CONTROLLING REVERSIBLE CELL DIFFERENTIATION FOR LABOR DIVISION IN MICROBIAL CONSORTIA

Davide Salzano¹

Davide Fiore²

Mario di Bernardo^{1,3,4,*}

SUMMARY

We address the problem of regulating and keeping at a desired balance the relative numbers between cells exhibiting a different phenotype within a monostrain microbial consortium. We propose a strategy based on the use of external control inputs, assuming each cell in the community is endowed with a reversible, bistable memory mechanism. Specifically, we provide a general analytical framework to guide the design of external feedback control strategies aimed at balancing the ratio between cells whose memory is stabilized at either one of two equilibria associated to different cell phenotypes. We demonstrate the stability and robustness properties of the control laws proposed and validate them in silico by means of a realistic agent-based implementation enabling cooperative bioproduction in microbial consortia. The proposed control framework may be used to allow long term coexistence of different populations, with both industrial and biotechnological applications.

¹Department of Electrical Engineering and Information Technology, University of Naples Federico II, Via Claudio 21, 80125 Naples, Italy

²Department of Mathematics and Applications "R. Caccioppoli", University of Naples Federico II, Via Cintia, Monte S. Angelo, 80126 Naples, Italy

³Scuola Superiore Meridionale, Largo S. Marcellino 10, 80138 Naples, Italy

⁴Lead contact

*Correspondence: mario.dibernardo@unina.it

Figure 1: Microbial consortia composed of reversible differentiable cells can be balanced in real-time by means of external feedback controllers to guarantee efficient labor division. (a) Reversible differentiable cells can carry out different roles by activating/deactivating specific set of genes, depending on which state of the internal bistable memory is currently active. Cells can change role in response to exogenous stimuli from the environment, e.g., injection of inducer molecules or light. (b) Cells can, for example, either grow and duplicate or produce some desired molecule, or they can produce two different molecules that react and produce the desired final bioproduct. (c) Cells expressing different genes grow also at different rates, and thus their coexistence can be compromised. Feedback control algorithms can be employed to regulate in real-time the relative number of cells in the two groups, so that a balance in the population numbers and in the expression of desired genes is always guaranteed.

1 Introduction

Synthetic Biology aims at engineering biological systems with new functionalities [1], with applications ranging from health treatments to bioremediation [2], production of biofuels and drugs in bioreactors [3]. This is made possible by embedding artificial genetic circuits into living cells, such as bacteria, yeast, and fungi, modifying their natural behavior [4]; that is, by synthetically modifying when and how much genes are expressed to produce proteins or other chemicals of interest. However, the level of complexity and the functions of such engineered genetic circuits are limited by intrinsic factors in the host cells, such as possible excessive metabolic burden, competition of limited resources and incompatible chemical reactions [5]. To overcome these limitations, a promising strategy is to distribute the required functionalities among multiple cell populations forming a microbial consortium, so that each cell strain embeds a smaller subset of engineered gene networks. [6–9]. In this way, each cell population carries out a specialized function and, by dividing labor with the others in the consortium, contributes more efficiently to the achievement of the overall final goal.

Unfortunately, this solution introduces an additional challenge; cells expressing different genes might also grow and divide at different rates. In particular, cells in the consortium whose function is associated to a lower metabolic burden will grow faster, eventually becoming dominant over the other populations; thus compromising the overall function of the consortium and giving rise to undesired dynamics, such as oscillations, or even extinction [10]. Therefore, it is crucial to develop methods to guarantee the stable coexistence between cell populations in a consortium by regulating and maintaining their relative numbers to some desired level, adjusting it to the requirements of the specific application of interest. This is possible, as suggested in [11], by using *feedback control algorithms* able to sense the relative size of all the populations involved and respond by applying appropriate stimuli to the cells in order to regulate their relative numbers. We proposed to term this problem as *ratio-metric control* of cell populations [11] as its overall goal is to achieve and maintain a certain desired ratio between the size of the populations in the consortium, despite differences in their growth rates, noise and perturbations (Figure 1.c). Examples of external stimuli that can be applied to this aim include changes in the concentration of some inducer molecules in the growth medium or light stimuli applied via optogenetics.

Different solutions to regulate the relative size of coexisting cell populations have been proposed in the literature, mostly based on embedding additional genetic circuitry in the cells that make them able to sense and respond to each other relative size [12–17]. Specifically, by sensing the density of the other group, cells can either increase their growth rate by producing some growth regulator protein (e.g., as in Ecolibrium project, iGEM 2016) or decrease their number by means of toxin-antitoxin mechanisms (e.g., as in [12]). Unfortunately, these embedded solutions cannot avoid the possible extinction of one of the two species, which causes either uncontrolled growth of the survivor species or its death, and are not flexible because the desired steady-state ratio of the two cell populations is hard-coded into the gene regulatory networks designed *ad hoc* and cannot be changed online. Moreover, in industrial applications where high efficiency is required, external control strategies [18–20] could be preferred to more sophisticated embedded solutions,

because additional synthetic genes in the cells can cause lower production rates of the desired chemicals due to excessive metabolic load on the cells.

In this paper, to solve this problem, we consider a microbial consortium composed by *reversible differentiable cells* [21], that is, cells that belong to the same strain and embed a genetic mechanism allowing them to keep memory of past states and adapt their behavior to external stimuli from the environment, for example by activating/deactivating specific set of genes. Specifically, we consider here the simplest case of cells that can switch between two states, mimicking a flip-flop or binary memory element (Figure 1.a). The state of this bistable memory encodes the current role played by the cells in the consortium, and therefore the set of genes they are expressing at that moment. For example, a cell can use its resources either to produce some molecule or to grow and divide sustaining the cell population number [9] (Figure 1.b, top panel). Also, in the case of a genetic pathway divided into two parts, a cell can switch from one state to the other so as to activate either two depending on the overall production levels in the consortium (Figure 1.b, bottom panel). We find that by endowing the cell population with a reversible bistable system, an external control strategy can be used to solve the ratiometric control problem. Specifically, by applying external stimuli to all the cells in the consortium it is possible to switch some cells from one state to the other so as to maintain the desired ratio. We show that this is possible in a number of different ways (namely, by using relay and PI controllers) and provide stability analysis of the resulting closed-loop system and exhaustive *in silico* validation of its performance and robustness. The validation is conducted by means of agent-based simulations in BSim [22, 23], a powerful platform for realistic *in silico* experiments in bacterial populations. As a representative example we consider the realistic agent-based implementation of the proposed ratiometric control strategy to enable cooperative bioproduction in microbial consortia, showing its effectiveness and flexibility when cell growth, cell-to-cell variability and other effects are appropriately modelled.

2 Results

Figure 2: Reversible cells can be switched from state A to B , and vice versa, by means of a common external input. (a) The scalar dynamical system $\dot{x}_i = \eta_i x_i - x_i^3 + u$, with $\eta_i > 0$, has two stable equilibrium points for $u = 0$, namely A_i and B_i , each one corresponding to one of the two possible *roles* the cell can play in the consortium. A cell is *controllable* if by varying the input u in its interval of definition $\mathcal{U}$, it can be moved from one group to the other and vice versa. That is, there exists an admissible value $u' \in \mathcal{U}$ ($u'' \in \mathcal{U}$) such that there is a unique positive (negative) stable solution to the equation $\dot{x}_i = 0$ when $u > u'$ ($u < u''$). Full and empty dots represent stable and saddle equilibria, respectively. (b) Not all cells might respond as desired due to their heterogeneity, captured here by different values of the parameter η_i (assumed to be drawn from some probability distribution with density function ϕ_η , here sketched as Gaussian, just for the sake of illustration). Only cells whose value of the parameter η_i is between 0 and $\hat{\eta}$ are *controllable* (Case 1), that is, they have two stable equilibria for $u = 0$ and a unique stable equilibrium for $u = \pm \bar{u}$. Others cells can either be memory-less, or *monostable*, that is, they have only one equilibrium point for all values of u (Case 2), or can be *unswitchable*, having two stable equilibria for every $u \in \mathcal{U}$, and therefore cannot change role in the consortium (Case 3).

2.1 Reversible differentiable cells can be controlled to a desired state via a common exogenous input

The cells we consider here are assumed to embed some mechanism that can store the memory of past events. In particular, we suppose they can be switched between two different states by appropriate external stimuli, so that their *bistable* behavior can be captured at the macroscopic level by the nonlinear dynamical system (see STAR Methods 7.1 for further details)

$$\dot{x}_i = \eta_i x_i - x_i^3 + u, \quad (1)$$

where $u \in \mathcal{U} \subset \mathbb{R}$ is an input signal and is in *common* to all cells.

In equation (1), $x_i \in \mathcal{X} \subset \mathbb{R}$ represents the macroscopic state, or role, of cell i and the value of parameter $\eta_i \in \mathbb{R}$ is assumed to be different for all cells, accounting for their heterogeneous responses to the common external input signal u . For positive values of η_i , the equation $\dot{x}_i = 0$ with $u = 0$ has two stable, non-trivial solutions, one negative and another positive, that we denote as A_i and B_i , respectively (Figure 2.a). These solutions are the stable equilibrium points of the dynamical system described in (1) when no input is applied. We define as $\mathcal{R}_{A_i} = \{x_i : x_i < 0\}$ and $\mathcal{R}_{B_i} = \{x_i : x_i > 0\}$ the regions of attraction [24, Sec. 8.2] of A_i and B_i , respectively. Each cell will asymptotically converge to either A_i or B_i depending on which region of attraction its initial condition belongs to. Moreover, we denote by $\mathcal{N}_t$ the finite set of all cells in the consortium at time t and with $N(t) = |\mathcal{N}_t|$ its cardinality, that is, the number of cells currently under observation (e.g., via a fluorescence microscope). Note that this number may vary in time as a consequence of cell growth or death, or because of their removal (e.g., flush away) from the culture chamber in which they are hosted. We define the sets $\mathcal{A}_t := \{i \in \mathcal{N}_t : x_i(t) \in \mathcal{R}_{A_i}\}$ and $\mathcal{B}_t := \{i \in \mathcal{N}_t : x_i(t) \in \mathcal{R}_{B_i}\}$, such that $\mathcal{A}_t \cup \mathcal{B}_t = \mathcal{N}_t$ and $\mathcal{A}_t \cap \mathcal{B}_t = \emptyset$, and denote with $n_A(t)$ and $n_B(t)$ their cardinality. These two sets represent the group of cells in the consortium that, at time instant t , in the absence of any control input u are expected to asymptotically converge to A_i and B_i , respectively. Note that, as $\mathcal{A}_t$ and $\mathcal{B}_t$ form a partition of $\mathcal{N}_t$, at any time it holds that $n_A(t) + n_B(t) = N(t)$.

We model cell-to-cell variability by assuming that the parameter η_i in (1) is drawn randomly for each cell from the real interval $[\underline{\eta}, \bar{\eta}]$ with some probability distribution (Figure 2.b). Also, we assume that the control input u is upper bounded by some maximum value $\bar{u} := \max_{u \in \mathcal{U}} |u|$, corresponding, for example, to the highest concentration of the exogenous input that can be provided by the experimental platform. In the presence of such a bound on the control signal, only cells whose parameter value η_i is smaller than the threshold value $\hat{\eta}$ defined as

$$\hat{\eta} = \sqrt[3]{\frac{27\bar{u}^2}{4}}, \quad (2)$$

can be switched from one state to the other by an admissible value of u . We define cells fulfilling the condition $\eta_i \in (0, \hat{\eta})$ as *controllable* cells.

All the cells not fulfilling this condition are instead uncontrollable because they are either *monostable* – their parameters differing so much from their nominal values that their bistable nature is lost – or *unswitchable*

because their parameters exceed the threshold value (2). In the former case η_i will be taking non-positive values in our model, that is, $\eta_i \leq 0$, while in the latter case $\eta_i \geq \hat{\eta}$.

Therefore all cells characterized by parameter values $\eta_i \in (0, \hat{\eta})$, can be switched from one state to the other by means of a common bounded external input u applied into the environment. It is therefore possible to design, for such subset of controllable cells, some feedback control law to automatically regulate their state and keep the balance in the consortium between the two groups $\mathcal{A}_t$ and $\mathcal{B}_t$ to some desired level. As we are going to show next, uncontrollable cells will contribute to a small residual error that can be precisely estimated as a function of the upper bound $\bar{u}$ on the control input and therefore appropriately taken into account in the applications.

Figure 3: Feedback control strategies are effective to balance two groups of controllable reversible cells to 1:1 ratio ($r_d = 0.5$). (a-c) Evolution of the error signal $e(t)$ and of the control input $u(t)$ for (a) the first implementation of the relay controller (22), (b) the second implementation of the relay controller (24), (c) the PI controller (26)-(27). (d-f) Distribution of the cells state at the beginning of the simulation ($t = 0$ a.u., gray histogram) and at steady state ($t = 1$ a.u., green and red histograms), for (d) the first implementation of the relay controller, (e) the second implementation of the relay controller, (f) the PI controller. The green and red bars in panels (d-f) correspond to cells being in the basin of attraction of A_i and B_i , respectively. The maximum control input is set to $\bar{u} = 5$ and the gains of the PI controller are set to $k_P = 30$ and $k_I = 10$. All cells ($N = 400$) have initial conditions $x_i(0)$ drawn from the normal random distribution $\mathcal{N}(0, 4)$, and the parameters η_i are drawn with uniform distribution from the interval $[1, 5]$, therefore all cells are controllable, as no monostable ($\underline{\eta} > 0$) and no unswitchable ($\bar{\eta} < \hat{\eta} \approx 5.53$) cells are present in the population. (See also Supplemental Figure 8 for more simulations with different desired ratios r_d .)

2.2 Ratiometric control of cell populations can be achieved by using external feedback strategies

The goal of ratiometric control is to regulate and maintain the *relative ratios* between the number of cells in $\mathcal{A}_t$ and the number of cells in $\mathcal{B}_t$, defined as

$$r_A(t) = \frac{n_A(t)}{N(t)}, \quad r_B(t) = \frac{n_B(t)}{N(t)}. \quad (3)$$

As by definition, $r_A(t) + r_B(t) = 1$ for all time, it suffices to control either r_A or r_B to control the other. Without loss of generality, we assume the ratio $r(t)$ to be controlled is $r_B(t)$.

More formally, the objective of the *ratiometric control problem* can be stated as follows.

Objective *Given a consortium of reversible differentiable cells whose macroscopic dynamics can be described by (1) and a desired ratio $r_d \in [0, 1]$, design a feedback control law $u = u(t, x)$, where $x = [x_1, \dots, x_{N(t)}]^\top$, such that at steady state the consortium is divided into two cell groups whose ratio converges to some desired value, r_d , that is,*

$$r(t) \rightarrow r_d \text{ as } t \rightarrow \infty. \quad (4)$$

The previous statement can be also reformulated in terms of the control error signal $e(t) := r_d - r(t)$, by requiring that it goes to zero at steady state, that is, $e_\infty = 0$, where $e_\infty := \lim_{t \rightarrow \infty} e(t)$.

Notice that the ratiometric control problem as defined in (4) can only be solved if all cells in the consortium are controllable, as described in Section 2.1.

We present here two different feedback control strategies to solve the ratiometric control problem (see Figure 3), an on-off relay controller and a proportional-integral (PI) controller. Both solutions are easy to implement, robust, and often used in other control applications of cell populations in microfluidic devices [25–28] (see STAR Methods 7.4).

Relay controllers (also known as bang-bang controllers) are simple yet effective feedback control laws that, by comparing the current output of the process of interest with its desired value, generate a piece-wise constant input signal $u_r(t)$ whose value belongs to a discrete set $\mathcal{U}_r$. Here, we propose the use of two alternative implementations of the relay controller, the former where the control input can also be set to zero, i.e., $\mathcal{U}'_r = \{0, \bar{u}, -\bar{u}\}$, and the latter where u_r is always non-zero, i.e., $\mathcal{U}'_r = \{\bar{u}, -\bar{u}\}$.

In the ideal case where all cells are controllable, the first implementation of the relay controller guarantees finite time convergence to zero of the error signal (Figure 3.a). The second implementation instead can only guarantee bounded convergence of the error to zero since the input signal cannot be turned off once the error reaches zero. Hence, when such an implementation is adopted, the control input will continue to oscillate between its possible values (Figure 3.b). As is common practice in applications where noise and uncertainties are unavoidable, a dead-zone or a delay can be added in the control loop to avoid high frequency oscillations of the control input that may cause excessive stress to cells and to the actuation system [29]. The details of the proof of convergence for the proposed relay controllers is reported in STAR Methods 7.5.

An alternative strategy is the use of a *PI controller* that generates a control input $u_{PI}(t)$ computed as the sum of one term proportional (P) to the error $e(t)$ and an other proportional to its integral (I) in time. In general, PI controllers guarantee zero regulation error at steady state in the presence of constant output disturbances [30]. In our implementation, this controller is complemented with a (anti-windup) reset condition that sets to zero the internal state of the integrator whenever the error signal $e(t)$ is equal to 0 or changes its sign (see STAR Methods 7.4). When all cells are controllable, this strategy was also proved to solve the ratiometric control problem and guarantee convergence of the error to zero (STAR Methods 7.5).

The evolution of the error signal $e(t)$ under the action of the PI controller is reported in Figure 3.c. The error converges to zero as expected and the control input $u_{PI}(t)$ also settles to zero after a short transient, similarly to what observed in the first implementation of the relay controller presented before.

Effective balancing of groups of reversible cells is also achieved by feedback control when the goal is to achieve groups of different sizes, that is, for r_d different from 0.5, e.g., equal to 0.75 or 0.25, corresponding to 1:3 and 3:1 ratios, respectively (Supplemental Figure 8).

Figure 4: Balance to 1:1 ratio is achieved with a residual steady-state error in the presence of uncontrollable cells ($r_d = 0.5$). (a-c) Evolution of the error signal $e(t)$ and of the control input $u(t)$ for (a) the first implementation of the relay controller (22), (b) the second implementation of the relay controller (24), (c) the PI controller (26)-(27). (d-f) Distribution of the cells' state at the beginning of the simulation ($t = 0$ a.u., gray histogram) and at steady state ($t = 3.0$ a.u., green and red histograms), for (d) the first implementation of the relay controller, (e) the second implementation of the relay controller, (f) the PI controller. The green and red bars in panels d-f correspond to cells being in the basin of attraction of A_i and B_i , respectively. The maximum control input is set to $\bar{u} = 5$ and the gains of the PI controller are set to $k_P = 30$ and $k_I = 10$. All cells ($N = 400$) have initial conditions $x_i(0)$ drawn from the normal random distribution $\mathcal{N}(-2, 1)$, and the parameters η_i are drawn with uniform distribution from the interval $[-1, 14]$, therefore both monostable ($\bar{\eta} < 0$) and unswitchable ($\bar{\eta} > \hat{\eta} \approx 5.53$) cells can be present in the population. The steady-state errors observed in the *in silico* experiment are equal to (a) 0.05, (b) 0.0225, and (c) 0.0825. Note that all the observed errors are below the theoretical upper bound on the control error that can be estimated using (5) as $e_r = e_r^0 + e_r^u \approx 0.07 + 0.06 = 0.13$ (depicted in the panels a-c as red dashed lines). (See also Supplemental Figure 9 for more simulations with different desired ratios r_d .)

2.3 Robust bounded regulation of the ratio is still possible in the presence of cell variability and physical constraints

When uncontrollable cells are present in the consortium, that is, cells that cannot be moved from one group to the other in response to any admissible inputs, the ratiometric control problem cannot be solved asymptotically, that is, we cannot guarantee that $e_\infty \rightarrow 0$ for an arbitrary initial configuration of the consortium. However, we can still guarantee that the absolute value of the steady-state error $|e_\infty|$ will be upper bounded by some positive quantity e_r , that is, $|e_\infty| \leq e_r$. This effect is well illustrated in Figure 4, where it is shown that, regardless of the control algorithm being used, the error $e(t)$ approaches, but does not converge exactly to, zero. The error bound at steady state will depend upon the interplay between heterogeneity of the cells dynamics, and the constraints on the maximum input value of $\bar{u}$ that can be applied to the cells, as discussed in Section 2.1.

For higher number of cells, the upper bound e_r can be estimated as being composed by two terms, that is,

$$e_r = e_r^0 + e_r^u, \quad (5)$$

each related to the probability of finding one of the two types of uncontrollable cells (i.e., monostable and unswitchable, respectively) in the consortium (Figure 2). The first term, denoted as e_r^0 , is related to the fraction of monostable cells, associated to a non-positive value of η_i , and so admitting only one stable equilibrium point for all values of u . Assuming that the probability distribution from where the parameters η_i are drawn is known, we can estimate e_r^0 as

$$e_r^0 = \Pr[\eta_i \leq 0], \quad (6)$$

where $\Pr$ denotes the probability measure. The second term, denoted as e_r^u , is related to the fraction of unswitchable cells, that is, cells that are bistable but cannot be switched by any admissible value of the control input $u \in \mathcal{U}$. For a given maximum input value $\bar{u}$, this fraction of uncontrollable cells can be estimated as the probability that the parameter η_i is greater than $\hat{\eta}$, that is, $\Pr[\eta_i > \hat{\eta}]$. Therefore, the residual error at steady state due to this second type of uncontrollable cells can be quantified as

$$e_r^u = \max \{0, \Pr[\eta_i > \hat{\eta}] - r_d, \Pr[\eta_i > \hat{\eta}] - (1 - r_d)\}. \quad (7)$$

This second residual error term is evaluated by considering the worst case scenario in which unswitchable cells are initially all in one group. Moreover, its value depends on the relationship between the desired ratio for that group (either r_d or $1 - r_d$) and the number of unswitchable cells therein, because they affect the error only when this number exceeds the desired value (see STAR Methods 7.6 for details).

Similar results are observed also when the desired goal is to split the cell population into groups of different sizes, e.g., with 3:1 or 1:3 ratios (Supplemental Figure 9), confirming that this undesired effect is not due to the particular control strategy adopted or to the chosen desired ratio.

Figure 5: Controlled cooperative bioproduction of a dimer in microfluidic devices. (a) Reversible differentiable cells are hosted in microfluidic chambers where they grow and produce a specific molecule related to the corresponding active state of their internal memory. The current ratio of the two cell groups in the chamber, and hence the production level of the corresponding monomers, is evaluated by measuring with a fluorescence microscope the expression of reporter proteins in each cell. This information is acquired by the feedback control algorithm that compares the current ratio $r(t)$ to the desired ratio r_d and computes online the corresponding control inputs. Finally, these signals are sent to the cells by actuating a pair of syringes connected to the microfluidic chambers and containing mixtures of growth medium and inducer molecules. (b) The required reversible bistable memory mechanism is implemented by using an inducible toggle-switch. Depending on which of the two repressor proteins, either LacI or TetR, is currently expressed, the cell produces the corresponding monomer and the reporter protein (either M1 and RFP or M2 and GFP, respectively). The state of the toggle-switch can be flipped by changing the concentration of the inducer molecules aTc and IPTG in the microfluidic chamber (denoted as u_{aTc} and u_{IPTG}), which diffuse through the cell membrane and bind to TetR and LacI, respectively.

Figure 6: Cooperative production of two monomers to 1:1 population ratio can be achieved by means of feedback ratiometric controllers in microfluidics. (a,d) Evolution in time of populations' ratio r_A (solid green line) and r_B (solid green line) with their respective desired reference values in dashed lines ($r_d = 0.5$), (b,e) of the error signals e_A (solid green line) and e_B (solid red line), and (c,f) inducer control signals u_{aTc} (solid red line) and u_{IPTG} (solid green line), normalized to their maximum values U_{aTc} and U_{IPTG} , respectively. (a-c) Parameters of the relay control (55): $U_{aTc} = 60$ ng/mL, $U_{IPTG} = 0.5$ mM. (d-f) Parameters of the PI controller (56): $U_{aTc} = 100$ ng/mL, $U_{IPTG} = 1$ mM, $k_{P,1} = 100$, $k_{P,2} = 1.5$, $k_{I,1} = 1.5$, $k_{I,2} = 0.05$. Cells (about 200) in the simulated microfluidic chamber (with dimensions $40 \mu\text{m} \times 50 \mu\text{m} \times 1 \mu\text{m}$) have all the same parameters' value, and their evolution has been obtained using the agent-based simulator BSim [22, 23] (See STAR Methods 7.8 for further details on the simulator setup). (See also Supplemental Video 1, and Supplemental Figure 10 for more *in silico* experiments with different desired ratios r_d .)

2.4 Ratiometric control enables cooperative bioproduction in microbial consortia

So far the analysis has been conducted by considering the scalar model (1) capturing the macroscopic bistable nature of the cells considered in this paper. As we are going to show by means of the representative application that follows, the behavior captured by the reduced model in (1) is qualitatively preserved also in more complex and realistic cell models exhibiting the required memory-like property. Therefore, albeit simple, we demonstrate that the model in (1) can be effectively used to design feedback control laws to solve the ratiometric control problem in realistic applications.

As a representative case of study, we consider the agent-based *in silico* implementation of ratiometric control for the bioproduction of protein dimers in microfluidic devices (Figure 5). In this scenario, according to its state (A or B), each cell in the consortium produces either one of two monomers. By acting on the available control inputs, we want to regulate the relative number of cells producing the two monomers so as to balance the overall production of the resulting dimer.

We assume that the mechanism required by the *E.coli* cells to guarantee their correct coordinated behavior is implemented by means of an inducible genetic toggle-switch [31]. Specifically, we consider the circuits design presented in [27] and further analyzed in [32–35]. This genetic regulatory network consists of two repressor proteins, LacI and TetR, both repressing each other’s promoter, so that only one protein is fully expressed at any time. The expression level of the two repressor proteins can be flipped by changing the concentration of two inducer molecules, aTc and IPTG. The former input, aTc, binds to TetR, increasing the rate of production of LacI, and therefore causing the cell to converge to the steady-state corresponding to high expression of LacI. Analogously, IPTG binds to LacI, causing the commitment of the cell to a steady-state corresponding to high expression of TetR.

The dynamical model of each cell is described in details in STAR Methods 7.7, in which the variables u_{aTc} and u_{IPTG} (as reported in Figure 5) denote the concentrations of the inducer molecules in the growth medium of the microfluidic chambers and they represent the control inputs that can be applied to all cells to change their production role in the consortium.

We further assume that the genes $m1$ and $m2$ encoding the two monomers of interest are each fused to one of the repressor genes $lacI$ and $tetR$ of the toggle-switch circuit. So that, at steady state, each cell fully produces only one monomer at the time and at a rate assumed to be proportional to the concentration of the corresponding repressor protein. Reporter genes of red and green fluorescent proteins (RFP and GFP) are also bound to the repressor genes to monitor the current level of production of the monomers by using fluorescence microscopy (Figure 5). Finally, we assume that the two monomers have equal transcription and translation rates. Therefore, for the dimer to be produced at high rate, the consortium must be split and maintained into two symmetric groups with a 1:1 ratio, that is, we set $r_d = 0.5$.

In the *in silico* experiments, we also take into account realistic physical and technological constraints of a possible implementation in the microfluidic experimental platform described in [11, 25]. Specifically, we

consider constraints on (i) the possible classes of input signals that can be generated by the actuators, (ii) an upper bound on the switching frequency of the inputs to limit osmotic stress to the cells, (iii) a time delay accounting for the time the chemical inducers take to flow from the reservoirs to the cell chambers, and (iv) a safety lower bound on the sampling time of the measurements to avoid excessive photo-toxicity (see STAR Methods 7.8).

By using a timescale separation argument (see STAR Methods 7.7), it is possible to capture the cell dynamics using the following reduced-order model

$$z_1' = k_1^0 + \frac{k_1}{1 + \left(z_2 \cdot \frac{1}{1+w_1}\right)^{\eta_{\text{TetR}}}} - z_1 \quad (8)$$

$$z_2' = k_2^0 + \frac{k_2}{1 + \left(z_1 \cdot \frac{1}{1+w_2}\right)^{\eta_{\text{LacI}}}} - z_2 \quad (9)$$

where the symbol (') denotes the derivative with respect to rescaled time t' ,

$$z_1 := \frac{[\text{LacI}]}{\theta_{\text{LacI}}}, \quad z_2 := \frac{[\text{TetR}]}{\theta_{\text{TetR}}}, \quad (10)$$

and

$$w_1 := \left(\frac{[\text{aTc}]}{\theta_{\text{aTc}}}\right)^{\eta_{\text{aTc}}}, \quad w_2 := \left(\frac{[\text{IPTG}]}{\theta_{\text{IPTG}}}\right)^{\eta_{\text{IPTG}}}, \quad (11)$$

where $[\cdot]$ denotes concentrations.

It is possible to verify that system (8)-(9) is memory-like as required, since in its domain of definition the phase portrait is qualitatively similar to that of the system in (14a) (Figure 7.a). Moreover, this dynamical system has also been numerically investigated in [27, Supplementary Figure 6] where it has been shown that its dynamics exhibit bistable behavior with hysteresis.

In silico control experiments have been conducted by using *ad hoc* implementations in BSim [22, 23] of the two feedback control algorithms presented in Section 2.2 (see STAR Methods 7.9). To test the relay control strategy, we assumed that the actuation of the inputs is realized using an ordinary T-junction [36], which allows only one inducer molecule at a time to be injected into the microfluidic chambers. For the PI controller we assumed that the actuation is realized by a Dial-A-Wave system, as described in [37]. This actuation system is more advanced than the previous one as it allows mixtures of the two inducers to be injected in different proportions into the chambers. Both feedback control algorithms take into account the characteristics of the experimental platform and in particular of the actuators. Full details about the control algorithms and the technological constraints of the platform are reported in STAR Methods 7.9.

The agent-based simulations in BSim accurately capture cells' reproduction, spatial distribution and geometry of the cells and of the microfluidic chambers, diffusion of chemicals into the environment and, more importantly, flush-out of the cells from the chambers. Further details on the stochastic simulation algorithm, geometry and other parameters used for *in silico* experiments in BSim are reported in STAR Methods 7.8.

We observed that both controllers can successfully regulate the populations' ratio to the desired value after relatively short transients (Figure 6 and Supplemental Video 1). The relay controller shows a faster response

with more severe oscillations (Figure 6.a-c), while the PI controller presents a smoother but slower response with higher accuracy at steady state (Figure 6.d-f). This is expected as it is well-known that the relay control strategy is in general more robust to uncertainties and noise affecting the controlled process but has poorer accuracy at steady state; the PI control strategy showing better steady-state performance thanks to the presence of an integral action.

The difference in the performance of the two strategies is also stressed by the different actuation systems employed in our experiments. Indeed, the Dial-A-Wave system allows for a finer regulation of the concentrations of the inducer molecules than the simpler (and cheaper) T-junction, allowing better accuracy of the control system at steady state. Similar performances are obtained also when the goal is changed to achieve different population ratios, e.g., a 1:3 ratio or a 3:1 ratio (Supplemental Figure 10). These scenarios may correspond, for example, to situations in which the two monomers have equal transcription rates but different translation rates, requiring the consortium to be split into two asymmetric groups, for efficient production of the dimer.

Besides biochemical noise, the fluctuations at steady state (Figure 6.a,d) are essentially due to cells being flushed out of the microfluidic chamber as they grow and duplicate. This effect is more relevant when cells are host in a small chamber and becomes less significant as the size of the growth chambers increases. Indeed, for the sake of simplicity, assuming the chamber to be a square of side ℓ , the magnitude ε of the fluctuations is proportional to $1/\ell$ (see STAR Methods 7.10), hence the fluctuations increase as the chamber size decreases (see Supplemental Figure 11 and Supplemental Video 2), and vice versa. Finally, the ratiometric controllers presented here also showed robustness to cell-to-cell variability in their response, as shown in Supplemental Figure 12.

3 Discussion

We presented a general framework to guide the design of external feedback controllers for labor division in microbial consortia. We showed that, by exploiting the memory-like property of reversible differentiable cells, a single-strain cell population can be divided into two groups, expressing different sets of genes, whose relative numbers, i.e., the ratio, can be regulated by means of common exogenous inputs. We showed by means of a representative example that ratiometric feedback controllers can robustly stabilize a cell population, endowed with a genetic toggle-switch functioning as a bistable memory, and can guarantee the balance between the two groups of cells even in the presence of realistic physical and technological constraints of the experimental microfluidic platform we considered.

A fundamental open problem in multicellular control applications is to guarantee coexistence of different microbial strains growing in the same environment. Although some solutions were proposed in the literature that rely on the use of additional genetic pathways embedded into the cells, the ratiometric control framework we presented here provides an alternative approach that might be more appropriate in other scenarios, for example, in industrial applications, where efficient production is strongly required. However, the scaling up of

the ratiometric control strategies we proposed here for the balancing of microbial communities in bioreactors is still not fully investigated; in previous work [19] we only considered the problem of regulating the relative number of two *different* microbial strains, while the case of a monostrain consortium composed by reversible cells has not yet been explored and is left for future work.

Finally, we wish to highlight that ratiometric control of a population of reversible cells by means of a common input signal is only made possible by the heterogeneity of their response to that input. Indeed, heterogeneous reversible cells characterized by different parameter values switch at different time instants when subject to the same input, and this is a crucial property that allows their state to be controlled by an external feedback action. Stochastic effects, such as biochemical noise or delays, by amplifying the cell-to-cell variability can indeed facilitate the stabilization of the reversible cells into different groups, as it was demonstrated in the *in silico* experiments we provided. A more in-depth analytical investigation of their effect is left for future work.

4 Acknowledgments

The authors wish to thank Prof. S. John Hogan, University of Bristol, U.K. for his insightful comments on the convergence of discontinuous maps, and also Dr. Lorena Postiglione, University of Bristol, U.K. for her advice about experimental microfluidic platforms.

M.d.B. and D.F. wish to acknowledge support from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 766840 (COSY-BIO). D.F. also wishes to acknowledge support from the research grant “BIOMASS” funded by the University of Naples Federico II - “Finanziamento della Ricerca di Ateneo (FRA) - Linea B”.

5 Author Contributions

D.F. and M.d.B. designed the research; D.F. and D.S. designed the controllers and carried out the modeling; D.S. implemented the controllers in BSim and carried out the *in silico* experiments; D.S. and D.F. developed the analysis; M.d.B. supervised the project; D.S., D.F. and M.d.B. analyzed the data and wrote the manuscript.

6 Declaration of Interests

The authors declare no competing interests.

7 STAR Methods

Lead contact and materials

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Mario di Bernardo (mario.dibernardo@unina.it).

Methods Details

7.1 Modeling reversible differentiable cells

We assume that the macroscopic behavior of any cell in the consortium we wish to control can be modeled in the domain of interest by a dynamical system of the form

$$\dot{z} = g(z, w), \quad (12)$$

with $g : \mathcal{Z} \times \mathcal{W} \mapsto \mathcal{Z}$ being a smooth vector field, $z \in \mathcal{Z} \subset \mathbb{R}^d$ the state variables and $w \in \mathcal{W} \subset \mathbb{R}^m$ the exogenous input variables, representing, for instance, the concentrations of chemical species inside the cell and those of control inducer molecules into the environment, respectively. We assume that stochastic effects, such as fluctuations due to biochemical reactions, do not significantly alter the behavior of the system at steady state; that is, the region of attraction of any stable equilibrium point of the dynamical system $\dot{z} = g(z, w)$ is large enough so that stochastic noise does not cause undesired switch from one equilibrium point to another.

As discussed in the Introduction and in Section 2.1, we are interested to robustly regulate the behavior of reversible cells, in particular we focus our attention to a specific class of cells whose dynamics satisfy the following fundamental property (Supplemental Figure 7.a).

Definition Consider a dynamical system of the form

$$\dot{z} = f(z, \alpha), \quad (13)$$

where $z \in \mathcal{Z}$ and the parameter $\alpha \in \mathcal{I} \subset \mathbb{R}$ depends on some exogenous input signal $w \in \mathcal{W}$, that is, $\alpha = \alpha(w)$. We say that system (13) is memory-like if for all $\alpha \in \mathcal{I}$ it is globally topologically equivalent, in its domain of definition $\mathcal{Z}$, to the dynamical system

$$\begin{cases} \dot{\xi} = \beta_1(\alpha) + \beta_2(\alpha)\xi - \xi^3, \\ \dot{\zeta} = -\zeta, \end{cases} \quad (14a)$$

$$(14b)$$

with $(\xi, \zeta) \in \mathcal{D} \subset \mathbb{R} \times \mathbb{R}^{d-1}$, β_1 and β_2 being in bijective correspondence with α , that is, there exists a homeomorphism $h_\alpha : \mathcal{Z} \rightarrow \mathcal{D}$ mapping orbits of the first system onto orbits of the second system, preserving the direction of time. Moreover, β_1 and β_2 are monotonic functions with respect to α , and $\beta_1(0) = 0$.

As state variables ξ and ζ in system (14a)-(14b) are uncoupled and $\zeta(t)$ asymptotically converges to 0 for all initial condition $\zeta(0)$, the qualitative behavior of the system can be completely described by the dynamics of ξ . Note that, (14a) is the topological normal form of the so-called cusp bifurcation [38, p.301], which can exhibit hysteretic behavior.

Under the assumption that system (12) is memory-like, hence being topological equivalent to system (14a), and by further assuming that (i) the value of $\beta_{i,1}$ is the same across the entire population, i.e., $\beta_{i,1} = u$, $\forall i$, and that (ii) $\beta_{i,2}$ is constant but different for each cell i , i.e., $\beta_{i,2} = \eta_i$, any reversible cell in the consortium can be modeled as in (1), where u is an *equivalent input* signal that depends on w . However, it can be showed that the previous assumptions, albeit simple, lead to more conservative results for the original model (13).

Indeed, by defining $\bar{\beta}_{i,2} := \sup_{\alpha} \beta_{i,2}(\alpha)$ and $\underline{\beta}_{i,2} := \inf_{\alpha} \beta_{i,2}(\alpha)$, a cell i whose dynamics is described by model (13), is controllable (see Section 2.1) if (i) $\underline{\beta}_{i,2} > 0$, and (ii) $\exists \bar{\alpha}$ such that $\beta_{i,1}^2(\alpha) > \frac{4}{27} \bar{\beta}_{i,2}^3$, for all $|\alpha| > \bar{\alpha}$. The first condition implies that the cell is not monostable, that is, the system has two stable equilibrium points for $\alpha = 0$. On the other hand, having assumed that $\beta_{i,1}$ and $\beta_{i,2}$ are monotonic functions with respect to α , the second condition implies that the cell is controllable, that is, there exists a value of the input α such that the bistable system becomes monostable (see STAR Methods 7.2).

Finally, by choosing the parameter η_i in model (1) as

$$\eta_i = \begin{cases} \inf_{\alpha} \beta_{i,2}(\alpha), & \text{if } \inf_{\alpha} \beta_{i,2}(\alpha) \leq 0 \\ \sup_{\alpha} \beta_{i,2}(\alpha), & \text{if } \inf_{\alpha} \beta_{i,2}(\alpha) > 0 \end{cases} \quad (15)$$

and by setting $\bar{u} = \sup_{i,\alpha} |\beta_{i,1}(\alpha)|$, the cells described by model (1) that are controllable are indeed a smaller subset of those described by model (13), that is, by making this choice for the parameters we are considering a worst-case scenario. Thus, all the results presented for model (1), included the proof of convergence of the proposed controllers, also hold for model (13).

7.2 Derivation of $\hat{\eta}$

In this section we show how the expression of $\hat{\eta}$ in (2) was derived. The quantity $\hat{\eta}$ represents the limit value of η_i such that the equation $\dot{x}_i = 0$ with $\eta_i > 0$ and $u = \bar{u}$, where $\bar{u} = \max_{u \in \mathcal{U}} |u|$, has a unique solution. That is, any cell whose behavior can be described by (1) with $\eta_i > \hat{\eta}$, cannot be switched from A_i to B_i , or vice versa, by any input value $u \in \mathcal{U}$, because it always admits at least two solutions and it is therefore uncontrollable.

To this aim, we have to characterize the bifurcation points at which two equilibrium points of (1) coalesce and disappear [38], under variations of the control input u and of the parameters η_i . The position of the limit points can be easily obtained as:

$$(x_{LP}^+, u_{LP}^+) = \left(\sqrt{\frac{\eta_i}{3}}, -\sqrt{\frac{4\eta_i^3}{27}} \right) \quad (16)$$

$$(x_{LP}^-, u_{LP}^-) = -(x_{LP}^+, u_{LP}^+) \quad (17)$$

Hence, equation $\dot{x}_i = 0$ admits at least two solutions if its parameters and the control input satisfy the relationship

$$-\sqrt{\frac{4\eta_i^3}{27}} < u < \sqrt{\frac{4\eta_i^3}{27}} \quad (18)$$

Therefore, it is easy to verify from (18) that, given the maximum value of the control input $\bar{u}$, the threshold value $\hat{\eta}$ over which a cell becomes uncontrollable is

$$\hat{\eta} := \left(\frac{27\bar{u}^2}{4} \right)^{\frac{1}{3}} \quad (19)$$

7.3 Event-driven modeling of the control error evolution

Here we derive an event-driven model for the evolution of the control error $e(t) = r_d - r(t)$, presented in Section 2.2. Recall that the finite set of all cells in the consortium and its cardinality at time t are denoted by $\mathcal{N}_t$ and $N(t)$, respectively. Now, we denote with $E_{A \rightarrow B}^i(t')$ the event at time instant $t = t'$, corresponding to when the state x_i of cell $i \in \mathcal{N}_t$ enters the basin of attraction of the equilibrium point B_i , that is, $x_i(t) \in \mathcal{R}_{B_i}$ for $t \in [t', t' + \varepsilon]$, and $x_i(t) \notin \mathcal{R}_{B_i}$ for $t \in [t' - \varepsilon, t')$, where ε is a small positive real number. Likewise, we denote by $E_{B \rightarrow A}^i(t'')$ the event at time instant $t = t''$, corresponding to when the state of cell $i \in \mathcal{N}_t$ enters the basin of attraction of A_i . Specifically, for solutions to dynamical system (1) with $u = 0$ and $\eta_i > 0$, an event $E_{A \rightarrow B}^i$ occurs when the state of cell i , x_i , crosses zero and becomes positive, while an event $E_{B \rightarrow A}^i$ occurs when x_i becomes negative. In this case the threshold at zero is defined by the unstable equilibrium point at the origin dividing the regions of attraction of A_i and B_i . Moreover, we denote by $\mathcal{E}_{A \rightarrow B}$ ($\mathcal{E}_{B \rightarrow A}$) the sets of all events $E_{A \rightarrow B}^i$ ($E_{B \rightarrow A}^i$) occurring for all i at any time t , and we denote by $\mathcal{E}$ the set of all events occurring in the population, that is, $\mathcal{E} := \mathcal{E}_{A \rightarrow B} \cup \mathcal{E}_{B \rightarrow A}$.

To derive the discrete (event-driven) dynamics of the control error $e(t)$, we make the following standing assumptions:

Assumption 1 *At any time t only one event in $\mathcal{E}$ can occur, that is, there exists a unique $i \in \mathcal{N}_t$ such that either $E_{A \rightarrow B}^i$ or $E_{B \rightarrow A}^i$, but not both, occurs at time t .*

Assumption 2 *The number of cells in the host chamber is assumed to be constant, that is, $N(t) = N$, for all t .*

Assumption 1 implies that any two events in $\mathcal{E}$ cannot occur simultaneously. Assumption 2 follows from the fact that, after a short transient from the beginning of the experiment, cells grow occupying the entire host chamber. From this time on, cells exceeding the maximum capacity of the chamber are flushed out. Therefore, the number of cells in the chamber can be assumed with a good approximation to be constant (except for a small, negligible oscillation due to flush-out, further discussed in STAR Methods 7.10).

Then, there exists a sequence of discrete time instants $\{t_k\}_{k \in \mathbb{N}}$, each one corresponding to the occurrence of an event in $\mathcal{E}$ and such that $t_{k+1} = t_k + \Delta t_k$, where $\Delta t_k > 0$ is the time interval between two consecutive events occurring at time t_k and t_{k+1} , respectively. Moreover, from Assumption 2 it follows that the functions $n_A(t)$ and $n_B(t)$, defining the number of cells converging to either A_i or B_i , respectively, are piece-wise constant functions, that is, $n_A(t) = n_A(t_k)$, and $n_B(t) = n_B(t_k)$, $\forall t \in [t_k, t_{k+1})$. Since $n_A(t)$ and $n_B(t)$ are constrained by the relation $n_A(t) + n_B(t) = N$, for all t , for the sake of brevity, we will refer to $n(t) := n_B(t)$ only; $n_A(t)$ being given by $N - n_B(t)$.

That said, being $n(t)$ the number of cells in the basin of attraction of B_i at time t , we can write the following discrete-time update law:

$$n(t_{k+1}) = \begin{cases} n(t_k) + 1, & \text{if an event in } \mathcal{E}_{A \rightarrow B} \text{ occurs} \\ n(t_k) - 1, & \text{if an event in } \mathcal{E}_{B \rightarrow A} \text{ occurs} \end{cases} \quad (20)$$

As a consequence, since $e(t) = r_d - \frac{n(t)}{N}$, we have that when an event in $\mathcal{E}_{A \rightarrow B}$ occurs, then

$$e(t_{k+1}) = r_d - \frac{n(t_{k+1})}{N} = \left(r_d - \frac{n(t_k)}{N} \right) - \frac{1}{N} = e(t_k) - \frac{1}{N}.$$

A similar reasoning holds when an event in $\mathcal{E}_{B \rightarrow A}$ occurs. Therefore, the discrete-time dynamics of the control error can be expressed as:

$$e(t_{k+1}) = \begin{cases} e(t_k) - \frac{1}{N}, & \text{if an event in } \mathcal{E}_{A \rightarrow B} \text{ occurs} \\ e(t_k) + \frac{1}{N}, & \text{if an event in } \mathcal{E}_{B \rightarrow A} \text{ occurs} \end{cases} \quad (21)$$

7.4 Implementation details of the proposed feedback control algorithms

Here we discuss the design of the control strategies proposed to solve the ratiometric control problem, namely, the relay controller and the PI controller.

Relay control algorithm. A *relay controller* is a feedback control law that generates a piece-wise constant input $u_r(t)$ by comparing an output measured from the plant to some desired reference value r_d . The input $u_r(t)$ takes value from a finite set of real values $\mathcal{U}_r$, generally composed by only two values, one chosen such that the control error $e(t) = r_d - r(t)$ decreases when $e(t) > 0$, and the other such that it increases when $e(t) < 0$.

We have considered two implementations of the relay controller. Specifically, the first implementation includes a control shutdown condition when $e(t) = 0$, and the second one does not.

Formally, the *first* implementation of the relay control input is defined as:

$$u'_r(t) = \begin{cases} \bar{u}, & e(t) > 0 \\ 0, & e(t) = 0 \\ -\bar{u}, & e(t) < 0 \end{cases} \quad (22)$$

The value $\bar{u} > 0$ is chosen such that (ideally) all cells are controllable, that is, for $u'_r = \bar{u}$ ($u'_r = -\bar{u}$), the equation $\eta_i x_i - x_i^3 + u'_r = 0$ has a unique stable solution, namely $\bar{x}_i = B_i(\bar{u})$ ($\bar{x}_i = A_i(\bar{u})$), for all i such that $\eta_i > 0$. This guarantees that when $e(t) > 0$ (i.e., $r_B(t) < r_d$), the next event occurring must belong to $\mathcal{E}_{A \rightarrow B}$, forcing the error to decrease, that is, $e(t_{k+1}) = e(t_k) - \frac{1}{N}$. Likewise, when $e(t) < 0$ (i.e., $r_B(t) > r_d$), $u'_r = -\bar{u}$ implies that the next event occurring belongs to $\mathcal{E}_{B \rightarrow A}$, and so $e(t_{k+1}) = e(t_k) + \frac{1}{N}$. Moreover, notice that, when no control is applied (i.e., $u'_r = 0$), each cell will converge to either A_i or B_i , depending on its current state, without any other event in $\mathcal{E}$ having to occur. Therefore, the shutdown condition ensures that, if there

exists a t^* such that $e(t^*) = 0$, then $e(t) = 0$ for all $t \geq t^*$. Combining (22) and (21), the discrete-time update law for the control error becomes:

$$e(t_{k+1}) = \begin{cases} e(t_k) - \frac{1}{N}, & \text{if } e(t_k) > 0 \\ e(t_k), & \text{if } e(t_k) = 0 \\ e(t_k) + \frac{1}{N}, & \text{if } e(t_k) < 0 \end{cases} \quad (23)$$

The previous discrete map can also be rewritten as $e(t_{k+1}) = e(t_k) - \frac{1}{N} \text{sgn}(e(t_k))$.

The *second* implementation of the relay control input, without the shutdown condition, is defined as:

$$u_r''(t) = \begin{cases} \bar{u}, & e(t) \geq 0 \\ -\bar{u}, & e(t) < 0 \end{cases} \quad (24)$$

In this case it is not possible to ensure that the error remains equal to zero indefinitely. By using a similar reasoning as before, (21) can be recast as:

$$e(t_{k+1}) = \begin{cases} e(t_k) - \frac{1}{N}, & \text{if } e(t_k) > 0 \\ -\frac{1}{N}, & \text{if } e(t_k) = 0 \\ e(t_k) + \frac{1}{N}, & \text{if } e(t_k) < 0 \end{cases} \quad (25)$$

PI control algorithm. The control input of the PI controller is defined as:

$$u_{\text{PI}}(t) = k_P e(t) + k_I z(t), \quad (26)$$

$$\dot{z}(t) = e(t), \quad z(0) = 0, \quad (27)$$

with k_P and k_I being positive constants. This control action is complemented with a (anti-windup) reset condition that sets the internal state z of the integrator to zero whenever the error becomes 0 or changes its sign. Furthermore, to take into account constraints on the actuation system of the experimental platform, the control input signal u_{PI} is assumed to be saturated at $\bar{u}$ and $-\bar{u}$.

The control algorithm guarantees that when $e(t) > 0$ the control input $u_{\text{PI}}(t)$ is positive and $\frac{d}{dt}u_{\text{PI}}(t) > 0$. So, $u_{\text{PI}}(t)$ will increase and reach some positive value $\hat{u}$ such that, for at least one cell, the equation $\eta_i x_i - x_i^3 + u = 0$ with $u = \hat{u}$ and $\eta_i > 0$ has a unique solution, namely $B_i(\hat{u})$ (see Figure 7). The cell will be attracted by this stable equilibrium point and, therefore, there will exist a time instant t' such that, for all $t \geq t'$, $u_{\text{PI}}(t) \geq \hat{u}$ and an event in $\mathcal{E}_{A \rightarrow B}$ will occur. A similar reasoning holds in the case $e(t) < 0$. Hence, it directly follows that the discrete-time update law for the control error $e(t)$ under the PI control law is the same as in (23).

7.5 Proof of convergence of the proposed controllers

Here we prove that the feedback control algorithms we proposed ensure convergence to the desired cell ratio and hence solve the ratiometric control problem stated in Section 2.2. To this aim we conduct a convergence analysis on the error map $e_{k+1} = f(e_k)$, where $f : D \rightarrow D$, with f being defined in (23) or (25), and D is the domain of definition of the error signal e_k . Specifically, we consider some positive definite scalar function

$V : D \rightarrow \mathbb{R}$, and we prove convergence of e_k to some invariant set by assessing the negative definiteness of the increment function $\Delta V := V(f(e_k)) - V(e_k)$ for every point outside this invariant set.

Notice that the domain of definition D is *finite*, indeed, from the definition of the control error signal, that is, $e(t) := r_d - r(t) = \frac{1}{N}(n_d - n(t))$, and from the fact that $n_d = N \cdot r_d$ is an integer constant and $n(t)$ is integer-valued function taking values in the finite set $\{0, 1, 2, \dots, N\}$, we have that

$$D := \left\{ -1, \dots, -\frac{2}{N}, -\frac{1}{N}, 0, \frac{1}{N}, \frac{2}{N}, \dots, 1 \right\}. \quad (28)$$

Relay control algorithm (22). To prove convergence of the error signal e to 0, let us consider the function $V(e) = e^2$. Using the discrete-time map in (23), we have that

$$\Delta V = V_{k+1} - V_k = e_{k+1}^2 - e_k^2 = e_k^2 - \frac{2}{N}|e_k| + \frac{1}{N^2} - e_k^2 = -\frac{1}{N^2}(2N|e_k| - 1). \quad (29)$$

The increment ΔV is negative definite everywhere except in the points such that

$$|e_k| \leq 1/(2N), \quad (30)$$

therefore any solution whose initial condition is outside this set converges towards it. Since the only possible value of e_k in D satisfying (30) is 0, we can therefore conclude that the error signal e_k converges to 0.

Relay control algorithm (24). From (25) it directly follows that the error dynamics does not admit any fixed point. However, a period-2 cycle exists whose stability we are going to prove by considering the second iterate of the error map in (25):

$$e_{k+2} = \begin{cases} e_k - \frac{2}{N}, & \text{if } e_k \geq \frac{1}{N} \\ e_k, & \text{if } e_k \in \{0, -\frac{1}{N}\} \\ e_k + \frac{2}{N}, & \text{if } e_k \leq -\frac{2}{N} \end{cases} \quad (31)$$

The periodic orbit of the first iterate map (25) corresponds to the pair of fixed points located in 0 and $-1/N$ of the second iterate map (31).

Let us consider the function $V(e) = |e|$. By using the second iterated map in (31) and taking into account (28) we have:

$$V(e_{k+2}) = \begin{cases} e_k - \frac{2}{N}, & \text{if } e_k \geq \frac{2}{N} \\ \frac{1}{N}, & \text{if } e_k = \frac{1}{N} \\ -e_k, & \text{if } e_k \in \{0, -\frac{1}{N}\} \\ -e_k - \frac{2}{N}, & \text{if } e_k \leq -\frac{2}{N} \end{cases} \quad (32)$$

Hence, after some algebra we can write:

$$\Delta V = V(e_{k+2}) - V(e_k) = \begin{cases} -\frac{2}{N}, & \text{if } |e_k| \geq \frac{2}{N} \\ 0, & \text{if } e_k \in \{0, -\frac{1}{N}, \frac{1}{N}\} \end{cases} \quad (33)$$

Therefore, we can conclude that all solutions converge to the set where $\Delta V = 0$, that is, such that $|e_k| \leq 1/N$, or, equivalently, to the set $\{-\frac{1}{N}, 0, \frac{1}{N}\}$. Now, by using the second iterate map (31), we can easily verify

that in the set $\{-\frac{1}{N}, 0, \frac{1}{N}\}$ there are two fixed points, namely $-\frac{1}{N}$ and 0. Therefore, these two fixed points correspond to period-2 cycles for the first iterated map (25). More precisely, both points correspond to the same cycle $M = \{-\frac{1}{N}, 0\}$, indeed, by applying (25) we have that $-\frac{1}{N} \mapsto 0$ and $0 \mapsto -\frac{1}{N}$. This proves that the error signal globally converges to the period-2 cycle M .

PI control algorithm (26)-(27). In this case the discrete-time map describing the dynamics of the control error is also defined by (23). Therefore, the proof of convergence for the PI controller is the same as that of the first implementation of the relay controller (22) presented earlier on.

7.6 Derivation of e_r^u

Here we illustrate how the expression of e_r^u in (7) was obtained. The quantity e_r^u is an upper bound on the residual steady-state error due to the presence of *unswitchable* cells, i.e., bistable cells that cannot be toggled by any admissible value of the control input u . The maximum value of this quantity can be estimated by considering the worst case scenarios in which all these uncontrollable cells are initially in one region of attraction and therefore cannot be moved to the other one by any admissible values of u .

Without loss of generality, assume that all uncontrollable cells are initially in the region of attraction of their equilibrium point B_i and recall that r_d denotes the desired ratio of cells converging to B_i at steady state. Denoting with r_u the ratio between the number of uncontrollable cells and the whole cell population, we have that if $r_u \leq r_d$, that is, the number of uncontrollable cells does not exceed the desired number in B_i , then uncontrollable cells do not cause any problem because we do not need to move any of them from where they are. Therefore, in this case, $e_\infty = 0$. On the other hand, if $r_u > r_d$, the fraction of uncontrollable cells exceeding the desired ratio, that is, $r_u - r_d$, will determine an error at steady state, that is, $|e_\infty| \leq r_u - r_d$. A similar argument can be stated for the dual scenario in which all uncontrollable cells are initially in the region of attraction of A_i . Recalling that in this case we want the ratio of cells converging to A_i to be equal to $(1 - r_d)$, the residual error at steady state in this scenario will satisfy $|e_\infty| \leq r_u - (1 - r_d)$.

From the previous analysis and exploiting that, for a high number of cells, r_u can be estimated as the probability of η_i being greater than $\hat{\eta}$, we finally obtain that

$$e_r^u = \max \{0, \Pr[\eta_i > \hat{\eta}] - r_d, \Pr[\eta_i > \hat{\eta}] - (1 - r_d)\}. \quad (34)$$

where $\Pr$ denotes the probability measure.

7.7 Mathematical model of the inducible toggle-switch

In this section we present the details of the mathematical model used to describe the time evolution of proteins' concentration in the inducible toggle-switch. The model of the toggle-switch we consider here was originally developed and parameterized from experimental data in [27]. It was further analyzed in [32–35]

The model captures the pseudo-reactions describing transcription

where

$$h_L(TetR, aTc) := \kappa_L^{m0} + \frac{\kappa_L^m}{1 + \left(\frac{[TetR]}{\theta_{TetR}} \cdot \frac{1}{1 + \left(\frac{[aTc]}{\theta_{aTc}} \right)^{\eta_{aTc}}} \right)^{\eta_{TetR}}}, \quad (37)$$

$$h_T(LacI, IPTG) := \kappa_T^{m0} + \frac{\kappa_T^m}{1 + \left(\frac{[LacI]}{\theta_{LacI}} \cdot \frac{1}{1 + \left(\frac{[IPTG]}{\theta_{IPTG}} \right)^{\eta_{IPTG}}} \right)^{\eta_{LacI}}}, \quad (38)$$

those describing translation

and those related to dilution/degradation

The repressive action of LacI and TetR on each other is captured in (37)-(38) by means of Hill functions (with parameters θ s and η s). Likewise, the repressions of aTc on TetR and of IPTG on LacI are both described by another Hill function nested in the previous two. In the above pseudo-reactions, the parameters $\kappa_{L/T}^{m0}$, $\kappa_{L/T}^m$, $\kappa_{L/T}^P$, $g_{L/T}^m$, $g_{L/T}^p$ are basal transcription, transcription, translation, mRNA degradation and protein degradation rates, respectively. The value of these parameters are reported in Table 1.

The pseudo-reactions listed above can be used to obtain the following *deterministic model* of the inducible toggle-switch:

$$\frac{d}{dt} [mRNA_{LacI}] = \kappa_L^{m0} + \frac{\kappa_L^m}{1 + \left(\frac{[TetR]}{\theta_{TetR}} \cdot \frac{1}{1 + \left(\frac{[aTc]}{\theta_{aTc}} \right)^{\eta_{aTc}}} \right)^{\eta_{TetR}}} - g_L^m [mRNA_{LacI}] \quad (45)$$

$$\frac{d}{dt} [mRNA_{TetR}] = \kappa_T^{m0} + \frac{\kappa_T^m}{1 + \left(\frac{[LacI]}{\theta_{LacI}} \cdot \frac{1}{1 + \left(\frac{[IPTG]}{\theta_{IPTG}} \right)^{\eta_{IPTG}}} \right)^{\eta_{LacI}}} - g_T^m [mRNA_{TetR}] \quad (46)$$

$$\frac{d}{dt} [LacI] = \kappa_L^P [mRNA_{LacI}] - g_L^p [LacI] \quad (47)$$

$$\frac{d}{dt} [TetR] = \kappa_T^P [mRNA_{TetR}] - g_T^p [TetR] \quad (48)$$

where the quantities in the form $[\cdot]$ denote concentrations of molecules inside each cell.

κ_L^{m0}	0.3045 mRNA min ⁻¹	g_L^p	0.0165 min ⁻¹
κ_T^{m0}	0.3313 mRNA min ⁻¹	g_T^p	0.0165 min ⁻¹
κ_L^m	13.01 mRNA min ⁻¹	θ_{LacI}	124.9
κ_T^m	5.055 mRNA min ⁻¹	η_{LacI}	2.00
κ_L^p	0.6606 a.u. mRNA ⁻¹ min ⁻¹	θ_{TetR}	76.40
κ_T^p	0.5098 a.u. mRNA ⁻¹ min ⁻¹	η_{TetR}	2.152
k_{aTc}	0.04 min ⁻¹	θ_{aTc}	35.98
k_{IPTG}	0.04 min ⁻¹	η_{aTc}	2.00
g_L^m	0.1386 min ⁻¹	θ_{IPTG}	0.2926
g_T^m	0.1386 min ⁻¹	η_{IPTG}	2.00

Table 1: Values of the parameters of the cell population models (taken from [27]).

The model is complemented by also taking into account the diffusion of aTc and IPTG across the cell membrane:

$$\frac{d}{dt} [aTc] = k_{aTc} (u_{aTc} - [aTc]), \quad (49)$$

$$\frac{d}{dt} [IPTG] = k_{IPTG} (u_{IPTG} - [IPTG]), \quad (50)$$

where the parameters k_{aTc}/k_{IPTG} are diffusion rates, whose values are reported in Table 1.

The deterministic model in (45)-(48) can be simplified by using a separation of time scales argument, as done in [32, 35]. Firstly, we can observe that equations (45) and (46) can be rewritten as

$$\begin{aligned} \frac{d}{dt} [mRNA_{LacI}] &= \kappa_L^{m0} + \frac{\kappa_L^m}{1 + \left(\frac{[TetR]}{\theta_{TetR}} \cdot \frac{1}{1+w_1} \right)^{\eta_{TetR}}} - g_L^m \cdot [mRNA_{LacI}] \\ \frac{d}{dt} [mRNA_{TetR}] &= \kappa_T^{m0} + \frac{\kappa_T^m}{1 + \left(\frac{[LacI]}{\theta_{LacI}} \cdot \frac{1}{1+w_2} \right)^{\eta_{LacI}}} - g_T^m \cdot [mRNA_{TetR}] \end{aligned}$$

where $w_1 = \left(\frac{[aTc]}{\theta_{aTc}} \right)^{\eta_{aTc}}$ and $w_2 = \left(\frac{[IPTG]}{\theta_{IPTG}} \right)^{\eta_{IPTG}}$. Now, taking into account that mRNA molecules are degraded faster than other molecules, we can obtain a quasi-steady state approximation by setting $\frac{d}{dt} [mRNA_{LacI}] = 0$ and $\frac{d}{dt} [mRNA_{TetR}] = 0$, yielding

$$\begin{aligned} [mRNA_{LacI}] &= \frac{\kappa_L^{m0}}{g_L^m} + \frac{\kappa_L^m}{g_L^m} \frac{1}{1 + \left(\frac{[TetR]}{\theta_{TetR}} \cdot \frac{1}{1+w_1} \right)^{\eta_{TetR}}} \\ [mRNA_{TetR}] &= \frac{\kappa_T^{m0}}{g_T^m} + \frac{\kappa_T^m}{g_T^m} \frac{1}{1 + \left(\frac{[LacI]}{\theta_{LacI}} \cdot \frac{1}{1+w_2} \right)^{\eta_{LacI}}} \end{aligned}$$

Assuming that LacI and TetR proteins degrade at the same rate, that is $g_L^p = g_T^p = g^p$, and using the dimensionless state variables $z_1 = \frac{[LacI]}{\theta_{LacI}}$ and $z_2 = \frac{[TetR]}{\theta_{TetR}}$, and dimensionless time $t' = g^p t$, we obtain from

(47):

$$\begin{aligned}\frac{dz_1}{dt'} &= \frac{\kappa_L^p}{\theta_{LacI} g^p} [mRNA_{LacI}] - z_1 \\ &= \frac{\kappa_L^p}{\theta_{LacI} g^p} \left(\frac{\kappa_L^{m0}}{g_L^m} + \frac{\kappa_L^m}{g_L^m} \frac{1}{1 + \left(z_2 \cdot \frac{1}{1+w_1} \right)^{\eta_{TetR}}} \right) - z_1 \\ &= k_1^0 + \frac{k_1}{1 + \left(z_2 \cdot \frac{1}{1+w_1} \right)^{\eta_{TetR}}} - z_1\end{aligned}$$

where

$$k_1^0 = \frac{\kappa_L^{m0} \kappa_L^p}{g_L^m \theta_{LacI} g^p}, \quad k_1 = \frac{\kappa_L^m \kappa_L^p}{g_L^m \theta_{LacI} g^p}, \quad (51)$$

and from (48):

$$\begin{aligned}\frac{dz_2}{dt'} &= \frac{\kappa_T^p}{\theta_{TetR} g^p} [mRNA_{TetR}] - z_2 \\ &= \frac{\kappa_T^p}{\theta_{TetR} g^p} \left(\frac{\kappa_T^{m0}}{g_T^m} + \frac{\kappa_T^m}{g_T^m} \frac{1}{1 + \left(z_1 \cdot \frac{1}{1+w_2} \right)^{\eta_{LacI}}} \right) - z_2 \\ &= k_2^0 + \frac{k_2}{1 + \left(z_1 \cdot \frac{1}{1+w_2} \right)^{\eta_{LacI}}} - z_2\end{aligned}$$

where

$$k_2^0 = \frac{\kappa_T^{m0} \kappa_T^p}{g_T^m \theta_{TetR} g^p}, \quad k_2 = \frac{\kappa_T^m \kappa_T^p}{g_T^m \theta_{TetR} g^p}. \quad (52)$$

7.8 Agent-based simulations in BSim

BSim is an agent-based simulator explicitly designed for the simulation of bacterial populations [22, 23]. In this environment the chemicals' spatial distribution and the bio-mechanics of the cells are also considered. Specifically, it is possible to mimic a microfluidic platform, where cells grow and move, and where chemicals diffuse into the environment.

The BSim implementation presented in [23] has been extended here to include also stochastic simulations of the biochemical processes taking place inside the cells. Specifically, we implemented the SDE-based algorithm presented in [39]. Formally, for each cell we solved a system of equations of the form

$$d\mathbf{x} = \mathbf{S} \cdot \mathbf{F}(\mathbf{x}, t)dt + \mathbf{S} \cdot \text{diag}(\mathbf{F}(\mathbf{x}, t))d\mathbf{w} + \Delta(\mathbf{x})dt, \quad (53)$$

where $\mathbf{x} = [mRNA_{LacI}, mRNA_{TetR}, [LacI], [TetR], [aTc], [IPTG]]^\top$ is the state vector, $\mathbf{w}$ is a vector of independent standard Wiener processes, $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_6]$ is the stoichiometry matrix and $\mathbf{F}(\mathbf{x}) = [a_1(\mathbf{x}), \dots, a_6(\mathbf{x})]^\top$, where $a_j(\mathbf{x})$ is the propensity function of the j -th pseudo-reaction defined as in STAR Methods 7.7. Notice that the term $\Delta(\mathbf{x})dt$ models the diffusion dynamics of chemicals across the cell membrane, assumed to be deterministic.

The microfluidic platform we considered consists of a microfluidic device, a microscopy, a computer and an actuation system. The microfluidic device contains a chamber where cells are trapped and grow. Specifically,

we simulated the microfluidic chip described in [40], where the growth chambers' dimensions are $40\mu\text{m} \times 30\mu\text{m} \times 1\mu\text{m}$. This allows bacteria to grow only in a mono-layer structure. The chamber is connected to two perfusion channels, which bring nutrients and inducers to the cellular population. The flow of the fluids sent to the chambers is governed by two syringes, each one containing a mixture of growth media and one inducer. By adjusting the relative height of the syringes, it is possible to change the amount of inducers delivered to the cells. The architecture is completed by a computer, controlling the syringes depending on the output of a control algorithm, and a microscopy, necessary to take images whose analysis results in the computation of the error.

The bio-mechanical parameters of the cells (i.e., cells' growth rate) have been set as in [23].

Moreover, we have taken into account the following realistic constraints on the experimental platform [29] :

1. the state of the cells cannot be measured more often than 5 min, to avoid excessive photo-toxicity;
2. there is a time delay of 40s on the actuation of the control inputs due to the time that the flow of the chemical inducers takes to reach the chambers in the microfluidic chip where cells are hosted;
3. the minimum time interval between two consecutive control inputs cannot be smaller than 15 min, to limit excessive osmotic stress on the cells;
4. the maximum duration of any experiment cannot exceed 24 hours (1440 min), to avoid substantial cell mutations during the experiments.

The specific implementation of the microfluidic device we adopted also introduces constraints on the possible classes of input signals $u(t) = [u_{\text{aTc}}(t), u_{\text{IPTG}}(t)]^\top$ that can be generated by the actuators. We consider two possible implementations:

1. a T-junction, which limits u_{aTc} and u_{IPTG} to be mutually exclusive and with fixed amplitudes, that is, u is either equals to $[U_{\text{aTc}}, 0]^\top$, or to $[0, U_{\text{IPTG}}]^\top$;
2. a Dial-A-Wave (DAW) system [37], which constraints u_{aTc} and u_{IPTG} to be in a convex combination. Namely, given $u_{\text{aTc}} \in [0, U_{\text{aTc}}]$ we have

$$u_{\text{IPTG}} = \left(1 - \frac{u_{\text{aTc}}}{U_{\text{aTc}}}\right) U_{\text{IPTG}} \quad (54)$$

In the above equations, U_{aTc} and U_{IPTG} denote the concentrations of the inducers in the reservoirs, and we assume that they can be varied at the beginning of the experiments in the intervals $[0, 100]$ and $[0, 1]$, respectively. The upper bounds for U_{aTc} and U_{IPTG} are selected to avoid excessive stress on cells and are the same as those used *in vivo* in [27].

7.9 Realistic *in silico* implementation of the controllers

To meet the specific dynamical features of the genetic toggle-switch implementing the required bistable memory mechanism in reversible differentiable cells (Section 2.4), we propose modified versions of the ideal relay controller and the PI controller proposed in Section 2.2.

Due to higher dimensional dynamics and unavoidable uncertainties affecting the system, it is not possible to define precise boundaries between the regions of attraction of the equilibrium points A_i and B_i for each cell. For this reason, we define conservative estimation of the regions of attraction common to all cells. Specifically, we defined $\mathcal{A}_t := \{i \in \mathcal{N}_t : 2 \text{LacI}_i(t) < \text{TetR}_i(t)\}$ and $\mathcal{B}_t := \{i \in \mathcal{N}_t : \text{LacI}_i(t) > 2 \text{TetR}_i(t)\}$, where LacI_i and TetR_i denote the concentrations of the molecules inside the i -th cell. Moreover, since in reality the union of these two sets does not cover the entire set $\mathcal{N}_t$ of cells in the consortium, cells not belonging to $\mathcal{A}_t$ and $\mathcal{B}_t$ are in an *uncertain* state. We define this uncertain set of cells as $\mathcal{C}_t := \{i \in \mathcal{N}_t : i \notin \mathcal{A}_t, i \notin \mathcal{B}_t\}$ and denote with $n_C(t) = |\mathcal{C}_t|$ its cardinality at time t . For this reason, the error signals e_A and e_B are not complementary (i.e., $e_A(t) \neq -e_B(t)$) and we need to consider both in the design of the control actions.

Relay controller. The relay controller consists of two mutually exclusive inputs with fixed amplitudes, which are applied to the system depending on the current value of the error signals $e_A(t)$ and $e_B(t)$. Specifically, at any time t , it is applied the input that causes $\max\{|e_A(t)|, |e_B(t)|\}$ to decrease. More formally, the control input $u(t) = [u_{\text{aTc}}(t), u_{\text{IPTG}}(t)]^\top$ is chosen as

$$u(t) = \begin{cases} u_1, & |e_B(t)| \geq |e_A(t)| \\ u_2, & |e_B(t)| < |e_A(t)| \end{cases}, \quad (55)$$

where

$$u_1 = \begin{cases} [0, U_{\text{IPTG}}]^\top, & e_B(t) \leq 0 \\ [U_{\text{aTc}}, 0]^\top, & e_B(t) > 0 \end{cases}, \quad u_2 = \begin{cases} [U_{\text{aTc}}, 0]^\top, & e_A(t) \leq 0 \\ [0, U_{\text{IPTG}}]^\top, & e_A(t) > 0 \end{cases}.$$

In this way, when $e_B > 0$ ($e_B \leq 0$) a high value of aTc (IPTG) is applied, that, by promoting LacI (TetR), causes e_B to decrease (increase). Similar arguments hold for e_A .

The control parameters U_{aTc} and U_{IPTG} were tuned empirically by trial-and-error. The numerical values used in each simulation are reported in the caption of the relative figure.

PI controller. The PI controller is implemented using a Dial-A-Wave system [37]. It consists of two control actions, each acting to decrease either of the control errors. Formally, we have:

$$u_{\text{aTc}}(t) = k_{\text{P},1}e_B(t) + k_{\text{I},1} \int_0^t e_B(\tau)d\tau - \left(k_{\text{P},2}e_A(t) + k_{\text{I},2} \int_0^t e_A(\tau)d\tau \right) \quad (56)$$

where $k_{\text{P},j}$ and $k_{\text{I},j}$, $j = 1, 2$, are the proportional and integral control gains. Note that, due to the constraints on the actuation system, u_{IPTG} is defined by (54).

The proportional and the integral gains, and the maximum concentrations of the inducers U_{aTc} and U_{IPTG} are empirically tuned by trial-and-error. The numerical values used in each simulation are reported in the caption of the relative figure.

Moreover, to improve performance, the control algorithm is complemented with an anti-windup scheme (that sets to zero the internal state of the integrator whenever the error signals are equal to 0 or change sign) and

a dynamic saturation defined as:

$$\begin{cases} u_{aTc} \in [0, 50], & \text{if } |e_B| < |e_A| \\ u_{aTc} \in [0, 100], & \text{otherwise} \end{cases} \quad (57)$$

Similarly for u_{IPTG} as defined by (54).

7.10 Noise magnitude due to flush-out effects

The fluctuations in the steady-state response of the closed-loop system (see for example Figure 6) are mostly due to cells flowing out of the microfluidic chamber as they grow. The magnitude ε of this noise is indeed proportional to the ratio between the flow rate Q of cells exiting the chamber and the total number N of cells therein. That is, $\varepsilon \propto Q/N$, where $Q = v \cdot A$; being A the cross-section area of the chamber apertures and v the flow velocity, that depends on the cell growth rate. Without loss of generality, assume that, as is usual in microfluidic applications, cells in the chamber are organized in a single-layer and that, on average, the area occupied by a cell can be approximated by a circle of radius r . Moreover, let us assume that the chamber has a square geometry with side ℓ with two open sides, as in Figure 5. Then, the flow rate Q can be approximated as $Q \approx v \cdot \frac{2\ell}{2r}$. Moreover, the total number of cells in the chamber is $N = \ell^2/(\pi r^2)$. Therefore, in conclusion, we have that

$$\varepsilon \propto \frac{Q}{N} = \frac{v \cdot \frac{\ell}{r}}{\frac{\ell^2}{\pi r^2}} = v \cdot \pi r \cdot \frac{1}{\ell},$$

that is, the magnitude of the noise is inversely proportional to the size of the culture chamber. We validated the formula numerically by comparing the noise levels between chambers of different sizes (compare Figure 6 and Supplemental Figure 11). Specifically, we observed that, as expected, the average variation of the ratios at steady state in the chamber $40\mu m \times 50\mu m$ wide (0.009) was three times lower than the one in a chamber where both the dimensions were scaled down to a third (0.03) (see also Supplemental Video 2).

Note that when the cells under observation flow out of the field of view from all the four sides of the chamber (a common situation when higher magnification factors are used on the microscope), the same result still holds but multiplied by a factor of 2.

Supplemental Videos

- Supplemental Video 1: Animations of the *in silico* control experiments in BSim reported in Figure 6.
- Supplemental Video 2: Animations of the *in silico* control experiments in BSim reported in Supplemental Figure 11, with a smaller growth chamber.

Data and code availability

The code used for all simulations is available at <https://github.com/diBernardoGroup/Ratiometric-Control-of-bacterial-populations>

8 Supporting Information

Figure 7: **Macroscopic dynamics of reversible differentiable cells with memory-like property can be effectively modeled as a scalar bistable dynamical system.** The dynamical system $\dot{z} = f(z, \alpha)$ describing the evolution of a reversible differentiable cell with *memory-like* property is topologically equivalent to the system $\dot{\xi} = \beta_1(\alpha) + \beta_2(\alpha)\xi - \xi^3$, $\dot{\zeta} = -\zeta$. That is, there exists a correspondence between solutions of the two dynamical systems, in particular they have the same number of equilibrium points of the same stability type. Moreover, this equivalence is preserved as the bifurcation parameter α varies, to which there correspond some values of the parameters $\beta_1(\alpha)$ and $\beta_2(\alpha)$ in the second system.

Figure 8: Feedback control strategies are effective to balance two *asymmetric* groups of controllable reversible cells to 1:3 ratio ($r_d = 0.75$) and 3:1 ratio ($r_d = 0.25$). (a-c, g-i) Evolution of the error signal $e(t)$ and of the control input $u(t)$ for (a, g) the first implementation of the relay controller (22), (b, h) the second implementation of the relay controller (24), (c, i) the PI controller (26)-(27). (d-f, j-l) Distribution of the cells state at the beginning of the simulation ($t = 0$ a.u., gray histogram) and at steady state ($t = 1.0$ a.u., green and red histograms), for (d, j) the first implementation of the relay controller, (e, k) the second implementation of the relay controller, (f, l) the PI controller. Desired ratios are set to (a-f) $r_d = 0.75$ (1:3 ratio) and (g-l) $r_d = 0.25$ (3:1 ratio). The green and red bars in panels d-f and j-l correspond to cells in the basin of attraction of A_i and B_i , respectively. The maximum control input is set to $\bar{u} = 5$ and the gains of the PI controller has been set to $k_P = 30$ and $k_I = 10$. All cells ($N = 400$) have initial conditions $x_i(0)$ drawn from the normal random distribution $\mathcal{N}(0, 4)$, and the parameters η_i are drawn with uniform distribution from the interval $[1, 5]$, therefore all cells are controllable, as no monostable ($\bar{\eta} > 0$) and no unswitchable ($\bar{\eta} < \hat{\eta} \approx 5.53$) cells are present in the population.

Figure 10: **Cooperative production of two monomers to *asymmetric* population ratios can be achieved by means of feedback ratiometric controllers in microfluidics.** (a, d, g, j) Evolution in time of populations' ratio r_A (solid green line) and r_B (solid green line) with their respective desired reference values in dashed lines, (b, e, h, k) of the error signals e_A (solid green line) and e_B (solid red line), and (c, f, i, l) inducer control signals u_{aTc} (solid red line) and u_{IPTG} (solid green line), normalized to their maximum values U_{aTc} and U_{IPTG} , respectively. Desired ratios are set to (a-f) $r_d = 0.75$ (1:3 ratio) and (g-l) $r_d = 0.25$ (3:1 ratio). (a-c, g-i) Parameters of the relay control (55): $U_{aTc} = 60$ ng/mL, $U_{IPTG} = 0.5$ mM. (d-f, j-l) Parameters of the PI controller (56): $U_{aTc} = 60$ ng/mL, $U_{IPTG} = 0.5$ mM, $k_{P,1} = 60$, $k_{P,2} = 0.75$, $k_{I,1} = 1.5$, $k_{I,2} = 0.05$. Cells (about 200) in the simulated microfluidic chamber (with dimensions $40 \mu\text{m} \times 50 \mu\text{m} \times 1 \mu\text{m}$) have all the same parameters' value, and their evolution has been obtained using the agent-based simulator BSim [22, 23].

Figure 11: **Cooperative production of two monomers to 1:1 population ratio can still be achieved in smaller growth chambers.** Larger noise fluctuations due to cells flowing out of the microfluidic chambers are expected in this situation. (a,d) Evolution in time of populations' ratio r_A (solid green line) and r_B (solid green line) with their respective desired reference values in dashed lines ($r_d = 0.5$), (b,e) of the error signals e_A (solid green line) and e_B (solid red line), and (c,f) inducer control signals u_{aTc} (solid red line) and u_{IPTG} (solid green line), normalized to their maximum values U_{aTc} and U_{IPTG} , respectively. (a-c) Parameters of the relay control (55): $U_{aTc} = 60$ ng/mL, $U_{IPTG} = 0.5$ mM. (d-f) Parameters of the PI controller (56): $U_{aTc} = 100$ ng/mL, $U_{IPTG} = 1$ mM, $k_{P,1} = 100$, $k_{P,2} = 1.5$, $k_{I,1} = 1.5$, $k_{I,2} = 0.05$. The simulated microfluidic chamber (with dimensions $13.3 \mu\text{m} \times 16.7 \mu\text{m} \times 1 \mu\text{m}$) can host a cell population of about 30 cells. All simulations were obtained using the agent-based simulator BSim [22,23]. (See also Supplemental Video 2.)

Figure 12: **Cooperative production of two monomers to 1:1 population ratio can be achieved in a heterogeneous cell population by means of feedback ratiometric controllers in microfluidics.** (a,d) Evolution in time of populations' ratio r_A (solid green line) and r_B (solid green line) with their respective desired reference values in dashed lines ($r_d = 0.5$), (b,e) of the error signals e_A (solid green line) and e_B (solid red line), and (c,f) inducer control signals u_{aTc} (solid red line) and u_{IPTG} (solid green line), normalized to their maximum values U_{aTc} and U_{IPTG} , respectively. (a-c) Parameters of the relay control (55): $U_{aTc} = 60$ ng/mL, $U_{IPTG} = 0.5$ mM. (d-f) Parameters of the PI controller (56): $U_{aTc} = 100$ ng/mL, $U_{IPTG} = 1$ mM, $k_{P,1} = 100$, $k_{P,2} = 1.5$, $k_{I,1} = 1.5$, $k_{I,2} = 0.05$. The simulated microfluidic chamber (with dimensions $40 \mu\text{m} \times 50 \mu\text{m} \times 1 \mu\text{m}$) can host a cell population of about 200 cells. Each time a cell divides, the parameters of the daughter are drawn from a normal distribution centered in the nominal value, and with variance equal to its 20%. All simulations were obtained using the agent-based simulator BSim [22,23].

References

- [1] Domitilla Del Vecchio, Yili Qian, Richard M Murray, and Eduardo D Sontag. Future systems and control research in synthetic biology. *Annual Reviews in Control*, 45:5–17, 2018.
- [2] Jan Roelof Van Der Meer and Shimshon Belkin. Where microbiology meets microengineering: design and applications of reporter bacteria. *Nature Reviews Microbiology*, 8:511–522, 2010.
- [3] Marco Mauri, Jean-Luc Gouzé, Hidde De Jong, and Eugenio Cinquemani. Enhanced production of heterologous proteins by a synthetic microbial community: Conditions and trade-offs. *PLoS Computational Biology*, 16(4):e1007795, 2020.
- [4] Danielle B. Pedrolli, Nathan V. Ribeiro, Patrick N. Squizzato, et al. Engineering microbial living therapeutics: The synthetic biology toolbox. *Trends in Biotechnology*, 37(1):100 – 115, 2019.
- [5] Domitilla Del Vecchio, Alexander J Ninfa, and Eduardo D Sontag. Modular cell biology: retroactivity and insulation. *Molecular Systems Biology*, 4(161), 2008.
- [6] Philip Bittihn, M Omar Din, Lev S Tsimring, and Jeff Hasty. Rational engineering of synthetic microbial systems: from single cells to consortia. *Current Opinion in Microbiology*, 45:92–99, 2018.
- [7] Gianfranco Fiore, Antoni Matyjaszkiewicz, Fabio Annunziata, Claire Grierson, Nigel J Savery, Lucia Marucci, and Mario di Bernardo. In-silico analysis and implementation of a multicellular feedback control strategy in a synthetic bacterial consortium. *ACS Synthetic Biology*, 6(3):507–517, 2016.
- [8] Davide Fiore, Davide Salzano, Enric Cristòbal-Cóppulo, Josep M Olm, and Mario di Bernardo. Multicellular feedback control of a genetic toggle-switch in microbial consortia. *IEEE Control Systems Letters*, 5(1):151–156, 2021.
- [9] Elise Weill-Duflos, Virgile Andreani, Chetan Aditya, Pierre Martinon, Gregory Batt, Frederic Bonnans, and Jakob Ruess. Optimal control of an artificial microbial differentiation system for protein bioproduction. In *Proc. of the 2019 European Control Conference*, pages 2663–2668, 2019.
- [10] Mehdi Sadeghpour, Alan Veliz-Cuba, Gábor Orosz, Krešimir Josić, and Matthew R Bennett. Bistability and oscillations in co-repressive synthetic microbial consortia. *Quantitative Biology*, 5(1):55–66, 2017.
- [11] Davide Salzano, Davide Fiore, and Mario di Bernardo. Ratiometric control for differentiation of cell populations endowed with synthetic toggle switches. In *Proc. of the 58th IEEE Conference on Decision and Control*, pages 927–932, 2019.
- [12] Xinying Ren, Ania-Ariadna Baetica, Anandh Swaminathan, and Richard M Murray. Population regulation in microbial consortia using dual feedback control. *Proc. of the 56th IEEE Conference on Decision and Control*, pages 5341–5347, 2017.
- [13] Spencer R Scott, M Omar Din, Philip Bittihn, Liyang Xiong, Lev S Tsimring, and Jeff Hasty. A stabilized microbial ecosystem of self-limiting bacteria using synthetic quorum-regulated lysis. *Nature Microbiology*, 2(17083), 2017.

- [14] Kristina Stephens, Maria Pozo, Chen-Yu Tsao, Pricila Hauk, and William E Bentley. Bacterial co-culture with cell signaling translator and growth controller modules for autonomously regulated culture composition. *Nature Communications*, 10(4129), 2019.
- [15] Alex JH Fedorec, Behzad D Karkaria, Michael Sulu, and Chris P Barnes. Single strain control of microbial consortia. *Nature Communications*, 12(1977), 2021.
- [16] Razan N Alnahhas, Mehdi Sadeghpour, Ye Chen, Alexis A Frey, William Ott, Krešimir Josić, and Matthew R Bennett. Majority sensing in synthetic microbial consortia. *Nature Communications*, 11(3659), 2020.
- [17] Virginia Fusco, Davide Salzano, Davide Fiore, and Mario di Bernardo. Embedded control of cell growth using tunable genetic systems. *bioRxiv*, 2021.
- [18] Patrick De Leenheer and Hal Smith. Feedback control for chemostat models. *Journal of Mathematical Biology*, 46(1):48–70, 2003.
- [19] Davide Fiore, Fabio Della Rossa, Agostino Guarino, and Mario di Bernardo. Feedback ratiometric control of two microbial populations in a single chemostat. *IEEE Control Systems Letters*, pages 1–1, 2021.
- [20] Alex dos Reis de Souza, Denis Efimov, Andrey Polyakov, and Jean-Luc Gouzé. Observer-based robust control of a continuous bioreactor with heterogeneous community. *IFAC-PapersOnLine*, 53(2):11800–11805, 2020.
- [21] Kevin D Litcofsky, Raffi B Afeyan, Russell J Krom, Ahmad S Khalil, and James J Collins. Iterative plug-and-play methodology for constructing and modifying synthetic gene networks. *Nature Methods*, 9:1077–1080, 2012.
- [22] Thomas E Gorochowski, Antoni Matyjaszkiewicz, Thomas Todd, et al. BSim: An agent-based tool for modeling bacterial populations in systems and synthetic biology. *PLoS One*, 7(8):e42790, 2012.
- [23] Antoni Matyjaszkiewicz, Gianfranco Fiore, Fabio Annunziata, Claire S Grierson, Nigel J Savery, Lucia Marucci, and Mario di Bernardo. BSim 2.0: an advanced agent-based cell simulator. *ACS Synthetic Biology*, 6(10):1969–1972, 2017.
- [24] Hassan K Khalil and Jessy W Grizzle. *Nonlinear systems*, volume 3. Prentice hall Upper Saddle River, NJ, 2002.
- [25] Filippo Menolascina, Gianfranco Fiore, Emanuele Orabona, Luca De Stefano, Mike Ferry, Jeff Hasty, Mario di Bernardo, and Diego di Bernardo. In-vivo real-time control of protein expression from endogenous and synthetic gene networks. *PLoS Computational Biology*, 10(5):e1003625, 2014.
- [26] Gianfranco Fiore, Giansimone Perrino, Mario di Bernardo, and Diego di Bernardo. In vivo real-time control of gene expression: a comparative analysis of feedback control strategies in yeast. *ACS Synthetic Biology*, 5(2):154–162, 2015.

- [27] Jean-Baptiste Lugagne, Sebastián Sosa Carrillo, Melanie Kirch, Agnes Köhler, Gregory Batt, and Pascal Hersen. Balancing a genetic toggle switch by real-time feedback control and periodic forcing. *Nature Communications*, 8(1671), 2017.
- [28] Giansimone Perrino, Sara Napolitano, Francesca Galdi, Antonella La Regina, Davide Fiore, Teresa Giuliano, Mario di Bernardo, and Diego di Bernardo. Automatic synchronisation of the cell cycle in budding yeast through closed-loop feedback control. *Nature Communications*, 12(2452), 2021.
- [29] Barbara Shannon, Criseida G Zamora-Chimal, Lorena Postiglione, Davide Salzano, Claire S Grierson, Lucia Marucci, Nigel J Savery, and Mario Di Bernardo. In vivo feedback control of an antithetic molecular-titration motif in escherichia coli using microfluidics. *ACS Synthetic Biology*, 9(10):2617–2624, 2020.
- [30] Karl Johan Aström and Richard M Murray. *Feedback systems: an introduction for scientists and engineers*. Princeton University Press, 2010.
- [31] Timothy S Gardner, Charles R Cantor, and James J Collins. Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767):339, 2000.
- [32] Davide Fiore, Agostino Guarino, and Mario di Bernardo. Analysis and control of genetic toggle switches subject to periodic multi-input stimulation. *IEEE Control Systems Letters*, 3(2):278–283, 2019.
- [33] Agostino Guarino, Davide Fiore, Davide Salzano, and Mario di Bernardo. Balancing cell populations endowed with a synthetic toggle switch via adaptive pulsatile feedback control. *ACS Synthetic Biology*, 9(4):793–803, 2020.
- [34] Agostino Guarino, Davide Fiore, and Mario di Bernardo. In-silico feedback control of a MIMO synthetic Toggle Switch via Pulse-Width Modulation. In *Proc. of the 2019 European Control Conference*, pages 680–685, 2019.
- [35] Alejandro F Villaverde, Nikolaos Tsiantis, and Julio R Banga. Full observability and estimation of unknown inputs, states and parameters of nonlinear biological models. *Journal of the Royal Society Interface*, 16(156):20190043, 2019.
- [36] Octavio Mondragón-Palomino, Tal Danino, Jangir Selimkhanov, Lev Tsimring, and Jeff Hasty. Entrainment of a population of synthetic genetic oscillators. *Science*, 333(6047):1315–1319, 2011.
- [37] Mike S Ferry, Ivan A Razinkov, and Jeff Hasty. Microfluidics for synthetic biology: from design to execution. In *Methods in Enzymology*, volume 497, pages 295–372. Elsevier, 2011.
- [38] Yuri A Kuznetsov. *Elements of applied bifurcation theory*. Springer-Verlag, 1998.
- [39] Eszter Lakatos. *Stochastic analysis and control methods for molecular cell biology*. PhD thesis, Imperial College London, 2017.
- [40] Tal Danino, Octavio Mondragón-Palomino, Lev Tsimring, and Jeff Hasty. A synchronized quorum of genetic clocks. *Nature*, 463(7279):326, 2010.