Quorum-Sensing Synchronization of Synthetic Toggle Switches: A Design based on Monotone Dynamical Systems Theory

Evgeni V. Nikolaev, Eduardo D. Sontag[*]

Department of Mathematics and Center for Quantitative Biology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8019, The United States

*Corresponding Author: eduardo.sontag@rutgers.edu

Abstract

Synthetic constructs in biotechnology, biocomputing, and modern gene therapy interventions are often based on plasmids or transfected circuits which implement some form of "on-off" switch. For example, the expression of a protein used for therapeutic purposes might be triggered by the recognition of a specific combination of inducers (e.g., antigens), and memory of this event should be maintained across a cell population until a specific stimulus commands a coordinated shut-off. The robustness of such a design is hampered by molecular ("intrinsic") or environmental ("extrinsic") noise, which may lead to spontaneous changes of state in a subset of the population and is reflected in the bimodality of protein expression, as measured for example using flow cytometry. In this context, a "majority-vote" correction circuit, which brings deviant cells back into the desired state, is highly desirable, and quorum-sensing has been suggested as a way for cells to broadcast their states to the population as a whole so as to facilitate consensus.

In this paper, we propose what we believe is the first such design that has mathematically guaranteed properties of stability and auto-correction. Our approach is guided by concepts and theory from the field of "monotone" dynamical systems developed by M. Hirsch, H. Smith, and others. We benchmark our design by comparing it to an existing design which has been the subject of experimental and theoretical studies, illustrating its superiority in stability and self-correction of synchronization errors. Our stability analysis, based on dynamical systems theory, guarantees global convergence to steady states, ruling out unpredictable ("chaotic") behaviors and even sustained oscillations. These results are valid no matter what are the values of parameters, and are based only on the wiring diagram. The theory is complemented by extensive computational bifurcation analysis, performed for a biochemically-detailed and biologically-relevant model that we developed. Another novel feature of our approach is that our theorems on stability of steady states for homogeneous or mixed populations are valid independently of the number of cells in the population, and depend only on the relative proportions of each type of state.

While monotone systems theory has been used previously for systems biology analysis, the current work illustrates its power for synthetic biology design, and thus has wider significance well beyond the application to the important problem of coordination of toggle switches.

Author Summary

For the last decade, outstanding progress has been made, and considerable practical experience has accumulated, in the construction of elementary genetic circuits that perform various tasks, such as memory storage and logical operations, in response to both exogenous and endogenous stimuli. Using modern molecular "plug-and-play" technologies, various (re-)programmable cellular populations can be engineered in a routine manner, and they can be combined into more complex cellular systems.

Among all engineered synthetic circuits, a *toggle*, a robust bistable switch leading to a binary response dynamics, is the simplest basic synthetic biology device, analogous to the "flip-flop" or latch in electronic design, and it plays a key role in biotechnology, biocomputing, and proposed gene therapies. However, despite many remarkable properties of the existing toggle designs, they must be tightly controlled in order to avoid spontaneous switching between different expression states (loss of long-term memory) or even the breakdown of stability through the generation of stable oscillations.

To address this concrete challenge, we have developed a new design for quorum-sensing synthetic toggles, based on monotone dynamical systems theory. Our design is endowed with strong theoretical guarantees that completely exclude unpredictable chaotic behaviors, as well as undesired stable oscillations, and leads to robust consensus states.

Introduction

In the short period since the pioneering milestones in synthetic biology [1,2], outstanding progress has been made, and considerable practical experience has accumulated, in the construction of genetic circuits that perform various tasks, such as memory storage and logical operations, as well as support biomedical interventions and biotechnological manipulations in response to both exogenous and endogenous stimuli [3–6]. These circuits often include plasmids or transfected circuits which implement some form of "on-off" binary device, generically referred to as a *toggle switch*. For example, the expression of a protein used for gene therapy could be triggered by the recognition of some combination of inducers such as antigens, and memory of this event should be maintained across a cell population until a specific stimulus commands a coordinated shut-off [1, 3, 4]. In this context, as well as in many others, it is desirable for populations of cells to achieve coordinated static and/or dynamic functionalities. However, this coordination is hampered by molecular ("intrinsic") or environmental ("extrinsic") noise, which may lead to spontaneous changes of state in a subset of the population and is reflected in the bimodality of protein expression, as measured for example using flow cytometry.

To achieve robustness across a population, one may implement a "majority-vote" correction circuit that brings deviant cells back into the desired state. Much synthetic biology research focuses on single-cell microorganisms, often bacteria [4,6]. Bacterial populations are relatively simple, and their sociality in all its complexity can be rationally understood [7], providing a foundation for building more complex cellular systems. For bacteria, quorum-sensing (QS) has been suggested as a way for cells to broadcast their states to the population as a whole so as to facilitate consensus. QS signaling pathways [8] can, for example, regulate bacterial gene expression in response to fluctuations in cell-population density. Bacteria produce and release various signaling molecules called autoinducers (AIs) [8–11]. The detection of a minimal threshold stimulatory concentration of an AI leads to an alteration in the host's gene expression. Both Gram-positive and Gram-negative bacteria use QS communication to regulate a diverse array of physiological activities. Synthetic biology design has adopted QS communication in its toolbox [12], because natural and artificially engineered QS modules can be used to interface synthetic circuits with exogenous and endogenous cues [4], and a systematic modular approach to standardize engineering toggle genetic circuits that would allow programmed cells to be designed for various specific purposes and to communicate their states to other cells was suggested as a bioengineering "plug-and-play" modular approach [4] . The design of such QS-toggle combinations is the focus of this paper.

A known design and its drawbacks

Kobayashi et al [4] consider a genetic toggle switch, interfaced with a QS signaling pathway. Their *E. coli* strain "B2" detects as well as produces (through the synthetase encoded by the expressed gene luxI, which converts common precursor metabolites) acyl-homoserine lactone (AHL) signaling molecules. AHL is a QS signaling pathway from *Vibrio fischeri*. Toggle B2 (Fig 1) enables an *E. coli* population to measure population density through AHL, because AHL signaling can be reversibly transported to the medium via diffusion, contributing to the AHL density in the culture [4].

To achieve an in-depth understanding of dynamic properties of coupled QS and toggle constructs, Kuznetsov et al. [13] developed and studied a mechanistic mathematical model of a population (or, equivalently, an ensemble) comprising N toggles, see Fig. 2 (bottom panel), corresponding
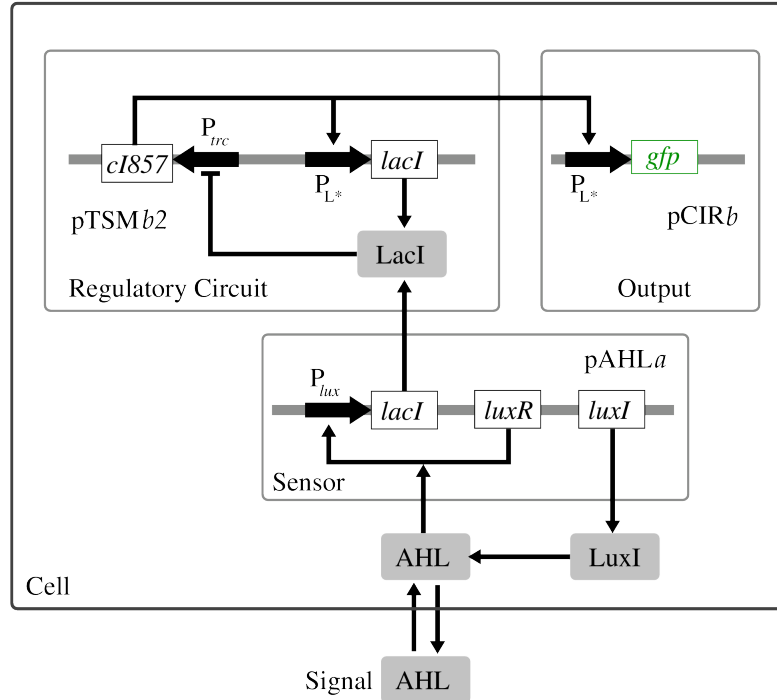
3

Figure 1. Toggle B2: Density-Dependent Gene Activation. *Notations and abbreviations.* Plasmids: pTSM*b2*, pCIR*b* and pAHL*a*; genes: *gfp, cI857, lacI, luxR*, and *luxI*; promoters: P$_{trc}$, P$_{L*}$, and P$_{lux}$; AHL, acyl-homoserine lactone. Figure adapted from Fig. 6(A) in [4].

to Toggle B2 (Fig. 1), $N > 1$. Their study revealed important multiple functions, namely bistability as well as stable oscillations, that an ensemble of Toggles B2 was capable of exhibiting. Analytical conditions for bistability were found, and a time separation was introduced to obtain a stable limit cycle for a population of interacting cells.

In bistable circuits (toggles), transitions such as those caused by fluctuations due to low copy numbers of species per cell, or due to local environmental "noise" can force individual cells to change expression state at random [4]. This noise effect can spontaneously lead to the emergence of heterogeneous (mixed) populations consisting of cells in different expression states, which appear as bimodal population distributions when the corresponding protein levels are measured [4]. To investigate the effect of a spontaneous toggle switching in single and coupled cellular systems, leading to bimodal population distributions, Wang et al. [14] developed models for a single cell and a multi-cellular toggle system comprising N cells, respectively. In their models, the dynamics of the repressor proteins LacI and λ CI is described by the two ODE equations developed in [1]. The AI-interfacing employed in the population model [14] corresponds to a signaling pathway which is slightly different from the signaling pathway in Toggle B2 (Fig. 1) as suggested in [4] and described earlier. The main difference is in the description of the expression of the gene that encodes LuxI, see [14] for more details.

Despite the remarkable properties of design B2 and its modifications, observed experimentally in controllable experimental settings [4], and studied theoretically [13, 14], the fact that their functional repertoire includes not only a bistable long-term memory but also the generation of stable oscillations suggests that the environment-toggle system must be tightly controlled in order to avoid spontaneous switching, not merely between different expression states but even

between different functions.

To address this challenge, we propose a novel design, which is endowed with mathematically guaranteed properties of stability and auto-correction. Our approach is closely guided by concepts and theory from the powerful framework of *monotone dynamical systems* pioneered by M. Hirsch and H. Smith [15–21].

We benchmark our design by comparing it to Toggle B2, which, as discussed, has been subject of experimental and theoretical studies, illustrating its superiority in stability and self-correction of synchronization errors. We employ monotone theory to provide guarantees of global convergence to steady states, thus ruling out unpredictable ("chaotic") behaviors and sustained oscillations. These theorems are valid no matter for all values of parameters and are based only on the network structure. We also provide an extensive computational bifurcation analysis for a biochemically-detailed and biologically-relevant. Our results for homogeneous or mixed populations are valid independently of the number of cells in the population, and depend only on the relative proportions of each type of state.

The components

As a basic design, we chose a genetic toggle switch consisting of two mutually repressing genes, *lacI* and *tetR* [1]. We use two acylated homoserine lactones (Acyl-HSLs), *(i)* N-butanoyl-l-homoserine lactone (C4-HSL) secreted by *Pseudomonas aeruginosa* [22], and *(ii)* N-(3-hydroxy-7-*cis*-tetradecenoyl)-L-homoserine lactone (3-OH-C14-HSL) produced by *Rhizobium leguminosarum* [10] as a means of coordinating toggle-host activity. Our design has two QS arms built-in the toggle in such a way that each promoter-repressor pair is controlled by its own QS signaling pathway symmetrically. Because of this "mirror-like" toggle symmetry, we call our design a symmetric toggle or an "S" design.

To benchmark the new S toggle design and the monotone systems approach, we compare the S design to the well-studied asymmetric B2-strain (Fig. 1) which has one QS arm only [4, 13]. In this work, we call the asymmetric B2-strain the "A" design. Our S design cannot be reduced to the A design by removing one QS arm, and, thus, the S design cannot be viewed as a straightforward extension of the A design. From a theoretical standpoint, it is worth remarking that the A design is non-monotone.

The S vs. A toggle comparative results obtained in this work can be summarized as follows:

- The monotone S toggle design completely excludes any unpredictable chaotic behaviors, as well as stable oscillations. Typical trajectories converge globally to stable equilibria. This conclusion is valid for all parameter values, and provides a strong theoretical guarantee missing from other synthetic biology designs.

- We refer to mixed states leading to bimodal distributions as *spontaneous synchronization errors*. We find that the S toggle design is able to self-correct (or, auto-correct) synchronization errors, while the non-monotone A toggle design is not.

- We show how monotone systems theory can predict not only the dynamics of an S toggle population but it also explains certain monotonically increasing or decreasing parametric dependencies of population steady states. We show that some of these predictions can facilitate self-synchronization and, thereby, reduce any chance for synchronization errors to emerge spontaneously.

Organization of paper

In Models and Methods, the S toggle and A toggle mathematical models are introduced. The basic formalism and fundamental mathematical results of monotone systems theory, including Strong Monotonicity and Hirsch's Theorem [15–19, 21] are reviewed there as well. Balanced graphs, relation to graph partitions, and order presentation by flows are explained there. Additionally, reference values of dimensionless parameters, a scaling procedure, and the selection and interpretation of bifurcation parameters are discussed, and the mirror symmetry of the S toggle is formalized. The symmetry will be used to explain symmetry breaking (pitchfork) bifurcations in the S toggle. We also formalize a concept of spontaneous synchronization errors. Here, three types of equilibrium populations are considered: One homogeneous population, and two heterogeneous (mixed) populations (bimodal distributions) with both equally (1:1)-mixed and not-equally (N_1:N_2)-mixed transcriptional signatures, $N_2 \ll N_1$, the latter giving rise to spontaneous synchronization errors, where $N = N_1 + N_2$, and N is the number of cells in the given population.

In Results and Discussion, we proceed to a comparative theoretical and computational analysis of the S toggle and A toggle design models. We begin this section with results on the application of monotone systems theory to the S design, as these results constitute the main conceptual and practical subject motivating this work (Application of Monotone Systems Theory to the S design). We start by explaining how monotone systems theory allows one to predict, based on qualitative knowledge only, that generically all solutions converge to equilibria, with no possible oscillations [13] nor chaotic behavior [23], no matter what the kinetic parameters are. This is in contrast to the A design, which may admit oscillations [13]. Next, we analyze single S and A toggles decoupled from the environment (Bistability in Single S-Design and A-Design Toggles Isolated from the Environment), and observe that the S toggle remains bistable even if "redundant" repressor genes are removed from the corresponding plasmids. To show how the S design is more robust than the A design, we carry out a comparative bifurcation analysis of populations consisting of coupled S or A toggles. We select a free (bifurcation) dimensionless parameter which can be interpreted in terms of experimental interventions [6] leading to *(a)* changes in the membrane permeability, or *(b)* changes in the half-lives of repressor proteins, or *(c)* changes in the specific growth rate of the host cell. We additionally test the toggle design capabilities to self-correct spontaneous synchronization errors by sampling the basin of attractor of the corresponding equilibrium solutions. We find that the S toggle design is able to self-correct synchronization errors far better than the A toggle design.

The paper also has three Supplemental Information (SI) materials. In S1 Model Derivation, we derive mathematical models and carry out a nondimensionalization (scaling) procedure, the conclusions of which are used in the main text (Scaling). In S2 Estimation of Parameter Values, we discuss ranges of biologically meaningful parameter values based on data available in the existing literature. Values of biologically meaningful parameters depend upon experimental conditions and other factors controlled by an experimenter, as reviewed in [6]. Therefore, we provide an example of a concrete estimation of values of dimensionless parameters, which we interpret in terms of interventions reviewed in [6]. Since bistability only holds for ranges of these dimensionless parameters, to increase the chances that both S and A-synthetic toggles will function in the bistable regime, we have tuned the estimated values of reference dimensionless parameters to pin a region in the parameter space from which we choose appropriate parameters values for subsequent modeling studies. In S3 Exponential Stability of Cellular Populations, we prove a number of general theorems to analyze exponential stability [24] of both homogeneous

6

and heterogeneous (mixed) population equilibrium states, independently of the number N of cells in the given population, which (*i.e.,* the value of $N \geq 2$) can be *a priori* unknown.

Models and Methods

Although our main objective in this paper is to present a conceptual and general organizing principle for the construction of self-correcting "majority-vote" multistable synthetic systems, we instantiate our ideas through a very concrete set of genes and protein products, all being standard molecular parts in synthetic biology [1, 2, 4, 25–29]. We do that in order to emphasize the fact that our constructs can be realistically implemented with currently available molecular components. However, replacing these components with others does not change the basic mathematical framework.

To facilitate a conceptual and quantitative comparison of the S and A toggle designs, the corresponding genetic circuits are assumed to be built from the same tightly controlled *lac-tet* transcription entities [30–35], which have been intensively used in a number of experimental and theoretic-modeling studies in the context of synthetic biology [1, 2, 4, 26–29]. Below, we briefly characterize relevant molecular details and then form the corresponding mathematical models.

Toggle Designs

For the sake of completeness of our description, we begin our discussion of the S toggle and A toggle designs (Fig. 2) with two classical orthogonal repressors (Table 1):

  I. LacI from *E. coli* which inhibits the transcription of the second repressor gene, tetR from the tetracycline-resistance transposon Tn10;

  II. TetR which represses the transcription of the first repressor gene *lacI*.

Next, the communication network among all toggles (Fig. 2) is built upon two quorum-sensing (QS) signaling molecules (Table 1):

  1. N-(3-hydroxy-7-*cis*-tetradecenoyl)-L-homoserine lactone (3-OH-C14-HSL);

  2. N-butanoyl-l-homoserine lactone (C4-HSL).

For the sake of brevity, the QS signaling molecules are called autoinducers G (C14-HSL) and R (C4-HSL). Note that the G- and R-signals (acylated homoserine lactones) are natural biological signals secreted by Gram-negative bacteria, including *E. coli*, as a means of coordinating cellular activity [4, 8].

Finally, to drive the autoinducer concentrations, two synthases are used (Table 1):

  (a) CinI, the gene product of *cinI*, driving the concentration of C14-HSL;

  (b) RhlI, the gene product of *rhlI*, driving the concentration of C4-HSL.

Using the above molecular species, we implement and study two different toggle designs called *symmetric* (S) and *asymmetric* (A) designs, respectively, (Fig. 2):
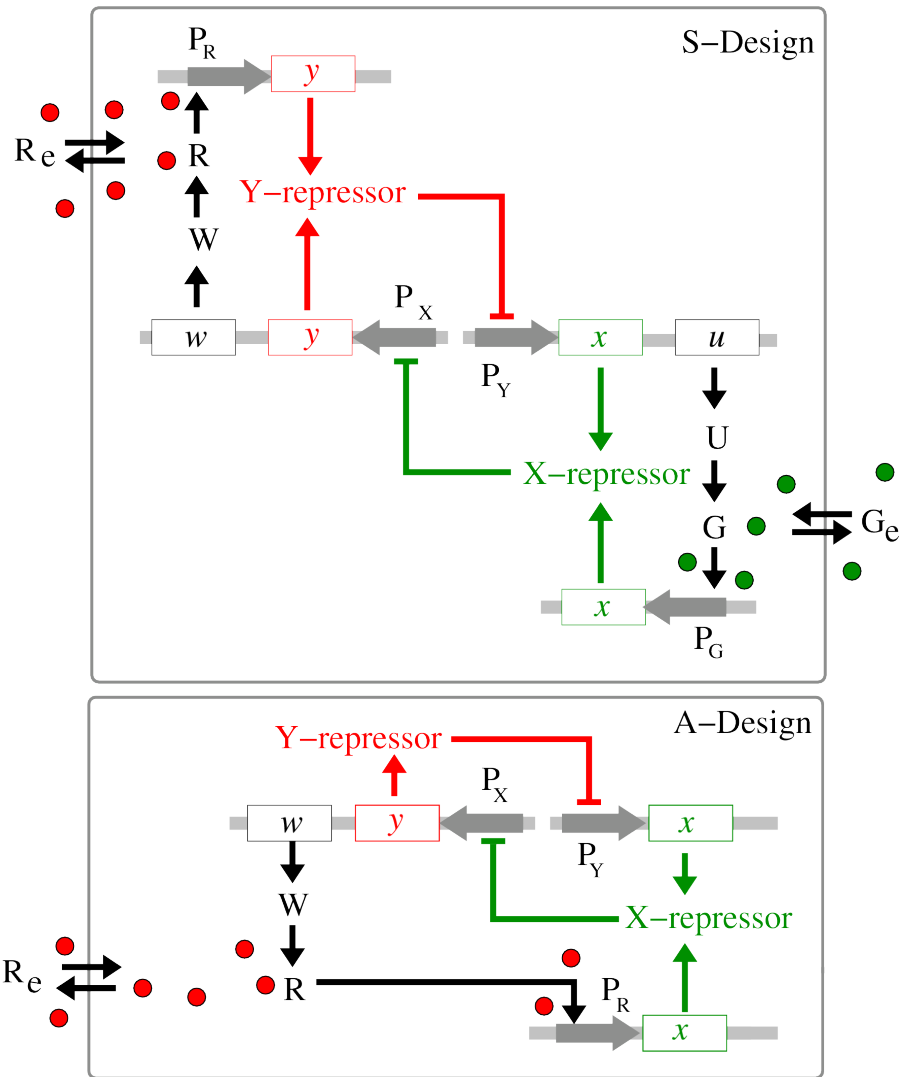
Figure 2. Monotone-symmetric and nonmonotone-asymmetric toggle designs. S design (*top panel*): Activation of the expression of gene x (*lacI*) occurs by binding of autoinducer G (C14-HSL) to promoter P_G (P_{cin}). Inhibition of the expression of both genes x (*lacI*) and u (*cinI*) occurs by binding of the gene product Y (TetR) of gene y (*tetR*) to a single promoter P_Y (P_{tet}). Symmetrically, activation of the expression of gene y (*tetR*) occurs by binding of autoinducer R (C4-HSL) to promoter P_R (P_{rhl}), while inhibition of the transcription of both genes y (*tetR*) and w (*rhlI*) occurs by binding of X (LacI) to a single promoter P_X (P_{lac}). **A design** (*bottom panel*): Activation of the expression of gene x (*lacI*) occurs by binding of autoinducer R (C4-HSL) to promoter P_R (P_{rhl}). Expression of genes y (*tetR*) and w (*rhlI*) is driven by a common single promoter P_X. Gene products U and W are synthases CinI and RhiI, respectively. Gray horizontal strips correspond to integration plasmids.

8

(i) In the S design, each of the two autoinducers activates symmetrically the transcription of the corresponding repressor gene through a single promoter, that is, promoter P_{cin} (P_G) for gene *lacI* (x) and promoter P_{rhl} (P_R) for gene *tetR* (y);

(ii) In the A design, the same repressor genes (as used in the S design) antagonistically repress one another directly, while there is only one autoinducer that asymmetrically facilitates communication between all toggles.

Table 1. A toggle molecular part catalog (explanations of variables are given in S Model).

Name	variable	Function	Description	References
lacI	−	repressor gene	lactose-inducible transcriptional repressor from *E. coli*	[1, 2, 6, 31]
tetR	−	repressor gene	from the tetracycline-resistance transposon Tn10	[1, 2, 6, 31, 34]
cinI	−	autoinducer gene	encodes protein CinI which synthesizes C14-HSL	[8, 10, 11]
rhlI	−	autoinducer gene	encodes protein RhlI which synthesizes C4-HSL	[8, 9]
LacI	x_i	lactose inhibitor	a DNA-binding protein encoded by *lacI*	[1, 2, 6, 30, 31, 33, 35]
TetR	y_i	repressor protein	a basic element of tetracycline-controlled regulation	[1, 2, 6, 32, 34]
CinI		synthase	the gene product of gene *cinI*	[8, 10, 11]
RhlI		synthase	the gene product of gene *rhlI*	[8, 9]
C14-HSL	g_i, g_e	autoinducer	N-(3-hydroxy-7-*cis*-tetradecenoyl)-L-Homoserine Lactone	[10]
C4-HSL	r_i, r_e	autoinducer	N-butyryl-L-Homoserine Lactone	[9, 22]

The genetic circuit topology used in the A design (Fig. 2) is taken from [13]. In order to keep making a fair comparison with the S design, we have replace the *luxA-luxI* system considered in [13] by the *lacI-tetR* system suggested in [1]. Note that both CinI and RhiI are homologous to LuxI [36].

To host the S and A toggles, we use *E. coli*, a bacterial cell which has been well-studied in a huge number of relevant experimental and modeling works [30, 37–46], and which has been widely used to implement and test various synthetic circuits [1, 2, 4, 12]. A practical modeling reason for this selection is narrowing-down our search for biologically-meaningful parameters to values known from the *E. coli* studies. However, our conclusions do not depend in any way on biological properties of the host.

As a readout of the toggle state in individual cells, we further assume that each *E. coli* cell contains a gene encoding a spectrally distinct fluorescent reporter, GFP for gene *lacI*, and RFP for gene *tetR*, driven by promoters that respond to the autoinducers C14-HSL and C4-HSL, respectively. We do not simulate the processes of bio-synthesis and degradation of the fluorescent proteins explicitly, using appropriate cascade models, for two reasons: *(i)* the "reporter" submodel does not affect the dynamics of the entire model, and *(ii)* the half-lives of the reporter proteins can be made similar to the half-lives of the repressor proteins [2].

Finally, because each toggle can either be in a state where *(a)* LacI protein is abundant, while TetR protein is scarce, or in a state where *(b)* TetR protein is abundant, while LacI protein is scarce, we call state *(a)* a green state or, simply, a G-state and state *(b)* a red state or, simply, an R-state, respectively.

9

S Model

A mathematical model describing a population of identical S toggles is

$$\frac{dx_i}{dt} = \frac{a_1}{1 + y_i^{n_Y}} + \frac{a_3\, g_i^{n_G}}{1 + g_i^{n_G}} - x_i, \tag{1a}$$

$$\frac{dy_i}{dt} = \frac{a_2}{1 + x_i^{n_X}} + \frac{a_4 r_i^{n_R}}{1 + r_i^{n_R}} - y_i, \tag{1b}$$

$$\frac{dg_i}{dt} = \frac{a_5}{1 + y_i^{n_Y}} + d\big(g_e - g_i\big) - \delta_g\, g_i, \tag{1c}$$

$$\frac{dr_i}{dt} = \frac{a_6}{1 + x_i^{n_X}} + d\big(r_e - r_i\big) - \delta_r\, r_i, \quad i = 1, \ldots, N, \tag{1d}$$

$$\frac{dg_e}{dt} = \frac{\rho}{N} \sum_{i=1}^{N} d\big(g_i - g_e\big) - \delta_e\, g_e, \quad 0 \le \rho \le 1, \tag{1e}$$

$$\frac{dr_e}{dt} = \frac{\rho}{N} \sum_{i=1}^{N} d\big(r_i - r_e\big) - \delta_e\, r_e. \tag{1f}$$

Here, all state variables and parameters are dimensionless, and are obtained from the corresponding biologically meaningful state variables and parameters describing the *lac-tet* system (Table 1) after an appropriate nondimensionalization carried out in S1 Model Derivation.

In the S model (1), t is dimensionless time; x_i and y_i are the dimensionless concentrations (levels) of intracellular repressor proteins LacI and TetR, respectively; g_i and r_i are the dimensionless concentrations of intracellular autoinducers C14-HSL and C4-HSL, respectively; g_e and r_e are the dimensionless concentrations of extracellular autoinducers C14-HSL and C4-HSL, respectively.

The dimensionless rate constants a_i, $i = 1, \ldots, 6$, depend on the copy numbers of the plasmids that bear the corresponding genes, see relationships (4) given in Scaling; n_X, n_Y, n_G, and n_R are the corresponding Hill coefficients; d is the dimensionless diffusion coefficient; δ_g and δ_r are the dimensionless lumped dilution-degradation rates due to the exponential cell growth and degradation of the corresponding species.

The degradation rate constants for repressor species x_i and y_i are scaled out to unity, as it is done in [1, 2, 13, 14], see S2 Estimation of Parameter Values; δ_e is the dilution rate due to flow in the medium; ρ is a population density; and N is the number of cells in the given population.

A-Model

A dimensionless mathematical model describing a population of identical A toggles is

$$\frac{dx_i}{dt} = \frac{a_1}{1 + y_i^{n_Y}} + \frac{a_4\, r_i^{n_R}}{1 + r_i^{n_R}} - x_i, \tag{2a}$$

$$\frac{dy_i}{dt} = \frac{a_2}{1 + x_i^{n_X}} - y_i, \tag{2b}$$

$$\frac{dr_i}{dt} = \frac{a_6}{1 + x_i^{n_X}} + d\big(r_e - r_i\big) - \delta_r\, r_i, \quad i = 1, \ldots, N, \tag{2c}$$

$$\frac{dr_e}{dt} = \frac{\rho}{N} \sum_{i=1}^{N} d\big(r_i - r_e\big) - \delta_e\, r_e. \tag{2d}$$

10

Here, all state variables and parameters are as defined for the S model (1). A similar model was developed in [13] for Toggle B2 shown in Fig. 1.

Model Parameters

Uncertainty about the values of parameters characterizing molecular components of synthetic circuits always presents a significant difficulty in circuit design [2]. Here, we discuss reference values of dimensionless parameters obtained using an appropriate scaling procedure. We also explain how we select and interpret parameters for our bifurcation analysis.

Reference Values of Model Parameters

Reference values of all parameters used in our modeling studies are estimated in S2 Estimation of Parameter Values, and these correspond to half-lives of all proteins in the range 4-10 min., which are close to a typical mRNA half-life in *E. coli* [2]. Also, to avoid competition for ribosomes [41], only a few plasmids bearing four promoters P_X, P_Y, P_G, and P_R can be used, and we use 1-2 copies per cell, see S2 Estimation of Parameter Values. The *E. coli* replication period is assumed to be around 25 min.

Despite the fact that much is known about *E. coli* [37–39, 43–46], it is not possible to model behavior in a quantitatively precise way, since not enough is yet known about molecular interactions between the toggle and the host cell to make such a description realistic [6]. Instead, we hope to identify classes of toggle designs and dynamic behaviors to determine which of the designs could lead to an improved self-synchronization reliability and an improved capability for self-correction of spontaneous synchronization errors, when a small fraction of cells flips to the opposite (undesirable) transcriptional signature state, see Spontaneous Synchronization Errors. We will also make some predictions that might help to facilitate engineering toggles with desired robust traits.

In our computational analysis, the following set of reference parameter values is used:

$$a_1 = a_2 = 20, \quad a_3 = a_4 = 10, \quad a_5 = a_6 = 3, \tag{3a}$$

$$n_X = n_Y = n_G = n_R = 3, \tag{3b}$$

$$\delta_g = \delta_r = 1.0, \quad \delta_e = 0.5, \quad \rho = 0.8. \tag{3c}$$

Groups of parameters with identically the same values are used to introduce the toggle mirror symmetry into the S model as discussed in Symmetry. We find that the working values of parameters estimated in (3) are within the range of equivalent parameters (rate constants, Hill coefficients, *etc.*) used earlier for genetic circuits built from similar (*e.g.*, homologous) molecular entities [1, 2, 4, 13, 14, 26–29, 47].

Finally, we note that the above choices of the parameter values can also be viewed biologically meaningful because both the strengths of the four different promoters, P_X, P_Y, P_G, and P_R, and the degradation/dilution rates can be set accordingly by appropriate experimental interventions [1, 2], reviewed in [6] and also discussed below, see Scaling and Selection and Interpretation of Bifurcation Parameters.

Scaling

One of the goals of a model nondimensionalization and scaling is to reduce the number of (correlated) parameters by lumping original parameters into a smaller parameter set. In this case, interpretation of changes in the values of dimensionless parameters should be done carefully, as the set of non-dimensionless parameters is usually not in one-to-one correspondence with the set of original parameters. For example, mathematical models used for synthetically engineered systems often contain parameters representing multiple biological parts and, so, tuning a dimensionless parameter in the corresponding mathematical model can be implemented experimentally in a number of different ways [6].

The dimensional and dimensionless parameters used in the S and A toggle models are related to one another by the following relationships (see S1 Model Derivation):

1. For the dimensionless rate parameters, we obtain:

$$a_1 = \frac{b_x \, k_x \, [\mathrm{P_Y}]}{K_\mathrm{X} \, (r_d + \mu)}, \quad a_2 = \frac{b_y \, k_y \, [\mathrm{P_X}]}{K_\mathrm{Y} \, (r_d + \mu)}, \quad a_3 = \frac{b_x \, k_x \, [\mathrm{P_G}]}{K_\mathrm{X} \, (r_d + \mu)}, \tag{4a}$$

$$a_4 = \frac{b_y \, k_y \, [\mathrm{P_R}]}{K_\mathrm{Y} \, (r_d + \mu)}, \quad a_5 = \frac{b_u \, k_u \, k_\mathrm{G} \, [\mathrm{P_Y}]}{K_\mathrm{G} \, (r_d + \mu)^2}, \quad a_6 = \frac{b_w \, k_w \, k_\mathrm{R}[\mathrm{P_X}]}{K_\mathrm{R} \, (r_d + \mu)^2}. \tag{4b}$$

2. For the dimensionless diffusion and degradation parameters, we obtain:

$$d_g = \frac{D_\mathrm{G}}{r_d + \mu}, \quad d_r = \frac{D_\mathrm{R}}{r_d + \mu}, \tag{5a}$$

$$\delta_g = \frac{r_\mathrm{G} + \mu}{r_d + \mu}, \quad \delta_r = \frac{r_\mathrm{R} + \mu}{r_d + \mu}, \quad \delta_e = \frac{\mu_e}{r_d + \mu}. \tag{5b}$$

Let us briefly discuss (4) and (5). Here, the burst parameter b_x for the protein X or, equivalently, LacI, depends on the efficiency of translation, controlled by strength of ribosome-binding sites (RBS) [1, 6], and the mRNA half-life time [48]; $[\mathrm{P_X}]$ is the number of promoters per cell for gene x; k_x is an average transcription rate for gene x (lacI); K_X is the number of LacI proteins required to half-maximally repress $\mathrm{P_{lac}}$; k_G is the maximum production rate of C14-HSL by CinI, D_G is the export rate of C14-HSL; μ is the intracellular specific dilution rate due to the host cell growth, $\mu = \ln 2/T$, T is the division period. Parameters for other proteins and QS signaling molecules are defined similarly, see S1 Model Derivation and S2 Estimation of Parameter Values. Based on the fact that N-Acyl Homoserine Lactone Lactonase (AHL-lactonase) hydrolyzes C4-HSL effectively [49], we also assume that specific degradation rate constants for the signaling molecules, C14-HSL and C4-HSL, can be set experimentally [6], corresponding to the parameter values used in our models. We pick these specific promoters and autoinducers in order to be concrete and to justify biologically meaningful values of the model parameters. However, we wish to emphasize that our results are generic for the architectures shown in Fig. 2.

Selection and Interpretation of Bifurcation Parameters

In our bifurcation analysis, we use almost all dimensionless parameters given in (3) as free parameters allowed to be varied to detect changes in stability of the corresponding solutions. Whenever a new bifurcation point is detected, we provide an appropriate interpretation in terms of interventions reviewed in [6], which can potentially lead to the corresponding effect.

For example, suppose that d is the free parameter used in bifurcation analysis. Due to the relationships (5), changes in the values of d may correspond to different and independent experimental interventions [6] leading to: *(a)* changes in the membrane permeability (*i.e.*, D_G and D_R), or *(b)* changes in the half-lives of repressor proteins, or *(c)* changes in the specific growth rate of the host cell. As such, both higher values of protein half-lives and diffusion permeability as well as lower values of the specific growth rate (longer replication periods) correspond to higher values of the parameter d. Recall that the value of the parameter d characterizes the strength of the interaction between cells in the given population, which facilitates self-synchronization [12–14, 23, 50].

More broadly, we can rely upon the fact that all dimensionless parameters are defined via appropriate combinations of the original dimensional parameters (4) and (5) in our interpretation of results obtained from bifurcation analysis as follows.

The values of dimensionless rate parameters (*i.e.*, a-parameters) can be changed by decreasing or increasing translational efficiency, which depends on the nucleotide sequence of the ribosome binding sites (RBS) located within the upstream noncoding part of the mRNA [46]. The RBS is encoded by the DNA sequence immediately upstream of the start codon of the gene and is an independent regulatory element that can be manipulated experimentally [1]. The values of dimensionless rate parameters can also be changed by decreasing or increasing the lifetime values of appropriate proteins. Indeed, a carboxy-terminal tag, based on ClpX, the ATP-dependent unfoldase/translocase of ClpXP recognizing specific protein substrates bearing *ssrA* tags [51], can be inserted at the 3W end of each repressor gene [2]. Proteases in *E. coli* recognize this tag and target the attached protein for destruction. Such tags are used to reduce the half-life of the proteins from more than 60 min to around 4 min, which makes it possible and (also convenient) to set the half-life times for all toggle proteins (approximately) equal to one another and close to the half-lives of mRNAs [2, 12]. To this end, both RBS and carboxy-terminal tags are the principal tools by which the parameters of an engineered gene network can be adjusted experimentally [1, 2, 6].

Symmetry

In this section, we formalize the symmetry of

- a single S toggle embedded into an environment ($N = 1$),

- a population (network) of N-identical S toggles interacting via a common environment ($N \geq 2$).

As is routine in physics and engineering, symmetry-based simplifications often lead to important insights into complex phenomena, and we also use symmetry to discuss bifurcations in S toggle populations. We observe that under a special condition imposed on the parameter values,

$$a_1 = a_2,\ a_3 = a_4,\ a_5 = a_6,\ d_g = d_r,\ \delta_g = \delta_r, \tag{6}$$

a single S toggle embedded into an external environment is described by the S model (1) with $N = 1$ which has a $\mathbb{Z}_2$-symmetry group generated by involution $\mathcal{I}$ [52, 53],

$$\mathcal{I} : (x, y, g, r, g_e, r_e) \longrightarrow (y, x, r, g, r_e, g_e). \tag{7}$$

13

Consider the fixed-point subspace $\text{Fix}\,(\mathbb{Z}_2) \in \mathbb{R}^6$ of the group $\mathbb{Z}_2$, see [52],

$$\text{Fix}\,(\mathbb{Z}_2) \;=\; \{z = (x, y, g, r, g_e, r_e) \in \mathbb{R}^6 \,|\, \mathcal{I}z \equiv z\}. \tag{8}$$

We ignore the trivial equilibria that belong to $\text{Fix}\,(\mathbb{Z}_2)$, that is, equilibria of the S model (1) for which the following equalities hold, $x = y$, $g = r$, and $g_e = r_e$, corresponding to identically the same levels of LacI and TetR, and C14-HSL and C4-HSL, respectively. Let us denote the equilibrium of the S model (1) by z_0, and let us assume that $z_0 \notin \text{Fix}\,(\mathbb{Z}_2)$. Now, because the S model is invariant with respect to the involution (7), and because $z_0 \notin \text{Fix}\,(\mathbb{Z}_2)$, we obtain that both z_0 and $\mathcal{I}\,z_0$, $\mathcal{I}\,z_0 \neq z_0$, are different equilibria of the S model (1), see [52,53]. The equilibria z_0 and $\mathcal{I}\,z_0$ are called relative equilibria [54]. All bifurcations for the relative equilibria occur simultaneously at the same values of free parameters. We generalize G- and R-homogeneous populations states as relative equilibria, which means that as soon as the S toggle has a G-state, it will also have the corresponding R-state, implying bistability.

The general case of $N \geq 2$ is slightly more complicated as a population of identical S toggles has a symmetry group obtained after combinations of permutations among all cells in the given population and the toggle involution (7), which we denote $\mathcal{G} = \mathbb{Z}_2 \times \mathbf{S}_N$ for brevity. Here, $\mathbf{S}_N$ is a symmetric group of order $N \geq 1$, and $\mathbb{Z}_2$ is the toggle involution (7) applied to all toggles simultaneously. For example, for $N = 2$, we will have one permutation,

$$\mathcal{P} \,:\, (x_1, y_1, g_1, r_1, x_2, y_2, g_2, r_2, g_e, r_e) \;\longrightarrow\; (x_2, y_2, g_2, r_2, x_1, y_1, g_1, r_1, g_e, r_e), \tag{9}$$

and the involution,

$$\mathcal{I} \,:\, (x_1, y_1, g_1, r_1, x_2, y_2, g_2, r_2, g_e, r_e) \;\longrightarrow\; (y_1, x_1, r_1, g_1, y_2, x_2, r_2, g_2, r_e, g_e). \tag{10}$$

Spontaneous Synchronization Errors

Capabilities of toggles to fail and recover from spontaneous synchronization errors can be formalized in terms of a *multistability* concept, that is, as a co-existence of bistable *homogeneous* populations and various *heterogeneous* populations (Fig. 3), also called *mixed* states, under the same conditions. Recall that mixed states are known to lead to bistable distributions [4].

Following [4], we call a population heterogeneous or, equivalently, mixed if it comprises toggles with *different* transcription signatures for the *same* genes: *(i)* the repressor gene *lacI* is active (G-state), while *tetR* is repressed, and *(ii) lacI* is repressed, while the repressor gene *tetR* is active (R-state), see Toggle Designs. In other words, a homogeneous population is fully characterized by either transcription signature *(i)* or *(ii)*, while a heterogeneous population is characterized by mixed signatures *(i)* or *(ii)* simultaneously present in the population (Fig. 3).

Different heterogeneous populations can be characterized by transcription signature "mixtures" with ratio $(p{:}q)$, $p + q = 1$, describing the fraction of toggles in the G-state versus the fraction of toggles in the R-state within the same population. For homogeneous populations, we, therefore, have either (1:0) or (0:1) transcriptional signature fulfilled by all toggles in the population (Fig. 3).

With these concepts, we can formulate more precisely our objective: to find conditions under which heterogeneous (mixed) population equilibrium solutions can loose their stability or can even be eliminated completely.

As a proof of concept, an example of an (9:1)-heterogeneous population (Fig. 3) will be used, where the number of toggles in the first, Green-subpopulation (G) (*tetR* is suppressed) is 9 times
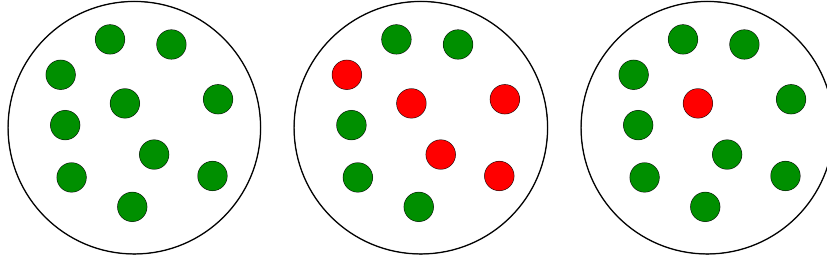
Figure 3. Homogeneous and heterogeneous (mixed) populations. An example of a population consisting of 10 cells is shown. The left panel demonstrates a homogeneous G-population. The center panel demonstrates a heterogeneous (1:1)-population, where the homogeneous G- and R-subpopulations have equal number of cells. The right panel demonstrates a heterogeneous (9:1)-population formed of two unequal subpopulations which represent a *spontaneous synchronization error*, when one or a few toggles spontaneously flip from green (G) to red (R) states. Therefore, our modeling goal is to find conditions under which the population can recover from the heterogeneous (right panel) case to the homogeneous (left panel) case. We will also try to eliminate cases similar to the one shown in the center panel which we, however, do not interpret as a synchronization error.

bigger that the number of toggles in the second, Red-subpopulation (R) (*lacI* is suppressed). In this simplest case, the G-subpopulation comprises 9 cells ($p = 0.9$ or 90%-fraction of all cells), while the R-subpopulation comprises one cell ($q = 0.1$ or 10%-fraction of all cells).

Note that our analysis of (9:1)-mixed states does not depend on the number of cells N in the entire population, which is usually unknown in experiments. In other words, our results hold for any integers N, N_1, and N_2, such that $N = N_1 + N_2$, and $N_1 : N_2 = 9 : 1$, where the fractions of cells with different transcription signatures are defined by the numbers $p = N_1/N$ and $q = N_2/N$, respectively, see S3 Exponential Stability of Cellular Populations.

Monotone Systems Formalism

The systems considered here are described by the evolution of states, which are time-dependent vectors $x(t) = (x_1(t), \ldots, x_n(t))$. The components x_i represent concentrations of chemical species (such as proteins, mRNA, metabolites, and so forth), the dynamics of which are given by a system of ODE's:

$$
\begin{aligned}
\frac{dx_1}{dt}(t) &= f_1(x_1(t), x_2(t), \ldots, x_n(t)), \\
\frac{dx_2}{dt}(t) &= f_2(x_1(t), x_2(t), \ldots, x_n(t)), \\
&\vdots \\
\frac{dx_n}{dt}(t) &= f_n(x_1(t), x_2(t), \ldots, x_n(t)).
\end{aligned}
$$

We also write simply $dx/dt = f(x)$, where f is a differentiable vector function with components f_i. The coordinates $x_i(t)$ are non-negative numbers. We write $\varphi(t, x_0)$ for the solution of the initial value problem $\dot{x}(t) = f(x(t))$ with $x(0) = x_0$, or just $x(t)$ if x_0 is clear from the context, and assume that this solution $x(t)$ exists and remains bounded for all $t \geq 0$.

15

Definition of Monotone Systems

A system is said to be *monotone* if there exists a partition of the set of indices of state variables:

$$\{1, 2, \ldots, n\} \;=\; S_+ \bigcup S_- \quad (S_+ \bigcap S_- = \emptyset)$$

with the following properties:

(1) for each pair of indices $\{i, j\} \in S_+$ $(i \neq j)$ and each pair of indices $\{i, j\} \in S_-$ $(i \neq j)$,

$$\frac{\partial f_i}{\partial x_j}(x) \geq 0 \quad \forall\, x$$

(2) and for each pair of indices $\{i, j\}$ such that $i \in S_+$ and $j \in S_-$ and each pair of indices $\{i, j\}$ such that $i \in S_-$ and $j \in S_+$,

$$\frac{\partial f_i}{\partial x_j}(x) \leq 0 \quad \forall\, x\,.$$

Observe that the definition does not impose any constrains on diagonal entries $\frac{\partial f_i}{\partial x_i}(x)$. These may have arbitrary signs, even depending on x.

Monotone systems [55–57] were introduced by Hirsch, and constitute a class of dynamical systems for which a rich theory exists. (To be precise, we have only defined the subclass of systems that are "monotone with respect to some orthant order" but the notion of monotone dynamics can be defined with respect to more general orders.)

We assume from now on that our system satisfies the following property: for each pair of distinct nodes i and j, one of these holds:

1. $\frac{\partial f_i}{\partial x_j}(x) > 0$ for all states x

2. $\frac{\partial f_i}{\partial x_j}(x) < 0$ for all states x

3. $\frac{\partial f_i}{\partial x_j}(x) = 0$ for all states x.

Of course, there are many models for which partial derivatives may change sign depending on the particular point x. With assumptions (1-3), however, the main results that we need from monotone dynamical systems theory will be particularly easy to state.

Monotone systems cannot admit any stable oscillations [16, 58, 59]. Under a stronger property, described next, only convergence to steady states is generically possible.

Strong Monotonicity

The directed *species influence graph* G associated to a system with n state variables is defined as follows. The graph G has n nodes (or "vertices"), which we denote by $v_1, \ldots, v_n$, one node for each species. If

$$\frac{\partial f_i}{\partial x_j} > 0 \ (\text{activation}),$$

16

we introduce an edge labeled "1" from v_j into v_i. If, instead,

$$\frac{\partial f_i}{\partial x_j} < 0 \text{ (inhibition)},$$

we introduce an edge labeled "-1" (or just "$-$") from v_j into v_i. Finally, no edge is drawn from node v_j into node v_i if the partial derivative $\frac{\partial f_i}{\partial x_j}(x)$ vanishes identically (no direct effect of the jth species upon the ith species). An alternative is to write a normal arrow "$\rightarrow$" or a blunted arrow "$\dashv$" (or an arrow labeled "$-$") respectively for the first two cases. The graph G is an example of a *signed graph* [60], meaning that its edges are labeled by signs.

No self-edges (edges from a node v_i to itself) are included in the graph G, whatever the sign of the diagonal entry $\partial f_i / \partial x_i$ of the Jacobian. The sign of this derivative may be positive, negative, or even be state-dependent. Results will not depend on signs of diagonals of the Jacobian of f.

The graph G is said to be *strongly connected* if, given an arbitrary pair of different indices $\{i, j\}$, there is a some, possibly indirect, effect of i on j. Formally, we ask that there is a sequence of indices $i = k_0, k_1, \ldots, k_r = j$ such that

$$\frac{\partial f_{k_{s+1}}}{\partial x_{k_s}} \neq 0 \quad \text{for } s = 0, \ldots, r-1 \,.$$

A system is said to be *strongly monotone* if it is monotone and, in addition, its species influence graph G is strongly connected. (As with the definition of monotonicity, one can extend strong monotonicity to far more general classes of systems, but we use a more restrictive notion that makes results less technical to state.) Even when there are multiple steady-states, the Hirsch Generic Convergence Theorem [55–57, 59] is a fundamental result.

Hirsch's Theorem

Even though they may have arbitrarily large dimensionality, monotone systems behave in many ways like one-dimensional systems: Hirsch's Theorem asserts that generic bounded solutions of strongly monotone differential equation systems must converge to the set of (stable) steady states. "Generic" means here "for every solution except for a measure-zero set of initial conditions." In particular, no nontrivial attractors arise. The genericity qualifier is needed in order to exclude the unstable manifolds of saddles as well as behavior on lower-dimensional sets [15].

Order Preservation by Flows of Monotone Systems

We defined monotonicity using graph partitions because this is the easiest way to present the concept. However, the usual definition found in textbooks is not phrased in that form. We explain here how our definition is equivalent to the usual one as well as to another property. For further remarks on these equivalences, see [20]. A signed graph (such as the species influence graph obtained by looking at signs of Jacobain entries) G is said to be *balanced* (see Harary [61]) if every undirected closed loop in the graph G has a net positive sign, that is to say, an even number, possibly zero, of negative arrows. Equivalently, any two (undirected) paths between two nodes must have the same net sign. By undirected loops or paths, we mean that one is allowed to transverse an edge either forward or backward. A *spin assignment* Σ for the graph G is an assignment, to each node v_i, of a number σ_i equal to "$+1$" or "-1" (a "spin," to borrow

from statistical mechanics terminology). If there is an edge from node v_j to node v_i, with label $J_{ij} \in \{\pm 1\}$, we say that this edge is *consistent with the spin assignment* Σ provided that:

$$J_{ij}\sigma_i\sigma_j = 1$$

which is the same as saying that $J_{ij} = \sigma_i\sigma_j$, or that $\sigma_i = J_{ij}\sigma_j$. An equivalent formalism is that in which edges are labeled by "0" or "1," instead of 1 and -1 respectively, and edge labels J_{ij} belong to the set $\{0, 1\}$, in which case consistency is the property that $J_{ij} \oplus \sigma_i \oplus \sigma_j = 0$ (sum modulo two). One says that Σ *is a consistent spin assignment for the graph* G (or simply that G is consistent) if every edge of G is consistent with Σ. In other words, for any pair of vertices v_i and v_j, if there is a positive edge from node v_j to node v_i, then v_j and v_i must have the same spin, and if there is a negative edge connecting v_j to v_i, then v_j and v_i must have opposite spins. (If there is no edge from v_j to v_i, this requirement imposes no restriction on their spins.) It is easy to see that if there is a consistent spin assignment for G, then the graph is balanced. Conversely, if G is balanced then there is a consistent spin assignment for G: to see this, simply label one node arbitrarily, and follow paths to label other nodes consistently. (If the graph is not connected, repeat the procedure in each connected component.)

For any spin assignment Σ, let A_1 be the subset of nodes labeled $+1$, and let A_{-1} be the subset of nodes labeled -1. The set of all nodes is partitioned into A_1 and A_{-1}. Conversely, any partition of the set of nodes into two subsets can be thought of as a spin assignment. With this interpretation, a consistent spin assignment is the same as a partition of the node set into two subsets A_1 and A_{-1} in such a manner that all edges between elements of A_1 are positive, all edges between elements of A_{-1} are positive, and all edges between a node in A_1 and a node in A_{-1} are negative. In summary, our definition of monotonicity, given in terms of partitions of state variables, amounts to the same as the requirement that there exist at least one consistent spin assignment for its associated graph G, or equivalently, that its graph G is balanced.

Supposing that a system is monotone, with a consistent spin assignment $\Sigma = \{\sigma_i, i = 1, \ldots, n\}$, we introduce following the relation among vectors $x \in \mathbb{R}^n_{\geq 0}$:

$$x \preceq y$$

means that

$$\sigma_i x_i \leq \sigma_i y_i \quad i = 1, \ldots, n \,.$$

This is a componentwise inequality that requires $x_i \leq y_i$ if node i has been assigned a positive spin, and $x_i \geq y_i$ if instead node i has been assigned a negative spin. Let $y(t)$ and $z(t)$ be any two solutions of the system $dx/dt = f(x)$, and suppose that $\sigma_i y_i(0) \leq \sigma_i z_i(0)$ for each $i = 1, \ldots, n$. Then, *Kamke's Theorem* states that $\sigma_i y_i(t) \leq \sigma_i z_i(t)$ for all $t \geq 0$ and coordinate $i = 1, \ldots, n$. This is the usual definition of monotonicity: if states start at time zero in a certain order, then they must remain forever in the same order. Conversely, a flow that preserves an order of this type must be monotone in the sense that we have defined the concept. See the textbook [57] for a proof, and [18] for extensions with systems with external inputs.

Results and Discussion

To carry out computational bifurcation analysis, MatCont [62, 63] has been used. A technical description of bifurcation points can be found in [53, 62–64].

Application of Monotone Systems Theory to the S design

To apply monotone systems theory to the S toggle model (1), we first rewrite the model in the following convenient general form with $4\,N + 2$ variables:

$$\frac{dx_i}{dt} = h_x(x_i, y_i, g_i),$$
$$\frac{dy_i}{dt} = h_y(x_i, y_i, r_i),$$
$$\frac{dg_i}{dt} = h_g(y_i, g_i, g_e),$$
$$\frac{dr_i}{dt} = h_r(x_i, r_i, r_e),$$
$$\frac{dg_e}{dt} = H_g(g_e, g_1, \ldots, g_n),$$
$$\frac{dr_e}{dt} = H_r(r_e, r_1, \ldots, r_n).$$

Here, $i = 1, \ldots, n$, all the functions in the right-hand side are differentiable, and the following signs hold for partial derivatives, everywhere in the state space:

$$\frac{\partial h_x}{\partial x_i} < 0, \quad \frac{\partial h_x}{\partial y_i} < 0, \quad \frac{\partial h_x}{\partial g_i} > 0, \quad \frac{\partial h_x}{\partial a_1} > 0, \quad \frac{\partial h_x}{\partial a_3} > 0, \tag{11a}$$

$$\frac{\partial h_y}{\partial x_i} < 0, \quad \frac{\partial h_y}{\partial y_i} < 0, \quad \frac{\partial h_y}{\partial r_i} > 0, \quad \frac{\partial h_y}{\partial a_2} > 0, \quad \frac{\partial h_y}{\partial a_6} > 0, \tag{11b}$$

$$\frac{\partial h_g}{\partial y_i} < 0, \quad \frac{\partial h_g}{\partial g_i} < 0, \quad \frac{\partial h_g}{\partial g_e} > 0, \quad \frac{\partial h_g}{\partial a_5} > 0, \quad \frac{\partial h_g}{\partial \delta} < 0, \tag{11c}$$

$$\frac{\partial h_r}{\partial x_i} < 0, \quad \frac{\partial h_r}{\partial r_i} < 0, \quad \frac{\partial h_r}{\partial r_e} > 0, \quad \frac{\partial h_g}{\partial a_6} > 0, \quad \frac{\partial h_r}{\partial \delta} < 0, \tag{11d}$$

$$\frac{\partial H_g}{\partial g_i} > 0, \quad \frac{\partial H_g}{\partial g_e} < 0, \quad \frac{\partial H_g}{\partial \delta_e} < 0, \tag{11e}$$

$$\frac{\partial H_r}{\partial r_i} > 0, \quad \frac{\partial H_r}{\partial r_e} < 0, \quad \frac{\partial H_r}{\partial \delta_e} < 0, \quad i = 1 \ldots, n. \tag{11f}$$

Next we observe that the S system is monotone, because we may partition its state variables as follows. One set consists of

$$x_i,\ g_i,\ g_e,\ i = 1, \ldots, n\,, \tag{12}$$

and another set consists of

$$y_i,\ r_i,\ r_e,\ i = 1, \ldots, n\,. \tag{13}$$

Moreover, the corresponding graph is strongly connected, as we have the following paths, for each two indices i, j:

$$x_j \dashv r_j \to r_e \to r_i \to y_i \dashv g_i \to g_e \to g_i \to x_i \tag{14}$$

which shows that one can reach any node from any other node by means of a directed path. Thus, the S model (1) is strongly monotone. We conclude as follows.

Theorem 1. *Typical solutions of the S model* (1) *converge to steady states.*

19

This fundamental result is robust to parameters as well as to the functional form of the equations. It insures that our proposed deign has theoretically guaranteed global stability properties. No stable oscillations [13] can exist, nor can other (for, example, "chaotic" [23]) solution regimes arise. In addition to these global properties, it is also possible to use the theory of monotone systems in order to make qualitative predictions about bifurcation diagrams as discussed in the next section.

The monotonicity property of the S system has important consequences regarding its transient as well as asymptotic behavior. We discuss in an appendix how Kamke's Theorem characterizes order relations for monotone systems. We explain now what these mean, explicitly, for the S system. Let $z_i(t)$ characterize the state of the i-th S toggle at time $t \geq 0$, that is, $z_i(t) = (x_i(t), y_i(t), g_i(t), r_i(t))$, $i = 1, \ldots, N$. Let $Z(t)$ characterize the state of the population of cells, $Z(t) = (z_1(t), \ldots, z_N(t), g_e(t), r_e(t))$. Suppose that we have two initial sets, $Z(0)$ and $\tilde{Z}(0)$, of values for the various expression levels of the repressor proteins, LacI and TerR, and we consider the behavior of $Z(t)$ and $\tilde{Z}(t)$ for $t > 0$.

Now suppose that we wish to understand what is the effect of a perturbation in one of the components of the initial state $z_i(0)$ for S toggle i with some fixed i, $1 \leq i \leq N$. (A similar argument can be applied to perturbations in other components of the initial state, or even simultaneous perturbations in all the components.) Suppose, for example, that we are interested in understanding the behavior starting from a state in which $\tilde{x}_3(0) \geq x_3(0)$ in the 3rd toggle z_3. This gives rise to a new population-wide solution $\tilde{Z}(t)$, and we use a tilde to denote its coordinates, that is, $\tilde{Z}(t) = (\tilde{z}_1(t), \ldots, \tilde{z}_N(t), \tilde{g}_e(t), \tilde{r}_e(t))$, where $\tilde{z}_i(t) = (x_i(t), y_i(t), g_i(t), r_i(t))$, $i = 1, \ldots, N$. Then, using the information provided by the partition shown in (12) and (13), we can predict that, for all $t > 0$: $\tilde{x}_i(t) \geq x_i(t)$, $\tilde{y}_i(t) \leq y_i(t)$, $\tilde{g}_i(t) \geq g_i(t)$, $\tilde{r}_i(t) \leq r_i(t)$, $\tilde{g}_e(t) \geq g_e(t)$, and $\tilde{r}_e(t) \leq r_e(t)$ for all $i = 1, \ldots, N$. As we will see shortly below, a similar conclusion can also be made with respect to perturbations in parameters, not merely initial states.

Monotone Parametric Dependencies in the S design

As a first step, we can include the nine parameters, a_i $(i = 1, \ldots, 6)$, δ_g, δ_r, and δ_e, as constant state variables by formally adding the corresponding equations $da_i/dt = 0$ $(i = 1, \ldots, 6)$, and $d\delta_g/dt = \delta_r/dt = d\delta_e/dt = 0$ to the S-model (1). The extended S-model is a monotone system. Dependencies between the S-model state variables and parameters (11) as shown in Fig. 4.

The extended S system has no strong monotonicity property, because the nodes corresponding to the parameters cannot be reached from other nodes, as the parametric extension violates the strong connectivity relationships (14). However, this is not of any consequence, as the global stability properties of the S system are determined by constant values of the parameters. We only introduced the extended system in the context of bifurcation analysis.

One might add additional constant variables to represent other parameters, such as the d's. These other parameters do not lead to monotonicity, and this lack of monotonicity will have important consequences in bifurcation analysis, as we discuss later.

The monotonicity of the extended system implies that stable loci in bifurcation diagrams depend monotonically on parameter variations. They will increase when the parameter being varied belongs to the component as the variable being analyzed, and will decrease if they are in different components. This property is a consequence of the general order preserving properties of monotone systems, as we explain now.
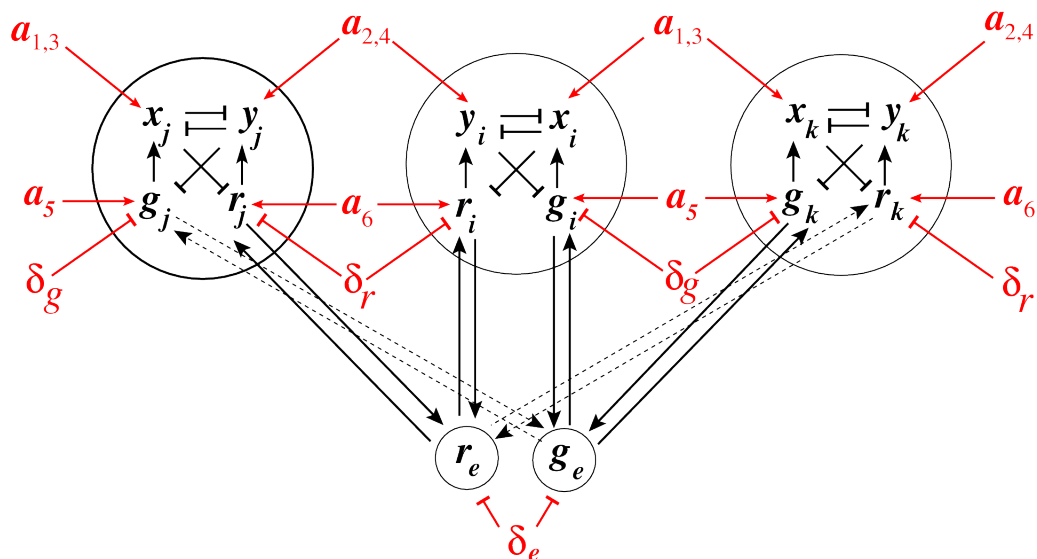
20

Figure 4. An example of three identical S toggles interacting via common autoinducers.
Solid arrows and lines highlighted in red correspond to monotone parameter dependencies.

Suppose that $\bar{x}_0$ is a steady state corresponding to a real parameter value p_0, that is to say, $f(\bar{x}_0, p_0) = 0$. Suppose that we now consider a real parameter p_1 that is very close to p_0 and larger than p_0, $p_1 > p_0$. Suppose in addition that $\bar{x}_1$ is a steady state for the parameter value p_1, $f(\bar{x}_1, p_1) = 0$, and that $\bar{x}_1$ is stable. Now pick the solution $x_1(t)$ of $\dot{x} = f(x, p_1)$ that has initial condition $x_1(0) = \bar{x}_0$. Suppose that the extended system $\dot{x} = f(x, p)$ and $\dot{p} = 0$ is monotone. Now, we may consider the following two initial states for the extended system: $(\bar{x}_0, p_0)$ and $(\bar{x}_0, p_1)$. Since the second state is larger (in the sense of Kamke's Theorem as earlier explained) in the monotone order, it follows that the solutions satisfy $x_1(t) \geq \bar{x}_0$ for all $t > 0$, and therefore, taking limits, we conclude that $\bar{x}_1 > \bar{x}_0$, as desired.

Using Fig. 4 in conjunction with the dimension analysis in terms of the relationships (4) and (5), certain qualitative predictions can be made about the parametric dependencies based on monotone systems theory. To benchmark the approach, we have selected, as an example, a subset of dependencies shown in Fig. 4, presented in Fig. 5.

To preserve the S toggle symmetry, we change the values of the corresponding parameters for all toggles simultaneously. Using the S-model (1) with the values of fixed parameters given in (3), we find that Fig. 4 predicts monotonically increasing dependencies of the x-state variable (describing levels of LacI) on the values of the dimensionless parameters a_1 and a_3, see Fig. 5 (A) and (B), respectively.

The loss of stability and disappearance of the mixed states shown in Fig. 5 (C) as a_5 increases can be interpreted intuitively by the fact that an increase in a_5 leads to an increase in the intracellular levels of the corresponding QS signaling molecules, which, in turn, lead to an increase of extracellular levels of the QS molecules via diffusion, thereby facilitating self-synchronization of the given population of all toggles under conditions corresponding to a stronger interaction among all toggles. In particular, the strong interaction and coupling condition eliminates spontaneous synchronization errors in terms of suppressing the emergence of undesired (9:1)-mixed states.

This result is similar to a well-known fact for oscillators coupled via a common medium that a transition from an unsynchronized to a synchronized regime emerges as the strength of
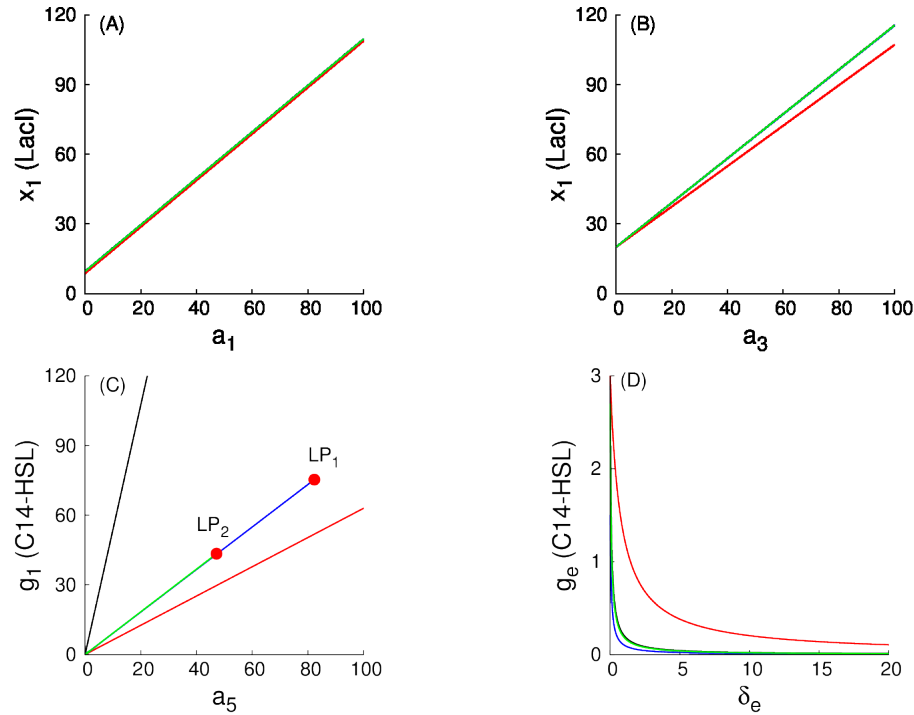
Figure 5. Examples of monotone parametric dependencies. Three different stable populations are chosen for the analysis of monotone parametric dependencies: (1) an G-homogeneous population; (2) an (1:1)-mixed population; here, levels of LacI and C14-HSL from one subpopulation (within which LacI is over-expressed) are shown; and (3) a (9:1)-mixed population (a spontaneous synchronization error); here, again, levels of LacI and C14-HSL from the largest subpopulation (within which LacI is over-expressed) are shown. The reference set of fixed parameter values is given by (3). The values of free parameters (a_1, a_3, a_5, and δ_e) are unique for the entire population of identical cells and are changed simultaneously for all toggles (subpopulations). A color-coding schema for all curves is described at the bottom of the caption. **Panel (A)** shows the monotone dependence of the values in the state variable x_1 (LacI) on the parameter a_1. **Panel (B)** shows the dependence of x_1 (LacI) on a_3. **Panel (C)** shows the dependence of g_1 (intracellular C14-HSL) on a_5; and panel (D) shows the dependence of g_e (extracellular C14-HSL) on δ_e. All detailed explanations, including biological interpretations, are given in the main text. Because the stable mixed populations do not exist for large values of the parameter d (see A (9:1)-Mixed Population Consisting of S-Toggles), we use both $d = 0.1$ (weak coupling) for all populations and, additionally, we use $d = 10$ (strong coupling) for the G-homogeneous population only. To distinguish between all the cases considered, we follow a color coding schema described below. However, due to mathematical properties of the models, projections of the corresponding plots on the 2D-planes often overlap, mixing different colors, which should not lead to any difficulty in recognizing similar monotone ("overlapping") dependencies. **The color-coding schema for all curves includes:** *(i)* black plots are used for G-homogeneous solutions at $d = 0.1$; *(ii)* red plots are used for G-homogeneous solutions at $d = 10$; *(iii)* blue plots are used for (1:1)-mixed states at $d = 0.1$; and *(iv)* green plots are used for (9:1)-mixed states at $d = 0.1$. Red filled circles in panel (C), labeled with LP_1 and LP_2, correspond to Limit Point (LP) (or, equivalently, Saddle-Node) bifurcation points [62]. Here, the blue curve connecting the origin $(0, 0)$ and the LP_1-point corresponds to the stable branch of the (1:1)-mixed state. The green curve connecting the origin $(0, 0)$ with the LP_2-point corresponds to the stable branch of the (9:1)-mixed state. Because the green curve was plotted after plotting the blue curve, a part of the blue curve is hidden beneath the green curve.

22

coupling increases [12–14, 23, 50]. Indeed, many microbial species accomplish this via quorum sensing, which entails the secretion and detection of diffusible molecules (autoinducers), whose concentration serves as a proxy for population density [7].

Using the expression for the dimensionless parameter a_5 given in (4), see Scaling, we can conclude that the increase in the values of the parameter a_5 leading to the bifurcation point LP_2 (Fig. 5) can be achieved by the following experimental interventions:

- stabilization of cell division with lower values of the specific growth rate μ (or, equivalently, higher division periods T);

- stabilization of relevant proteins, using lower values of r_d (or, equivalently, higher half-lives);

- an increase in the maximum production rate (k_G) of C14-HSL by enzyme CinI, see S2 Estimation of Parameter Values;

- an increase in the sensitivity (K_G) of promoter P_{cin} with respect to the number of molecules C14-HSL to half-activate P_{cin}, see Table S2.2 given in S2 Estimation of Parameter Values.

We have used bifurcation analysis with respect to changes in the values of the parameter a_5 as a way to illustrate predictions from monotone systems theory, and in the process we obtained conclusions regarding improvements of S toggle self-synchronization properties by eliminating the (9:1)-mixed state. To this end, we note that there is no need to further increase values of a_5 to move the system to the bifurcation point LP_1 at which the (1:1)-mixed state loses it stability and disappears, because we do not interpret the (1:1)-mixed state as a spontaneous synchronization error, see Spontaneous Synchronization Errors.

We next consider parametric dependencies with respect to changes in the parameters δ_g and δ_r (Fig. 6). As predicted by the diagram shown in Fig. 4, an increase in the values of δ_g should lead to suppressed levels in x_1 (LacI) as well as to elevated levels in y_1 (TetR). This is illustrated in Fig. 6. The almost constant dependencies in Fig. 6 (C) and (D) can be explained by suppressed levels of TetR and C4-HSL in the G-population.

Analogously (Fig. 6), an increase in the values of δ_g should also lead to suppressed levels in g_1 (C14-HSL) as well as to elevated levels in r_1 (C4-HSL), while an increase in the values of δ_r should lead to elevated levels in g_1 (C14-HSL) and, simultaneously, to decreased levels in r_1 (C4-HSL). This is illustrated in Fig. 7. Constant dependencies in Fig. 7 (C) can be explained by suppressed levels of TetR and C4-HSL in the G-population.

We then repeat the analysis of the same parametric dependencies for a (1:1)-mixed state, illustrated in Fig. 8 and Fig. 9. Like in the previous case, we observe that all dependencies are in line with the predictions suggested by Fig. 4.

Additionally, the LP-bifurcation point (Fig. 8) can be interpreted as follows. Decreasing values of both parameters δ_g and δ_r leads to an increase in the intracellular and extracellular levels of the corresponding QS signaling molecules, which, in turn, leads to stronger interactions among all toggles. Indeed, it follows from (5) (see Scaling) that the described changes in the values of dimensionless parameters δ_g and δ_r can be achieved by increasing half-lives of the corresponding QS signaling molecules.

To this end and similarly to the interpretation provided earlier, as the values of the parameters δ_g and δ_r decrease, the (1:1)-mixed state loses its stability and disappear via an LP-bifurcation
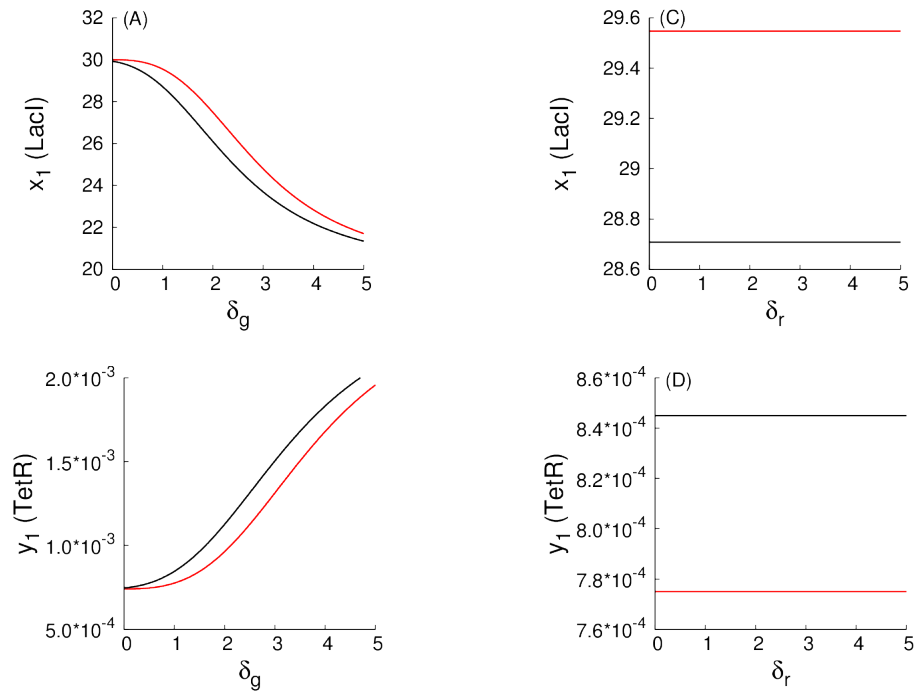
Figure 6. Examples of monotone parametric dependencies for the repressor-protein levels in the G-homogeneous state. Panels (A) and (B) correspond to δ_g, while panels (C) and (D) correspond to δ_r. Red solid curves correspond to a weak coupling among all toggles ($d = 0.1$), while black solid curves correspond to a strong coupling among all toggles ($d = 10$). The values of all other fixed parameters are given in (3), see Reference Values of Model Parameters.
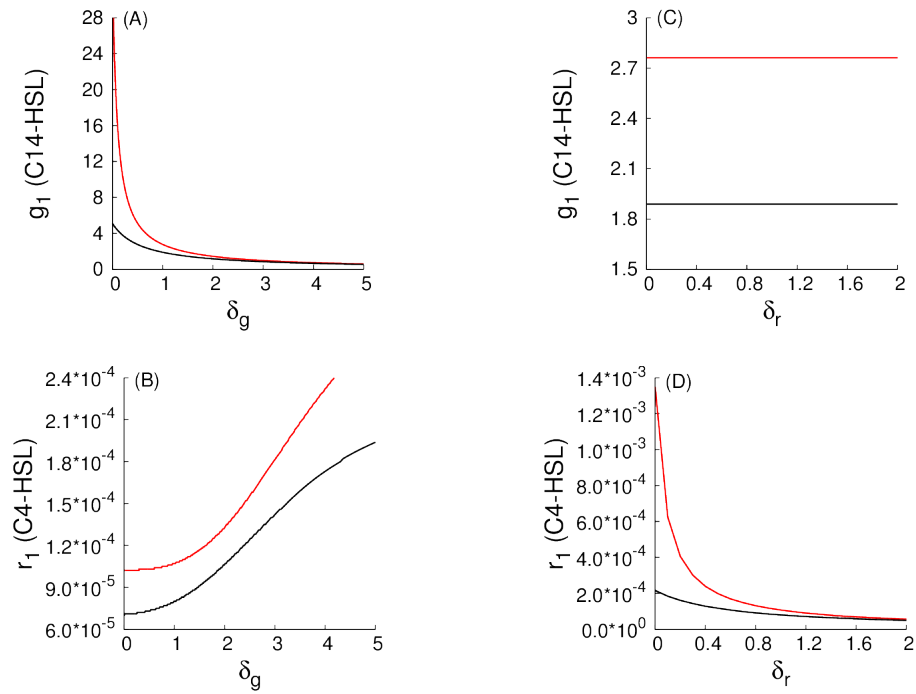
24

Figure 7. Examples of monotone parametric dependencies for the signaling species levels in the G-homogeneous state. Panels (A) and (B) correspond to δ_g, while panels (C) and (D) correspond to δ_r. Red solid curves correspond to a weak coupling among all toggles ($d = 0.1$), while black solid curves correspond to a strong coupling among all toggles ($d = 10$). The values of all other fixed parameters are given in (3), see Reference Values of Model Parameters.
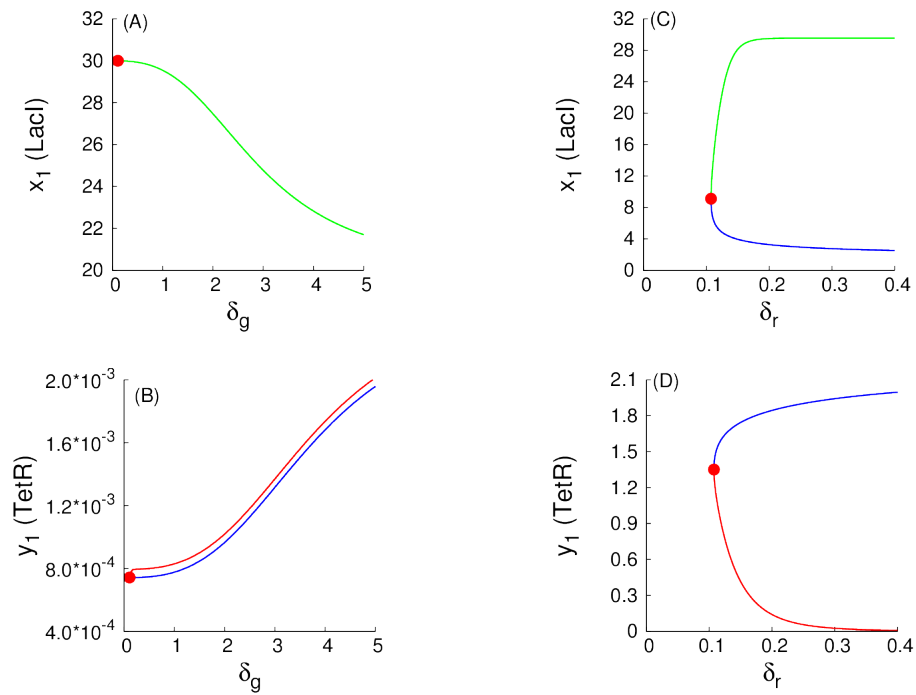
Figure 8. Examples of monotone parametric dependencies for the repressor-protein levels in the (1:1)-mixed state. As an example, the dependencies for the G-subpopulation are shown only, within which LacI is activated, while TetR is repressed. Panels (A) and (B) correspond to δ_g, while panels (C) and (D) correspond to δ_r. Green and red solid curves correspond to stable branches of (1:1)-equilibrium solutions, while all blue curves correspond to unstable branches of the solutions. Red filled circles correspond to an LP-bifurcation point. In panel (A), projections of stable and unstable branches coincide and, so, only the stable branch is shown.
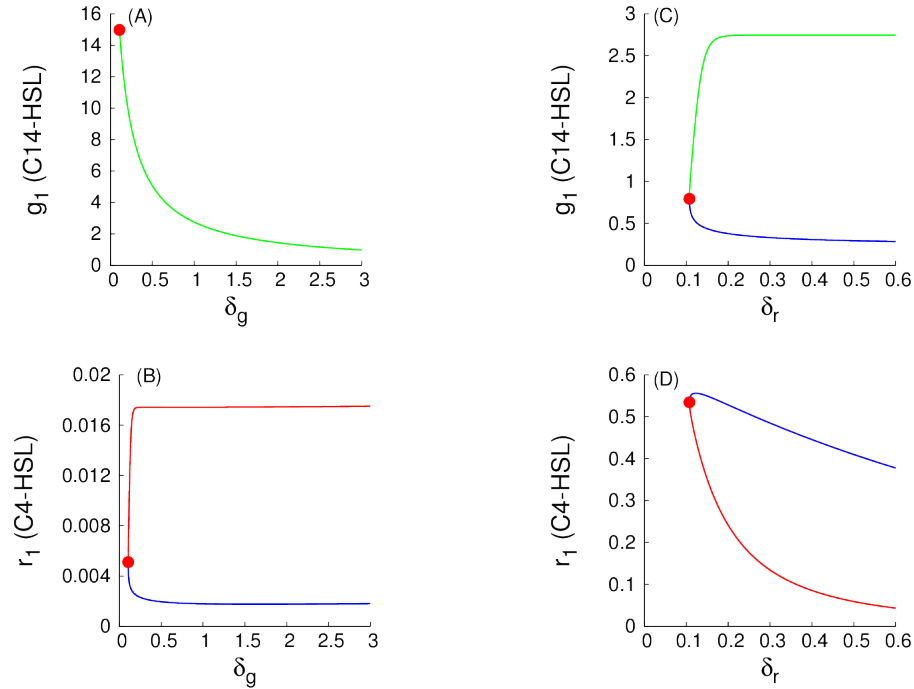
Figure 9. Examples of monotone parametric dependencies for the signaling species levels in the (1:1)-mixed state. Panels (A) and (B) correspond to δ_g, while panels (C) and (D) correspond to δ_r. All explanations are as in Fig. 8.

(Fig. 8), the effect which is similar to the well-known fact that oscillators coupled via common medium synchronize as the strength of coupling increases [12, 23, 50].

We note that the parametric dependencies for unstable solutions are not described by Fig. 4. To explain this observation, we recall that our proof of monotone dependence on parameters applies to stable solutions only, see above.

The monotone parametric dependencies for a (9:1)-mixed state corresponding to a spontaneous synchronization error are illustrated in Fig. 10 for a large G-subpopulation comprising 90% of all cells, and in Fig. 11 for a small R-subpopulation comprising 10% of all cells in the given (9:1)-mixed state.

We observe that LP-bifurcation points are present in both panels (A) and (B), and are absent from both panels (C) and (D) in Fig. 10 and Fig. 11. To explain this observation we have to recall the difference between parameters δ_g and δ_r. As discussed earlier, a decrease in the values of δ_g can be interpreted in terms of the improved communication between the toggles within the large subpopulation, while a decrease in the values of δ_r can be interpreted in terms of the improved communication between the toggles within the small subpopulation. We can thus conclude that by increasing the strength of interactions between the toggles from the large subpopulation, the spontaneous error can be eliminated, corresponding to the existence of the LP-points in panels (A) and (B) of Fig. 10 and Fig. 11. At the same time, increasing the strength of interactions between the toggles from the small population, the corresponding spontaneous error cannot be eliminated.
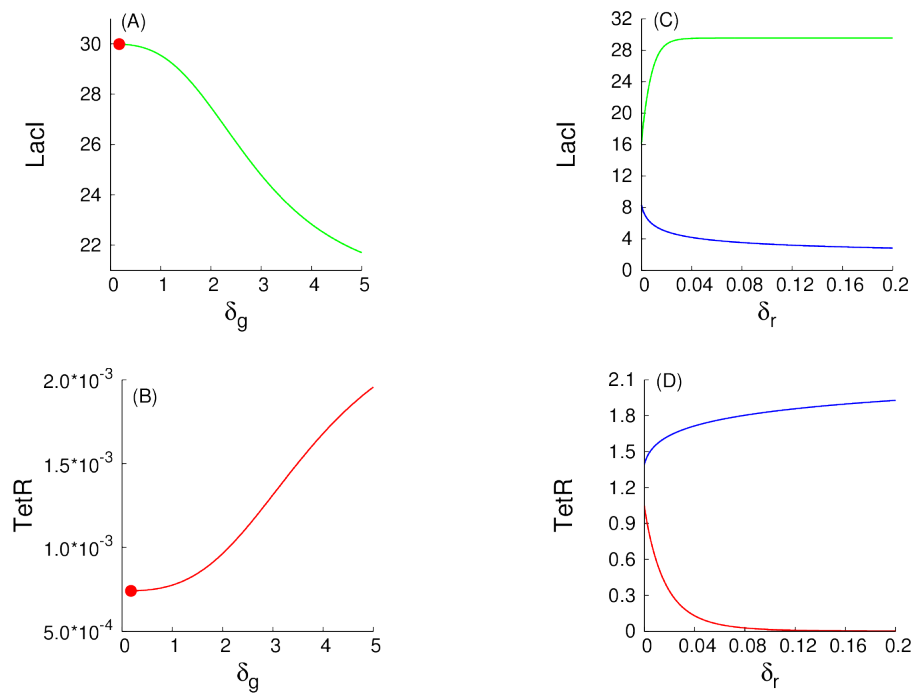
27

Figure 10. Examples of monotone parametric dependencies for the repressor-protein levels in the (9:1)-mixed state (a 90% large G-subpopulation.) The curves for the large G-subpopulation are shown in panels (A) - (D), within which TetR is activated, while LacI is repressed. Panels (A) and (B) correspond to δ_g. Panels (C) and (D) correspond to δ_r. Green and red solid curves correspond to stable solution branches, while all blue curves correspond to unstable solution branches. Red filled circles correspond to an LP-bifurcation point. In panels (A) and (B), projections of stable and unstable solution branches coincide and, so, only the stable solution branches are shown.
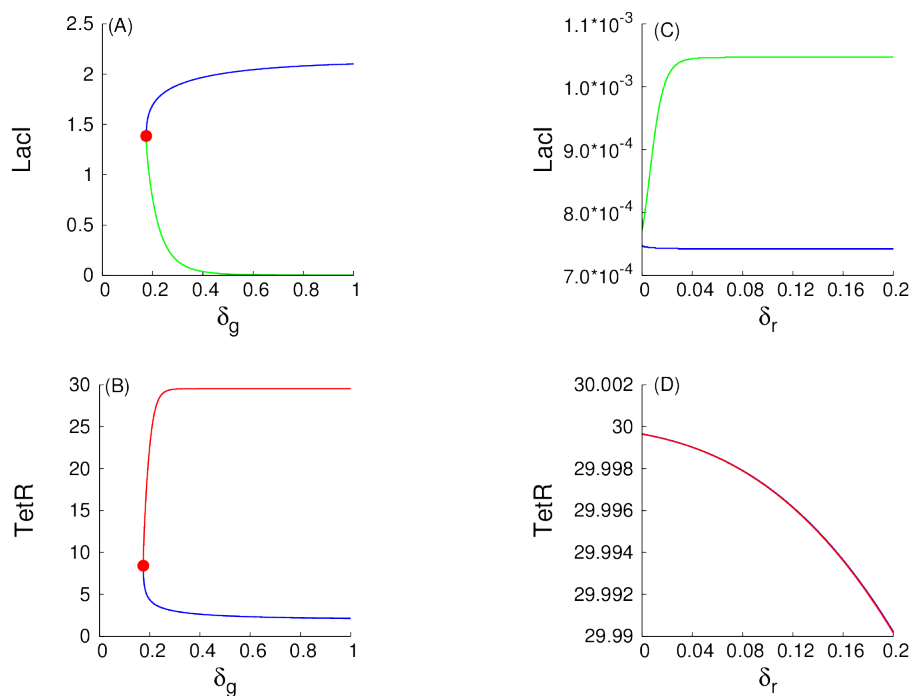
Figure 11. Examples of monotone parametric dependencies for the repressor-protein levels in the (9:1)-mixed state (a 10% small R-subpopulation.) The curves for the small R-subpopulation are shown in panels (A) - (D), within which TetR is activated, while LacI is repressed. Panels (A) and (B) correspond to δ_g. Panels (C) and (D) correspond to δ_r. Green and red solid curves correspond to stable solution branches, while all blue curves correspond to unstable solution branches. In panel (D), projections of stable and unstable solution branches coincide. Red filled circles in panels (A) and (B) correspond to an LP-bifurcation point.

Bistability in Single S-Design and A-Design Toggles Isolated from the Environment

Before comparing population properties of our S design to those of the A design, we remark that, even for isolated cells (when the diffusion constant d is zero), there is a larger range of bistability for the S design compared to the A design. Specifically, a bistability region for a single A toggle in the plane (a_1, a_2) at $d = 0$ is shown in Fig. 12. Similar regions were found in [1, 13].

We observe that the entire quadrant, $a_1 \geq 0$ and $a_2 \geq 0$, spans a bistability region for the S-model at the fixed parameter values given in (3).
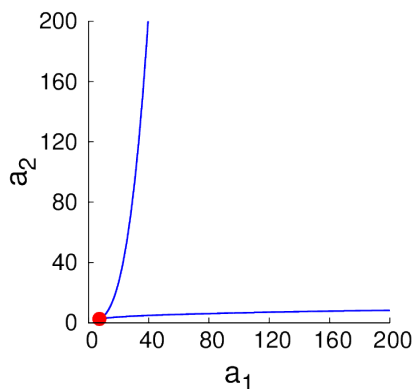

Figure 12. Bistability regions for single S toggle and A toggle models. The region between two blue color coded LP-bifurcation loci corresponds to a bistability region for the A toggle model (2) at $d = 0$. A red filled circle corresponds to a cusp point (CP). For the S toggle model, bistability exists for all parameter values $a_1 \geq 0$ and $a_2 \geq 0$ at $d = 0$. Other fixed parameter values are given in (3).

An important observation that follows immediately from Fig. 12 is that in the case of the S toggle, bistability exists at the origin of the non-negative quadrant in the plane (a_1, a_2), that is, at $a_1 = a_2 = 0$. This simply means that the genes *lacI* and *tetR* can be removed from the corresponding plasmids bearing promoters P_Y and P_X, respectively (Fig. 2). In this case, it is enough to keep the genes on the plasmids bearing the corresponding promoters P_G and P_R (Fig. 2). We view the reduced S toggle as a minimal design that could be implemented experimentally. The fuller construct S is interesting too, in so far as it is based on the well-characterized and studied Cantor-Collins switch, coupled to quorum-sensing components [4]. We find that the full and reduced designs do not differ much in performance, and, so, we do not consider the minimal design in the rest of the paper.

Bistable Homogeneous Populations Consisting of S toggles

Fig. 13 shows scaled levels of LacI and C14-HSL for a homogeneous population in the G-state, depending on the values of the parameter d. We observe that the intracellular and extracellular levels of the QS signaling molecule C14-HSL become asymptotically indistinguishable from one another as $d \rightarrow \infty$.

The asymptotic behavior for large values of d can be analytically understood after introducing a small parameter $\varepsilon = d^{-1}$ into the S-model (1) which becomes a singularly-perturbed problem [65].
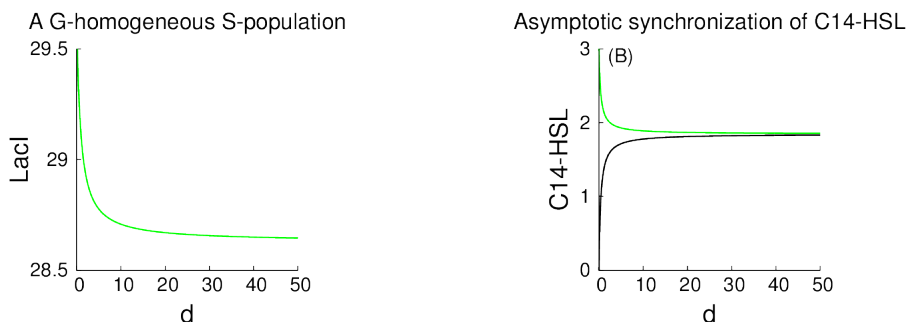
Figure 13. A G-homogeneous population of S toggles. Dependence of a G-homogeneous population on the values of the parameter d is shown. Panel (A) presents dimensionless levels of the activated LacI. Dimensionless levels of the repressed TetR are of order of magnitude about 10^{-3} and are not shown. Panel (B) presents dimensionless levels of C14-HSL. Here, the green curve corresponds to the intracellular levels, while the black plot corresponds to extracellular levels of C14-HSL in the medium, respectively.

Setting formally $\varepsilon = 0$ in the singularly-perturbed problem as required by the theory of singular perturbations [65], the differential equations (1c) and (1d) can be reduced to elementary algebraic equations $g = g_e$ and $r = r_e$, respectively.

Bistable Homogeneous Populations Consisting of A toggles

Fig. 14 shows two stable homogeneous populations of A toggles which coexist while the parameter d is allowed to vary.

Because the A toggle design does not have typically any intrinsic symmetry, the levels of the activated repressor proteins, LacI for the G-homogeneous population shown in the Fig. 14(A,B) and TetR for the R-homogeneous population shown in Fig. 14(C,D), differ significantly from one another. Recall that the levels of LacI and TetR in the corresponding G- and R-homogeneous populations consisting of S toggles (Fig. 13) are identically the same due to the mirror ($\mathbb{Z}_2$) symmetry.

The combination of the analyses in Figs. 13 and 14 can be summarized by saying that under each one of the two designs, S and A, bistable homogeneous stable populations are possible, in either "Red" or "Green" consensus states, and with the same order of magnitude of expression. The difference between these designs is evident when we study heterogeneous (mixed) populations, as discussed next.

A (1:1)-Mixed Population Consisting of S toggles

Fig. 15 shows a richness of dynamic effects (bifurcations) for a (1:1)-mixed population of S toggles. We see that as soon as the parameter d takes on larger values, the (1:1)-mixed state loses its stability via a Branch Point (BP) bifurcation (alternatively called "pitchfork" or "symmetry-breaking" bifurcation), giving rise to two stable (1:1)-mixed non-symmetric states at $d \approx 1.43$. The bifurcation has co-dimension one for all typical systems with $\mathbb{Z}_2$-symmetry [52, 53].

To understand the symmetry-breaking phenomenon occurring at the BP-point shown in Fig. 15, we need first to define the symmetry of the original symmetric (1:1)-mixed state. We observe

31
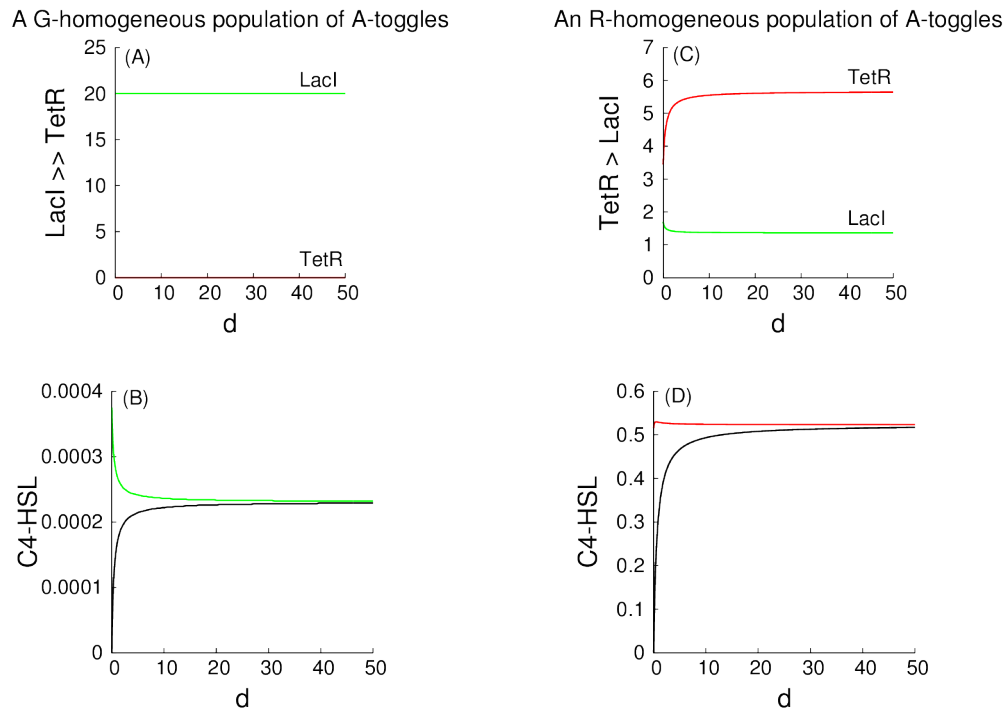
Figure 14. G- and R-homogeneous populations of A toggles. Panels (A) and (B) present dimensionless levels of the activated LacI and C4-HSL obtained for the A toggle settled at the G-state (LacI > TetR), while the other two panels (C) and (D) present dimensionless levels of the activated TetR and C4-HSL obtained for the A toggle settled at the R-state (TetR > LacI). Recall that C4-HSL is a unique QS signaling molecule facilitating communication between the A toggles, and this is why C4-HSL is present in both (B) and (D) panels.

that the original symmetric (1:1)-mixed state is invariant with respect to transformation g,

$$g = \mathcal{P} \circ \mathcal{I} = \mathcal{I} \circ \mathcal{P}, \quad g \circ g = \text{id} \tag{15}$$

where id is the identity. Here, $\mathcal{P}$ and $\mathcal{I}$ are as defined in (9) and (10), respectively.
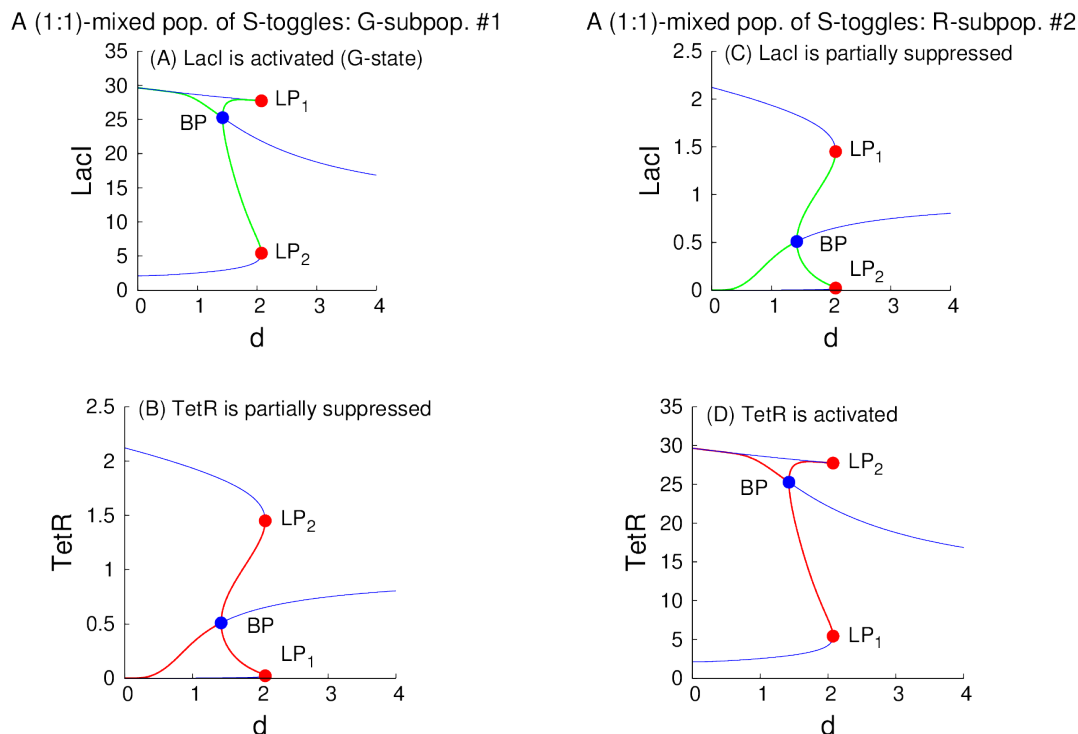

Figure 15. Symmetry breaking in a (1:1)-mixed population of S toggles. The given (1:1)-mixed population consists of two equal subpopulations, (1) a G-subpopulation and (2) an R-subpopulation, see the center panel in Fig. 3. Panels (A) and (B) show dimensionless levels of *activated* LacI (A) and *repressed* TetR (B), respectively, within the G-subpopulation. Panels (C) and (D) show dimensionless levels of *repressed* LacI (A) and *activated* TetR (B), respectively, within the R-subpopulation. Blue color-coded plots correspond to all *unstable* equilibrium solution branches, while green and red color-coded plots correspond to all *stable* equilibrium solution branches. All blue filled BP-labeled points correspond to $d \approx 1.43$. All red filled LP-labeled points correspond to $d \approx 2.07$. Both labels LP$_1$ and LP$_2$ correspond to the same bifurcation and are used for convenience of understanding the diagrams. Due to the *isotropy* subgroup $\Sigma_{(1:1)}$ of the original (1:1)-mixed state, the isotropy involution g defined in (15) of the main text maps panel (A) to panel (D), and panel (B) to panel (C). We thus have, correspondingly, $g : \text{LP}_1 \to \text{LP}_2$, $g : \text{LP}_2 \to \text{LP}_1$, and $g : \text{BP} \to \text{BP}$, see the coordinates of the three critical points, BP, LP$_1$, and LP$_2$, at the end of the caption. *Abbreviations:* pop., population; subpop., subpopulation. The coordinates of the critical points are given in the format $(x_1, y_1, x_2, y_2; d)$, where index i corresponds to the subpopulation i, $i = 1, 2$,
BP $= (25.265, 0.510, 0.510, 25.265; 1.424)$,
LP$_1 = (27.730, 0.023, 1.450, 5.427; \ 2.073)$,
LP$_2 = (5.427, \ 1.450, 0.023, 27.730; 2.073)$.

The transformation (15) forms subgroup $\Sigma_{(1:1)}$ of the group $\mathbb{Z}_2 \times \mathbf{S}_2$ (see Symmetry), which consists of two elements, that is, $\Sigma_{(1:1)} = \{\text{id}, g\}$. The subgroup $\Sigma_{(1:1)}$ is called the *isotropy* subgroup [52] of the original (1:1)-mixed state. Because the two (1:1)-mixed states bifurcating from the original $\Sigma_{(1:1)}$-symmetric (1:1)-mixed state at the BP-point (Fig. 15) are not invariant with respect to the isotropy subgroup $\Sigma_{(1:1)}$, and, instead, are mapped to one another by the

33

transformation (15), this motivates using the "symmetry-breaking" terminology [52] with respect to the loss of the isotropy symmetry by the (1:1)-mixed state at the BP-point.

As the value of the parameter d is further allowed to increase above the critical value $d \approx 1.43$ (Fig. 15), one branch of the pair of the stable non-symmetric (1:1)-states evolves along the locus connecting the points BP and LP_1, while another branch of the same pair evolves along the locus connecting the points BP and LP_2 (Fig. 15).

As the value of the parameter d is further increased, the (1:1)-mixed non-symmetric equilibria lose their stability and disappear via an LP-bifurcation at $d \approx 2.07$ (Fig. 15) simultaneously as relative states (see Symmetry). Indeed, we have $g : LP_1 \longrightarrow LP_2$, where g is defined in (15).

We can describe the above symmetry-breaking scenario intuitively as follows. Suppose that we start with a mixed population in which 50% of the cells are in "green" state and 50% of the cells are in "red" state, and the nondimensional diffusion coefficient d (which, as we saw, in fact incorporates many of the kinetic parameters in the original system) has a low value. Suppose that we now slowly increase the value of d, and ask what happens to the (1:1)-mixed state. The first event that is observed, at $d \approx 1.43$, is that this "pure 50-50 mixed state" loses its stability. A new mixed state arises, in which there are *two* subpopulations, one in which green gene-expression dominates (but with different expression levels of LacI in each of them), and another one which red gene-expression dominates (also with different TetR levels).

Furthermore, as d is increased a bit more (past $d \approx 2.07$), even these mixed states disappear. Thus, even with moderate diffusion, heterogeneous populations cannot be sustained, emphasizing the consensus-forming character of the S design. This is in marked contrast to the A design, as shown next.

The loss of stability by the (1:1)-mixed state increases the robustness of the S toggle design towards its self-synchronization by reducing the number of alternative stable states to which the toggle state can settle.

A (1:1)-Mixed Population Consisting of A toggles

In contrast to (1:1)-mixed populations of S toggles, we observe from Fig. 16, that the original (1:1)-mixed A-population cannot be eliminated (made unstable) by increasing the values of the parameter d within a very large parameter interval. In other words, increasing the strength of interactions between the cells does not help to establish synchronization across the given population of identical A toggles.

A (9:1)-Mixed Population Consisting of S-Toggles

Next, we consider bistable (9:1)-mixed populations, which as discussed in the introduction, we think of as arising from random synchronization errors. We observe that populations quickly become unstable as a function of the nondimensional diffusion parameter d (Fig. 17).

(9:1)- and (1:9)-Mixed Populations Consisting of A toggles

In contrast to the S design, in the A design, the mixed (9:1)- and (1:9)-heterogeneous populations that might arise from random state switching cannot be eliminated by changes in the values of the parameter d (Fig 18).
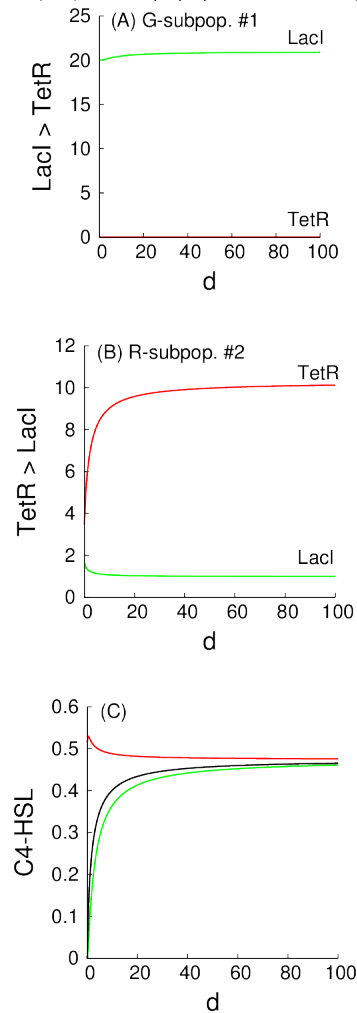
Figure 16. A (1:1)-mixed population of A toggles. Green and red color coded curves in panels (A) and (B) correspond to the intracellular concentrations of LacI and TetR, respectively, while a black color-coded curve in panel (C) corresponds to the extracellular concentration of C4-HSL. In panel (C), the green color-coded curve corresponds to the concentration of C4-HSL within the G-subpopulation, that is, LacI > TetR as in panel (A), while the red color-coded curve corresponds to the R-subpopulation, that is, TetR > LacI as in panel (B).

Figure 17. A (9:1)-mixed population of S toggles. Here, all notations and color-coding schemes are as in Fig. 15. Red filled circles correspond to the same LP-bifurcation point. In panel (A), projections of stable and unstable solution branches overlap. Because TetR is totally suppressed in the large (90%) subpopulation (G-subpopulation in the left panel), the levels of TetR are not shown. Instead, we zoom-in to a small range of the levels for LacI to obtain a better graphical visua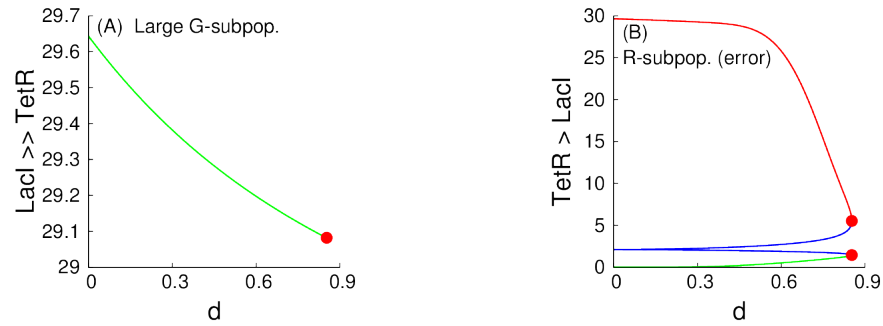lization of the evolution of the LacI levels as d increases. Contrarily to panel (A), both TetR and LacI levels are plotted in panel (B) since LacI is only moderately suppressed in the small (10%) R-subpopulation.
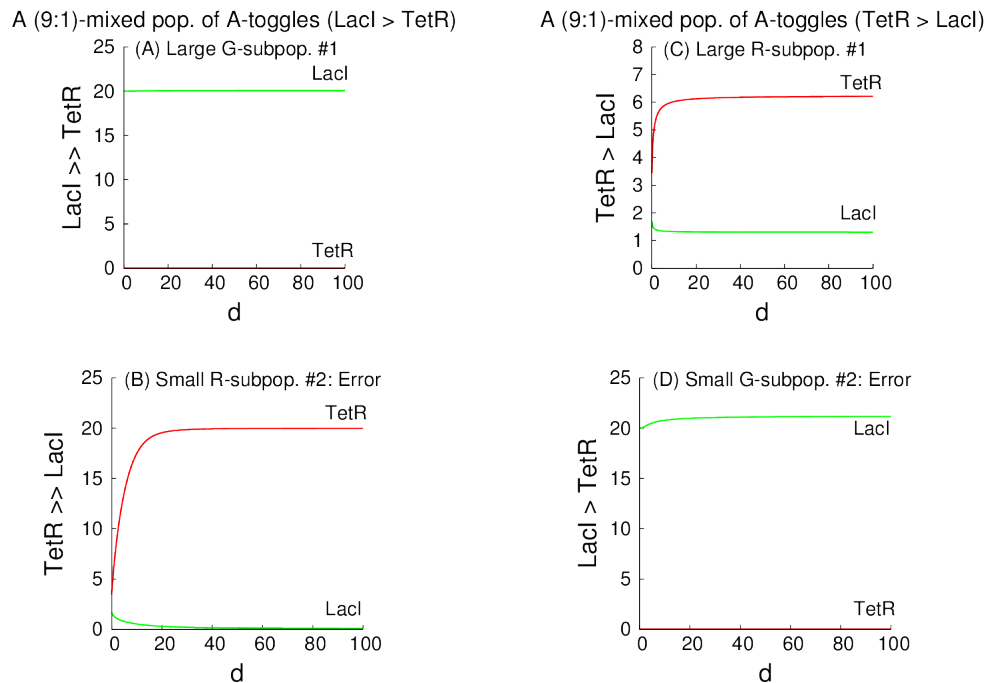
Figure 18. A (9:1)- and (1:9)-mixed population of A toggles. Here, all notations and color-coding schemes are as in Fig. 17. Panels (A) and (B) correspond to the (9:1)-mixed population, within which the transcription signature LacI $\gg$ TetR dominates in proportion 9:1 (*i.e.*, with 90% of green cells and 10% of red cells), while panels (C) and (D) correspond to the (1:9)-mixed population, within which the opposite transcription signature TetR $\gg$ LacI as well dominates in proportion 9:1 (*i.e.*, with 90% of red cells and 10% of green cells.)

Probing Capabilities of S- and A-Toggle Designs for Self-Correction of Spontaneous Synchronization Errors

To probe and compare capabilities of the S toggle and A toggle designs to correct "spontaneous synchronization errors" caused by a random flip of one toggle (or a small fraction of toggles) from a homogeneous population to the state opposite to the transcription signature adopted by the majority of the cells, we have performed simple random tests which will be discussed below. In mathematical and computational terms, these random tests can be interpreted as an elementary numerical procedure to evaluate the size of the basin of attraction for the corresponding equilibrium solutions by sampling the corresponding small neighborhoods of the solutions, using random initial conditions.

Our numerical procedure is organized as follows. For each of the parameter values in the following set: $d \in \{0.01, 10, 100\}$, two stable G- and R-homogeneous states have been found numerically, where the stability is computed for any value of $N > 1$, see S3 Exponential Stability of Cellular Populations. Due to the mirror symmetry of the S design, the analysis of the G-homogeneous state is sufficient as all trajectories are mapped to the corresponding trajectories for the alternative R-homogeneous state under the action of the involution defined in (7). In contrast, the property of error self-correction should be analyzed for two different A-homogeneous stable equilibria (Fig. 14) separately.

The random initial conditions at $t = 0$ have been selected as follows. First, we take $Z = \left(x^0, y^0, g^0, r^0, \ldots, x^0, y^0, g^0, r^0, g_e^0, r_e^0\right)$ to be a (10:0)-homogeneous solution (G-state with $x^0 > y^0$). Then, the perturbed initial conditions for 9 cells were set to the G-state. The initial condition for the last, 10th cell was set at the R-state taken from the same the (10:0)-homogeneous solution Z, using the involution $x_{10}(0) \leftarrow y^0$, $y_{10}(0) \leftarrow x^0$, $g_{10}(0) \leftarrow r^0$, and $r_{10}(0) \leftarrow g^0$, that is, we obtained $x_{10}(0) < y_{10}(0)$ (R-state). We also used $g_e(0) = g_e^0$ and $r_e(0) = r_e^0$. Each initial condition was then perturbed by adding a random term with a magnitude of 10% of the initial condition value for the corresponding state variable.

For the A design, two different sets of initial conditions were used as follows. First, let $Z^\omega = (x^\omega, y^\omega, r^\omega, \ldots, x^\omega, y^\omega, r^\omega, r_e^\omega)$, where $\omega \in \{G, R\}$. Then, to form the fist set of initial conditions, the initial conditions for 9 cells were set at the G-state taken from the (10:0)-homogeneous solution, Z^G, that is, we used $x_i(0) = x^G$, $y_i(0) = y^G$, $r_i(0) = r^G$ ($i = 1, \ldots, 9$), and $r_e(0) = r_e^G$, while the initial condition for the last, 10th cell was set at the R-state taken from the alternative, stable (0:10)-homogeneous solution, Z^R, that is, we had $x_{10}(0) = x^R$, $y_{10}(0) = y^R$, and $r_{10}(0) = r^R$. The second set of initial conditions was formed, using similar steps, resulting in $\left(x^R, y^R, r^R, \ldots, x^R, y^R, r^R, x^G, y^G, r^G, r_e^R\right)$. Finally, each initial condition was perturbed by adding a random term with a magnitude of 10% of the initial condition value for the corresponding state variable.

We can conclude from Fig. 19 and Fig. 20 that the A toggle does not have any capability for self-correction of spontaneous errors for all tested values of the parameter d (Fig. 20). The S toggle can self-correct spontaneous synchronization errors for medium and large values of d (Fig. 19, that is, for all parameters values for which the mixed state becomes unstable, see Fig. 17 (A (9:1)-Mixed Population Consisting of S-Toggles.)
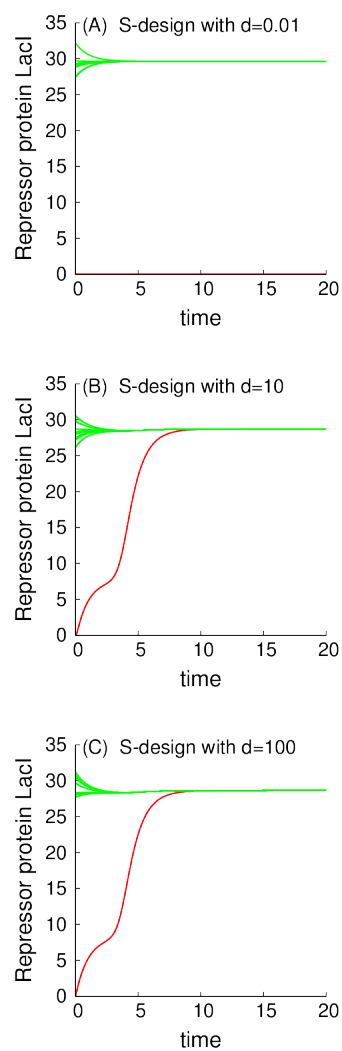
Figure 19. Probing self-correction capabilities of spontaneous errors by S toggles. All explanations are given in the main text.
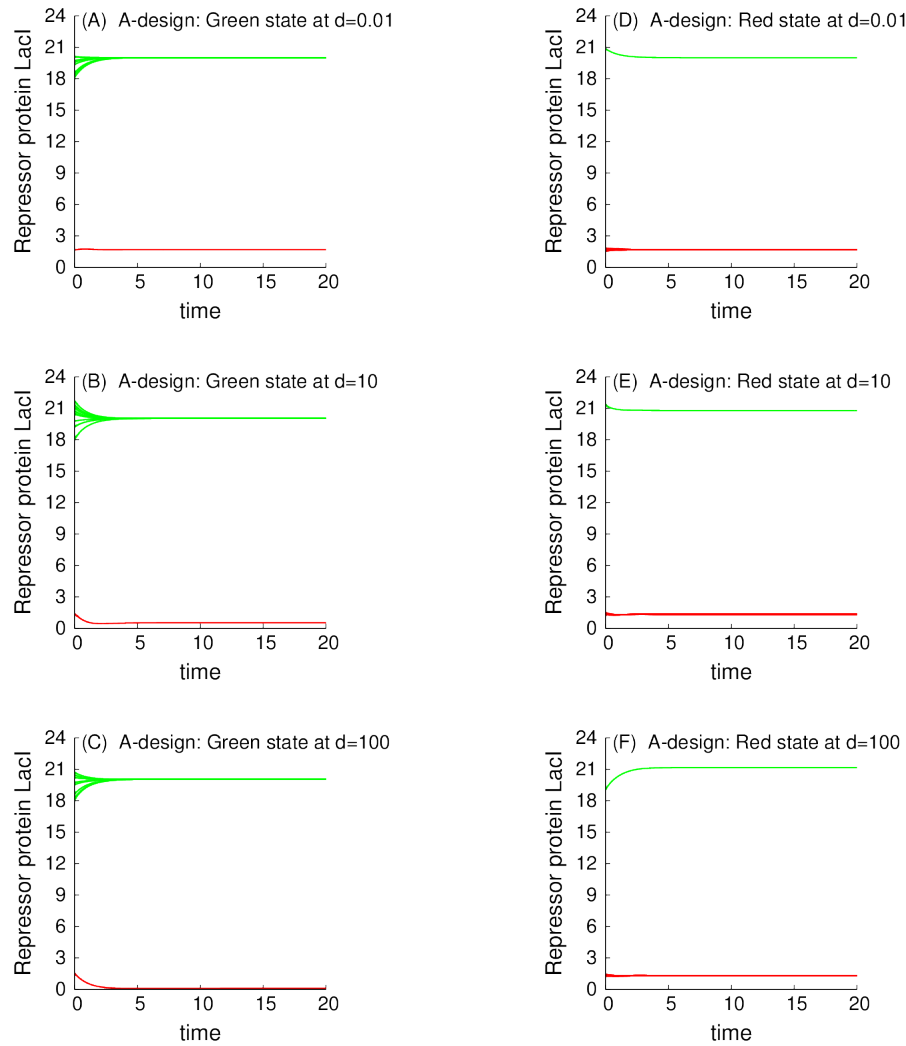
Figure 20. Lack of any self-correction capability for spontaneous errors by A toggles. The left panels correspond to the reference (10:0)-state (G-state), while the right panels correspond to the reference (0:10)-state (R-state). The initial conditions in the left panels correspond to nine "green" cells and one "red" cell. The initial conditions in the right panels correspond to one "green" cell and nine "red" cells.

Conclusions

In this study, we have shown how synthetic bistable circuits (toggles), and hosting them, programmable cellular populations, can be designed so as to solve a robust molecular task, the maintenance of a coordinated state, and a "majority-vote" auto-correction of deviations, of a binary switch. Our design was guided by concepts from monotone systems theory [15–21]. Specifically, we have shown how this concept can be used for the design of a new class of monotone synthetic biological toggles, including predictive capabilities describing both dynamic state variables and monotone parametric tendencies caused by parameter perturbations.

To benchmark the new toggle design, termed the S design, and the monotone systems approach, we have compared the S design with the known (and non-monotone) asymmetric B2-strain from [4], termed the asymmetric or A design in this work. The B2-strain has been previously studied both experimentally [4] and theoretically [13, 14]. Despite a number of remarkable properties of the B2-strain (A design), the A toggle multifunctionality suggests that the design must be tightly controlled to avoid spontaneous switching not only between different expression states, but, as well, between different functions such as a bistable memory and an oscillatory phenotype.

In this respect, modern gene therapy interventions are currently limited to transfected genes to be either in an "on" or "off" state, when the expression of the transfected gene needs to be regulated tightly for the effective treatment of many diseases. To address this challenge, the monotone S toggle design completely excludes any unpredictable chaotic behaviors, as well as undesired stable oscillations. This conclusion is valid for all parameter values, and provides a strong theoretical guarantee missing from other synthetic biology designs.

To achieve an in-depth understanding of dynamic properties of the S toggle design, we have developed biochemically-detailed and biologically-relevant mathematical models to test predictions of monotone systems theory by employing computational bifurcation analysis. To have all results biologically grounded, concrete molecular entities have been used, though the results are general and independent of any specific details.

Additionally, to investigate the effect of a spontaneous toggle switching within cellular populations, leading to bimodal distributions, we have formalized a concept of spontaneous synchronization errors and tested the toggle design capabilities to self-correct spontaneous synchronization errors by sampling the basin of attraction of the corresponding equilibrium solutions. We found that the S toggle design was able to self-correct (or, auto-correct) synchronization errors, while the non-monotone A toggle design was not.

Finally, because the number of cells in populations is *a priori* unknown, all the above results and conclusions can make sense only if they are made independently of the population size. To justify the above assertion, we have proved a few general theorems on the exponential stability of the equilibrium solutions corresponding to both homogeneous and mixed populations. The simple exponential stability results are independent of the number of cells in the populations and are based on basic first principles of stability analysis resulting from the Schur's formula [66], which allow the characteristic polynomials for the corresponding model linearizations to be computed explicitly.

Supporting Information

S1 Model Derivation

Here, we describe the main assumptions and steps used to derive mass-balance equations for the S- and A-models formulated in the main text.

S1.1   Mass-Balance Equations

The derivation of the S- and A-models includes the development of two modules:

   I. A transcription-translation module describing biosyntheis of repressor proteins.

   II. A metabolic module describing biosynthesis of autoinducers.

A general and systematic discussion of both modules can be found in [67, 68]. The derivation of the first module for the A-model, relevant to our work, is given in [13]. Because one of our modeling objectives is to ultimately describe how the analysis of the mathematical models can be mechanistically interpreted in terms of tuning synthetic toggle "dials" by implementable experimental interventions as reviewed in [6], including modifications of ribosome-binding sites (RBS), carboxy-terminal tags, $etc.$, [1, 2, 51], we will derive mass-balance equations at the level of molecular detail sufficient to suggest plausible modeling predictions.

A transcription-translation module can be described by a basic two-stage model [13, 67, 68],

$$\frac{dn_x}{dt} = n_A k_x - r_x n_x, \tag{S1.1a}$$

$$\frac{dn_X}{dt} = k_X n_x - r_X n_X. \tag{S1.1b}$$

Here, n_x is the number of mRNA transcripts per cell for gene x, and n_X is the number of protein molecules per cell; n_A is the number of active promoters from which the mRNA of gene x is transcribed at an average rate k_x; k_X is the averaged translation rate; r_x and r_X are the effective first-order rate constants associated with degradation of the mRNA and proteins, respectively.

Since mRNA molecules are usually degraded rapidly compared to other cellular processes, a quasi-steady state for the equation (S1.1a) can often be assumed [13], yielding

$$n_x = \frac{n_A k_x}{r_x}. \tag{S1.2}$$

Using (S1.2) in the right-hand side of the equation (S1.1b), we obtain

$$\frac{dn_X}{dt} = k_X \frac{k_x\, n_A}{r_x} - r_X\, n_X. \tag{S1.3}$$

The ratio $b_x = k_X/r_x$ in equation (S1.3) is called a burst parameter of the protein X [13]. Using b_x in (S1.3) yields

$$\frac{dn_X}{dt} = b_x k_x n_A - r_X n_X. \tag{S1.4}$$

Assuming log-phase growth of *E. coli*, the volume $V(t)$ of the growing bacterium can be approximated by the expression $V(t) = V_0 \exp(\lambda t)$, and equation (S1.4) can be rewritten in a concentration form,

$$\frac{d[\text{X}]}{dt} = b_x k_x [\text{P}_\text{A}] - (r_\text{X} + \mu)[\text{X}]. \tag{S1.5}$$

Here, $[\text{X}](t) = n_\text{X}(t)/V(t)$ and $[\text{P}_\text{A}](t) = n_A(t)/V(t)$.

The concentration of activate promoters, $[\text{P}_\text{A}]$, can be computed, using an appropriate Hill function [68]. For example, we use

$$[\text{P}_\text{YA}] = \frac{[\text{P}_\text{Y}]}{1 + ([\text{Y}]/K_\text{Y})^{n_\text{Y}}} \tag{S1.6}$$

for the repressor protein Y binding to the promoter P_Y with the dissociation constant $K_\text{Y}^{n_\text{Y}}$. In (S1.6), $[\text{P}_\text{Y}]$ is the total concentration of all promoters P_Y, while $[\text{P}_\text{YA}]$ is the concentration of active promoters not bound with the repressor protein Y. Recall that the cooperativity described by the Hill exponent n_Y can arise from [1, 67, 68]:

(i). Multimerization of represssor proteins;

(ii). Cooperative binding of repressor multimers to multiple operator sites in the promoter.

Analogously, we use a Hill-function

$$[\text{P}_\text{GA}] = [\text{P}_\text{G}] \frac{([\text{G}]/K_\text{G})^{n_\text{G}}}{1 + ([\text{G}]/K_\text{G})^{n_\text{G}}} \tag{S1.7}$$

for the autoinducer G binding to the promoter P_G with the dissociation constant $K_\text{G}^{n_\text{G}}$ and the Hill exponent n_G. The concentration of all active promoters P_A can now be obtained from (S1.6) and (S1.7) as

$$\text{P}_\text{A} = \text{P}_\text{YA} + \text{P}_\text{GA}. \tag{S1.8}$$

Using expression (S1.8), the equation (S1.5) can finally be updated as

$$\frac{d[\text{X}]}{dt} = \frac{b_x k_x [\text{P}_\text{Y}]}{1 + ([\text{Y}]/K_\text{Y})^{n_\text{Y}}} + b_x k_x [\text{P}_\text{G}] \frac{([\text{G}]/K_\text{G})^{n_\text{G}}}{1 + ([\text{G}]/K_\text{G})^{n_\text{G}}} - (r_\text{X} + \mu)[\text{X}]. \tag{S1.9}$$

Here, all parameters are described below in Table S2.2.

Similar mass balanced equations can be derived for the repressor protein Y, and synthases U and W. For example, the mass balance equation for the synthase U is

$$\frac{d[\text{U}]}{dt} = \frac{b_u k_u [\text{P}_\text{Y}]}{1 + ([\text{Y}]/K_\text{Y})^{n_\text{Y}}} - (r_\text{U} + \mu)[\text{U}]. \tag{S1.10}$$

Analogously, we can write down a mass balance equation for the autoinducer concentration [G], that is, [C14-HSL], governed by the synthase U (CinI),

$$\frac{d[\text{G}]}{dt} = k_\text{G}[\text{U}] + D_\text{G}(\text{G}_\text{e} - \text{G}) - (r_\text{G} + \mu)[\text{G}]. \tag{S1.11}$$

Here, k_G is the maximum production rate of C14-HSL by CinI (Table S2.4), D_G is the export rate of C14-HSL (Table S2.4), and $[\text{G}_\text{e}]$ is the extracellular concentration of C14-HSL.

42

Assuming that the concentration of the enzyme U reaches its quasi-steady state rapidly [69], one can obtain from (S1.10) that

$$[U] = \frac{b_u k_u}{r_U + \mu} \times \frac{[P_Y]}{1 + ([Y]/K_Y)^{n_Y}}. \tag{S1.12}$$

Using (S1.12) in the equation (S1.11) yields

$$\frac{d[G]}{dt} = \frac{b_u\, k_u\, k_G}{r_U + \mu} \times \frac{[P_Y]}{1 + ([Y]/K_Y)^{n_Y}} - (r_G + \mu)\,[G]. \tag{S1.13}$$

Here, the definitions and the values of all parameters are given in Table S2.4. A similar mass balanced equation can be derived for the second autoinducer R (C4-HSL), and we omit the details.

S1.2   Nondimensionalization (Scaling)

To nondimensionalize mass balance equations, as for example, the mass balance equations (S1.9) and (S1.13), we use the following dimensionless state variables, which are similar to those introduced in [13],

$$t' = (r_d + \mu)\,t, \quad x = \frac{[X]}{K_X}, \quad y = \frac{[Y]}{K_Y}, \quad g = \frac{[G]}{K_G}, \quad g_e = \frac{[G_e]}{K_G}, \quad r = \frac{[R]}{K_R}, \quad r_e = \frac{[R_e]}{K_R}. \tag{S1.14}$$

Here, we assume that all protein degradation rates can be set experimentally so that the following equalities can be obtained *approximately* [2],

$$r_X = r_X = r_U = r_W = r_d = \frac{\ln 2}{\tau_{1/2}}, \quad \tau_{1/2} = 4\,\mathrm{min}. \tag{S1.15}$$

The procedure of setting all protein degradation rates or, equivalently, all protein half-lives approximately equal to a prescribed value close to mRNA half-lives [2] is required to balance the toggle [1]. We discuss the balancing procedure and relevant experimental interventions in Sect. S1.3.

Using the dimensionless variables (S1.14), all original modeling mass balances can be nondimensionalized, yielding the S- and A-models formulated in the main text, where the prime is dropped from dimensionless time t'. In this case, dimensional and dimensionless parameters are related to one another as:

1. For the dimensionless rates, we obtain:

$$a_1 = \frac{b_x\, k_x\, [P_Y]}{K_X\,(r_d + \mu)}, \quad a_2 = \frac{b_y\, k_y\, [P_X]}{K_Y\,(r_d + \mu)}, \quad a_3 = \frac{b_x\, k_x\, [P_G]}{K_X\,(r_d + \mu)}, \tag{S1.16a}$$

$$a_4 = \frac{b_y\, k_y\, [P_R]}{K_Y\,(r_d + \mu)}, \quad a_5 = \frac{b_u\, k_u\, k_G\, [P_Y]}{K_G\,(r_d + \mu)^2}, \quad a_6 = \frac{b_w\, k_w\, k_R[P_X]}{K_R\,(r_d + \mu)^2}. \tag{S1.16b}$$

2. For dimensionless diffusion and degradation parameters, we obtain:

$$d_g = \frac{D_G}{r_d + \mu}, \quad d_r = \frac{D_R}{r_d + \mu}, \quad \delta_g = \frac{r_G + \mu}{r_d + \mu}, \quad \delta_r = \frac{r_R + \mu}{r_d + \mu}, \quad \delta_e = \frac{\mu_e}{r_d + \mu}, \tag{S1.17a}$$

Molecular and biophysical parameter values used in the expressions (4) - (5) will be estimated in Sect. S1.3, while dimensionless parameters will be estimated in Sect. S1.3 In this section, we only mention that due to [2], we can set

$$K_{\mathrm{X}} = K_{\mathrm{Y}} = 40 \, \text{monomers per cell.} \tag{S1.18}$$

We could not find any estimation of values for the two parameters K_{G} for C14-HSL and K_{R} for C4-HSL in the literature despite the fact that more and more precise measurements of kinetic parameters become available [27]. We estimate the order of magnitude of K_{G} and K_{R} as follows.

In the detailed experimental results on the C4-HSL-mediated quorum sensing regulatory system of the opportunistic Gram-negative bacterium *Aeromonas hydrophila*, the concentration of C4-HSL was found to be of order of magnitude equal to $10 \, \mu\text{M}$ [70]. In *E.coli* biology, it is convenient to use nM units [38], because relative to the effective *E. coli* volume [43], the value of 1 nM corresponds to one molecule per cell. This fact is widely used in the literature [2]. Therefore, the above estimate of $10 \, \mu\text{M}$ corresponds to 10^4 C4-HSL signaling molecules per cell.

Another *ad-hoc* rule of *E.coli* biology used in a number of studies with the Cornell *E.coli* computer model [37, 44, 71, 72], resulting in a number of relevant predictions such as ribosomal-protein limitations, *lac*-control, plasmid stability, and *etc.* [30, 39–43, 45], is that, the coarse-grained estimation for the dissociation equilibrium constant to be used in the Hill function can be calculated as 25% of the intracellular modifier (reference) concentration. In our case, this yields a coarse-grade estimate of 0.25×10^4 C4-HSL signaling molecules per cell,

$$K_{\mathrm{G}} = K_{\mathrm{R}} = 2.5 \times 10^3 \, \text{molecules per cell.} \tag{S1.19}$$

The values for other parameters will be estimated below.

S1.3   Toggle Balancing (Symmetrization)

As it was observed experimentally [1,2], synthetic circuits can operate and deliver the engineered traits only if special molecular constrains are fulfilled,

(i) repressor protein half-lives are close to mRNA halflifes [2];

(ii) repressor protein half-lives are approximately equal [1,2].

Constrain *(ii)* is required for "balancing" the given circuit [1]. Moreover, both works suggest experimental interventions to fulfill the above constraints [1,2]. Such and similar interventions are termed a "tuning dials" in the review [6].

It is mathematically convenient for us to generalize the above balancing procedure by the procedure of "symmetrization" of two antagonistic, mutually repressing toggle subsystems by selecting synthetic (tuned) parameter values that would make two antagonist subsystems symmetric to one another. In other words, we assume that an *ideal* S toggle has a mirror $\mathbb{Z}_2$-symmetry group of permutations between the two antagonistic subsystems. Symmetry usually helps with analytical analysis of nonlinear mathematical models.

Specifically, we "symmetrize" (balance) biosynthesis kinetic rates, using constraints

$$a_1 = a_2, \quad a_3 = a_4 \quad \text{and} \quad a_5 = a_6. \tag{S1.20}$$

44

Appropriate molecular interventions, which can be used to set the relationships (S1.20) approximately under certain experimental conditions, are reviewed in [6]. Similarly, we symmertize "diffusion" parameters,

$$d_g = d_r = d, \tag{S1.21}$$

and the autoinducer "degradation" or "utilization" ("load") parameters,

$$\delta_g = \delta_r = \delta_e = \delta. \tag{S1.22}$$

S2 Estimation of Parameter Values

Toggle Parameters

In our estimations, we use general biochemical calculations [73, 74]. First, we collect generic prokaryotic and specific *E.coli* parameter values in Table S2.1.

Table S2.1. Generic prokaryotic and *E. coli* specific values of model parameters

Name	Description	Value	Unit	Reference
r_{mRNA}	Rate of transcription by RNA polymerase in prokaryotes	80	bp/sec	[46]
τ_{mRNA}	Typical half-life time for 80% of genes in *E. coli*	3 - 8	min	[75]
r_{aa}	Rate of translation by the ribosome in prokaryotes	20	aa/sec	[46][a]
k_{P}	Rate of translation by the ribosome in prokaryotes	1.71	sec^{-1}	estimated in (S2.3)
T	*E. coli* replication period under specific nutrition conditions	25	min	[43]
μ	Intracellular specific dilution rate due to *E. coli* cell growth	$\ln 2/T$	min^{-1}	[43]
μ_e	Extracellular dilution rate due to flow	0.1	min^{-1}	[76]
ρ	Total volume fraction of cells in chamber	0.8	—	[77]
N	Number of *E. coli* cells in an overnight population culture	10^9 (OD600 $= 1$)	cells/ml	[43]

[a]This estimate is smaller than the estimate 33 aa/sec used in [27].

A general rate of translation of protein P in prokaryotes (k_{P}). Suppose that an mRNA transcript of protein P contains n_{aa} amino acids. Then, for one ribosome to transcribe P from its mRNA transcript, assuming a translation rate of 20 amino acids per second (Table S2.1), it will take time

$$t_{\text{P}} = \frac{n_{\text{aa}}}{20}. \tag{S2.1}$$

The above estimates yields the rate per ribosome which is

$$\frac{1 \text{ molecule}}{t_{\text{P}}} = \frac{20}{n_{\text{aa}}} \text{ sec}^{-1}. \tag{S2.2}$$

Given that the coding region of protein P is $n_{\text{aa}} \times 3$ nucleotides long, and that a ribosome can attach every 35 nucleotides, we can estimate that $n_{\text{aa}} \times 3/35$ ribosomes can be attached per mRNA molecule. We, thus, obtain

$$k_{\text{P}} = \frac{20}{n_{\text{aa}}} \times \frac{n_{\text{aa}} \times 3}{35} = 1.71 \text{ protein molecules sec}^{-1}. \tag{S2.3}$$

To illustrate our parameter estimation procedure, we derive parameter values for the Lac-repressor subsystem only. Parameter values for all other subsystems can be derived similarly.

k_x : **Fully induced strength of promoters** P_Y (P_{tet}) **and** P_G. One *lacI* mRNA transcript is 1204 bases long (Table S2.2). To transcribe one molecule of *lacI* mRNA from one gene with a rate of 80 bases per second (Table S2.2) takes

$$\frac{1204 \text{ bases}}{80 \text{ bases/sec}} = 15.05 \text{ sec.} \tag{S2.4}$$

Table S2.2. Parameter values of the LacI-repressor subsystem.

Name	Description	Value	Unit	Reference
Transcription:				
L_{lacI}	Size of *lacI* gene	1204	bp	[78]
Γ_x	Repressed strength of promoter P_{tet}	5.0×10^{-4}	molecules/cell/sec	[2]
k_x	Fully induced transcription rate[a]	6.65×10^{-2}	species/cell/sec	estimated in (S2.5)
K_Y	The number of TetR to repress[b] P_{tet}	40	monomers/cell	estimated in (S1.18)
K_G	The number of C14-HSL to activate[c] P_{cin}	2.5×10^3	molecules/cell	estimated in (S1.19)
$\tau_{1/2,lacI}$	Half-life of *lacI* mRNA	3.8	min	[35]
r_x	Rate of *lacI* mRNA degradation	3.04×10^{-3}	sec^{-1}	estimated in (S2.7)
n_Y	The number of subunits in TetR	2		[34]
n_G	Hill coefficient of C14-HSL	4		[1]
Translation:				
L_{LacI}	Size of one subunit in tetrameric LacI	360	aa/subunit	[33]
k_X	Rate of LacI translation	1.71	molecules/cell/sec	estimated in (S2.3)
$\tau_{1/2,LacI}$	Half-life of LacI protein	4	min	[2]
r_X	Rate of LacI degradation	1.16×10^{-3}	sec^{-1}	estimated in (S2.7)

[a]Fully induced strength of promoters P_{tet} and P_{cin},
[b]For the sake of brevity, "to repress" means "to half-maximally repress."
[c]For the sake of brevity, "to activate" means "to half-maximally activate."

Then, per gene the estimate (S2.4) yields

$$k_x = \frac{1}{15.05 \text{ sec}} = 6.65 \times 10^{-2} \text{ lacI mRNA (molecules/cell/sec)}. \tag{S2.5}$$

The estimate (S2.4) is one order of magnitude less than the estimate 0.5 lacI mRNA (molecules/cell/sec) provided in [2].

r_x : **Rate of *lacI* mRNA degradation.**   The calculation of degradation rates for proteins is based on the known protein half-lives,

$$r_x = \frac{\ln 2}{t_{1/2}}. \tag{S2.6}$$

We obtain (Table S2.2),

$$r_x = \frac{\ln 2}{3.8 \times 60 \text{ sec}} = 3.04 \times 10^{-3} \text{ sec}^{-1}. \tag{S2.7}$$

r_X : **Rate of LacI (X) degradation.**   We obtain (Table S2.2),

$$r_d = r_X = \frac{\ln 2}{4 \times 60 \text{ sec}} = 2.89 \times 10^{-3} \text{ sec}^{-1}. \tag{S2.8}$$

We use the estimate (S2.8) for all proteins in the model.

Dimensionless parameter values

Using data from Table S2.1 and the estimate (S2.8), we obtain

$$\mu = \frac{\ln 2}{25 \times 60} = 0.46 \times 10^{-3} \text{ sec}^{-1}, \ r_d + \mu = 3.35 \times 10^{-3} \text{ sec}^{-1}. \tag{S2.9}$$

Table S2.3. Parameter values of the TetA-repressor subsystem.

Name	Description	Value	Unit	Reference
Transcription:				
L_{tetR}	Size of $tetR$ gene	905	bp	[79]
Γ_y	Repressed strength of promoter P_{lac}	5×10^{-4}	molecules/cell/sec	estimated in Table S2.2
k_y	Fully induced transcription rate[a]	8.87×10^{-2}	molecules/cell/sec	estimated
K_X	The number of LacI to repress[b] P_{lac}	40	monomers/cell	estimated in Table S2.2[d]
K_R	The number of C4-HSL to activate[c] P_{rhl}	5×10^4	monomers/cell	estimated in Table S2.2
$\tau_{1/2,tetR}$	Half-life of $tetR$ mRNA	0.5	min	[32]
r_y	Rate of $tetR$ mRNA degradation	2.31×10^{-2}	sec^{-1}	estimated
n_X	The number of subunits in LacI	2		[34]
n_R	Hill coefficient of C4-HSL	4		estimated in Table S2.2
Translation:				
L_{TetR}	Size of one subunit in tetrameric TetR	207	aa/subunit	[79]
k_Y	Rate of TetR translation	1.71	molecules/cell/sec	estimated in (S2.3)
$\tau_{1/2,TetR}$	Half-life of TetR protein	4	min	estimated Table S2.2
r_Y	Rate of TetR degradation	1.16×10^{-3}	sec^{-1}	estimated

[a]Fully induced strength of promoters P_{lac} and P_{tet},
[b]For the sake of brevity, "to repress" means "to half-maximally repress."
[c]For the sake of brevity, "to activate" means "to half-maximally activate."
[d] Equilibrium dissociation constant for LacI is 7.7×10^{-8} M [80].

Table S2.4. Parameter values of the 3-OH-C14-HSL/CinI-signaling subsystem.

Name	Description	Value	Unit	Reference
Transcription:				
L_{cinI}	Size of $cinI$ gene[a]	663	bp	estimated
Γ_u	Repressed strength of promoter P_{lac}	5×10^{-4}	molecules/cell/sec	estimated in Table S2.2
k_u	Fully induced strength of promoter P_{tet}	13.27×10^{-2}	molecules/cell/sec	estimated
$\tau_{1/2,cinI}$	Half-life of $cinI$ mRNA	6.6	min	arbitrary[b]
r_u	Rate of $cinI$ mRNA degradation	1.75×10^{-3}	sec^{-1}	estimated
Translation:				
L_{CinI}	Size of CinI aa-sequence	221	aa	http://string-db.org
k_U	Rate of CinI translation	1.71	molecules/cell/sec	estimated in Table S2.2
$\tau_{1/2,CinI}$	Half-life[a] of CinI protein	4	min	[2]
r_U	Rate of protein CinI degradation	1.16×10^{-3}	sec^{-1}	estimated in Table S2.2
Signaling:				
k_G	Maximal production rate of CinI	2	min^{-1}	[81]
r_G	Degradation rate of C14-HSL	0.002	hr^{-1}	[82][c]
D_G	Export rate of C14-HSL	2.1	min^{-1}	[83]

[a]The coding region of the gene has been estimated from its protein sequence size provided in the same table as $221 \times 3 = 663$.
[b]This estimate corresponds to a general (or typical) pattern for mRNA half-lives in $E.\ coli$ [84]. Note that [2] use a generic half-life parameter value of 2 min.
[c]Data for 3-OH-C12-HSL is used.

Table S2.5. Parameter values of the C4-HSL/RhlI-signaling subsystem.

Name	Description	Value	Unit	Reference
Transcription:				
L_{rhlI}	Length of *rhlI* gene	603	bp	[85]
Γ_w	Repressed strength of promoter P_{lac}	5×10^{-4}	molecules/cell/sec	estimated in Table S2.2
k_w	Fully induced strength of promoter P_{lac}	13.27×10^{-2}	molecules/cell/sec	estimated
$\tau_{1/2,rhII}$	Half-life of *rhlI* mRNA	6.6	min	[86][a]
r_w	Rate of *lacI* mRNA degradation	1.75×10^{-3}	sec^{-1}	estimated
Translation:				
L_{LacI}	Length of RhlI protein aa-sequence	196	aa	[87]
k_W	Rate of RhlI translation	1.71	molecules/cell/sec	estimated in Table S2.2
$\tau_{1/2,RhlI}$	Half-life of RhlI protein	4	min	estimated in Table S2.2
r_X	Rate of LacI degradation	1.16×10^{-3}	sec^{-1}	estimated
Signaling:				
k_R	Maximal production rate of C4-HSL by RhlI	16	min^{-1}	[81]
r_R	Degradation rate of C4-HSL	0.02	hr^{-1}	[82]
D_R	Export rate of C4-HSL	3.0	min^{-1}	[83]

[a]The half-life data for *lasI* mRNA is used because the degradation of *rhlI* is positively regulated by *LasI* [86] and, so, could have a longer half-life. This estimate is in line with a general (or typical) pattern for mRNA half-lives in *E. coli* [84]. Note that [2] use a genetic half-life parameter value of 2 min (Table S2.4).

To estimate rates a_i, $i = 1 \ldots 6$, defined in (4), we assume that the equalities $b_x = b_y = b_u = b_w = 10$ can be approximately set by using RBS-related interventions [6]. Also, to avoid competition for ribosomes, only a few plasmids bearing promoters P_X, P_Y, P_G, and P_R can be used. By selecting $[P_X] = [P_Y] = [P_G] = [P_R] = 1$ copies per cell, we obtain

$$a_1 = a_3 = \frac{10 \times \left(6.65 \times 10^{-2}\right) \times 1}{40 \times (3.35 \times 10^{-3})} = 4.96 \approx 5, \tag{S2.10a}$$

$$a_2 = a_4 = \frac{10 \times \left(8.87 \times 10^{-2}\right) \times 2}{40 \times (3.35 \times 10^{-3})} = 6.61 \approx 7, \tag{S2.10b}$$

$$a_5 = a_6 = \frac{10 \times \left(13.27 \times 10^{-2}\right) \times (2/60) \times 2}{(2.5 \times 10^3) \times (3.35 \times 10^{-3})^2} = 3.15 \approx 3. \tag{S2.10c}$$

Next, from (5), we obtain

$$d_g = \frac{2.1/60}{3.35 \times 10^{-3}} = 10.44, \tag{S2.11a}$$

$$d_r = \frac{3/60}{3.35 \times 10^{-3}} = 14.40, \tag{S2.11b}$$

$$\delta_g = \delta_r \approx \frac{\mu}{r_d + \mu} = \frac{0.46 \times 10^{-3}}{3.35 \times 10^{-3}} = 0.14, \tag{S2.11c}$$

$$\delta_e = \frac{0.1/60}{3.35 \times 10^{-3}} = 0.50. \tag{S2.11d}$$

We find the estimated values of the parameters to be of the same order of magnitude as the corresponding parameter values estimated and used in $[1, 2, 13, 14]$. Not enough is yet known about molecular interactions inside host cells to obtain highly precise descriptions [6]; it is common to computationally evaluate the effect of different values for rate parameters and even

for Hill exponents $[1, 2, 13, 14]$. Following $[1, 2, 13, 14]$, where genetic circuits built from similar elements have been studied, we have explored sets of parameter values which are close to the estimates given in (S2.10) and (S2.11), which ensure bistability in both S- and A-models, see the main text.

S3 Exponential Stability of Cellular Populations

A systematic analysis of dynamical mathematical models begins with finding equilibrium solutions followed by the analysis of their exponential stability [24]. The next step is often to carry out (local) bifurcation analysis of the equilibrium solutions, allowing for the exploration of "stability boundaries" in the parameter space [53]. Both stability and bifurcation analyses rely on the computation of the eigenvalues from the corresponding model linearizations [53].

A nontrivial specificity of the computation of eigenvalues for the stability and bifurcation analyses of the A- and S-population models is that both models with $N > 1$ are invariant with respect to the action or the given linear representation of the *symmetric* group $\mathbf{S}_N$ of permutations among N-cells [52]. It is known that irreducible representations of groups enforce *multiple* eigenvalues of matrices that commute with their linear representations, a well-known fact following from Schur's Lemma in the representation theory of Lie groups [52].

To take into account the necessity to deal with multiple eigenvalues in the situations when the value of N is *a priori* unknown, we have developed a general approach to the analysis of exponential stability [24] in arbitrary populations of identical cells, independently of N, as described below.

A General Population Model of Identical Cells

In this SI, we use Schur's formula [66] to compute explicitly the characteristic polynomials for the corresponding model linearizations. The most important implication of Schur's formula is that it can be easily seen that the values of the eigenvalues are independent of $N \geq 2$. A conceptually similar result on the exponential orbital stability of limit cycles in a system of both identical and slightly different oscillators coupled via a medium was obtained by E. E. Shnol [88] in 1987. In his work, an averaging technique over the entire cellular population was used for both homogeneous and heterogeneous populations. Some results obtained in [88] were independently rediscovered in 2008 by G. Katriel [89] for homogeneous populations only, using Floquet Theory.

To describe the general exponential stability analysis, we first introduce an appropriate notation as follows. Let S and z be "generalized" global (extracellular) and local (intracellular) state variables, respectively, $\dim S = m \geq 1$ and $\dim z = k \geq 2$. Using the generalized variables, both the S-model (1) and the A-model (2) can then be rewritten in the following general form, which we call a G-model,

$$\frac{dS}{dt} = H_0(S) + \frac{\rho}{N} \sum_{i=1}^{N} H(S, z_i), \qquad 0 \leq \rho \leq 1, \tag{S3.1a}$$

$$\frac{dz_i}{dt} = h(S, z_i), \quad i = 1, \ldots N. \tag{S3.1b}$$

The G-model (S3.1) includes $m + Nk$ equations.

A Homogeneous Population

In the case of a homogeneous population of identical cells, we have $z_i(t) \equiv z(t)$. As a result, the G-model (S3.1) reduces to a system of $(m + k)$-differential equations,

$$\frac{dS}{dt} = H_0(S) + \rho H(S, z), \qquad 0 \leq \rho \leq 1, \tag{S3.2a}$$

$$\frac{dz}{dt} = h(S, z). \tag{S3.2b}$$

Observe that the model (S3.2) describes a *single* cell placed in a "free", non-constant medium.

Definition 1. *Let (S_0, z_0) be an equilibrium solution of the model (S3.2). Then, (S_0, z_0) corresponds to a homogeneous population equilibrium solution,*

$$(S_0, z_0, \ldots, z_0) = (S_0, N z_0), \tag{S3.3}$$

of the full G-model (S3.1) for any $N \geq 2$. Notation $N z_0$ means that z_0 is repeated N-times in $(S_0, z_0, \ldots, z_0)$.

Although the model (S3.2) is sufficient to study the *existence* of homogeneous population equilibrium solutions (S3.3), it is not enough to establish the exponential *stability* of the corresponding solutions (S3.3). Let $(S_0, N z_0)$ be a homogeneous population equilibrium solution of the G-model (S3.1) with any fixed $N \geq 2$. To analyze the exponential stability of $(S_0, N z_0)$ in the "full" G-model (S3.1), we need to compute the eigenvalues of the corresponding Jacobian matrix $\mathbf{J}_N$,

$$\mathbf{J}_N = \begin{pmatrix} \mathbf{A} & \frac{\rho}{N}\mathbf{B} & \frac{\rho}{N}\mathbf{B} & \cdots & \frac{\rho}{N}\mathbf{B} \\ \mathbf{C} & \mathbf{D} & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{C} & \mathbf{O} & \mathbf{D} & \cdots & \mathbf{O} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C} & \mathbf{O} & \mathbf{O} & \cdots & \mathbf{D} \end{pmatrix}. \tag{S3.4}$$

In (S3.4), each of three matrices, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$, is repeated N-times; $\mathbf{A}$ and $\mathbf{D}$ are square matrices of sizes m and k, respectively; $\mathbf{B}$ and $\mathbf{C}$ are rectangular matrices of sizes $m \times k$ and $k \times m$, respectively,

$$\mathbf{A} = \frac{\partial H_0}{\partial S} + \rho \frac{\partial H}{\partial S}, \quad \mathbf{B} = \frac{\partial H}{\partial z}, \quad \mathbf{C} = \frac{\partial h}{\partial S}, \quad \mathbf{D} = \frac{\partial h}{\partial z}. \tag{S3.5}$$

All partial derivatives in the expressions (S3.5) are evaluated at (S_0, z_0) which depends on all G-model parameters with the one important exception that they are *independent* of N because (S_0, z_0) is obtained using (S3.2). Notation $\mathbf{O}$ corresponds to zero submatrices of appropriate sizes.

We call a square matrix *stable* if all its eigenvalues have strictly negative real parts. The following theorem holds for $\mathbf{J}_N$.

Theorem 2. (I). *Statements* (a), (b), *and* (c) *are equivalent.*

(a). *The matrix $\mathbf{J}_N$ is stable for all $N \geq 2$.*

(b). *The matrix $\mathbf{J}_1$ and its submatrix $\mathbf{D}$ are both stable.*

(c). *The matrix $\mathbf{J}_2$ is stable.*

(II). *The matrix $\mathbf{J}_N$ has typically k different eigenvalues, each of multiplicity $N-1$ in the following sense. Let $\{\lambda_1,\ldots,\lambda_{m+k}\}$ be the set of eigenvalues of matrix $\mathbf{J}_1$, and let $\{\mu_1,\ldots,\mu_k\}$ be the set of eigenvalues of its submatrix $\mathbf{D}$, Then,*

$$\{\lambda_1,\ldots,\lambda_{m+k},(N-1)(\mu_1,\ldots,\mu_k)\} \tag{S3.6}$$

is the set of all eigenvalues of matrix $\mathbf{J}_N$ for any $N \geq 2$, where $\{\mu_1,\ldots,\mu_k\}$ is repeated $(N-1)$-times.

Proof. Let λ be a complex number, $\lambda \in \mathbb{C}$. Consider a new matrix $\mathbf{M}_\lambda = \mathbf{J}_N - \lambda\mathbf{I}_{m+Nk}$, where $\mathbf{I}_{m+Nk}$ is the identity matrix of size $m + nk$. To find eigenvalues of $\mathbf{J}_N$, we need to write down the corresponding characteristic equation $P(\lambda) = 0$, $P(\lambda) = \det\mathbf{M}_\lambda$. Let us represent matrix $\mathbf{M}_\lambda$ in the form

$$\mathbf{M}_\lambda = \begin{pmatrix} \mathbf{A}_\lambda & \mathbb{B} \\ \mathbb{C} & \mathbb{D}_\lambda \end{pmatrix}. \tag{S3.7}$$

Here, matrices $\mathbf{A}_\lambda = \mathbf{A} - \lambda\mathbf{I}_m$, $\mathbb{B} = \frac{1}{N}(\mathbf{B},\ldots,\mathbf{B})$, $\mathbb{C} = (\mathbf{C},\ldots,\mathbf{C})^{\mathrm{T}}$, and $\mathbb{D}_\lambda = \mathrm{diag}(\mathbf{D}_\lambda,\ldots,\mathbf{D}_\lambda)$ with $\mathbf{D}_\lambda = \mathbf{D} - \lambda\mathbf{I}_k$. Next, assume for a moment that $\mathbb{D}_\lambda^{-1}$ exists. Then, Schur's formula can be used to compute $\det\mathbf{M}_\lambda$ [66],

$$\det\mathbf{M}_\lambda = \det\mathbb{D}_\lambda \cdot \det\left(\mathbf{A}_\lambda - \mathbb{B}\mathbb{D}_\lambda^{-1}\mathbb{C}\right). \tag{S3.8}$$

Next, we compute

$$\mathbb{B}\mathbb{D}_\lambda^{-1}\mathbb{C} = \mathbb{B}\left(\mathbb{D}_\lambda^{-1}\mathbb{C}\right) = \frac{1}{N}(\mathbf{B},\ldots,\mathbf{B})\begin{pmatrix} \mathbf{D}_\lambda^{-1}\mathbf{C} \\ \mathbf{D}_\lambda^{-1}\mathbf{C} \\ \ldots \\ \mathbf{D}_\lambda^{-1}\mathbf{C} \end{pmatrix} = \mathbf{B}\mathbf{D}_\lambda^{-1}\mathbf{C}. \tag{S3.9}$$

For the determinant of the block diagonal $\mathbb{D}_\lambda$, we obtain $\det\mathbb{D}_\lambda = (\det\mathbf{D}_\lambda)^N$. Substituting (S3.9) into (S3.8) yields

$$\det\mathbf{M}_\lambda = (\det\mathbf{D})^n \cdot \det\left(\mathbf{A}_\lambda - \mathbf{B}\mathbf{D}_\lambda^{-1}\mathbf{C}\right). \tag{S3.10}$$

Using the Schur's formula for the product $\det\mathbf{D}\cdot\det\left(\mathbf{A}_\lambda - \mathbf{B}\mathbf{D}_\lambda^{-1}\mathbf{C}\right)$ in the "backward" direction, we can rewrite (S3.10) in the following equivalent form

$$\det\mathbf{M}_\lambda = (\det\mathbf{D}_\lambda)^{N-1} \cdot \det\begin{pmatrix} \mathbf{A}_\lambda & \mathbf{B} \\ \mathbf{C} & \mathbf{D}_\lambda \end{pmatrix}. \tag{S3.11}$$

The expression (S3.11) can now be rewritten simply as

$$P(\lambda) = (\det\mathbf{D} - \lambda\mathbf{I}_k)^{N-1} \cdot \det(\mathbf{J}_1 - \lambda\mathbf{I}_{m+k}). \tag{S3.12}$$

Recall that the expression (S3.12) has been proven under a restrictive condition $\det\mathbb{D}_\lambda \neq 0$, see above, which means that λ is not an eigenvalue of the matrix $\mathbf{D}$. This restriction can be removed, for example, as follows. Let λ_0 be an eigenvalue of the matrix $\mathbf{D}$. Then, we obtain for the polynomial $P(\lambda)$ by continuity

$$P(\lambda_0) = \lim_{\lambda\to\lambda_0} P(\lambda) = \lim_{\lambda\to\lambda_0} (\det\mathbf{D} - \lambda\mathbf{I}_k)^{N-1} \cdot \det(\mathbf{J}_1 - \lambda\mathbf{I}_{m+k}) = 0. \tag{S3.13}$$

53

It follows from (S3.13) that (S3.12) holds for all $\lambda \in \mathbb{C}$.

Finally, we observe from (S3.12) that to compute all eigenvalues of the Jacobian matrix $\mathbf{J}_N$ for any $N \geq 2$, it is sufficient to compute the eigenvalues of either two smaller matrices, $\mathbf{D}$ and $\mathbf{J}_1$, or one matrix $\mathbf{J}_2$. The latter may be practically slightly easier than computing the eigenvalues for $\mathbf{D}$ and $\mathbf{J}_1$ separately. The proof of the theorem follows. $\qquad\square$

Consider a differential equation

$$\frac{dz}{dt} = h(S_0, z), \tag{S3.14}$$

where S_0 is a fixed parameter corresponding to the equilibrium $(S_0, N \times z_0)$ of the full G-model (S3.1). In contrast to equation (S3.2), equation (S3.14) describes a single cell placed into a constant environment, which can be interpreted as an environment shaped by the large population of cells and which does not "sense" any changes in a single cell. Additionally, consider a cascade model

$$\frac{dS}{dt} = H_0(S) + \rho H(S, z_1), \quad 0 \leq \rho \leq 1, \tag{S3.15a}$$

$$\frac{dz_j}{dt} = h(S, z_j), \quad j = 1, 2. \tag{S3.15b}$$

Observe that the variable z_2 is absent from the first equation (S3.15a) and, hence, (S3.15) cannot be obtained from (S3.1) by simply setting $N = 2$.

Then, using the definition of exponential stability [24], the first statement of Theorem 2 can be reformulated as the following corollary which admits an intuitive interpretation of the fact why the case of $N = 2$ is sufficient to study the exponential stability of homogeneous population solutions.

Corollary 1. *Let $(S_0, N \times z_0)$ be an equilibrium solution of the G-model (S3.1). Then, statements (a) - (d) are equivalent.*

(a). *$(S_0, N \times z_0)$ is exponentially stable in the G-model (S3.1) for any $N \geq 2$.*

(b). *(S_0, z_0) is exponentially stable in the reduced model (S3.2), and z_0 is exponentially stable in the single-cell model (S3.14).*

(c). *(S_0, z_0, z_0) is exponentially stable in the G-model (S3.1) at $N = 2$.*

(d). *(S_0, z_0, z_0) is exponentially stable in the cascade model (S3.15).*

A comparison of Statements *(a)* and *(b)* of Corollary 1 leads to a conclusion that the given population consisting of identical cells is stable with respect to any small perturbation if and only if *(i)* the population is stable with respect to any small *uniform* perturbation of the entire population described by system (S3.1) and, simultaneously, *(ii)* a majority of *unperturbed* cells forces a single *slightly perturbed* cell to re-join back the unperturbed majority.

Indeed, system (S3.14) used in Statement *(b)* means that the entire population does not sense small perturbations in a single cell because S_0 is fixed in (S3.14).

Note that both conditions in Statement *(b)* can be reformulated, using the cascade model (S3.15) from statement *(d)*. Finally, because the stability property is independent of the number N of identical cells in the population, the simple case of $N = 2$ can be used as given by statement *(c)*.

A Mixed Population Split into Two Subpopulations

Suppose now that the given population consisting of N, $N \geq 4$, identical cells is split into two different subpopulations of sizes $N_1 \geq 2$ and $N_2 \geq 2$, respectively, where $N = N_1 + N_2$. We always assume that each subpopulation consists of at least two cells. Then, the two different homogeneous subpopulations can be described by two state variables z_1 and z_2, respectively, where $z_1 \neq z_2$, that is, $z_{i_p}(t) \equiv z_1(t)$ for some subset of indexes i_p, $p = 1, \ldots, N_1$, and $z_{i_q}(t) \equiv z_2(t)$, for another subset of indexes i_q, $q = 1, \ldots, N_2$. It follows that the equation (S3.1a) from the G-model (S3.1) simplifies as follows

$$\frac{dS}{dt} = H_0(S) + \frac{\rho}{N} \sum_{i=1}^{N} H(S, z_i) = \rho\Big(\beta_1 H(S, z_1) + \beta_2 H(S, z_2)\Big). \qquad \text{(S3.16)}$$

In (S3.16), β_j is the fraction of the j-th subpopulation, $\beta_j = N_j/N$, $j = 1, 2$, $\beta_1 + \beta_2 = 1$. In this case, the entire G-model (S3.1) reduces to the following three equations

$$\dot{S} = H_0(S) + \rho\Big(\beta_1 H(S, z_1) + \beta_2 H(S, z_2)\Big), \quad \beta_j \in \mathbb{Q}, \quad \beta_1 + \beta_2 = 1, \qquad \text{(S3.17a)}$$

$$\dot{z}_j = h(S, z_j), \quad j = 1, 2. \qquad \text{(S3.17b)}$$

Definition 2. *Let (S_0, z_{10}, z_{20}), $z_{10} \neq z_{20}$, be a non-uniform equilibrium solution of the reduced system (S3.17). Then, (S_0, z_{10}, z_{20}), $z_{10} \neq z_{20}$ corresponds to a mixed population equilibrium solution,*

$$(S_0, z_{10}, \ldots, z_{10}, z_{20}, \ldots, z_{20}) = (S_0, N_1 z_{10}, N_2 z_{20}), \qquad \text{(S3.18)}$$

of the full G-model (S3.1). The solution (S3.18) describes a mixed population of N identical cells, split into two (non-identical) subpopulations of sizes $N_1 > 0$ and $N_2 > 0$, respectively, $N_1 + N_2 = N$. Notation $N_j z_{j0}$ means that z_{j0} is repeated N_j-times in $(S_0, z_{10}, \ldots, z_{10}, z_{20}, \ldots, z_{20})$, $j = 1, 2$.

Due to the condition $\beta_1 + \beta_2 = 1$ used in (S3.17a), there formally exists a *continuum* of different fractions $\beta_1 : \beta_2$, $\beta_1 \in \mathbb{R}$ and $\beta_2 \in \mathbb{R}$. Of course, in the *biological* sense, only rational values $\beta_1 \in \mathbb{Q}$ and $\beta_2 \in \mathbb{Q}$ are allowable, leading to infinitely many fractional $(\beta_1 : \beta_2)$-configurations in the subdivision of the original population into two different subpopulations. Simple examples of such situations can be easily presented (Fig. S3.1).
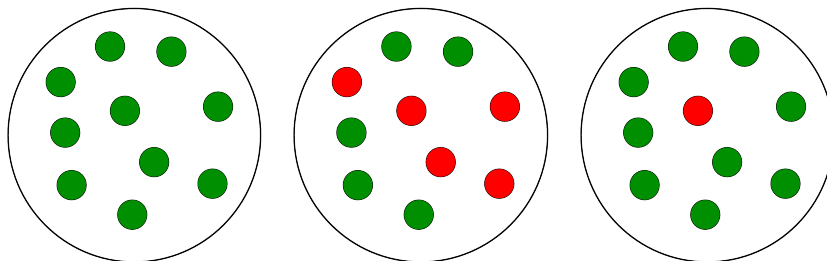

Figure S3.1. Examples of $(p : q)$-populations. The left panel corresponds to the case of a (10:0)-homogeneous population; the middle panel corresponds to the case of a (5:5)- or, equivalently, (1:1)-mixed population, and the right panel corresponds to a (9:1)-mixed population.

For the sake simplicity of the exponential stability analysis, we will always assume that both β_1 and β_2 are real numbers, that is, $\beta_j \in \mathbb{R}$, $j = 1, 2$.

Let $(S_0, N_1 z_{10}, N_2 z_{20})$ be a mixed population equilibrium solution of the G-model (S3.1) with any fixed $N \geq 4$, see (S3.3). To analyze the exponential stability of $(S_0, N_1 z_{10}, N_2 z_{20})$, we need to compute the eigenvalues of the corresponding Jacobian matrix $\mathbf{J}_N$ obtained from the G-model (S3.1),

$$\mathbf{J}_N = \begin{pmatrix} \mathbf{A} & \frac{\rho}{N}\mathbf{B}_1 & \cdots & \frac{\rho}{N}\mathbf{B}_1 & \frac{\rho}{N}\mathbf{B}_2 & \cdots & \frac{\rho}{N}\mathbf{B}_2 \\ \mathbf{C_1} & \mathbf{D}_1 & \cdots & \mathbf{D}_1 & \mathbf{O} & \cdots & \mathbf{O} \\ \vdots & \cdots & \ddots & \cdots & \vdots & \ddots & \vdots \\ \mathbf{C_1} & \mathbf{O} & \cdots & \mathbf{D}_1 & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{C_2} & \mathbf{O} & \cdots & \mathbf{O} & \mathbf{D}_2 & \cdots & \mathbf{O} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C_2} & \mathbf{O} & \cdots & \mathbf{O} & \mathbf{O} & & \mathbf{D}_2 \end{pmatrix}. \tag{S3.19}$$

In the matrix (S3.19), submatrices $\mathbf{B}_i$, $\mathbf{C}_i$, and $\mathbf{D}_i$, are repeated N_j-times; $\mathbf{A}$ and $\mathbf{D}_i$ are square matrices of sizes m, and k, respectively; $\mathbf{B}_i$ and $\mathbf{C}_i$ are rectangular matrices of sizes $m \times k$ and $k \times m$, respectively, and

$$\mathbf{A} = \frac{\partial H_0}{\partial S} + \sum_{j=1}^{2} \rho_j \frac{\partial H(S_0, z_{j0})}{\partial S}, \tag{S3.20a}$$

$$\mathbf{B}_j = \frac{\partial H(S_0, z_{j0})}{\partial z}, \quad \mathbf{C}_j = \frac{\partial h(S_0, z_{j0})}{\partial S}, \quad \mathbf{D}_j = \frac{\partial h(S_0, z_{j0})}{\partial z}, \quad j = 1, 2. \tag{S3.20b}$$

Consider the Jacobian matrix $\mathbf{Q}_2$ of size $m + 2k$ for the system (S3.17), computed at (S_0, z_{10}, z_{20}),

$$\mathbf{Q}_2 = \begin{pmatrix} \mathbf{A} & \rho_1 \mathbf{B}_1 & \rho_2 \mathbf{B}_2 \\ \mathbf{C}_1 & \mathbf{D}_1 & \mathbf{O} \\ \mathbf{C}_2 & \mathbf{O} & \mathbf{D}_2 \end{pmatrix} \tag{S3.21}$$

Theorem 3. (I). *Statements (a) and (b) are equivalent.*

(a). *The matrix $\mathbf{J}_N$ is stable for all $N \geq 4$, and with any $N_1 \geq 2$ and $N_2 \geq 2$ such that $N_1 + N_2 = N$.*

(b). *Matrix $\mathbf{Q}_2$, and its two submatrices, $\mathbf{D}_1$ and $\mathbf{D}_2$, are stable.*

(II). *Matrix $\mathbf{J}_N$ has typically $2k$ different multiple eigenvalues in the following sense. Let $\{\lambda_1, \ldots, \lambda_{m+2k}\}$ be the set of eigenvalues of $\mathbf{Q}_2$, let $\{\mu_1, \ldots, \mu_k\}$ be the set of eigenvalues of $\mathbf{D}_1$, and let $\{\sigma_1, \ldots, \sigma_k\}$ be the set of eigenvalues of $\mathbf{D}_2$. Then,*

$$\{\lambda_1, \ldots, \lambda_{m+k}, \ (N_1 - 1)(\mu_1, \ldots, \mu_k), \ (N_2 - 1)(\sigma_1, \ldots, \sigma_k)\}, \tag{S3.22}$$

is the set of all eigenvalues of the matrix $\mathbf{J}_N$ for any $N \geq 4$. In (S3.22), the set $\{\mu_1, \ldots, \mu_k\}$ is repeated $(N_1 - 1)$-times, and the set $\{\sigma_1, \ldots, \sigma_k\}$ is repeated $(N_2 - 1)$-times. To have a nonzero value of multiplicity $N_j - 1$ in (S3.22), condition $N_j \geq 2$ and, hence, $N \geq 4$, are natural requirements, $j = 1, 2$. In other words, the latter two conditions guarantee that both matrices D_1 and D_2 exist. Otherwise, Theorem 3 does not make any sense.

Proof. The proof Theorem 3 can be carried out, using a simple modification of the proof of Theorem 2. For this reason, we only provide a brief sketch of the proof for Theorem 3. Similarly

to the proof of Theorem 2, we need to write down a characteristic equation $P(\lambda) = 0$ Here $P(\lambda) = \det \mathbf{M}_\lambda$, and matrix $\mathbf{M}_\lambda$ can be defined as in (S3.7), using appropriate submatrices,

$$\mathbf{A}_\lambda = \mathbf{A} - \lambda \mathbf{I}_m, \tag{S3.23a}$$

$$\mathbb{B} = \frac{1}{N}\left(\mathbf{B}_1, \ldots, \mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_2\right), \tag{S3.23b}$$

$$\mathbb{C} = \left(\mathbf{C}_1, \ldots, \mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_2\right)^{\mathrm{T}}, \tag{S3.23c}$$

$$\mathbb{D}_\lambda = \operatorname{diag}\left(\mathbf{D}_{\lambda 1}, \ldots, \mathbf{D}_{1\lambda}, \mathbf{D}_{2\lambda}, \ldots, \mathbf{D}_{2\lambda}\right), \quad \mathbf{D}_{i\lambda} = \mathbf{D}_i - \lambda \mathbf{I}_k. \tag{S3.23d}$$

In the above submatrix definitions, the matrices with index j are repeated N_j-times, $i, j = 1, 2$. In this case, $\det \mathbf{M}_\lambda$ can also be computed using Schur's formula, see (S3.8). However, (S3.9) should be replaced by

$$\mathbb{B}\mathbb{D}_\lambda^{-1}\mathbb{C} = \rho_1 \mathbf{B}_1 \mathbf{D}_{1\lambda}^{-1}\mathbf{C}_1 + \rho_2 \mathbf{B}_2 \mathbf{D}_{2\lambda}^{-1}\mathbf{C}_2. \tag{S3.24}$$

For the block diagonal matrix $\mathbb{D}_\lambda$, we obtain $\det \mathbb{D}_\lambda = (\det \mathbf{D}_{1\lambda})^{N_1} \cdot (\det \mathbf{D}_{2\lambda})^{N_2}$. Now, similarly to (S3.10), we will have

$$\det \mathbf{M}_\lambda = (\det \mathbf{D}_1)^{N_1} \cdot (\det \mathbf{D}_2)^{N_2} \cdot \det\left(\mathbf{A}_\lambda - \rho_1 \mathbf{B}_1 \mathbf{D}_{1\lambda}^{-1}\mathbf{C}_1 - \rho_2 \mathbf{B}_2 \mathbf{D}_{2\lambda}^{-1}\mathbf{C}_2\right). \tag{S3.25}$$

Using the Schur's formula in the "backward" direction, we will then have

$$P(\lambda) = (\det \mathbf{D}_1 - \lambda \mathbf{I}_k)^{N_1 - 1} \cdot (\det \mathbf{D}_2 - \lambda \mathbf{I}_k)^{N_2 - 1} \cdot \det \mathbf{Q}_2. \tag{S3.26}$$

The rest can be proved as in the proof for Theorem 2. The proof of Theorem 3 follows. $\qquad \square$

Consider the following cascade model

$$\frac{dS}{dt} = H_0(S) + \rho\Big(\beta_1 H(S, z_1) + \beta_2 H(S, z_3)\Big), \quad \beta_1 + \beta_2 = 1, \tag{S3.27a}$$

$$\frac{dz_j}{dt} = h(S, z_j), \quad j = 1, \ldots, 4. \tag{S3.27b}$$

Variables z_2 and z_4 are absent from the first equation (S3.27a) and, hence, the cascade system (S3.27) cannot be obtained from the G-model (S3.1) by simply setting $N = 4$. Now, Theorem 3 can be reformulated in terms its Corollary 2 as follows.

Corollary 2. *Let $(S_0, N_1 \times z_{10}, N_2 \times z_{20})$ be a mixed population equilibrium solution of the G-model (S3.1). Then, Statements (a) - (c) are equivalent.*

(a). *$(S_0, N_1 \times z_{10}, N_2 \times z_{20})$ is exponentially stable in the G-model (S3.1) for any $N \geq 4$, and with any $N_1 \geq 2$ and $N_2 \geq 2$ such that $N_1 + N_2 = N$.*

(b). *(S_0, z_{10}, z_{20}) is exponentially stable in the reduced model (S3.17), and each z_{j0} is exponentially stable in the single-cell model (S3.14), $j = 1, 2$.*

(c). *$(S_0, z_{10}, z_{10}, z_{20}, z_{20})$ is exponentially stable in the cascade model (S3.27).*

A Mixed Population Split into L Subpopulations

The case of a mixed population split into two subpopulations with densities $\rho_1 = \beta_1 \rho$ and $\rho_2 = \beta_2 \rho$ can be generalized to the case of a mixed population split into L-different subpopulations with densities $\rho_1, \ldots, \rho_L$, where $\rho_1 + \ldots + \rho_L = \rho$, $L \geq 3$ as follows.

Let subpopulation j consist of N_j cells, and let subpopulation j correspond to variable z_j, that is, we have $z_{j'}(t) \equiv z_j(t)$, where $j' \in \{i_1, i_2, \ldots, i_{N_j}\} \subset \{1, 2, \ldots N\}$, $j = 1, \ldots, L$. In this case, the G-model (S3.1) reduces to the following equations

$$\dot{S} = H_0(S) + \sum_{j=1}^{L} \rho_j H(S, z_j), \quad \sum_{j=1}^{L} \rho_j = \rho, \tag{S3.28a}$$

$$\dot{z}_j = h(S, z_j), \quad \rho_j = \beta_j \rho, \quad \beta_j = \frac{N_j}{N}, \quad j = 1, \ldots, L. \tag{S3.28b}$$

Definition 3. *Let $(S_0, z_{10}, \ldots, z_{L0})$ be a non-uniform equilibrium solution of the system (S3.28), where $z_{j0} \neq z_{j'0}$ for all $j \neq j'$. Then, $(S_0, z_{10}, \ldots, z_{L0})$, corresponds to a mixed population equilibrium solution,*

$$(S_0, N_1 z_{10}, \ldots, N_L z_{L0}), \tag{S3.29}$$

of the full G-model (S3.1). The solution (S3.29) describes a mixed population of N identical cells, which is split into L subpopulations of the corresponding sizes $N_j \geq 2$, $N_1 + \ldots + N_L = N$. Notation $N_j z_{j0}$ means that z_{j0} is repeated N_j-times in the vector-form solution of the the full G-model (S3.1), $N_j \geq 2$, $j = 1, \ldots, L$.

Consider the Jacobian matrix $\mathbf{Q}_L$ for the reduced system (S3.28), computed at $(S_0, z_{10}, \ldots, z_{L0})$,

$$\mathbf{Q}_L = \begin{pmatrix} \mathbf{A} & \rho_1 \mathbf{B}_1 & \ldots & \rho_L \mathbf{B}_L \\ \mathbf{C}_1 & \mathbf{D}_1 & \ldots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_L & \mathbf{O} & \ldots & \mathbf{D}_L \end{pmatrix}. \tag{S3.30}$$

In (S3.30), all submatrices are defined as in (S3.20), where $j = 1, 2$ should be replaced by $j = 1, \ldots, L$. Below, we formulate Theorem 4 and Corollary 3 without any proof because they are similar to Theorem 3 and Corollary 1, respectively.

Theorem 4. (I). *Statements* (a) *and* (b) *are equivalent.*

(a). *The Jacobian $\mathbf{J}_N$ computed for the G-model at the given equilibrium (S3.29) is stable for all $N \geq 2L$, and with any $N_j \geq 2$, $j = 1, \ldots, L$, such that $N_1 + \ldots N_L = N$.*

(b). *Matrix $\mathbf{Q}_L$ and its submatrices $\mathbf{D}_j$, $j = 1, \ldots, L$, are stable.*

(II). *Matrix $\mathbf{J}_N$ has typically kL different multiple eigenvalues in the following sense. Let $\{\lambda_1, \ldots, \lambda_{m+kL}\}$ be the set of eigenvalues of $\mathbf{Q}_L$, and let $\{\mu_1^{(j)}, \ldots, \mu_k^{(j)}\}$ be the set of eigenvalues of $\mathbf{D}_j$, $j = 1, \ldots, L$. Then,*

$$\left\{ \lambda_1, \ldots, \lambda_{m+k}, \ (N_1 - 1) \left(\mu_1^{(1)}, \ldots, \mu_k^{(1)} \right), \ldots, (N_L - 1) \left(\mu_1^{(N_L)}, \ldots, \mu_k^{(N_L)} \right) \right\} \tag{S3.31}$$

is the set of all eigenvalues of the matrix $\mathbf{J}_N$ for any $N \geq 2L$. In (S3.31), each set $\{\mu_1^{(j)}, \ldots, \mu_k^{(j)}\}$ is repeated $(N_j - 1)$-times with all $N_j \geq 2$, $j = 1, \ldots, K$.

Consider the following cascade model

$$\frac{dS}{dt} = H_0(S) + \sum_{j=1}^{L} \rho_j H(S, z_{2j-1}), \quad \sum_{j=1}^{L} \rho_j = \rho, \tag{S3.32a}$$

$$\frac{dz_j}{dt} = h(S, z_j), \quad j = 1, \ldots, 2L. \tag{S3.32b}$$

State variables z_{2j} with even indeces are absent from the first equation (S3.32a) of the cascade model (S3.32).

Corollary 3. *Let $(S_0, N_1 z_{10}, \ldots, N_L z_{L0})$ be a mixed equilibrium solution of the G-model (S3.1), where $N_{j0} \neq N_{j'0}$ for all $j \neq j'$. Then, Statements (a) - (c) are equivalent.*

(a). *$(S_0, N_1 z_{10}, \ldots, N_L z_{L0})$ is exponentially stable in the G-model (S3.1) for any $N \geq 2L$, and with any $N_j \geq 2$, $j = 1, \ldots, L$, such that $N_1 + \ldots N_L = N$.*

(b). *$(S_0, z_{10}, \ldots, z_{L0})$ is exponentially stable in the reduced model (S3.28), and each z_{j0} is exponentially stable in the single-cell model (S3.14), $j = 1, \ldots, L$.*

(c). *$(S_0, z_{10}, z_{10}, z_{20}, z_{20}, \ldots, z_{L0}, z_{L0})$ is exponentially stable in the cascade model (S3.32).*

Acknowledgments

References

1. Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. Nature. 2000;403(6767):339–342.

2. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. Nature. 2000;403(6767):335–338.

3. Hasty J, Pradines J, Dolnik M, Collins JJ. Noise-based switches and amplifiers for gene expression. Proceedings of the National Academy of Sciences. 2000;97(5):2075–2080.

4. Kobayashi H, Kærn M, Araki M, Chung K, Gardner TS, Cantor CR, et al. Programmable cells: interfacing natural and engineered gene networks. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(22):8414–8419.

5. Purnick PE, Weiss R. The second wave of synthetic biology: from modules to systems. Nature reviews Molecular cell biology. 2009;10(6):410–422.

6. Arpino JA, Hancock EJ, Anderson J, Barahona M, Stan GBV, Papachristodoulou A, et al. Tuning the dials of synthetic biology. Microbiology. 2013;159(Pt 7):1236–1253.

7. Nadell CD, Bassler BL, Levin SA. Observing bacteria through the lens of social evolution. Journal of biology. 2008;7:27.

8. Miller MB, Bassler BL. Quorum sensing in bacteria. Annual Reviews in Microbiology. 2001;55(1):165–199.

9. Pesci EC, Pearson JP, Seed PC, Iglewski BH. Regulation of *las* and *rhl* quorum sensing in Pseudomonas aeruginosa. Journal of bacteriology. 1997;179(10):3127–3132.

10. Lithgow JK, Wilkinson A, Hardman A, Rodelas B, Wisniewski-Dyé F, Williams P, et al. The regulatory locus cinRI in *Rhizobium leguminosarum* controls a network of quorum-sensing loci. Molecular microbiology. 2000;37(1):81–97.

11. McAnulla C, Edwards A, Sanchez-Contreras M, Sawers RG, Downie JA. Quorum-sensing-regulated transcriptional initiation of plasmid transfer and replication genes in *Rhizobium leguminosarum biovar viciae*. Microbiology. 2007;153(7):2074–2082.

12. Garcia-Ojalvo J, Elowitz MB, Strogatz SH. Modeling a synthetic multicellular clock: repressilators coupled by quorum sensing. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(30):10955–10960.

13. Kuznetsov A, Kærn M, Kopell N. Synchrony in a population of hysteresis-based genetic oscillators. SIAM Journal on Applied Mathematics. 2004;65(2):392–425.

14. Wang J, Zhang J, Yuan Z, Zhou T. Noise-induced switches in network systems of the genetic toggle switch. BMC systems biology. 2007;1(1):50.

15. Smale S. On the differential equations of species in competition. Journal of Mathematical Biology. 1976;3:5–7.

16. Hirsch MW. The dynamical systems approach to differential equations. Bull AMS. 1984;11:1–64.

17. Smith HL. Monotone Dynamical Systems, Vol. 41. Mathematical Surveys and Monographs. 1995;.

18. Angeli D, Sontag ED. Monotone control systems. IEEE Trans Automat Control. 2003;48(10):1684–1698.

19. Hirsch M, Smith H. Monotone dynamical systems. Handbook of differential equations: ordinary differential equations. 2005;2:239–357.

20. Sontag ED. Monotone and near-monotone biochemical networks. Systems and Synthetic Biology. 2007;1:59–87.

21. Smith HL. Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems. vol. 41. American Mathematical Soc.; 2008.

22. Parsek MR, Greenberg EP. Acyl-homoserine lactone quorum sensing in gram-negative bacteria: a signaling mechanism involved in associations with higher organisms. Proceedings of the National Academy of Sciences. 2000;97(16):8789–8793.

23. Resmi V, Ambika G, Amritkar R. Synchronized states in chaotic systems coupled indirectly through a dynamic environment. Physical Review E. 2010;81(4):046216.

24. Sontag ED. Mathematical control theory: deterministic finite dimensional systems. vol. 6. Springer; 1998.

25. Parsek MR, Val DL, Hanzelka BL, Cronan JE, Greenberg E. Acyl homoserine-lactone quorum-sensing signal generation. Proceedings of the National Academy of Sciences. 1999;96(8):4360–4365.

26. Hooshangi S, Thiberge S, Weiss R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(10):3581–3586.

27. Tuttle LM, Salis H, Tomshine J, Kaznessis YN. Model-driven designs of an oscillating gene network. Biophysical journal. 2005;89(6):3873–3883.

28. Stricker J, Cookson S, Bennett MR, Mather WH, Tsimring LS, Hasty J. A fast, robust and tunable synthetic gene oscillator. Nature. 2008;456(7221):516–519.

29. Hussain F, Gupta C, Hirning AJ, Ott W, Matthews KS, Josić K, et al. Engineered temperature compensation in a synthetic genetic clock. Proceedings of the National Academy of Sciences. 2014;111(3):972–977.

30. Laffend L, Shuler M. Structured model of genetic control via the *lac* promoter in *Escherichia coli*. Biotechnology and bioengineering. 1994;43(5):399–410.

31. Lutz R, Bujard H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic acids research. 1997;25(6):1203–1210.

32. Baumeister R, Flache P, Melefors Ö, von Gabain A, Hillen W. Lack of a 5'non-coding region in Tn 1721 encoded *tetR* mRNA is associated with a low efficiency of translation and a short half-life in *Escherichia coil*. Nucleic acids research. 1991;19(17):4595–4600.

33. Markiewicz P, Kleina LG, Cruz C, Ehret S, Miller JH. Genetic studies of the lac repressor. XIV. Analysis of 4000 altered *Escherichia coli* lac repressors reveals essential and non-essential residues, as well as" spacers" which do not require a specific sequence. Journal of molecular biology. 1994;240(5):421–433.

34. Ramos JL, Martínez-Bueno M, Molina-Henares AJ, Terán W, Watanabe K, Zhang X, et al. The TetR family of transcriptional repressors. Microbiology and Molecular Biology Reviews. 2005;69(2):326–356.

35. Semsey S, Jauffred L, Csiszovszki Z, Erdőssy J, Stéger V, Hansen S, et al. The effect of LacI autoregulation on the performance of the lactose utilization system in *Escherichia coli*. Nucleic acids research. 2013;p. gkt351.

36. Gray KM, Garey JR. The evolution of bacterial LuxI and LuxR quorum sensing regulators. Microbiology. 2001;147(8):2379–2387.

37. Shuler M, Leung S, Dick C. A mathematical model for the growth of a single bacterial cell. Annals of the New York Academy of Sciences. 1979;326(1):35–52.

38. Neidhardt FC, Ingraham JL, Schaechter M. Physiology of the Bacterial Cell: A Molecular Approach. Sinauer Associates Sunderland, MA; 1990.

39. Kim BG, Shuler M. A structured, segregated model for genetically modified *Escherichia coli* cells and its use for prediction of plasmid stability. Biotechnology and bioengineering. 1990;36(6):581–592.

40. Shu J, Shuler M. Prediction of effects of amino acid supplementation on growth of *E. coli* B/r. Biotechnology and bioengineering. 1991;37(8):708–715.

41. Laffend L, Shuler M. Ribosomal protein limitations in *Escherichia coli* under conditions of high translational activity. Biotechnology and bioengineering. 1994;43(5):388–398.

42. Bailey JE. Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. Biotechnology progress. 1998;14(1):8–20.

43. Domach M, Leung S, Cahn R, Cocks G, Shuler M. Computer model for glucose-limited growth of a single cell of *Escherichia coli* B/r-A. Biotechnology and bioengineering. 2000;67(6):827–840.

44. Nikolaev EV, Atlas JC, Shuler ML. Computer models of bacterial cells: from generalized coarsegrained to genome-specific modular models. In: Journal of Physics: Conference Series. vol. 46. IOP Publishing; 2006. p. 322.

45. Atlas J, Nikolaev E, Browning S, Shuler M. Incorporating genome-wide DNA sequence information into a dynamic whole-cell model of *Escherichia coli*: Application to DNA replication. Systems Biology, IET. 2008;2(5):369–382.

46. Lodish H, Berk A, Kaiser CA, Krieger M, Bretscher A, Ploegh H, et al. Molecular cell biology. Macmillan; 2012.

47. Dockery JD, Keener JP. A mathematical model for quorum sensing in *Pseudomonas aeruginosa*. Bulletin of mathematical biology. 2001;63(1):95–116.

48. Bressloff PC. Stochastic processes in cell biology. vol. 41. Springer; 2014.

49. Dong YH, Wang LH, Xu JL, Zhang HB, Zhang XF, Zhang LH. Quenching quorum-sensing-dependent bacterial infection by an N-acyl homoserine lactonase. Nature. 2001;411(6839):813–817.

50. Waldherr S, Allgöwer F. Network-level dynamics of diffusively coupled cells. 51st IEEE Conference on Decision and Control December 10-13, Maui, Hawaii, USA. 2012;.

51. Farrell CM, Baker TA, Sauer RT. Altered specificity of a AAA+ protease. Molecular cell. 2007;25(1):161–166.

52. Golubitsky M, Schaeffer DG, Stewart I. Singularities and groups in bifurcation theory. vol. 2. Springer New York; 1988.

53. Kuznetsov YA. Elements of applied bifurcation theory. vol. 112. Springer Science & Business Media; 2013.

54. Krupa M. Bifurcations of relative equilibria. SIAM journal on mathematical analysis. 1990;21(6):1453–1486.

55. Hirsch M. Differential equations and convergence almost everywhere in strongly monotone flows. Contemporary Mathematics. 1983;17:267–285.

56. Hirsch M. Systems of differential equations that are competitive or cooperative II: Convergence almost everywhere. SIAM J Mathematical Analysis. 1985;16:423–439.

57. Smith H. Monotone dynamical systems: An introduction to the theory of competitive and cooperative systems, Mathematical Surveys and Monographs, vol. 41. Providence, RI: AMS; 1995.

58. Hadeler K, Glas D. Quasimonotone systems and convergence to equilibrium in a population genetics model. J Math Anal Appl. 1983;95:297–303.

59. Hirsch M, Smith H, et al. Monotone dynamical systems. Handbook of differential equations: ordinary differential equations. 2005;2:239–357.

60. Zaslavsky T. Bibliography of signed and gain graphs. Electronic Journal of Combinatorics. 1998;DS8.

61. Harary F. On the notion of balance of a signed graph. Michigan Mathematical Journal. 1953;2:143?146.

62. Dhooge A, Govaerts W, Kuznetsov YA. MATCONT: a MATLAB package for numerical bifurcation analysis of ODEs. ACM Transactions on Mathematical Software (TOMS). 2003;29(2):141–164.

63. Govaerts W, Kuznetsov YA, Dhooge A. Numerical continuation of bifurcations of limit cycles in MATLAB. SIAM journal on scientific computing. 2005;27(1):231–252.

64. Khibnik AI, Kuznetsov YA, Levitin VV, Nikolaev EV. Continuation techniques and interactive software for bifurcation analysis of ODEs and iterated maps. Physica D: Nonlinear Phenomena. 1993;62(1):360–371.

65. Vasil'eva AB, Butuzov VF, Kalachev LV. The boundary function method for singular perturbation problems. SIAM; 1995.

66. Gantmakher FR. The theory of matrices. vol. 131. American Mathematical Soc.; 1959.

67. Bolouri H. Computational modeling of gene regulatory networks: a primer. World Scientific; 2008.

68. Sauro HM. Enzyme kinetics for systems biology. Future Skill Software; 2011.

69. Cornish-Bowden A. Fundamentals of enzyme kinetics. John Wiley & Sons; 2013.

70. Garde C, Bjarnsholt T, Givskov M, Jakobsen TH, Hentzer M, Claussen A, et al. Quorum sensing regulation in *Aeromonas hydrophila*. Journal of molecular biology. 2010;396(4):849–857.

71. Shuler M. Single-cell models: promise and limitations. Journal of biotechnology. 1999;71(1):225–228.

72. Castellanos M, Wilson DB, Shuler ML. A modular minimal cell model: purine and pyrimidine transport and metabolism. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(17):6681–6686.

73. Segel IH. Biochemical Calculations: How To Solve Mathematical Problems In General Biochemistry. Wiley; 1976.

74. Bhattacharya BS, Sweby PK, Minihane AM, Jackson KG, Tindall MJ. A mathematical model of the sterol regulatory element binding protein 2 cholesterol biosynthesis pathway. Journal of Theoretical Biology. 2014;p. 150–162.

75. Bernstein JA, Khodursky AB, Lin PH, Lin-Chao S, Cohen SN. Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. Proceedings of the National Academy of Sciences. 2002;99(15):9697–9702.

76. Prindle A, Selimkhanov J, Li H, Razinkov I, Tsimring LS, Hasty J. Rapid and tunable post-translational coupling of genetic circuits. Nature. 2014;.

77. Thomas PW, Stone EM, Costello AL, Tierney DL, Fast W. The quorum-quenching lactonase from *Bacillus thuringiensis* is a metalloprotein. Biochemistry. 2005;44(20):7559–7569.

78. Raj M. PCR Amplification, and Sequence Comparison of lacI gene in WT *E. coli* C29 cells and a presumptive *lacI* Knockout *E. coli* C29 cells to Determine the Difference in the Basal Expression Level of *lacZ* in Lac Operon. J Exp Microbiol Immunol. 2004;6:13–19.

79. Postle K, Nguyen TT, Bertrand KP. Nucleotide sequence of the repressor gene of the TN10 tetracycline resistance determinant. Nucleic acids research. 1984;12(12):4849–4863.

80. Wang Y, Tegenfeldt JO, Reisner W, Riehn R, Guan XJ, Guo L, et al. Single-molecule studies of repressor–DNA interactions show long-range interactions. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(28):9796–9801.

81. Pai A, You L. Optimal tuning of bacterial sensing potential. Molecular systems biology. 2009;5(1).

82. Chen CC, Riadi L, Suh SJ, Ohman DE, Ju LK. Degradation and synthesis kinetics of quorum-sensing autoinducer in *Pseudomonas aeruginosa* cultivation. Journal of biotechnology. 2005;117(1):1–10.

83. Flynn JM, Levchenko I, Seidel M, Wickner SH, Sauer RT, Baker TA. Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. Proceedings of the National Academy of Sciences. 2001;98(19):10584–10589.

84. Selinger DW, Saxena RM, Cheung KJ, Church GM, Rosenow C. Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. Genome research. 2003;13(2):216–223.

85. Ochsner UA, Reiser J. Autoinducer-mediated regulation of rhamnolipid biosurfactant synthesis in *Pseudomonas aeruginosa*. Proceedings of the National Academy of Sciences. 1995;92(14):6424–6428.

86. Takaya A, Tabuchi F, Tsuchiya H, Isogai E, Yamamoto T. Negative regulation of quorum-sensing systems in *Pseudomonas aeruginosa* by ATP-dependent Lon protease. Journal of bacteriology. 2008;190(12):4181–4188.

87. Brint JM, Ohman DE. Synthesis of multiple exoproducts in *Pseudomonas aeruginosa* is under the control of RhlR-RhlI, another set of regulators in strain PAO1 with homology to the autoinducer-responsive LuxR-LuxI family. Journal of Bacteriology. 1995;177(24):7155–7163.

88. Shnol E. The synchronization of oscillators which interact via a medium. Journal of Applied Mathematics and Mechanics. 1987;51(1):9–13.

89. Katriel G. Synchronization of oscillators coupled through an environment. Physica D: Nonlinear Phenomena. 2008;237(22):2933–2944.