In Vivo Real-Time Control of Gene Expression: A Comparative Analysis of Feedback Control Strategies in Yeast

Gianfranco Fiore,†‡§ Gian Simone Perrino,†‡§ Mario di Bernardo*,†‡§ and Diego di Bernardo*†‡§

†Tele’thon Institute of Genetics and Medicine (TIGEM), 80078 Pozzuoli, Italy
‡Department of Electrical Engineering and Information Technology, University of Naples Federico II, 80125 Naples, Italy
§Department of Engineering Mathematics, University of Bristol, Bristol BS8 1TH, United Kingdom

ABSTRACT: Real-time automatic regulation of gene expression is a key technology for synthetic biology enabling, for example, synthetic circuit’s components to operate in an optimal range. Computer-guided control of gene expression from a variety of inducible promoters has been only recently successfully demonstrated. Here we compared, in silico and in vivo, three different control algorithms: the Proportional-Integral (PI) and Model Predictive Control (MPC) controllers, which have already been used to control gene expression, and the Zero Average Dynamics (ZAD), a control technique used to regulate electrical power systems. We chose as an experimental testbed the most commonly used inducible promoter in yeast: the galactose-responsive GAL1 promoter. We set two control tasks: either cells to express a desired constant fluorescence level of a reporter protein downstream of the GAL1 promoter (set-point) or a time-varying fluorescence (tracking). Using a microfluidics-based experimental platform, in which either glucose or galactose can be provided to the cells, we demonstrated that both the MPC and ZAD control strategies can successfully regulate gene expression from the GAL1 promoter in living cells for thousands of minutes. The MPC controller can track fast reference signals better than ZAD but with a higher actuation effort due to the large number of input switches it requires. Conversely, the PI controller’s performance is comparable to that achieved by the MPC and the ZAD controllers only for the set-point regulation.

KEYWORDS: synthetic biology, control engineering, microfluidics, gene expression, yeast

Control Engineering aims at driving a physical system in order to attain a specific value of a quantity of interest (such as a boiler that needs to warm water to a desired temperature, or a car cruise-control maintaining a constant speed) despite the presence of disturbances. This is achieved by appropriately varying its inputs (switch on or off a heater in the case of the boiler, or accelerating or braking in the cruise-control) as a function of the difference between the measured value of the output and its desired target value (control error). At the core of most control schemes lies a negative feedback loop, as shown in Figure 1A. The variable to be controlled (system output $y$) is measured, and its value is subtracted from the desired value (control reference $r$). The quantity that is obtained, the control error $e$, is minimized by the controller, a set of logical and mathematical rules through which an appropriate value of the input $u$ is chosen in order to guarantee that the output $y$ matches the desired reference $r$.

Feedback control has been extensively applied to control growing conditions of cells in chemostats in terms of temperature and/or CO$_2$, and it is a current feature of benchtop and industrial chemostats.1,2,3 Only recently, however, the application of Control Engineering principles has been exploited to regulate molecular events in living cells, thanks to innovative microfluidics and optogenetics platforms.4–8 In refs 4,5, we built a completely automated microfluidic platform to control in real-time gene expression in yeast cells. We demonstrated the ability of the platform to reach and maintain a desired value of gene expression, measured in terms of the fluorescence intensity of a reporter protein expressed from the endogenous GAL1 promoter.

Other successful attempts to control gene expression, or even signaling pathways, have been described in the literature. They mainly differ in the control input (osmotic pressure, light, small-molecules) and the control strategy adopted. Optogenetics-based light inducible systems have been exploited to control gene expression in yeasts,9,10 to regulate intracellular signaling dynamics in mammalian cells,7 and to drive protein levels by using light-switchable two-component systems in bacteria.10 Microfluidic-based devices, allowing a tight control of cellular growing medium and the administration of inducer small-molecules, have been successfully employed to investigate synchronization properties of synthetic biological clocks in bacterial cells,11 to control the transcription from the HOG1 promoter in yeast S. cerevisiae by varying the osmotic pressure8

Received: July 23, 2015

© XXXX American Chemical Society

DOI: 10.1021/acs.synbio.0b00135
ACS Synth. Biol. XXXX, XXX, XXX–XXX
Yeast cells will first consume all the available glucose in the medium before starting metabolizing galactose. Hence, the control input can either be glucose (switch off signal) or galactose (switch on signal), but not an intermediate concentration of the two, because cells will not respond to galactose when glucose is present.

We thus decided to use the GAL1 promoter upstream of a reporter gene (Gfp fused with the Gal1 protein) as a testbed for comparing and assessing the performance of the different control strategies. When dealing with living cells, one of the major issues is represented by the uncertainty affecting transcriptional and translational processes, introducing a remarkable cell-to-cell variability in mRNA and protein production. Rather than trying to control stochastic behavior of cells, here we addressed the simpler problem of regulating the average fluorescence intensity expressed by all cells as the quantity to be controlled \( \langle y \rangle \), thus averaging out the effects due to intrinsic and extrinsic sources of noise.

To carry out in vivo control experiments, we used the same integrated experimental setup presented in previous works, comprising a microfluidic device, a time-lapse microscope, and a set of automated syringes, all controlled by a computer. As depicted in Figure 1B, the computer runs the control algorithm, which at each sampling interval: (i) processes images acquired by the microscope to estimate the fluorescence \( \langle y \rangle \); (ii) executes the control algorithm to derive the input \( u \) for the next sampling period; (iii) controls the automated syringes to provide the calculated input (i.e., galactose or glucose) to the cells. We already demonstrated that the average fluorescence level of a yeast population can be effectively regulated with this platform using a simple Proportional-Integral control strategy.

Controlling Gene Expression from the GAL1 Promoter: Set-Point and Tracking Control Tasks. We compared the performance of three control algorithms (PI, MPC and ZAD) when performing two different tasks, as shown in Figure 2: (i) set-point control, where the average Gfp fluorescence must reach and maintain a desired reference level, and (ii) signal tracking control where the average fluorescence must follow (or track) a desired time-varying signal. Specifically, in the set-point control (Figure 2A), the desired fluorescence \( r \) was set equal to 50% of the initial average fluorescence expressed by the cells during the calibration phase of 180 min. During the calibration phase, cells are kept in galactose, in order to let cells adapt to the microfluidic environment, and to set the unit of measure of fluorescence, which may vary due to technical and biological variability in each experiment. In the signal tracking control, we used three different references: (i) a descending staircase function (Figure 2B) where each step lasts 500 min, beginning at 75% of the calibration phase average fluorescence, then stepping down to 50%, and then 25%; (ii) a linear descending ramp of 1500 min (Figure 2C) starting at the 100% of average fluorescence measured in the calibration phase, and decreasing down to 25%; and (iii) a sinusoidal wave of period \( T = 2000 \) min (Figure 2D) defined as \( z(t) = 0.5 + 0.25 \sin \left( \frac{2 \pi}{T} (t - 100) \right) + \frac{1}{2} \).

Control Algorithms. PI and MPC have been previously applied to control gene expression and protein activation. Toettcher and colleagues applied a Proportional-Integral (PI) control to regulate protein signaling in mammalian cells using light as control input in an optogenetics framework; we have applied the same PI control scheme to regulate gene expression from the GAL1 promoter in yeast using galactose and glucose.
The proportionality constants to the sum of the past values of the error (the integral term).

HOG1 promoter in yeast using osmotic pressure as the control input. We therefore compared the performance of PI, MPC, and a new ZAD controller when applied to the regulation of gene expression from the GAL1 promoter in yeast cells.

We identified two major constraints affecting the control algorithms: the sampling-time and the admissible values of the control input. We set the sampling time \( T = 5 \) min; this is the time interval at which images are acquired from the microscope, and it is an ideal trade off to avoid phototoxicity and capture the dynamics of the Gfp protein expression. The control input \( u \) can assume only two values (Galactose-ON, Glucose-OFF). Thus, at each sampling time \( kT \), the control algorithms can only choose the duration of galactose pulse (ON), which can vary from 0 to 5 min, and it is defined as the duty-cycle \( d \) (i.e., the percentage of the time interval during which galactose is provided to the cells).

The Proportional-Integral (PI) control algorithm uses the error \( e(t) = r(t) - y(t) \) to choose, at each sampling time \( kT \), the duty-cycle value \( d_k \). Specifically \( d_k \) has a value proportional to the weighted sum of two contributions, one proportional to the actual error \( e(t) \) and the other proportional to the sum of the past values of the error (the integral term). The proportionality constants \( K_r \) and \( K_i \) are called, respectively, proportional and integral gains, and their values were chosen by simple empirical rules (Methods and Supporting Information).

The Model Predictive Control (MPC) algorithm is an optimization-based technique which uses a mathematical model of the process being controlled to predict the future values of the control error and to find the best value of the duty-cycle value \( d_k \) that minimizes it (Methods and Supporting Information).

The Zero Average Dynamics (ZAD) algorithm relies on a feedback strategy devised for the regulation of power converters, and it is a modified version of Sliding Mode Control. Specifically, the ZAD calculates, at each sampling time, the best value of \( d_k \) which minimizes the actual control error \( e(t) \) and its predicted future value (estimated by the derivative \( \dot{e}(t) \) over the next time interval (refer to Methods section and Supporting Information for further details).

**Set-Point Control Experiments.** We first tested in silico the PI, MPC, and ZAD control strategies described above, by simulating the behavior of yeast cells in a computer (Methods and Supporting Information). In silico results are shown in Figure 3. PI (Figure 3A), MPC (Figure 3B) and ZAD (Figure 3C) are able to reach and maintain the desired fluorescence value without exhibiting oscillations at steady-state. Performance indices (ISE, IAE, ITAE in Figure 3D) are of the same order of magnitude for all the control strategies; interestingly, the ZAD controller is able to achieve satisfying results with a reduced number of input switches (5- and 6-fold less than respectively MPC and PI). This is advantageous in the experimental setting because it reduces unnecessary stress to cells.

In vivo control experiments, shown in Figure 4, mirror in silico results, showing that the three strategies are indeed all able to reach and maintain the desired fluorescence level. As predicted by the in silico simulations, the ZAD controller employs fewer galactose pulses (Figure 4C) and displays smaller oscillations around the set-point than the MPC feedback strategy (Figure 4B).

**Signal Tracking Control Experiments.** In silico simulation of the descending staircase reference shows that the three control strategies have different performances. The PI is not able to properly follow the reference signal (Figure 5A). This is to be expected, because the PI controller was designed specifically to solve set-point control tasks. The MPC control algorithm, with its intrinsic predictive ability, achieves a good performance, specifically noticeable in the proximity of the steps’ edges (Figure 5B). Indeed, the MPC is able to foresee changes in the reference signal and to adjust the control input accordingly, by starting to "switch off" the system in advance. The ZAD control algorithm (Figure 5C) achieves satisfying results, comparable to the MPC controller (except in the
proximity of the steps’ edges), but with a smaller number of galactose pulses.

In vivo experiments for the descending staircase reference (Figure 6) confirm in silico results. The PI controller (Figure 6A) poorly tracks the reference r, despite the high number of input switches. The MPC, as already demonstrated in silico, has a much better performance, quantitatively confirmed by the performance indices (Figure 6 B,D). As in the case of the in silico simulations, the ZAD controller (Figure 6B,D) achieves a performance comparable to that of the MPC (even if not as good in the proximity of the steps’ edges) employing fewer galactose pulses than the MPC.

Because of the poor tracking results achieved by the PI controller, we further compared only the MPC and ZAD controllers.
Figure 5. In silico staircase tracking control task. The blue line is the reference signal \( r \). The green line is the simulated fluorescence level \( y \). The red line is the control input \( u \). (A–C) Three in silico staircase tracking control experiments performed on the GALI promoter mathematical model by the means of the PI (A), MPC (B), and ZAD (C) controllers. The initial level of fluorescence is assumed to be equal to 1. The control action starts at time \( t = 0 \) min and ends at \( t = 1000 \) min. (D) Performance indices: Integral Square Error (ISE), Integral Absolute Error (IAE), Integral Time Absolute Error (ITAE), number of switches of the control input, and the percentage of time during which the model is provided with the “ON” input.

<table>
<thead>
<tr>
<th>Controller</th>
<th>ISE</th>
<th>IAE</th>
<th>ITAE</th>
<th>Switches</th>
<th>Time in GaL(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI</td>
<td>2.68</td>
<td>19.38</td>
<td>5.83E3</td>
<td>404</td>
<td>42.67</td>
</tr>
<tr>
<td>MPC</td>
<td>1.02</td>
<td>7.31</td>
<td>2.20E3</td>
<td>358</td>
<td>41.87</td>
</tr>
<tr>
<td>ZAD</td>
<td>2.05</td>
<td>12.36</td>
<td>3.70E3</td>
<td>60</td>
<td>42.40</td>
</tr>
</tbody>
</table>

Figure 6. In vivo staircase tracking control task. The black line is the average fluorescence intensity during the calibration phase of 180 min. The blue line is the reference signal \( r \). The green line is the measured fluorescence level \( y \) across the yeast population. The red line is the control input \( u \). (A–C) Three in vivo staircase tracking control experiments performed by the means of the PI (A), MPC (B), and ZAD (C) controllers. The control action starts at time \( t = 0 \) min and ends at \( t = 1000 \) min. (D) Performance indices: Integral Square Error (ISE), Integral Absolute Error (IAE), Integral Time Absolute Error (ITAE), number of switches of the control input, and the percentage of time during which the cells are provided with the “ON” input.

<table>
<thead>
<tr>
<th>Controller</th>
<th>ISE</th>
<th>IAE</th>
<th>ITAE</th>
<th>Switches</th>
<th>Time in GaL(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI</td>
<td>1.71</td>
<td>18.18</td>
<td>5.11E3</td>
<td>300</td>
<td>46.05</td>
</tr>
<tr>
<td>MPC</td>
<td>1.23</td>
<td>12.04</td>
<td>3.62E3</td>
<td>70</td>
<td>42.53</td>
</tr>
<tr>
<td>ZAD</td>
<td>1.43</td>
<td>13.54</td>
<td>4.08E3</td>
<td>94</td>
<td>35.67</td>
</tr>
</tbody>
</table>
controllers when tracking the ramp and the sinusoidal signal. Both in silico (Figure 7) and in vivo (Figure 8) observations confirm that the ZAD controller is able to guarantee a performance (Figure 7E and Figure 8E), similar to that of the MPC strategy, but again with a reduced number of control input switches.

Conclusions. Precise and quantitative regulation of gene expression from inducible promoters is a key technology for current and future Synthetic Biology applications. It can be used to quantitatively characterize biological "parts" in a single experiment by generating a desired time-varying concentration of an effector protein and measuring the activity of the target. For example, the level of a Transcription Factor can be controlled to follow a descending staircase reference while, at the same time, measuring the level of a report protein downstream of the promoter to be characterized, in order to derive a quantitative dose–response curve to be used for modeling.

A second application of automatic control of gene expression is to ensure that a synthetic circuit works in the optimal operating conditions in terms of expression of its constituent proteins, similarly to what happens in a modern computer where the operating temperature is automatically controlled by switching on/off a fan in order to keep the electronic circuits from overheating.

Here we provided a comparative analysis, in silico and in vivo, of three different strategies to control gene expression from the GAL1 inducible promoter, whose advantages and disadvantages are summarized in Table 1.

To this end we implemented and compared PI and MPC controllers, which have been previously reported in the literature and proposed an additional strategy, the ZAD controller.

We demonstrated that both MPC and ZAD control strategies can be successfully employed to control gene expression from the GAL1 promoter to generate any desired time-varying concentration of the reporter protein. These controllers require a quantitative model of the system to be controlled. This is not a strong limitation, because it is possible to identify a dynamical input-output model of the biological system under investigation using standard system identification techniques, which work very well at least for simple inducible promoters.
The PI controller, as expected from control theory and from our *in silico* predictions, performs similarly to the MPC and ZAD strategies only in the set-point control task, whereas it is the worst performer in the case of signal-tracking experiments. The MPC and ZAD controller perform similarly well in all the control tasks. The main differences are that the MPC performs slightly better than ZAD for fast varying references (such as the staircase signal in Figure 6); however, it requires a higher number of input switches when compared to the ZAD controller. The ZAD technique may be advantageous in those applications in which a high cost is associated with the actuation such as when the input administration can cause stress to the cells (e.g., light stimuli, antibiotic, osmotic shocks, etc.).

In conclusion, automatic control of gene expression from inducible promoters is mature enough to be applied routinely in synthetic biology and more generally in quantitative biology applications. Although we showed the experimental application of these control strategies to the GAL1 promoter, the same techniques can be applied to other inducible promoters and to different cellular models.

The choice of the control strategy to employ will depend on which kind of control task needs to be achieved (set-point or tracking), the complexity of the synthetic circuit to be controlled, the availability of a descriptive mathematical model of the circuit, and the cost associated with the actuation. The PI controller, as expected from control theory and from our *in silico* predictions, performs similarly to the MPC and ZAD strategies only in the set-point control task, whereas it is the worst performer in the case of signal-tracking experiments.
the circuit to be controlled, the cost associated with the actuation effort and, whether a minimal stress to the cells is required (i.e., a small number of input switches).

**METHODS**

**Yeast Strain and Cell Culture.** Control experiments were performed in the yeast strain (yGIL337, Gal1-GFP::KanMX, Gal10-mCherry::NatMX) provided to us by Lang et al. In this strain, the Gal1 protein, expressed by the GAL1 promoter, was fused to a green fluorescent protein (Gfp). Before each in vivo control experiment started, cells were inoculated overnight in 10 mL synthetic complete medium + galactose/raffinose (2%); the culture was then repeatedly diluted to achieve a final OD$_{600}$ of 0.01 on the day the cells were injected into the microfluidic device (Supporting Information for further details).

**Microfluidics.** The MFD0050sa device was used for all the microfluidics experiments. This device houses a microchannel (height: 3.5 μm) which “traps” yeast cells, that can only grow in a monolayer, thus allowing easier automated image analysis. Microfluidics devices were fabricated as described in ref 17. Details of the microfluidics setup and of the galactose and glucose growing medium used in the experiments can be found in ref 5 and Supporting Information.

**Microscopy and Image Analysis.** To monitor cellular processes dynamics, as well as to check for the right administration of external inputs to trapped cells, we took advantage of an inverted fluorescence microscope (Nikon Eclipse Ti) equipped with an automated and programmable stage, an incubator to guarantee fixed temperature and gases to cell environment and a high sensitivity Electron Multiplying CCD (EMCCD) Camera (Andor iXON Ultra897). The microscope and the camera were programmed to acquire, at 5 min intervals, two types of images: (a) a bright field image (phase contrast) and (b) fluorescence images (with the appropriate filters) to monitor cell fluorescence and to track the dye (sulforhodamine) added to the inducer compound in order to evaluate the control input administered to the cells. Fluorescence was quantified using a previously developed custom image processing algorithm. The algorithm is able to locate cells within each Phase Contrast image thus identifying all pixels belonging to the cells. This information is used to calculate the fluorescence being expressed by the entire population as well as by each single cell. The actual measured fluorescence is mainly affected by the efficiency of the fluorescent lamp and by the background light in the surrounding microscope environment. We observed that as the fluorescent lamp nears its lifetime, its brightness decreases and this affects the measured fluorescence in the cells. Indeed, the measurement units for the fluorescence are considered arbitrary, and thus, a calibration phase at the beginning of each experiment is needed to calculate a reference value for the fluorescence.

**Control Strategies Implementation.** The control input is described as follows, where ON means galactose administration and OFF glucose administration:

$$u(t) = \begin{cases} u_{\text{ON}} = \text{ON} & \text{kT} \leq t < (k + d_k)T \\ u_{\text{OFF}} = \text{OFF} & (k + d_k)T \leq t < (k + 1)T \end{cases}$$

Pl. Proportional and integral gains, $K_p$ and $K_i$, were calculated with the Ziegler–Nichols' open-loop tuning method applied to the mathematical model of the GAL1 promoter described in the Supporting Information. Thus, the gains were set to $K_p = 13.49$ and $K_i = 0.17$.

Given the constraints on the control input as well as on the sampling time described above, a modulation on the PI output was implemented to calculate the duty cycle $d_k$ as

$$d_k = \frac{\hat{u} - u_{\text{min}}}{u_{\text{max}} - u_{\text{min}}}$$

where $\hat{u}$ is the output of the PI regulator saturated between $u_{\text{min}} = 0$ and $u_{\text{max}} = 2$. To avoid delays and overshoots introduced by the saturation of the regulator output, an antiwindup block, described in the Supporting Information, was added to the feedback loop.

**MPC.** The MPC strategy chooses, at each sampling time $kT$, the optimal control input that minimizes the sum of the squared control error (SSE):

$$\text{SSE} = \sum_{i=1}^{kN} (N + 1 + k - i)c_i^2 = \sum_{i=1}^{kN} (N + 1 + k - i)(\hat{y} - y)_i^2$$

where $\hat{y}$ is the output provided by the dynamical model of the GAL1 promoter (Supporting Information), which is computed by a Kalman state estimator, able to reconstruct system states from the measured output $y$, as shown Figure 1A. The integer $N = 12$ (corresponding to 60 min) defines the length of the prediction horizon in terms of sampling intervals. The forgetting factor $(N + 1 + k - i)$ weights the error samples more at the beginning of the prediction horizon than at the end; this guarantees faster corrections of output deviations from the reference. The optimization was carried out by adopting the Matlab implementation of the Genetic Algorithm described in ref 22.

The result of the optimization is an array of $N$ optimal duty cycles $d_{k,i}$, $i \in [1, N]$. In the absence of external disturbances and other sources of uncertainty, the optimal input computed by the MPC could, in principle, be applied to the yeast cells over the entire prediction horizon. However, in order to make the control action robust to any source of uncertainty and variability, the feedback loop is closed by applying only the first element of the calculated control input and at the next sampling time $(k + 1)T$ when the entire procedure is repeated.

**ZAD. Zero Average Dynamics (ZAD) control relies on a feedback strategy devised for the regulation of power converters and allows to directly calculate the duty cycle $d_k$ of a switching control input.** ZAD control is a practical implementation of sliding mode control, where the control objective consists in attracting and then maintaining onto a fixed surface $s(x) = 0$ (denoted as the sliding surface) the states of the system by appropriately switching the available inputs. In the ZAD control approach, the sliding condition has to be fulfilled only on average over each sampling period $kT$, thus allowing to directly calculate the duty cycle $d_k$ via the solution of the following integral equation:

$$E_T[s(x(t))] = \frac{1}{T} \int_{t=1}^{(k+1)T} s(x(t))dt = 0$$

where mathbbE_T indicates the operator taking the average over the time interval T.

To control GAL1 promoter dynamics onto the desired reference signal, we considered the following sliding surface, which was derived using the dynamical model of the GAL1 promoter as described in the Supporting Information:

$$s(x(t)) = (x_2(t) - x_{\text{set}}(t)) + (\dot{x}_2(t) - \dot{x}_{\text{set}}(t))$$

H

DOI: 10.1021/acssynbio.5b00135

ACS Synth. Biol. XXX, XXX, XXX--XXX
where \( x_t \) is the state variable describing the dynamics of the fluorescent reporter note that \( \dot{x}_t (t) = 0 \) in the case of set-point regulation. For further details on the implementation of the ZAD controller refer to Supporting Information.

**ASSOCIATED CONTENT**

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.5b00135.

Additional text and figures referenced in this article (PDF)

**AUTHOR INFORMATION**

**Corresponding Authors**

*E-mail: mario.dibernardo@unina.it.*

*E-mail: dibernardo@tigem.it.*

**Author Contributions**

G.F. and G.P. contributed equally to this work

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

We are thankful to Prof. Botstein for providing yeast strains and Prof. Jeff Hasty for microfluidic device. This work was supported by a Human Frontier Science Program HFSP RGP0020/2011 to DdB and by the Italian Fondazione Telethon.

**REFERENCES**


