Supporting Information: "Balancing cell populations endowed with a synthetic toggle switch via adaptive pulsatile feedback control"

Agostino Guarino,[†] Davide Fiore,[‡] Davide Salzano,[†] and Mario di Bernardo^{†,¶}

†Department of Electrical Engineering and Information Technology, University of Naples Federico II, Italy ‡Department of Mathematics and Applications "R. Caccioppoli", University of Naples

Federico II, Italy

¶Department of Engineering Mathematics, University of Bristol, UK

Mathematical models

Pseudo-reactions and deterministic model

The model of the synthetic toggle switch we considered in our analysis was originally developed in Lugagne et al.^{S1} The model captures the pseudo-reactions describing transcription

$$\begin{split} & \emptyset \xrightarrow{f_{\rm L}^{\rm m}(TetR, aTc)} mRNA_{\rm LacI}, \\ & \emptyset \xrightarrow{f_{\rm T}^{\rm m}(LacI, IPTG)} mRNA_{\rm TetR}, \end{split}$$

those describing translation

$$mRNA_{\text{LacI}} \xrightarrow{\kappa_{\text{L}}^{\text{p}}} mRNA_{\text{LacI}} + LacI,$$
$$mRNA_{\text{TetR}} \xrightarrow{\kappa_{\text{T}}^{\text{p}}} mRNA_{\text{TetR}} + TetR,$$

and those related to dilution/degradation

$$\begin{split} mRNA_{\text{LacI}} & \xrightarrow{g_{\text{L}}^{\text{m}}} \emptyset, \quad mRNA_{\text{TetR}} \xrightarrow{g_{\text{T}}^{\text{m}}} \emptyset, \\ LacI & \xrightarrow{g_{\text{L}}^{\text{p}}} \emptyset, \quad TetR \xrightarrow{g_{\text{T}}^{\text{p}}} \emptyset. \end{split}$$

In the above equations, $f_{\rm L}(TetR, aTc)$ and $f_{\rm T}(LacI, IPTG)$ are the gene regulation functions defined as:

$$f_{\rm L}(TetR, aTc) := \kappa_{\rm L}^{\rm m0} + \kappa_{\rm L}^{\rm m} \cdot h^{-}(TetR \cdot h^{-}(aTc, \theta_{\rm aTc}, \eta_{\rm aTc}), \theta_{\rm TetR}, \eta_{\rm TetR}),$$

$$f_{\rm T}(LacI, IPTG) := \kappa_{\rm T}^{\rm m0} + \kappa_{\rm T}^{\rm m} \cdot h^{-}(LacI \cdot h^{-}(IPTG, \theta_{\rm IPTG}, \eta_{\rm IPTG}), \theta_{\rm LacI}, \eta_{\rm LacI}),$$

the parameters $\kappa_{L/T}^{m0}$, $\kappa_{L/T}^{m}$, $\kappa_{L/T}^{p}$, $g_{L/T}^{m}$, $g_{L/T}^{p}$ are leakage transcription, transcription, translation, mRNA degradation, and protein degradation rates, respectively, and $h^{-}(x, \theta, \eta) = 1/(1 + (x/\theta)^{\eta})$ represents a decreasing Hill function.

The pseudo-reactions listed above can be put together to obtain the following *deterministic model* of the toggle switch dynamics:

$$\frac{d\,mRNA_{\text{LacI}}}{dt} = \kappa_{\text{L}}^{\text{m0}} + \frac{\kappa_{\text{L}}^{\text{m}}}{1 + \left(\frac{TetR}{\theta_{\text{TetR}}} \cdot \frac{1}{1 + (aTc/\theta_{\text{aTC}})^{\eta_{\text{aTc}}}}\right)^{\eta_{\text{TetR}}}} - g_{\text{L}}^{\text{m}} \cdot mRNA_{\text{LacI}} \tag{1}$$

$$\frac{d\,mRNA_{\text{TetR}}}{dt} = \kappa_{\text{T}}^{\text{m0}} + \frac{\kappa_{\text{T}}^{\text{m}}}{1 + \left(\frac{LacI}{\theta_{\text{LacI}}} \cdot \frac{1}{1 + (IPTG/\theta_{\text{IPTG}})^{\eta_{\text{IPTG}}}}\right)^{\eta_{\text{LacI}}} - g_{\text{T}}^{\text{m}} \cdot mRNA_{\text{TetR}}$$
(2)

$$\frac{dLacI}{dt} = \kappa_{\rm L}^{\rm p} \cdot mRNA_{\rm LacI} - g_{\rm L}^{\rm p} \cdot LacI \tag{3}$$

$$\frac{dTetR}{dt} = \kappa_{\rm T}^{\rm p} \cdot mRNA_{\rm TetR} - g_{\rm T}^{\rm p} \cdot TetR$$
(4)

The model is completed by considering the diffusion dynamics of the inducer molecules, aTc and IPTG, across the cells' membranes with a non-symmetrical exchange dynamics^{S1} given by:

$$\frac{d \, aTc}{dt} = \begin{cases} k_{aTc}^{in}(u_{aTc} - aTc), & \text{if } u_{aTc} > aTc \\ k_{aTc}^{out}(u_{aTc} - aTc), & \text{if } u_{aTc} \le aTc \end{cases},$$

$$\frac{d \, IPTG}{dt} = \begin{cases} k_{IPTG}^{in}(u_{IPTG} - IPTG), & \text{if } u_{IPTG} > IPTG \\ k_{IPTG}^{out}(u_{IPTG} - IPTG), & \text{if } u_{IPTG} \le IPTG \end{cases},$$
(5)

where aTc and IPTG denote the concentrations of the inducer molecules inside the cell, while u_{aTc} and u_{IPTG} those in the growth medium of the cells.

The values of all model parameters^{S1} are listed in Table S1.

$\kappa_{ m L}^{ m m0}$	$3.20e-2 \text{ mRNA} \min^{-1}$	$g_{ m L}^{ m m},g_{ m T}^{ m m}$	$1.386e-1 \text{ min}^{-1}$
$\kappa_{\mathrm{T}}^{\mathrm{m0}}$	$1.19e-1 \text{ mRNA} \min^{-1}$	$g_{ m L}^{ m p},g_{ m T}^{ m p}$	$1.65e-2 \text{ min}^{-1}$
$\kappa_{ m L}^{ m m}$	$8.30 \ \mathrm{mRNA} \ \mathrm{min}^{-1}$	$\theta_{\rm LacI}$	31.94 a.u.
$\kappa_{\mathrm{T}}^{\mathrm{m}}$	$2.06 \text{ mRNA} \min^{-1}$	$\eta_{\rm LacI}$	2.00
$\kappa^{\mathrm{p}}_{\mathrm{L}}$	$9.726e-1 a.u. mRNA min^{-1}$	$ heta_{ m TetR}$	30.00 a.u.
$\kappa^{\mathrm{p}}_{\mathrm{L}}$	$9.726e-1 a.u. mRNA min^{-1}$	η_{TetR}	2.00
$k_{\rm IPTG}^{\rm in}$	$2.75e-2 \text{ min}^{-1}$	$ heta_{ m IPTG}$	$9.06\mathrm{e}\text{-}2~\mathrm{mM}$
$k_{\rm IPTG}^{\rm out}$	$1.11e-1 min^{-1}$	$\eta_{ m IPTG}$	2.00
$k_{\mathrm{aTc}}^{\mathrm{in}}$	$1.62e-1 \text{ min}^{-1}$	θ_{aTc}	11.65 ng/ml
$k_{\mathrm{aTc}}^{\mathrm{out}}$	$2.00e-2 \text{ min}^{-1}$	$\eta_{ m aTc}$	2.00

Table S1: Value of the parameters of the model (1)-(6).

Average Model

By assuming (i) instantaneous diffusion of the inducers across the cell membrane, (ii) equal degradation rates for LacI and TetR (that is, $g_{\rm L}^{\rm p} = g_{\rm T}^{\rm p} = g^{\rm p}$), and (iii) exploiting the fact that the time scales of the mRNA dynamics are notably faster than those of the proteins,^{S2}

we can obtain the following non-dimensional quasi-steady state model of the toggle switch:

$$\frac{dx_1}{dt'} = k_1^0 + \frac{k_1}{1 + x_2^2 \cdot w_1(t'/g^p)} - x_1
\frac{dx_2}{dt'} = k_2^0 + \frac{k_2}{1 + x_1^2 \cdot w_2(t'/g^p)} - x_2$$
(7)

where

$$t' = g^{\mathrm{p}} t, \quad x_1 = \frac{LacI}{\theta_{\mathrm{LacI}}}, \quad x_2 = \frac{TetR}{\theta_{\mathrm{TetR}}},$$
(8)

are *rescaled* time and states, and the dimensionless parameters are defined as

$$k_1^0 = \frac{\kappa_{\rm L}^{\rm m0} \kappa_{\rm L}^{\rm p}}{g_{\rm L}^{\rm m} \theta_{\rm LacI} g^{\rm p}}, \quad k_1 = \frac{\kappa_{\rm L}^{\rm m} \kappa_{\rm L}^{\rm p}}{g_{\rm L}^{\rm m} \theta_{\rm LacI} g^{\rm p}},$$
$$k_2^0 = \frac{\kappa_{\rm T}^{\rm m0} \kappa_{\rm T}^{\rm p}}{g_{\rm T}^{\rm m} \theta_{\rm TetR} g^{\rm p}}, \quad k_2 = \frac{\kappa_{\rm T}^{\rm m} \kappa_{\rm T}^{\rm p}}{g_{\rm T}^{\rm m} \theta_{\rm TetR} g^{\rm p}}.$$

The nonlinear functions $w_1(t)$ and $w_2(t)$ take into account the static relationship between each repressor protein (TetR or LacI) and its corresponding regulator molecule (aTc or IPTG, respectively). They are defined as

$$w_1(aTc(t)) := \frac{1}{\left(1 + \left(\frac{aTc(t)}{\theta_{\rm aTC}}\right)^{\eta_{\rm aTc}}\right)^{\eta_{\rm TetR}}}$$
(9)

$$w_2(IPTG(t)) := \frac{1}{\left(1 + \left(\frac{IPTG(t)}{\theta_{\rm IPTG}}\right)^{\eta_{\rm IPTG}}\right)^{\eta_{\rm LacI}}}$$
(10)

System (7) can be averaged when fed with two *mutually exclusive pulsatile inputs*, of the form

$$u_{\rm aTc}(t) = \bar{u}_{\rm aTc} \cdot (1 - s_q(t/T))$$

$$u_{\rm IPTG}(t) = \bar{u}_{\rm IPTG} \cdot s_q(t/T)$$
(11)

where $s_q(t/T)$ is a periodic square wave of period T with duty-cycle $d \in [0, 1]$. Such averaging

analysis yields the following average model:

$$\frac{dx_1}{d\tau} = \varepsilon \left[k_1^0 + k_1 \left(\frac{d}{1 + x_2^2} + \frac{1 - d}{1 + x_2^2 \cdot \bar{w}_1(\bar{u}_{aTc})} \right) - x_1 \right]$$

$$\frac{dx_2}{d\tau} = \varepsilon \left[k_2^0 + k_2 \left(\frac{d}{1 + x_1^2 \cdot \bar{w}_2(\bar{u}_{IPTG})} + \frac{1 - d}{1 + x_1^2} \right) - x_2 \right]$$
(12)

where $\tau = t'/g_{\rm p}T$ and $\varepsilon = Tg^{\rm p}$.

The most relevant property of model (12) is that when it possesses a unique exponentially stable equilibrium point \bar{x}_{av} , then the solutions of the original time-varying system (7), from which (12) is derived, will converge at steady-state to a neighborhood of \bar{x}_{av} . Therefore, \bar{x}_{av} can be used as good approximation of the average value of the response of (7) when subject to mutually exclusive pulsatile inputs (11).^{S2}

Curves of equilibria of the average model

The number and position in state space of the equilibrium points $\bar{x}_{av} = [\bar{x}_1, \bar{x}_2]$ of the average vector field (12) depend on the specific choice of the amplitudes \bar{u}_{aTc} and \bar{u}_{IPTG} of the mutually exclusive pulsatile inputs, and on the value of the duty-cycle d. For example, for the reference values $\bar{u}_{aTc} = 50 \text{ ng/ml}$ and $\bar{u}_{IPTG} = 0.5 \text{ mM}$, system (12) is monostable and the position of the equilibrium point \bar{x}_{av} varies monotonically with d as reported in Figure 5 in our previous work^{S2} (blue dots). Hence, given certain values of \bar{u}_{aTc} and \bar{u}_{IPTG} , it is possible to move the position of \bar{x}_{av} on the corresponding curve by varying d, as reported in Figure S1 in our previous work.^{S2}

Supplementary details on designed controllers and simulations

PI-PWM

The PI-PWM relies on an ensemble of analytical approximations.^{S2} Firstly, the curves of equilibria $\Gamma_{\bar{u}_{aTe},\bar{u}_{IPTG}}$ of the average model used by the projector II are computed by assuming quasi-steady state of the transcription dynamics of mRNAs and instantaneous diffusion of inducer molecules through the cell membrane. Secondly, the equilibrium point \bar{x}_{av} of the average model is an approximation of the mean value of the oscillations of LacI and TetR. This accuracy depends on the parameter $\varepsilon = T g^{p}$ in the average model equations, where T is the period of the forcing inputs, and we fixed its value to 240 min that represented a good trade-off between the time scales of the toggle switch itself and diffusion effects of the cell membrane. Moreover, the desired setpoint $\bar{x}_{ref} = [LacI_{ref}/\theta_{LacI}, TetR_{ref}/\theta_{TetR}]$, to which we want regulate the measured mean value $\langle x_k \rangle$ of the toggle switch response, does not exactly lie on the curve $\Gamma_{\bar{u}_{aTe},\bar{u}_{IPTG}}$ returned by the Model Based Inversion algorithm and employed to compute the error signal e_k^{π} for the PI (see Supplementary Figure S3). Therefore, the curve represents an additional constraint to the performance of the control system.

The tuning of the PI gains was carried out heuristically via numerical simulations in MATLAB. Specifically, the closed loop system was simulated for 50 periods for 40,000 pairs of gain values $k_{\rm P}$ and $k_{\rm I}$ selected uniformly in the ranges $k_{\rm P} \in [10^{-4}, 1]$ and $k_{\rm I} \in [10^{-5}, 0.1]$; both intervals were divided in 200 uniformly distributed samples. Fig. S1 shows the value of the settling time of the duty-cycle d_k and the norm of steady-state projected error e_{π}^{∞} for each pair of gain values. The values of $k_{\rm P} = 0.0101$ and $k_{\rm I} = 0.0401$ were selected as those giving the best compromise between speed of the transient and residual steady-state error.

Model Predictive Control

Genetic algorithms¹ were used to numerically find the (sub)optimal control solution at each step. We adopted the MATLAB genetic algorithm toolbox by setting the initial population to 50 individuals randomly chosen in the interval $[0, d_{ref}]$. The maximum number of generations was set to 150, while the maximum number of stall generations was set to 30. All the other parameters were kept to their default values. Using the parallel computation toolbox and utilizing 12 logical cores of an INTEL XEON E5-2640v3 CPU, the optimization routine (setting $T_p = 720$ minutes, that is, N = 3) takes about 3 minutes to converge towards a solution.

The control parameters K_{LacI} and K_{TetR} in the cost function J_k were selected heuristically to $K_{\text{LacI}} = 1$ and $K_{\text{TetR}} = 4$, after an extensive numerical search in MATLAB. Specifically, the control evolution was simulated for 18 periods fixing $K_{\text{LacI}} = 1$ and varying K_{TetR} over 37 values chosen uniformly in the interval [0.01, 100], so as to vary the ratio between the two gains. The best values given above were then selected for the in-silico experiments reported in the main text.

Simulations

Stochastic simulations were also performed in MATLAB using the Gillespie's Stochastic Simulation Algorithm. 2

¹https://it.mathworks.com/help/gads/ga.html

 $^{^{2}}$ https://it.mathworks.com/matlabcentral/fileexchange/34707-gillespie-stochastic-simulation-algorithm

Supplementary Figures



Figure S1: Tuning of the PI controller. (a) Settling time of the duty-cycle at the 10% of its final value, computed as a number of periods, for all pairs $(k_{\rm P}, k_{\rm I}) \in [10^{-4}, 1] \times [10^{-5}, 0.1]$. Note that the performance was evaluated over a simulation time of 50 periods and yellow colored squares denote values of settling time ≥ 50 periods. (b) Zoom of the most significant parameter region in panel a (highlighted within the red box); (c) Norm of the steady-state projected error e_{π}^{∞} for the same range of values of control gains as in panel b. The red box in panels b and c indicates the values of PI gains that were selected and used for all in-silico control experiments.



Figure S2: Agent-based simulation in BSim of the cells evolution in open-loop. The pulsatile inputs' amplitudes were set to $\bar{u}_{aTc} = 35$, $\bar{u}_{IPTG} = 0.35$, while the duty-cycle was kept constant (without any adaptation) to $d_{ref} = 0.4$. The period was selected as usual to be T = 240 min. Total simulation time is 72 hours. We considered E. coli cells growing in a single chamber of a "mother machine"-like microfluidic device: ^{S3} the simulations start with a single cell located at the bottom of the chamber; as the cell grows and duplicates, it pushes outside of the chamber new cells that exceed the maximum capacity of the chamber (around 10 cells). (a) Evolution over time of *LacI*; the dashed line representing the setpoint *LacI*_{ref} = 750, while lighter lines the evolution of the state for each cell in the simulation, and the darker solid line the mean trajectory computed over the population, evaluated through a moving window of period *T*. Panel (b) Evolution of the state for each cell in the simulation the state for each cell in the setpoint $TetR_{ref} = 300$, lighter lines are the evolution of the state for each cell in the solid line represents the evolution of the mean trajectory across the population in the period, evaluated using a moving window of period *T*.



Figure S3: Working principle of the nonlinear projector block. (a) The red curve $\Gamma_{\bar{u}_{aTc},\bar{u}_{IPTG}}$ represents the closest one to the setpoint \bar{x}_{ref} selected by the Model Based Inversion algorithm; black curves are other equilibrium curves that are farther from the setpoint. The setpoint \bar{x}_{ref} and the mean value of the state in the k-th period $\langle x_k \rangle$ are projected onto the curve on the points \hat{x}_{ref} and $\langle \hat{x}_k \rangle$, respectively. The length of the curve between \hat{x}_{ref} and $\langle \hat{x}_k \rangle$, highlighted in blue, is the projection error e_k^{π} at the time instant k. (b) Even if the projected error e_k^{π} is zero, that is $\|\langle \hat{x}_k \rangle - \hat{x}_{ref}\| = 0$, this does not necessarily imply zero regulation error, indeed in the case represented here $\|\langle x_k \rangle - \bar{x}_{ref}\| \neq 0$.

References

- (S1) Lugagne, J.-B., Sosa Carrillo, S., Kirch, M., Köhler, A., Batt, G., and Hersen, P. (2017) Balancing a genetic toggle switch by real-time feedback control and periodic forcing. *Nat. Commun.* 8, 1671.
- (S2) Fiore, D., Guarino, A., and di Bernardo, M. (2019) Analysis and Control of Genetic Toggle Switches Subject to Periodic Multi-Input Stimulation. *IEEE Control Syst. Lett.* 3(2), 278–283.
- (S3) Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., and Jun, S.
 (2010) Robust growth of Escherichia coli. *Curr. Biol.* 20(12), 1099–1103.