Control-based continuation: a new approach to prototype synthetic gene networks

Irene de Cesare[1,2], Davide Salzano[1,2], Mario di Bernardo[2,†], Ludovic Renson [3,†,#] and Lucia Marucci[1,4,5,†,#]

[1]Engineering Mathematics Department, University of Bristol, Bristol BS8 1TW, UK.
[2]Department of Electrical Engineering and Information Technologies, University of Naples Federico II, 80125 Naples, Italy.
[3]Department of Mechanical Engineering, Imperial College London, UK.
[4] BrisSynBio, University of Bristol, Bristol BS8 1TQ, UK.
[5] School of Cellular and Molecular Medicine, University of Bristol, Bristol BS8 1UB, UK.
[†] Co-senior authors.
[#] Co-corresponding authors; l.renson@imperial.ac.uk; lucia.marucci@bristol.ac.uk

Abstract

Control-Based Continuation (CBC) is a general and systematic method to carry out the bifurcation analysis of physical experiments. CBC does not rely on a mathematical model and thus overcomes the uncertainty introduced when identifying bifurcation curves indirectly through modelling and parameter estimation. We demonstrate, *in silico*, CBC applicability to biochemical processes by tracking the equilibrium curve of a toggle switch which includes additive process noise and exhibits bistability. We compare results obtained when CBC uses a model-free and model-based control strategy and show that both can track stable and unstable solutions, revealing bistability. We then demonstrate CBC in conditions more representative of a real experiment using an agent-based simulator describing cells growth and division, cell-to-cell variability, spatial distribution, and diffusion of chemicals. We further show how the identified curves can be used for parameter estimation and discuss how CBC can significantly accelerate the prototyping of synthetic gene regulatory networks.

Keywords— Control-Based Continuation, Synthetic Biology, Bifurcations, Toggle Switch

Introduction

The complexity of biological systems is widely acknowledged. In native organisms, multi-scale intracellular interactions often result in complex nonlinear dynamics. Consider, for example, switches and oscillations in gene expression, which are used by cells to process external inputs and program

specific cell outputs. Synthetic biology aims at engineering biological computation by recreating such dynamics using circuits embedded into living cells [1, 2, 3, 4, 5, 6, 7, 8, 9, 10].

Mathematical modelling is widely used within synthetic biology design-build-test-learn cycles. In the context of engineered gene regulatory networks, models can both support the design phase (indicating the parameter space which allows the emergence of the desired dynamics, such as oscillations), and the testing upon experimental implementation. Moreover, modelling is currently the only way in synthetic biology to study the relationship between physical parameters variations and bifurcations; the latter represent stability boundaries where qualitative and quantitative changes to the system's dynamics occur. For instance, saddle-node and Hopf bifurcations (responsible for hysteresis and oscillatory behaviours, respectively) that are commonly observed in native biological systems [11] can only be studied by nonlinear model analysis. This means that, firstly, a mathematical model needs to be derived and fitted to often *ad hoc* generated experimental data. Then, nonlinear model analysis can be performed to identify the bifurcation behaviour being observed.

The derivation of biochemical models can however be challenging, both in terms of model structure (which depends on underlying hypothesis on the system), and parameter identification and validation (which can be troublesome in the case of incomplete/noisy experimental datasets). Model uncertainties will inevitably result in misleading conclusions [12, 13]. As a consequence, the design and testing of engineering synthetic biochemical circuits that perform as intended is extremely difficult, unless various design, build, test and learn iterations are performed [14].

Control-based continuation (CBC), originally proposed by Sieber and Krauskopf [15], is a general, systematic and model-free testing method that applies the principles of numerical continuation (a computational method for bifurcation analysis) to physical experiments. The fundamental principles of CBC are well established and the method has been applied to a wide range of mechanical systems. For instance, Barton et al. [16] studied the periodic responses of a bistable energy harvester. A similar system was studied by Renson et al. [17]. The continuation method was also demonstrated on a multi-degree-of-freedom structure exhibiting isolated curves of periodic responses and quasi-periodic oscillations [18]. CBC exploits feedback control to explore the nonlinear dynamics of a system, detect bifurcations, and eventually trace their evolution with respect to adjustable parameters directly during experimental tests.

Recently, external feedback controllers have been exploited for controlling gene expression in living cells by means of microfluidics/microscopy or optogenetics platforms [19, 20, 21, 22, 23, 24, 25, 26, 27]. From a control theory standpoint, genetic networks are the processes to be controlled, while the controller is implemented on a computer. Biosensors are used to measure the state of the process, usually by means of fluorescent reporter proteins. The fluorescence evaluation can be done either at single-cell or at cell population level, and is fed to a control algorithm that evaluates the appropriate control input aiming to steer the process to a selected reference signal. Inputs are then actuated on the process via light or chemical compounds [28].

By employing external feedback controllers to steer gene expression it should be possible to apply CBC to track nonlinear dynamics in biochemical systems, overcoming the need for a model

identification step and defining a shorter way for bifurcation tracking which includes the isolation of unstable equilibria (see in Fig. 1A path 1 compared to path 2).

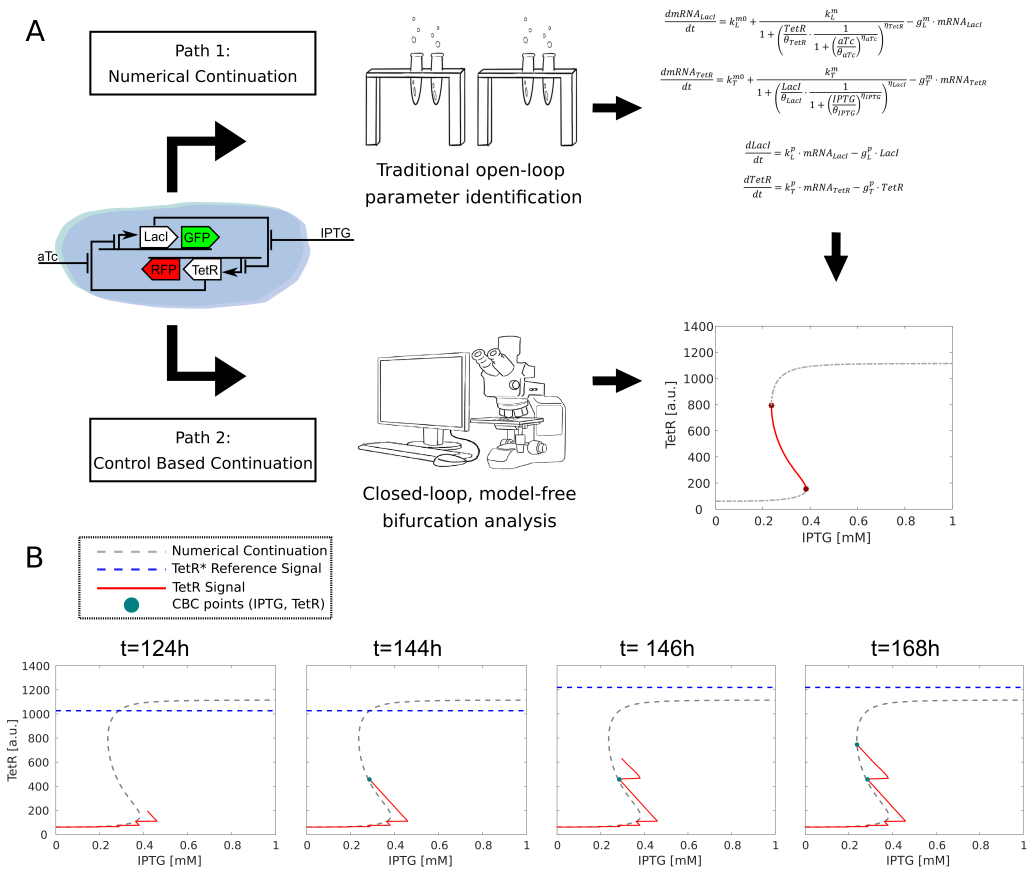

Figure 1: A) A schematic Toggle Switch from Lugagne et al. [6]; two different paths to the bifurcation diagram (- -): numerical continuation (path 1), and control-based continuation (path 2). Unstable branch of the bifurcation curve (–) and bifurcation points (*) are highlighted. B) Steps of the CBC algorithm: at time $t = 124h$ the control pushes the system's trajectory $TetR(IPTG)$ (–) towards the new reference $TetR^*$ (- -). At time $t = 144h$ transient dynamics are extinguished and a point $(\overline{IPTG}, \overline{TetR})$ is saved (.). Then, the reference signal is increased and the process is repeated.

Here, we demonstrate, for the first time, the applicability of CBC to prototype the dynamics of a synthetic gene regulatory network and to estimate the parameters of a fixed model structure, if available, using data collected after the CBC routine. We use as a test bed the toggle switch, a bistable biological circuit, firstly embedded in *Escherichia coli* cells by Gardner et al. [10], which is often used to benchmark new control strategies and is considered a fundamental tool for cell computation [11, 29]. Specifically, we run *in silico* experiments on a recent version of the toggle switch, for which external feedback control was shown to be successful [6]. Moreover, we leveraged an agent-based simulator called BSim [30] to test the performance of the designed algorithms in conditions more representative of a real experiment. Our *in silico* demonstration of CBC gives us the confidence

that the method should be exploitable *in vivo* to fully explore the parameter space which allows the emergence of the desired dynamics (in our example, hysteresis), leading to rapid and model-free characterization, and also parameter estimation, of engineered synthetic modules.

Results and Discussion

CBC retrieves stable and unstable invariant solutions of a dynamical system through the application of an external control action. In order to acquire the bifurcation diagram, the controller must not modify the position in the parameter space of the uncontrolled system's invariant solutions. For example, the equilibria of a controlled system are in general different from the one of the uncontrolled system and, to recover the response of the underlying uncontrolled system of interest, CBC seeks a control signal that decays asymptotically to zero, i.e.,

$$\lim_{t\to\infty} u(x(t), x_{ref}) \equiv 0 \tag{1}$$

where $x \in \mathbb{R}$ is the system's state, $x_{ref} \in \mathbb{R}$ is the control reference signal and t is time. Although the control signal is asymptotically converging to zero, it is synthetized in order to stabilise local dynamics of the system such that unstable equilibria become stable and hence detectable forward in time. A control strategy that satisfies (1) is called noninvasive and does not modify the system's equilibrium positions in the parameter space. Finding a noninvasive control signal usually requires finding the "right" reference input (x_{ref}) for the controller such that (1) is satisfied, while the solution of interest is stabilised.
When the control input can be chosen as the bifurcation parameter of interest, the methodology can be significantly simplified as the control signal is only required to settle to a constant value. Indeed, when this condition is achieved, the non-zero constant control signal can be viewed as a mere shift in the bifurcation parameter value. This simplified method is used in this paper, and more details about its implementation can be found in the Methods section. Furthermore, CBC can also be extended to characterise systems exhibiting oscillations and other bifurcations as described in [17, 18, 31, 32].

The applicability of CBC is demonstrated here *in silico* on a synthetic gene network, the toggle switch (Fig. 1). The mathematical model developed by Lugagne et al. [6] is used. From a control system standpoint, it can be seen as a 2-input (aTc and $IPTG$) 2-output ($LacI$ and $TetR$) nonlinear dynamical system. The inputs correspond to drugs that can be added to the cells environment, while the outputs are proteins synthesized by cells and detected by tagging them with fluorescent reporters. Here the control signal (specifically $IPTG$) plays a dual role: not only it pushes the system into a new state, but it keeps the system stable allowing it to explore the nearby unstable dynamics which would not be collected otherwise (see the red branch of the bifurcation diagram in Fig. 1 A). The model derivation and further information about the network can be found in the Methods section.

Fig. 1 B illustrates an experiment conducted with CBC; data points collected with CBC (blue point) are compared to the actual bifurcation curve obtained using standard model-based numerical continuation algorithms (gray dashed line). First, a particular control reference value ($x_{ref} = TetR^*$) is selected (dashed-blue line). The reference signal is compared to the actual expression of the gene of interest to compute a control error. This is fed to the controller, which evaluates a controlling signal ($IPTG$ concentration), translated into an input to the toggle switch that will therefore change its state (the red line in Fig. 1 B shows the system transient dynamic). At every sampling time a new measure

is acquired and the entire process is repeated until a steady state value is reached. The steady state value of the state variables, together with the associated control signal, are then saved to define a point $(\overline{IPTG}, \overline{TetR})$ in the bifurcation diagram. Subsequently, a new reference value is picked (see Methods) and the entire process is repeated. To trace out the entire equilibrium curve, the continuation algorithm requires a set of control reference values, broadly covering the range of gene expression of interest (in our case the $TetR$ fluorescence expression).

CBC is a model-free method because it does not require the knowledge of a mathematical model of the system to be applied to, and the results' accuracy does not depend on the structure and parameter values of a model. Furthermore, the controller within CBC is only required to be stabilizing and noninvasive. As such, CBC is not tied to a specific type of feedback controller and control laws that do not require a mathematical model of the system can be used. However, the design, parameter tuning and overall performance of the controller can be improved if some knowledge (like a model) of the system dynamics is available. Here, the use of a model-free proportional (P) controller and a model-predictive controller (MPC) is compared. The former controller was chosen as it is widely used in CBC applications, while the latter is commonly exploited in synthetic biology [20]. When using Model Predictive control strategies, simple linear models work well, showing that detailed mathematical representations are not needed for CBC to be work. A sampling time of 5 minutes is considered in this work; this is a realistic time interval that we previously used to image and externally control bacterial cells [22, 25]. Note that, with the simplified CBC approach, the controller is not required to take the control error to zero but only to stabilise the system to the nearest equilibrium point. This does not affect the accuracy of results, as further explained in Methods.

CBC can reconstruct the toggle switch bifurcation diagram under deterministic simulation settings

We first applied CBC on the deterministic model of the toggle switch. During the *in silico* experiment, 30 steady state points $(\overline{IPTG}, \overline{TetR})$ were acquired by varying the reference input to the control strategy between $TetR^* = 1800[a.u.]$ (or $1200[a.u.]$) and $TetR^* = 0[a.u]$. The maximum duration of a single equilibrium acquisition was constrained to 9 hours and 55 minutes, after which the control reference was modified. The steady states values were computed by taking the average of the samples recorded in the last 60 minutes (12 samples) of the simulation carried out for each value of the reference input. Averaging is not essential in a deterministic scenario, but becomes fundamental in a noisy environment as a way to filter out some noise. Points collected with the CBC perfectly overlap with the numerical continuation diagram (Fig. 2 A and 3 A). Both control strategies have comparable tracking performances and proved able to stabilise the unstable branch of equilibria delimited by the Saddle-Node bifurcations, capturing the whole bifurcation curve.
The proportional control signal is linearly dependent on the error signal. Every time the reference is modified (Fig. 2 B), the error changes and consequently the control signal as well changes (Fig. 2 C). As the controller does not include an integral action, the error never nullifies, but it becomes constant once the system reaches the equilibrium. When this happens, the control signal and bifurcation parameter $IPTG$ becomes constant as well (because the proportional contribution is constant), and the steady state can be collected.

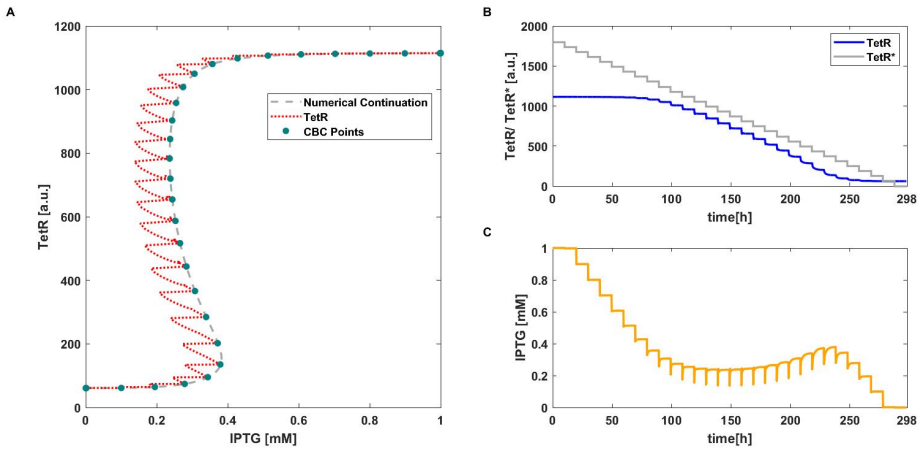

Figure 2: CBC with P controller applied to the deterministic toggle switch model (2). A) Equilibrium curve measured using CBC (·). $TetR(IPTG)$ transient trajectories (..). Reference equilibrium curve obtained using numerical continuation (- -). B) Time evolution of $TetR$ and the control reference signal $TetR^*$ (–). C) Time evolution of $IPTG$ (i.e. control signal). Parameter values: $k_p = 0.0016$ and $aTc = 25ngmL^-1$.

With MPC, the control action is computed as the optimal signal that minimizes the error, and therefore its contribution pushes the system's trajectories towards the reference signal, generating some early oscillations (Fig. 3 B, C). As the bifurcation parameter is not directly proportional to the control error, the latter does not have to be different from zero. The steady state error introduced by the MPC algorithm is much lower than the proportional one (see Fig. 2 B and Fig. 3 B); thus, it is easier to define the range of reference values for the MPC as it corresponds to the effective dynamical range explored by the continuation algorithm. Instead, when using a proportional controller, the range of the control reference may need to be varied significantly (as in Fig. 2 B, where the maximum reference was set to $TetR^* = 1800[a.u.]$ in order for the algorithm to get a full coverage of the bifurcation diagram). This should not be considered as a fault of the proportional algorithm but only as a difference between the two control strategies. Results in Fig. 2 and Fig. 3 are also represented in Supplementary Movies S1 and S2.

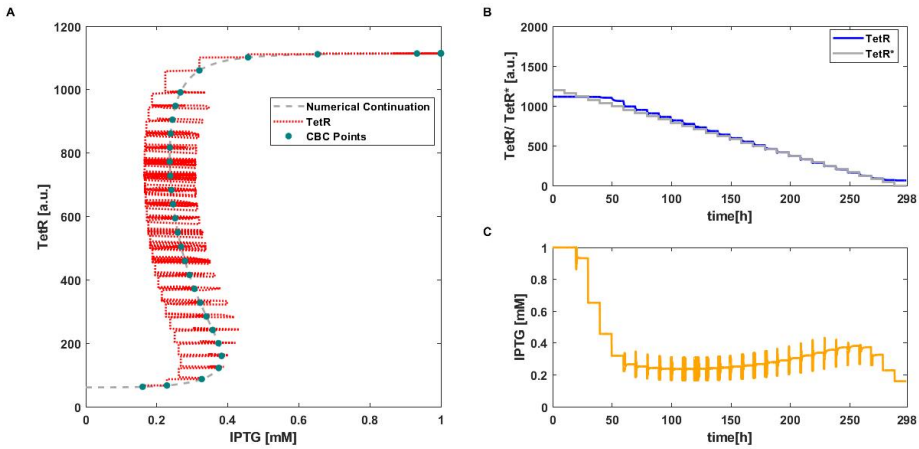

Figure 3: CBC with MPC applied to the deterministic toggle switch model (2). A) Equilibrium curve measured using CBC (·). $TetR(IPTG)$ transient trajectories (..). Reference equilibrium curve obtained using numerical continuation (- -). B) Time evolution of $TetR$ and the control reference signal $TetR^*$ (–). C) Time evolution of $IPTG$ (i.e. control signal). Parameter values: $\gamma = 0.3$ (Eq. (25)) and $aTc = 25ngmL^{-}1$.

To decrease the total time of each simulation we also implemented a steady state detection routine that changes the reference each time the system reaches steady state (Methods). This, together with a smaller number of collected points, allowed to decrease the experimental time to approximately 55 hours for both the controllers, see Fig. S1 and S2.

CBC can reconstruct the toggle switch bifurcation diagram under a stochastic simulation setting

CBC is now demonstrated on a stochastic toggle switch model to reproduce conditions more similar to a real experiment. As noise can strongly modify results of a single continuation experiment, we consider 10 repetitions of the same experiment. In each we collect 30 points, with the steady states values computed as described for the ODE case.

The resulting plot is a density map made of all the collected steady states $(\overline{IPTG}, \overline{TetR})$ over the 10 simulations, that we can compare with the numerically computed bifurcation diagram (Fig. 4 A and 5 A). The results show dense clouds of points falling in the area where the branch of equilibria is, confirming the ability of the strategy to work also in a stochastic setting: both controllers were able to track the unstable branch of equilibria, preventing the system from jumping between the two stable steady states. For simplicity, we only show one out of 10 computed trajectories of $TetR$ and $IPTG$ in Fig. 4 B, C and 5 B, C.

For both the proportional controller and MPC, the output signal shows oscillations due to noise, that also affect the control signal u. However, these oscillations do not seem to strongly alter the average values we take as steady state points. Similar observations to the deterministic case can be made. For the proportional controller, the error does not nullify and the reference signal is chosen in a wider range of values than for the MPC in order to uncover the full bifurcation curve (Fig. 4 B).

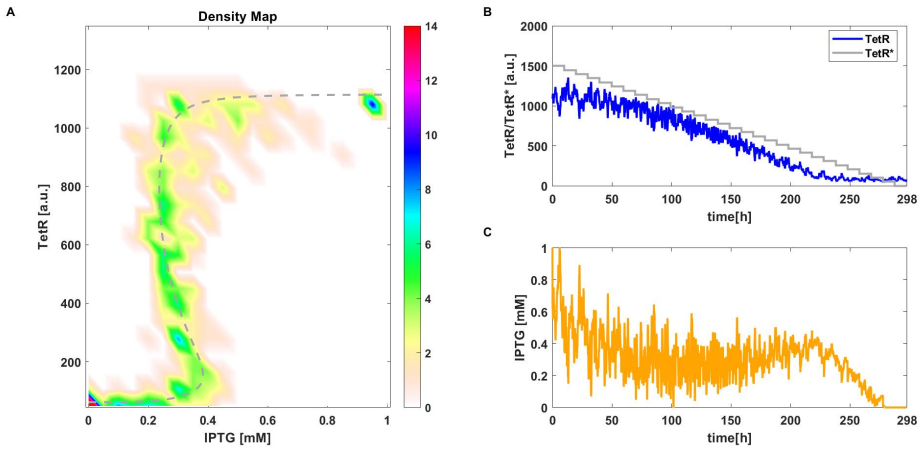

Figure 4: CBC with P controller applied to the stochastic toggle switch model (4). A) Density plot of equilibrium curve measured using CBC. Reference equilibrium curve obtained using numerical continuation (- -). B) Time evolution of one simulation of $TetR$ and the control reference signal $TetR^*$ (–). C) Time evolution of $IPTG$ (i.e. control signal). Parameter values: $k_p = 0.0016$ and $aTc = 25ngmL^-1$.

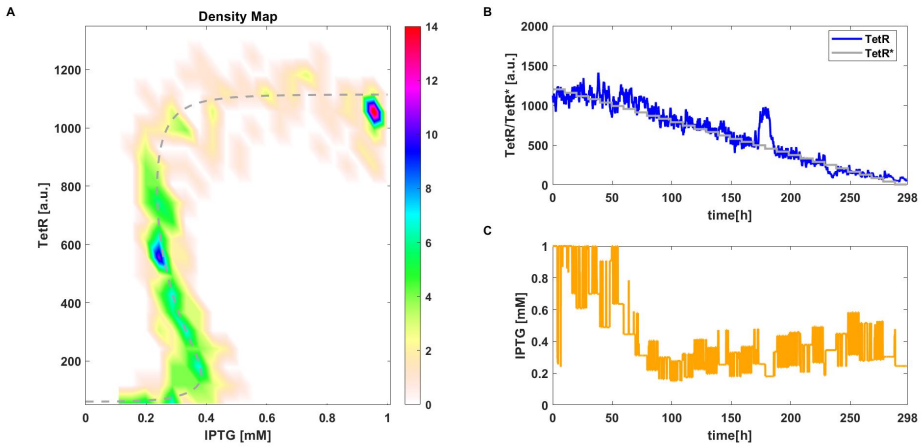

Figure 5: CBC with MPC applied to the stochastic toggle switch model (4). A) Density plot of equilibrium curve measured using CBC. Reference equilibrium curve obtained using numerical continuation (- -). B) Time evolution of one simulation of $TetR$ and the control reference signal $TetR^*$ (–). C) Time evolution of $IPTG$ (i.e. control signal). Parameter values: $\gamma = 0.3$ (Eq. (25)) and $aTc = 25ngmL^-1$.

As in the deterministic scenario, the simulation time can be drastically decreased considering a variable step time associated to the computation of the steady states and a reduced amount of collected points. The resulting bifurcation curves can be seen in Fig. S3 A and S4 A, while panels B and C show

just one representative trajectory of $TetR$ and $IPTG$, respectively. Furthermore, for the stochastic scenario, we also considered the case of steady state check with full amount of points (30). We found that the reference shift guided by steady state detection could only reduce the total time to 228 hours for a proportional controller and 233 hours in case of MPC algorithm (see Figure S5 and S6), i.e. a 23.4% and a 21.8% reduction, respectively. Reducing the total testing time requires to decrease the number of points collected, which can be achieved without losing accuracy in the diagram retrieved as shown in Figs. S3 and S4. Nonetheless, the present method could be improved by developing a algorithm for automatic reference stepping with variable numbers of point collected.

Single-cell CBC using agent-based simulations

To further prove the applicability of CBC to synthetic biology, we validated the method in BSim, a Java-based bacteria simulator [30, 33], where it is possible to consider a scenario more representative of a real experiment, including spatial distribution, movement, growth and division of cells as well as spatial distribution and diffusion of chemicals. We implemented single cell control considering the same model used for the stochastic scenario (Equation (4)). Simulation settings (i.e. sampling time, number of collected points, gains, etc.) where kept the same as above, with minor differences due to changes in the programming language.

In figures 6 B, 7 B and Supplementary Movies S3 and S4 representative examples of $TetR$ and $IPTG$ trajectories are shown. The bifurcation curves that are obtained, shown in Fig. 6 A and 7 A, are comparable with previous results obtained on the stochastic model. However, we observed worse control performances for the proportional controller in BSim. This is mainly due to additional factors introduced in simulations. Specifically, the cells biomechanics and their flush out from the microfluidic chamber are simulated, as well as the chemicals' spatial distribution and diffusion that introduce additional delay in the control input delivery. Additionally, in BSim we explicitly simulated an actuation delay due to the time the media takes to be delivered to the cells within the microfluidic device. All these factors might contribute to the performances exhibited by the proportional controller. A similar result when using more realistic simulation environments was also found in [34], where the authors showed that the performance of a PI controller considerably deteriorated when the algorithm was tested on an agent based model, while an MPC algorithm was able to maintain similar performance. Further information about BSim can be found in the Methods section.

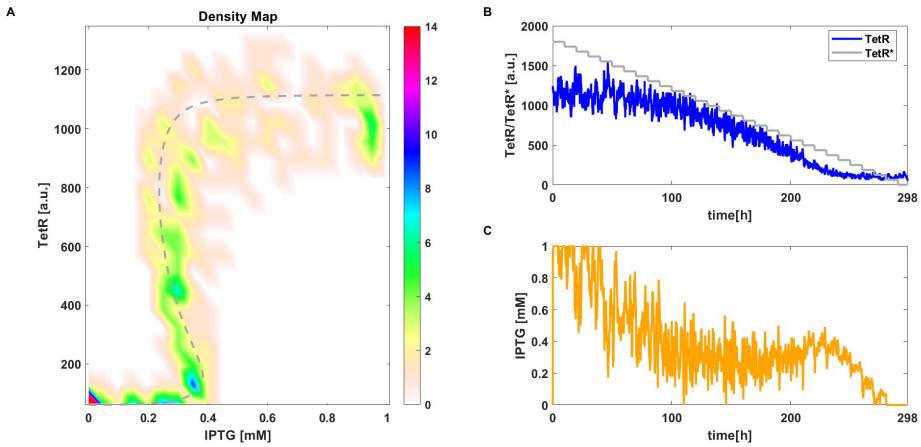

Figure 6: CBC with P controller applied to the stochastic toggle switch model (4) in BSim. A) Density plot of equilibrium curve measured using CBC. Reference equilibrium curve obtained using numerical continuation (- -). B) Time evolution of one simulation of $TetR$ and the control reference signal $TetR^*$ (–). C) Time evolution of $IPTG$ (i.e. control signal). Parameter values: $k_p = 0.0016$ and $aTc = 25ngmL^-1$.

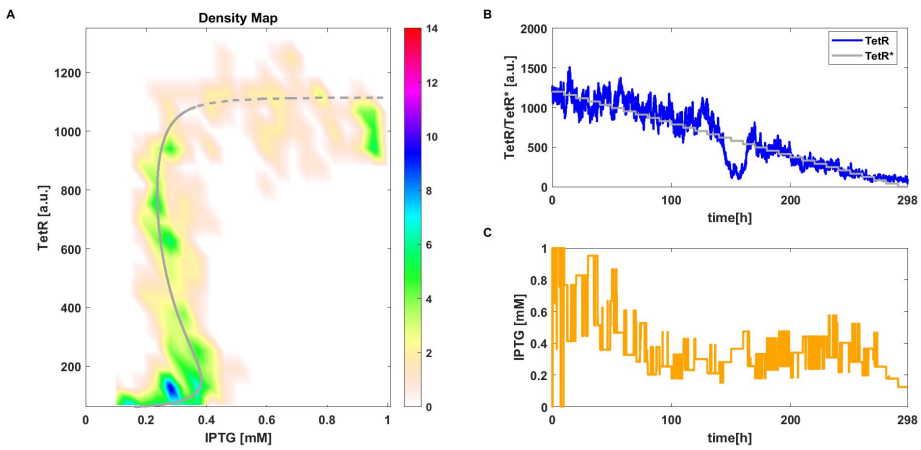

Figure 7: CBC with MPC applied to the stochastic toggle switch model (4) in BSim. A) Density plot of equilibrium curve measured using CBC. Reference equilibrium curve obtained using numerical continuation (- -). B) Time evolution of one simulation of $TetR$ and the control reference signal $TetR^*$ (–). C) Time evolution of $IPTG$ (i.e. control signal). Parameter values: $\gamma = 0.3$ (Eq. (25)) and $aTc = 25ngmL^-1$.

Results with reduced points and steady state detection obtained in BSim (Fig. S7 and Fig. S8) were similar to short stochastic simulations in Matlab. Here, the increased complexity of simulations takes into account many sources of noise (such as spatial diffusion of chemicals); for this reason,

results are more scattered around the numerically computed bifurcation diagram (see Fig. S7 A and Fig. S8 A). Nonetheless, implemented controllers manage to stabilise trajectories along the unstable branch of equilibria.
Considering the stochasticity of the experiments, there might be variability in the amount of time needed to uncover the bifurcation curve but, in general, it can be reduced up to less than a third of the original simulations.

Model calibration via data collected using CBC in a noisy scenario

Methods for parameters identification in nonlinear systems often rely on optimization routines: given a model, parameters identification allows to calibrate its output in order to match experimental data (refer to [35, 36] for further information).
Using the approach by Beregi et al. [32], we show how the equilibria measured using CBC can be exploited to estimate the toggle switch model parameters, and how data corresponding to unstable equilibria helps retrieving a model that captures the bistable behaviour of the system. Further information about the identification process can be found in Methods.
The equilibrium curves obtained from parameters identified using CBC data are shown in Fig. S9 D for the deterministic scenario and in Fig. S9 E, F for the stochastic scenario (in green and red data associated to the proportional controller and to the MPC, respectively). For the sake of comparison, model parameters were also estimated from points collected using a traditional, open-loop parameter sweep where different levels of input parameter ($IPTG$) were tested for both noise-free and noisy scenarios (Fig. S9 A and B, C, respectively). The equilibrium curve obtained in this way generally does not reproduce the bistable behaviour of the system, except slightly in the ideal case of the deterministic simulations (Fig. S9 A). On the other hand, data collected with CBC still capture bistability even in stochastic simulations (Fig. S9 E, F), although without being able to reproduce the original numerical bifurcation curve.
Parameter estimation results show that the use of data including unstable equilibria helps estimating parameters that reproduce bistability. As such, it is thought that the MPC control approach is preferable compared to the proportional control as it offers greater control over the distribution of data points along the equilibrium curve. Estimated parameters are shown in Table S1 for all different collected data. Interestingly, the variation of some parameters does not significantly affect the resulting fitted curve. Take, for example, the red line in S9 D, which is in excellent agreement with the reference equilibrium curve but was calculated with some identified parameters values being up to 41% different from their nominal value, hinting to non-identifiability of some parameters. We confirmed this by performing a structural identifiability analysis on the system, using the STRIKE-GOLDD toolbox [37], which showed unidentifiability of 9 parameters out of 14 when we considered two measurable outputs ($LacI$ and $TetR$), one known control input ($IPTG$) and one other input as a fixed constant (aTc). For further reference, structural unidentifiability of the 2 ODEs toggle switch system for fixed inputs was also highlighted by Villaverde et al. [38].
We note that bifurcation curves estimated from sweeps can be improved to reproduce bistability if additional knowledge about the presence of horizontal asymptotes for the stable branches is introduced in the cost function used for parameter estimation (Fig. S10 A-C). However, obtaining qualitatively different results for similar estimation approaches would typically lead to poor confidence in the results. This modified approach is also not systematic as the true behaviour of the system is assumed to be known. Using this new approach with CBC data further improves the agreement between the reference and identified models (Fig. S10 D-F).

Concluding Remarks

This paper presents the first application, *in silico*, of CBC to biochemical systems. In particular, CBC can track the stable and unstable equilibria of a synthetic gene network behaving as a toggle switch. In the absence of noise, a perfect agreement between equilibrium curves measured using CBC and standard numerical continuation methods is found. A notable challenge in biological applications, as compared to the mechanical systems studied so far with CBC, is the significant presence of (process) noise. Our results on the stochastic models (Matlab and Bsim) clearly show that CBC is still able to uncover the bistable nature of the system, as recorded data points qualitatively follow the equilibrium curve of the underlying deterministic model.

Besides exploring the nonlinear dynamics of the physical system directly in the experiment, CBC provides informative data for parameter estimation. Our results show that the measured data, especially those falling in the unstable region, was critical for the identification of model parameters that can reproduce bistability.

CBC appears therefore as a valuable tool for uncovering the dynamics of potentially more complex synthetic gene networks even in the presence of significant levels of noise. CBC can also be used to characterise oscillations [39] and other bifurcations [17, 40]. The information gathered with CBC has the potential to enable a more accurate understanding of biochemical interactions and thus a more precise prototyping of those behaviours into novel synthetic gene circuits.

An important feature of CBC is that it is a model-free method since it does not require a model of the system to work, and the accuracy of the results is independent from any modelling assumption. CBC also does not rely on a particular control strategy, provided the controller can be made noninvasive and is able to stabilise unstable responses. In this regard, the use of a mathematical model (even approximate) can improve the control performance. For instance, the initial estimate of a proportional controller gain can be obtained based on a model and then further refined by trial and error during experiments. We also showed that a model-predictive controller, based on a simple linear model, can provide a smaller steady-state control error, resulting in better discretisation of the equilibrium curve and beneficial for parameter estimation. MPC is a controller broadly used in biology [20] which generally performs best for fast varying references [41]. We hope that the use of MPC in the present paper paves the way for a wide range of applications of CBC in systems and synthetic biology.

Future developments

In future studies we aim to apply the strategy *in-vivo*, using the microfluidic experimental set-up already described in [20, 22, 24], which proved to be successful for control experiments involving both mammalian and bacterial cells.
Several challenges will have to be addressed to enable a wider uptake of CBC:

1. *Noise*. The presence of noise can make stabilization more difficult, as observed here and also highlighted by Lugagne et al. [6]. Moreover, the presence of noise affects the noninvasiveness associated to the control input. Indeed, noise introduces variations in the control signal which cannot be suppressed completely. While this appears to have a mild effect on the system considered here, this is arguably not the case in general. As of now, some noise-related effects

can be compensated by considering signal processing techniques, but more rigorous methods are required.

2. *Control*. As of now, there are no general and systematic methods to design the controller used in CBC. Currently, control parameters are found by trials and errors. While this works for simple systems, this approach does not scale to more complex systems where more sophisticated control strategies including, for instance, multiple inputs and outputs, will be needed. A step in that direction has recently been made by Li and Dankowicz [42] who proposed adaptive control design strategies for CBC.

3. *Experimental time*. Our initial simulations were particularly long (298 hours in Fig. 4-7). While such experiments can be achieved with some experimental set-ups – Balagaddé et al. [43] performed experiments on an *Escherichia coli* prey-predator system that lasted between 200 and 500 hours using a microchemostat platform – this is currently not possible using microscopy/microfluidic platform [6, 22]. It is in principle possible to significantly reduce testing time by improving the performance of the controller to reduce the transient. However, this can prove difficult due to the significant presence of noise in biochemical systems. Nonetheless, we showed that, for the presented gene network, it is possible to adjust the number of recorded points and CBC parameters to significantly reduce experimental time (2-3 days in the deterministic case, and 3-4 days in the stochastic case) while preserving a fine discretisation of the equilibrium curve.

Methods

Model of the toggle switch

The genetic toggle switch is a well-characterized bistable biological circuit, first embedded in *Escherichia coli* by Gardner et al. [10]. The system consists of two repressors (*LacI* and *TetR*), which mutually repress each other, and two inducers (Atc and $IPTG$), which can externally tune the genes' production (Fig. 1 A). In the absence of inducers, the circuit exhibits two stable equilibria. Levels of *LacI* and *TetR* can be measured via fluorescent reporters: for example, *mKate2*(RFP) and *mEGFP*(GFP) can be used to monitor *LacI* and *TetR*, respectively [6].

Here, we will refer to the Hill-type model developed by Lugagne et al. [6], representing mRNA transcription, translation and degradation/dilution due to growth. Mathematically, the system is described by a set of four ordinary differential equations:

$$\begin{aligned}
\frac{dmRNA_{LacI}}{dt} &= k_L^{m0} + \frac{k_L^m}{1+\left(\frac{TetR}{\theta_{TetR}} \times \frac{1}{1+\left(\frac{aTc}{\theta_{aTc}}\right)^{\eta_{aTc}}}\right)^{\eta_{TetR}}} - g_L^m \times mRNA_{LacI} \\
\frac{dmRNA_{TetR}}{dt} &= k_T^{m0} + \frac{k_T^m}{1+\left(\frac{LacI}{\theta_{LacI}} \times \frac{1}{1+\left(\frac{IPTG}{\theta_{IPTG}}\right)^{\eta_{IPTG}}}\right)^{\eta_{LacI}}} - g_T^m \times mRNA_{TetR} \\
\frac{dLacI}{dt} &= k_L^p \times mRNA_{LacI} - g_L^p \times LacI \\
\frac{dTetR}{dt} &= k_T^p \times mRNA_{TetR} - g_T^p \times TetR
\end{aligned} \tag{2}$$

where the system states are given by the genes production *LacI* and *TetR* and the associated *mRNAs* ($mRNA_{LacI}, mRNA_{TetR}$), $k^{m0}_{L/T}$ is the leakage transcription rate, $k^{m}_{L/T}$ is the transcription rate, $k^{p}_{L/T}$ the translation rate, $g^{m}_{L/T}$ the m*RNA* degradation rate and $g^{p}_{L/T}$ is the protein degradation rate. Parameters are listed in table 1.

Parameters	**Description**	**Value**
k^{m0}_{L}	Transcription rates (mRNA min^{-1})	$3.20e-2$
k^{m0}_{T}	-	$1.19e-1$
k^{m}_{L}	-	8.30
k^{m}_{T}	-	2.06
k^{p}_{L}	Translation rates (a.u. $mRNA^{-1}min^{-1}$)	$9.726e-1$
k^{p}_{T}	-	1.170
g^{m}_{L}	Degradation rates (min^{-1})	$1.386e-1$
g^{m}_{T}	-	$1.386e-1$
g^{p}_{L}	-	$1.65e-2$
g^{p}_{T}	-	$1.65e-2$
θ_{LacI}	*plac* regulation by TetR (-)	31.94
η_{LacI}	-	2.00
θ_{IPTG}	-	$9.06e-2$
η_{IPTG}	-	2.00
θ_{TetR}	*ptet* regulation by LacI (-)	30.00
η_{TetR}	-	2.00
θ_{aTc}	-	11.65
η_{aTc}	-	2.00
k^{in}_{IPTG}	IPTG exchange rate (min^{-1})	$2.75e-2$
k^{out}_{IPTG}	-	$1.11e-1$
k^{in}_{aTc}	aTc exchange rate (min^{-1})	$1.62e-1$
k^{out}_{aTc}	-	$2.00e-2$

Table 1: Parameters for the toggle switch model

The authors assumed that the repressors could be modelled with Hill function ($h(x,\theta,\eta) = 1/(1+x/\theta)^{\eta}$), where θ represents a threshold parameter for the Hill function and η is the Hill coefficient. Two more equations were added to take into account inducers diffusion through the cell membrane. Formally:

$$\frac{daTC}{dt} = \begin{cases} k^{in}_{aTc}(u_{aTc} - aTc), & if \quad u_{aTc} > aTc \\ k^{out}_{aTc}(u_{aTc} - aTc), & if \quad u_{aTc} \leqslant aTc \end{cases}$$

$$\frac{dIPTG}{dt} = \begin{cases} k^{in}_{IPTG}(u_{IPTG} - IPTG), & if \quad u_{IPTG} > IPTG \\ k^{out}_{IPTG}(u_{IPTG} - IPTG), & if \quad u_{IPTG} \leqslant IPTG \end{cases} \tag{3}$$

The deterministic version of the model does not capture the intrinsic stochasticity of biochemical processes. A more accurate description is provided by the stochastic modelling procedure, based on the pseudo-reactions listed in table 2.

Description	**Pseudo-reaction**
transcription	$\emptyset \xrightarrow{f^m_L(TetR,aTc)} mRNA_L$ $\emptyset \xrightarrow{f^m_T(LacI,IPTG)} mRNA_T$
translation	$mRNA_L \xrightarrow{k^p_L} mRNA_L + LacI$ $mRNA_T \xrightarrow{k^p_T} mRNA_T + TetR$
degradation/dilution	$mRNA_L \xrightarrow{g^m_L} \emptyset$ $mRNA_T \xrightarrow{g^m_T} \emptyset$ $LacI \xrightarrow{g^p_L} \emptyset$ $TetR \xrightarrow{g^p_T} \emptyset$

Table 2: Table of pseudo-reactions: here $f^m_L(TetR, aTc) = k^{m0}_L + k^m_L \cdot h(TetR \cdot h(aTc, \theta_{aTc}, \eta_{aTc}), \theta_{TetR}, \eta_{TetR})$ and $f^m_T(LacI, IPTG) = k^{m0}_T + k^m_T \cdot h(LacI \cdot h(IPTG, \theta_{IPTG}, \eta_{IPTG}), \theta_{LacI}, \eta_{LacI})$ are gene regulation function, and $h(x, \theta, \eta) = 1/(1 + x/\theta)^\eta$ is the decreasing Hill function. The others arrows superscripts are parameters that can be found in Table 1.

Specifically, we employed a SDE based model for the description of biochemical systems [44]:

$$dx = \underline{\underline{S}}\underline{f}(x)dt + \underline{\underline{S}}diag(\sqrt{\underline{f}(x)})d\underline{W} \tag{4}$$

Here:

- $\underline{\underline{S}}$ is the stoichiometry matrix where each row is associated to a state variable and each column i includes stoichiometry coefficients associated to the reaction i.
- $\underline{f}$ is a vector where each element i corresponds to a propensity function. The latter is a function describing the probability of a certain reaction given the concentration of the chemical species ([45]).
- $d\underline{W}$ is a vector containing the Wiener Process increments.

Pseudo-reactions and propensity functions were all given in [6], and here reported in Table 2.
All simulations were performed using Matlab R2019b. To solve the system (2) we used the function

ode45, while for the stochastic model in eq. (4) we implemented the Euler-Maruyama method [44], that is known to be slightly less accurate than the Stochastic Simulation Algorithm [45], but more computationally efficient.
Throughout the paper we have assumed aTc to be fixed to $25ngmL^{-}1$ for all simulations.

Control-based continuation theory: using control to track equilibria of the underlying uncontrolled system

CBC seeks to define a bifurcation diagram by the application of a noninvasive control action, that does not modify the underlying uncontrolled system's equilibria position in parameter space. To do so, the original method looks for a control action whose contribution vanishes asymptotically (eq. (1)). However, the same noninvasiveness can be achieved with a slightly simplified method, provided that the control action enters the system as bifurcation parameter [31]. The fundamental principles of the simplified CBC method used in this paper can be explained using the scalar differential equation:

$$\dot{x} = \mu + x^2 \tag{5}$$

where $\mu \in \mathbb{R}$ is the bifurcation parameter. Eq. (5) corresponds to the normal form of a fold (or saddle-node) bifurcation. The equilibria of the above system are given by the formula:

$$\bar{x} = \pm\sqrt{-\mu}. \tag{6}$$

To trace out the equilibrium curve of (5), including both stable and unstable equilibria, CBC relies on the presence of a stabilising feedback controller. The equation of motion of the system including feedback control is given by:

$$\dot{x} = \mu + x^2 + u \tag{7}$$

where u is the control signal given by a suitable, i.e. stabilizing, control law. When the control signal and bifurcation parameter affect the system in the same way (as in Eq. (7)), the static component of the control signal can be interpreted as a shift of the bifurcation parameter μ. In this paper, a simple linear proportional (P) law will be considered such that u is given by

$$u = K_{\mathrm{p}}(x_{\mathrm{ref}}(t) - x(t)). \tag{8}$$

where $x_{\mathrm{ref}}(t)$ is the control reference (or target) and K_p is the proportional gain. Using (8) as control input, (7) becomes:

$$\dot{x} = \mu + x^2 + K_{\mathrm{p}}(x_{\mathrm{ref}}(t) - x(t)) \tag{9}$$

that can be reordered as:

$$\dot{x} = \left(x^2 - K_{\mathrm{p}}x + (\mu + K_{\mathrm{p}}x_{\mathrm{ref}})\right). \tag{10}$$

Thus, we can write the equilibria of (10) and the control signal at steady state as:

$$\bar{x}_{\pm} = \frac{K_{\mathrm{p}} \pm \sqrt{K_{\mathrm{p}}^2 - 4(\mu + K_{\mathrm{p}}x_{\mathrm{ref}})}}{2} \tag{11}$$

$$\bar{u} = K_{\mathrm{p}}(x_{\mathrm{ref}} - \bar{x}_{\pm}). \tag{12}$$

Writing x_{ref} as function of the steady state control action u, we can substitute Eq. (12) in Eq. (11) to obtain:

$$\bar{x}_{\pm} = \frac{K_{\mathrm{p}} \pm \sqrt{K_{\mathrm{p}}^2 - 4(\mu + K_{\mathrm{p}}(\frac{\bar{u}}{K_{\mathrm{p}}} + \bar{x}_{\pm})}}{2} \tag{13}$$

$$2\bar{x}_{\pm} - K_{\mathrm{p}} = \pm\sqrt{K_{\mathrm{p}}^2 - 4(\mu + K_{\mathrm{p}}(\frac{\bar{u}}{K_{\mathrm{p}}} + \bar{x}_{\pm})} \tag{14}$$

$$(2\bar{x}_{\pm} - K_{\mathrm{p}})^2 = K_{\mathrm{p}}^2 - 4(\mu + K_{\mathrm{p}}(\frac{\bar{u}}{K_{\mathrm{p}}} + \bar{x}_{\pm}) \tag{15}$$

Solving the squared term and simplifying the equation we can thus obtain:

$$\bar{x}_{\pm} = \pm\sqrt{-(\mu + \bar{u})}. \tag{16}$$

Eq. (16) defines the same equilibrium curve as Eq. (6) for a parameter $\mu^* = \mu + \bar{u}$. Therefore, the steady state contribution of the control signal does not have to be removed and it can simply be considered as a shift in parameter space. The control strategy so defined is noninvasive (it does not modify the position of system's equilibria in parameter space) and the requirement of eq. (1) can be translated into:

$$\lim_{t \to +\infty} \dot{u}(t) \equiv 0. \tag{17}$$

which is equivalent to converging to a steady state value.
Here we demonstrated how CBC can be applied to track equilibria of a system by the application of a stabilizing proportional controller. However, the same results could be achieved with other control strategies as long as the controller stabilises the system's unstable equilibria. In fact, CBC is not restricted to a particular form of control and thus the control law can in principle take a general form:

$$u(t) = g(x_{\mathrm{ref}}(t) - x(t)). \tag{18}$$

Main steps of the control-based continuation algorithm

The algorithm applied for control-based continuation can be briefly summarized in the following steps:

1. Set a new control reference. A reference value $x_{ref} = TetR^*$ is used to evaluate the error (the difference between the measured $TetR$ and the reference target), and it is given to the controller. The range where the reference is picked normally covers the minimum and maximum expression of the protein of interest, but it can be modified according to the need (Fig. 2 B shows a maximum reference value of $TetR^* = 1800[a.u]$, although the maximum expression of $TetR$ is $1200[a.u]$).

2. Compute the control action. Depending on the control strategy, either the MPC performs optimization and defines the optimal control input or the proportional controller evaluates the control action depending on the error value.

3. Feed the control action to the system. The computed control action u is fed to the system for 5 minutes continuously, as it is considered to be the minimum sampling time for measurements and actuation in an hypothetical microfluidics/microscopy experimental set-up. After 5 minutes, the system output is measured and a new error is computed. Steps 2 and 3 are repeated until the process is considered at steady state.

4. Check for steady state. The initial implementation of the algorithm considers the system at steady state after a fixed amount of time ($9h$ and $55min$ for experiments shown in the Results section). The algorithm allows a variable time for steady state computation. In this case, to compute the steady state of $TetR$ and $IPTG$, the slope of the linear curve fitting the last 12 samples after 3 hours of simulation is calculated for both the error and the control signals (where the slope corresponds to the angle the curve makes with respect to the $x-axis$). If the computed value is below a user-defined tolerance, the system is assumed at steady state. Otherwise, the simulation continues and the steady state is checked every new sample, until the slope is sufficiently small. A threshold is set on both the error and the control slopes, which may vary between a deterministic and a stochastic simulation.
5. Acquire the steady state. The steady state values of $TetR$ and $IPTG$ ($\overline{TetR}$ and $\overline{IPTG}$) are saved as the average value computed over the last 12 samples of each signal, and the algorithm goes back to step 1.

All steps are repeated until enough steady states are collected, or until the range of output and parameter values has been covered. An initial guess is given as initial state for the algorithm to start with. A schematic of few steps can be seen in Fig. 1 B.

Results of the variable stepping reference algorithm can be seen in Fig. S1-S8 for both the deterministic and the stochastic scenarios.

Data availability

The code used in this publication is available at: https://github.com/lrenson/cbc-synbio-paper, https://github.com/diBernardoGroup/Control-Based-Continuation-of-a-genetic-Toggle-Switch. Movies S1, S2, S3 and S4 are provided as supplementary files.

Proportional Controller

The simplest controller that can satisfy CBC requirements is a proportional controller, formally described by the control law of Equation (8). The error signal is computed by subtracting the measured output to the control reference ($TetR^* - TetR$). The error is then fed to the controller to evaluate the control input $IPTG$ given to the toggle switch system. A schematic of the control loop can be seen in Fig. S11 A.
Despite its simplicity, the proportional controller has proved to be a solid choice for CBC. As long as the controller can stabilise the system, the control action is noninvasive and thus the points collected through the CBC routine correspond to the equilibria of the underlying uncontrolled system.
To tune the proportional gain K_p we linearized the system around an unstable equilibria and then we used the root locus to find the minimum gain able to stabilise the linearized system. Depending on the type of simulation, deterministic or stochastic, the K_p gain is then adjusted by trial and error.

Model Predictive Controller

Model predictive control is a control scheme based on two main features: prediction and optimization. At each step, a model reproducing the process behaviour is used to predict the process outputs to given

input signals. The input minimizing a cost function is computed using an optimization algorithm and fed to the controlled process. We chose a linear model to reproduce behaviours we expect from the real process. A further requirement for the prediction procedure is the process current state which, if not fully measurable, can be estimated using a Kalman filter. A schematic of the control loop can be seen in Fig. S9 B.

System identification to reproduce the process experimental behaviour through a simplified model

The MPC strategy requires a model to predict the process trajectories and carry out the optimization routine. We artificially reproduced the experimental behaviours using stochastic numerical simulations. Specifically, we measured the response of the system's outputs (*LacI* and *TetR*) to different randomly generated pulses of aTc and $IPTG$, both with fixed and variable amplitude. This data-set was used to identify a LTI system via noniterative subspace estimation algorithm, using the Matlab System Identification toolbox. The result was a continuous-time identified state space model of the following form:

$$\begin{aligned} \tfrac{dx}{dt}(t) &= Ax(t) + Bu(t) + Kd(t) \\ y(t) &= Cx(t) + Du(t) + d(t) \end{aligned} \tag{19}$$

where $x(t)$ is the state, $u(t)$ is the input to the system, $y(t)$ is the system output, $d(t)$ is the state disturbance and K is its associated matrix [46].
Here, A, B, C and K are free parameters to estimate, while D was set to 0. Equation (19) is called innovation form of the state-space model and the matrix K corresponds to the Kalman gain matrix associated to the identified system.

Definition of the Kalman predictor to reconstruct unmeasurable states

Considering the Toggle Switch system (2), we measure only *LacI* and *TetR* concentrations.
To estimate the remaining states we opted for a Kalman filter, defined as:

$$\begin{aligned} \hat{x}(t+1) &= A\hat{x}(t) + Bu(t) + K[y(t) - C\hat{x}(t) - Du(t)] \\ \hat{y}(t) &= C\hat{x}(t) + Du(t) \end{aligned} \tag{20}$$

where $\hat{x}$ is the state estimate, $\hat{y}$ is the output estimate and all the matrices A, B, C, D, K correspond to those obtained through the identification process.
Eq. (20) can be rewritten as follows:

$$\begin{aligned} \hat{x}(t+1) &= \begin{bmatrix} A - KC \end{bmatrix} \hat{x}(t) + \begin{bmatrix} B - KD & K \end{bmatrix} \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} = A_k \hat{x}(t) + B_k U_k \\ \hat{y}(t) &= C\hat{x}(t) + Du(t) \end{aligned} \tag{21}$$

The addition of a correction based on the measured outputs *LacI* and *TetR* ($U_k = [u(t); [LacI, TetR]]$) enables a better fit of the original process outputs [47].
The ability to predict the data recorded is evaluated using the percentage of the output variation that is reproduced by the model. This is defined as:

$$fit = \left(1 - \frac{|y - \hat{y}|}{|y - \bar{y}|}\right) \times 100 \tag{22}$$

where y is the measured output, $\hat{y}$ is the simulated predicted model output and $\bar{y}$ is the mean value of y. Fit percentages of *LacI* and *TetR* outputs for the system in eq. (21) are 96.4 and 92.2 respectively.

Definition of a cost function to compute an optimal control action

To compute the optimal control action we employed a genetic algorithm [48]. Commonly, MPC algorithms use quadratic cost functions weighting both the state of the system and the control action. Considering a scalar example, a quadratic cost function could be defined as:

$$J(i) = \sum_{k=1}^{N} \alpha e(i+k)^2 + \sum_{k=1}^{N} \beta u(i+k)^2 \tag{23}$$

Here, α and β correspond to weighting coefficients reflecting the relative importance of the error e and the control input u; N represents the prediction horizon, and i the current time. To adapt the cost function to CBC and guarantee noninvasiveness of the control input we initially replaced the control input u by its variation Δu as follows:

$$J(i) = \sum_{k=1}^{N} \alpha e(i+k)^2 + \sum_{k=1}^{N} \beta \Delta u(i+k)^2 \tag{24}$$

In this way, the optimization finds the best trade off between regulation accuracy and control variation. Although eq. (24) granted us good results in most of the simulations, the tuning of both α and β was difficult, as small changes into the CBC algorithm variables (such as the number of points collected during a single simulation) affected the quality of the results and demanded for retuning of the cost function parameters.

To simplify the cost function, we set $\beta = 0$ and $\alpha = N - k + 1$ and forced the optimizer to limit both the control input amplitude and variation at each iteration point. Formally we defined the following constraints:

$$\begin{gathered} 0 < u_{init}(t) < 1 \\ \bar{u}_{prev} - \gamma \bar{u}_{prev} < u_{curr}(t) < \bar{u}_{prev} + \gamma \bar{u}_{prev} \quad \forall u_{curr} \neq u_{init} \end{gathered} \tag{25}$$

where $u_{init}(t)$ is the control input during the first step of the algorithm; $u_{curr}(t)$ is the control input for all other steps except for the first one and $\bar{u}_{prev}$ is the control input steady state value, saved before stepping the reference in the algorithm.
The presence of external constraints on the control action allows to have a noninvasive contribution without the need for the $\beta \Delta u(i+k)^2$ term in the cost function. The parameter γ defines a percentage bandwidth that we consider acceptable for the control action to vary between. In other words, at each iteration of the CBC algorithm -by which we mean every time we change the reference signal- the control action cannot vary more than γ with respect to the previous registered steady state input. In our simulations γ was set to 0.3 for 30 points (Fig. 3, 5) and to 0.5 for 11 and 12 points (Fig. S2, S4). Other settings for γ might be found in Figures captions. Fewer points result in larger control reference steps, which can be associated with bigger steps in the control input, and for this reason γ can be more easily calibrated just by considering the amount of points to be collected in a simulation.

Parameters estimation process

For the estimation process, we calculated the analytical solutions of our toggle switch model (2) that will be used to build the cost function for parameter estimation. In particular, we want to define the

function:

$$IPTG = g(TetR, \Theta) \tag{26}$$

that has a unique real solution for each admissible value of $TetR$ and for a chosen set of parameters $\Theta = (klm_0, klm, \theta_{aTc}, \eta_{aTc}, \theta_{TetR}, \eta_{TetR}, ktm_0, ktm, \theta_{IPTG}, \eta_{IPTG}, \theta_{LacI}, \eta_{LacI}, klp, ktp)$, which is the subject of the estimation process. The full Equation (26) is presented in the Supplementary (S1). Here we considered the parameters (gtm, glm, glp, gtp) to be fixed to the values computed in [6] and therefore we did not estimate them.
Once we computed the analytical definition of $g(TetR, \Theta)$, we built the following cost function:

$$J = \sum_i (\min_k ||[g(\overline{TetR}, \Theta), \overline{TetR}]_k - [\overline{IPTG}, \overline{TetR}]_i||_2) \tag{27}$$

where $[\overline{IPTG}, \overline{TetR}]_i$ are the measured steady state values of control input and control output out of the CBC routine for the i point of the bifurcation curve. The k index looks for the estimated point $[g(\overline{TetR}, \Theta), \overline{TetR}]$ with minimum distance (L^2 norm) from the measured steady state considered by the index i. $\overline{TetR}$ is normalized between 0 and 1 to be comparable with $\overline{IPTG}$.
Minimizing the cost function (27) for all points collected during an experiment, it is possible to estimate the best set of parameters Θ that allows to characterize a model of the toggle switch reproducing the bistability behaviour seen via experiments. Once a set of parameters is found (using a genetic algorithm), it is possible to build the bifurcation curve by solving Equation (26) for fixed values of $TetR$. Sets Θ that gave complex values of $IPTG$ were directly discarded. Estimates of parameters for different experimental settings and percentage variation with respect to nominal values can be seen in Table S1.

In order to validate the results of the estimation using data collected with the CBC algorithm, a comparison with the parameter sweeps method was made. Parameter sweeps is a simple experimental design to collect data at different values of a parameter of interest and is commonly used for parameter estimation. Considering the example of the toggle switch, parameter sweeps is implemented registering the steady state signal $\overline{TetR}$ for fixed values of the input $\overline{IPTG}$. Collected points $(\overline{IPTG}, \overline{TetR})$ fall on the stable branches of our bifurcation curve, as shown in Fig. S11 A, B, C. This method does not allow to collect unstable steady states as there is no control over the system.
We showed that CBC grants a more robust estimation by collecting data in the unstable branch, which is not possible through simple parameter sweeps, and allows to retrieve the bistability of the system (see Fig. S9 D, E, F).
As highlighted in the main text, assuming an *a priori* knowledge of the system's bistability, we could improve the performances of our calibration by releasing the constraint on sets Θ that generate complex values of $IPTG$. Furthermore, more complex cost function could be considered (for example, the minimum distance of measured point from the calibrated bifurcation curve). Fig. S10 shows this latter improved case, with very good fitting performances compared to the initial estimation. Although the noticeable improvement, the modified cost function requires a much longer computational time and knowledge of the system of interest's dynamics, which is often not the case when CBC is applied experimentally.

Agent-based simulations

All agent based simulations were performed using BSim, an agent-based environment designed to simulate bacterial populations [30, 33]. Here, cell bio-mechanics are explicitly modelled, together

with cell growth and division. The spatial configuration and dynamics of the chemical inducers added in the culture environment are also simulated, introducing an extra time delay in the delivery of the inducer molecules to the cells. In addition, it is possible to mimic microfluidic-based experimental platforms, including related physical and technological constraints (e.g. dimensions and shape of microfluidic chambers, flush out of cells). Specifically, we can define the geometry and dimension of the microfluidic chamber, and the physical interactions of the cells with the device. We can account for cells movement and collision inside the growth chamber and simulate the flushing out of the cells from the chamber with their consequent removal from the set of analysable cells. The CBC algorithm was implemented by adding to the simulation environment both the proportional and the MPC controllers.
As previously done in Shannon et al. [22] and Salzano et al. [49], the simulation environment has been complemented with the SDE solver based on the Euler-Maruyama method (4).
BSim simulates experiments via a microfluidic-based system composed by a microfluidic device, a microscopy, a computer and an actuation system. Dimensions of the growth chamber were initially taken from Lugagne et al. [6] but then set to $16\mu M \times 1.5\mu M \times 1\mu M$ to reduce the simulation time. The proportional controller was implemented directly in Bsim, while the MPC was implemented in Matlab and externally called by the agent-based simulator. All bio-mechanical parameters of the cells were set to the same values used by Matyjaszkiewicz et al. [30]. Moreover, realistic constraints on the sampling and actuation time were considered, similarly to what already done for Matlab CBC simulations. More precisely, we enforced the state of the cells to be measured every 5 min and we assumed the input provided to be changed at most every 5 min. Finally, we assumed a 40 s delay on the actuation of the control inputs, accounting for the time the inducers would take to reach the microfluidics cell culture chambers in a physical experiment.
All the factors taken into account in BSim added additional sources of variability to the simulations, sometimes affecting the performance of the control algorithms, as discussed in the Results and Discussion sections.

Associated Content

Supporting Information. Figures of short CBC simulations in a deterministic setting (Fig. S1 and S2) and in a stochastic setting (Fig. S3 and S4) with both controllers and reduced points. Figures of short CBC simulations in a stochastic setting (Fig. S5 and S6) with both controllers and all points. Figures of short CBC simulations in a stochastic setting with BSIM (Fig. S7 and S8). Equilibrium curve computed for model parameters estimated with constraints on IPTG complexity using parameter sweep data, CBC with MPC data and CBC with P controller data (Fig. S9). Equilibrium curve computed for model parameters estimated without constraints on IPTG complexity using parameter sweep data, CBC with MPC data and CBC with P controller data (Fig. S10). Control Feedback Loops, both MPC and P, implemented for CBC (Fig. S11). Estimated parameters for the toggle switch model depending on the method used to collect data (Tab. S1). Full IPTG $= g(TetR, \Theta)$ equation for parameter estimation. CBC movies legends.

Abbreviation

CBC, control-based continuation; MPC, model predctive control; P, proportional control.

Author Information

Corresponding Authors

Ludovic Renson - *Department of Mechanical Engineering, Imperial College London, UK.* Co-senior author. Email: l.renson@imperial.ac.uk
Lucia Marucci - *Engineering Mathematics Department, School of Cellular and Molecular Medicine and BrisSynBio, University of Bristol, UK.* Co-senior author. Email: lucia.marucci@bristol.ac.uk

Authors

Irene de Cesare - *Engineering Mathematics Department, University of Bristol, UK.*
Davide Salzano - *Department of Electrical Engineering and Information Technologies, University of Naples Federico II, Italy.*
Mario di Bernardo - *Department of Electrical Engineering and Information Technologies, University of Naples Federico II, Italy.* Co-senior author.

Author Contributions

I.d.C. designed and carried out the simulations in Matlab2019b and wrote the manuscript with inputs from all the authors. D.S. developed the BSim code and wrote the agent-based simulations section. M.d.B supervised the initial development of CBC and the agent-based implementation. L.R. provided the initial framework for the CBC code. L.R. and L.M. conceived the project and supervised the entire work.

Declaration of Interests

Nothing to declare.

Acknowledgments

I.d.C. is supported by an EPSRC DTP Scholarship. L.M. is supported by the Medical Research Council grant MR/N021444/1, by the Engineering and Physical Sciences Research Council grants EP/R041695/1 and EP/S01876X/1, and by the European Union's Horizon 2020 under Grant Agreement No. 766840. L.R. is funded by a Research Fellowship from the Royal Academy of Engineering (RF1516/15/11).

References

[1] M Omar Din, Tal Danino, Arthur Prindle, Matt Skalak, Jangir Selimkhanov, Kaitlin Allen, Ellixis Julio, Eta Atolia, Lev S Tsimring, Sangeeta N Bhatia, et al. Synchronized cycles of bacterial lysis for in vivo delivery. Nature, 536(7614):81–85, 2016.

[2] Michael B Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. Nature, 403(6767):335–338, 2000.

[3] Lucia Marucci, David AW Barton, Irene Cantone, Maria Aurelia Ricci, Maria Pia Cosma, Stefania Santini, Diego di Bernardo, and Mario di Bernardo. How to turn a genetic circuit into a synthetic tunable oscillator, or a bistable switch. PloS one, 4(12):e8083, 2009.

[4] Jesse Stricker, Scott Cookson, Matthew R Bennett, William H Mather, Lev S Tsimring, and Jeff Hasty. A fast, robust and tunable synthetic gene oscillator. Nature, 456(7221):516–519, 2008.

[5] Marcel Tigges, Nicolas Dénervaud, David Greber, Joerg Stelling, and Martin Fussenegger. A synthetic low-frequency mammalian oscillator. Nucleic acids research, 38(8):2702–2711, 2010.

[6] 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(1):1–8, 2017.

[7] Beat P Kramer, Alessandro Usseglio Viretta, Marie Daoud-El Baba, Dominique Aubel, Wilfried Weber, and Martin Fussenegger. An engineered epigenetic transgene switch in mammalian cells. Nature biotechnology, 22(7):867–870, 2004.

[8] Ruben Perez-Carrasco, Chris P Barnes, Yolanda Schaerli, Mark Isalan, James Briscoe, and Karen M Page. Combining a toggle switch and a repressilator within the ac-dc circuit generates distinct dynamical behaviors. Cell systems, 6(4):521–530, 2018.

[9] Velia Siciliano, Filippo Menolascina, Lucia Marucci, Chiara Fracassi, Immacolata Garzilli, Maria Nicoletta Moretti, and Diego di Bernardo. Construction and modelling of an inducible positive feedback loop stably integrated in a mammalian cell-line. PLoS Comput Biol, 7(6):e1002074, 2011.

[10] Timothy S Gardner, Charles R Cantor, and James J Collins. Construction of a genetic toggle switch in escherichia coli. Nature, 403(6767):339–342, 2000.

[11] Gábor Balázsi, Alexander van Oudenaarden, and James J Collins. Cellular decision making and biological noise: from microbes to mammals. Cell, 144(6):910–925, 2011.

[12] Lucia Marucci. Nanog dynamics in mouse embryonic stem cells: results from systems biology approaches. Stem cells international, 2017, 2017.

[13] Ann C Babtie, Paul Kirk, and Michael PH Stumpf. Topological sensitivity analysis for systems biology. Proceedings of the National Academy of Sciences, 111(52):18507–18512, 2014.

[14] Victoria Hsiao, Anandh Swaminathan, and Richard M Murray. Control theory for synthetic biology: recent advances in system characterization, control design, and controller implementation for synthetic biology. IEEE Control Systems Magazine, 38(3):32–62, 2018.

[15] Jan Sieber and Bernd Krauskopf. Control based bifurcation analysis for experiments. Nonlinear Dynamics, 51(3):365–377, 2008.

[16] David AW Barton, Brian P Mann, and Stephen G Burrow. Control-based continuation for investigating nonlinear experiments. Journal of Vibration and Control, 18(4):509–520, 2012.

[17] Ludovic Renson, David AW Barton, and Simon A Neild. Experimental tracking of limit-point bifurcations and backbone curves using control-based continuation. International Journal of Bifurcation and Chaos, 27(01):1730002, 2017.

[18] Ludovic Renson, AD Shaw, DAW Barton, and SA Neild. Application of control-based continuation to a nonlinear structure with harmonically coupled modes. Mechanical Systems and Signal Processing, 120:449–464, 2019.

[19] 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 Comput Biol, 10(5):e1003625, 2014.

[20] Lorena Postiglione, Sara Napolitano, Elisa Pedone, Daniel L Rocca, Francesco Aulicino, Marco Santorelli, Barbara Tumaini, Lucia Marucci, and Diego di Bernardo. Regulation of gene expression and signaling pathway activity in mammalian cells by automated microfluidics feedback control. ACS synthetic biology, 7(11):2558–2565, 2018.

[21] Elisa Pedone, Lorena Postiglione, Francesco Aulicino, Dan L Rocca, Sandra Montes-Olivas, Mahmoud Khazim, Diego di Bernardo, Maria Pia Cosma, and Lucia Marucci. A tunable dual-input system for on-demand dynamic gene expression regulation. Nature communications, 10(1):1–13, 2019.

[22] 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.

[23] Jannis Uhlendorf, Agnès Miermont, Thierry Delaveau, Gilles Charvin, François Fages, Samuel Bottani, Gregory Batt, and Pascal Hersen. Long-term model predictive control of gene expression at the population and single-cell levels. Proceedings of the National Academy of Sciences, 109(35):14271–14276, 2012.

[24] Elisa Pedone, Irene De Cesare, Criseida G Zamora-Chimal, David Haener, Lorena Postiglione, Antonella La Regina, Barbara Shannon, Nigel J Savery, Claire S Grierson, Mario Di Bernardo, et al. Cheetah: a computational toolkit for cybergenetic control. ACS Synthetic Biology, 10(5):979–989, 2021.

[25] Irene De Cesare, Criseida G Zamora-Chimal, Lorena Postiglione, Mahmoud Khazim, Elisa Pedone, Barbara Shannon, Gianfranco Fiore, Giansimone Perrino, Sara Napolitano, Diego di Bernardo, et al. Chipseg: An automatic tool to segment bacterial and mammalian cells cultured in microfluidic devices. ACS omega, 6(4):2473–2476, 2021.

[26] Mahmoud Khazim, Elisa Pedone, Lorena Postiglione, Diego di Bernardo, and Lucia Marucci. A microfluidic/microscopy-based platform for on-chip controlled gene expression in mammalian cells. In Synthetic Gene Circuits, pages 205–219. Springer, 2021.

[27] Mahmoud Khazim, Lorena Postiglione, Elisa Pedone, Dan L Rocca, Carine Zahra, and Lucia Marucci. Towards automated control of embryonic stem cell pluripotency. IFAC-PapersOnLine, 52(26):82–87, 2019.

[28] Domitilla Del Vecchio, Aaron J Dy, and Yili Qian. Control theory meets synthetic biology. Journal of The Royal Society Interface, 13(120):20160380, 2016.

[29] Min Wu, Ri-Qi Su, Xiaohui Li, Tom Ellis, Ying-Cheng Lai, and Xiao Wang. Engineering of regulated stochastic cell fate determination. Proceedings of the National Academy of Sciences, 110(26):10610–10615, 2013.

[30] 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.

[31] David AW Barton and Jan Sieber. Systematic experimental exploration of bifurcations with noninvasive control. Physical Review E, 87(5):052916, 2013.

[32] Sandor Beregi, David AW Barton, Djamel Rezgui, and Simon A Neild. Robustness of nonlinear parameter identification in the presence of process noise using control-based continuation. Nonlinear Dynamics, 104(2):885–900, 2021.

[33] Thomas E Gorochowski, Antoni Matyjaszkiewicz, Thomas Todd, Neeraj Oak, Kira Kowalska, Stephen Reid, Krasimira T Tsaneva-Atanasova, Nigel J Savery, Claire S Grierson, and Mario Di Bernardo. Bsim: an agent-based tool for modeling bacterial populations in systems and synthetic biology. 2012.

[34] 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.

[35] Eric Walter and Luc Pronzato. Identification of parametric models: from experimental data. Springer Verlag, 1997.

[36] Lennart Ljung. Perspectives on system identification. Annual Reviews in Control, 34(1):1–12, 2010.

[37] Alejandro F Villaverde, Antonio Barreiro, and Antonis Papachristodoulou. Structural identifiability of dynamic systems biology models. PLoS computational biology, 12(10):e1005153, 2016.

[38] 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.

[39] Jan Sieber, Bernd Krauskopf, David Wagg, Simon Neild, and Alicia Gonzalez-Buelga. Control-based continuation of unstable periodic orbits. Journal of Computational and Nonlinear Dynamics, 6(1), 2011.

[40] Ludovic Renson, Jan Sieber, David AW Barton, AD Shaw, and SA Neild. Numerical continuation in nonlinear experiments using local gaussian process regression. Nonlinear Dynamics, 98(4):2811–2826, 2019.

[41] 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, 2016.

[42] Yang Li and Harry Dankowicz. Adaptive control designs for control-based continuation in a class of uncertain discrete-time dynamical systems. Journal of Vibration and Control, page 1077546320913377, 2020.

[43] Frederick K Balagaddé, Hao Song, Jun Ozaki, Cynthia H Collins, Matthew Barnet, Frances H Arnold, Stephen R Quake, and Lingchong You. A synthetic escherichia coli predator–prey ecosystem. Molecular systems biology, 4(1):187, 2008.

[44] Eszter Lakatos. Stochastic analysis and control methods for molecular cell biology. 2017.

[45] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. The journal of physical chemistry, 81(25):2340–2361, 1977.

[46] MATHWORKS Repository. https://it.mathworks.com/help/ident/ug/identifying-state-space-models-with-independent-process-and-measurement-noise.html.

[47] MATHWORKS Repository. https://it.mathworks.com/help/ident/ug/definition-simulation-and-prediction.html.

[48] Lawrence Davis. Handbook of genetic algorithms. 1991.

[49] Davide Salzano, Davide Fiore, and Mario di Bernardo. Controlling reversible cell differentiation for labor division in microbial consortia. bioRxiv, 2021.