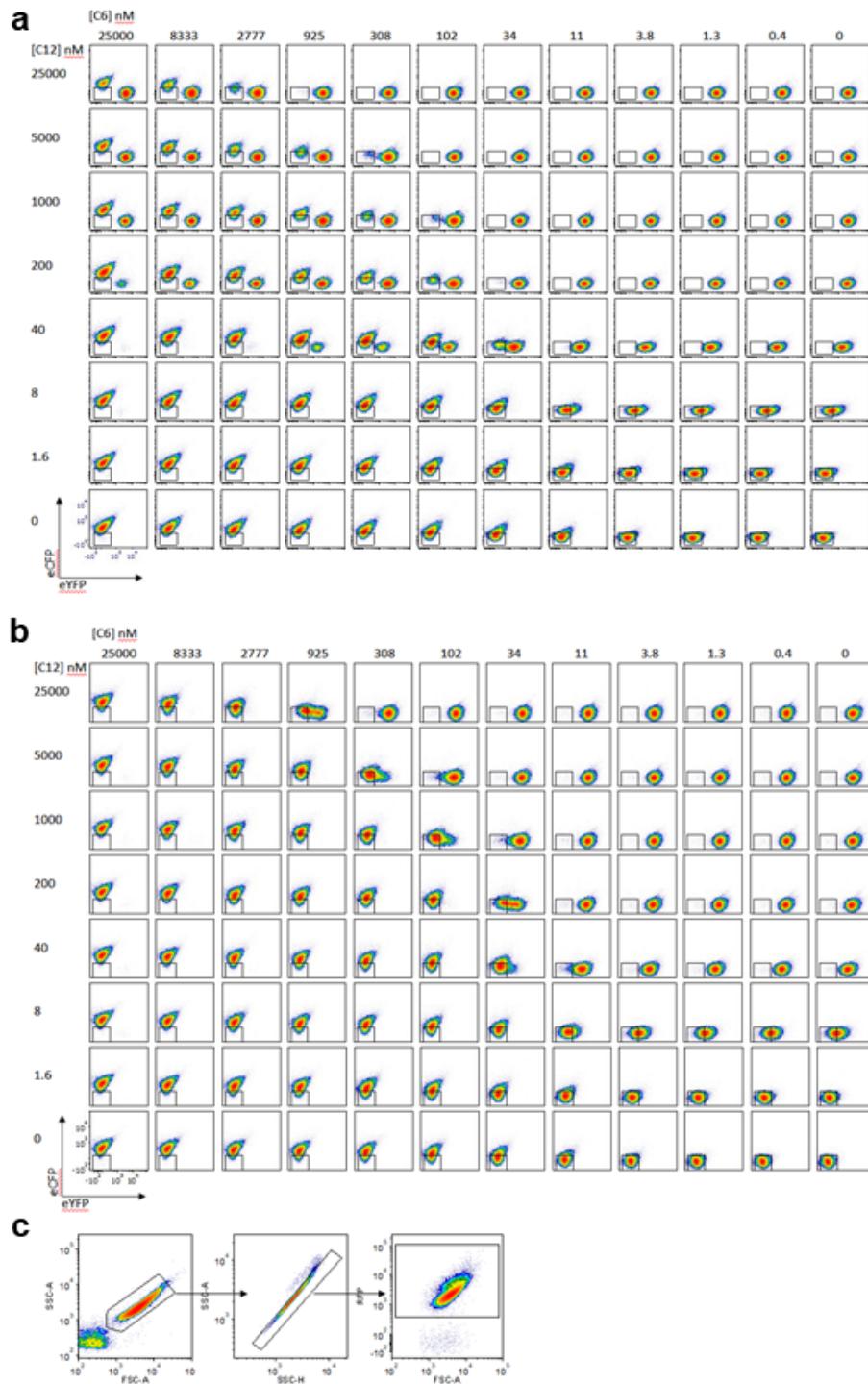


Supplementary Figure 2: Raw timecourse fluorescence traces. Fluorescence output, measured in microplate fluorometer assays, of the Exclusive Receiver circuit, plotted as CFP (dark cyan) and YFP (dark yellow) fluorescence against time, at the concentrations of C6, C12, ATC, and IPTG indicated. Model simulations (cyan and yellow) of the maximum likelihood parameters are overlaid. Source data are provided as a Source Data file.

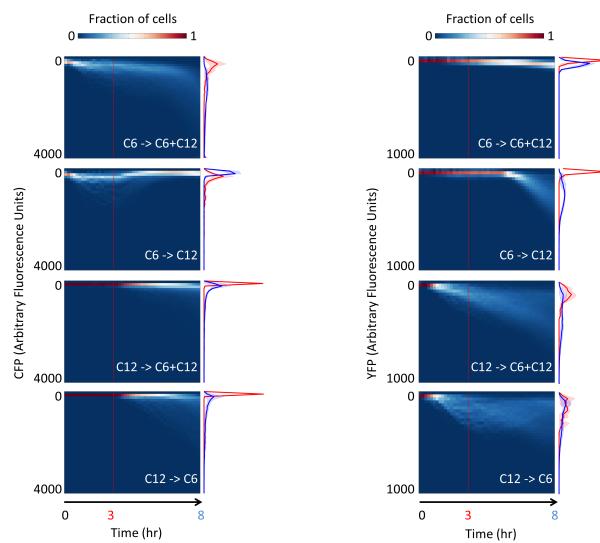
a**b****c**

Supplementary Figure 3: Flow cytometry measurements of naive cells a, Exclusive Receiver cells cultured in the indicated concentrations of C6 and C12. YFP fluorescence is plotted on the X-axis while CFP fluorescence is plotted on the Y-axis. b, Receiver cells cultured in the indicated concentrations of C6 and C12. Square indicates the position of untreated cells. c, Gating strategy. Cells also constitutively express mRFP1 via a genomic transgene. Only RFP⁺ cells were used for analysis and electronic noise, cell debris and cell clusters were excluded sequentially. Source data are provided as a Source Data file.

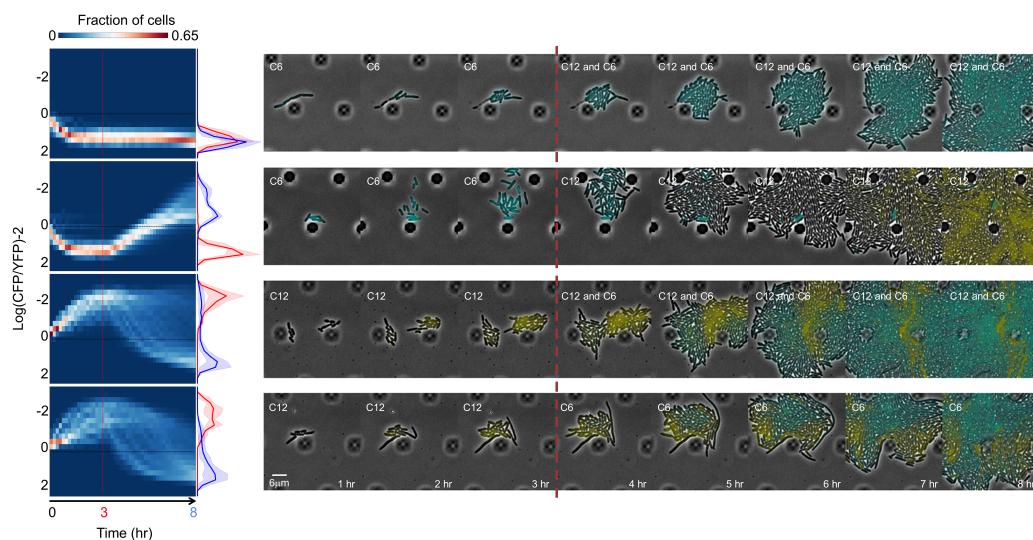




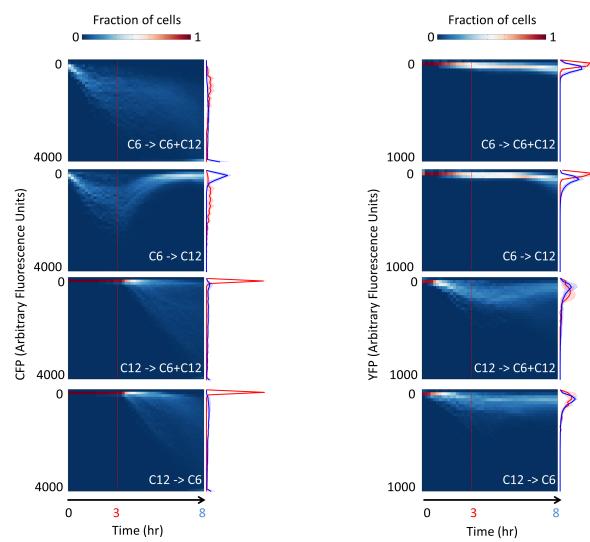
Supplementary Figure 5: Flow cytometry measurements of conditioned cells a, Cells conditioned in 500 nM C12 then exposed to the indicated concentrations of C6 and C12. YFP fluorescence is plotted on the X-axis while CFP fluorescence is plotted on the Y-axis. b, Cells conditioned in 500 nM C6 then exposed to the indicated concentrations of C6 and C12. Square indicates the position of untreated cells. c, Gating strategy. Cells also constitutively express mRFP1 via a genomic transgene. Only RFP⁺ cells were used for analysis and electronic noise, cell debris and cell clusters were excluded sequentially. Source data are provided as a Source Data file.



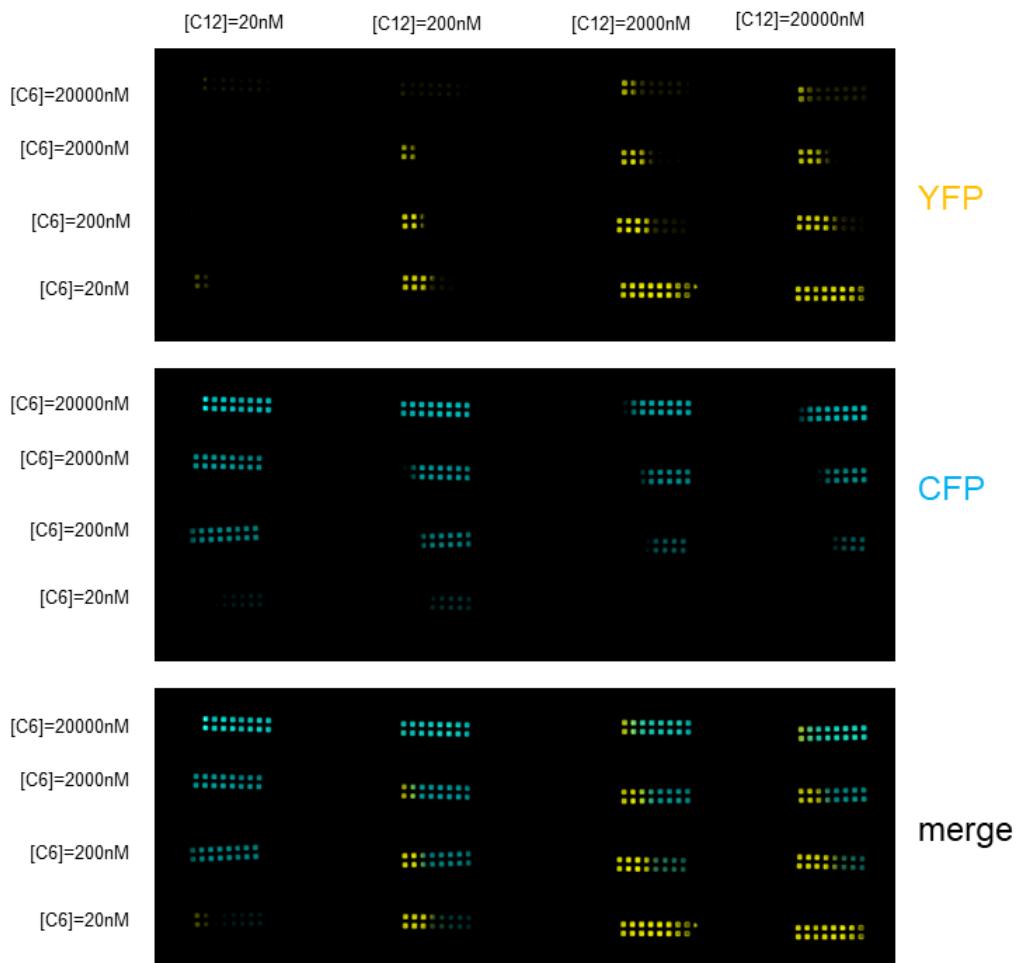
Supplementary Figure 6: Microfluidics measurements of bistability. Single-cell data used to compute ratios in Figure 2c. Cells were grown in microfluidic chips for 3 hours in the presence of either 37 nM C6 (rows 1 and 2) or 100 nM C12 (rows 3 and 4). Then media was changed to 100 nM C12 + 37 nM C6 (rows 1 and 3) or 100 nM C12 (row 2) or 37 nM C6 (row 4). Cells were imaged with a frame rate of (1 frame/10 minutes). Left panels in each column are kymographs of the CFP (left column) or YFP (right column) expression per-cell, and fraction of cells as a heat map. Histograms represent the populations at 3 hours (red) and 8 hours (blue). Lines and shaded region represent the mean and standard deviation, respectively, over $n = 4$ biological replicates performed over 4 different days.



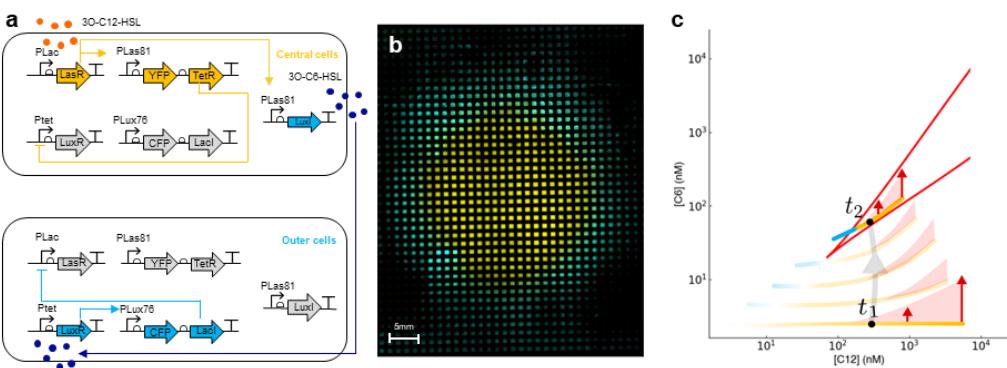
Supplementary Figure 7: Bistability and switching of single cells is robust to high C6 signal concentration. Cells were grown in microfluidic chips for 3 hours in the presence of either 1 μM C6 (rows 1 and 2) or 100 nM C12 (rows 3 and 4). Then media was changed to 100 nM C12 + 1 μM C6 (rows 1 and 3) or 100 nM C12 (row 2) or 1 μM C6 (row 4). Cells were imaged with a frame rate of (1 frame/10 minutes). Left panels are kymographs of the log-ratio of CFP expression per-cell to YFP expression per-cell, and fraction of cells as a heat map. Histograms represent the populations at 3 hours (red) and 8 hours (blue). Lines and shaded region represent the mean and standard deviation, respectively, $n = 4$ biological replicates performed over 4 different days. Right panels are sample montages of cells switching state (rows 3 and 4) or exhibiting bistability (rows 1 and 2); phase contrast and fluorescence channel ranges chosen for display.



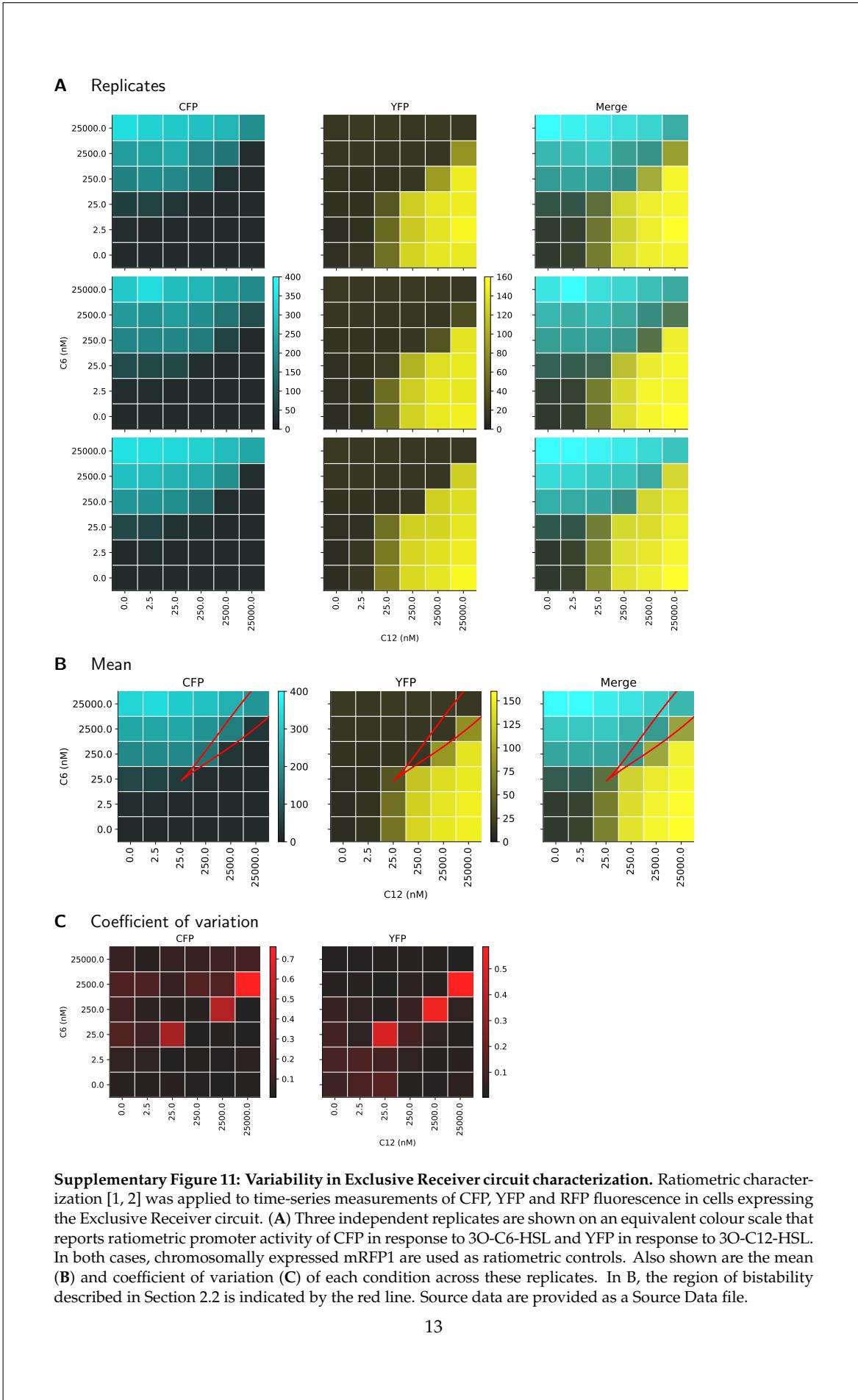
Supplementary Figure 8: Microfluidics measurements of bistability with high C6 signal concentration. Single-cell data used to compute ratios in 7. Cells were grown in microfluidic chips for 3 hours in the presence of either 1 μ M C6 (rows 1 and 2) or 100 nM C12 (rows 3 and 4). Then media was changed to 100 nM C12 + 1 μ M C6 (rows 1 and 3) or 100 nM C12 (row 2) or 1 μ M C6 (row 4). Cells were imaged with a frame rate of (1 frame/10 minutes). Left panels in each column are kymographs of the CFP (left column) or YFP (right column) expression per-cell, and fraction of cells as a heat map. Histograms represent the populations at 3 hours (red) and 8 hours (blue). Lines and shaded region represent the mean and standard deviation, respectively, over $n = 4$ biological replicates performed over 4 different days.

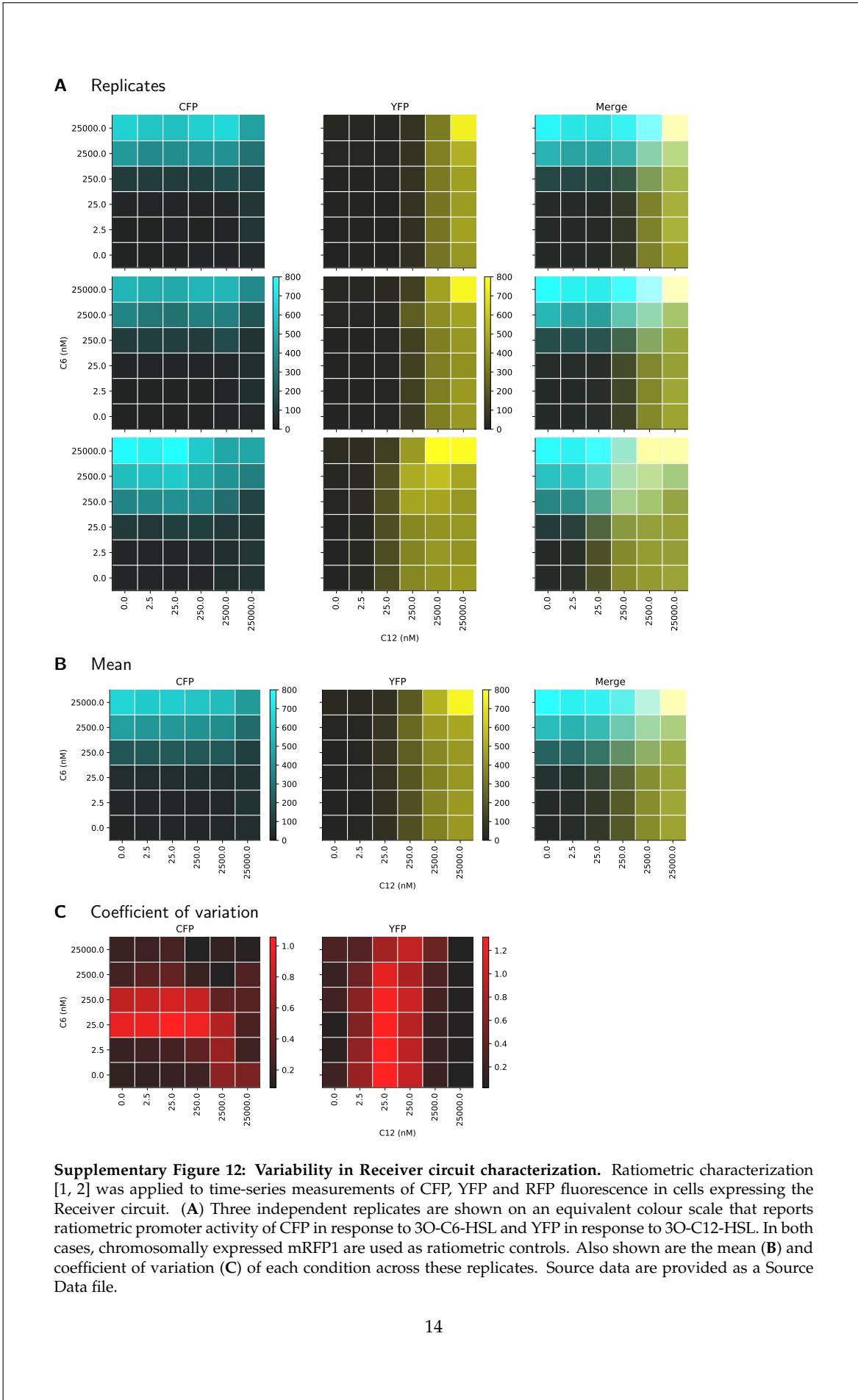


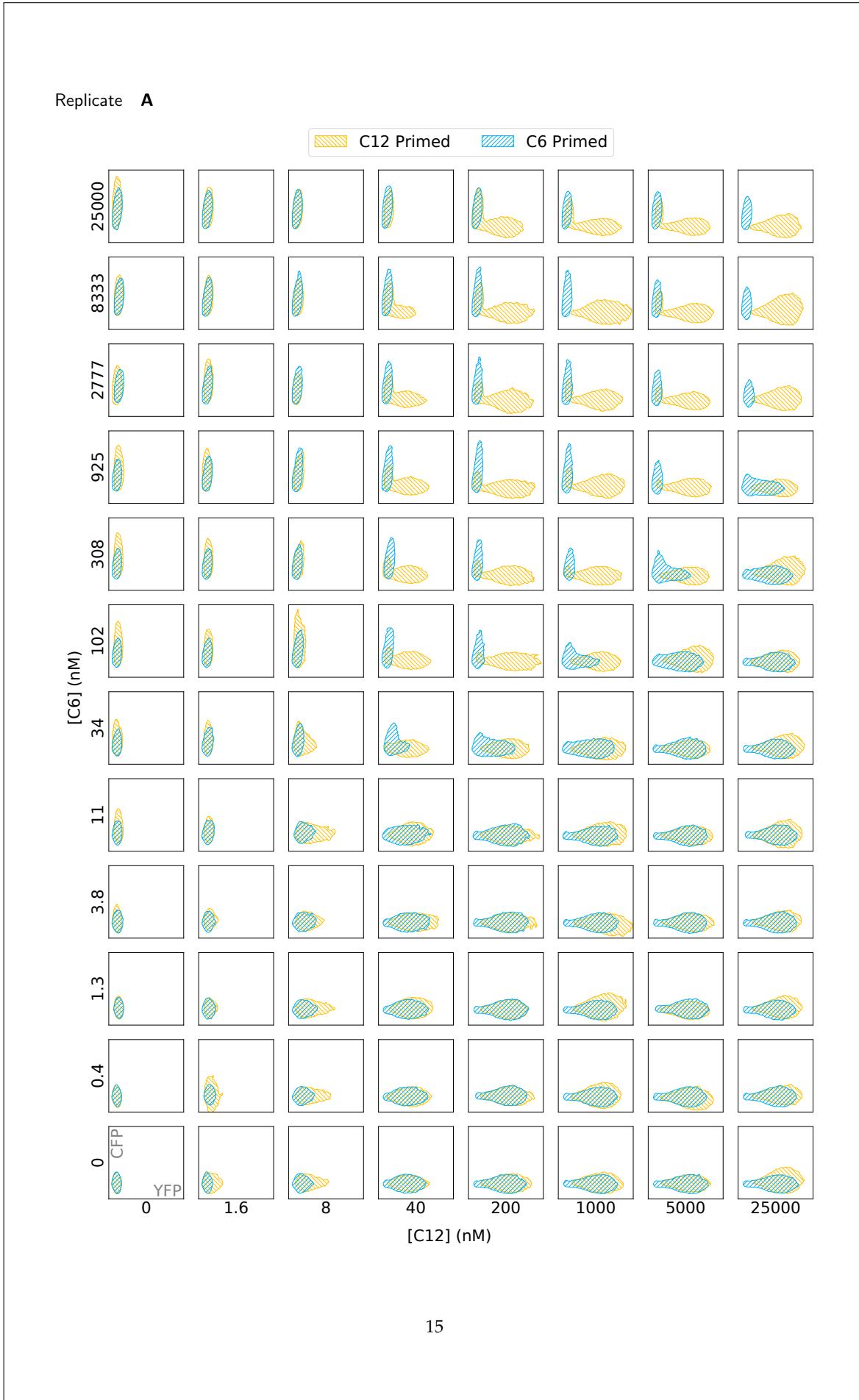
Supplementary Figure 9: Boundaries summarized in Figure 3c Endpoint fluorescence microscopy of Exclusive Receiver cells grown in transient gradients of signals (C12 diffusing from the left, C6 diffusing from the right) at the spatial average concentrations indicated and in the context of 10 μ M IPTG throughout.

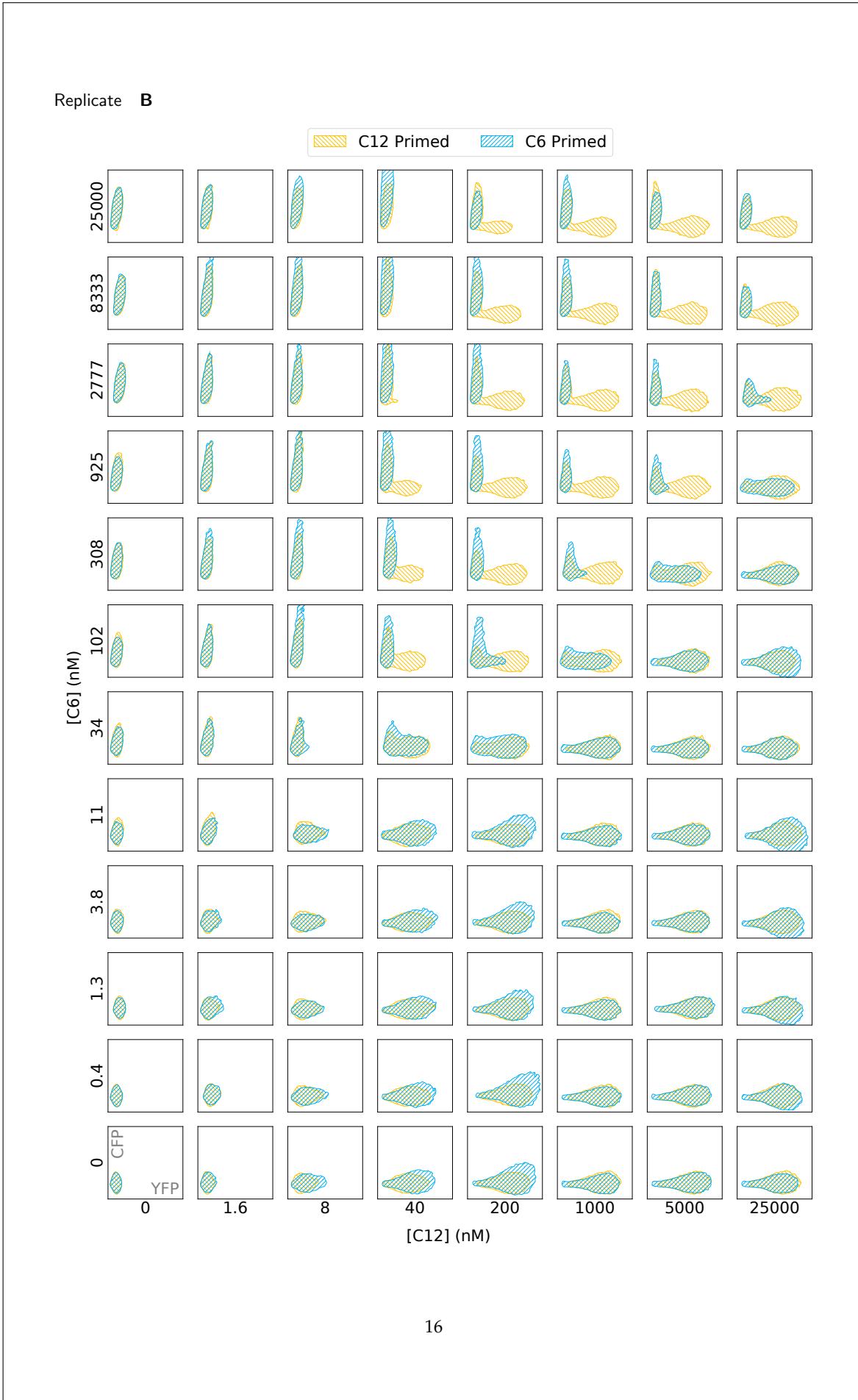


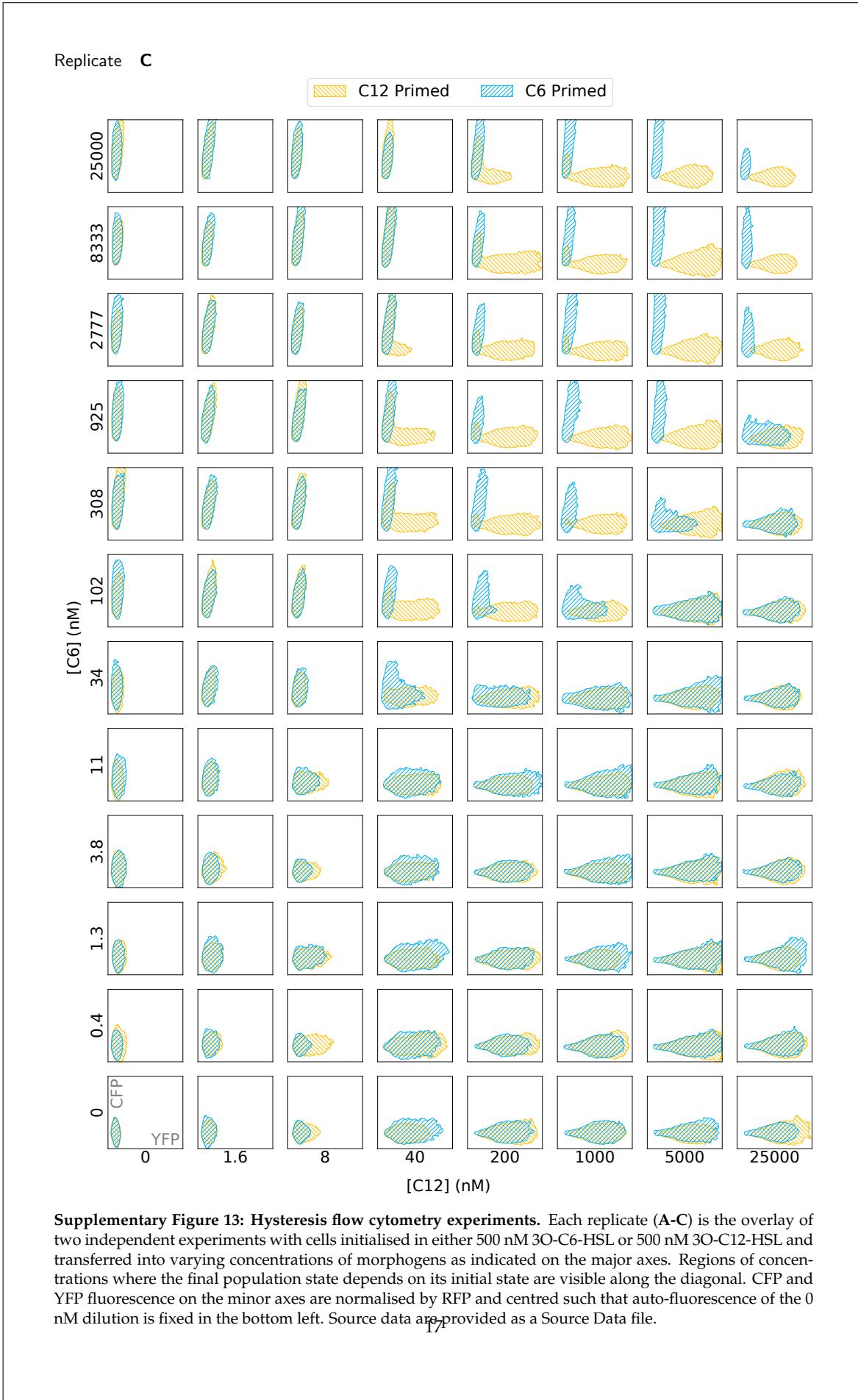
Supplementary Figure 10: Swapping primary and secondary morphogens also produces patterning. **a**, A circuit diagram of exclusive reporter cells co-transformed with the P81-LuxI relay device that responds to C12 by producing C6. Cells that experience high levels of C12 (central cells) will express YFP, TetR, and LuxI, causing them to produce C6 but be unable to sense it. Neighbouring cells (outer cells) that do not experience C12 will sense C6 and express CFP and LacI. **b**, Isogenic cells transformed with the circuit shown in (a) grown for 24 hours in the presence of a gradient of C12 diffusing from the centre express CFP and YFP in mutually exclusive domains of gene expression. **c**, A simulation in C6-C12 space over time (t_1 - t_2) labelling points in physical space by their CFP and YFP expression (cyan and yellow points), and showing the production of C6 as vectors (red arrows) that move the spatial average (black point) toward increasing C6. The bistable region is outlined in red.











2 Supplementary Methods

Supplementary Table 1: Primers used for Gibson assembly of the exclusive receiver circuit

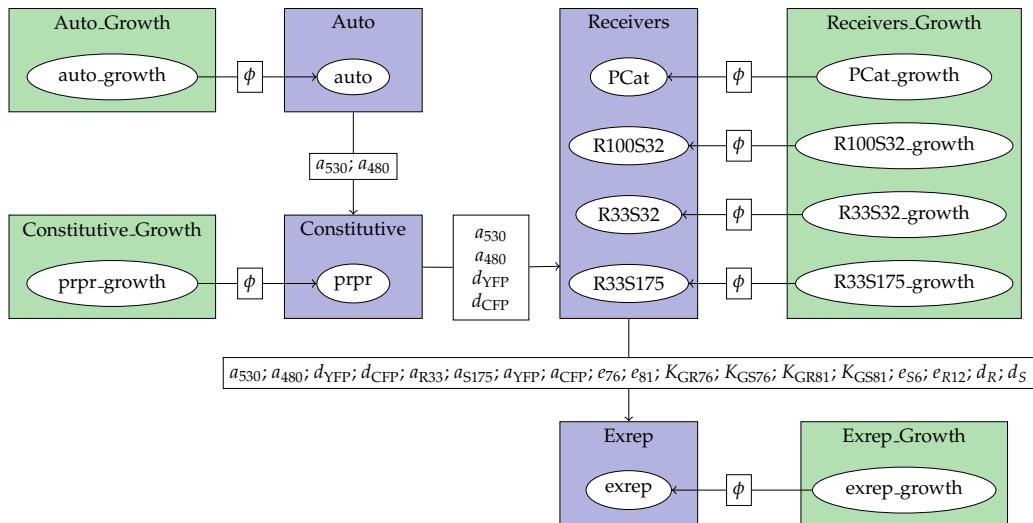
Primer name	Sequence
PGMSR014 LacIND F	TTTATTTGATGCCTGGCTTATTACTGCCCGCTTCCAGTCGG
PGMSR015 LacIND R	GCGGGCAGTAATAAAGCCAGGCATCAAATAAAACGAAAGG
PGMSR038 LacI300 F	TTTAGGACTGGACGGCGAAGGCCTGATGGAGTTCTAGTAT-
	TATTACTGTACAGCTCGTCCATGCC
PGMSR039 LacI300 R	AGAACTCCCATCAGGCCCTCGCCGTCCAGTCCTAAAATGGT-
	GAATGTGAAACCAGTAACGTTAT
PG081 Kan F	GGAATCGAACATGCAACCGGGC
PG082 Kan R	TGATGCGCTGGCAGTGTTC
PG342 Tet33 F	TCACACAGGACTACTAGATGTCCAGATTAGATAAAAG-
	TAAAGTGATTAACAGCGC
PG343 Tet33 R	ATCTAACATGGACATCTAGTAGTCCTGTGTGATTATTACTG-
	TACAGCTCGTCCATGCCG
PG373 TetAAV F	CGAAAAACTACGCTGCTGCTGTTAATAACCAGGCAT-
	CAAATAAAACGAAAGGC
PG374 TetAAV R	TGGTTATTAAACAGCAGCAGCGTAGTTTCGTCGTTGCTGCG-
	GACCCAC
JS3F Ori	GAGTGTATACTGGCTACTATGTTGGCACTG
JS3R Ori	CAGTGCCAACATAGTAAGCCAGTACACTC

2.1 Differential Equation Models and Parameter Inference

In this section, we derive ordinary differential equation (ODE) models for the reaction kinetics underlying the Exclusive Receiver circuit. These derivations broadly follow the derivations of the Receiver circuit in [3] and [4]. Importantly, we introduce differences in that original derivation that lead to changes in the location of bifurcations in (C12,C6) space, when those derivations are extended to incorporate dynamics of the repressor proteins LacI and TetR, and their chemical inhibitors IPTG and ATC.

Dynamic characterization with inference graphs

In order to infer the parameters of the Exclusive Receiver circuit, we adopt the strategy described in [4], evaluating parameters of sub-circuits first, and propagating their inferred values to larger circuits that embed those same parameters. This results in an inference graph, where we infer parameters over a sequence of models and corresponding datasets.



Supplementary Figure 14: Inference graph for characterizing the Exclusive Receiver circuit. Blue rectangular nodes are inference problems for collections of synthetic gene circuit models, compared with CFP and YFP measurements, while green rectangular nodes are inference problems for the growth models of the circuits in the downstream blue nodes, compared with OD₆₀₀ measurements. Internal to the coloured nodes are white elliptical nodes, which correspond to individual synthetic gene circuits and one or more associated dataset(s). White rectangular nodes are the sets of inferred parameters that are propagated between nodes. To simplify the notation, ϕ has been used to denote the set of growth model parameters, which are culture-specific (local) values for r , K and t_{lag} , and are propagated as maximum likelihood estimates. All other parameters are global (non-culture-specific) values and propagated as marginal posterior estimates.

The simplest possible cell line to characterize is one in which there is no synthetic gene circuit at all. Applying dynamic characterization in this context enables us to quantify autofluorescence, and so we name this circuit *auto*. Therefore, we measured cells under a range of conditions to explore how gene expression capacity influenced time-series measurements at fluorescence wavelengths corresponding to CFP and YFP. Subsequently, we characterize a circuit (*prpr*) in which CFP and YFP are driven by constitutive promoters (PR), enabling us to characterize the rates of degradation of the fluorescent proteins. Next, we use four variants of simple HSL receiver (PCat, R100S32, R33S32 and R33S175) to characterize the genetic parts associated with LuxR and LasR receiver proteins and their interactions with HSL molecules 3OC6HSL (C6) and 3OC12HSL (C12). Finally, having obtained parameter estimates for parts associated with CFP, YFP, LuxR and LasR, we characterize the Exclusive Receiver circuit (*exrep*) itself, establishing quantitative estimates for the parts associated with TetR and LacI repressor proteins.

In the following, we introduce the models for each of the circuits just mentioned, define their parameters and present results of the inference.

Autofluorescence model

The model we used for autofluorescence assumes that the rate of autofluorescence is constant, and that the fluorescent material dilutes with cell growth. As such, the equations for intracellular autofluorescence corresponding to CFP and YFP are

$$\frac{d\rho}{dt} = \gamma(\rho) \cdot \rho \quad (1a)$$

$$\frac{dc_{480}}{dt} = a_{480} - \gamma(\rho) \cdot c_{480} \quad (1b)$$

$$\frac{dc_{530}}{dt} = a_{530} - \gamma(\rho) \cdot c_{530} \quad (1c)$$

where $\gamma(\rho) = r(1 - \frac{\rho}{K})$ is the specific growth rate of the cell culture with density ρ .

To compare with experimental measurements, we consider the bulk fluorescence given by

$$B_{480} = \rho \cdot c_{480} + b_{480} \quad (2a)$$

$$B_{530} = \rho \cdot c_{530} + b_{530} \quad (2b)$$

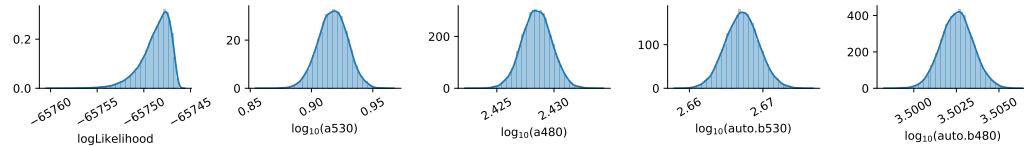
where c_{480} and c_{530} are modelled as in (1). Here, the quantities b_{480} and b_{530} represent background fluorescence at 480 nm and 530 nm, corresponding to CFP and YFP respectively.

Inference. We use the data in [4] to infer the parameters of the auto circuit. Cells were treated with EtOH to perturb cell growth, enabling us to determine how autofluorescence changes with different cellular growth rates. The priors used are detailed in the following table.

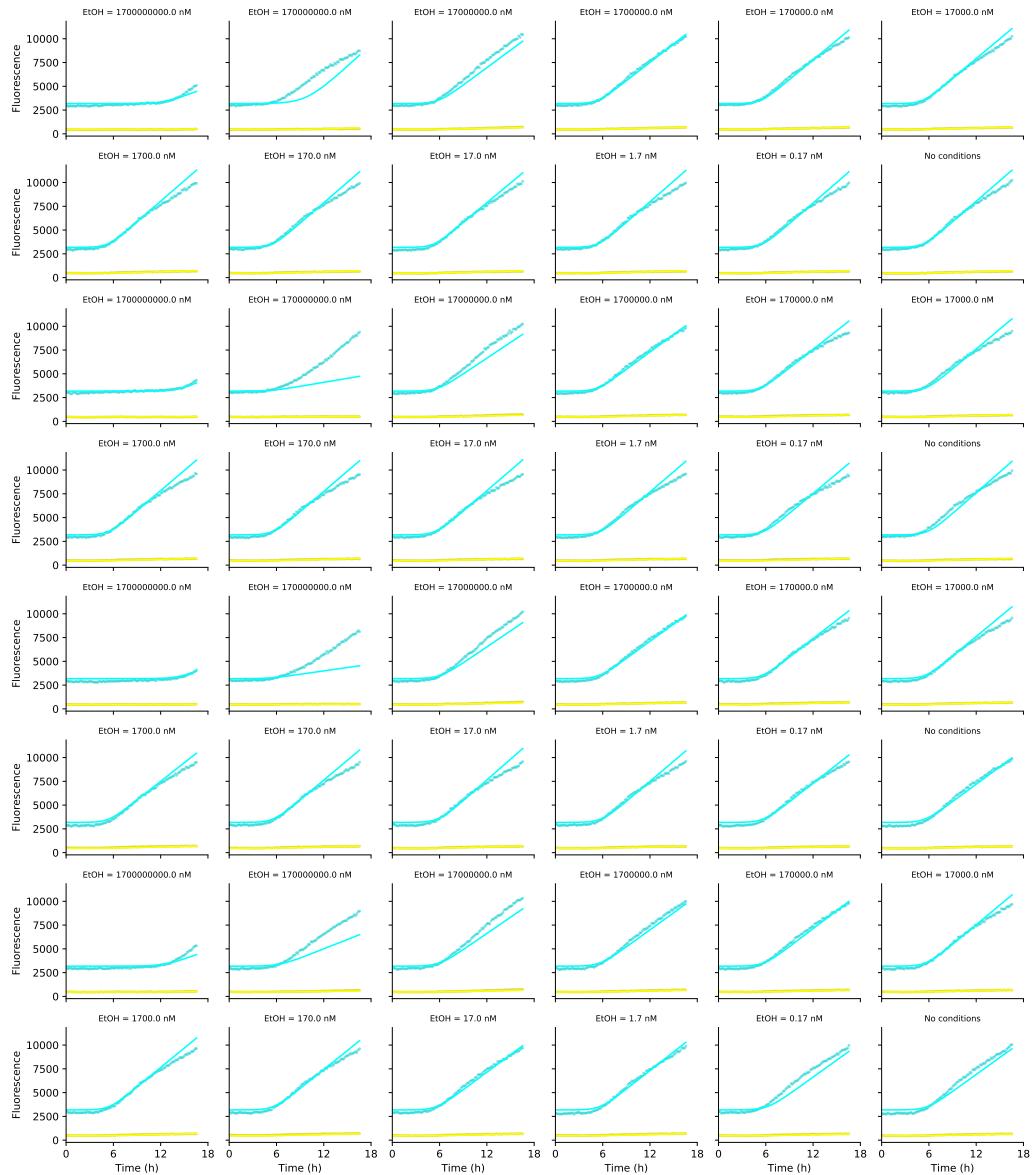
Supplementary Table 2: Priors for the auto circuit characterization. In the unit column, Fl denotes fluorescence units. The scaling column refers to the proposal distribution used to generate new parameters during MCMC.

Parameter	Description	Unit	Distribution	Scaling
a_{480}	Autofluorescence (CFP)	Fl.cell ⁻¹ .h ⁻¹	$U(10^{-3}, 10^3)$	Log
a_{530}	Autofluorescence (YFP)	Fl.cell ⁻¹ .h ⁻¹	$U(10^{-3}, 10^3)$	Log
b_{480}	Background fluor. (CFP)	Fl	$U(0, 10^4)$	Real
b_{530}	Background fluor. (YFP)	Fl	$U(0, 5 \times 10^3)$	Real

The marginal posterior estimates are shown in Figure 15. Simulation of the maximum likelihood estimate is shown in Figure 16.



Supplementary Figure 15: Marginal parameter posterior estimates of the auto model parameters. The marginal distributions are computed from 20 independent MCMC chains.



Supplementary Figure 16: Comparison of auto model with fluorescence measurements. Simulations are for B_{480} (cyan lines) and B_{530} (yellow lines), evaluated with the maximum likelihood parameter set. Measurements of bulk culture fluorescence are shown for the emission wavelengths corresponding to CFP (cyan circles) and YFP (yellow circles). EtOH treatment concentrations are indicated atop each panel.

Constitutive (prpr) model

The prpr circuit described above uses the constitutive PR promoter to drive CFP and YFP expression, in two separate operons. Following the derivation in [4], we arrive at a system of equations that describe the time-evolution of the intracellular concentrations of CFP and YFP as

$$\frac{dc}{dt} = \gamma(\rho).c \quad (3a)$$

$$\frac{dc_{\text{CFP}}}{dt} = a_{\text{CFP}} - (d_{\text{CFP}} + \gamma(\rho))c_{\text{CFP}} \quad (3b)$$

$$\frac{dc_{\text{YFP}}}{dt} = a_{\text{YFP}} - (d_{\text{YFP}} + \gamma(\rho))c_{\text{YFP}} \quad (3c)$$

where a_{CFP} and a_{YFP} are aggregated parameters that incorporate the rate of transcription and translation of CFP and YFP respectively. To compare with bulk culture fluorescence data, we use the observer model

$$B_{480} = \rho.(c_{\text{CFP}} + c_{480}) + b_{480} \quad (4a)$$

$$B_{530} = \rho.(c_{\text{YFP}} + c_{530}) + b_{530} \quad (4b)$$

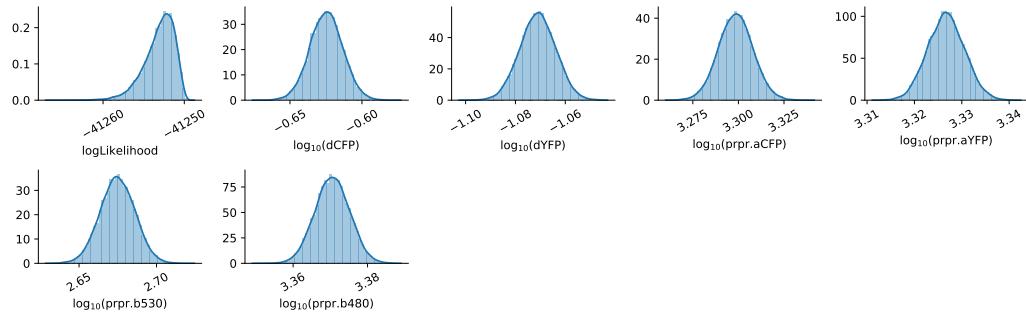
where the dynamics of c_{480} and c_{530} are governed by equations (1) above.

Inference. We use the data in [4] to infer the parameters of the prpr circuit. Cells were treated with chloramphenicol to perturb cell growth, enabling us to establish how constitutively expressed proteins are altered with different cellular growth rates. While the model above does not explicitly describe any explicit functional response to chloramphenicol, our general strategy of allowing the cell growth parameters to vary across different measurements enables the effect of chloramphenicol on growth rates to be implicitly captured. The quantification of autofluorescence was reused from the auto circuit (upstream in the inference graph), but the background fluorescence parameters were re-inferred. The priors used are detailed in Table 3. The marginal posterior estimates are shown in Figure 17.

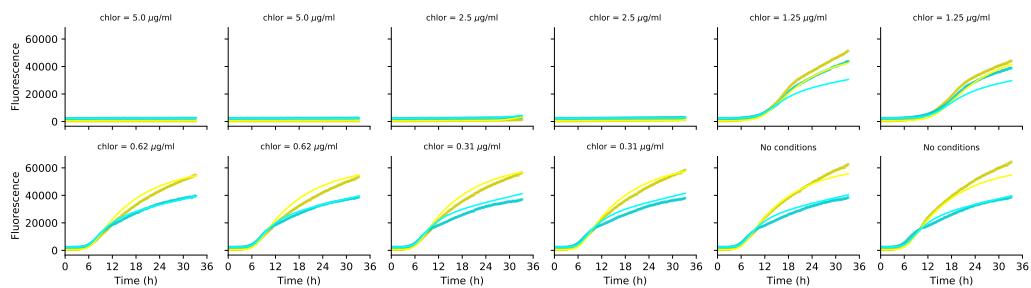
Supplementary Table 3: Priors for the prpr circuit characterization. In the unit column, Fl denotes fluorescence units. The scaling column refers to the proposal distribution used to generate new parameters during MCMC.

Parameter	Description	Unit	Distribution	Scaling
d_{CFP}	CFP degradation	h^{-1}	$U(10^{-3}, 10^0)$	Log
d_{YFP}	YFP degradation	h^{-1}	$U(10^{-3}, 10^0)$	Log
a_{CFP}	CFP synthesis	$\text{Fl.cell}^{-1}.\text{h}^{-1}$	$U(10^0, 10^5)$	Log
a_{YFP}	YFP synthesis	$\text{Fl.cell}^{-1}.\text{h}^{-1}$	$U(10^0, 10^5)$	Log
b_{480}	Background fluor. (CFP)	Fl	$U(0, 10^4)$	Real
b_{530}	Background fluor. (YFP)	Fl	$U(0, 5 \times 10^3)$	Real

Simulation of the maximum likelihood estimate is shown in Figure 18. The simulated CFP and YFP largely agree with the measured fluorescence at the culture level. As the effect of chloramphenicol is not explicitly modelled here, this comparison indicates that the majority of the effect of chloramphenicol can be described via its effect on cell growth. Any additional direct effect on CFP and YFP expression directly, is relatively minor.



Supplementary Figure 17: Marginal parameter posterior estimates of the prpr model parameters. The marginal distributions are computed from 20 independent MCMC chains.



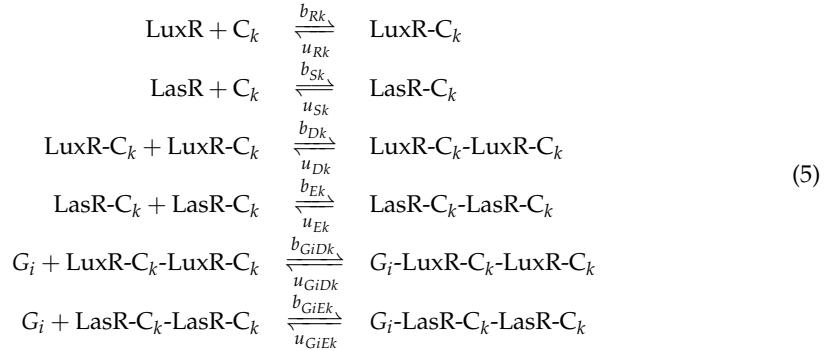
Supplementary Figure 18: Comparison of prpr model with fluorescence measurements. Simulations are of bulk CFP (cyan lines) and YFP (yellow lines) fluorescence, evaluated with the maximum likelihood parameter set. Measurements of bulk culture fluorescence are shown in for CFP (cyan circles) and YFP (yellow circles). Chloramphenicol treatment concentrations are indicated atop each panel.

Receiver model

We consider the dynamic characterization of the HSL Receiver circuit introduced in [3] and modelled dynamically in [4, 5]. In this circuit, which we refer to in the main text as the Receiver, two variations of the wild-type PLux promoter, PLux76 and PLas81, were engineered to bind preferentially to activated LuxR and LasR complexes respectively. As LuxR favours binding of C6 and LasR favours binding of C12, optimized expression of LuxR and LasR can lead to near-orthogonal intracellular detection of C6 and C12. The Receiver device was originally measured with PLux76 upstream of the coding sequence for CFP, and PLas81 upstream of the coding sequence for YFP.

Version 1 - Uniform degradation. The first version of the model we introduce is derived in [4]. It is based on the assumption that all degradation processes are of the same order, e.g. LuxR monomers are degraded at a similar rate as LuxR Dimers bound to HSLs. Subsequently, we will introduce an alternative derivation where we assume that complexes are protected from degradation, i.e. that degradation mainly occurs on the monomer level. We start by repeating some of the derivation from [4].

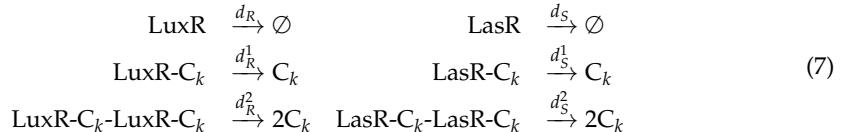
We denote by C_k the HSL molecule with length k carbon chain, and by G_i the PLux76 and PLas81 promoters. Then similar to the derivation in [3], we can specify all of the reactions between the HSLs, LuxR and LasR, and eventual binding of transcriptional regulators to PLux76/PLas81.



Constitutive expression of LuxR and LasR is described by

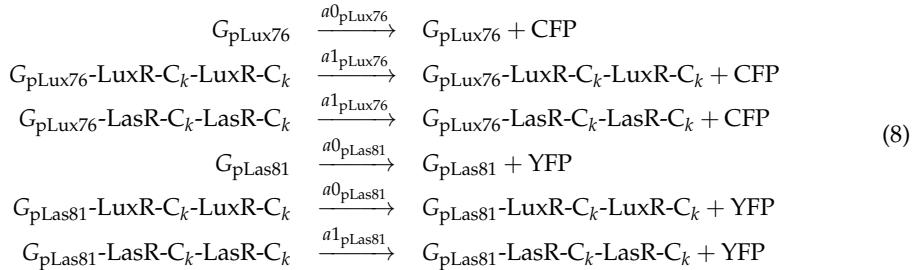


Degradation of LuxR and LasR is described by



where d_R^1 and d_R^2 are distinguishable from d_R to describe the effect of HSL molecules protecting receiver proteins from degradation.

Inducible expression of CFP and YFP by P_{OLux} and P_{OLas} respectively is described by



To produce a simplified ODE model amenable to parameter inference, we start with the equations describing LuxR and LasR protein, their complexes involving C₆ and C₁₂, and the bound/unbound promoters. Crucially, in this first derivation, we make a *rapid equilibrium* assumption for the binding reactions (5), and obtain the following relationships

$$[G_i\text{-LuxR-C}_k\text{-LuxR-C}_k]^* = K_{GDk}[G_i][\text{LuxR-C}_k\text{-LuxR-C}_k] \quad (9a)$$

$$[\text{LuxR-C}_k\text{-LuxR-C}_k]^* = K_{Dk}[\text{LuxR-C}_k]^2 \quad (9b)$$

$$[\text{LuxR-C}_k]^* = K_{Rk}[c_R][C_k] \quad (9c)$$

where $K_{Rk} = \frac{b_{Rk}}{\gamma + u_{Rk}}$, $K_{Dk} = \frac{b_{Dk}}{u_{Dk}}$ and $K_{GiDk} = \frac{b_{GiDk}}{u_{GiDk}}$. Therefore (also symmetry of LuxR and LasR),

$$[G_i\text{-LuxR-C}_k\text{-LuxR-C}_k]^* = K_{GiDk}K_{Dk}(K_{Rk}[C_k][c_R])^2 \quad (10a)$$

$$[G_i\text{-LasR-C}_k\text{-LasR-C}_k]^* = K_{GiEk}K_{Ek}(K_{Sk}[C_k][c_S])^2 \quad (10b)$$

where the new K's are defined as above.

To reduce the system to fewer variables, we consider the evolution of total LuxR, and seek to co-ordinate this with the rapid equilibrium relationships above. By denoting the total concentration of LuxR as c_R, we can write down its time-evolution as

$$\frac{d[R]_T}{dt} = \frac{d[c_R]}{dt} + \sum_k \left(\frac{d[\text{LuxR-C}_k]}{dt} + 2 \frac{d[\text{LuxR-C}_k\text{-LuxR-C}_k]}{dt} + 2 \sum_i \frac{d[G_i\text{-LuxR-C}_k\text{-LuxR-C}_k]}{dt} \right) \quad (11a)$$

$$= a_R - \gamma[c_R] - d_R[c_R] \dots \\ - \sum_k \left((\gamma + d_R^1)[\text{LuxR-C}_k] + (\gamma + d_R^2)[\text{LuxR-C}_k\text{-LuxR-C}_k] + \gamma \sum_i [G_i\text{-LuxR-C}_k\text{-LuxR-C}_k] \right) \quad (11b)$$

If there are differences between d_R, d_R¹ and d_R², then further reduction is complicated. However, by assuming that HSL is not protective of receiver protein (d_R¹ = d_R² = d_R), we obtain the simplification

$$\frac{dc_R}{dt} = a_R - (\gamma + d_R)c_R \quad (12)$$

Now using a conservation relationship for LuxR, we can obtain

$$c_R = [\text{LuxR}] + \sum_k \left([\text{LuxR-C}_k] + 2[\text{LuxR-C}_k\text{-LuxR-C}_k] + 2 \sum_i [G_i\text{-LuxR-C}_k\text{-LuxR-C}_k] \right) \\ = [\text{LuxR}] + \sum_k \left(K_{Rk}[\text{LuxR}][C_k] + 2K_{Dk}K_{Rk}^2[\text{LuxR}]^2[C_k]^2 + \sum_i 2K_{GiDk}[G_i]K_{Dk}K_{Rk}^2[\text{LuxR}]^2[C_k]^2 \right)$$

When C_k is low, total LuxR is closely approximated by free LuxR, c_R ≈ [LuxR]. But when C_k is high, c_R should be partitioned between the [LuxR-C_k-LuxR-C_k] and [G_i-LuxR-C_k-LuxR-C_k] species. Therefore, to simplify the analysis, we propose modelling this by using the assumption

$$c_R \approx [\text{LuxR}] + \sum_k [\text{LuxR-C}_k] = [\text{LuxR}] \left(1 + \sum_k K_{Rk}[C_k] \right) \quad (13)$$

which still captures the saturation of LuxR by C_k , using the approximations

$$[\text{LuxR}] \approx c_R \cdot \frac{1}{1 + \sum_k K_{Rk}[C_k]} \quad (14\text{a})$$

$$[\text{LuxR-C}_k] \approx c_R \cdot \frac{K_{Rk}[C_k]}{1 + \sum_k K_{Rk}[C_k]} \quad (14\text{b})$$

$$[\text{LuxR-C}_k\text{-LuxR-C}_k] \approx K_{Dk}c_R^2 \left(\frac{K_{Rk}[C_k]}{1 + \sum_k K_{Rk}[C_k]} \right)^2 \quad (14\text{c})$$

$$[G_i\text{-LuxR-C}_k\text{-LuxR-C}_k] \approx K_{GR}^{(i)}[G_i]c_R^2 \left(\frac{K_{Rk}[C_k]}{1 + \sum_k K_{Rk}[C_k]} \right)^2 \quad (14\text{d})$$

where $K_{GR}^{(i)} = K_{GiDk}K_{Dk}$ is assumed to be independent of which signal is bound (k), equivalent to the derivation in [3]. By symmetry, we immediately obtain equivalent expressions for interactions between LasR, HSL and PLux promoters. We denote the total concentration of LasR as c_S .

In addition to the saturation of LuxR and LasR, our reduced model also allows for saturation of G_i . By taking advantage of the conservation law

$$\begin{aligned} N_i &= [G_i] + [G_i\text{-LuxR-C}_6\text{-LuxR-C}_6] + [G_i\text{-LuxR-C}_{12}\text{-LuxR-C}_{12}] + \dots \\ &\quad [G_i\text{-LasR-C}_6\text{-LasR-C}_6] + [G_i\text{-LasR-C}_{12}\text{-LasR-C}_{12}], \end{aligned}$$

we can derive the rate of production of mRNA as a function of c_R , c_S [C_6] and [C_{12}]. For notational convenience we write $c_6 := [C_6]$ and $c_{12} := [C_{12}]$. Then,

$$P_i(c_6, c_{12}, c_R, c_S) = \frac{\epsilon^{(i)} + K_{GR}^{(i)}B_R + K_{GS}^{(i)}B_S}{1 + K_{GR}^{(i)}B_R + K_{GS}^{(i)}B_S} \quad (15)$$

where the fractions of bound LuxR and LasR are defined by

$$B_R := c_R^2 \left(\frac{K_{Rk}c_k}{1 + \sum_k K_{Rk}c_k} \right)^{n_R}, \quad B_S := c_S^2 \left(\frac{K_{Sk}c_k}{1 + \sum_k K_{Sk}c_k} \right)^{n_S} \quad (16)$$

Here, we have introduce alternative exponents n_R and n_S , analogous to the usage of n in [3]. Accordingly, we obtain the following system of equations

$$\frac{d\rho}{dt} = \gamma(\rho) \cdot \rho \quad (17\text{a})$$

$$\frac{dc_R}{dt} = a_R - (d_R + \gamma(\rho)) \cdot c_R \quad (17\text{b})$$

$$\frac{dc_S}{dt} = a_S - (d_S + \gamma(\rho)) \cdot c_S \quad (17\text{c})$$

$$\frac{dc_{\text{CFP}}}{dt} = a_{\text{CFP}} \cdot P_{76}(c_6, c_{12}, c_R, c_S) - (d_{\text{CFP}} + \gamma(\rho)) \cdot c_{\text{CFP}} \quad (17\text{d})$$

$$\frac{dc_{\text{YFP}}}{dt} = a_{\text{YFP}} \cdot P_{81}(c_6, c_{12}, c_R, c_S) - (d_{\text{YFP}} + \gamma(\rho)) \cdot c_{\text{YFP}} \quad (17\text{e})$$

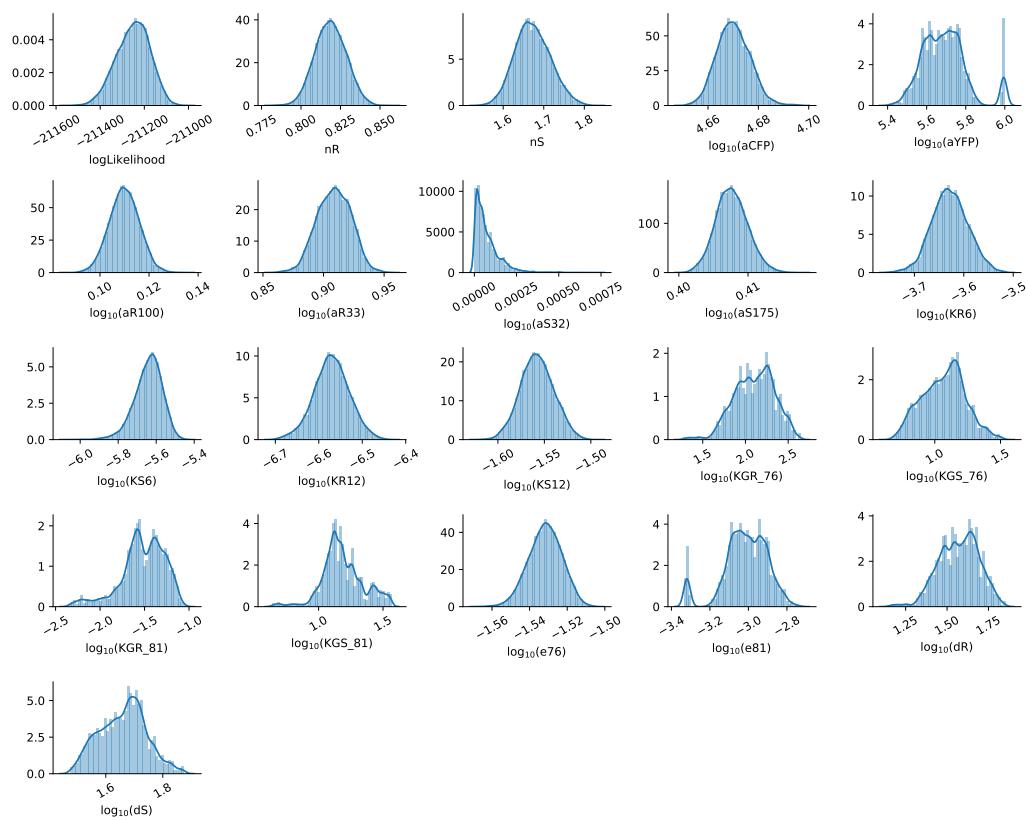
Inference for version 1 (uniform degradation). To characterize the LuxR and LasR signalling components, we used measurements of the response of four Receiver circuits from [3] to treatment with C6 and C12 over 3-fold dilutions. The maximum LuxR and LasR production rates were normalized to the values corresponding to the Pcat promoters, as done previously [3], thus setting the scale for unobserved concentrations of LuxR and LasR.

We used (uninformative) uniform priors on the previously uncharacterized parameters, and (informative) truncated Gaussian priors on f_{480} , f_{530} , d_{CFP} and d_{YFP} with mean and standard deviation taken from the marginal posteriors of the prpr circuit characterization. We did not propagate the marginal posteriors of a_{CFP} and a_{YFP} as the promoter involved differed between the prpr circuit and Receiver circuits. The priors used are detailed in the following table.

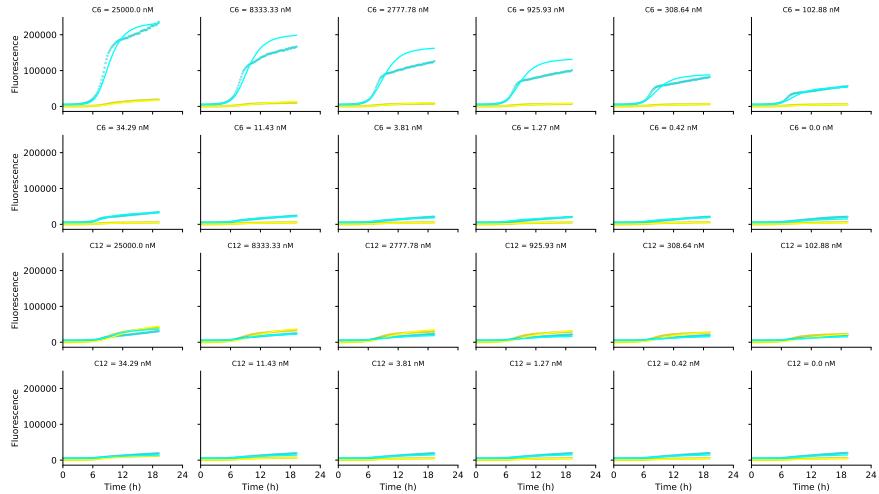
The marginal posterior estimates are shown in Figure 19. Simulation of the maximum likelihood estimate is shown in Figure 20.

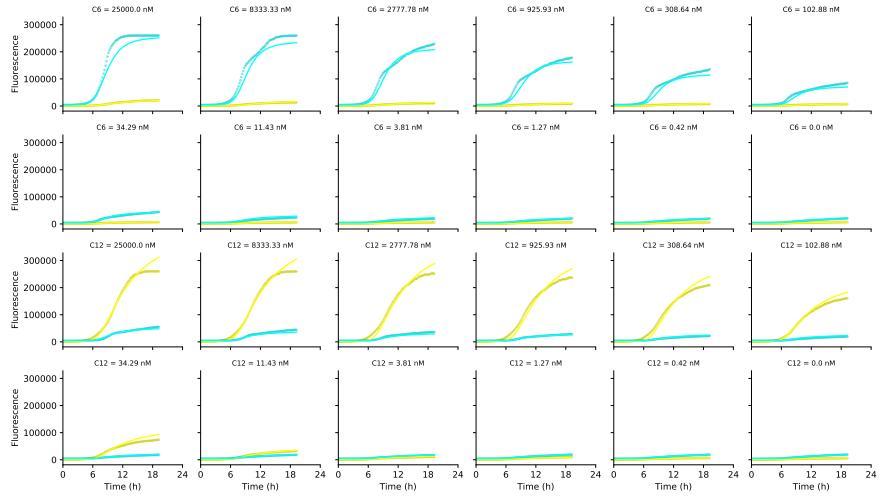
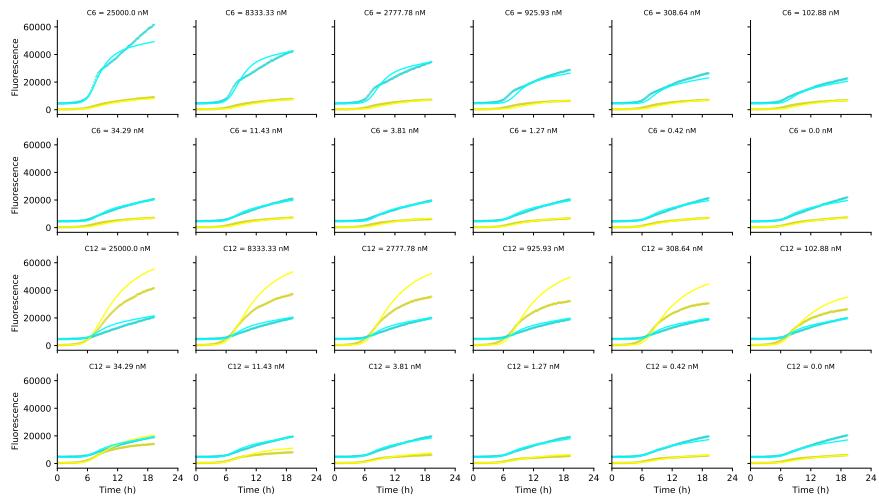
Supplementary Table 4: Priors for version 1 of the Receiver circuit characterization. In the unit column, M_R and M_S denote the normalised mass units for LuxR and LasR, and Fl denotes fluorescence units. The scaling column refers to the proposal distribution used to generate new parameters during MCMC.

Parameter	Description	Unit	Distribution	Scaling
a_{R100}	LuxR synthesis	$M_R \cdot h^{-1}$	$U(1, 10^2)$	Log
a_{R33}	LuxR synthesis	$M_R \cdot h^{-1}$	$U(1, 10^2)$	Log
a_{S32}	LasR synthesis	$M_S \cdot h^{-1}$	$U(1, 10^2)$	Log
a_{S175}	LasR synthesis	$M_S \cdot h^{-1}$	$U(1, 10^2)$	Log
d_R	LuxR degradation	h^{-1}	$U(10^{-2}, 10^2)$	Log
d_S	LasR degradation	h^{-1}	$U(10^{-2}, 10^2)$	Log
K_{R6}	Dissociation (LuxR-C6)	nM^{-1}	$U(10^{-8}, 1)$	Log
K_{S6}	Dissociation (LasR-C6)	nM^{-1}	$U(10^{-8}, 1)$	Log
K_{R12}	Dissociation (LuxR-C12)	nM^{-1}	$U(10^{-8}, 1)$	Log
K_{S12}	Dissociation (LasR-C12)	nM^{-1}	$U(10^{-8}, 1)$	Log
$K_{GR}^{(76)}$	Dissociation (P76-LuxR)	M_R^{-2}	$U(10^{-4}, 10^3)$	Log
$K_{GS}^{(76)}$	Dissociation (P76-LasR)	M_S^{-2}	$U(10^{-8}, 10^3)$	Log
$K_{GR}^{(81)}$	Dissociation (P81-LuxR)	M_R^{-2}	$U(10^{-8}, 10^3)$	Log
$K_{GS}^{(81)}$	Dissociation (P81-LasR)	M_S^{-2}	$U(10^{-4}, 10^3)$	Log
n_R	Hill constant (LuxR-HSL)	-	$U(0.5, 2)$	Real
n_S	Hill constant (LasR-HSL)	-	$U(0.5, 2)$	Real
$e^{(76)}$	Leak production (P76)	-	$U(10^{-4}, 1)$	Log
$e^{(81)}$	Leak production (P81)	-	$U(10^{-4}, 1)$	Log
a_{CFP}	CFP synthesis	$Fl \cdot cell^{-1} \cdot h^{-1}$	$U(10^2, 10^6)$	Log
a_{YFP}	YFP synthesis	$Fl \cdot cell^{-1} \cdot h^{-1}$	$U(10^2, 10^6)$	Log
b_{480}	Background fluor. (CFP)	Fl	$U(0, 10^4)$	Real
b_{530}	Background fluor. (YFP)	Fl	$U(0, 5 \times 10^3)$	Real



Supplementary Figure 19: Marginal parameter posterior estimates of the Receiver model (version 1) parameters. The marginal distributions are computed from 20 independent MCMC chains.

A Pcat circuit**B R100S32 circuit**

C R33S32 circuit**D R33S175 circuit**

Supplementary Figure 20: Comparison of Receiver models with fluorescence measurements. Simulations are of bulk CFP (cyan lines) and YFP (yellow lines) fluorescence, evaluated with the maximum likelihood parameter set. Measurements of bulk culture fluorescence are shown in for CFP (cyan circles) and YFP (yellow circles). C6 and C12 treatment concentrations are indicated atop each panel.

Version 2 - Protected degradation. Here we provide an alternative derivation for the model reduction based on the assumption that HSL complexes and dimers are protected from degradation, i.e. that degradation predominantly occurs on the level of LuxR/LasR monomers. This is supported by in vitro analysis of purified LuxR suggesting that LuxR protein is unstable in the absence of 3OC6-HSL [6]. Furthermore, we assume that all dilution effects occur on a slow time scale as compared to the kinetic. With this derivation, we obtain the same functional structure of the promoter activities of PLux76 and PLas81 from equation (15), but with the definitions of the bound fraction of LuxR and LasR given instead by

$$B_R := c_R^2 (c_6^{n_R} + (e_{R12} c_{12})^{n_R}), \quad B_S := c_S^2 ((e_{S6} c_6)^{n_S} + c_{12}^{n_S}) \quad (18)$$

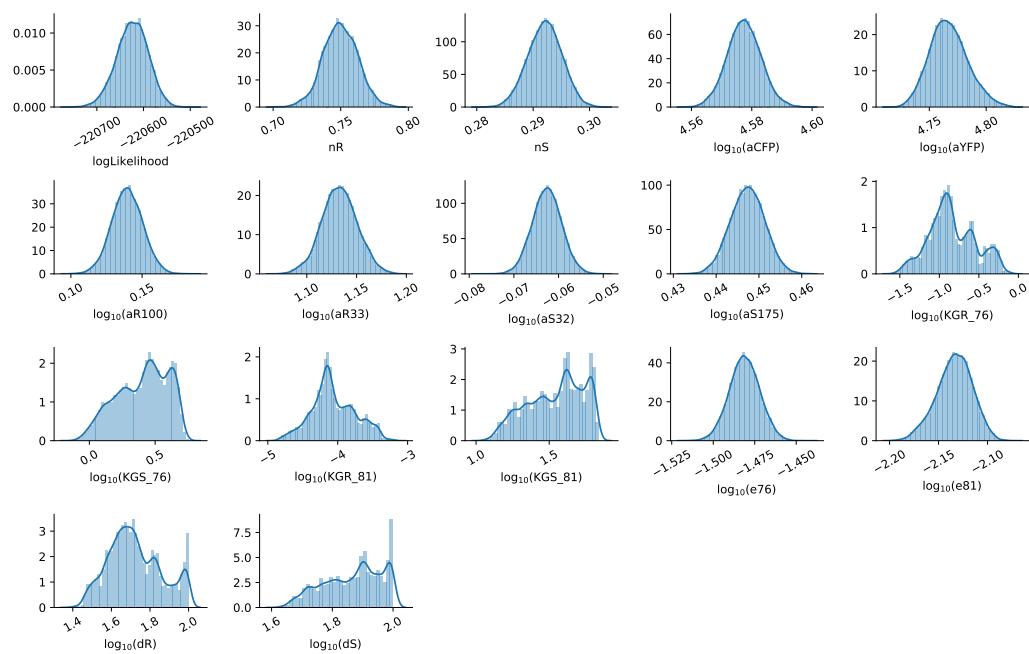
where $e_{R12} = \frac{K_{R12}}{K_{R6}}$ and $e_{S6} = \frac{K_{S6}}{K_{S12}}$ result from dividing by K_{R6} and K_{S12} . Consequently, the parameters $K_{GR}^{(i)}$ and $K_{GS}^{(i)}$ in equation (15) are rescaled by K_{R6} and K_{S12} respectively.

Inference for version 2 (Protected degradation). We carried out parameter inference for version 2 of the model directly equivalent to that done for version 1. The complete list of prior distributions for the parameters is as follows.

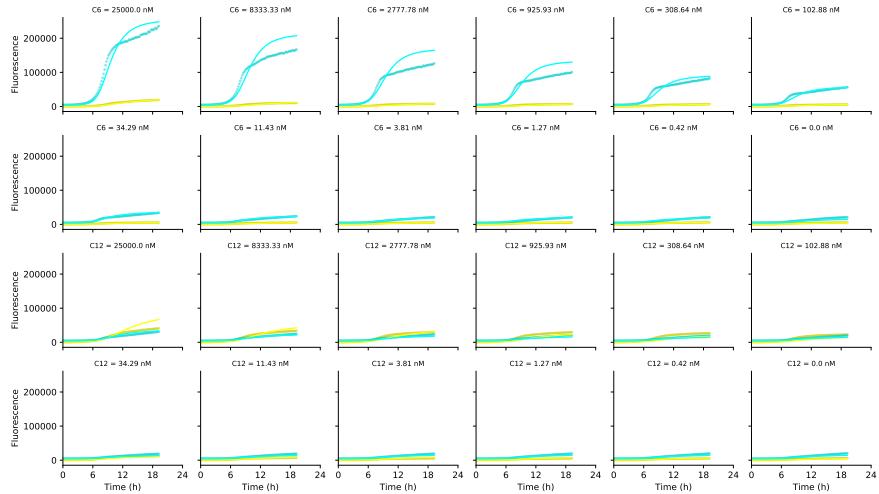
Supplementary Table 5: Priors for version 2 of the Receiver circuit characterization. In the unit column, M_R and M_S denote the normalised mass units for LuxR and LasR, and Fl denotes fluorescence units. The scaling column refers to the proposal distribution used to generate new parameters during MCMC.

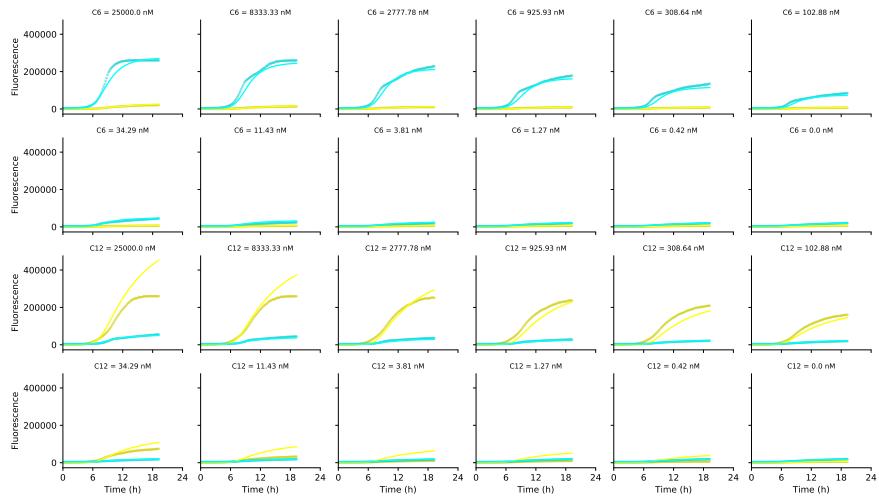
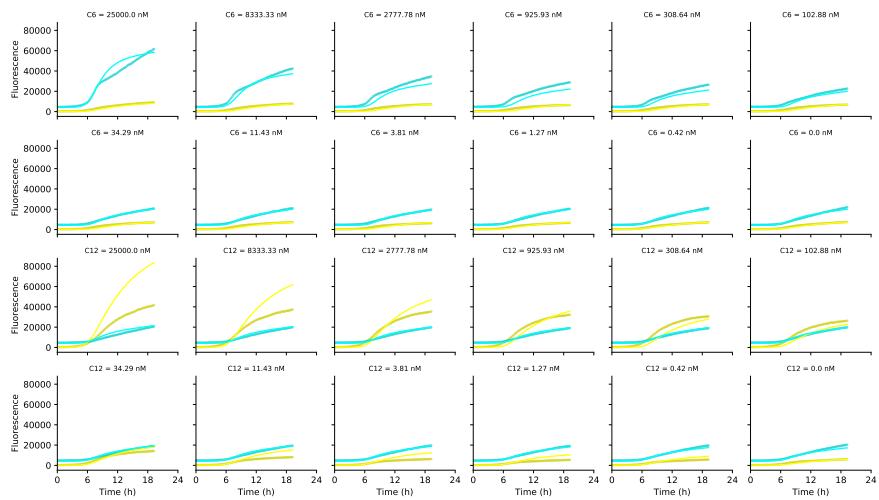
Parameter	Description	Unit	Distribution	Scaling
a_{R100}	LuxR synthesis	$M_R \cdot h^{-1}$	$U(1, 10^2)$	Log
a_{R33}	LuxR synthesis	$M_R \cdot h^{-1}$	$U(1, 10^2)$	Log
a_{S32}	LasR synthesis	$M_S \cdot h^{-1}$	$U(1, 10^2)$	Log
a_{S175}	LasR synthesis	$M_S \cdot h^{-1}$	$U(1, 10^2)$	Log
d_R	LuxR degradation	h^{-1}	$U(10^{-2}, 10^2)$	Log
d_S	LasR degradation	h^{-1}	$U(10^{-2}, 10^2)$	Log
e_{R12}	Chemical crosstalk (LuxR)	-	$U(10^{-8}, 1)$	Log
e_{S6}	Chemical crosstalk (LasR)	-	$U(10^{-8}, 1)$	Log
$K_{GR}^{(76)}$	Dissociation (P76-LuxR)	$M_R^{-(2+n_R)}$	$U(10^{-4}, 10^3)$	Log
$K_{GS}^{(76)}$	Dissociation (P76-LasR)	$M_S^{-(2+n_S)}$	$U(10^{-8}, 10^3)$	Log
$K_{GR}^{(81)}$	Dissociation (P81-LuxR)	$M_R^{-(2+n_R)}$	$U(10^{-8}, 10^3)$	Log
$K_{GS}^{(81)}$	Dissociation (P81-LasR)	$M_S^{-(2+n_S)}$	$U(10^{-4}, 10^3)$	Log
n_R	Hill constant (LuxR-HSL)	-	$U(0.2, 2)$	Real
n_S	Hill constant (LasR-HSL)	-	$U(0.2, 2)$	Real
$e^{(76)}$	Leak production (P76)	-	$U(10^{-4}, 1)$	Log
$e^{(81)}$	Leak production (P81)	-	$U(10^{-4}, 1)$	Log
a_{CFP}	CFP synthesis	$Fl \cdot cell^{-1} \cdot h^{-1}$	$U(10^2, 10^6)$	Log
a_{YFP}	YFP synthesis	$Fl \cdot cell^{-1} \cdot h^{-1}$	$U(10^2, 10^6)$	Log
B_{480}	Background fluorescence (CFP)	Fl	$U(0, 10^4)$	Real
B_{530}	Background fluorescence (YFP)	Fl	$U(0, 5 \times 10^3)$	Real

The marginal posterior estimates are shown in Figure 21. Simulation of the maximum likelihood estimate is shown in Figure 22.



Supplementary Figure 21: Marginal parameter posterior estimates of the Receiver model (version 2) parameters. The marginal distributions are computed from 20 independent MCMC chains.

A Pcat circuit**B R100S32 circuit**

C R33S32 circuit**D R33S175 circuit**

Supplementary Figure 22: Comparison of Receiver models with fluorescence measurements. Simulations are of bulk CFP (cyan lines) and YFP (yellow lines) fluorescence, evaluated with the maximum likelihood parameter set. Measurements of bulk culture fluorescence are shown in for CFP (cyan circles) and YFP (yellow circles). C6 and C12 treatment concentrations are indicated atop each panel.

Exclusive Receiver model

To model the Exclusive Receiver, we consider the inhibition of PTet by TetR and PLac by LacI, and the mechanism of chemical inhibition by IPTG and ATC. For promoter regulation, we use the inhibition Hill function

$$H_I(x, n) := \frac{1}{1 + x^n} \quad (19)$$

Typically, these functions would include a parameter for the half-saturation concentration, but we omit that here because, without loss of generality, [LacI] and [TetR] can be arbitrarily scaled by those half-saturation concentrations. In such a rescaling, the half-saturation constants get embedded within the maximal production rates, a_L and a_T . For the chemical inhibitors, we assume a reaction of the form



Correspondingly, the action of IPTG and ATC is proportional to the product of its concentration and its target repressor protein concentration.

$$\frac{d\rho}{dt} = \gamma(\rho).\rho \quad (21a)$$

$$\frac{dc_{YFP}}{dt} = a_{YFP}.f_{81}(c_6, c_{12}, c_R, c_S) - (\gamma(\rho) + d_{YFP})c_{YFP} \quad (21b)$$

$$\frac{dc_{CFP}}{dt} = a_{CFP}.f_{76}(c_6, c_{12}, c_R, c_S) - (\gamma(\rho) + d_{CFP})c_{CFP} \quad (21c)$$

$$\frac{dc_{530}}{dt} = b_Y - \gamma(\rho).c_{530} \quad (21d)$$

$$\frac{dc_{480}}{dt} = b_C - \gamma(\rho).c_{480} \quad (21e)$$

$$\frac{dc_R}{dt} = a_R.H_I(c_T, n_T) - (\gamma(\rho) + d_R).c_R \quad (21f)$$

$$\frac{dc_S}{dt} = a_S.H_I(c_L, n_L) - (\gamma(\rho) + d_S).c_S \quad (21g)$$

$$\frac{dc_L}{dt} = a_L.P_{76}(c_6, c_{12}, c_R, c_S) - (\gamma(\rho) + d_L + i_I.c_I).c_L \quad (21h)$$

$$\frac{dc_T}{dt} = a_T.P_{81}(c_6, c_{12}, c_R, c_S) - (\gamma(\rho) + d_T + i_A.c_A).c_T \quad (21i)$$

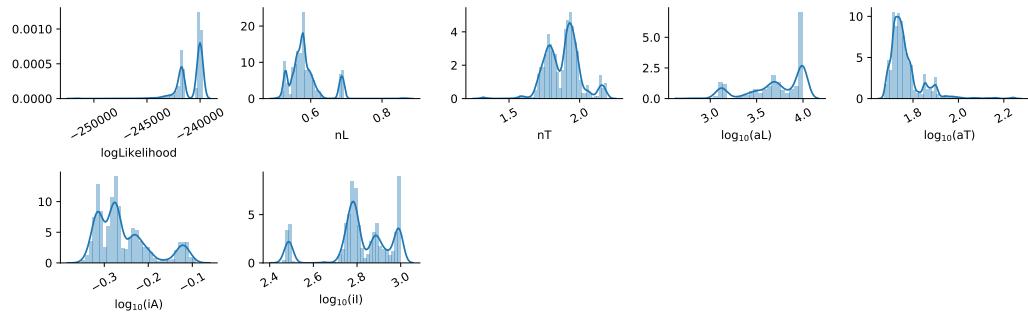
where P_{76} and P_{81} are defined in (15).

Inference for version 1 (uniform degradation). The inference procedure was less robust for the Exclusive Receiver, as compared with upstream circuits in the inference graph. In particular, we found it was not possible to identify a unique mode within the parameter space when all parameters were allowed to be flexible. Specifically, there was strong interdependency within the triplet $\{a_L, d_L, i_I\}$ and the triplet $\{a_T, d_T, i_A\}$. Our interpretation is that we are unable to fully recover the time-scales of variations in c_L and c_T , as they are likely to not vary much during the experiments we used for characterization. It's likely that c_L and c_T quickly stabilise to equilibria when the cells are transferred to the media containing the treatments (specific concentrations of C6, C12, IPTG and ATC). In which case, only those equilibrium values will be identifiable, and not the production and degradation rates separately. Therefore, in the final version of the inference results presented here, we have fixed the degradation rates of LacI and TetR to 1 h^{-1} .

Even when running the inference with d_L and d_T fixed, we found that chain convergence was not perfect, and some chains got stuck in local optima. Therefore, in our marginal posterior estimates, we have only included chains that converged to relatively good likelihood scores Figure 23. The marginals clearly indicate some additional flexibility in the inferred parameter values, possibly resulting from the larger parameter space being navigated, which includes uninformative priors for the parameters listed in Table 6, but also some flexibility in all of the parameters associated with the Receiver module, despite them having a strong prior.

Supplementary Table 6: Priors for the Exclusive Receiver circuit characterization. In the unit column, M_L and M_T denote the normalised mass units for LacI and TetR, and Fl denotes fluorescence units. The scaling column refers to the proposal distribution used to generate new parameters during MCMC.

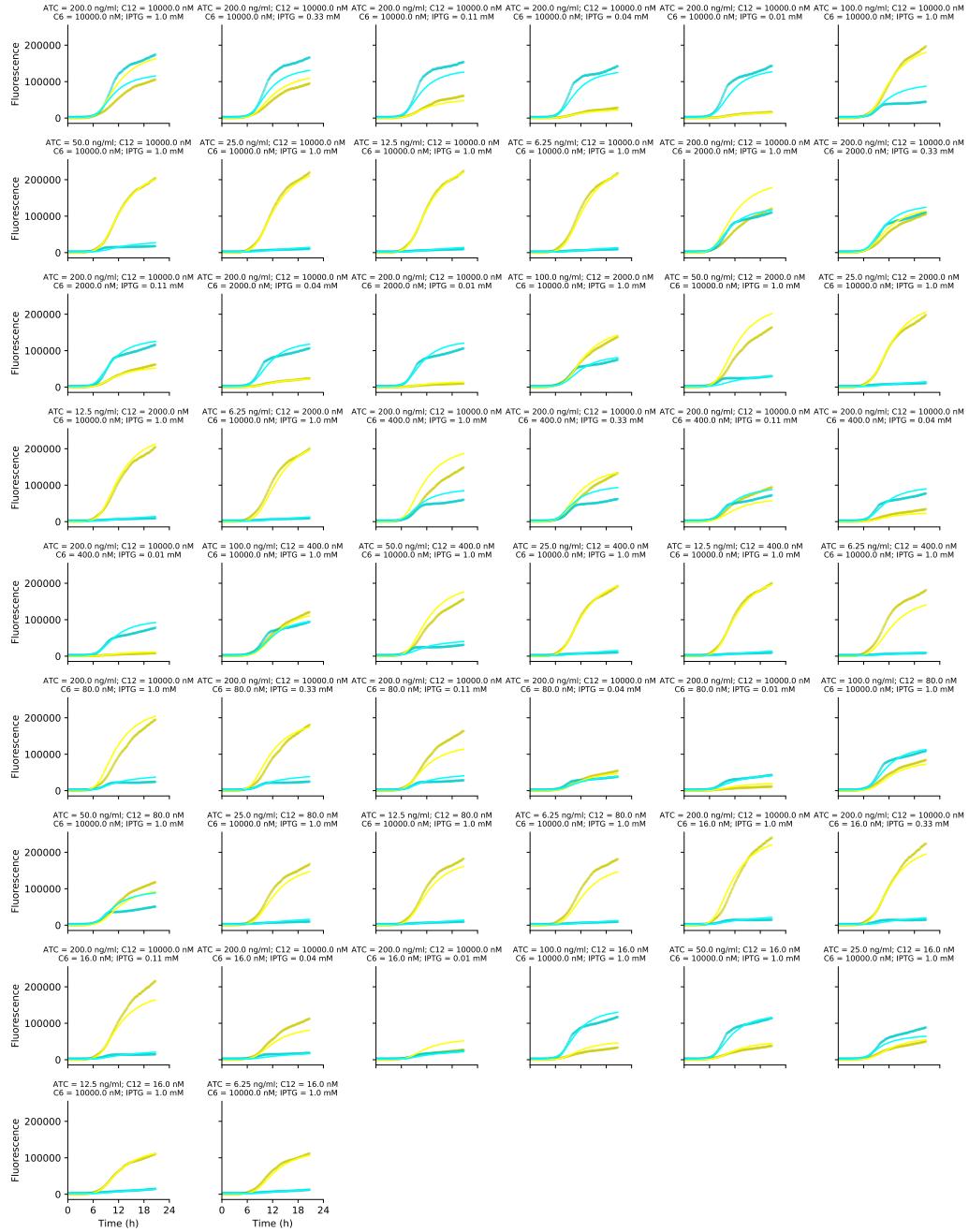
Parameter	Description	Unit	Distribution	Scaling
a_L	LacI synthesis	$M_L \cdot h^{-1}$	$U(1, 10^4)$	Log
a_T	TetR synthesis	$M_T \cdot h^{-1}$	$U(1, 10^4)$	Log
d_L	LacI degradation	h^{-1}	1 (Fixed)	N/A
d_T	TetR degradation	h^{-1}	1 (Fixed)	N/A
n_L	Hill constant (Laci)	-	$U(0.5, 4.0)$	Real
n_T	Hill constant (TetR)	-	$U(0.5, 4.0)$	Real
i_A	ATC inhibition of TetR	$M_T^{-1} \cdot (\text{ng/ml ATC})^{-1} \cdot h^{-1}$	$U(10^{-2}, 10^3)$	Log
i_I	IPTG inhibition of LacI	$M_L^{-1} \cdot (\text{mM IPTG})^{-1} \cdot h^{-1}$	$U(10^{-2}, 10^3)$	Log
B_{480}	Background fluor. (CFP)	Fl	$U(0, 10^4)$	Real
B_{530}	Background fluor. (YFP)	Fl	$U(0, 5 \times 10^3)$	Real



Supplementary Figure 23: Marginal parameter posterior estimates of the Exclusive Receiver model (version 1) parameters. The marginal distributions are computed from 9 of 20 independent MCMC chains. Chains were discarded that did not converge to high log-likelihood regions.

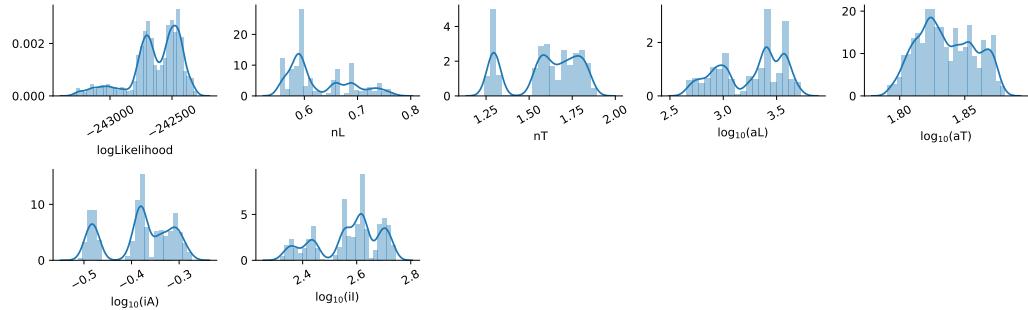
Simulation of the maximum likelihood estimate is shown in Figure 23.

A 2d dilution of C₆ and C₁₂

B Characterization of IPTG and ATC

Supplementary Figure 23: Comparison of Exclusive Receiver model (version 1) with fluorescence measurements. Simulations are of bulk CFP (cyan lines) and YFP (yellow lines) fluorescence, evaluated with the maximum likelihood parameter set. Measurements of bulk culture fluorescence are shown in for CFP (cyan circles) and YFP (yellow circles). C6, C12, ATC and IPTG treatment concentrations are indicated atop each panel.

Inference for version 2 (Protected degradation). The inference procedure was also not completely robust for version 2 of the Exclusive Receiver model. We used the same uninformative priors for the parameters specific to the Exclusive Receiver model as in version 1, including fixing d_L and d_T . Again, in our marginal posterior estimates, we have only included chains that converged to relatively good likelihood scores Figure 24.



Supplementary Figure 24: Marginal parameter posterior estimates of the Exclusive Receiver model (version 2) parameters. The marginal distributions are computed from 8 of 20 independent MCMC chains. Chains were discarded that did not converge to high log-likelihood regions.

Simulation of the maximum likelihood estimate is shown in Extended Data Figure 2.

2.2 Bistability Analysis

In this section we outline computations used to create Figure 2b in the main text which compares regions of bistability indicated by hysteresis experiments in flow cytometry to that given by the differential equation model for the exclusive receiver circuit.

To characterize the region in the (c_{12}, c_6) plane where bistability occurs, we used numerical continuation to calculate a co-dimension two limit curve. The code in `get_bifurcations.jl` in our repository takes advantage of the Julia package `PseudoArcLengthContinuation.jl` [7]. To calculate the bifurcations, we only need consider the steady states of the model. The auto-fluorescence equations are independent of the others and CFP and YFP are simply readouts of c_L and c_T respectively, leaving only four coupled equations to solve, defined by state vector $\mathbf{c} = (c_R, c_S, c_L, c_T)$. For simplicity of analysis, we assume that cell density ρ is constant, and consequently the specific growth rate $\gamma(\rho) =: \gamma_0$ is constant. Later, we check this assumption in Figure 26. With this simplification, the model given by (21) can be represented compactly as

$$\frac{d\mathbf{c}}{dt} = \mathbf{F}_\theta(\mathbf{c}, \mathbf{u}) \quad (22)$$

where $\mathbf{u} = (c_6, c_{12}, c_I, c_A)$ are the experimental control parameters, θ is the vector containing the inferred parameters from Section S1 and γ_0 .

To improve the numerical stability of numerical continuation, we transform the model into \log_{10} coordinates via the element-wise transformations $\mathbf{c} \rightarrow 10^\mathbf{c}$ and $\mathbf{u} \rightarrow 10^\mathbf{u}$ yielding

$$\frac{d\mathbf{c}}{dt} = \frac{\mathbf{F}_\theta(10^\mathbf{c}, 10^\mathbf{u}) 10^{-\mathbf{c}}}{\ln(10)} \quad (23)$$

The steady states are defined by zeros of the numerator of the right-hand side. We can immediately see that this transformation induced a zero at $\mathbf{c} \rightarrow \infty$ which we are not interested in and thus simply seek to solve

$$\mathbf{F}_\theta(10^\mathbf{c}, 10^\mathbf{u}) = 0 \quad (24)$$

To further increase numerical stability of finding 24 we explicitly calculate the Jacobian in the log coordinate system. Luckily the Jacobian in the new coordinates can be expressed in terms of the matrix product between the Jacobian in the original coordinates and the Jacobian of the coordinate transformation

$$\frac{\partial \mathbf{F}_\theta}{\partial \mathbf{c}} \rightarrow \ln(10) \frac{\partial \mathbf{F}_\theta}{\partial \mathbf{c}} \Big|_{\mathbf{c} \rightarrow 10^\mathbf{c}} \text{Diag}[10^\mathbf{c}] \quad (25)$$

where $\text{Diag}[\mathbf{v}]$ is a diagonal matrix with the vector components of \mathbf{v} along the diagonal. The Jacobian in the original coordinates is

$$\frac{\partial \mathbf{F}_\theta}{\partial \mathbf{c}} = \begin{pmatrix} 0 & 0 & 0 & a_{R33} \frac{\partial H_T}{\partial T} \\ 0 & 0 & a_{S175} \frac{\partial H_L}{\partial L} & 0 \\ a_L \frac{\partial P_{76}}{\partial R} & a_L \frac{\partial P_{76}}{\partial S} & 0 & 0 \\ a_T \frac{\partial P_{81}}{\partial R} & a_T \frac{\partial P_{81}}{\partial S} & 0 & 0 \end{pmatrix} - \text{Diag} \left[\gamma_0 + \begin{pmatrix} d_R \\ d_S \\ d_L \\ d_T \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ i_I c_I \\ i_A c_A \end{pmatrix} \right] \quad (26)$$

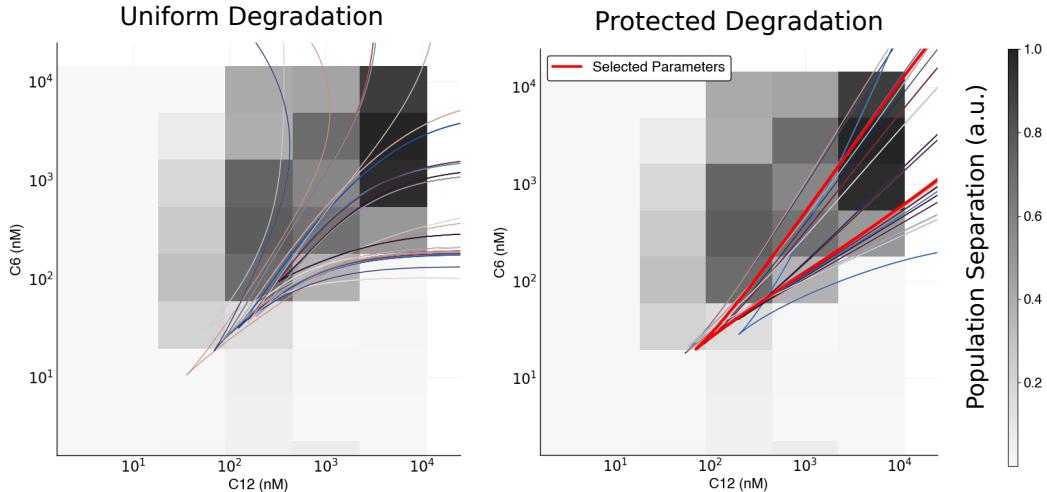
where partials of inhibitory hill functions and promoter activities are

$$\frac{\partial H_X}{\partial X} = \frac{-n_X c_X^{(n_X-1)}}{(1+c_X^{n_X})^2} \quad \frac{\partial P_N}{\partial X} = \frac{K_{GX}^N (1-e_N)}{[1+K_{GR}^N B_R(c_R) + K_{GS}^N B_S(c_S)]^2} \frac{\partial B_X}{\partial X} \quad (27)$$

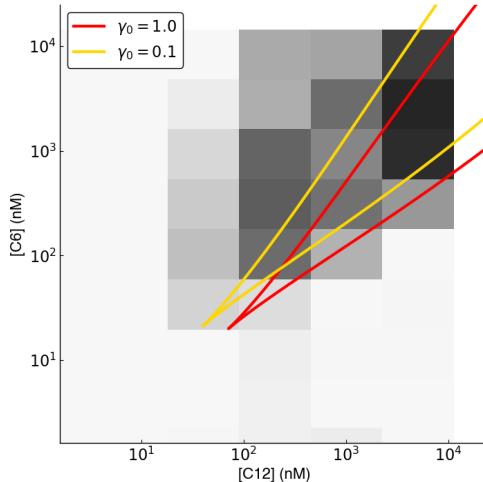
and partial bound molecules for uniform and protected degradation models are respectively

$$\frac{\partial B_X}{\partial X} = 2c_X \frac{(K_{X6}c_6)^{n_X} + (K_{X12}c_{12})^{n_X}}{(1+K_{X6}c_6+K_{X12}c_{12})^{n_X}} \quad \frac{\partial B_X}{\partial X} = 2c_X \begin{cases} c_6^{n_X} + (E_{R12}c_{12})^{n_X} & X = R \\ c_{12}^{n_X} + (E_{S6}c_6)^{n_X} & X = S \end{cases}$$

Given the rate function \mathbf{F}_θ and its Jacobian $\frac{\partial \mathbf{F}_\theta}{\partial c}$ in log coordinates we can perform a co-dimension one parameter continuation for a fixed value of c_{12} along the c_6 direction to find a limit point. Then the solution can be continued along a limit curve in the (c_6, c_{12}) plane along both directions until the limits of the observation region are met. Figure 25 reveals these curves for different models and inferred maximum likelihood parameter sets θ and Figure 26 reveals that the chosen model for the main text is insensitive to changes in growth γ_0 and therefore we can safely assume that the qualitative behaviour of the model will not change if the cell density is constant. For simplicity, this is what is done in the spatial simulations in Sections 2.3 and 2.4.



Supplementary Figure 25: Bifurcation curves for uniform and protected degradation models. Curves are plotted on top of bimodal population separation heatmaps generated from flow cytometry data. The selected maximum likelihood parameters are used in the main text.



Supplementary Figure 26: Bifurcation curve insensitivity specific growth rate γ_0 . The cell density is a monotonically increasing function of time that saturates at the carrying capacity. Concomitantly, the specific growth rate declines towards 0. By evaluating the bifurcation diagram at different values of γ_0 , we can see how the cusp would move over time. Here the maximum likelihood parameters from the main text reveal that the cusp does not move much as the cell density increases over time.

2.3 Boundary Experiments

Simulation and local equilibria

For spatial simulations a simple forward-Euler method is implemented in `get_movie.py`. The bacterial colonies for the spatial experiments were placed on top of agar with no signalling molecules inside it. Then additional volumes of agar were attached either side of the width of the experiment with different concentrations of c_6 and c_{12} . The signalling molecules then diffuse in the agar and established a cross-gradient felt by the bacterial colonies. These are governed by diffusion

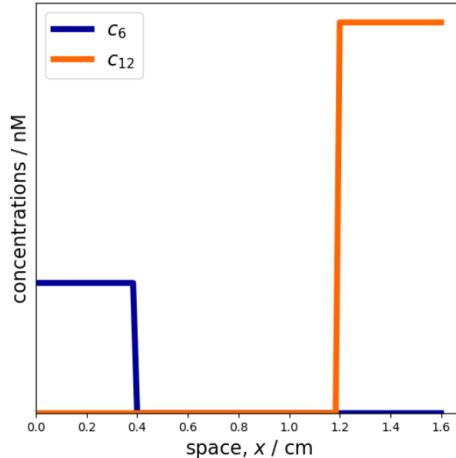
$$\frac{\partial c_X(x, t)}{\partial t} = D_X \frac{\partial^2 c_X(x, t)}{\partial x^2} \quad (28)$$

where $D_6 = 1.8 \cdot 10^{-6} \text{ m}^2 \cdot \text{h}^{-1}$ and $D_{12} = 0.9 \cdot 10^{-6} \text{ m}^2 \cdot \text{h}^{-1}$.

The initial conditions for spatial simulations are zero everywhere except for $c_6(x, t)|_{t=0}$ and $c_{12}(x, t)|_{t=0}$ initialised in small regions widths w on opposite sides of the experiment of width W at concentrations such that the homogeneous equilibrium after diffusion would be $c_6(x, t)|_{t \rightarrow \infty} = \bar{C}_6$ and $c_{12}(x, t)|_{t \rightarrow \infty} = \bar{C}_{12}$ with zero-flux boundary conditions. Therefore

$$c_6(x, t)|_{t=0} = \frac{W\bar{C}_6}{w} H(w - x) \quad c_{12}(x, t)|_{t=0} = \frac{W\bar{C}_{12}}{w} H(x - W + w) \quad (29)$$

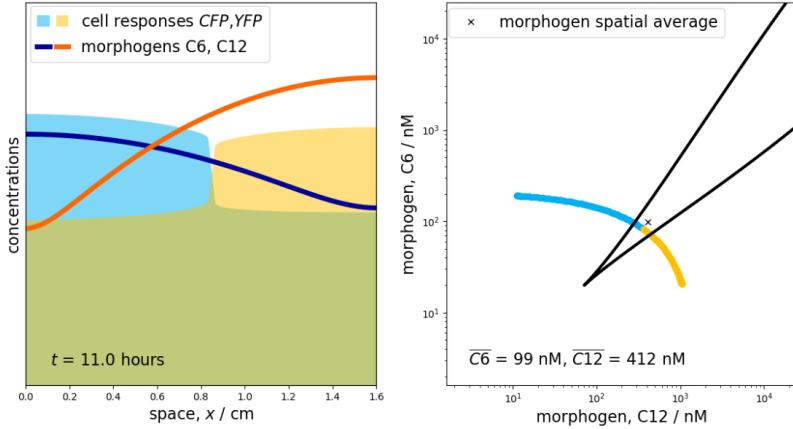
where $H(x)$ is a unit step function.



Supplementary Figure 27: Initial conditions $c_6(x, t)|_{t=0}$ and $c_{12}(x, t)|_{t=0}$ with $W = 1.6 \text{ cm}$ and $w = 0.4 \text{ cm}$

Each location x experiences concentrations c_6, c_{12} which define a local equilibrium for the remaining state variables. These local equilibria evolve over time and are chased by the actual concentrations of protein in the cells at that location. As shown by Figure 25, at some concentrations c_6, c_{12} there may two stable equilibria rather than one. Therefore it becomes useful to not only display the dynamics in the one dimensional spatial domain but also in the (c_6, c_{12}) plane. How and when local equilibria bifurcate reveals the eventual fate of the spatial pattern.

Figure 28 shows a snapshot of the dynamics in the spatial domain x and the (c_6, c_{12}) plane. A sharp boundary in space forms when the state density in (c_6, c_{12}) moves into the bistable region enclosed by the limit point curve. This means that, given the cross-gradient initial conditions 29, if the homogeneous equilibrium \bar{C}_6, \bar{C}_{12} lies within the bistable region, a sharp stationary boundary will form. If \bar{C}_6, \bar{C}_{12} lies below the cusp of the limit point curve, only soft boundaries will form. In other regions outside the bistable cone the sharp boundary has a finite velocity and will eventually leave the experimental region. See supplementary movies 2-5 for examples of the above. This motivated the experimental exploration of the space of \bar{C}_6, \bar{C}_{12} and measurements of boundary velocity, which are described in the following section.



Supplementary Figure 28: Dynamics in the spatial domain x on the left and (c_6, c_{12}) plane on the right. Spatial averages of morphogens \bar{C}_6, \bar{C}_{12} are 99 nM and 412 nM respectively for both panels

Computation of the boundary velocity

The velocity of the boundary is determined using `get_movement.py` from the TIFF image stack $M[t, x, y, s]$ obtained by the fluorescence microscope from one experiment, set up with a chosen \bar{C}_6, \bar{C}_{12} combination. Here t indexes the time point, x and y index the width and height and s indexes the three channels: CFP, YFP and RFP.

First the data are normalised by the RFP channel. This way the location of the boundary can be defined by comparing the pixel values of one channel against the other. The pixels are masked for the colony grid squares $(x, y) \in \Omega$ which are otherwise surrounded by hydrophobic ink filter paper, on which no colonies grow. The grid squares are detected by thresholding the constituent RFP channel at the `end` time point.

$$X[t, x, y, s] := \frac{M[t, x, y, s]}{M[t, x, y, \bullet]} \quad \text{where } (x, y) \in \Omega \quad (30)$$

$$\text{where } \Omega := \left\{ (x, y) : M[\text{end}, x, y, \bullet] > \frac{1}{2} \right\} \quad (31)$$

Next, a sigmoidal basis function Bayesian Ridge regressor is applied to each channel and time point along the width x of the preprocessed data $X[t, x, y, s]$. Sigmoidal basis functions are used because they satisfy the zero-flux boundary conditions as well as our expectations that the fluorescence profile across the width of the experiment will mostly be flat with a transient step due to the cross-gradient. The unknown weights ϕ of the basis functions are obtained by minimising the objective

$$J_\phi[t, s] := \sum_{(x, y) \in \Omega} ||X[t, x, y, s] - f_\phi(x)[t, s]||^2 \quad (32)$$

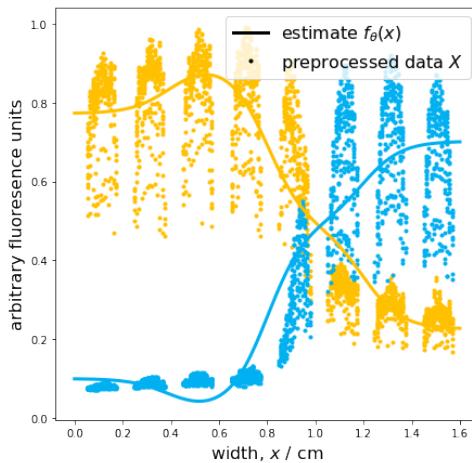
$$\text{where } f_\phi(x) := \sum_{(\alpha, \mu, \sigma) \in \phi} \frac{\alpha}{1 + e^{-\frac{x-\mu}{\sigma}}} \quad (33)$$

With the continuous estimate $f_{\phi^*}(x)$ for each $[t, s]$, the location of the boundary can be obtained even if it was estimated to lie between two colonies as seen in Figure 29. Since this continuous estimate is obtained for each point in time, the position of the boundary can be tracked in a smooth kymograph as shown in Figure 30. We define the boundary location $\beta[t]$ to be where the estimate of the CFP channel \bullet is equal to that of the YFP channel \circ ,

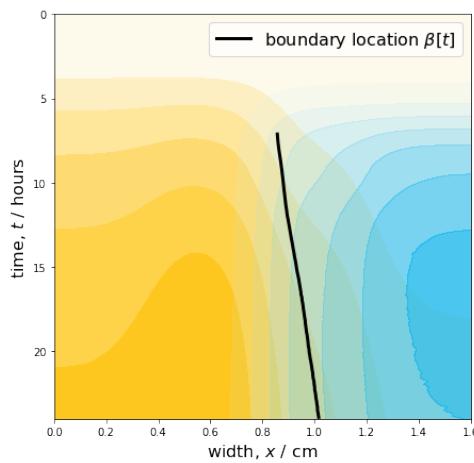
$$f_{\phi^*}(\beta)[t, \bullet] = f_{\phi^*}(\beta)[t, \bullet] \quad (34)$$

The distance travelled by the boundary $\Delta\beta$ from its formation time t^* to the end as a fraction of the size of the experiment W can now be computed. The formation time t^* was judged by eye and seems to lie between 3 – 5 h, at which fluorescence values are sufficiently steep to form a sharp boundary. The boundary should have travelled at least 10% along the width - which is the approximate size of one colony grid square - in order to be classified as moving.

$$\Delta\beta = \frac{\beta[t^*] - \beta[\text{end}]}{W} \quad (35)$$

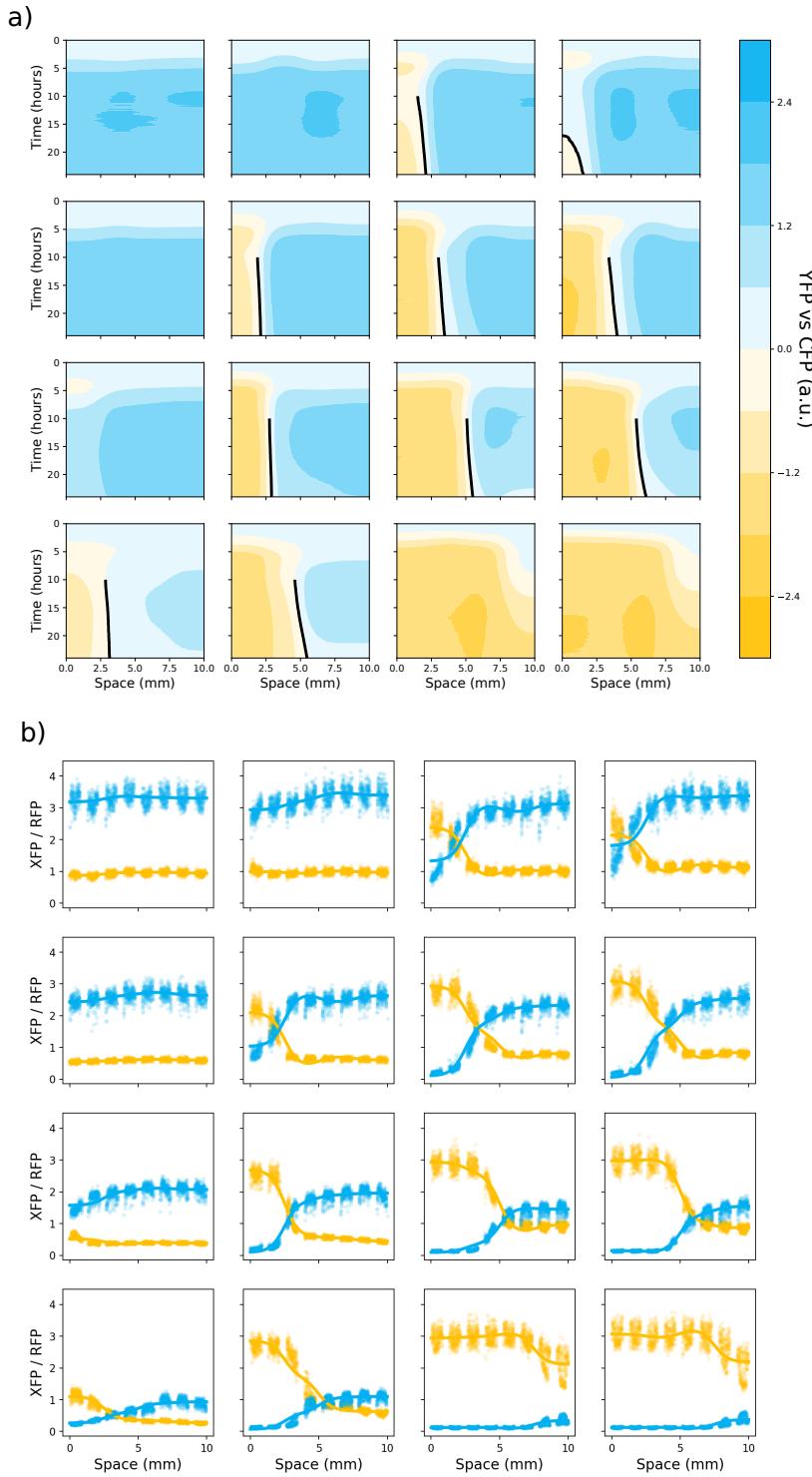


Supplementary Figure 29: Bayesian Ridge regression estimate $f_{\phi}(x)$ from preprocessed data X

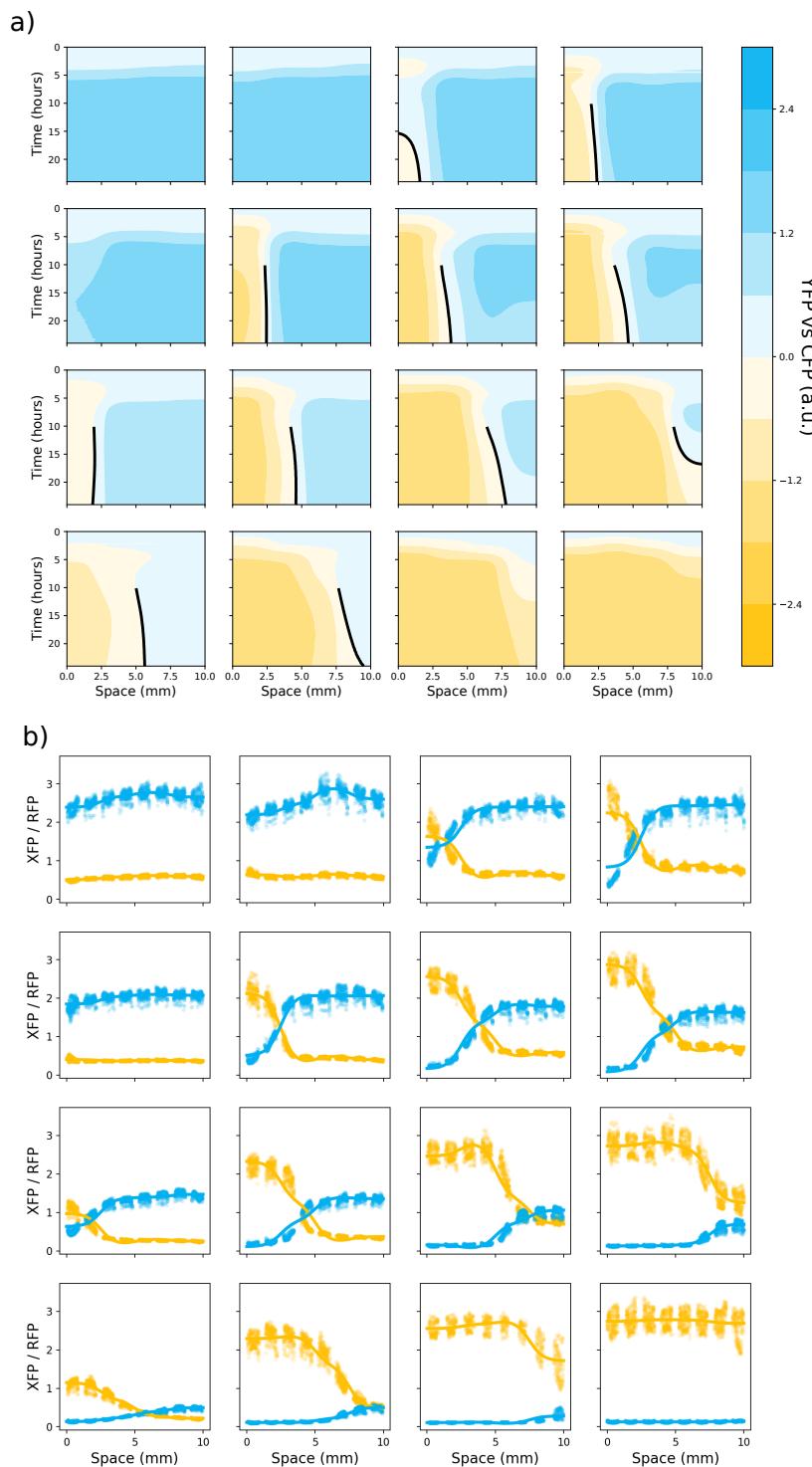


Supplementary Figure 30: Boundary location $\beta[t]$ by equating estimates from two channels

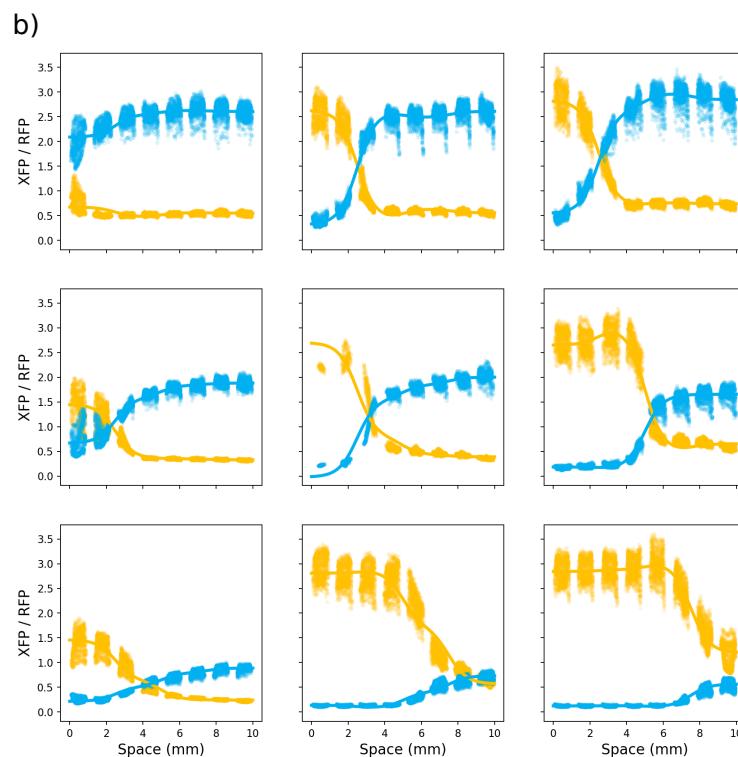
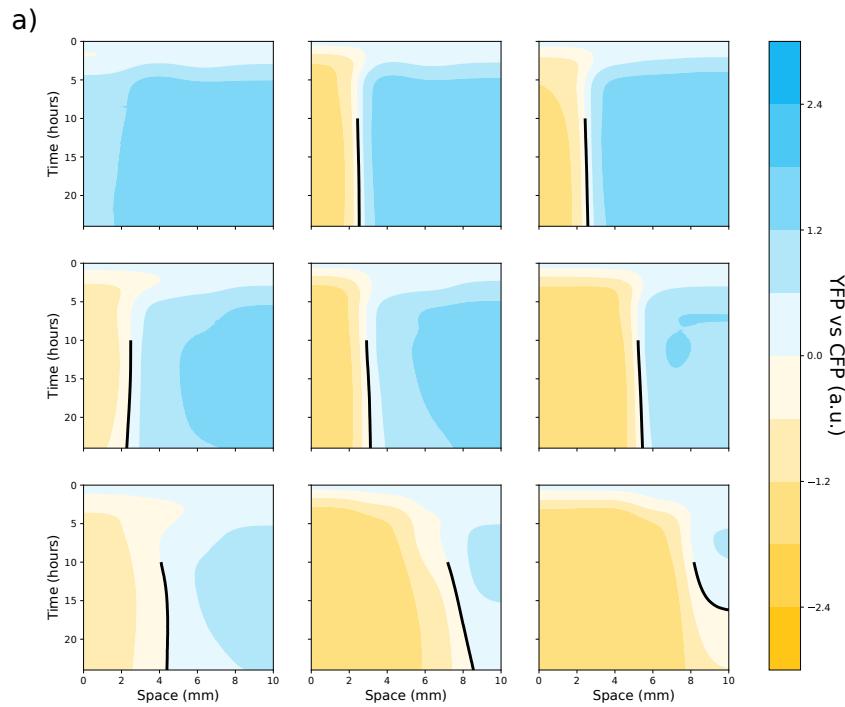
The distance travelled $\Delta\beta$ can be investigated for different equilibrium concentrations $\overline{C6}$ and $\overline{C12}$. Figure 31 shows results for a two dimensional dilution between 5 nM and 25000 nM and fixed 10 μ M IPTG. The subsequent classification using $\Delta\beta$ is shown as Figure 3c in the main text.



Supplementary Figure 31: a) Boundary movements $\beta[t]$ for a grid of equilibrium concentrations $\overline{C}_6, \overline{C}_{12}$ and fixed 10 μM IPTG. This is an expanded version of Figure 3c in the main text, showing the kymographs that were used for classification. b) Bayesian Ridge regression estimates for final time points of preprocessed data X



Supplementary Figure 32: a) Replicate of Figure 31 for a two dimensional dilution of $\overline{C}_6, \overline{C}_{12}$ between 5 nM and 25000 nM and fixed 10 μ M IPTG. b) Bayesian Ridge estimates for final time points of preprocessed data X

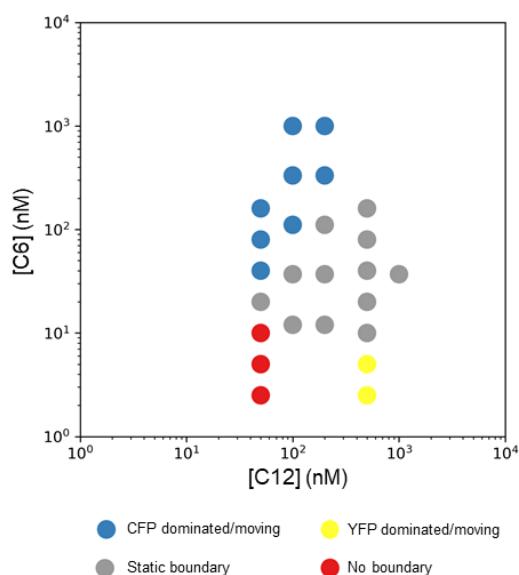


Supplementary Figure 33: a) Replicate of Figure 31 for a two dimensional dilution of $\overline{C6}, \overline{C12}$ between 20 nM and 2000 nM and fixed 10 μ M IPTG. b) Bayesian Ridge estimates for final time points of preprocessed data X

Use of IPTG to influence bifurcation curve

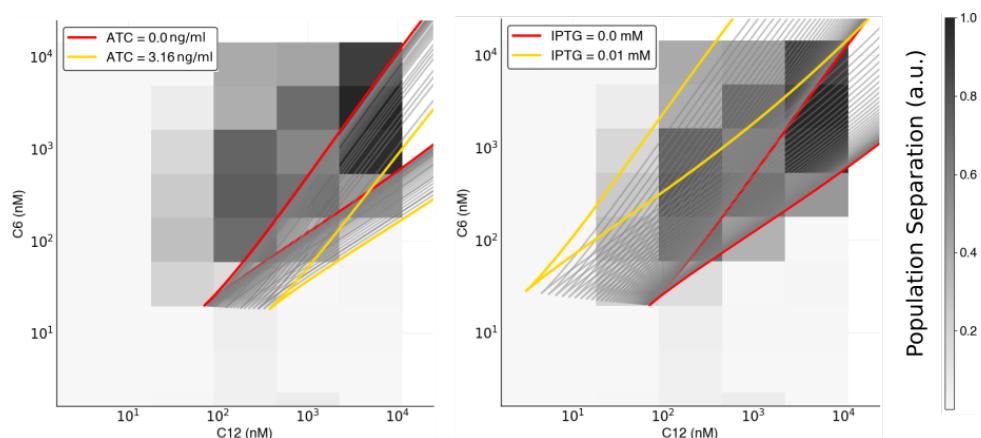
Solid culture experiments on boundary movement were initially performed in the absence of IPTG. We observed stationary boundaries at the concentrations labelled as grey points (Figure S24). The shape of the region encompassing the grey points was qualitatively similar to the conical bistability region we observed in previous experiments and in our models, but was quantitatively shifted such that even very low concentrations of C6 enabled bistability. We hypothesized that this was due to minor differences in culture conditions between solid and liquid cultures. We hypothesized that the addition of a low concentration of IPTG would partially derepress LacI, thereby making the YFP-dominant region larger and more like that seen in liquid culture. We used IPTG to shift the bistability region back (Figure S25) to coincide with the region in liquid culture. This allowed us to perform the solid culture experiments in a regime in which we could observe the transition from stable boundary to moving boundary with YFP dominance, by using higher concentrations of C12 while keeping C6 constant.

Experiments with the relay circuit (Fig. 4b) performed as expected without the need for addition of IPTG or ATC.



Supplementary Figure 34: Boundary movement in the absence of IPTG Boundaries were classified by eye at the spatial average concentrations indicated.

Experiments with the relay circuit (4a) performed as expected without the addition of IPTG or ATC.



Supplementary Figure 35: Bifurcation curve dependence on ATC and IPTG. Increasing the concentrations c_A or c_I shifts the cusp point keeping the general shape of the bistable region. Here the maximum likelihood parameters from the main text are used as the reference cusp.

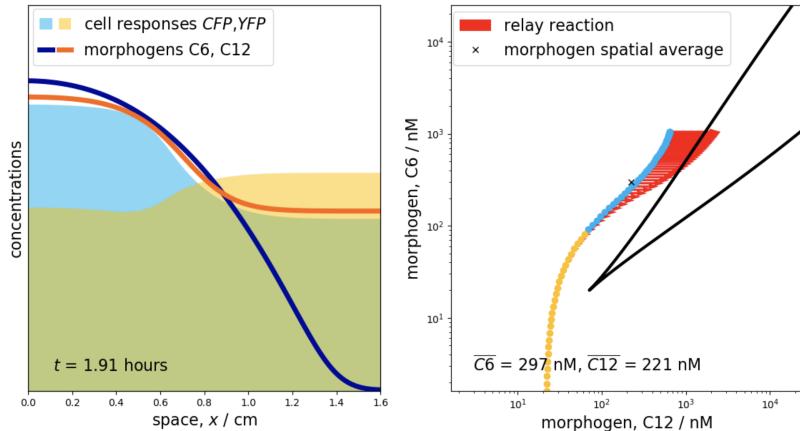
2.4 Models of the Exclusive Receiver Relay Circuits

The Exclusive Receiver Relay circuits were described in Figure 4 of the main text. The models for these circuits are simple extensions of the constant density model of the Exclusive Receiver circuit (22) with $\rho = \rho_0$, $\gamma = \gamma_0$ and production terms in the signal diffusion equations

$$\frac{\partial c_6(x, t)}{\partial t} = D_6 \frac{\partial^2 c_6(x, t)}{\partial x^2} + k_{C6}\rho_0 c_{\text{luxI}} \quad \frac{\partial c_{12}(x, t)}{\partial t} = D_{12} \frac{\partial^2 c_{12}(x, t)}{\partial x^2} + k_{C12}\rho_0 c_{\text{lasI}} \quad (36)$$

$$\frac{dc_{\text{luxI}}}{dt} = P_{81} - (\gamma_0 + d_{\text{luxI}})c_{\text{luxI}} \quad \frac{dc_{\text{lasI}}}{dt} = P_{76} - (\gamma_0 + d_{\text{lasI}})c_{\text{lasI}} \quad (37)$$

Here, the additional production terms $k_{C6}\rho_0 c_{\text{luxI}}$ and $k_{C12}\rho_0 c_{\text{lasI}}$ break mass conservation of the signalling molecules and thus may increase the spatial averages \bar{C}_6 , \bar{C}_{12} . These terms can be visualised as vertical and horizontal vector fields components in the (c_6, c_{12}) plane respectively, which induce a drift on the local equilibria and the spatial average \bar{C}_6 , \bar{C}_{12} . Since there is no saturation of signal production, the spatial average \bar{C}_6 , \bar{C}_{12} will eventually always move outside of the bistable region, giving rise to a uniform dominant CFP or YFP profile. Within the finite observation time $t < 24$ h however, we may observe sharp boundaries forming, as the spatial average \bar{C}_6 , \bar{C}_{12} passes through the bistable region. Figure 36 shows how even monotonic gradients can facilitate boundary formation.



Supplementary Figure 36: Dynamics of the relay circuit for $k_{C12} > 0$, $k_{C6} = 0$ in the spatial domain x on the left and (c_6, c_{12}) plane on the right, showing additional reaction terms as red vectors driving the dynamics. This way even monotonic gradients can facilitate boundary formation

The additional parameters $k_{C6}, k_{C12}, d_{\text{luxI}}, d_{\text{lasI}}$ are tuned by hand. Simulations in Section 2.3 suggest that local cellular responses approach local equilibria faster than the diffusion timescales of signalling molecules. This means we can reasonably assume that our system is diffusion-limited and so reactive dynamics (37) are much faster than diffusive dynamics (36) and we can apply the quasi-steady state assumption to (36) yielding

$$\frac{\partial c_6(x, t)}{\partial t} = D_6 \frac{\partial^2 c_6(x, t)}{\partial x^2} + \frac{k_{C6}\rho_0 P_{81}}{\gamma_0 + d_{\text{luxI}}}, \quad \frac{\partial c_{12}(x, t)}{\partial t} = D_{12} \frac{\partial^2 c_{12}(x, t)}{\partial x^2} + \frac{k_{C12}\rho_0 P_{76}}{\gamma_0 + d_{\text{lasI}}}, \quad (38)$$

and therefore in effect only two ratio parameter ratios need to be tuned. These ratios determine the magnitude of the relay reactions in the (c_6, c_{12}) plane, which determine whether diffusion had contracted the local equilibria to the spatial average \bar{C}_6 , \bar{C}_{12} before or after having crossed into bistable region. Boundaries only form if contraction to \bar{C}_6 , \bar{C}_{12} occurred after entering the bistable region. See supplementary movies 6-7 for examples of these cases.

References

- [1] Boyan Yordanov, Neil Dalchau, Paul K Grant, Michael Pedersen, Stephen Emmott, Jim Haseloff, and Andrew Phillips. A computational method for automated characterization of genetic components. *ACS synthetic biology*, 3(8):578–588, 2014.
- [2] Timothy J. Rudge, James R. Brown, Fernan Federici, Neil Dalchau, Andrew Phillips, James W. Ajioka, and Jim Haseloff. Characterization of Intrinsic Properties of Promoters. *ACS Synthetic Biology*, 5(1):89–98, jan 2016.
- [3] Paul K Grant, Neil Dalchau, James R Brown, Fernan Federici, Timothy J Rudge, Boyan Yordanov, Om Patange, Andrew Phillips, and Jim Haseloff. Orthogonal intercellular signaling for programmed spatial behavior. *Molecular Systems Biology*, 12(1):849, 2016.
- [4] Neil Dalchau, Paul K. Grant, Prashant Vaidyanathan, Carlo Spaccasassi, Colin Gravill, and Andrew Phillips. Scalable dynamic characterization of synthetic gene circuits. *bioRxiv*, page 635672, aug 2019.
- [5] Geoffrey Roeder, Paul K Grant, Andrew Phillips, Neil Dalchau, and Edwards Meeds. Efficient amortised bayesian inference for hierarchical and nonlinear dynamical systems. In *International Conference on Machine Learning (ICML 2019)*, 2019.
- [6] ML Urbanowski, CP Lostroh, and EP Greenberg. Reversible acyl-homoserine lactone binding to purified vibrio fischeri luxr protein. *Journal of Bacteriology*, 186(3):631–637, 2004.
- [7] Romain Veltz. PseudoArcLengthContinuation.jl, March 2019.

Appendix B

Parameter Inference with Bifurcation Diagrams

Supplementary Material

A Bifurcation Diagrams as Tangent Fields

Let each component of the vector function F_θ in the model (1) implicitly define a surface embedded in \mathbb{R}^{N+1} . Let's assume that the intersection of these N surfaces exists and is not null or degenerate, then the steady states of (1) must be a set of one dimensional space curves in $z \in \mathbb{R}^{N+1}$ defined by

$$F_\theta(z) = 0 \quad (\text{A.1})$$

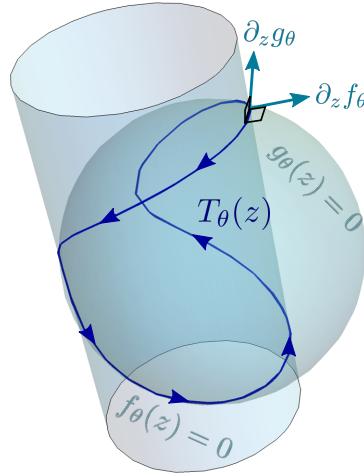


Figure A.1: Two implicit surfaces $f_\theta(z) = 0$ and $g_\theta(z) = 0$ in \mathbb{R}^3 intersecting to form a space curve which is tangent to field $T_\theta(z)$ and perpendicular to gradients $\partial_z f_\theta$ and $\partial_z g_\theta$

An expression for the field $T_\theta(z)$ tangent to the set of curves would allow us to take derivatives and integrals along the bifurcation curve. This is exactly what we need to do to evaluate our cost function 8. Fortunately the tangent field can be constructed by ensuring it is perpendicular to the gradient ∂_z of each component of F_θ as illustrated by an example two component system in Figure A.1. The tangent field $T_\theta(z)$ can be constructed perpendicular to all gradient vectors using the properties of the determinant [35]

$$T_\theta(z) := \begin{vmatrix} \hat{z} \\ \partial_z F_\theta \end{vmatrix} \quad T_\theta : \mathbb{R}^{N+1} \rightarrow \mathbb{R}^{N+1} \quad (\text{A.2})$$

$$= \sum_{i=1}^{N+1} \hat{z}_i (-1)^{i+1} \left| \frac{\partial F_\theta}{\partial (z \setminus z_i)} \right| \quad (\text{A.3})$$

where \hat{z} is a collection of unit basis vectors in the \mathbb{R}^{N+1} space and $\partial_z F_\theta$ is an $N \times (N+1)$ rectangular Jacobian matrix of partial derivatives and $z \setminus z_i$ denotes the N dimensional vector z with component z_i removed. This construction ensures perpendicularity to any gradients of F_θ

$$T_\theta(z) \cdot \partial_z f_\theta = \left| \frac{\partial_z f_\theta}{\partial_z F_\theta} \right| = 0 \quad \forall f_\theta \in F_\theta \quad (\text{A.4})$$

since the determinant of any matrix with two identical rows or columns is zero. Note that the tangent field $T_\theta(z)$ is actually defined for all values of z where adjacent field lines trace out other level sets where $F_\theta(z) \neq 0$. Furthermore deformations with respect to θ are always orthogonal to the tangent

$$T_\theta(z) \cdot \frac{dT_\theta}{d\theta} = 0 \quad (\text{A.5})$$

Figure A.2 shows how the bifurcation curve defined by $F_\theta(z) = 0$ picks out one of many level sets or traces in tangent field $T_\theta(z)$ for the saddle and pitchfork. The tangent field $T_\theta(z)$ can always be

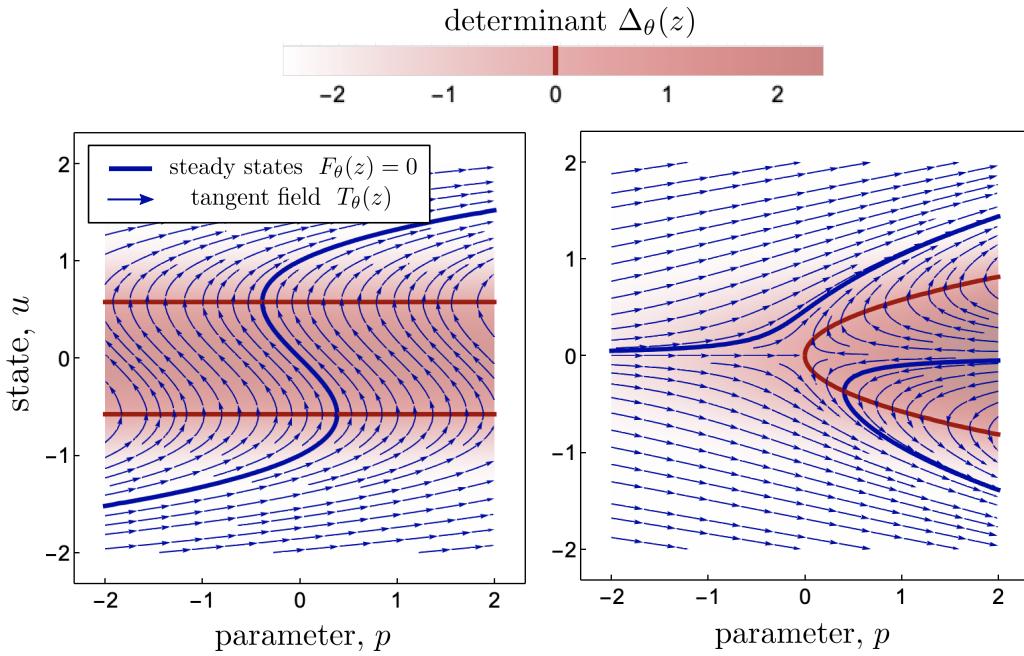


Figure A.2: Left/Right : Determinant $|\frac{\partial F_\theta}{\partial u}|$ and tangent field $T_\theta(z)$ for the saddle-node/pitchfork models for some set values of θ revealing that $|\frac{\partial F_\theta}{\partial u}| = 0$ defines bifurcations

analytically evaluated by taking the determinant in (A.2). We will proceed with calculations on $T_\theta(z)$ in the whole space z and pick out a single trace by solving $F_\theta(z) = 0$ later. For our two models

$$\begin{aligned} T_\theta(z) &= \hat{u} - (3\theta_2 u^2 + \theta_1) \hat{p} & T_\theta(z) &= u \hat{u} - (3\theta_2 u^2 + p) \hat{p} \\ \text{saddle-node model} & & \text{pitchfork model} & \end{aligned} \quad (\text{A.6})$$

Figure A.2 reveals that $|\frac{\partial F_\theta}{\partial u}| = 0$ is also a level set and that the intersection with level set $F_\theta(z) = 0$ defines the bifurcations at specific parameter θ . In this particular setting we can see that the tangent field $T_\theta(z)$ only folds when $|\frac{\partial F_\theta}{\partial u}| = 0$. Plotting the value of the determinant along $F_\theta(z) = 0$ from Figure A.2 would give rise to Figures 1. The directional derivative of the determinant $|\frac{\partial F_\theta}{\partial u}|$ along the tangent field $T_\theta(z)$ is defined as

$$\frac{d}{ds} \left| \frac{\partial F_\theta}{\partial u} \right| := \hat{T}_\theta(z) \cdot \frac{\partial}{\partial z} \left| \frac{\partial F_\theta}{\partial u} \right| \quad (\text{A.7})$$

where $\hat{T}_\theta(z)$ is the unit tangent field.

B Bifurcation Measure Properties

Consider a vector $v(s) \in \mathbb{R}^N$ parametrised by $s \in \mathbb{R}$ that is tangent to an equilibrium manifold defined by $F_\theta(u) = 0$. The conditions for a non-degenerate static bifurcation at s^* along such a tangent can be expressed in terms of an eigenvalue $\lambda(s)$ of the state-space Jacobian crossing zero with a finite slope. A bifurcation exists at s^* if

$$\frac{\partial F_\theta}{\partial u} v(s) = \lambda(s) v(s) \quad \exists \lambda : \quad \lambda(s)|_{s=s^*} = 0 \quad \frac{d\lambda}{ds} \Big|_{s=s^*} \neq 0 \quad (\text{B.1})$$

These conditions are necessary and sufficient for a non-degenerate static local breakdown of stability. For now we do not consider dynamic bifurcations involving limit cycles or imaginary parts of eigenvalues and restrict $\lambda \in \mathbb{R}$. Cases where both $\lambda(s)|_{s=s^*} = 0$ and $\frac{d\lambda}{ds}|_{s=s^*} = 0$ require investigation into higher order derivatives $\frac{d^n \lambda}{ds^n}$. These are the cases we refer to as *degenerate* and are not considered here.

Instead of considering conditions on each eigenvalue individually it is possible to use the determinant of the state-space Jacobian to detect whether the conditions (B.1) are satisfied. The determinant can

be expressed as the product of eigenvalues

$$\left| \frac{\partial F_\theta}{\partial u} \right| = \prod_{n=1}^N \lambda_n(s) \quad (\text{B.2})$$

Applying the product rule when differentiating yields

$$\frac{d}{ds} \left| \frac{\partial F_\theta}{\partial u} \right| = \sum_{n=1}^N \frac{d\lambda_n}{ds} \prod_{n' \neq n} \lambda_{n'}(s) \quad (\text{B.3})$$

$$= \left| \frac{\partial F_\theta}{\partial u} \right| \sum_{n=1}^N \frac{d\lambda_n}{ds} \lambda_n(s)^{-1} \quad (\text{B.4})$$

Substituting this expression into measure (5)

$$\varphi_\theta(s) = \left(1 + \left| \sum_{n=1}^N \frac{d\lambda_n}{ds} \lambda_n(s)^{-1} \right|^{-1} \right)^{-1} \quad (\text{B.5})$$

Which implies the following

$$\exists \lambda : \begin{cases} \lambda(s) = 0 & \frac{d\lambda}{ds} \neq 0 \\ \lambda(s) \neq 0 & \frac{d\lambda}{ds} \rightarrow \pm\infty \end{cases} \implies \varphi_\theta(s) = 1 \quad (\text{B.6})$$

If there exists an eigenvalue that satisfies conditions (B.1) then the measure is equal to one. The measure also approaches one in cases where the rate of change of an eigenvalue with respect to a manifold s location diverges while not crossing zero. This gives rise to finite gradients in the eigenvalue term in regimes far away from any bifurcation.

C Leibniz Rule for Space Curves

Suppose there exists a one dimensional space curve $\mathcal{C}(\theta)$ embedded in $z \in \mathbb{R}^{N+1}$ whose geometry changes depending on input parameters $\theta \in \mathbb{R}^M$. This curve could be open or closed and changes in θ could change the curve topology as well. Let the function $\gamma_\theta : \mathbb{R} \rightarrow \mathbb{R}^{N+1}$ be a parametrisation of the position vector along the curve within a fixed domain $s \in \mathcal{S}$. Note that the choice of parametrisation is arbitrary and our results should not depend on this choice. Furthermore, if we parametrise the curve $\mathcal{C}(\theta)$ with respect to a fixed domain \mathcal{S} the dependence on θ is picked up by the parametrisation $\gamma_\theta(s)$. We can write a line integral of any scalar function $L_\theta : \mathbb{R}^{N+1} \rightarrow \mathbb{R}$ on the curve as

$$L(\theta) := \int_{\mathcal{C}(\theta)} L_\theta(z) dz = \int_{\mathcal{S}} L_\theta(z) \left| \frac{d\gamma_\theta}{ds} \right| ds \Big|_{z=\gamma_\theta(s)} \quad (\text{C.1})$$

where $\left| \frac{d\gamma_\theta}{ds} \right|$ is the magnitude of tangent vectors to the space curve and we remind ourselves that the integrand is evaluated at $z = \gamma_\theta(s)$. We would like to track how this integral changes with respect to θ . The total derivative with respect to θ can be propagated into the integrand [34] as long as we keep track of implicit dependencies

$$\frac{dL}{d\theta} = \int_{\mathcal{S}} \left| \frac{d\gamma_\theta}{ds} \right| \left(\frac{\partial L}{\partial \theta} + \frac{\partial L}{\partial z} \cdot \frac{dz}{d\theta} \right) + L_\theta(z) \frac{d}{d\theta} \left| \frac{d\gamma_\theta}{ds} \right| ds \Big|_{z=\gamma_\theta(s)} \quad (\text{C.2})$$

Here we applied the total derivative rule in the first term due to the implicit dependence of z on θ through $z = \gamma_\theta(s)$. Applying the chain rule to the second term

$$\frac{d}{d\theta} \left| \frac{d\gamma_\theta}{ds} \right| = \left| \frac{d\gamma_\theta}{ds} \right|^{-1} \frac{d\gamma_\theta}{ds} \cdot \frac{d}{d\theta} \left(\frac{d\gamma_\theta}{ds} \right) \quad (\text{C.3})$$

By choosing an s that has no implicit θ dependence we can commute derivatives

$$\frac{d}{d\theta} \left(\frac{d\gamma_\theta}{ds} \right) = \frac{d}{ds} \left(\frac{d\gamma_\theta}{d\theta} \right) \implies \frac{d}{d\theta} \left| \frac{d\gamma_\theta}{ds} \right| = \left| \frac{d\gamma_\theta}{ds} \right|^{-1} \frac{d\gamma_\theta}{ds} \cdot \frac{d}{ds} \left(\frac{d\gamma_\theta}{d\theta} \right) \quad (\text{C.4})$$

To proceed we note that the unit tangent vector can be written as an evaluation of a tangent field $\hat{T}_\theta(z)$ defined in the whole domain $z \in \mathbb{R}^{N+1}$ along the parametric curve $z = \gamma_\theta(s)$. The unit tangent field may disagree with the tangent given by $\frac{d\gamma_\theta}{ds}$ up to a sign

$$\hat{T}_\theta(z) \Big|_{z=\gamma_\theta(s)} = \pm \left| \frac{d\gamma_\theta}{ds} \right|^{-1} \frac{d\gamma_\theta}{ds} \quad (\text{C.5})$$

this leads to

$$\frac{d}{d\theta} \left| \frac{d\gamma_\theta}{ds} \right| = \left| \frac{d\gamma_\theta}{ds} \right| \left(\hat{T}_\theta(z) \cdot \frac{\partial}{\partial z} \left(\frac{d\Gamma_\theta}{d\theta} \right) \cdot \hat{T}_\theta(z) \right)_{z=\gamma_\theta(s)} \quad (\text{C.6})$$

It is possible to find the normal deformation of the implicit space curves due to changes in θ . This can be done by taking the total derivative of the implicit equation defining the level set

$$\frac{dF_\theta(z)}{d\theta} = \frac{\partial F}{\partial \theta} + \frac{\partial F}{\partial z} \cdot \frac{dz}{d\theta} \quad (\text{C.7})$$

We can rearrange for $\frac{dz}{d\theta}$ using the Moore-Penrose inverse of the rectangular Jacobian matrix $\frac{\partial F}{\partial z}$ which appeared in equation (A.2). Since the level set is defined by $F_\theta(z) = 0$ the total derivative along the level set $dF_\theta(z) = 0$ and we arrive at an expression for the deformation field [28]

$$\frac{dz}{d\theta} = -\frac{\partial F^\top}{\partial z} \left(\frac{\partial F}{\partial z} \frac{\partial F^\top}{\partial z} \right)^{-1} \frac{\partial F}{\partial \theta} \quad (\text{C.8})$$

The tangential component of the deformation field is not uniquely determined because there is no unique way of parametrising a surface. This is the subject of many computer graphics papers [28, 36, 37]. We are however not interested in the continuous propagation of a mesh - as is the subject of those papers. In fact we are looking for a deformation field that is orthogonal to the tangent vector $\hat{T}_\theta(z) \cdot \frac{dz}{d\theta} = 0$ for the space curve, and therefore letting the tangential component of the deformation equal zero is a valid choice and we can it instead of the parametrised deformation

$$\frac{d\gamma_\theta}{d\theta} \rightarrow \frac{dz}{d\theta} \quad (\text{C.9})$$

To summarise we now have the gradient of our line integral only in terms of the implicit function defining the integration region.

$$\frac{dL}{d\theta} = \int_{F_\theta(z)=0} \frac{\partial L}{\partial \theta} + \frac{\partial L}{\partial z} \cdot \varphi_\theta(z) + L_\theta(z) \hat{T}_\theta(z) \cdot \frac{\partial \varphi}{\partial z} \cdot \hat{T}_\theta(z) dz \quad (\text{C.10})$$

$$\text{where } \hat{T}_\theta(z) := \frac{T_\theta(z)}{|T_\theta(z)|} \quad T_\theta(z) := \left| \begin{array}{c} \hat{z} \\ \partial_z F_\theta \end{array} \right| \quad \varphi_\theta(z) := -\frac{\partial F^\top}{\partial z} \left(\frac{\partial F}{\partial z} \frac{\partial F^\top}{\partial z} \right)^{-1} \frac{\partial F}{\partial \theta} \quad (\text{C.11})$$

We have settled on choosing normal deformations which we will call $\varphi_\theta(z)$. The above result can be seen as the generalised Leibniz rule [34] for the case of line integration regions. The last integrand term can be seen as the divergence the vector field $\varphi_\theta(z)$ projected onto the one dimensional space curve.

D Application of Bifurcation Inference to a Complex Model

To demonstrate the wider reaching applicability of our method we optimise the *double exclusive reporter* [3], a synthetic gene circuit in *E. coli* that was designed to exhibit a cusp bifurcation. The circuit behaviour is observed by measuring a fluorescent protein whose expression is controlled by transcription factors (regulatory proteins) LacI (L) and TetR (T), whose expression is in turn controlled by externally controllable *input* signals c_6 and c_{12} . To apply the method, we consider one of the input signals be the control condition $c_6 = p$, with the other packed together with the remaining 20 parameters into vector θ . Once the optima θ^* have been obtained, we perform dimensionality reduction using GigaSOM.jl [38] so that the results can be visualised in a two dimensional embedding (Figure D.1A).

The embedding reveals four optimal parameter regions. We find that, as with the two-state model in the main text (11), there are two qualitatively distinct regimes: mutual activation (region 1) and inhibition (regions 2-4). The mutual inhibition region can be further subdivided into three regions that are geometrically equivalent, but kinetically distinct: region 3 has swapped kinetic roles for regulatory proteins LacI and TetR compared to region 2, and region 4 has additional damped oscillations in the dynamics across the whole range of *input* c_6 (Figure D.1B). The two dimensional embedding of sampled optima θ^* enables navigation the space of qualitative behaviours of the *double exclusive reporter* and organisation in terms of geometric and kinetic equivalence.

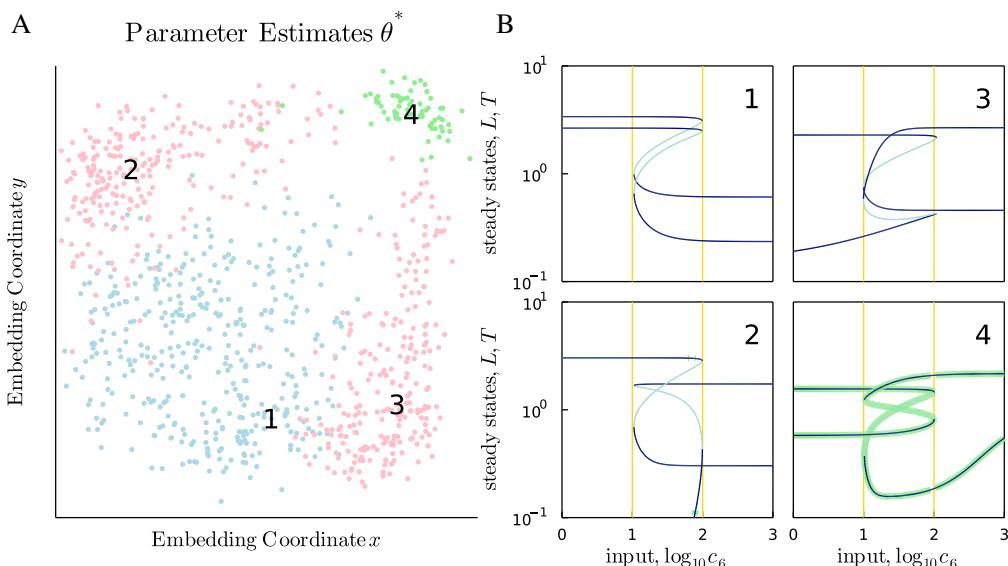


Figure D.1: Bifurcation inference for the *double exclusive reporter*. A. Optimal parameter estimates θ^* for the targets $\mathcal{D} = \{1, 2\}$ (indicated by yellow lines in panel B) reveal four regions with two geometrically different regimes: mutual activation (region 1) and mutual inhibition (regions 2-4). B. Example bifurcation diagrams indicate that region 2 has swapped kinetics between L and T to region 3. Region 4 has models with non-zero imaginary parts to eigenvalues indicating damped oscillations (shown in light green).

These results were obtained with a modification of the bifurcation measure (5) to improve convergence rates. In parameter regimes where bifurcations are not present, according to conditions (B.6), maximising the measure $\varphi_\theta(s)$ can lead to a divergence in directional derivative $\frac{d\lambda}{ds} \rightarrow \pm\infty$ rather than a creation of a bifurcation. To discourage this from happening we can flatten out the gradients in that regime by applying the tanh non-linearity to the determinant. This leads to

$$\varphi_\theta(s) := \left(1 + \left| \frac{\tanh \left| \frac{\partial F_\theta}{\partial u} \right|}{\frac{d}{ds} \tanh \left| \frac{\partial F_\theta}{\partial u} \right|} \right| \right)^{-1} \quad (\text{D.1})$$

E Extension for Hopf Bifurcations

In order to detect bifurcations involving limit cycles, the measure must be extended to detect changes in the real part $\Re[\lambda(s)]$ for any eigenvalue of the Jacobian. These conditions can no longer be compactly written in terms of the determinant. Instead, the measure can be defined as the sum of eigenvalue terms

$$\varphi_\theta(s) := \sum_{\lambda(s) \in \frac{\partial F_\theta}{\partial u}} \left(\left| \frac{d}{ds} \log \Re[\lambda(s)] \right|^{-1} + 1 \right)^{-1} \quad (\text{E.1})$$

The directional derivative of the logarithm diverges under two conditions: when eigenvalues vanish $\lambda(s) = 0$ and when the directional derivative $\frac{d}{ds} \Re[\lambda(s)]$ diverges. These properties are sufficient for detecting the onset of damped oscillations and emergence of limit cycles via Hopf bifurcation as shown in Figure E.1. Eigenvalues with negative real part which gain a finite imaginary part give rise to damped oscillations. At this onset we observe a discontinuity in the derivative $\frac{d}{ds} \Re[\lambda(s)]$ which is detected by equation (E.1). Once damped oscillations exist, flipping the stability of the stable fixed point gives rise to a limit cycle, which can be detected by inspecting $\Re[\lambda(s)]$.

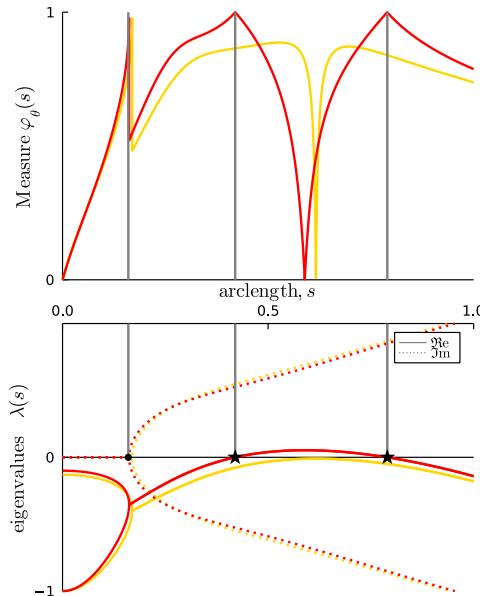


Figure E.1: Bifurcation measure $\varphi_\theta(s)$ and eigenvalues $\lambda(s)$ along the arclength s for two different bifurcation curves demonstrating how the measure detects non-zero imaginary parts $\Im[\lambda]$ (onset of damped oscillations marked by circle) and sign changes in real parts $\Re[\lambda]$ (Hopf bifurcations marked by stars)

In principle it is possible to construct measures to detect a variety of bifurcations as long as the conditions can be expressed in terms of derivatives with respect to fixed-point manifold direction s . Measures can be used sequentially or in parallel to encourage optimisers to run through a sequence of bifurcations or place specific bifurcation types next to each other.

Bibliography

- [1] A M Turing. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1):230–265, 1 1937.
- [2] Nicholas Metropolis. The beginning of the Monte Carlo method. *Los Alamos Science*, 15(584):125–130, 1987.
- [3] Dermot Turing. *X, Y and Z : The Real Story of How Enigma Was Broken*. The History Press, 2018.
- [4] Ashkan Shafiee and Anthony Atala. Tissue Engineering: Toward a New Era of Medicine. *Annual Reviews Medicine*, 68:29–40, 2017.
- [5] Kerstin Göpfrich, Ilia Platzman, and Joachim P. Spatz. Mastering Complexity: Towards Bottom-up Construction of Multifunctional Eukaryotic Synthetic Cells. *Trends in Biotechnology*, 36(9):938–951, 9 2018.
- [6] James Sharpe. Computer modeling in developmental biology: growing today, essential tomorrow. *Development (Cambridge, England)*, 144(23):4214–4225, 12 2017.
- [7] Sofia Triantafillou, Vincenzo Lagani, Christina Heinze-Deml, Angelika Schmidt, Jesper Tegner, and Ioannis Tsamardinos. Predicting Causal Relationships from Biological Data: Applying Automated Causal Discovery on Mass Cytometry Data of Human Immune Cells. *Scientific Reports*, 7(1):12724, 12 2017.

- [8] Jack W. Scannell, Alex Blanckley, Helen Boldon, and Brian Warrington. Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery*, 11(3):191–200, 3 2012.
- [9] John P A Ioannidis. Why most published research findings are false. *PLoS medicine*, 2(8):e124, 2005.
- [10] Alessandro Abate, Luca Cardelli, Marta Kwiatkowska, Luca Laurenti, and Boyan Yordanov. Experimental Biological Protocols with Formal Semantics. In *Lecture Notes in Computer Science*, volume 11095 LNBI, pages 165–182. Springer, Cham, 9 2018.
- [11] Shota Gugushvili and Chris A.J. Klaassen. root n-consistent parameter estimation for systems of ordinary differential equations: bypassing numerical integration via smoothing. *Bernoulli*, 18(3):1061–1098, 8 2012.
- [12] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Sparse Identification of Nonlinear Dynamics with Control (SINDYc). *IFAC-PapersOnLine*, 49(18):710–715, 5 2016.
- [13] Nico S Gorbach, Stefan Bauer, and Joachim M Buhmann. Scalable Variational Inference for Dynamical Systems. *ArXiv*, 2017.
- [14] Oana-Teodora Chis, Alejandro F. Villaverde, Julio R. Banga, and Eva Balsa-Canto. On the relationship between sloppiness and identifiability. *Mathematical Biosciences*, 282:147–161, 12 2016.
- [15] Attila Gábor, Alejandro F. Villaverde, and Julio R. Banga. Parameter identifiability analysis and visualization in large-scale kinetic models of biosystems. *BMC Systems Biology*, 11(1):54, 12 2017.
- [16] Alejandro F. Villaverde, Antonio Barreiro, and Antonis Papachristodoulou. Structural Identifiability of Dynamic Systems Biology Models. *PLOS Computational Biology*, 12(10):e1005153, 10 2016.

- [17] Luca Cardelli. Morphisms of reaction networks that couple structure to function. *BMC Systems Biology*, 8(1):84, 2014.
- [18] Luca Cardelli, Mirco Tribastone, Max Tschaikowski, and Andrea Vandin. Maximal aggregation of polynomial dynamical systems. *Proceedings of the National Academy of Sciences*, 114(38):10029–10034, 9 2017.
- [19] Luca Cardelli, Mirco Tribastone, Max Tschaikowski, and Andrea Vandin. ERODE: A Tool for the Evaluation and Reduction of Ordinary Differential Equations. In *Lecture Notes in Computer Science*, volume 10206 LNCS, pages 310–328. Springer, Berlin, Heidelberg, 2017.
- [20] J. Halatek and E. Frey. Rethinking pattern formation in reaction-diffusion systems. *Nature Physics*, page 1, 2 2018.
- [21] Geoffrey Roeder, Paul K Grant, Andrew Phillips, Neil Dalchau, and Edward Meeds. Efficient Amortised Bayesian Inference for Hierarchical and Nonlinear Dynamical Systems. *ICML*, pages 4445–4455, 5 2019.
- [22] Christopher Rackauckas and Qing Nie. Differentialequations.jl - a performant and feature-rich ecosystem for solving differential equations in julia. *Journal of Open Research Software*, 5(1), 2017.
- [23] Romain Veltz. BifurcationKit.jl. Technical report, Inria Sophia-Antipolis, 7 2020.
- [24] Matthias Seeger. Gaussian processes for machine learning. *International journal of neural systems*, 14(2):69–106, 2004.
- [25] Edward J Fuselier and Grady B Wright. A radial basis function method for computing Helmholtz-Hodge decompositions. *IMA Journal of Numerical Analysis*, 37(2):774–797, 2017.

- [26] P. K. Grant, N. Dalchau, J. R. Brown, F. Federici, T. J. Rudge, B. Yordanov, O. Patange, A. Phillips, and J. Haseloff. Orthogonal intercellular signaling for programmed spatial behavior. *Molecular Systems Biology*, 12(1):849–849, 1 2016.
- [27] L Wolpert. Positional information and the spatial pattern of cellular differentiation. *Journal of theoretical biology*, 25(1):1–47, 10 1969.
- [28] Melissa B. Miller and Bonnie L. Bassler. Quorum Sensing in Bacteria. *Annual Review of Microbiology*, 55(1):165–199, 10 2002.
- [29] J. B. A. Green and J. Sharpe. Positional information and reaction-diffusion: two big ideas in developmental biology combine. *Development*, 142(7):1203–1211, 4 2015.
- [30] James Lu, Heinz W. Engl, and Peter Schuster. Inverse bifurcation analysis: Application to simple gene systems. *Algorithms for Molecular Biology*, 1(1):1–16, 7 2006.
- [31] Attila Csikász-Nagy, Bela Novák, and John Tyson. Reverse engineering models of cell cycle regulation. *Advances in experimental medicine and biology*, 641:88–97, 2008.
- [32] Emery D Conrad, John Tyson, Reinhard Laubenbacher, John Phillips, and Michael Renardy. Bifurcation Analysis and Qualitative Optimization of Models in Molecular Cell Biology with Applications to the Circadian Clock. *Virginia Tech*, 4 2006.
- [33] A. M. Turing. The Chemical Basis of Morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 237(641):37–72, 8 1952.
- [34] Neil Dalchau, Gregory Szép, Rosa Hernansaiz-Ballesteros, Chris P. Barnes, Luca Cardelli, Andrew Phillips, and Attila Csikász-Nagy.

Computing with biological switches and clocks. *Natural Computing*,
page 1–19, 6 2018.