Article

A blueprint for a synthetic genetic feedback optimizer

Received: 18 August 2022

Accepted: 5 April 2023

Published online: 03 May 2023

Check for updates

Andras Gyorgy [●]¹[∞], Amor Menezes [●]² & Murat Arcak³

Biomolecular control enables leveraging cells as biomanufacturing factories. Despite recent advancements, we currently lack genetically encoded modules that can be deployed to dynamically fine-tune and optimize cellular performance. Here, we address this shortcoming by presenting the blueprint of a genetic feedback module to optimize a broadly defined performance metric by adjusting the production and decay rate of a (set of) regulator species. We demonstrate that the optimizer can be implemented by combining available synthetic biology parts and components, and that it can be readily integrated with existing pathways and genetically encoded biosensors to ensure its successful deployment in a variety of settings. We further illustrate that the optimizer successfully locates and tracks the optimum in diverse contexts when relying on mass action kinetics-based dynamics and parameter values typical in *Escherichia coli*.

Biological systems continuously sense and respond to changes in their environment by relying on sophisticated regulatory mechanisms to ensure both optimality and robustness, eventually leading to considerable complexity¹. Synthetic biology seeks to (re)program cellular processes by designing and combining standardized biological parts in a modular fashion², often enhanced by computational approaches³, to deploy synthetic gene circuits in a variety of contexts that range from biotherapeutics to environmental sciences⁴.

Unfortunately, genetic modules regularly fail and require multiple design iterations⁵ as a result of unmodeled interactions among circuit components and the host organism^{6,7}, frequently resulting in perplexing behaviors^{8,9}. Further exacerbating this problem, genetic modules developed in one strain can behave fundamentally differently when deployed in other organism chassis^{10,11}. Therefore, context-dependence presents a major obstacle to rationally and robustly controlling cellular processes, to creating novel functionalities, and to fine-tuning existing ones^{6,12}. Metabolic burden, for instance, couples the expression of independent proteins and can significantly reduce cellular growth rate¹³⁻¹⁹ (Fig. 1a).

While screening for optimal genetic realizations is now wellestablished by varying a diverse array of biochemical properties^{20–23}, such static approaches are unable to adaptively respond to disturbances and shifts in the environment. To address this challenge, dynamic gene expression control^{24–26} offers a promising solution, using a system-level approach that combines quantitative tools from a wide array of disciplines^{27–30}. Recently, an especially fruitful direction has been to borrow ideas from control theory to analyze and design genetic feedback systems^{31,32}. Notably, integral feedback has been proposed³³ to ensure perfect adaptation, that is, to return to a desired setpoint after a perturbation. Although it is not without limitations³⁴, this simple biological feedback control module has been successfully implemented in a variety of settings: in vivo^{35,36} using both RNA-based and protein-based mechanisms, and in vitro^{37,38} in bacteria, as well as in mammalian cells³⁹, and the design has even been extended in the form of a proportional-integralderivative controller⁴⁰.

While these results represent a major leap in our ability to robustly control cellular behavior, they rely on an explicitly defined reference signal. Often times, however, the desired reference signal is defined only implicitly (i.e., where performance is maximized), and it continuously shifts due to fluctuations in the cellular environment, for instance, to ensure optimal resource re-allocation for maximal growth⁴¹. Therefore, we must develop genetic modules that can locate and track these implicit reference signals to create synthetic circuits

¹Division of Engineering, New York University Abu Dhabi, Abu Dhabi, UAE. ²Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL, USA. ³Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA. is e-mail: andras.gyorgy@nyu.edu



Fig. 1 | Genetic optimizers can ensure maximal cellular performance. Simulation parameters and further details are provided in Supplementary Section 5. a Population-level production of a target gene is maximized when growth rate and cellular synthesis rate are balanced93. The corresponding optimal concentration of a regulator may depend on both cellular and environmental conditions, and can be automatically adjusted by a genetic optimizer. b Gradient-based optimization can successfully track the time-varying optimum, but cannot be immediately

that can sense and respond to environmental changes and adjust their activity for optimal performance^{25,26}.

To address this central challenge, here we develop and present a blueprint for a genetic optimizer module that dynamically adjusts the production and decay rates of a set of regulator species to ensure optimal cellular performance. Our approach relies on interconnecting common biomolecular sensors and actuators (e.g., logic gates, bistable switches, oscillators) that are readily available, using the traditional feedback loop architecture (Fig. 1a). As we illustrate through multiple application examples, our optimizer successfully locates and tracks an implicitly defined and time-varying optimum in diverse contexts when relying on mass action kinetics-based dynamics and parameter values typical in Escherichia coli (E. coli). While the basic idea underpinning our results can be most easily understood considering idealized

translated to a genetic circuit because it may result in infeasible negative quantities. Decreasing ϵ_x yields faster convergence at the expense of greater control inputs u_1 and u_2 . **c** Calculating u_1 and u_2 based on the trend of x and y ensures convergence to the optimum x[']. In the four panels at right, ϵ_v increases by an order of magnitude going from left to right (leading to slower y dynamics), and the delay t_d increases by an order of magnitude going from top to bottom.

time

time

modules (e.g., completely symmetric toggle switches, absence of dilution), we also probe these assumptions and show that our approach works well even when considering non-ideal implementations. We further demonstrate that closed loop performance is robust to stochastic noise and disturbances (e.g., time-varying parameters due to various sources of context-dependence), and that our proposed optimizer can be easily integrated with existing pathways and genetically encoded biosensors to ensure its successful deployment in a variety of settings.

Results

Problem formulation

Our goal is to construct a genetic network to control the concentration of a set of regulator species x so that cellular performance $F(x, \theta_y)$ is maximized, where θ_y are (possibly time-dependent) biophysical parameters. We only assume that this performance is measurable via the reporter signal *y*, but it is otherwise unknown. We take the optimal concentration x^* as a reference signal, and we seek to implement reference tracking where the time-varying setpoint is defined only implicitly (i.e., where performance is maximized). Finally, we wish to achieve this tracking (i) robustly, that is, in the presence of disturbances, uncertainties, and noise; and (ii) by relying on standard biological parts that are readily available and can be easily combined.

Let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$ denote the concentration vector of the controlled regulator species and the reporter (proxy for cellular performance) that we wish to maximize by adjusting the control signal $u \in \mathbb{R}^k$. The dynamics of the regulator species take the general form of

$$\epsilon_x \dot{x} = f(x, u, \theta_x), \tag{1}$$

where dot denotes differentiation with respect to time, θ_x is a vector of (possibly time-dependent) parameters (e.g., transcription and translation rates, dissociation constants, degradation rates), and ϵ_x characterizes the timescale on which *x* evolves. Regarding the reporter, we consider the general formulation of

$$\epsilon_{v} \dot{y} = F(x, \theta_{v}) - y \tag{2}$$

to account for potential regulatory delays, where θ_y contains timedependent parameters as above, and ϵ_y characterizes the timescale on which *y* evolves. The lag between *x* and *y* increases with ϵ_y/ϵ_x : for instance, $\epsilon_y/\epsilon_x \rightarrow 0$ corresponds to the case when changes in *x* are observable in *y* instantaneously according to $y = F(x, \theta_y)$.

Considering the dynamics (1)-(2), we seek a feedback optimizer (Fig. 1a) of the form

$$\epsilon_z \dot{z} = g_z(z, x, y, \theta_z), \qquad u = g_u(z, \theta_u),$$
 (3)

where *z* and *u* are the internal state and the output of the optimizer, respectively, together with their time-dependent parameter vectors θ_z and θ_u , and ϵ_z characterizes the timescale of the optimizer species. Although any one of ϵ_x , ϵ_y , and ϵ_z can be eliminated by rescaling time (Supplementary Section 1.1), we keep them all to illustrate their effects throughout the paper.

Assuming that $F(x, \theta_y)$ has a unique global time-dependent optimum at x^* (if this is not the case, our approach only yields local optimality), we seek to design the optimizer to ensure that (i) trajectories of the closed loop system (1)–(3) converge towards the unique optimum x^* ; and (ii) the integrated system is biologically realizable and relies only on non-negative signals.

Gradient ascent and approximate gradient ascent

To illustrate the main idea underpinning the proposed optimizer module, we first consider a single species *x* that is produced and degraded according to $\epsilon_x \dot{x} = u_1 - u_2 x$ (Supplementary Section 1.1), where u_1 and u_2 represent control signals regulating the production and degradation of *x*, respectively. In the absence of biological constraints, a potential choice for the optimizer is a gradient-based system with the feedback law $u_1 = \nabla_x F(x, \theta_y)$ and $u_2 = -\nabla_x F(x, \theta_y)$, where $\nabla_x F(x, \theta_y)$ denotes the gradient of $F(x, \theta_y)$. The closed loop dynamics $\epsilon_x \dot{x} = (1+x)\nabla_x F(x, \theta_y)$ realizes gradient ascent by converging to $x = x^*$ where $\nabla_x F(x, \theta_y) = 0$, thus tracking the time-varying optimum (Fig. 1b). Moreover, it is sufficient to simply swap the signs in the control laws above to implement minimization (gradient descent). Unfortunately, this approach has two fundamental limitations: (i) it requires explicit knowledge of the gradient $\nabla_x F(x, \theta_y)$; and (ii) u_1 and u_2 can become negative (Fig. 1b, gray regions).

The gradient-based control law above can be phrased as: *x* should be increased if the gradient of cellular performance is positive, and

decreased otherwise. To overcome the first issue (reliance on the gradient), we can rephrase the control law as: x should be increased if x and y are increasing or decreasing together, and decreased otherwise. To implement this, we must keep track of the change in both x and y: let $w_d(t) = w(t - t_d)$ for the time delay $t_d > 0$ so that $\Delta x = x - x_d$ and $\Delta y = y - y_d$ represent the changes in x and y, respectively. With this, x should be increased when $\Delta x \Delta y > 0$, and decreased otherwise (Fig. 1a), corresponding to $u_1 > 0$, $u_2 = 0$ and $u_1 = 0$, $u_2 > 0$, respectively. Unfortunately, this solution still suffers from the second issue highlighted above, as Δx and Δy can take on negative values. To avoid this, we must keep track of whether x is increasing or decreasing via the non-negative indicator signals $x_+ = h(x - x_d)$ and $x_- = h(x_d - x)$ where $h(\cdot)$ is defined as h(w) = 1for w > 0 and h(w) = 0 otherwise. The indicator signals $y_+, y_- \ge 0$ are defined similarly. With this, x needs to be increased $(u_1 > 0, u_2 = 0)$ when either $x_+, y_+ > 0$ or $x_-, y_- > 0$, and it needs to be decreased $(u_1 = 0, u_2 > 0)$ otherwise. This can be implemented via the control law

$$u_1 = x_+ y_+ + x_- y_-, \qquad u_2 = x_+ y_- + x_- y_+ \tag{4}$$

to ensure that trajectories of the closed loop system converge towards the optimum (Fig. 1c), largely unaffected by the value of ϵ_y and t_d (Supplementary Fig. 1). As expected, however, once the lag between xand y (governed by ϵ_y), or the delay between x and x_d or y and y_d (governed by t_d) approach the timescale of the regulator dynamics (governed by ϵ_x), closed loop performance quickly deteriorates.

In summary, the proposed optimizer overcomes the major issues of the gradient-based solution (reliance on the gradient and negative signals) without compromising closed loop performance. As illustrated in Fig. 1c, this is achieved by combining three crucial modules: (i) delay of x and y to obtain x_d and y_d , respectively; (ii) comparison of x with x_d to obtain the indicator signals x_+ and x_- , and similarly for y; and (iii) computation of the control signals u_1 and u_2 based on the indicators according to (4). Before presenting the whole integrated system, we next detail how these idealized and abstract functions can be realized individually by relying on standard synthetic biology modules, and how their biophysical parameters should be selected for optimal performance that is robust to stochastic noise and disturbances.

Delay module

The delayed signal x_d that tracks x with a lag can be implemented using a variety of modules that selectively synthesize/activate the former only in the presence of the latter (and similarly for y_d and y), as illustrated in Fig. 2. Such a positive relationship can be captured via the dynamics

$$\epsilon_d \dot{x}_d = \alpha_d x - x_d, \qquad \epsilon_d \dot{y}_d = \alpha_d y - y_d,$$
 (5)

where $\alpha_d \approx 1$ denotes the production rate constant (following nondimensionalization, see Supplementary Section 1.2), whereas ϵ_d regulates the timescale of the delay module (e.g., greater ϵ_d yields slower dynamics). When $\alpha_d = 1$, we obtain that $x_d \rightarrow x$ and $y_d \rightarrow y$. Thus, the signals x_d and y_d track x and y without steady state error (in case of constant reference signals). We first assume that the delay module is tuned to guarantee perfect tracking ($\alpha_d = 1$). Later, we will demonstrate that the optimizer works even when this is not ensured (provided that $\alpha_d \approx 1$).

The value of ϵ_d needs to be sufficiently small so that tracking is fast enough and the optimizer remains responsive to changes in *x* (and *y*). This observation from Fig. 1c is echoed in the simulation data presented in Fig. 2: as long as the timescale on which the reporter and the delayed signals evolve is sufficiently fast compared to the dynamics of the regulator (i.e., ϵ_d , $\epsilon_y \ll \epsilon_x$), closed loop performance of the optimizer module is largely unaffected by the exact value of ϵ_d and ϵ_y . Unsurprisingly, there is a sharp decline in closed loop performance as the timescale of either the reporter or the delay module approaches that of the regulator (Supplementary Fig. 2).

time

time

time

 $\epsilon_y = 1.8$

= 4.2

= 10.0

 ϵ_u

 $10\epsilon_{\rm v}$

10*e*

10*e*



increases towards the lower right panel where $\epsilon_y = \epsilon_x/10$ (sample points are spaced equidistantly on a logarithmic scale). Simulation parameters and further details are provided in Supplementary Section 5.



Fig. 2 | The delay module ensures tracking of the regulator and the reporter

signals. Light, medium, and dark red correspond to $\epsilon_d = \epsilon_v/2$, $\epsilon_d = \epsilon_v$, and $\epsilon_d = 2\epsilon_v$,

respectively. The panel in the top left corner corresponds to $\epsilon_v = \epsilon_x/100$, and ϵ_v

Fig. 3 | **The comparator module generates the indicator signals based on the actual and delayed signals for both the regulator and the reporter.** Simulation parameters and further details are provided in Supplementary Section 5. **a** The signal *c* alternates between two states (c = 0 and c = 1) with period τ , activating two different sets of regulatory interactions⁹⁴. **b** During phase 1, (x_+ , x_-) tracks the



reference (x, x_d), whereas during phase 2, (x_+ , x_-) converges to either of the stable fixed points based on the sign of $x - x_d$. **c** The signals x_+ and x_- switch between their ON and OFF states depending on whether $x < x_d$ or $x > x_d$ (phase 1 is depicted in gray). **d** Closed loop performance is largely unaffected by the value of $\alpha_{c,2}$.

Comparator module

We next need to compare x with x_d and y with y_d to obtain the indicator signals x_+ , x_- , y_+ , and y_- (Fig. 1c). The comparator module provides these signals by combining three parts: the delay module from the previous section, bistable switches⁴², and an oscillator^{43,44} to periodically activate/deactivate regulatory interactions (Fig. 3a).

For the sake of simplicity, here we assume that the output *c* of a genetic oscillator alternates between two values: c = 0 and c = 1 during phase 1 and phase 2, respectively. We rely on this periodic signal to switch the comparator with inputs x_d and x and outputs x_- and x_+ between two modes according to

$$\begin{split} & \epsilon_{c} \dot{x}_{-} = (1-c) \left(\alpha_{c,1} x_{d} - x_{-} \right) + c \left(\frac{\alpha_{c,2}}{1+x_{+}^{2}} - x_{-} \right), \\ & \epsilon_{c} \dot{x}_{+} = (1-c) \left(\alpha_{c,1} x - x_{+} \right) + c \left(\frac{\alpha_{c,2}}{1+x_{-}^{2}} - x_{+} \right), \end{split}$$

where $\alpha_{c,1}$ and $\alpha_{c,2}$ are production rate constants (de-dimensionalized, see Supplementary Section 1.3), and ϵ_c characterizes the timescale of the comparator dynamics.

During phase 1 (c = 0), the system behaves as the delay module in Fig. 2: the indicator signals x_{+} and x_{-} track the reference signals x and

 x_d , respectively (Fig. 3b). Assuming that this tracking occurs on a timescale faster than that of x and x_d , during phase 1 x and x_d are effectively frozen and x_+ and x_- converge to these values if $\alpha_{c,1}=1$ (Fig. 3c). Therefore, at the end of phase 1, the values of x and x_d are stored in x_+ and x_- , respectively.

During phase 2 (c = 1), the comparator behaves as a bistable switch to implement a memory module, such that the initial conditions of x_{+} and x_{-} are the frozen values of x and x_{d} , respectively. A balanced toggle switch (i.e., $\alpha_{c,2}$ is identical for x_+ and x_-) forces convergence to the x_+ dominated stable fixed point if $x_+ > x_-$ at the beginning of phase 2 (i.e., if $x > x_d$ at the beginning of phase 1), and to the x_- -dominated stable fixed point if $x_- > x_+$ at the beginning of phase 2 (i.e., if $x_d > x$ at the beginning of phase 1). Importantly, the location of these stable fixed points is independent of the difference between x and x_d at the beginning of phase 1 (Fig. 3b and Supplementary Section 1.3). Therefore, at the end of phase 2, the output of the comparator depends only on the sign of $x - x_d$ (at the beginning of phase 1), but not on its magnitude (Fig. 3c), a crucial feature of the proposed design to ensure the compatibility of dynamic ranges between the comparator and the logic module (Fig. 1c). Another comparator with inputs y_d and y and outputs y_{-} and y_{+} behaves similarly.



Fig. 4 | **The logic module combines the indicator signals to generate the control signals.** The dynamic range of the signals in the input, middle (between the AND and OR gates), and output layer of the logic module is denoted by ρ_{Λ} , ρ_{\vee} , and ρ_{u} , respectively. While selecting the dissociation constants K_{Λ} and K_{\vee} in the geometric mean of the respective dynamic ranges may be an intuitive choice (yielding the dark gray lines), the performance of the optimizer displays significant robustness to

deviations from this particular baseline choice when tracking the optimum value \vec{x} (blue). Colored circles correspond to different choices of $\alpha_{c,2}$ (affecting the input dynamic range ρ_{\wedge}), together with substantial perturbations in the dissociation constants compared to the above specified baseline choice. Simulation parameters and further details are provided in Supplementary Section 5.

While $\alpha_{c,1} \approx 1$ is essential for precise tracking, the behavior of the optimizer is considerably less impacted by the value of $\alpha_{c,2}$ as long as bistability is ensured ($\alpha_{c,2} > 2$). For instance, closed loop performance and the time it takes to reach either of the stable fixed points display negligible dependence on $\alpha_{c,2}$ (Fig. 3d and Supplementary Fig. 3b). Although here we assume that the bistable switch is perfectly balanced, we later demonstrate that closed loop performance does not appreciably deteriorate in the presence of moderate levels of imbalance.

Logic module

The final module in the optimizer (Fig. 1c) combines the indicator signals x_+ , x_- , y_+ , and y_- to generate u_1 and u_2 by approximating the control law in (4). This can be done by relying on standard logic gates^{45,46}, as illustrated in Fig. 4. For the sake of simplicity, here we assume that these logic gates operate on a faster timescale than any other module in Fig. 1c (Supplementary Section 1.4). Therefore, the behavior of the logic gates can be approximated by utilizing their quasi-steady state input-output mappings⁴⁵. In particular, considering the common mathematical model of AND (\land) and OR (\lor) gates with input signals *A* and *B* (Supplementary Section 1.4), the outputs are given by

$$H_{\wedge}(A,B) = \frac{A^n}{A^n + K^n_{\wedge}} \frac{B^n}{B^n + K^n_{\wedge}}, \qquad H_{\vee}(A,B) = \frac{A^n}{A^n + K^n_{\vee}} + \frac{B^n}{B^n + K^n_{\vee}},$$

with Hill coefficient *n* and dissociation constants K_{\wedge} and K_{\vee} for the AND and OR gates, respectively. The input-output mapping of the proposed circuit in Fig. 4 is thus given by $u_1 = H_{\vee}(H_{\wedge}(x_+,y_+),H_{\wedge}(x_-,y_-))$ and $u_2 = H_{\vee}(H_{\wedge}(x_+,y_-),H_{\wedge}(x_-,y_+))$.

We next investigate how the dynamic range of the indicator signals, the Hill coefficient, and the dissociation constants of the logic gates affect closed loop performance. As illustrated in Fig. 4, the behavior is insensitive to the dissociation constants K_{\wedge} and K_{\vee} , to the input dynamic range (determined by $\alpha_{c,2}$, see Supplementary Section 1.3), and to the Hill coefficient *n* (Supplementary Fig. 4). Thus, the value of these parameters represent non-critical design choices, unlike the timescale on which the logic module evolves. As expected, the optimizer successfully locates the optimum as long as the dynamics of the logic module evolve on a timescale faster than that of the comparator module (so that the quasi-steady state approximation above is accurate on the slower timescale of the system), however, further speed reduction causes closed loop performance to quickly deteriorate (Supplementary Fig. 5).

Closed loop performance of the simplified optimizer

Before outlining a concrete molecular implementation of the whole integrated system, here we analyze closed loop performance and accuracy when considering the abstract modules introduced above. We first assume that the delay module ensures reference tracking with no error and that the comparator module realizes perfect comparison of its input signals when generating the output indicator signals. We then focus on a less ideal but more practical realization of the optimizer module, and reveal how the presence of errors in delay and comparison affect closed loop performance.

The time it takes to approach the optimum x^{i} depends on the position of the initial value x_0 relative to x^{i} . However, closed loop performance after this initial transient is virtually identical for all (x_0, x^{i}) pairs in the absence of tracking and comparison error (Supplementary Fig. 6a). Importantly, closed loop performance is robust to environmental disturbances that cause drastic shifts in the optimum and also to noise (Fig. 5a). Furthermore, as long as the delay module evolves on a timescale faster than the regulator ($\epsilon_d \ll \epsilon_x$), closed loop performance is insensitive to the value of ϵ_d , as expected: e.g., a 10-fold increase in the value of ϵ_d yields virtually no change (red and green curves in Fig. 5a).

Once the optimum is reached, trajectories oscillate around it due to the periodic behavior of the comparator module. The amplitude of these oscillations increases with τ/ϵ_x , where τ is the period of the oscillator in the comparator module and ϵ_x characterizes the timescale of the regulator dynamics (Supplementary Section 1.5). This result can be interpreted as follows: lower values of ϵ_x correspond to faster control (yielding more rapid changes in *x* over the same time interval), whereas greater values of τ correspond to longer durations when the control action is maintained (but not recalculated and adjusted even if the optimum is reached and crossed). Therefore, decreasing ϵ_x or



Fig. 5 | **Closed loop performance and accuracy with the simplified dynamics.** Simulation parameters and further details are provided in Supplementary Section 5. **a** In the absence of additive noise (dark red and dark green), trajectories are confined within the gray region around the time-varying optimum x^{*} (blue) when $\epsilon_{y} = 0$ (Supplementary Section 1.5). In the presence of stochastic noise (light red and

light green), closed loop trajectories may temporarily leave this region. The value of ϵ_d is 10-times greater for (light and dark) green than for (light and dark) red. **b** Performance decreases as the tracking error in the delay module increases. **c** Shaded regions correspond to initial conditions such that trajectories converge to an incorrect stable fixed point.

increasing τ , thus increasing the ratio τ/ϵ_{xr} eventually leads to more pronounced (occasional) steps in the wrong direction, before these errors are inevitably corrected. These unwanted oscillations can be mitigated by rendering the dynamics of the regulator slower (increasing ϵ_x), or by selecting a clock with shorter period (decreasing τ).

The results so far are underpinned by two crucial assumptions: (i) perfect tracking by the delay module ($\alpha_d = 1$); and (ii) perfect comparison by the comparator module ($\alpha_{c,2}$ is identical for both species). Therefore, we next explore how closed loop performance is affected when these assumptions do not hold true. Tracking occurs with steady state error when $\alpha_d \neq 1$ (Fig. 5b), thus we may erroneously conclude that *x* is increasing even though $x_d > x$, simply because $\alpha_d > 1$ (Fig. 5b). Similarly, unless $\alpha_{c,2}$ is identical for both species in the comparators, comparison of the indicator signals x_{-} and x_{+} can introduce error as trajectories may incorrectly converge to the x_+ -dominated stable fixed point even when $x_- > x_+$ at the start of phase 2 (red in Fig. 5c). These errors then propagate through the logic module, eventually leading to incorrect control signals. Tracking error $(\alpha_d \neq 1)$ leads to more pronounced oscillations around the optimum, whereas substantial asymmetry in $\alpha_{c,2}$ can render the optimizer unreliable (red in Fig. 5c). However, reducing this asymmetry to more modest levels (for instance, via RBS and promoter engineering, or by leveraging decoy sites, see Supplementary Section 1.3) eliminates such performance degradation (green in Fig. 5c). Therefore, while closed loop performance may deteriorate with overwhelming errors in either tracking or comparison, the optimizer displays robust behavior in the presence of moderate levels of error.

Implementation with existing synthetic biology modules

The proposed genetic optimizer can be readily implemented by combining existing synthetic biology modules relying on standard parts and components, as we illustrate in Fig. 6. To ensure correct functioning of the optimizer, modules evolve on two different timescales: the regulator, the reporter, the delay module, and the phase selector oscillator are governed by dynamics that are slower than those of the toggle switch-based memory modules and the logic gates, which guides the design choices in Fig. 6.

Production and removal of the regulator *x* is coordinated by the outputs of the logic module: while u_1 activates synthesis of *x*, u_2 stimulates its degradation by upregulating the transcription of the protease *v* that targets *x* by recognizing the degradation tag fused to it. For instance, appending the *Mesoplasma florum* (*M. florum*) ssrA tag (mf-ssrA) to the C terminus of *x* results in its inducible degradation upon expression of the *M. florum* Lon protease (mf-Lon) *v*, operating independently from and orthogonally to the host degradation pathways^{47,48}. Alternatively, instead of expressing the mf-Lon protease, *v* can encode the adapter SspB, tethering protein targets fused with a modified ssrA tag to ClpXP for controlled degradation⁴⁹. To ensure that the removal of *x* is primarily due to controlled degradation via *v*, the regulator is also equipped with weak self-activation to balance the impact of dilution due to cell growth.

The rest of the genetic optimizer operating on the slower timescale can be implemented using common transcription factors⁵⁰. For instance, the phase selector oscillator can be realized using the repressilator^{43,44} that consists of three repressor proteins, one of them (*r*) co-expressed with an activator (*a*), so that these two can together coordinate periodic shifting between phase 1 and phase 2 of the comparators. Similarly, transcriptional control underpins the implementation of delaying *x* and *y* to obtain x_d and y_d . While in Fig. 6 we assume that the regulator *x* impacts *y* via transcriptional activation and that *y* acts as an activator, crucially, the proposed optimizer module works even when the impact of *x* on *y* is mediated via more complex pathways, or when the reporter acts as a repressor, as we demonstrate in the subsequent section.

As RNA-based solutions enable considerably faster information processing, the indicator signals are implemented by relying on CRISPRi, thus x_+ , x_- , y_+ , and y_- are gRNAs. During phase 1, x_+ and x_- are



Fig. 6 | **Genetic layout of the optimizer module.** The realization relies on genetic parts and modules that are already available, in particular: (i) protein-based transcriptional control⁵⁰; (ii) inducible degradation via the *M. florum* Lon protease and ssrA tag^{47,48}; (iii) a repressilator-based oscillator^{43,44}; (iv) CRISPRi-based toggle switches⁵¹; and (v) STAR-based logic gates^{53,54}. The mass action kinetics-based mathematical model underpinning the dynamics of the integrated system is

included in Supplementary Section 2.1, together with detailed discussion of the typical range of model parameters in Supplementary Section 2.2, and their selected values in Supplementary Table 1. Here, we assume that the host genome is already equipped with a dCas9 expression cassette⁹⁵, otherwise the optimizer must also include it.

expressed from promoters under the induction of *x* and x_d to track them with a delay. During phase 2, x_+ and x_- mutually repress each other to form a CRISPRi-based bistable switch by recruiting dCas9⁵¹. Comparison of *y* and y_d is realized similarly via y_+ and y_- . To alternate between phase 1 and phase 2, promoters of the gRNAs are under the coordinated control of the repressor *r* and the activator *a* (Fig. 6). To ensure reliable switching between the phases, promoter leakiness could be minimized via stringent multi-level control of gene expression⁵².

Finally, the logic module is realized using STAR-based AND and OR gates due to their highly programmable and orthogonal nature^{53,54}. To interface these with the CRISPRi-based indicator signals, we first co-express STARs (\bar{x}_+ , \bar{x}_- , \bar{y}_+ , and \bar{y}_-) with the gRNAs of the comparator module (x_+ , x_- , y_+ , and y_-) during phase 2. This can be achieved by duplicating the transcriptional units encoding gRNAs, only this time expressing STARs (Fig. 6). As a result, the AND gates can take these STARs that serve as proxy for the indicator signals to generate the intermediate STARs q_{++} , q_{-+} , q_{-+} , and q_{--} , either by performing signal integration at the level of the target RNA⁵³, or at the level of the STAR

by splitting it into two halves corresponding to the linear and the terminator hairpin binding regions alongside with an interaction sequence to promote their assembly⁵⁴. Similarly, the OR gates take these intermediate STARs to regulate the expression of the transcription factors u_1 and u_2 by interacting with the intrinsic terminator hairpins that prevent their transcription (either q_{++} or q_{--} for the former, and either q_{+-} or q_{-+} for the latter). Production of x is then modulated by u_1 , whereas its controlled removal by u_2 (via v), thus closing the feedback loop.

Each module featured in Fig. 6 has already been successfully implemented, and considering the mass action kinetics-based model of the whole integrated system (Supplementary Section 2.1), the optimizer locates and tracks the optimum when relying on parameter values typical in *E. coli* (Supplementary Section 2.2). Furthermore, closed loop performance is robust to considerable parameter variations, ranging from changes in the period of the phase selector oscillator through shifts in the production and degradation rates of the regulator to the shape of the application-specific objective function (Supplementary Figs. 8–14). Next, we demonstrate that the optimizer

ensures equally robust closed loop performance in a wide variety of contexts.

The optimizer tracks the optimum in diverse contexts

In Fig. 6, we considered the case when the regulator *x* directly activates the production of the reporter *y*. Here, we illustrate not only that closed loop performance remains similar when considering more complex regulatory schemes and different objective functions (detailed mathematical models are provided in Supplementary Section 3), but also that the optimizer requires only minor modifications to implement minimization instead of maximization, or even multispecies optimization (Fig. 7). As a concrete application example, in Fig. 8 we demonstrate that the genetic optimizer can be deployed for optimal growth rate regulation in the presence of cellular stress. In all examples featured in Figs. 7 and 8, we rely on the mass action kinetics-based model of the whole integrated system outlined in Fig. 6 without any changes to the typical biophysical parameter values featured in Supplementary Table 1.

Starting with the case of simple activation in Fig. 7a, while the production rate of v initially increases with x, the overall relationship between the regulator and the reporter may become non-monotonic due to metabolic burden⁹. Considering either a non-smooth objective function or a smooth bell shaped curve (Supplementary Figs. 11-14), the optimizer successfully locates the optimum, demonstrating that the shape of the objective function has negligible impact on closed loop performance. Closed loop performance also remains practically identical when the impact of x on y is mediated via a regulatory cascade that may even contain self-regulation (Fig. 7a). The behavior is robust to changes in the parameter values of the regulated pathway as well. To illustrate this, consider first the regulatory cascade in Fig. 7b equipped with a negative feedback loop. The optimizer quickly locates the new optimum and reliably tracks it subsequent to abrupt changes in the dissociation constants κ_{x} , κ_{w} , and κ_{z} . The optimum is also rapidly reached when x impacts the expression of y via an incoherent type-1 feed-forward loop after substantial perturbations in the dissociation constant $\kappa_{\rm r}$, the inducer concentration w, or the production rate constant β_{γ} (Fig. 7b).

To locate and track the minimum instead of the maximum, gradient descent can be realized by simply swapping u_1 and u_2 either in the output stage of the logic module so that the former promotes removal of x (via the protease v) and the latter its synthesis (alternative #1 in Fig. 7c), or when activating the expression of x and v, respectively (alternative #2 in Fig. 7c). This may prove helpful, for instance, when we wish to maximize the concentration of the repressor \tilde{y} , thus it may not be used to directly obtain the delayed signal y_d . While we could coexpress an activator together with \tilde{y} and rely on the latter to maximize the former, this may result in additional metabolic burden and issues related to compositional context⁸. Alternatively, we can minimize a target y of the repressor \tilde{y} , thus indirectly maximizing \tilde{y} without any negative impact on closed loop performance (Fig. 7c).

To demonstrate that the proposed optimizer can be deployed even when multiple regulators need to be tuned together to locate the optimum, consider the multi-species optimization problem depicted in Fig. 7d. Instead of seeking the optimum $x^* = (x_1^*, x_2^*, ..., x_N^*)$ in all coordinates simultaneously, we focus on each controlled variable x_i separately. By decomposing the multi-dimensional optimization problem into a collection of *N* one-dimensional problems, each of them takes the form of the previously considered scalar case. Thus, we can deploy a realization of the controller from Fig. 6 for the regulation of each x_i (their independently controlled removal can be realized by complementing the *M. florum*-based solution in Fig. 6 with orthogonal protease:cleavage site pairs from *Potyvirus*⁵⁵), activated one-by-one in a cyclical structure, for instance, by using a ring oscillator with *N* species⁵⁶, as illustrated in Fig. 7e. Data in Fig. 7d confirm that closed loop performance does not deteriorate even

Finally, consider growth rate control in the presence of cellular stress (Fig. 8). Termed as the speedometer of growth rate⁵⁷, the alarmone (p)ppGpp is the primary regulator of both growth and RNA synthesis during exponential growth^{58,59} via a strong inverse relationship^{60,61}. To ensure optimal (p)ppGpp concentration, cells carefully balance the expression levels of RelA/SpoT Homolog (RSH) proteins to catalyze the synthesis and hydrolysis of (p)ppGpp⁶². Leveraging this, the feedforward controller developed in ref. 63 and outlined in Fig. 8a can be deployed to exogenously adjust growth rate via SpoTH. While expression of SpoTH has a positive impact on growth rate at moderate levels as a result of (p)ppGpp hydrolysis⁶⁴, its overexpression can result in considerable metabolic burden, thus giving rise to the non-monotonic relationship between SpoTH expression and growth rate observed in ref. 63 and illustrated in Fig. 8b. By placing the protein y under the control of an rrn P1 promoter, we create a proxy signal for ribosome production and growth rate^{65,66}, and ensure that the optimizer indirectly maximizes growth rate via y by leveraging that rRNA synthesis is the rate-limiting step in ribosome production⁵⁸, yielding a strong inverse relationship between (p)ppGpp concentration and both ribosome synthesis and growth⁶⁷. Data presented in Fig. 8c, d demonstrate that (i) optimal regulation of SpoTH, and thus of (p)ppGpp expression can substantially mitigate the negative impact of cellular stress on growth rate; (ii) performance with the optimizer approaches the theoretical optimum that can be achieved by careful exogenous tuning of SpoTH expression; and (iii) closed loop behavior is robust to considerable stochastic noise and parameter variations (Supplementary Figs. 16-20).

Taken together, these results demonstrate that the optimizer can be successfully deployed in a variety of contexts when relying on (i) an implementation combining existing synthetic biology parts and components; (ii) mass action kinetics-based dynamics; and (iii) biologically plausible parameter values typical in *E. coli*.

Discussion

Synthetic biology promises to revolutionize multiple sectors ranging from metabolic engineering²⁵ to sustainable biomanufacturing⁶⁸. By rearranging existing regulatory linkages and introducing novel ones, experimental techniques complemented by computational tools^{69,70} offer a particularly lucrative research direction to utilize cells as microscopic factories for the dynamic regulation of metabolic pathways. Accordingly, recent years have seen an explosion of studies focusing on controller design for genetic systems^{31-40,71} to ensure that synthetic circuits function robustly in different host organisms and cellular contexts. Complementing these efforts, the optimizer module developed here acts as a feedback controller and steers the regulated genetic system towards the optimum, but crucially, without having explicit knowledge of either its location, the objective function itself, or the pathway that needs to be regulated.

Our results also highlight how existing regulatory solutions based on exogenous control can be enhanced by combining them with the proposed genetic optimizer without compromising performance or requiring any modification to already developed pathways. Both the regulator *x* and the reporter *y* can be implemented using widely available transcription factors⁵⁰, and synthesis of *y* can be made responsive to a multitude of signals and analytes relying on the vast collection of protein-based and RNA-based biosensors^{72–75}. For instance, transcription of *y* can be modulated by expressing it from the cognate promoter of a ligand inducible transcription factor^{73,76}, whereas metabolite-sensing riboswitches can be deployed to directly impact translation initiation^{77,78}. The optimizer can thus be easily interfaced with the vast collection of biosensors that is continuously expanded via genome mining, rational design, and screening based on directed evolution^{74,75,79,80}. Consequently, we expect that our



Fig. 7 | **The genetic optimizer can be successfully deployed in diverse contexts.** Detailed mathematical models and additional data are provided in Supplementary Section 3, together with simulation parameters in Supplementary Section 5. In **a**, **c**, mean and error bars denote the average of *x* and its standard deviation, averaged over 100 independent simulations with randomly selected initial conditions. In **b**, **d**, thin red curves correspond to 30 independent closed loop trajectories with random initial points. **a** The optimizer locates the static optimum. **b** The

optimizer tracks the time-varying optimum (blue) as parameters fluctuate (indicated by the arrowheads). **c** The expression of \tilde{y} can be maximized by minimizing *y*. **d** The optimizer tracks the time-varying optimum (blue) even when *y* is regulated by multiple species. The thick red curves denote the average of 30 independent simulations. **e** Genetic layout of the multi-dimensional optimizer re-using and modifying the modules originally featured in Fig. 6.

genetic optimizer can be deployed in a variety of domains to regulate cellular performance along multiple dimensions for dynamic and optimal control of microbial cell factories, to ensure the affordable and sustainable synthesis of a vast array of products ranging from biofuels to pharmaceuticals^{81,82}. We anticipate the main limitation to stem from the delay between changes in the regulator x and its impact on the reporter y, as excessive lag can lead to deteriorating performance (Supplementary Fig. 15). Additionally, while the optimizer is unable to respond to shifts that are faster than the oscillator period, it can successfully track changes that occur less rapidly (Figs. 7b, d and 8d and Supplementary Fig. 17). As a result, we expect

that the proposed implementation of the optimizer is especially well-suited in contexts where shifts in the culture environment occur slowly (on a timescale ranging from hours to days), which is typical in many fed-batch culture-based metabolic engineering applications⁸³⁻⁸⁵.

The biomolecular circuit outlined in Fig. 6 is underpinned by existing synthetic biology modules (although their integration may require considerable experimental fine-tuning), and it can be successfully deployed in diverse settings as we illustrate through multiple application examples in Figs. 7 and 8 to ensure optimal closed loop performance that is robust to parameter variations and stochastic



Fig. 8 | The genetic optimizer can be deployed to maximize cellular growth rate. Detailed mathematical model and additional data are provided in Supplementary Section 3, together with simulation parameters in Supplementary Section 5. Blue curves indicate performance when SpoTH is exogenously optimized by adjusting the inducer concentration. Green and purple curves denote trajectories with zero and maximal induction of SpoTH. Closed loop performance is evaluated in the presence of stochastic noise impacting the kinetics of all species. Cellular stress is modulated via β_p^{63} . **a** Growth rate is negatively impacted by rising levels of the alarmone (p)ppGpp (*p*) as it downregulates the production of ribosomes (*z*). Cellular stress results in elevated RelA (*r*) expression, which upregulates the synthesis of (p)ppGpp via the increased production rate constant β_p . Conversely, (p)ppGpp concentration can be decreased via SpoTH (*s*) by activating its expression either exogenously⁶³ or by placing its promoter under the control of the regulator *x*. The dashed red flat headed arrows from SpoTH

noise, without requiring any application-specific tuning of its parameters. Considering the parameter values summarized in Supplementary Table 1 and their typical ranges discussed in Supplementary Section 2.2, the concrete molecular implementation of the whole integrated system in Fig. 6 is expected to have a bioenergetic cost similar to genetic circuits⁸⁶ that have already been implemented without negatively impacting growth rate (Supplementary Section 4). Finally, as the coarse-grained mechanistic model underpinning our approach reveals fundamental principles that ensure correct functioning, we envision that other realizations of the proposed genetic optimizer will be implemented by incorporating synthetic components from a diverse set of kits similar to recent approaches^{52,87}, combining plasmid copy number control⁸⁸, transcriptional regulation^{42,43}, post-translational modification⁸⁹, RNA-based interactions^{36,71}, and especially fast phosphotransfer processes⁹⁰.

Biomolecular controllers offer a promising avenue towards robust and optimal gene regulation, an essential feature of complex synthetic systems. The rationally designed pathway-independent optimizer presented here builds on ideas rooted in control theory to implement a genetic feedback module that ensures convergence to, and tracking of, the time-varying optimum even in the presence of disturbances and stochastic noise. The blueprint of this module can be realized in a versatile fashion, relying on the large variety of already existing bioengineering parts.

Methods

Deterministic simulations

Data corresponding to deterministic dynamics were obtained using the ode45 solver of MATLAB.

Stochastic simulations

Data corresponding to stochastic dynamics were obtained as follows. Define the discretization $0 = t_0 < t_1 < \cdots < t_N = T$ over the time interval [0, T] with $t_i = i\delta$ where $\delta = T/N$ is the stepsize, and introduce $\Delta W_n = W(t_{n+1}) - W(t_n)$ where *W* is a *d*-dimensional standard Wiener process. The solution of the *d*-dimensional system of (Itô) stochastic differential equations

 $dX(t) = \mu(X(t), \Theta)dt + \sigma(X(t), \Theta)dW(t), \qquad X(0) = x_0,$

represent the load that SpoTH expression places on ribosomes as its mRNA is translated. **b** Expression of SpoTH results in sequestration of shared cellular resources, thus the metabolic burden due to SpoTH overexpression can counteract the positive impact of (p)ppGpp removal on growth rate, resulting in a nonmonotonic relationship. **c** Closed loop performance is evaluated based on 100 independent simulations with random initial conditions during the second half of each simulation by considering the average of *y* and its standard deviation. Red curve and red shaded region denote the mean of these averages and standard deviations, respectively. **d** The optimizer successfully tracks the time-varying optimum in response to both abrupt and gradual changes in β_p representing cellular stress. Red curves and shaded regions correspond to the mean and standard deviation of trajectories considering 100 independent simulations with random initial conditions, see Supplementary Fig. 17.

with $\mu \in \mathbb{R}^d$ drift and $\sigma \in \mathbb{R}^{d \times d}$ diffusion coefficient according to the multi-dimensional Euler–Maruyama scheme^{91,92} is given by

$$X_{n+1} = X_n + \delta \mu(X_n, \Theta) + \sigma(X_n, \Theta) \Delta W_n, \qquad X_0 = x_0,$$

where X_i denotes the numerical approximation of $X(t_i)$.

Data collection

Simulation data were generated using MATLAB (version R2022b).

Data analysis

Simulation data were analyzed and plotted using MATLAB (version R2022b). Figures were created using Adobe Illustrator (version 26.5). Overleaf (version v2) online LaTeX editor was used to prepare and compile the manuscript.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Source data are provided with this paper.

Code availability

The manuscript does not rely on custom mathematical algorithms or software. Simulation data were generated and analyzed as described in the "Methods" by considering the models detailed in the Supplementary Information using built-in MATLAB functions. The MATLAB code for implementing and deploying the optimizer for the examples featured in the manuscript, along with instructions on how to modify it for other application examples, are publicly available at https://github.com/qbionet/genetic-optimizer. Additional information is available from the corresponding author upon request.

References

- Csete, M. & Doyle, J. Reverse engineering of biological complexity. Science 295, 1664–1669 (2002).
- Purnick, P. E. M. & Weiss, R. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* 10, 410–422 (2009).

Article

- 3. Del Vecchio, D., Dy, A. J. & Qian, Y. Control theory meets synthetic biology. J. R. Soc. Interface **13**, 20160380 (2016).
- Brophy, J. A. N. & Voigt, C. A. Principles of genetic circuit design. Nat. Methods 11, 508–520 (2014).
- Smanski, M. J. et al. Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* 32, 1241–1249 (2014).
- 6. Cardinale, S. & Arkin, A. P. Contextualizing context for synthetic biology-identifying causes of failure of synthetic biological systems. *Biotechnol. J.* **7**, 856–866 (2012).
- Borkowski, O., Ceroni, F., Stan, G. & Ellis, T. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Curr. Opin. Microbiol.* **33**, 123–130 (2016).
- Yeung, E. et al. Biophysical constraints arising from compositional context in synthetic gene networks. *Cell Syst.* 5, 11.e12–24.e12 (2017).
- Qian, Y., Huang, H.-H., Jiménez, J. I. & Vecchio, D. D. Resource competition shapes the response of genetic circuits. ACS Synth. Biol. 6, 1263–1272 (2017).
- Moser, F. et al. Genetic circuit performance under conditions relevant for industrial bioreactors. ACS Synth. Biol. 1, 555–564 (2012).
- Vilanova, C. et al. Standards not that standard. J. Biol. Eng. 9, 17 (2015).
- 12. Gyorgy, A. & Del Vecchio, D. Modular composition of gene transcription networks. *PLoS Comput. Biol.* **10**, 1–16 (2014).
- Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102 (2010).
- Ceroni, F., Algar, R., Stan, G.-B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* 12, 415–418 (2015).
- 15. Gyorgy, A. et al. Isocost lines describe the cellular economy of genetic circuits. *Biophys. J* **109**, 639–646 (2015).
- Weiße, A. Y., Oyarzún, D. A., Danos, V. & Swain, P. S. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proc. Natl Acad. Sci. USA* **112**, E1038–E1047 (2015).
- Raveh, A., Margaliot, M., Sontag, E. D. & Tuller, T. A model for competition for ribosomes in the cell. J. R. Soc. Interface 13, 20151062 (2016).
- Gorochowski, T. E., Avcilar-Kucukgoze, I., Bovenberg, R. A. L., Roubos, J. A. & Ignatova, Z. A Minimal model of ribosome allocation dynamics captures trade-offs in expression between endogenous and synthetic genes. ACS Synth. Biol. 5, 710–720 (2016).
- 19. Frei, T. et al. Characterization and mitigation of gene expression burden in mammalian cells. *Nat. Commun.* **11**, 4641 (2020).
- Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* 27, 946–950 (2009).
- Mutalik, V. K. et al. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* 10, 354–360 (2013).
- Pfleger, B. F., Pitera, D. J., Smolke, C. D. & Keasling, J. D. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat. Biotechnol.* 24, 1027–1032 (2006).
- Alper, H., Fischer, C., Nevoigt, E. & Stephanopoulos, G. Tuning genetic control through promoter engineering. *Proc. Natl Acad. Sci.* USA 102, 12678–12683 (2005).
- Holtz, W. J. & Keasling, J. D. Engineering static and dynamic control of synthetic pathways. *Cell* **140**, 19–23 (2010).
- Brockman, I. M. & Prather, K. L. J. Dynamic metabolic engineering: New strategies for developing responsive cell factories. *Biotechnol.* J. 10, 1360–1369 (2015).
- Tan, S. Z. & Prather, K. L. Dynamic pathway regulation: recent advances and methods of construction. *Curr. Opin. Chem. Biol.* 41, 28–35 (2017).

- 27. Bandiera, L. et al. On-line optimal input design increases the efficiency and accuracy of the modelling of an inducible synthetic promoter. *Processes* **6**, 148 (2018).
- Sakurai, Y. & Hori, Y. Optimization-based synthesis of stochastic biocircuits with statistical specifications. J. R. Soc. Interface 15, 20170709 (2018).
- Bokes, P., Hojcka, M. & Singh, A. MicroRNA based feedforward control of intrinsic gene expression noise. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 18, 272–282 (2021).
- Gorochowski, T. E., Grierson, C. S. & di Bernardo, M. Organization of feed-forward loop motifs reveals architectural principles in natural and engineered networks. *Sci. Adv.* 4, eaap9751 (2018).
- Perrino, G., Hadjimitsis, A., Ledesma-Amaro, R. & Stan, G.-B. Control engineering and synthetic biology: working in synergy for the analysis and control of microbial systems. *Curr. Opin. Microbiol.* 62, 68–75 (2021).
- 32. Hu, C. Y. & Murray, R. M. Layered feedback control overcomes performance trade-off in synthetic biomolecular networks. *Nat. Commun.* **13**, 5393 (2022).
- Briat, C., Gupta, A. & Khammash, M. Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell Syst.* 2, 15–26 (2016).
- Olsman, N. et al. Hard limits and performance tradeoffs in a class of antithetic integral feedback networks. *Cell Syst.* 9, 49.e16–63.e16 (2019).
- Aoki, S. K. et al. A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature* 570, 533–537 (2019).
- Huang, H.-H., Qian, Y. & Del Vecchio, D. A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *Nat. Commun.* 9, 5415 (2018).
- Franco, E., Giordano, G., Forsberg, P.-O. & Murray, R. M. Negative autoregulation matches production and demand in synthetic transcriptional networks. ACS Synth. Biol. 3, 589–599 (2014).
- Agrawal, D. K., Marshall, R., Noireaux, V. & Sontag, E. D. In vitro implementation of robust gene regulation in a synthetic biomolecular integral controller. *Nat. Commun.* **10**, 5760 (2019).
- Lillacci, G., Benenson, Y. & Khammash, M. Synthetic control systems for high performance gene expression in mammalian cells. *Nucleic Acids Res.* 46, 9855–9863 (2018).
- Chevalier, M., Gómez-Schiavon, M., Ng, A. H. & El-Samad, H. Design and analysis of a proportional-integral-derivative controller with biological molecules. *Cell Syst.* 9, 338.e10–353.e10 (2019).
- Hidalgo, D., Martínez-Ortiz, C. A., Palsson, B. O., Jiménez, J. I. & Utrilla, J. Regulatory perturbations of ribosome allocation in bacteria reshape the growth proteome with a trade-off in adaptation capacity. *iScience* 25, 103879 (2022).
- Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *E. coli. Nature* **403**, 339–342 (2000).
- 43. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
- Potvin-Trottier, L., Lord, N. D., Vinnicombe, G. & Paulsson, J. Synchronous long-term oscillations in a synthetic gene circuit. *Nature* 538, 514–517 (2016).
- Andrews, L. B., Nielsen, A. A. K. & Voigt, C. A. Cellular checkpoint control using programmable sequential logic. *Science* 361, eaap8987 (2018).
- 46. Matsuura, S. et al. Synthetic RNA-based logic computation in mammalian cells. *Nat. Commun.* **9**, 4847 (2018).
- 47. Cameron, D. E. & Collins, J. J. Tunable protein degradation in bacteria. *Nat. Biotechnol.* **32**, 1276–1281 (2014).
- Szydlo, K., Ignatova, Z. & Gorochowski, T. E. Improving the robustness of engineered bacteria to nutrient stress using programmed proteolysis. ACS Synth. Biol. 11, 1049–1059 (2022).
- McGinness, K. E., Baker, T. A. & Sauer, R. T. Engineering controllable protein degradation. *Mol. Cell* 22, 701–707 (2006).

- Meyer, A. J., Segall-Shapiro, T. H., Glassey, E., Zhang, J. & Voigt, C. A. Escherichia coli "Marionette" strains with 12 highly optimized small-molecule sensors. *Nat. Chem. Biol.* **15**, 196–204 (2019).
- Santos-Moreno, J., Tasiudi, E., Stelling, J. & Schaerli, Y. Multistable and dynamic CRISPRi-based synthetic circuits. *Nat. Commun.* 11, 2746 (2020).
- Greco, F. V., Pandi, A., Erb, T. J., Grierson, C. S. & Gorochowski, T. E. Harnessing the central dogma for stringent multi-level control of gene expression. *Nat. Commun.* 12, 1738 (2021).
- 53. Chappell, J., Takahashi, M. K. & Lucks, J. B. Creating small transcription activating RNAs. *Nat. Chem. Biol.* **11**, 214–220 (2015).
- Chappell, J., Westbrook, A., Verosloff, M. & Lucks, J. B. Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. *Nat. Commun.* 8, 1051 (2017).
- Fernandez-Rodriguez, J. & Voigt, C. A. Post-translational control of genetic circuits using Potyvirus proteases. *Nucleic Acids Res.* 44, 6493–6502 (2016).
- 56. Niederholtmeyer, H. et al. Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife* **4**, e09771 (2015).
- Imholz, N. C. E., Noga, M. J., Broek, N. J. Fvd & Bokinsky, G. Calibrating the bacterial growth rate speedometer: a re-evaluation of the relationship between basal ppGpp, growth, and RNA synthesis in *Escherichia coli*. Front. Microbiol. **11**, 574872 (2020).
- Paul, B. J., Ross, W., Gaal, T. & Gourse, R. L. rRNA transcription in Escherichia coli. Annu. Rev. Genet. 38, 749–770 (2004).
- Potrykus, K., Murphy, H., Philippe, N. & Cashel, M. ppGpp is the major source of growth rate control in *E. coli. Environ. Microbiol.* 13, 563–575 (2011).
- Bremer, H. & Dennis, P. P. Modulation of chemical composition and other parameters of the cell by growth rate. *Escherichia Coli Salmonella Cell. Mol. Biol.* 2, 1553–1566 (1996).
- Dai, X. & Zhu, M. Coupling of ribosome synthesis and translational capacity with cell growth. *Trends Biochem. Sci.* 45, 681–692 (2020).
- Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T. & Gerdes, K. Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat. Rev. Microbiol.* **13**, 298–309 (2015).
- Barajas, C., Huang, H.-H., Gibson, J., Sandoval, L. & Vecchio, D. D. Feedforward growth rate control mitigates gene activation burden. *Nat. Commun.* 13, 7054 (2022).
- Gentry, D. R. & Cashel, M. Mutational analysis of the Escherichia coli spoT gene identifies distinct but overlapping regions involved in ppGpp synthesis and degradation. *Mol. Microbiol.* **19**, 1373–1384 (1996).
- Sarubbi, E., Rudd, K. E. & Cashel, M. Basal ppGpp level adjustment shown by new spoT mutants affect steady state growth rates and rrnA ribosomal promoter regulation in *Escherichia coli. Mol. Gen. Genet.* 213, 214–222 (1988).
- Murray, H. D., Schneider, D. A. & Gourse, R. L. Control of rRNA expression by small molecules is dynamic and nonredundant. *Mol. Cell* 12, 125–134 (2003).
- Hernandez, V. J. & Bremer, H. Guanosine tetraphosphate (ppGpp) dependence of the growth rate control of rrnB P1 promoter activity in *Escherichia coli*. J. Biol. Chem. 265, 11605–11614 (1990).
- Peralta-Yahya, P. P., Zhang, F., del Cardayre, S. B. & Keasling, J. D. Microbial engineering for the production of advanced biofuels. *Nature* 488, 320–328 (2012).
- Wang, X., Rai, N., Merchel Piovesan Pereira, B., Eetemadi, A. & Tagkopoulos, I. Accelerated knowledge discovery from omics data by optimal experimental design. *Nat. Commun.* **11**, 5026 (2020).
- Angenent-Mari, N. M., Garruss, A. S., Soenksen, L. R., Church, G. & Collins, J. J. A deep learning approach to programmable RNA switches. *Nat. Commun.* **11**, 5057 (2020).
- Kelly, C. L. et al. Synthetic negative feedback circuits using engineered small RNAs. *Nucleic Acids Res.* 46, 9875–9889 (2018).

- Eggeling, L., Bott, M. & Marienhagen, J. Novel screening methods-biosensors. *Curr. Opin. Biotechnol.* 35, 30–36 (2015).
- Rogers, J. K., Taylor, N. D. & Church, G. M. Biosensor-based engineering of biosynthetic pathways. *Curr. Opin. Biotechnol.* 42, 84–91 (2016).
- Hanko, E. K. R. et al. A genome-wide approach for identification and characterisation of metabolite-inducible systems. *Nat. Commun.* 11, 1213 (2020).
- Qin, L., Liu, X., Xu, K. & Li, C. Mining and design of biosensors for engineering microbial cell factory. *Curr. Opin. Biotechnol.* **75**, 102694 (2022).
- Chen, Y. et al. Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. *Nat. Commun.* 9, 64 (2018).
- 77. Serganov, A. & Nudler, E. A decade of riboswitches. *Cell* **152**, 17–24 (2013).
- 78. Chauvier, A. et al. Monitoring RNA dynamics in native transcriptional complexes. *Proc. Natl Acad. Sci. USA* **118**, e2106564118 (2021).
- Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. Computational design of receptor and sensor proteins with novel functions. *Nature* 423, 185–190 (2003).
- Boussebayle, A. et al. Next-level riboswitch developmentimplementation of Capture-SELEX facilitates identification of a new synthetic riboswitch. *Nucleic Acids Res.* 47, 4883–4895 (2019).
- Ro, D.-K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440, 940–943 (2006).
- Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* 451, 86–89 (2008).
- Antoniewicz, M. R. et al. Metabolic flux analysis in a nonstationary system: Fed-batch fermentation of a high yielding strain of *E. coli* producing 1,3-propanediol. *Metab. Eng.* 9, 277–292 (2007).
- Toussaint, C., Henry, O. & Durocher, Y. Metabolic engineering of CHO cells to alter lactate metabolism during fed-batch cultures. J. Biotechnol. 217, 122–131 (2016).
- Mulukutla, B. C. et al. Metabolic engineering of Chinese hamster ovary cells towards reduced biosynthesis and accumulation of novel growth inhibitors in fed-batch cultures. *Metab. Eng.* 54, 54–68 (2019).
- Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C. & Voigt, C. A. Genetic programs constructed from layered logic gates in single cells. *Nature* 491, 249–253 (2012).
- Darlington, A. P. & Bates, D. G. Architectures for combined transcriptional and translational resource allocation controllers. *Cell* Syst. 11, 382.e9–392.e9 (2020).
- Joshi, S. H.-N., Yong, C. & Gyorgy, A. Inducible plasmid copy number control for synthetic biology in commonly used *E. coli* strains. *Nat. Commun.* 13, 6691 (2022).
- 89. Barkai, N. & Leibler, S. Robustness in simple biochemical networks. *Nature* **387**, 913–917 (1997).
- 90. Mishra, D. et al. An engineered protein-phosphorylation toggle network with implications for endogenous network discovery. *Science* **373**, eaav0780 (2021).
- Higham, D. J. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.* 43, 525–546 (2001).
- Leobacher, G. & Szölgyenyi, M. Convergence of the Euler-Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient. *Numer. Math.* **138**, 219–239 (2018).
- 93. Jones, K. L., Kim, S.-W. & Keasling, J. Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. *Metab. Eng.* **2**, 328–338 (2000).
- Bordoy, A. E., O'Connor, N. J. & Chatterjee, A. Construction of twoinput logic gates using transcriptional interference. ACS Synth. Biol. 8, 2428–2441 (2019).

 Rousset, F. et al. Genome-wide CRISPR-dCas9 screens in *E. coli* identify essential genes and phage host factors. *PLoS Genet.* 14, e1007749 (2018).

Acknowledgements

This work was supported by research funds from New York University Abu Dhabi and by DARPA HRO011-16-2-0044. The authors would like to thank Adam P. Arkin for posing the genetic optimization problem and facilitating its initial study under the DARPA award, and Shivang H. N. Joshi for helpful comments about the parameter values used in this study.

Author contributions

A.G., A.M., and M.A. contributed to the conception and design of the genetic optimizer. A.G. developed the mathematical model of the genetic optimizer with inputs from A.M. and M.A. A.G. analyzed the mathematical model, designed the molecular implementation, collected and analyzed the simulation data, created the figures, and drafted the paper and the revision. A.G., A.M., and M.A. jointly edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-37903-0.

Correspondence and requests for materials should be addressed to Andras Gyorgy.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/ licenses/by/4.0/.

© The Author(s) 2023