How Cell-to-Cell Heterogeneity and Scarce Resources Shape the Population-Level Stability Profile of Toggle Switches

Andras Gyorgy

***Abstract*— The lack of modularity in synthetic biology presents one of the major bottlenecks in the scalability of complex gene circuits. One source of this context-dependent behavior is the scarcity of shared transcriptional and translational resources. To overcome this issue, predictive computational tools must account for the resulting competition phenomenon both when studying individual cells and at the population-level considering cell-to-cell heterogeneity. Since toggle switches are one of the most widely used genetic modules, here we focus on how shared resources affect the stability profile of toggle switches even in the presence of loading from their context. Modeling the parameters of the toggle switch as random variables reveals how cellular context, noise and correlation between key parameters shape the population-level stability distribution. To demonstrate the relevance of our results, we illustrate that detrimental effects of even unknown contexts can be bounded, thus enabling the design of genetic modules that are robust to disturbances due to unknown loading effects.**

I. INTRODUCTION

The rational engineering of complex systems relies on predictive and quantitative models. While the goal of synthetic biology is to create large-scale genetic circuits, currently the development of even simple circuit components requires expensive and time consuming iterative processes [1], [2]. One major factor contributing to this is that gene circuits exhibit context-dependent behavior [3], [4], [5], [6], [7], for instance, due to the scarcity of shared transcriptional and translational resources [8], [9], [10], [11], [12], [13].

To enable the modular creation of large-scale systems in synthetic biology, resource competition effects must be accounted for when designing and analyzing genetic components. Since one of the most widely used genetic modules is the toggle switch [14], [15], with applications ranging from clocks [16] to frequency multipliers [17], understanding how the scarcity of shared cellular resources affects the stability profile of the toggle switch is crucial. Leveraging a mechanistic model that accurately predicts loading effects due to shared cellular resources both *in vitro* and *in vivo* [9], [11], we recently revealed that resource competition acts against bistability. In particular, a bistable toggle switch will become monostable when resource sequestration exceeds a critical threshold [18], and this phenomenon is further amplified by the presence of parameter asymmetries [19].

Cellular behavior exhibits significant variance at the population-level due to multiple factors (noise, difference in growth rate and plasmid copy number, fluctuations in the environment, etc.), often leading to the emergence of distinct subpopulations [20]. In case of the toggle switch, this could mean that genetically identical cells can exhibit a mixture of monostable and bistable behavior at the population-level. Understanding how variance of key parameters and the correlation among them affect the monostable fraction $\epsilon \in [0,1]$ of the population is the focus of this paper, thus revealing the interplay between cell-to-cell heterogeneity, scarce resources, and the stability profile at the population-level.

To this end, we model parameters of the toggle switch as random variables drawn from a probability distribution. We first show that ϵ can be determined by studying the ratio distribution of random variables. The analytic expression of ϵ unfortunately conceals the role that each parameter plays, thus failing to provide design guidelines to control ϵ. To overcome this issue, we next derive an approximation of ϵ which introduces negligible error considering the typical range of biophysical parameters. Leveraging this description, we reveal the role that each parameter plays in shaping the population-level stability profile through ϵ.

These results provide design guidelines for creating bistable toggle switches that are robust to loading effects due to resource competition. These guidelines not only apply at the cellular level [18], [19], but the results derived here also reveal how to design toggle switches that show minimal variation at the population level. For instance, we show that the higher the correlation between production rate and the parameter measuring resource sequestration, the lower ϵ (higher fraction of bistable toggles at the population-level). The results presented here not only apply to isolated toggle switches, but we also consider loading from their cellular context. Importantly, we illustrate that the effect of the resulting disturbance on ϵ can be bounded even in the presence of unknown parameters. In addition to revealing how to design toggle switches that are robust to such perturbations, our results also provide guidelines ensuring homogeneous population-level behavior despite cell-to-cell heterogeneity.

This paper is structured as follows. First, we briefly introduce the mathematical model of the toggle switch explicitly accounting for the scarcity of shared cellular resources. In this description, the key parameters are random variables, leading to cell-to-cell heterogeneity at the population level. Following this, we first derive the analytic expression characterizing the bistable/monostable composition at the population level, then we introduce an approximation that reveals how each parameter affects ϵ, the fraction of the population that is monostable. Finally, we illustrate how our results can be applied when designing toggle switches that exhibit robust behavior at the population-level.

A. Gyorgy is with the Department of Electrical and Computer Engineering, New York University Abu Dhabi, UAE andras.gyorgy@nyu.edu

II. MODEL AND PROBLEM FORMULATION

Without accounting for the scarcity of shared transcriptional/translational resources, the behavior of the toggle switch [14] is governed by

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{a_y}{1+z^2} - y, \qquad \frac{\mathrm{d}z}{\mathrm{d}t} = \frac{a_z}{1+y^2} - z, \tag{1}$$

where y and z denote the concentrations of the two proteins repressing each other's expression, whereas a_y and a_z are the non-dimensional production rate constants. These rate constants encompass a collection of complex processes required for protein production (e.g., transcriptional/translational initiation, elongation, etc.). As detailed in [9], [11], a_w for $w \in \{y, z\}$ is given by

$$a_w = \frac{\lambda_w^{\mathrm{TX}} \lambda_w^{\mathrm{TL}} D_w}{\delta_w \kappa_w k_w}, \qquad w \in \{y, z\}, \tag{2}$$

where λ_w^{TX} and λ_w^{TL} represent transcriptional and translational rate constants, respectively, D_w denotes the DNA concentration encoding w, δ_w stands for mRNA decay rate, whereas κ_w and k_w are the dissociation constants of RNA polymerase (RNAP) and ribosomes to the promoters and ribosome binding sites, respectively. Typical values of these parameters can be found in the Appendix.

While the model in (1) captures the dynamics of the toggle switch when shared transcriptional/translational resources are abundant, it does not account for coupling effects that arise when these resources become scarce [4], [9], [10], [11]. In particular, when accounting for the limited availability of shared transcriptional/translational resources, the dynamics can be written as

$$\begin{aligned} \dot{y} &= \frac{a_y \frac{1}{1+z^2}}{1 + b_y \frac{1}{1+z^2} + b_z \frac{1}{1+y^2} + l} - y, \\ \dot{z} &= \frac{a_z \frac{1}{1+y^2}}{1 + b_y \frac{1}{1+z^2} + b_z \frac{1}{1+y^2} + l} - z, \end{aligned} \tag{3}$$

where b_y and b_z can be interpreted as a measure of resource sequestration associated with the production of y and z, respectively [11], [18], [19]. As detailed in [9], [11], b_w for $w \in \{y, z\}$ is given by

$$b_w = \frac{D_w}{\kappa_w}\left(1 + \frac{\lambda_w^{\mathrm{TX}}}{\delta_w k_w}\right), \qquad w \in \{y, z\}, \tag{4}$$

and similarly, l represents loading due to the context of the toggle switch: resource sequestration by the expression of proteins other than y and z (see [18] for details).

According to [18], when $a_y = a_z =: a$ and $b_y = b_z =: b$ the system in (3) is bistable if $a > 2(1 + b + l)$ and monostable if $a < 2(1 + b + l)$, and similar results can be obtained even when $a_y \neq a_z$ and $b_y \neq b_z$ (see [18], [19] for details). That is, there is a critical threshold for the resource sequestration represented by $b + l$, above which a bistable toggle switch becomes monostable. While this result characterizes the stability profile of the toggle switch in an individual cell, it fails to account for variability at the population-level due to differences in growth rate, plasmid copy number, local fluctuations in the environment, etc.

To model this variability, we assume that a, b, and l are realizations of random variables drawn from a distribution. In particular, we consider the dynamics

$$\begin{aligned} \dot{Y} &= \frac{A \frac{1}{1+Z^2}}{1 + B \frac{1}{1+Y^2} + B \frac{1}{1+Z^2} + L} - Y, \\ \dot{Z} &= \frac{A \frac{1}{1+Y^2}}{1 + B \frac{1}{1+Y^2} + B \frac{1}{1+Z^2} + L} - Z, \end{aligned} \tag{5}$$

where the random vector $X = (A, B, L)$ follows the multivariate normal distribution $\mathcal{N}(\mu_X, \Sigma_X)$ with mean vector $\mu_X = (\mu_A, \mu_B, \mu_L)^\top$ and covariance matrix

$$\Sigma_X = \begin{bmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B & \rho_1\sigma_A\sigma_L \\ \rho\sigma_A\sigma_B & \sigma_B^2 & \rho_2\sigma_B\sigma_L \\ \rho_1\sigma_A\sigma_L & \rho_2\sigma_B\sigma_L & \sigma_L^2 \end{bmatrix}. \tag{6}$$

Define ϵ as the probability of the system in (5) being monostable, so that $1 - \epsilon$ is the probability of bistability.

In this paper, we seek to characterize how parameters of μ_X and Σ_X shape $\epsilon \in [0, 1]$: the fraction of toggle switches that are monostable, or alternatively, the probability that the system in (5) is monostable when (A, B, L) are distributed according to $\mathcal{N}(\mu_X, \Sigma_X)$. We first focus on the case when loading from the context is negligible, then show how the presence of L in (5) affects ϵ, even when the probability distribution of L is not completely specified, illustrated in the application example.

III. RESULTS

To find the probability $\epsilon \in [0, 1]$ of (5) being monostable, introduce the random variables

$$R := \frac{A}{2(1+B)}, \qquad Q := \frac{A}{2(1+B+L)}. \tag{7}$$

From [18] it follows that (i) $\epsilon = \mathbf{P}(R < 1)$ when loading from the context is negligible, corresponding to the condition $a < 2(1 + b)$; and (ii) $\epsilon = \mathbf{P}(Q < 1)$ when loading from the context is accounted for, corresponding to the condition $a < 2(1 + b + l)$. Therefore, in what follows, we seek to compute the above probabilities.

A. Analytic Solution for the Population-Level Distribution

To find the probability $\mathbf{P}(R < r)$ for $r \in \mathbb{R}$ analytically, introduce the functions

$$\phi(t) := \frac{1}{\sqrt{2\pi}} e^{-t^2/2}, \qquad \Phi(t) := \int_{-\infty}^{t} \phi(u)\mathrm{d}u. \tag{8}$$

Proposition 1. *For R in (7) with $\mu_{A|B}(b) := \frac{\sigma_B\mu_A + \sigma_A\rho(b-\mu_B)}{\sigma_B}$ and $\sigma_{A|B}(b) := \sqrt{1-\rho^2}\sigma_A$ we have that*

$$\begin{aligned} \mathbf{P}(R < r) = &\frac{1}{\sigma_B}\int_{-1}^{\infty} \Phi\left(\frac{2r(1+b) - \mu_{A|B}(b)}{\sigma_{A|B}(b)}\right)\phi\left(\frac{b-\mu_B}{\sigma_B}\right) db \\ &- \frac{1}{\sigma_B}\int_{-\infty}^{-1} \Phi\left(\frac{2r(1+b) - \mu_{A|B}(b)}{\sigma_{A|B}(b)}\right)\phi\left(\frac{b-\mu_B}{\sigma_B}\right) db \\ &+ \Phi\left(-\frac{1+\mu_B}{\sigma_B}\right). \end{aligned} \tag{9}$$

Proof. Marginalization and Bayes' theorem yields

$$\begin{aligned}\mathbf{P}(R<r) &= \int_{-\infty}^{\infty} \mathbf{P}\left(\tfrac{A}{2(1+B)} \le r, B=b\right) \mathrm{d}b \\ &= \int_{-\infty}^{-1} \mathbf{P}\left(\tfrac{A}{2(1+b)} \le r | B=b\right) \mathbf{P}(B=b) \mathrm{d}b \\ &+ \int_{-1}^{\infty} \mathbf{P}\left(\tfrac{A}{2(1+b)} \le r | B=b\right) \mathbf{P}(B=b) \mathrm{d}b.\end{aligned}$$

Furthermore, since for $b<-1$

$$\mathbf{P}\left(\tfrac{A}{2(1+b)} < r | B=b\right) = 1 - \mathbf{P}\left(A < 2r(1+b) | B=b\right),$$

whereas for $b>-1$ we obtain

$$\mathbf{P}\left(\tfrac{A}{2(1+b)} < r | B=b\right) = \mathbf{P}\left(A < 2r(1+b) | B=b\right),$$

we have with $\Phi'(\frac{x-\mu}{\sigma}) = \phi(\frac{x-\mu}{\sigma})/\sigma$ that

$$\begin{aligned}\sigma_B \mathbf{P}(R<r) = &\int_{-\infty}^{-1} \phi\left(\tfrac{b-\mu_B}{\sigma_B}\right) \mathrm{d}b \\ &- \int_{-\infty}^{-1} \Phi\left(\tfrac{2r(1+b)-\mu_{A|B}(b)}{\sigma_{A|B}(b)}\right) \phi\left(\tfrac{b-\mu_B}{\sigma_B}\right) \mathrm{d}b \\ &+ \int_{-1}^{\infty} \Phi\left(\tfrac{2r(1+b)-\mu_{A|B}(b)}{\sigma_{A|B}(b)}\right) \phi\left(\tfrac{b-\mu_B}{\sigma_B}\right) \mathrm{d}b,\end{aligned}$$

where we used the fact that

$$\mathbf{P}\left(A < 2r(1+b) | B=b\right) = \Phi\left(\tfrac{2r(1+b)-\mu_{A|B}(b)}{\sigma_{A|B}(b)}\right),$$

since $A|B=b \sim \mathcal{N}(\mu_{A|B}(b), \sigma_{A|B}(b))$. Finally, we conclude the proof by noting that

$$\int_{-\infty}^{-1} \phi\left(\tfrac{b-\mu_B}{\sigma_B}\right) \mathrm{d}b = \sigma_B \Phi\left(-\tfrac{1+\mu_B}{\sigma_B}\right).$$

□

While the above result can be used to compute the probability $\epsilon = \mathbf{P}(R<1)$, it provides little insight into how μ_A, μ_B, σ_A, σ_B, and ρ affect ϵ. As a result, we would need to resort to numerical simulations to study these parameter effects. Additionally, the above result focuses on the case when resource sequestration of the context is negligible. Including these additional loading effects would further obfuscate parameter dependencies, presenting a bottleneck in the forward engineering of genetic circuits.

B. Approximation to Reveal Parameter Dependencies

One way to overcome the above issue is to consider the results in [21], [22]: for certain parameter regions, R can be approximated with a normal random variable. This way, we could study how parameters of μ_X and Σ_X affect the mean and variance of this approximate distribution to reveal their role in shaping ϵ. Unfortunately, considering the typical parameters in the Appendix, the conditions outlined in [21], [22] do not necessarily apply in our case. Therefore, we next consider an approximation to study both R and Q defined in (7) by presenting the result of Hinkley [23] based on the work of Geary [24] in a more general format.

Proposition 2. *Let $X \in \mathbb{R}^N$ be a jointly normal random variable with mean vector $\mu \in \mathbb{R}^N$ and covariance matrix $\Sigma \in \mathbb{R}^{N\times N}$. Introduce $\alpha, \gamma \in \mathbb{R}^N$ and $\beta, \delta \in \mathbb{R}$, together with the random variable $W \coloneqq (\alpha^\top X + \beta)/(\gamma^\top X + \delta)$ and*

$$F(w) \coloneqq \Phi\left(\frac{(w\delta-\beta) - (\alpha - w\gamma)^\top \mu}{\sqrt{(\alpha-w\gamma)^\top \Sigma (\alpha - w\gamma)}}\right). \tag{10}$$

With this, we have that

$$|\mathbf{P}(W<w) - F(w)| \le \Phi\left(-\frac{\gamma^\top \mu + \delta}{\sqrt{\gamma^\top \Sigma \gamma}}\right). \tag{11}$$

Proof. From (10) with $Y \coloneqq \gamma^\top X + \delta$ we obtain that

$$F(w) = \mathbf{P}([\alpha - w\gamma]^\top X < w\delta - \beta) = \mathbf{P}(\alpha^\top X + \beta < wY)$$

and with $d \coloneqq \mathbf{P}(\alpha^\top X + \beta < wY,\ Y \le 0)$ we can write

$$\begin{aligned}F(w) =& \mathbf{P}(\alpha^\top X + \beta < wY) \\ =& d + \mathbf{P}(\alpha^\top X + \beta < wY,\ Y>0) \\ =& d + \mathbf{P}\left(\frac{\alpha^\top X + \beta}{\gamma^\top X + \delta} < w,\ Y>0\right) \\ =& d + \mathbf{P}(W<w,\ Y>0).\end{aligned} \tag{12}$$

Since from the law of total probability we also have that

$$\mathbf{P}(W<w) = \mathbf{P}(W<w,\ Y\le 0) + \mathbf{P}(W<w,\ Y>0),$$

from (12) it follows that

$$\mathbf{P}(W<w) - F(w) = \mathbf{P}(W<w, Y\le 0) - d, \tag{13}$$

yielding

$$|\mathbf{P}(W<w) - F(w)| \le \mathbf{P}(Y\le 0) = \Phi\left(-\frac{\gamma^\top \mu + \delta}{\sqrt{\gamma^\top \Sigma \gamma}}\right)$$

since $0 \le d, \mathbf{P}(W<w, Y\le 0) \le \mathbf{P}(Y\le 0)$. □

The above result can be interpreted as follows. With $Y \coloneqq \gamma^\top X + \delta$ we have that $\mu_Y \coloneqq \mathbb{E}[Y] = \gamma^\top \mu + \delta$ and $\sigma_Y \coloneqq \mathrm{var}(Y) = \sqrt{\gamma^\top \Sigma \gamma}$. Therefore, the difference between $\mathbf{P}(W<w)$ and $F(w)$ is smaller than the probability of $Y<0$. In other words, the error when using $F(w)$ to approximate $\mathbf{P}(W<w)$ vanishes if Y is unlikely to take on negative values.

C. Parameter Effects in the Absence of the Context

Building on Proposition 2, we now approximate $\epsilon = \mathbf{P}(R<1)$ with $F(\cdot)$ introduced in (10) to reveal the role that each parameter in μ_X and Σ_X plays in the absence of the context. Furthermore, we show that the approximation error is negligible for typical parameter values.

Proposition 3. *For the random variable R in* (7) *we have that*

$$\mathbf{P}(R<1) = \Phi\left(\tfrac{2-\mu}{\sigma}\right) + \Delta, \qquad |\Delta| \le \Phi\left(-\tfrac{1+\mu_B}{\sigma_B}\right), \tag{14}$$

where $\mu \coloneqq \mu_A - 2\mu_B$ and $\sigma \coloneqq \sqrt{\sigma_A^2 + 4\sigma_B^2 - 4\rho\sigma_A\sigma_B}$.

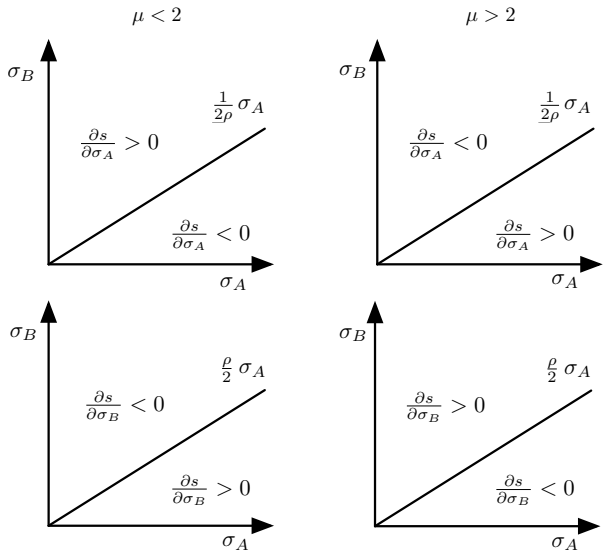

Fig. 1. The monostable fraction $\epsilon \approx \Phi(s)$ of the population increases with $s = (2-\mu)/\sigma$. Decreasing μ_A and increasing μ_B always increases s, thus pushes the population towards monostability. Additionally, the effect of the correlation coefficient ρ depends on whether $\mu < 2$ or $\mu > 2$. Finally, the effect of σ_A and σ_B on s, thus on ϵ depends on their relative value, the correlation coefficient ρ, and whether $\mu < 2$ or $\mu > 2$.

Proof. Let $\alpha := (1 \quad 0)^\top$, $\beta := (0 \quad 2)^\top$, $\gamma := 0$, and $\delta := 2$, so that from Proposition 2 we obtain that

$$|\mathbf{P}(R \leq r) - F(r)| \leq \Phi\left(-\tfrac{1+\mu_B}{\sigma_B}\right) \tag{15}$$

with

$$F(r) = \Phi\left(\frac{2r - \mu_R(r)}{\sigma_R(r)}\right),$$

where $\sigma_R(r) := \sqrt{\sigma_A^2 + 4r^2\sigma_B^2 - 4r\rho\sigma_A\sigma_B}$ together with $\mu_R(r) := \mu_A - 2r\mu_B$. Evaluating (15) at $r = 1$ so that $\mu = \mu_R(1)$ and $\sigma = \sigma_R(1)$ concludes the proof. □

Unlike Proposition 1, the above result reveals the role that each parameter plays in establishing the fraction ϵ of the population that is expected to be monostable via the scalar $s := (2-\mu)/\sigma$ as $\epsilon \approx \Phi(s)$. In particular, we have that

$$\begin{aligned}
\frac{\partial s}{\partial \mu_A} &= \frac{-1}{\sqrt{\sigma_A^2 + 4\sigma_B^2 - 4\rho\sigma_A\sigma_B}},\\
\frac{\partial s}{\partial \mu_B} &= \frac{2}{\sqrt{\sigma_A^2 + 4\sigma_B^2 - 4\rho\sigma_A\sigma_B}},\\
\frac{\partial s}{\partial \sigma_A} &= \frac{(2-\mu)(2\rho\sigma_B - \sigma_A)}{(\sigma_A^2 + 4\sigma_B^2 - 4\rho\sigma_A\sigma_B)^{3/2}},\\
\frac{\partial s}{\partial \sigma_B} &= \frac{2(2-\mu)(\rho\sigma_A - 2\sigma_B)}{(\sigma_A^2 + 4\sigma_B^2 - 4\rho\sigma_A\sigma_B)^{3/2}},\\
\frac{\partial s}{\partial \rho} &= \frac{2(2-\mu)\sigma_A\sigma_B}{(\sigma_A^2 + 4\sigma_B^2 - 4\rho\sigma_A\sigma_B)^{3/2}},
\end{aligned}$$

and since $\Phi(\cdot)$ is monotonically increasing, ϵ is decreasing with μ_A and it is increasing with μ_B, as expected from [18]. Additionally, the above result also reveals the effect that σ_A, σ_B, and ρ play, depending on whether $\mu < 2$ or $\mu > 2$. These relationships are illustrated in Fig. 1, together with simulation results confirming the expected behavior in Fig. 2. The approximation $\epsilon \approx \Phi(s)$ introduces the error $\Delta \leq \Phi\left(-\frac{1+\mu_B}{\sigma_B}\right) = \mathbf{P}(B < -1)$, which is negligible as B is a lumped parameter expected to take on non-negative values (e.g., the error is less than 0.1% if $\mu_B \geq 3\sigma_B$).

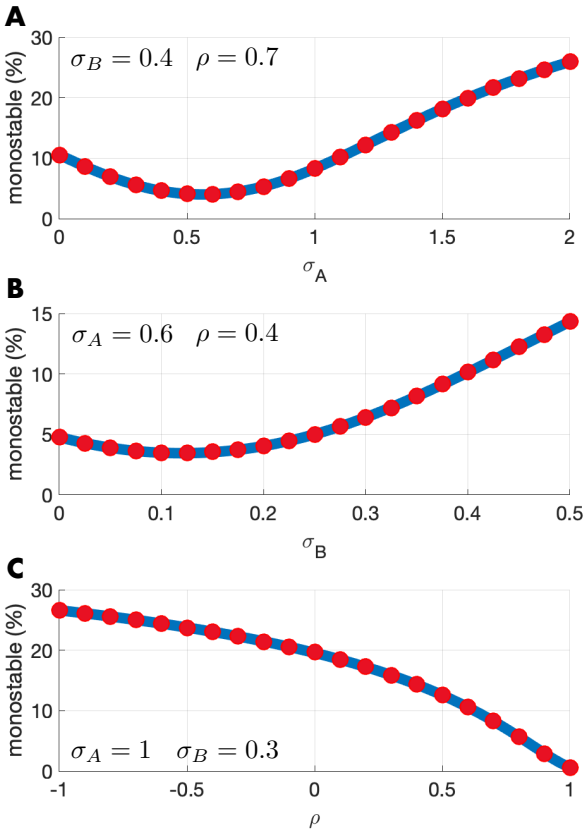

Fig. 2. Since B is expected to be non-negative, $\epsilon \approx \Phi(s)$ in blue well approximates the probability $\mathbf{P}(R < 1)$ in red as $\Delta = \mathbf{P}(B < -1) \approx 0$. In all plots $\mu = \mu_A - 2\mu_B = 3 > 2$. (A) The monostable fraction ϵ decreases when $\sigma_A < 2\rho\sigma_B = 0.56$, then it increases since $\mu > 2$. (B) The monostable fraction ϵ decreases when $\sigma_B < \rho\sigma_A/2 = 0.12$, then it increases since $\mu > 2$. (C) The monostable fraction ϵ decreases with ρ since $\mu > 2$. In all simulations $\mu_A = 5$ and $\mu_B = 1$.

D. Effects of Loading due to the Context

After revealing how each parameter of μ_X and Σ_X affects ϵ when the effects of L are neglected, we next focus on how the context of the toggle switch shapes the composition of the population via ϵ. In particular, we are interested how ϵ depends on the parameters μ_L, σ_L, ρ_1, and ρ_2.

Proposition 4. *For Q defined in (7) we have that*

$$\mathbf{P}(Q < 1) = \Phi\left(\tfrac{2-\mu}{\sigma}\right) + \Delta, \qquad |\Delta| \leq \Phi\left(-\tfrac{1+\mu_0}{\sigma_0}\right),$$

where $\mu := \mu_A - 2\mu_0$ and $\sigma := \sqrt{\sigma_A^2 + 4\sigma_0^2 - 4\rho_0\sigma_A\sigma_0}$ with $\mu_0 := \mu_B + \mu_L$, $\sigma_0 := \sqrt{\sigma_B^2 + 2\rho_2\sigma_B\sigma_L + \sigma_L^2}$, together with $\rho_0 := (\rho\sigma_B + \rho_1\sigma_L)/\sigma_0$.

Proof. Let $\alpha := (1 \quad 0 \quad 0)^\top$, $\beta := (0 \quad 2 \quad 2)^\top$, $\gamma := 0$, and $\delta := 2$, thus with $\sigma_Q(q) := \sqrt{\sigma_A^2 + 4q^2\sigma_0^2 - 4q\rho_0\sigma_A\sigma_0}$ and

$\mu_Q(q) \coloneqq \mu_A - 2q\mu_0$ we obtain that

$$\frac{(q\delta-\beta)-(\alpha-q\gamma)^\top\mu}{\sqrt{(\alpha-r\gamma)^\top\Sigma(\alpha-q\gamma)}} = \frac{2q-\mu_Q(q)}{\sqrt{\sigma_Q(q)}},$$
$$\frac{\gamma^\top\mu+\delta}{\sqrt{\gamma^\top\Sigma\gamma}} = \frac{1+\mu_0}{\sigma_0}.$$

Therefore, from Proposition 2 it follows that

$$|\mathbf{P}(Q \le q) - F_Q(q)| \le \Phi\left(-\tfrac{1+\mu_0}{\sigma_0}\right) \tag{16}$$

with

$$F(q) \coloneqq \Phi\left(\frac{2q-\mu(q)}{\sigma(q)}\right). \tag{17}$$

Evaluating (16) at $q = 1$ so that $\mu = \mu_Q(1)$ and $\sigma = \sigma_Q(q)$ concludes the proof. □

Therefore, loading from the context of the toggle switch can be treated by modifying the distribution of B (loading internal to the toggle) to that of $B+L$ (loading both internal and external to the toggle) by simply changing (μ_B, σ_B, ρ) to $(\mu_0, \sigma_0, \rho_0)$. As a result, parameters of μ_X and Σ_X associated with L affect ϵ via $(\mu_0, \sigma_0, \rho_0)$, which has been detailed in the previous section and in Fig. 1.

IV. APPLICATION EXAMPLE

The results presented in this paper not only allow us to reveal the role that each parameter in μ_X and Σ_X plays in shaping the population-level stability profile of the toggle switch, but also to develop design guidelines. To illustrate this, consider the following example.

For a given bistable toggle switch, we seek to quantify the maximal loading μ_L so that the resulting degradation in performance (increase of ϵ) does not exceed a critical threshold. Therefore, the parameters μ_A, μ_B, σ_A, σ_B, and ρ are known and fixed, whereas σ_L, ρ_1, and ρ_2 are unknown in (6). This way we can ensure robustness of the toggle switch even to unknown disturbances.

To this end, we first construct the correlation matrix

$$\begin{bmatrix} 1 & \rho & \rho_1 \\ \rho & 1 & \rho_2 \\ \rho_1 & \rho_2 & 1 \end{bmatrix}$$

where ρ_1 and ρ_2 are unknown, whereas ρ is known. Since the above correlation matrix must be positive semi-definite, its determinant must be non-negative. Thus, we must have $1+2\rho\rho_1\rho_2-\rho^2-\rho_1^2-\rho_2^2 \ge 0$, so that ρ_1 and ρ_2 must satisfy the constraint

$$\rho_1\rho_2 - \sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2} \le \rho \le \rho_1\rho_2 + \sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}.$$

With this, we can use $\epsilon \approx \Phi\left(\frac{2-\mu}{\sigma}\right)$ from Proposition 4 to determine the possible range of ϵ for any μ_L and σ_L, as illustrated in Fig. 2. For example, if we want to have $\epsilon \le 0.25$, then we must have $\mu_L \le 0.1$ if $\rho = -0.9$, whereas we can tolerate four times as much loading from the context if $\rho = 0.9$, regardless of its exact nature. Interestingly, ϵ can be smaller when $\sigma_L > 0$ than when $\sigma_L = 0$ (blue in Fig. 3), that is, noise from the context can be beneficial by increasing the fraction of the population that is bistable.

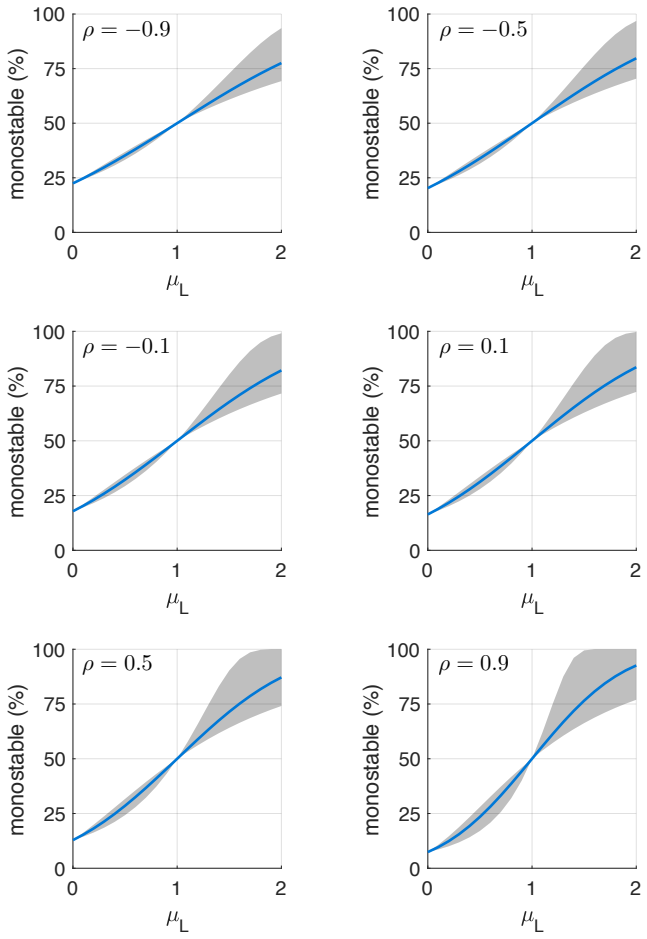

Fig. 3. The monostable fraction of the population of toggle switches increases with loading from the context. Blue curves correspond to the case when $\sigma_L = 0$, whereas the grey band shows the range that can be achieved by varying (ρ_1, ρ_2). Simulation parameters: $\mu_A = 6$, $\mu_B = 1$, $\sigma_A = \mu_A/3$, $\sigma_B = \mu_B/3$, $\sigma_L \le \mu_L/3$. Since $\Phi(0) = 0.5$, all curves must cross $\mu_L = 1$, $\epsilon = 0.5$.

V. DISCUSSION

The rational forward engineering of large-scale systems requires quantitative models. Therefore, for synthetic biology to realize its full potential, predictive computational tools need to be developed that reveal the role that each parameter plays [25], [26], [27], [28]. Since cellular behavior often demonstrates significant variance due to several factors (differences in growth rate, plasmid copy number, environmental fluctuations, etc.), it is fundamental to understand not only the nominal behavior of synthetic circuits, but also their population-level distribution.

Therefore, in this paper, we quantified how the population-level stability profile of toggle switches is affected by key parameters in the presence of shared cellular resources, considering a mechanistic model that shows strong agreement with experimental data both *in vitro* [11] and *in vivo* [9]. Competition for these resources is a major source of context-dependence, hindering the modular design of complex gene circuits.

In particular, we first derived an analytic formula quantifying the fraction ϵ of the population that is monostable. As this description conceals the role that each parameter plays in shaping ϵ, we next derived an approximation with negligible error considering the typical range of biophysical parameters. This description reveals how the population-level stability profile is affected by the parameters in both μ_X and Σ_X. For instance, in the absence of parameter variations (i.e., $\sigma_A = \sigma_B = 0$), the system in (5) is bistable if $\mu_A > 2(1 + \mu_B)$ according to [18], thus the population is completely bistable ($\epsilon = 0$). Conversely, in the presence of parameter variations we obtain that if $s \approx 0$ in Proposition 3 then $\epsilon \approx \Phi(0) = 0.5$, that is, the population is approximately 50% bistable and 50% monostable. This happens if $0 < \mu_A - 2(1 + \mu_B) \ll \sigma$ where $\sigma = \sqrt{\sigma_A^2 + 4\sigma_B^2 - 4\rho\sigma_A\sigma_B}$. Therefore, to avoid such undesirable outcome it is advisable to design toggle switches such that not only $\mu_A - 2(1+\mu_B) > 0$ ensuring bistability, but also $\mu_A - 2(1+\mu_B) \gg \sigma$ ensuring homogeneous population-level behavior ($\epsilon \approx 0$).

Complementing the results presented here, we next seek to extend them to non-Gaussian parameter distributions. Additionally, future research directions include estimating the distribution of random switching times between stable steady states both at the cellular and at the population level, especially the effects that scarce resources and the context of the toggle switch play.

APPENDIX

Based on data in Tab. I, we estimate $\mu_A \approx 1 \dots 10$ and $\mu_B \approx 0.1 \dots 1$ to be typical from (2) and (4), respectively (see [18] for details on the typical parameter values). Naturally, this range can be be decreased/increased by tuning promoters, ribosome binding sites, degradation tags, etc.

TABLE I

REPRESENTATIVE VALUE OF MODEL PARAMETERS

symbol	meaning	typical value	unit
D	DNA concentration	100-1000	nM
κ	RNAP dissociation constant	1	μM
λ^{TX}	transcriptional rate constant	100	1/h
λ^{TL}	translational rate constant	1000	1/h
k	ribosome dissociation constant	10	μM
K	repressor dissociation constant	0.1	nM
δ	mRNA decay rate	10	1/h

REFERENCES

[1] M. Xie and M. Fussenegger, "Designing cell function: assembly of synthetic gene circuits for cell biology applications," pp. 507–525, 2018.

[2] P. E. Purnick and R. Weiss, "The second wave of synthetic biology: From modules to systems," pp. 410–422, 2009.

[3] E. Yeung, A. J. Dy, K. B. Martin, A. H. Ng, D. Del Vecchio, J. L. Beck, J. J. Collins, and R. M. Murray, "Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks," *Cell Systems*, vol. 5, no. 1, pp. 11–24.e12, jul 2017.

[4] F. Ceroni, R. Algar, G.-B. Stan, and T. Ellis, "Quantifying cellular capacity identifies gene expression designs with reduced burden." *Nature Methods*, vol. 12, no. 5, pp. 415–418, 2015.

[5] V. H. Nagaraj, J. M. Greene, A. M. Sengupta, and E. D. Sontag, "Translation inhibition and resource balance in the TX-TL cell-free gene expression system," *Synt Biol*, vol. 2, no. 1, pp. 1–7, 2017.

[6] 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, dec 2018.

[7] C. Briat, A. Gupta, and M. Khammash, "Antithetic Integral Feedback Ensures Robust Perfect Adaptation in Noisy Bimolecular Networks," *Cell Systems*, vol. 2, no. 1, pp. 15–26, 2016.

[8] P. M. Caveney, S. E. Norred, C. W. Chin, J. B. Boreyko, B. S. Razooky, S. T. Retterer, C. P. Collier, and M. L. Simpson, "Resource Sharing Controls Gene Expression Bursting," *ACS Synthetic Biology*, vol. 6, no. 2, pp. 334–343, feb 2017.

[9] A. Gyorgy, J. I. Jiménez, J. Yazbek, H.-H. Huang, H. Chung, R. Weiss, and D. Del Vecchio, "Isocost Lines Describe the Cellular Economy of Genetic Circuits," *Biophys J*, vol. 109, no. 3, pp. 639–646, 2015.

[10] D. Siegal-Gaskins, Z. A. Tuza, J. Kim, V. Noireaux, and R. M. Murray, "Gene Circuit Performance Characterization and Resource Usage in a Cell-Free "Breadboard"," *ACS Synth Biol*, vol. 3, pp. 416–425, 2014.

[11] A. Gyorgy and R. M. Murray, "Quantifying resource competition and its effects in the TX-TL system," in *55th IEEE Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 3363–3368.

[12] O. Borkowski, C. Bricio, M. Murgiano, B. Rothschild-Mancinelli, G. B. Stan, and T. Ellis, "Cell-free prediction of protein expression costs for growing cells," *Nature Communications*, vol. 9, no. 1, 2018.

[13] A. Boo, G.-B. Stan, O. Borkowski, T. E. Gorochowski, S. Furini, T. Ellis, Y. N. Ladak, C. Gilbert, F. Ceroni, and A. R. Awan, "Burden-driven feedback control of gene expression," *Nature Methods*, 2018.

[14] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of a genetic toggle switch in Escherichia coli," *Nature*, vol. 403, no. 6767, pp. 339–342, 2000.

[15] P. Hersen, G. Batt, J.-B. Lugagne, A. Köhler, M. Kirch, and S. Sosa Carrillo, "Balancing a genetic toggle switch by real-time feedback control and periodic forcing," *Nature Communications*, vol. 8, no. 1, p. 1671, dec 2017.

[16] O. Purcell, M. di Bernardo, C. S. Grierson, and N. J. Savery, "A multi-functional synthetic gene network: A frequency multiplier, oscillator and switch," *PLOS ONE*, vol. 6, no. 2, pp. 1–12, 2011.

[17] C. Cuba Samaniego and E. Franco, "A robust molecular network motif for period-doubling devices," *ACS Synthetic Biology*, vol. 7, no. 1, pp. 75–85, 2018, pMID: 29227103.

[18] A. Gyorgy, "Sharing resources can lead to monostability in a network of bistable toggle switches," *IEEE Control Systems Letters*, vol. 3, no. 2, pp. 308–313, 2018.

[19] A. Gyorgy, "Bistability requires better balanced toggle switches in the presence of competition for shared cellular resources," in *American Control Conference (ACC)*. IFAC, 2019.

[20] A. E. Vasdekis, H. Alanazi, A. M. Silverman, C. J. Williams, A. J. Canul, J. B. Cliff, A. C. Dohnalkova, and G. Stephanopoulos, "Eliciting the impacts of cellular noise on metabolic trade-offs by quantitative mass imaging," *Nature Communications*, vol. 10, no. 1, p. 848, dec 2019.

[21] G. Marsaglia, "Ratios of Normal Variables and Ratios of Sums of Uniform Variables," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 193–204, may 1965.

[22] G. . Marsaglia, "Ratios of Normal Variables," *Journal of Statistical Software*, 2006.

[23] D.V. Hinkley, "On the Ratio of Two Correlated Normal Random Variables," *Biometrika*, vol. 56 (3), no. 3, pp. 635–639, 1969.

[24] R. C. Geary, "The Frequency Distribution of the Quotient of Two Normal Variates, Tech. Rep. 3, 1930.

[25] D. D. Lewis, F. D. Villarreal, F. Wu, and C. Tan, "Synthetic Biology Outside the Cell: Linking Computational Tools to Cell-Free Systems," *Frontiers in Bioengineering and Biotechnology*, vol. 2, p. 66, dec 2014.

[26] M. Tomazou, M. Barahona, K. M. Polizzi, and G. B. Stan, "Computational Re-design of Synthetic Genetic Oscillators for Independent Amplitude and Frequency Modulation," *Cell Systems*, vol. 6, no. 4, pp. 508–520.e5, 2018.

[27] N. Kylilis, Z. A. Tuza, G. B. Stan, and K. M. Polizzi, "Tools for engineering coordinated system behaviour in synthetic microbial consortia," *Nature Communications*, vol. 9, no. 1, 2018.

[28] F. Blanchini, C. Cuba Samaniego, E. Franco, and G. Giordano, "Homogeneous Time Constants Promote Oscillations in Negative Feedback Loops," *ACS Synthetic Biology*, vol. 7, no. 6, pp. 1481–1487, 2018.