2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Selection of control inputs for enhancing the long-term performance of synthetic gene circuits

Daniel P. Byrom, Declan G. Bates and Alexander P. S. Darlington

***Abstract*— Engineered biological systems fail when mutations arise that inhibit their intended function. Such mutations introduce uncertainty into the system parameters, making negative feedback control an attractive strategy to improve the evolutionary longevity of synthetic gene circuits. Here we propose three classes of controller that improve evolutionary performance in repeated batch culture. Each controller takes a different biological input: (i) the gene product in the cell, (ii) the host growth rate and (iii) the total population output. Using a multi-scale model of host-circuit interactions, we demonstrate that these different modes of action differentially influence the growth dynamics between mutant strains, driving significant differences in the long-term performance of gene circuits. We show that population-based feedback is least effective and that, whilst direct feedback inhibition is effective in the short-term, growth-based feedback can enhance long-term performance to a greater degree. We propose a novel control strategy which combines two of the strategies and improves both short- and long-term performance.**

I. Introduction

The goal of synthetic biology is to engineer living systems with novel and useful functionalities. Genetic programs (commonly referred to as "gene circuits") can be encoded in DNA molecules and engineered into "host" cells where they are "executed" using the host's existing metabolic resources and gene expression machinery. In doing so, resources are diverted away from essential host processes, such as growth, and towards gene circuit expression. As a result, engineered host cells typically exhibit reduced growth compared with their non-engineered counterparts. This reduction in growth rate is often referred to as "burden" [1], [2]. In the long term, burden results in the loss of circuit function through evolution; error-prone DNA replication leads to the emergence of new, variant "mutant" strains, and where such mutations reduce circuit resource utilisation, burden is relieved, increasing their growth rate and giving them a competitive advantage over their functional ancestors. Over time, strains carrying mutated circuits come to dominate the population and circuit function is inevitably lost [3].

Negative feedback confers synthetic biological systems with robustness in the face of parametric uncertainty and environmental perturbation [4]. It has been demonstrated both experimentally and theoretically that negative feedback has the potential to confer gene circuits with robustness against mutation and improve their performance over evolutionary timescales [5]–[7].

DPB is funded by EPSRC Standard Research Studentship (EP/W52455/1). APSD is funded by the Royal Academy of Engineering under the Research Fellowship scheme. The authors are with the School of Engineering, University of Warwick, Coventry, UK. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising. (correspondence to a.darlington.1@warwick.ac.uk)

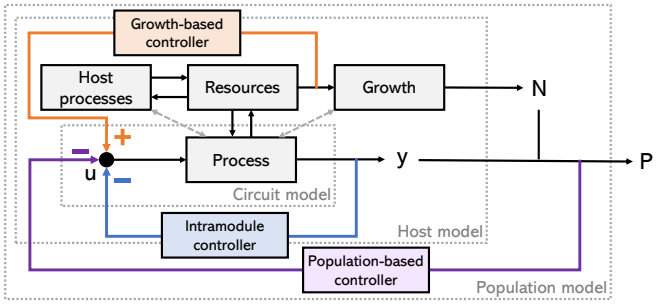

Fig. 1. A block diagram outlining the model and the three control inputs: intramodule control (blue), growth-based control (orange) and population-based control (purple).

Here we consider the design of negative feedback controllers which maintain the *population-level* output of a simple gene circuit (i.e. the sum of circuit output across all cells in a population) over evolutionary time. We investigate three primary classes of controller based on their input (Fig. 1). The first (blue) corresponds to sensing the circuit output at the single cell level. As circuits mutate and synthetic protein production falls, inhibition is relieved, enabling the production level to rise again so cells can maintain output despite disturbances/uncertainties. This can be implemented using repressive transcription factors or small RNAs [8], [9]. The second class (orange) involves sensing the host cell's growth rate. In the presence of burden (i.e. low growth), native genes are activated which can be used to drive circuit repression, creating a burden-based feedback system [10]. In response to mutations which reduce circuit function, burden is alleviated and the host growth rate increases. This increased growth rate relieves inhibition so that production can be maintained. The third class integrates negative feedback using the total population-wide protein output (purple) and could be implemented experimentally using cell-cell communication systems such as quorum sensing [11].

In this paper we compare these approaches using a multi-scale model of circuit gene expression, host physiology and mutation. In Section II, we outline a mathematical model of an evolving population of engineered cellular processes and simulate a simple open-loop gene circuit in repeated batch conditions. In Section III, we develop phenomenological models of the three controllers and investigate their performance. In Section IV, we use a parameter sampling approach to compare performance at the topological level.

In Section V, we discuss the short- and long-term dynamics of the controllers to understand differences in controller performance. Finally, in Section VI, we propose a new control system that combines two of the control inputs to achieve greater improvements in evolutionary longevity.

II. Modelling gene circuit evolution

A. A model of a simple gene circuit

We consider a simple gene circuit consisting of a single gene A that produces protein p_A, the cellular circuit output. The circuit dynamics follow an ordinary differential equation (ODE) model tracking three variables: mRNA m_A, translation complexes c_A (i.e. ribosomes actively translating mRNA A) and protein p_A, each measured in molecules per cell (mc). mRNA molecules m_A are spawned according to the transcription rate $T_{X\,A}(e,u)$, dependent on the energy supply e and the control input of choice u. The mRNA binds host ribosomes R at a binding rate b_A to form translating complexes c_A, and decays at a rate δ_{m_A}. The translating complexes can either dissociate (at rate u_A to produce mRNA m_A without completing translation) or successfully translate a protein, producing p_A alongside m_A at the translation rate $T_{L\,A}(c_A,e)$, dependent on the energy supply e and number of translating complexes c_A. The model exploits the host's supply of ribosomes R and energy e. All variables are diluted by the host growth rate λ. The model is defined as follows:

$$\dot{m}_A = T_{X_A}(e,u) + T_{L_A}(c_A,e) - b_A R m_A + u_A c_A - (\lambda + \delta_{m_A}) m_A, \quad (1)$$

$$\dot{c}_A = -T_{L_A}(c_A,e) + b_A R m_A - u_A c_A - \lambda c_A, \quad (2)$$

$$\dot{p}_A = T_{L_A}(c_A,e) - \lambda p_A. \quad (3)$$

The transcription and translation rates are given by:

$$T_{X_A}(e,u) = \frac{\omega_A e}{\pi_A + e}\Theta_A(u), \quad (4)$$

$$T_{L_A}(c_A,e) = \frac{c_A}{n_A}\frac{\gamma_{\max} e}{K_\gamma + e}, \quad (5)$$

where ω_A is the maximal transcription rate and π_A is the transcription energy threshold. n_A is the protein length, $\gamma_{\max}$ is the maximum rate of protein elongation and K_γ is the elongation energy threshold. Control is enacted through the regulatory function $\Theta_A(u)$ which impacts the transcription rate. For an open-loop system, $\Theta_A(\cdot) = 1$. In Section III, we choose the control input u and regulatory function $\Theta_A(u)$ to maximise evolutionary longevity. Throughout, we fix the following parameters, based on the parameterisation of host genes in the host model [12]: $\pi_A = 4.38$ mc, $n_A = 300$ aa (average length of *E. coli* gene), $b_A = 0.1$ mc min^{-1}, $u_A = 0.01$ min^{-1}, $\delta_{m_A} = 0.1$ min^{-1}. We vary ω_A between 10^{-1} and 10^3 mc min^{-1} representing a wide, biologically feasible range of gene expression.

B. Modelling interactions between host and circuit

This simple three-state gene circuit model is embedded into an established model of *E. coli* physiology [12]. The combined host-circuit model contains 19 ordinary differential equations (ODEs) that explicitly capture the dynamics of cell metabolism, gene expression and dynamic growth. In addition to synthetic protein p_A, the model tracks a coarse-grained proteome consisting of transport proteins p_T, enzymes p_E, ribosomal protein p_R and other host "housekeeping" proteins p_H. Growth is calculated dynamically as a function of the total number of translating complexes:

$$\lambda = \frac{\gamma_{\max} e}{M(K_\gamma + e)} \sum_x c_x, \quad (6)$$

where M defines the total mass of proteins in the cell. The model contains a simplified metabolism with the following dynamics:

$$\dot{s}_I = v_{\text{imp}}(p_T, s_X) - v_{\text{cat}}(p_E, s_I) - \lambda s_I \quad (7)$$

$$\dot{e} = \phi_e v_{\text{cat}}(p_E, s_I) - \sum_x [n_x T_{L_x}(c_x, e)] - \lambda e. \quad (8)$$

Here, s_X is the external substrate which is consumed by the cell. It is imported (i.e. converted to internal substrate s_I) at a rate $v_{\text{imp}}(p_T, s_X)$ and metabolised (i.e. converted to energy e) at a rate $v_{\text{cat}}(p_E, s_I)$ with the nutrient efficiency ϕ_e determining the amount of energy produced per molecule of substrate. Both v_{imp} and v_{cat} have a Michaelis-Menten form. The energy e is assumed to influence both the transcription rate $T_{X\,x}(e,u)$ and the translation rate $T_{Lx}(c_x,e)$, but is consumed only through translation [12]. The energy consumption by translation is scaled by the protein length n_x. The transcription and translation rates for host protein types obey the same form as those for synthetic protein (Eq. 4, 5). Housekeeping proteins self-regulate via $\Theta_H(p_H) = k_H^4/(k_H^4 + p_H^4)$ with $k_H = 152219$ mc [12]. All other protein types are unregulated ($\Theta_x(\cdot) = 1$).

Gene expression of most protein types is modelled identically to synthetic protein (Eq. 1, 2, 3). Ribosomes, composed of rRNA r and proteins p_R, undergo biogenesis via the multi-step process discussed in detail elsewhere [12]. The dynamics of the free ribosome pool is:

$$\dot{R} = -\sum_x [T_{X_x}(c_x,e) - b_x R m_x + u_x c_x] + \underbrace{\beta_r p_R r - \mu_r R}_{\text{Ribo. biogenesis}} - \lambda R \quad (9)$$

The dynamics of the host and circuit are interlinked, with interactions in both directions, capturing the phenomenon of burden. The host influences the circuit by supplying energy and ribosomes, and diluting all circuit components according to its growth rate (Eq. 1, 2, 3). The circuit influences the host by altering the growth rate (Eq. 6) and the division of energy and ribosomes between protein types (Eq. 8, 9).

C. Modelling a population of evolving circuit-bearing cells

To model an evolving population of circuit-bearing cells, we extend the approach we used in [6], similar to one recently proposed by Ingram and Stan [13]. Simulations of chemostat growth (as in [6]) limit the exploration of the design space; if a given circuit has a growth rate slower than the chemostat dilution rate, all population dynamics collapse to zero. We therefore model population dynamics in a batch setting. To simulate a heterogeneous population

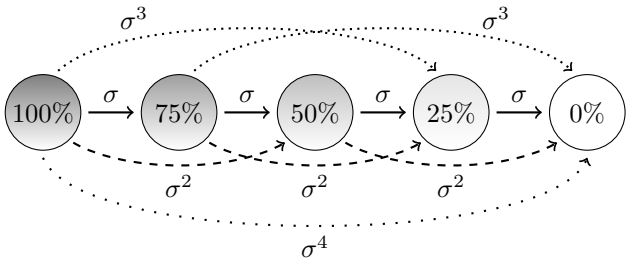

Fig. 2. The mutation scheme used for the model. Each circle represents a "mutation state" - a strain of cells distinguished by different maximal transcription rates ω_A. Percentages describe the level of ω_A relative to the designed level. Mutations are represented by arrows between states, with labeled values describing the rate of transition between states. Only mutations which reduce function are allowed. Here, $\sigma = 10^{-6}$ min^{-1}.

without a combinatorial explosion in complexity, we define a population N comprised of five discrete "mutation states" so that $N = \sum_{j=1}^{5} N_j$. These mutation states represent distinct mutant strains (here defined as percentages of the designed transcription rate of synthetic protein ω_A: 100%, 75%, 50%, 25%, 0%). The 100% state represents a fully functional circuit and the 0% state represents a non-functional circuit. Mutation is modelled via transitions between mutation states. The population of an engineered strain i follows:

$$\dot{N}_i = \lambda_i N_i + \sum_{j=1, j\neq i}^{5} \left[\sigma_{ji} N_j - \sigma_{ij} N_i\right], \qquad (10)$$

where λ_i is the growth rate of state i and σ_{ij} is the transition rate from state i to state j. Transition rates are defined in reference to a single parameter $\sigma = 10^{-6}$ min^{-1}, with more extreme mutations assumed to be less frequent (Fig. 2). All mutation states consume the extracellular substrate s_X:

$$\dot{s_X} = -\sum_{j=1}^{5} \left[v_{\text{imp}}(p_{T_j}, s_X) N_j\right], \qquad (11)$$

where p_{T_i} is the transporter of the ith state. The combined model consists of 101 ODEs, 20 per state in addition to s_X. Host parameters are consistent with previous work [6], [12], although we alter the nutrient efficiency ($\phi_e = 20$) and elongation energy threshold ($K_\gamma = 8 \times 10^8$ mc) so that the population reaches steady state within 24 hours [14].

D. Quantifying circuit performance over evolutionary time

To verify the ability of the model to capture evolutionary dynamics, we simulate an open-loop system in the absence of control with $\omega_A = 5$ mc min^{-1} (Fig. 3). To resemble a realistic laboratory study, external substrate is replenished ($s_X = 10^{12}$ molecules) every 24 hours and the population size is reset to 1000, maintaining the ratio between mutation states (as in [14]) (Fig. 3a). We consider the total process output across the combined population over time:

$$P = \sum_{j=1}^{5} \left(N_j p_{A_j}\right).$$

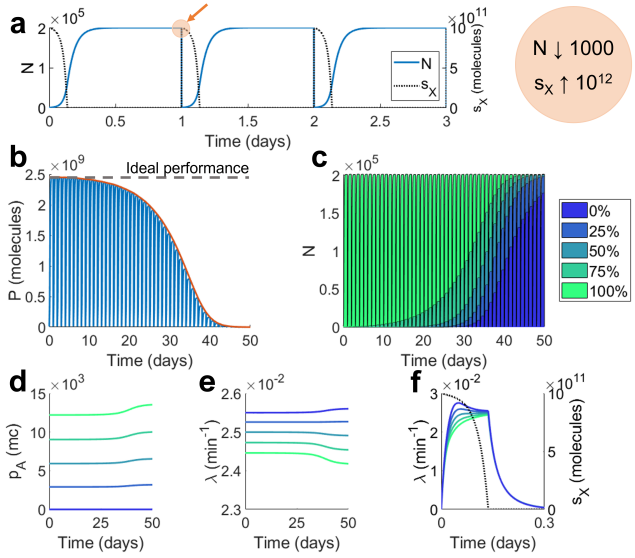

Fig. 3. Time-series outputs of an open-loop process with maximal transcription rate $\omega_A = 5$ mc min^{-1}. (a) Population size N and external substrate s_X for the first three days of the simulation. Every 24 hours, the population size is reduced to 1000 cells and substrate is replenished. (b) Total population-wide circuit output P, plotted in full (blue) and at the end of each simulation day (red). The grey dashed line shows how a theoretically ideal system would perform. (c) Population size N, distributed according to mutation state. (d) Output per cell p_A according to mutation state, measured once per day. (e) Growth rate λ according to mutation state, measured once per day. (e) Left axis: A snapshot of the entire growth dynamics for a single day of the simulation. Right axis: External substrate s_X.

An ideal system would maintain P over long time scales. To quantify evolutionary performance, we define three metrics:

1) The initial population-wide protein output P_0.
2) The 90%-life τ_{90}, the time taken for P to fall to 90% of its initial value.
3) The 50%-life τ_{50}, the time taken for P to halve.

τ_{50} is a widely-used measure of longevity [15]. τ_{90} quantifies the duration of "good" performance, useful in applications where maintaining high production is key. The output P fluctuates widely in repeated batch conditions. Therefore, we calculate these metrics by interpolating between the output values at the end of each day. Likewise, to track long-term protein production and growth dynamics, we measure outputs p_A and growth rates λ once per day at the instant where substrate runs out. (Minor variations over time as in Fig. 3d,e are therefore possible as perturbation to substrate consumption may lead to it running out at different times of day.) The initial/nominal circuit design has $\tau_{90} = 20.5$ days and $\tau_{50} = 32.2$ days, and function is completely eradicated within 50 days (Fig. 3b). This loss of function corresponds to a transition of the population distribution from one initially comprised entirely of cells with fully functional circuits to one comprised entirely of cells with non-functional circuits, with intermediate mutation states arising and dying out during the simulation (Fig. 3c). This transition occurs because states with reduced transcription produce less synthetic protein and have lower burden/faster growth than states with greater production (Fig. 3d,e,f).

III. Phenomenological Models of Feedback Controllers

We develop phenomenological models of controllers which act to influence the circuit transcription rate via different control inputs u and regulatory functions $\Theta_A(u)$ (Eq. 4). We consider three forms for $\Theta_A(u)$ which correspond to distinct biological inputs. We call them (i) intramodule feedback, (ii) growth-based feedback and (iii) population-based feedback (Fig. 1). In each case, we implement schemes which act where burden is high in order to reduce the selective advantage of mutants.

A. Intramodule feedback

Intramodule feedback arises from feedback inhibition by the process's protein product p_A on a per-cell basis. This approach is analogous to control designs proposed by e.g. Shopera *et al.* [8] and Huang *et al.* [9], where each cell senses a proxy for its own output. Self-inhibition of p_A production is implemented using a Hill function:

$$\Theta_A(p_A) = \frac{{k_A}^2}{{k_A}^2 + {p_A}^2} \tag{12}$$

We simulate the action of this controller for three nominal processes with various burdens generated by maximal transcription rates ω_A set as 25, 5 and 1 mc min^{-1} (corresponding to ancestral growth rates of 0.0223, 0.0242 and 0.0253 min^{-1}). Increasing the controller strength by decreasing k_A across its biologically feasible range increases both τ_{90} and τ_{50} at a loss of initial output P_0 (Fig. 4a,d,e).

B. Growth-based feedback

Since the rate of function loss in an engineered population is primarily determined by the difference in growth rates between functional and non-functional strains, using control to explicitly react to growth rates is an appealing strategy. In [10], Ceroni *et al.* proposed a burden-activated controller that exploits native promoters that are upregulated when cells experience burden. This controller acts to inhibit synthetic protein production at low growth and relieves repression at high growth. We model this function as:

$$\Theta_A(\lambda) = \frac{\lambda^2}{{k_\lambda}^2 + \lambda^2}. \tag{13}$$

Nominal controller designs successfully improve both τ_{90} and τ_{50} at a loss of initial output P_0 as controller strength increases (Fig. 4b,d,e).

C. Population-based feedback

Since our goal is to maintain population-wide output P over time, we consider using this as the input for a negative feedback controller. A population-sensitive system like this could be implemented biologically using cell-cell communication systems such as quorum sensing.

$$\Theta_A(P) = \frac{{k_P}^2}{{k_P}^2 + P^2} \tag{14}$$

We again simulate a range of controllers by varying k_P and see an improvement in both τ_{90} and τ_{50} at the expense of initial output P_0 as controller strength increases (Fig. 4c,d,e).

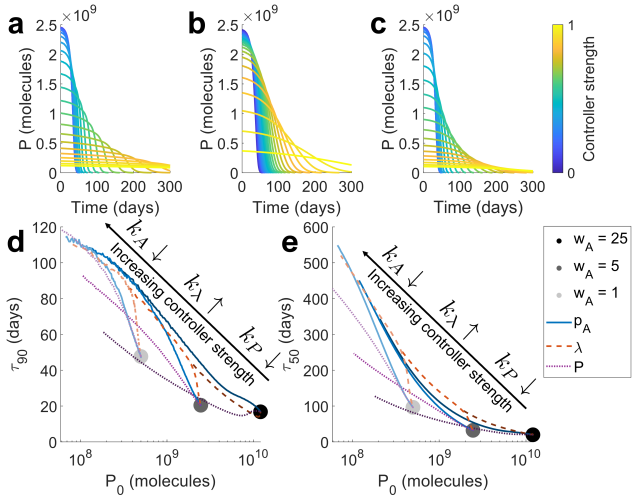

Fig. 4. Three open-loop processes are considered with maximal transcription rates $\omega_A = 25, 5, 1$ mc min^{-1}. (a,b,c) For the process with $\omega_A = 5$ mc min^{-1}, time-series outputs are presented with increasing control strengths for each controller: (a) intramodule, (b) growth-based, (c) population-based. Line colour indicates the strength of control, determined by varying the dissociation parameters k_A, k_λ, k_P over biologically feasible ranges and normalised to be on a scale from 0 to 1. (d,e) Grey circles mark the initial outputs P_0 of the three open-loop processes against (d) 90%-life τ_{90} and (e) 50%-life τ_{50}. Lines indicate how this relationship changes as controller strength is increased for each control scheme.

IV. Performance Comparison of Controllers

We now study controller performance by varying both process activity (via the maximum transcription rate ω_A) and controller strength (via dissociation parameters k_A, k_λ, k_P) to compare their topological behaviour (Fig. 5). In each case, controller performance is determined by comparison with an open-loop system of equal initial output P_0, generated by separately varying ω_A.

Intramodule feedback always outperforms open-loop with improvements versus open-loop of up to 100% for τ_{50} and an even greater 128% for τ_{90}. Across samples of equivalent output, those with the highest process activity and strongest control provide the largest improvement, although continually increasing control strength beyond a certain point offers diminishing returns (Fig. 5a,b). When control is very weak, while τ_{50} is unchanged versus open-loop, τ_{90} can still noticeably improve, particularly in systems with large initial outputs. Directly maximising controller strength significantly reduces the range of possible initial outputs achievable. For this reason, when designing such a controller, it is recommended to first maximise the base transcription rate via ω_A, then adjust the controller strength to achieve the preferred initial output.

While intramodule feedback is still effective at very large initial outputs ($> 10^{10}$), growth-based feedback is unable to yield improvement in this range. However, for the majority of the performance space, growth-based feedback does improve both τ_{90} and τ_{50} (Fig. 5c,d). Unlike intramodule feedback, the potential improvement in τ_{90} (up to 99%) is exceeded by that of τ_{50} (up to 145%). Among systems of

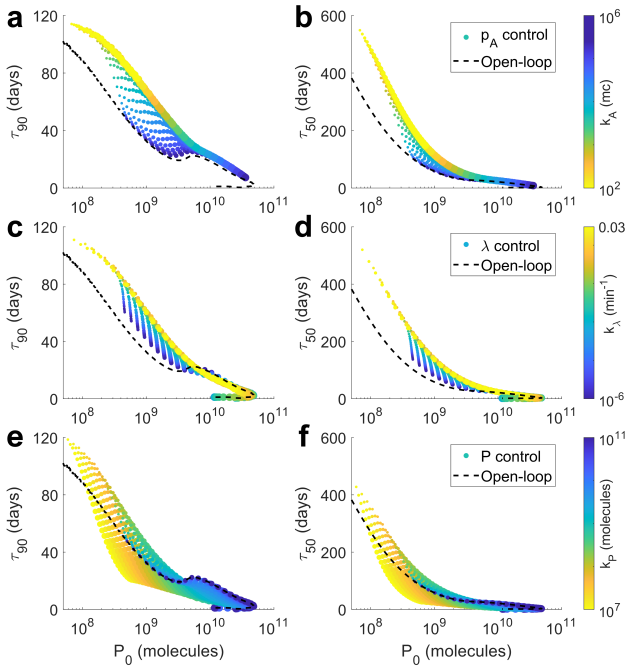

Fig. 5. All plots show the relationship between initial output P_0 and 90%-life τ_{90} or 50%-life τ_{50} for various process-controller systems. Each marker represents the output of a simulation with a different parameterisation. Dotted black lines indicate open-loop relationships. Markers indicate different process-controller parameterisations, where marker colour indicates the strength of the controller (k_A, k_λ, k_P) and marker size indicates the activity of the process (ω_A). (a,c,e) τ_{90}, (b,d,f) τ_{50}. (a,b) Intramodule feedback, (c,d) Growth-based feedback, (e,f) Population-based feedback.

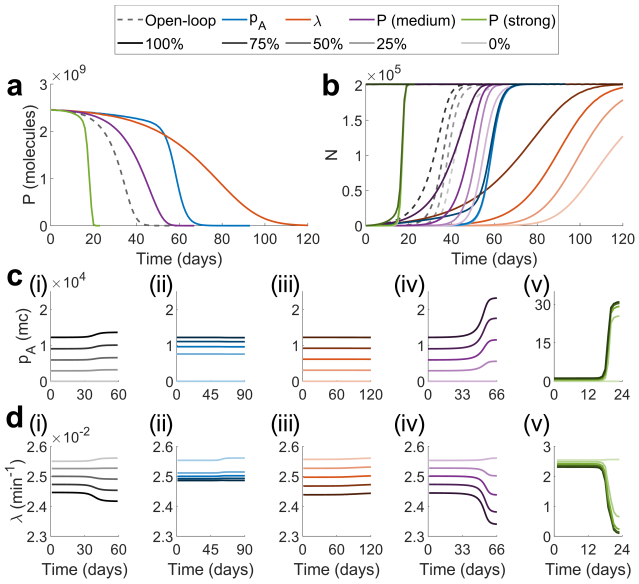

Fig. 6. Time series outputs for five individual systems: (i) Open-loop, (ii) Strong intramodule feedback, (iii) Strong growth-based feedback, (iv) Medium-strength population-based feedback, (v) Strong population-based feedback. All quantities are measured once per day. (a) Total output P, (b) Population N distributed according to mutation state, (c) Protein output per cell p_A according to mutation state, (d) Growth rate λ for each mutation state.

equivalent initial output, the strongest controllers provide the largest improvement, though again there is little to separate controller performance once the strength passes a certain point. While the weakest controllers have little impact at high initial outputs, they are capable of improving both the τ_{90} and τ_{50} when P_0 is low. Contrary to intramodule feedback, a wide range of P_0 is achievable with all but the strongest controllers. Therefore, it is recommended to first choose a strong controller, and then adjust the process activity to achieve the preferred initial output.

For population-based feedback, the majority (536/900 for τ_{90} and 526/900 for τ_{50}) of samples perform *worse* than open-loop. Further, across systems of equal initial output, controllers with the strongest action actually yield the worst performance. When improvement is possible, performance is instead maximised at some intermediate control strength (Fig. 5e,f).

V. Understanding Differences in Controller Performance

Our results suggest that, at the topological level, population-based feedback is vastly outperformed by intramodule feedback and growth-based feedback. Intramodule feedback particularly excels at extending the 90%-life, whereas growth-based feedback is most effective at extending the 50%-life. To better understand these differences in performance and their mechanistic cause, we consider the dynamics of a series of individual controller designs, whose maximal transcription rates ω_A have been tuned so that their initial outputs P_0 align with the original open-loop system (Fig. 3). In total, we compare five systems:

(i) Open-loop ($\omega_A = 5$ mc min^{-1}).
(ii) Strong intramodule feedback ($\omega_A = 390$ mc min^{-1}, $k_A = 1000$ mc).
(iii) Strong growth-based feedback ($\omega_A = 21$ mc min^{-1}, $k_\lambda = 0.02$ min^{-1}).
(iv) Medium-strength population-based feedback, which outperforms the open-loop system ($\omega_A = 8.5$ mc min^{-1}, $k_P = 10^8$molecules).
(v) Strong population-based feedback, which does *not* outperform the open-loop system ($\omega_A = 705$ mc min^{-1}, $k_P = 1.4 \times 10^9$molecules).

To explain the differences in their performance, we look at dynamics over the long term (with quantities measured once per day) and in the short term (considering quantities over the course of one day).

A. Long-term dynamics explain intramodule 90%-life performance and accelerated population-based loss-of-function

We have previously demonstrated that intramodule feedback is capable of improving both the 90%-life and 50%-life of a simple system in a chemostat, with the improvement in 90%-life being more significant because the fully functional (100%) state endures for longer but the non-functional (0%) state dominates quickly when it arises [6]. Results are consistent in the repeated batch setting (Fig. 6a,b, blue), with τ_{90} and τ_{50} improving by 122% and 79% versus open-loop. When intermediate states arise, they produce less output per

cell p_A than the fully functional (100%) state. Correspondingly, control is alleviated, so p_A rises and remains closer to the designed level (Fig. 6c(ii)). This means that mutation into intermediate states provides only a small growth advantage, but mutation into the non-functional (0%) state provides a larger growth advantage (Fig. 6d(ii)). Intramodule feedback would therefore be even more effective under alternative mutation schemes where complete loss-of-function mutations are less likely, and vice versa.

Despite having a smaller 90%-life, the growth-based controller outlasts the intramodule controller, with improvements versus open-loop of τ_{90} by 89% and τ_{50} by 125% (Fig. 6a, red). This corresponds to a more steady transition in population distribution through the intermediate states, as opposed to a sudden rise in the non-functional (0%) state (Fig. 6b, red). The long-term protein production and growth dynamics do not appear to change significantly versus open-loop (Fig. 6c(iii),d(iii)).

The medium-strength population-based feedback controller improves τ_{90} by 30% and τ_{50} by 26% versus open-loop, with a steady transition in population distribution (Fig. 6a,b, purple). The strong population-based controller significantly worsens both metrics, corresponding to a sharp transition in population distribution from fully functional (100%) to non-functional (0%) (Fig. 6a,b, green). As the population distribution changes over the course of the simulation, the control input P, which is the same for all mutation states, falls. This means that the internal dynamics of cells change *even if they haven't mutated*. Since P decreases over time, the strength of control reduces over time. Thus, for each state, the per-cell production rate increases and growth rate decreases (Fig. 6c(iv),c(v),d(iv),d(v)). This means that functional circuits are put at an even greater selective disadvantage than non-functional circuits, accelerating loss-of-function.

B. Short-term dynamics drive enhanced growth-based performance and poor population-based performance

We now consider the first day of the simulations from Section V-A. Here, mutation has negligible influence, and the whole population dynamics can be captured by considering only the fully functional (100%) mutation states. In response to external substrate being replenished, there is a spike in the per-cell production of p_A. This protein production overshoot damps growth. The best performing controllers reduce the height of this spike, with growth-based feedback capable of flattening it altogether. The strong population-based controller actually heightens this overshoot (Fig. 7a).

By avoiding this spike in protein production, the growth-based controller is able to divert more resources towards growth, leading to a faster-growing population. On the other hand, the strong population-based controller diverts more resources away from growth, leading to a slower-growing population (Fig. 7b). As a result, the population-based controller is able to produce more protein in the first 3 hours, using a small number of high-producing cells. However, production by the growth-based system then overtakes it,

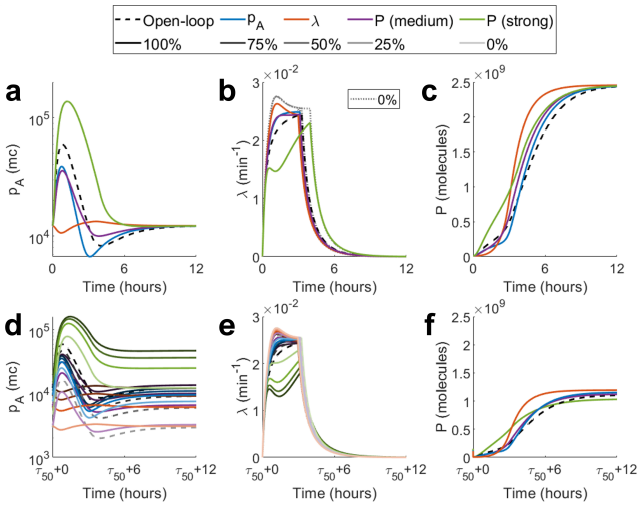

Fig. 7. Time series outputs from a single day of simulation for five individual systems: (i) Open-loop, (ii) Strong intramodule feedback, (iii) Strong growth-based feedback, (iv) Medium-strength population-based feedback, (v) Strong population-based feedback. (a,b,c) First day of simulation. Only 100% (fully functional) mutation states are plotted. (a) Protein output per cell p_A, (b) Growth rate λ (dotted grey lines show growth rates of 0% (non-functional) mutation states), (c) Total output P. (d,e,f) Day of simulation when τ_{50} is reached. Analogous to (a,b,c) with all mutation states displayed.

making the most of its increased population size to produce output from a larger number of less burdened cells (Fig. 7c). Similar dynamics are observed across the simulation in the presence of mutation. (The day when systems reach their 50%-life is shown in Fig. 7d,e,f.)

VI. Combining Intramodule and Growth-Based Feedback Boosts Evolutionary Longevity

We have demonstrated that both intramodule feedback and growth-based feedback can be effective at improving evolutionary longevity, the first by "condensing" the growth rates of the intermediate states so that they have less of a selective advantage over the fully functional state and the second by improving the growth rate of the fully functional state in the short-term response to replenished substrate. We propose a new controller which combines both intramodule and growth-based feedback. We employ the following regulatory function:

$$\Theta_A(p_A, \lambda) = \left[\frac{{k_A}^2}{{k_A}^2 + {p_A}^2}\right]\left[\frac{\lambda^2}{{k_\lambda}^2 + \lambda^2}\right]. \qquad (15)$$

We sample a large number of controller designs by varying ω_A, k_A and k_λ over biologically feasible ranges. We select those samples which simultaneously maximise initial output, 50%-life and 90%-life. This is equivalent to solving the following multi-objective optimisation problem:

$$\begin{aligned} &\underset{\omega_A,\ k_A,\ k_\lambda}{\text{maximise}} \quad \left(P_0,\ \tau_{90},\ \tau_{50}\right) \\ &\text{subject to} \\ &\qquad 10^0 \le \omega_A \le 10^3 \text{ mc min}^{-1} \\ &\qquad 10^2 \le k_A \le 10^6 \text{ mc} \\ &\qquad 10^{-6} \le k_\lambda \le 0.03 \text{ min}^{-1}. \end{aligned} \qquad (16)$$

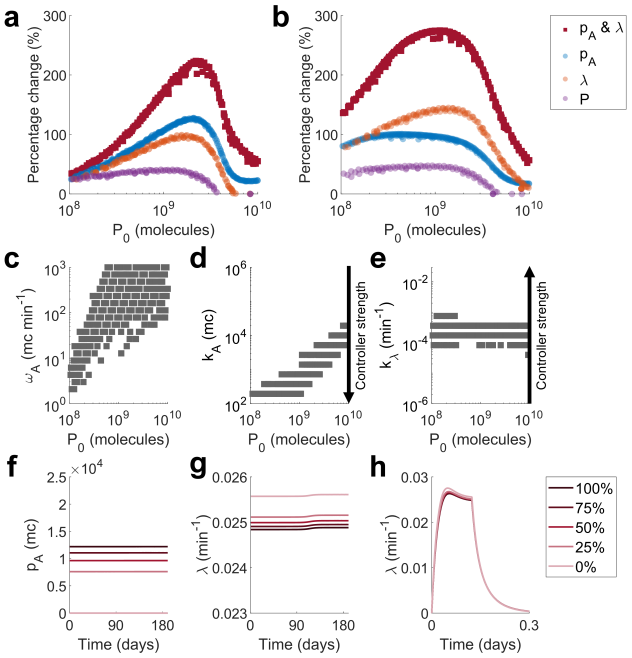

Fig. 8. Combining intramodule and growth-based control. (a,b,c,d,e) Samples were generated by varying process activity (ω_A) and control strength (k_A, k_λ). Samples which simultaneously maximise initial output 90%-life τ_{90} and 50%-life τ_{50} are plotted with initial output P_0 on the x axis. (a,b) Percentage change in (a) τ_{90} and (b) τ_{50} versus open-loop. Other controllers also presented for comparison. (c) Process maximal transcription rate ω_A. (d) Control strength from the intramodule component k_A. (e) Control strength from the growth-based component k_λ. (f,g,h) A single controller design was selected with quantities plotted over time. (f) Output per cell p_A, measured once per day. (g) Growth rate λ, measured once per day. (h) Growth rate λ over the course of the first day.

This control topology significantly outperforms all previous systems, with improvements of up to 224% for τ_{90} and 275% for τ_{50} versus open-loop (Fig. 8a,b). Parameterisations of optimal controllers suggest that process activity and intramodule control strength should be large, while growth-based control strength should be kept at a medium value (Fig. 8c,d,e.) Excessively increasing the growth-based control strength beyond this level significantly reduces the initial output P_0.

To understand the dynamics of this control scheme, we select a close-to-optimal parameterisation and tune the process activity via ω_A to align with the original open-loop system, as in Section V-B ($w_A = 502$ mc min^{-1}, $k_A = 1000$ mc, $k_\lambda = 10^{-3}$ min^{-1}). Versus open-loop, this controller improves τ_{90} by 221% and τ_{50} by 243%. As expected, the production rates and growth rates of intermediate states are condensed (Fig. 8f,g). Furthermore, the growth rate of the fully functional state is able to closely track that of the non-functional state in response to fresh medium (Fig. 8h).

VII. Conclusion

In this work, we have developed a mathematical model to compare the ability of controllers to improve the evolutionary longevity of a simple gene circuit. When the strength of control is maximised, population-based feedback actually performs worse than open-loop. Even when optimised, it performs far worse than the other two inputs considered. Intramodule feedback is particularly effective at improving the 90%-life, because it enables intermediate mutation states to better maintain synthetic protein production and reduces their selective advantage. Growth-based feedback is most effective at improving the 50%-life, because it prevents burdensome protein production from overshooting when extracellular substrates are replenished. We demonstrate that differences in controller performance are a result of differences in both long- and short-term dynamics, with evolutionary performance dictated by the relative growth rates of competing strains. We propose a novel control strategy based on both intramodule and growth-based feedback which significantly improves evolutionary longevity. We are now investigating the performance of biologically feasible controller designs inspired by these theoretical comparisons. In particular, it will be important to evaluate the extent to which the theoretical benefits provided by negative feedback remain if the controller itself introduces additional burden.

References

[1] O. Borkowski, F. Ceroni, G.-B. Stan, and T. Ellis, "Overloaded and stressed: Whole-cell considerations for bacterial synthetic biology," *Current opinion in microbiology*, vol. 33, pp. 123–130, 2016.

[2] A. Boo, T. Ellis, and G.-B. Stan, "Host-aware synthetic biology," *Current Opinion in Systems Biology*, vol. 14, pp. 66–72, 2019.

[3] S. C. Sleight, B. A. Bartley, J. A. Lieviant, and H. M. Sauro, "Designing and engineering evolutionary robust genetic circuits," *Journal of biological engineering*, vol. 4, no. 1, pp. 1–20, 2010.

[4] A. Becskei and L. Serrano, "Engineering stability in gene networks by autoregulation," *Nature*, vol. 405, no. 6786, pp. 590–593, 2000.

[5] D. C. Marciano, R. C. Lua, P. Katsonis, S. R. Amin, C. Herman, and O. Lichtarge, "Negative feedback in genetic circuits confers evolutionary resilience and capacitance," *Cell reports*, vol. 7, no. 6, pp. 1789–1795, 2014.

[6] D. P. Byrom and A. P. Darlington, "On the implications of controller resource consumption for the long-term performance of synthetic gene circuits," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, IEEE, 2023, pp. 8844–8850.

[7] A. Gyorgy, "Competition and evolutionary selection among core regulatory motifs in gene expression control," *Nature Communications*, vol. 14, no. 1, 2023.

[8] T. Shopera, L. He, T. Oyetunde, Y. J. Tang, and T. S. Moon, "Decoupling resource-coupled gene expression in living cells," *ACS synthetic biology*, vol. 6, no. 8, pp. 1596–1604, 2017.

[9] H.-H. Huang, Y. Qian, and D. Del Vecchio, "A quasi-integral controller for adaptation of genetic modules to variable ribosome demand," *Nature communications*, vol. 9, no. 1, p. 5415, 2018.

[10] F. Ceroni, A. Boo, S. Furini, *et al.*, "Burden-driven feedback control of gene expression," *Nature methods*, vol. 15, no. 5, pp. 387–393, 2018.

[11] C. McBride and D. Del Vecchio, "Trade-offs in robustness to perturbations of bacterial population controllers," in *2020 American Control Conference (ACC)*, IEEE, 2020, pp. 4276–4282.

[12] A. P. Darlington, J. Kim, J. I. Jiménez, and D. G. Bates, "Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes," *Nature communications*, vol. 9, no. 1, p. 695, 2018.

[13] D. Ingram and G.-B. Stan, "Modelling genetic stability in engineered cell populations," *Nature Communications*, vol. 14, no. 1, p. 3471, 2023.

[14] S. C. Sleight and H. M. Sauro, "Visualization of evolutionary stability dynamics and competitive fitness of escherichia coli engineered with randomized multigene circuits," *ACS synthetic biology*, vol. 2, no. 9, pp. 519–528, 2013.

[15] B. A. Renda, M. J. Hammerling, and J. E. Barrick, "Engineering reduced evolutionary potential for synthetic biology," *Molecular BioSystems*, vol. 10, no. 7, pp. 1668–1678, 2014.