Genetic Optimizer Module for Synthetic Biology ⋆

Andras Gyorgy * **Amor Menezes** ** **Murat Arcak** ***

* *New York University Abu Dhabi, UAE (e-mail: andras.gyorgy@nyu.edu)*
** *University of Florida, Gainesville, 32611 FL, USA, (e-mail: amormenezes@ufl.edu)*
*** *University of California, Berkeley, CA, 94720 USA (e-mail: arcak@berkeley.edu).*

Abstract: For optimal performance, living organisms must fine-tune a vast array of complex processes such as metabolism, respiration, and growth. Due to the limited availability of available resources, these decisions are fundamentally intertwined and involve a large collection of unknown trade-offs. While synthetic biology enables us to control the expression of any gene, we currently lack the ability to automatically tune it to its optimal level, thus maximizing/minimizing some user-defined performance metric (e.g., biomass production, growth rate). To obtain this goal, here we present an optimizer module that can be constructed using standard biological parts. This feedback controller module is inspired by classical gradient-based numerical methods, however, it conforms to the unique constraints due to the biological context (e.g., non-negative signals). Importantly, the performance of the optimizer module is robust to parameter variations in the unknown and time-varying plant and objective function, as well as disturbances in the optimizer itself.

Keywords: Optimization, timescale separation, stability analysis, synthetic biology, modularity.

1. INTRODUCTION

Synthetic biology currently lacks rationally engineered cell-based genetic control systems capable of dynamically adjusting the activity profile of gene networks to optimize their performance. Such performance metric could be as simple as maximizing biomass production or growth rate, or more complex, such as the optimal resource allocation between growth and toxin production for outcompeting an invader species Ahmad et al. (2019). This task is especially challenging due to the uncertain and context-dependent nature of biological circuit design.

To tackle this challenge, here we present a feedback controller system that can successfully optimize the performance of an unknown time-varying process with an unknown time-varying objective function, while adhering to biological constraints (e.g., non-negative signals). The developed optimizer module is inspired by gradient-based optimization methods, comprising three main steps: delay of state and output signals, detection of the change in these signals, and the integration of these information to generate the control inputs.

This paper is organized as follows: following the introduction of the mathematical model and the biological constraints, we provide a brief overview of the proposed optimizer module. Finally, we illustrate that performance of the developed optimizer is robust to parameter variations, thus it is well-suited for the cellular context dominated by uncertainty.

⋆ This work was supported by DARPA Grant #???.

2. MATHEMATICAL MODEL AND PROBLEM FORMULATION

2.1 System of Interest and Optimizer

Consider the system $\dot{x} = f(x, u, t)$ with $y = F(x, t)$, where $x, y \in \mathbb{R}$, $u \in \mathbb{R}^2$, and assume that $F(x, t) = 0$ only at the unique optimum $x^*(t)$. Furthermore, we consider the scalar dynamics with $f(x, u, t) = u_1(t)\alpha(t) - u_2(t)\gamma(t)x(t)$, where the rest of the dynamics (e.g., context, disturbances) is captured by the time-varying production rate $\alpha(t)$ and degradation rate $\gamma(t)$.

For the above system, we seek the general feedback optimizer of the form $\dot{z} = g(z, x, y)$, together with $u_i = h_i(z)$ for $i = 1, 2$, such that the closed-loop system x approaches x^* as $t \to \infty$. Importantly, we must have $u_1, u_2 \geq 0$ as the production/degradation rate constants must be non-negative.

2.2 Limitations

In the absence of biological constraints, one solution for the optimizer would be a gradient-based system with a choice of $u_i(t) = -(-1)^i \lambda_i \nabla_x F(x, t)$ where $\nabla_x F(x, t)$ denotes the gradient of $F(x, t)$, and $\lambda_1, \lambda_2 > 0$ for maximization, and $\lambda_1, \lambda_2 < 0$ for minimization. Within the biological context, there are two major problems with this: (i) we do not have access to $\nabla_x F(x, t)$ and to $-\nabla_x F(x, t)$; and (ii) both $u_1(t)$ and $u_2(t)$ can take negative values, which is not physically realizable.

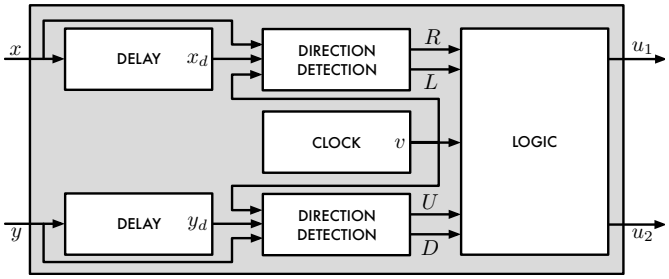

Fig. 1. The optimizer block comprises three main stages.

3. RESULTS

To overcome to above issues, define $\Delta x(t) = x(t) - x(t-\tau)$, $\Delta y(t) = y(t) - y(t-\tau)$, and $I(t) = \Delta x(t)\Delta y(t)$. With this the control law

$$u_1(t) = \begin{cases} \lambda_1 & I(t) > 0, \\ 0 & I(t) < 0, \end{cases}, \qquad u_2(t) = \begin{cases} 0 & I(t) > 0, \\ \lambda_2 & I(t) < 0 \end{cases} \tag{1}$$

ensures not only that the closed-loop dynamics approach the optimal value x^* where $F(x,t)$ is maximized (for minimization, switch the conditions above), but also that it stays close to it following convregence.

To realize the control law in (1), the optimizer relies on three crucial steps (Fig. 1): (*i*) delay of $x(t)$ and $y(t)$; (*ii*) comparison of these signals and their delayed versions; and (*iii*) logic integration to obtain the indicator I. Next, we introduce the modules realizing these function, relying only on standard genetic parts.

1) Signal delay Considering the dynamics $\epsilon_1 \dot{x}_d = x - x_d$ and $\epsilon_1 \dot{y}_d = y - y_d$, the signals x_d and y_d track x and y, respectively, with a time delay that increases with ϵ_1.

2) Direction detection Consider a periodic signal $v(t)$ switching between two values: 0 and 1 (for instance, normalized output of an activator-repressor clock Guantes and Poyatos (2006)). With this, introduce the dynamics

$$\begin{aligned} \epsilon_2 \dot{L} &= v\,(x_d - L) + (1-v)\left(\tfrac{\beta}{1+R^2} - L\right), \\ \epsilon_2 \dot{R} &= v\,(x - R) + (1-v)\left(\tfrac{\beta}{1+L^2} - R\right), \end{aligned} \tag{2}$$

and note that when $v = 1$, we have that $L \to x_d$ and $R \to x$ as $t \to \infty$. Conversely, when $v = 0$, the dynamics in (2) become that of the standard toggle switch Gardner et al. (2000), so that when $\beta \gg 2$ we have that the two stable equilibria are given by approximately β and 0, independent of the initial conditions ($L \approx x_d$, $R \approx x$). Therefore, the initial conditions only determine which stable state we converge to: if L is greater than R (i.e., $x_d > x$) at the end of the first phase (when $v = 1$), then in the second phase (when $v = 0$) we have that $L \to \beta$ and $R \to 0$, and vice versa. Therefore, L and R denote if we are moving Left or Right, respectively, along the x-axis (in the (x, y)-plane). Considering a dynamics like in (2) for D and U, these signals denote if we are moving Down or Up, respectively, along the y-axis (in the (x, y)-plane).

3) Logic integration Finally, at the last stage of the optimizer, we are using standard AND and OR gates to integrate the direction signals L, R, D, and U. To this end, introduce $H(a) = a^n/(1 + a^n)$ and $g(s_1, s_2) = H(s_1/K)H(s_2/K)$ with $K > 0$. With this, the dynamics

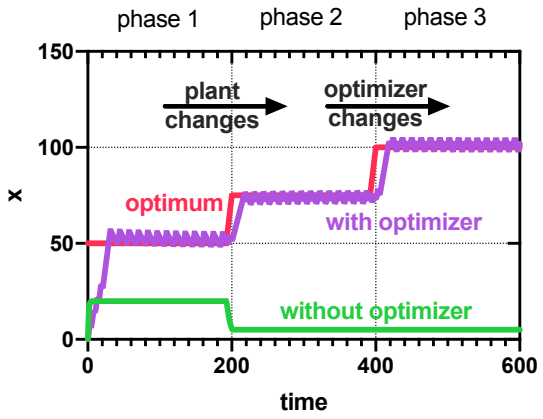

Fig. 2. Without the optimizer, the plant behaves independently of the value of y, whereas with the optimizer the state x tracks the optimum x^*. Throughout the simulation $\beta = 100$, $K = \beta/2$, $\sigma = 10$, $\epsilon_1 = 1$, $\epsilon_2 = \epsilon_3 = 0.01$, $n = 2$, $\lambda_0 = 1$, and in the three phases $\alpha = (20, 10, 10)$, $\gamma = (1, 2, 2)$, $\mu = (50, 75, 100)$, $\lambda_1 = (0.5, 0.25, 0.75)$, $\lambda_2 = (0.05, 0.05, 0.025)$.

$$\begin{aligned} \epsilon_3 \dot{A}_1 &= \lambda_0 g(R, U) - A_1, & \epsilon_3 \dot{B}_1 &= \lambda_0 g(L, D) - B_1, \\ \epsilon_3 \dot{A}_2 &= \lambda_0 g(R, D) - A_2, & \epsilon_3 \dot{B}_2 &= \lambda_0 g(L, U) - B_2, \end{aligned}$$

yield that $A_1 \to \lambda_0$ when R and U are ON, and zero otherwise, and similarly for the other signals. This, together with $\epsilon_3 \dot{u}_i = (1 - v)\left[\lambda_i \left(H\left(A_i\right) + H\left(B_i\right)\right) - u_i\right]$ for $i = 1, 2$ yields (1), the control law for maximizing $F(x,t)$. For minimizing $F(x,t)$, swap U and D above.

4. SIMULATION RESULTS

Consider the objective function $F(x,t) = \exp[(x - \mu(t))^2/(2\sigma^2(t))]$. Without the optimizer, the plant converges to α/γ as $t \to \infty$, which is far from the optimal state $x^* = \mu$ (Fig. 2). Conversely, the optimizer displays robust performance, even in the presence of considerable parameter variations (including those of the plant, the objective function, and even the optimizer) and noise (Fig. 2).

5. CONCLUSION

Here, we outlined the structure of an optimizer module for a wide array of synthetic biology applications, even in the presence of significant parameter uncertainties and noise. The developed optimizer module is inspired by gradient-based optimization methods and it can be constructed using standard and readily available genetic parts.

REFERENCES

Ahmad, S., Wang, B., Walker, M.D., Tran, H.K.R., Stogios, P.J., Savchenko, A., Grant, R.A., McArthur, A.G., Laub, M.T., and Whitney, J.C. (2019). An interbacterial toxin inhibits target cell growth by synthesizing (p)ppApp. *Nature*, 1–5. doi:10.1038/s41586-019-1735-9.

Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767), 339–342.

Guantes, R. and Poyatos, J.F. (2006). Dynamical principles of two-component genetic oscillators. *PLOS Computational Biology*, 2(3), 1–10. doi: 10.1371/journal.pcbi.0020030. .