Architectural Principles for Characterizing the Performance of Sequestration Feedback Networks

Noah Olsman*,†, Fangzhou Xiao‡, John C. Doyle*

September 26, 2018

Summary

As we begin to design increasingly complex synthetic biomolecular systems, it is essential to develop rational design methodologies that yield predictable circuit performance. Here we apply theoretical tools from the theory of control and dynamical systems to yield practical insights into the architecture and function of a particular class of biological feedback circuit. Specifically, we show that it is possible to analytically characterize both the operating regime and performance tradeoffs of a sequestration feedback circuit architecture. Further, we demonstrate how these principles can be applied to inform the design process of a particular synthetic feedback circuit.

1 Introduction

A present challenge in synthetic biology is to design circuits that not only perform a desired function, but do so robustly. The difficulty in doing this arises in large part from the enormous amount of variability between both intracellular and extracellular environments [1–5]. While it is becoming easier to quickly implement a given circuit architecture [6, 7], ensuring that its performance is robust to parameter variations is still a time-consuming and challenging task [8]. As synthetic circuits grow in size and complexity, it will be essential to develop a rational design methodology that allows biological engineers to easily identify the important design constraints for a given circuit and determine whether or not their desired behavior is feasible [9]. We can draw inspiration from the study of natural biological circuits, where cells are frequently confronted with a large amount of variability, yet exhibit robust behavior at the system level [10–16].

In the design of electrical and mechanical systems this problem is often solved with the implementation of feedback control [17, 18], where the dynamics of a process are adjusted based on measurements of the system’s state with the aim of achieving some performance goals. For example, in a commercial cruise control system an engineer may want the car to be able to rapidly track whatever reference speed the user desires by measuring the car’s current velocity, while accelerating at a safe rate and not being too sensitive to small disturbances (e.g. road conditions). Similarly, a biological engineer may want the output concentration of a molecular species in a circuit to track an input signal (e.g. an inducer) with dynamics that are robust to parametric variability in reaction rates and the inherent noisiness of chemical kinetics [19].

Our primary focus here is a circuit architecture proposed by Briat et al. [20] that uses a sequestration mechanism to implement feedback control in a biomolecular circuit. This circuit immediately had a broad impact on the study of biological feedback systems, as sequestration is both abundant in natural biological contexts [21–23] and appears to be feasible to implement in synthetic networks [24–28]. For example, sequestration feedback can be implemented using sense-antisense mRNA pairs [29], sigma-antisigma factor pairs [30], or scaffold-antiscaffold pairs [19].

---

*Department of Control and Dynamical Systems, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA.
†Correspondence: Noah Olsman, nolsman@caltech.edu
‡Division of Biology and Biological Engineering, California Institute of Technology
The purpose of this circuit is to control a process, composed of the molecular species $X_1$ and $X_2$ with two control species $Z_1$ and $Z_2$ (figure 1A). The goal is to set an external reference $\mu$ and have the concentration of the output species $X_2$ robustly track it (figure 1B). The key property of this circuit is that it is able to implement robust perfect adaptation. This means that, at steady state, the concentration of $X_2$ will be proportional to the reference $\mu$ (specifically, $X_2 = \mu/\theta_2$ on average). Importantly, the steady-state value of $X_2$ will be independent of every other parameter of the network, implying its steady-state behavior is robust. We will focus first on studying a deterministic ordinary differential equation model of the circuit:

$$\dot{x}_1 = \theta_1 z_1 - \gamma_p x_1, \quad \text{(1a)}$$
$$\dot{x}_2 = kx_1 - \gamma_p x_2, \quad \text{(1b)}$$
$$\dot{z}_1 = \mu - \eta z_1 z_2 - \gamma_c z_1, \quad \text{(1c)}$$
$$\dot{z}_2 = \theta_2 x_2 - \eta z_1 z_2 - \gamma_c z_2. \quad \text{(1d)}$$

Here $\theta_1$, $\theta_2$, and $k$ are production rates, $\gamma_p$ and $\gamma_c$ are degradation rates for the process species and controller species, respectively, $\eta$ is the rate at which $z_1$ and $z_2$ sequester each other, and $\mu$ is the reference input that sets the synthesis rate of $z_1$. We will refer to $x_1$ and $x_2$ as process species, and will focus in particular on $x_2$ as the controlled output of the circuit. We can think of $z_2$ as making measurements of $x_2$, which are then propagated to $z_1$ which can indirectly affect the production rate of $x_2$ through $x_1$. The use of lower-case letters for the species in equation (1) indicates that we are referring to a deterministic quantity (later in section 2.6 we will use upper-case variables to denote random variables).

![Figure 1: Characterizing Performance in the Sequestration Feedback Circuit. A) The circuit diagram for a class of sequestration feedback circuits, adapted from the model presented in [20]. Here we take $X_1$ and $X_2$ as the process species we are trying to control, and $Z_1$ and $Z_2$ as the controller species. One notable addition is that we explicitly model degradation of the control species $Z_1$ and $Z_2$ at the rate $\gamma_c$. $\theta_1$ and $\theta_2$ represent the interconnection between the process species and the control species, $k$ represents the $X_1$-dependent synthesis rate of $X_2$, and $\gamma_p$ is the degradation rate of the process species. Finally, $\mu$ acts as an external reference that determines production of $Z_1$ through $\theta_1$ which we’d like to ultimately control the steady-state concentration of $Z_2$, and $\eta$ is the rate at which $Z_1$ and $Z_2$ irreversibly bind to each other. B) A representative plot of the type of behavior we expect from the circuit, where the concentration of $X_2$ tracks a changing reference $\mu$. We see that different cells have the same overall behavior, but with slight variations due to noise. This plot highlights the performance characteristics of this particular implementation of the circuit. For example we see that, when tracking the reference, $X_2$ has some overshoot of the target (red), a period of time it takes to respond to changes (blue), random fluctuations due to noise (yellow), and steady-state error (purple). Ideally, we would like to a rational methodology to tune the circuit-level parameters of panel A to predictably control the system-level characteristics of panel B.](image-url)
If we take the viewpoint of an engineer designing this circuit for a particular application, for example controlling the expression of a downstream gene target, then we likely will have performance requirements that we would like the circuit to follow, such as responding to a change in reference within a given time frame and not transiently overshooting the reference by more than some set amount. The problem lies in understanding how these types of high-level system performance metric relate to the individual parameters of the circuit. Alternatively, if we consider natural circuits then we can take an evolutionary perspective, and study how different fitness constraints might shape the structure and function of existing biological feedback systems. Our work here uses methods from the field of control theory to shed light on this mapping between parameters and performance.

The structure and content of this article is somewhat unconventional, in that we present both a non-technical treatment of some of the core results in the companion piece [31], and also standalone research results that are not discussed in detail elsewhere. Where the content of [31] focuses on the application of classical tools in control theory to study the mathematical properties of the sequestration feedback circuit, the goal of this article is to outline practical guidelines that are accessible to an audience interested in utilizing these theoretical results to inform future experimental work in biological engineering.

In [20], the authors assume that the controller degradation rate $\gamma_c = 0$ in equations (1c) and (1d), which yield the robust precise adaptation mentioned earlier. Our analysis of circuit performance and tradeoffs in sections 2.1 to 2.3 makes the same assumption, and we address the case of $\gamma_c > 0$ in section 2.4. In section 2.5 we apply this analysis to study a particular synthetic growth control circuit that makes use of sequestration feedback. Finally, in section 2.6 we will investigate the effects of noise on the circuit.

2 Results

2.1 So Long As Sequestration Is Fast, It Can Be Ignored

The most obvious parameter to investigate in the sequestration feedback system is the sequestration rate $\eta$. To simplify our analysis, we will ignore controller species degradation for the time being ($\gamma_c = 0$ in equations (1c) and (1d)), and analyze the effect of $\gamma_c$ in later sections. Because sequestration is a bimolecular interaction, $\eta$ is a rate of association and has units of the form nM$^{-1}$h$^{-1}$. Since the binding of $z_1$ and $z_2$ at the rate $\eta$ in equations (1c) and (1d) encapsulates feedback in the system, we know that $\eta$ cannot be too small. If sequestration is sufficiently slow, it will be as if there is no feedback in the circuit at all. More generally, a small $\eta$ corresponds to a slow feedback action that tends to lead to a large overshoot of the desired steady-state, as can be seen in figure 2A.

The question then becomes, how large should $\eta$ be? Briat et al. in their original work on this system observed that some sets of parameters lead to unstable behavior (which we discuss in greater depth in section 2.2), so it is important to investigate whether or not a large $\eta$ could ever destabilize the circuit [20]. Ideally, we would like the find a regime of parameters where the system’s behavior is easy to predict and we do not have to worry about fine-tuning $\eta$. While it would be possible to get a sense for the behavior of $\eta$ by simulating a broad parameter sweep and analyzing the resulting dynamics, we find that it is possible to gain a precise understanding of effects of $\eta$ via theoretical analysis (described in detail in [31]).

Before we can analyze the effects of the sequestration rate on the circuit, we first need some notion of to what quantity it even makes sense to compare $\eta$. If nothing else, this quantity must have the same units as $\eta$, which immediately rules out a direct comparison with any other rate parameter in equation (1), as $\eta$ is the only association rate in the system. Since $\mu$ is the only other parameter that has units of concentration (specifically nM h$^{-1}$), it must be involved in the comparison. Dimensional analysis can get us as far as noting that a quantity of the form $\alpha^2/\mu$, with $\alpha$ taking units of h$^{-1}$, would at least have same units as $\eta$. This is as far as dimensional analysis alone can take us, since the other parameters in the system ($\theta_1$, $\theta_2$, $k$, and $\gamma_p$) are all rates that have units consistent with $\alpha$.

In [31], we find that $\alpha$ should take the form

$$\alpha = \frac{\theta_1 \theta_2 k}{\gamma_p^2}. \quad (2)$$
Figure 2: The Effects of Sequestration Rate on Dynamics. A) Here we show simulations of the circuit in equation (1) with $k = \theta_1 = \theta_2 = \gamma_p = 1 \text{h}^{-1}$, $\gamma_e = 0 \text{h}^{-1}$, and $\mu = 100 \text{nM h}^{-1}$. This leads to a value of $\alpha^2/\mu = 10^{-2} \text{nM}^{-1} \text{h}^{-1}$. We vary $\eta$ over the range $10^{-3} - 1 \times 10^{3} \text{nM}^{-1} \text{h}^{-1}$ and see that for small $\eta$ the system’s response is highly sensitive on the sequestration rate. Once $\eta$ is in the regime described by inequality (3), its dynamics are independent of $\eta$. B) A parametric plot showing this phenomenon, using the overshoot of the desired steady state $x^*_2 = \mu/\theta_2$ as a proxy for system performance. The black curve is generated by a simulated sweep of $\eta$ values over the range $10^{-3} - 1 \times 10^{3} \text{nM}^{-1} \text{h}^{-1}$, demonstrating that the overshoot becomes almost entirely invariant to $\eta$ in the blue region where $\eta > 10 \cdot \alpha^2/\mu$. The colored dots correspond to the parameters of the simulations in A.

While the proof of this results is somewhat technical, we can derive from it a simple guideline for what it means to have $\eta$ be large enough. The aggregate quantity $\alpha$ describes the steady-state rate at which $z_2$ molecules are produced relative to the concentration of $z_1$. In control theory, we would describe $\alpha$ as the open-loop gain of the system, since it captures the amplification between input and output in the absence of feedback. Characterizing the circuit in terms of $\alpha$ may yield practical benefits, as it allows us to sidestep the problem of individually measuring each of the four rate parameters in equation (2). We will see in later sections that many important features of the circuit can be written either in terms of the individual parameters in equation (1), or in terms of $\alpha$.

Combining these ideas, we can now state precisely both what it means for $\eta$ to be large and what effect this has on the circuit. We find that, though the behavior of the circuit is highly dependent on the value of all other parameters in the network, it is surprisingly insensitive to $\eta$, so long as

$$\eta \gg \frac{\alpha^2}{\mu}. \quad (3)$$

This inequality characterizes a separation of timescales between the production and degradation dynamics of the system (captured by $\alpha$ and $\mu$), and the sequestration reaction. Intuitively, so long as sequestration is sufficiently fast, it does not affect the stability and performance of the circuit. This is demonstrated in figure 2A and B, where the system’s response becomes independent of $\eta$ when inequality (3) holds. Here we use the amount of relative overshoot of $x_2$ (defined as $x_2^{\text{max}}/x_2^* - 1$) as a proxy for characterizing the circuit’s behavior. By simulating the circuit with $\alpha^2/\mu = 10^{-2} \text{nM}^{-1} \text{h}^{-1}$ as $\eta$ varies logarithmically between $10^{-3} - 10^{3} \text{nM}^{-1} \text{h}^{-1}$, we see in figure 2B that the behavior of the circuit is insensitive to $\eta$ when it is sufficiently large. The vertical line corresponds to $\eta = 10 \cdot \alpha^2/\mu$.

The circuit’s stability and performance are also robust to changes in the $z_1$ synthesis rate $\mu$. Additionally, we observe that the circuit should be functional for any biologically plausible value of $\mu$. One consequence of inequality (3) is that the steady-state values of $z_1$ and $z_2$ (denoted by $^*$) must be have the relationship $z_1^* \gg z_2^*$. If measuring and comparing the rates in inequality (3) is not feasible, it is possible to tell if a circuit described by equation (1) is in the regime of strong feedback by simply comparing the concentrations of $z_1$ and $z_2$ at steady state.

We should emphasize that the circuit can still be functional when inequality (3) does not hold,
however our analytic techniques yield less insight into what should be expected from the circuit in this regime. It is also important to note that inequality (3) is not sufficient to guarantee good performance (a concept into which we will delve more deeply in later sections). It merely implies that, once $\eta$ is sufficiently large, the qualitative behavior of the circuit (good or bad) will not be affected by varying it. In the next sections, our analysis will focus on the parameter regime where inequality (3) holds and study how the rest of the rates in equation (1) affect the system’s dynamics.

### 2.2 Instability Arises from Production Outpacing Degradation

A central question for all feedback systems is whether or not the closed-loop circuit is stable. In many engineering applications, it can be the case that poorly implemented feedback control can destabilize an otherwise stable process [17]. This effect is quite salient in the sequestration feedback circuit, which was shown to have unstable oscillatory dynamics (known as limit cycles) for some parameter values. Before we can begin to consider how well a given set of parameters perform, we first need a guarantee that the corresponding dynamics have the baseline functionality of being stable. Ideally, we would be able to find system-level constraints that determine a priori when the circuit will be stable or unstable. While there exist straightforward numerical methods to predict whether or not a system with a given set of parameters is stable, it is substantially more difficult to derive general parametric conditions that characterize stability.

![Figure 3: Dynamics can be Either Stable or Unstable](image)

**Figure 3: Dynamics can be Either Stable or Unstable.** Here we demonstrate how inequality (4) affects the dynamics of the circuit in equation (1). A) This simulation uses $k = \theta_1 = \theta_2 = \gamma_p = 1 \text{ h}^{-1}$, $\gamma_c = 0 \text{ h}^{-1}$, $\mu = 100 \text{ nM h}^{-1}$, and $\eta = 10 \text{ nM}^{-1} \text{ h}^{-1}$. We see that $x_2$ shows some transient oscillatory behavior, but ultimately adapts the the steady-state value $\mu/\theta_2$, since $\alpha = 1 \text{ h}^{-1}$ and $\gamma_p = 1 \text{ h}^{-1}$, inequality (4) holds and the system is stable. B) Now we run the simulation with the same parameters, except $\theta_1 = 3 \text{ h}^{-1}$. This implies that $\alpha = 3 \text{ h}^{-1}$, which tells us that inequality (4) no longer holds and the system will become unstable. In this case, the instability takes the form of indefinite oscillations. This figure is adapted from one presented in [31].

Take, for example, the relatively simple circuit described in equation (1), which has 7 free parameters. While it is straightforward to numerically simulate the circuit and investigate different parameter regimes, it is not at all obvious at first glance how the circuit will perform for a particular set of parameter values. We find that, in the limit of strong feedback and no degradation of $z_1$ and $z_2$, there exists a surprisingly simple relationship that determines stability. Using the same notation as in section 2.1, we find that the system is stable if and only if $\alpha < 2\gamma_p$ (as shown in figure 3). This says that we need the net production rate in the circuit $\alpha$ to be slower than twice the degradation rate $\gamma_p$. We can
rewrite this result, using the fact that $\alpha = \theta_1 \theta_2 k / \gamma_p^2$, in the form

$$\frac{3}{\theta_1 \theta_2 k} < \gamma_p.$$

We note that the left-hand term is proportional to the geometric mean of all the production rates in the circuit. This is another perspective from which we can observe that production needs to be, on average, slower than degradation.

An interesting property of this inequality is that $\mu$ and $\eta$ are conspicuously absent. The lack of dependence on $\eta$ echoes the results we discussed in section 2.1, where we showed that the system’s performance is independent of the sequestration rate in the limit of large $\eta$. The fact that the system does not depend on $\mu$ makes sense because it is the only production rate in the system with units of concentration, so it must set the concentration scale for the circuit. We should expect that simply changing the units of concentration in the model should not affect stability. It then follows that varying $\mu$ is analogous to changing the units of concentration, and consequently should not affect stability and performance. We also find that the system is intrinsically stable when there is only one process species in figure 1A, which we prove in [31]. We find that for the case of $n > 2$ species we can derive results analogous to inequality (4), implying that there is a qualitative difference performance between $n = 1$ species and $n \geq 2$ species.

To further investigate inequality (4), we will assume that both the process parameters ($\gamma_p$ and $k$) and the desired set point (determined by $\mu / \theta_2$) are fixed. Since we assumed that we are in the regime where $\eta$ is large enough to not matter, the only remaining parameter to tune is $\theta_1$. We can interpret $\theta_1$ as quantifying the strength of interaction between the control species $z_1$ and the process species $x_1$. In this sense, inequality (4) tells us that there is a limit on how strong the connection between the controller and the process can be. It is then natural to ask how varying $\theta_1$ affects the circuits performance. To do this, we first need to define the relevant system-level performance metrics.

2.3 There is a Performance Tradeoff Between Speed and Robustness

In the previous section we noted that, for a given set of parameters, there is a maximum value that $\theta_1$ can take such that the controller remains stable. Picking the rate $\theta_1$ to be fast will speed up the response of the circuit, however if there is variability in $\theta_1$ the circuit may inadvertently exceed the limit set by inequality (4) and consequently become unstable. In this context, one way to characterize a circuit’s robustness is to determine not only if it is stable, but also that there are no small parameter changes that would result in it becoming unstable. A robust circuit is one that is far from instability, whereas a fragile circuit will be very near the stability boundary in inequality (4).

These observations yield a tradeoff: in order for the response time of the circuit to be as short as possible $\theta_1$ should be fast, however this necessarily makes the circuit more fragile. This notion of robustness is distinct from robustness of steady state, which is more commonly studied in biological contexts [13,14]. While steady-state robustness is fairly easy to define, as it is simply the error between the actual steady state of a system and a desired steady state, this notion of fragility requires more sophisticated mathematical tools and is generally difficult to solve for analytically. We discuss these theoretical methods in detail in [31], where we show that a good approximation of the fragility of the system has the form

$$F = \frac{2\gamma_p^3 + \theta_1 \theta_2 k}{2\gamma_p^3 - \theta_1 \theta_2 k} = \frac{1 + \frac{\alpha}{2}}{1 - \frac{\alpha}{2}}.$$

An equivalent interpretation of $F$ is as quantifying the system’s worst-case amplification of disturbances. As a sanity check we can see that, when we have equality in inequality (4) (i.e. $\theta_1 \theta_2 k = 2\gamma_p^3$), the fragility $F = \infty$ corresponding to the circuit becoming unstable. When $\theta_1 \theta_2 k < 2\gamma_p^3$, $F$ increases monotonically as $\theta_1$ increases. If $\theta_1 = 0$, then $F = 1$ corresponding to no disturbance amplification (but also no control). We see in figure 4A that equation (5) yields a tradeoff curve that concisely capture the relationship between $F$ and $\theta_1$.

The upshot of this characterization is that we can now precisely quantify the performance tradeoff between speed and robustness. Where inequality (4) gave us a binary condition for stability, equation (5) provides a more nuanced measure of the circuit’s performance. Figure 4B demonstrates the
Figure 4: A Performance Tradeoff Between Response Time and Fragility. A) Here we show a tradeoff curve demonstrating the relationship between response time (parametrized by $\theta_1^{-1}$) and fragility (as described in equation (5)). While an ideal system would have both a fast response and be minimally fragile (i.e. robust), this curve shows that, given all other parameters the system only has so much freedom to simultaneously optimize its performance. B) These trajectories corresponding to parameters associated with the colored dots in A, showing how fragility and response time relate to the actual dynamics of the circuit. We see that the blue curve rises quickly, but is highly oscillatory and takes a long time to settle. Conversely, the purple curve has a slow rise time, but is quite robust and settles quickly. In each plot we vary $\theta_1$ and use $k = \theta_2 = \gamma_p = 1 \, \text{h}^{-1}$, $\gamma_c = 0 \, \text{h}^{-1}$, $\mu = 100 \, \text{nM} \, \text{h}^{-1}$, and $\eta = 10 \, \text{nM}^{-1} \, \text{h}^{-1}$. This figure is adapted from one presented in [31].

effects of this tradeoff on the system’s dynamics. We see that, as the initial response time of the system decrease ($\theta_1$ increases), the system begins to oscillate and takes a much longer time to settle in to its steady-state value. These oscillations are indicative of the system approaching instability, a topic we explore in more depth in [31].

2.4 Controller Degradation Improves Stability but Introduces Steady-State Error

So far we have neglected the effects of degradation on the control species $z_1$ and $z_2$, assuming that they are only removed from the system by the sequestration reaction. Under this assumption, the results of the previous section tells us that the system is, in a sense, fundamentally constrained in its performance. We might hope, then, that nonzero controller degradation ($\gamma_c > 0$) might give us an addition knob to tune in designing sequestration feedback circuits.

This turns out to be precisely the case, as is shown in figure 5. If all other parameters are held constant, the control species degradation rate $\gamma_c$ decreases the fragility of the system at the cost of introducing steady-state error. Again assuming the limit of large $\eta$, we show in [31] that there is a simple expression for this error, which we denote by $\varepsilon$:

$$\varepsilon = \frac{\mu/\theta_2 - x^*_2}{\mu/\theta_2} = \frac{1}{1 + \frac{\gamma_c}{\mu/\theta_2}}. \tag{6}$$

We can think of $\varepsilon$ as capturing the steady-state error of $x_2$ relative to the desired set point we would expect in the absence of controller degradation ($\mu/\theta_2$). It is useful to think of controller degradation as adding ‘leakiness’ into the feedback mechanism. Imagine a scenario where $x_2 > \mu/\theta_2$. Ideally, this would cause an increase in $z_2$ that precisely compensates for the mismatch. This increase in $z_2$ will reduce the amount of $z_1$ via sequestration. If, however, $z_1$ is also degraded, there will be two
mechanisms through which $z_1$ is removed, namely degradation and sequestration. If we imagine a scenario in which $\gamma_c$ is extremely large, then every $z_1$ molecule would likely be degraded before having a chance to be sequestered by a $z_2$ molecule. If this were the case, then $x_2$ cannot possibly be used. Feedback to track the set point $\mu/\theta_2$, because there is effectively no way for the increase in $z_2$ to affect the concentration of $z_1$. In equation (6), we see that $\epsilon = 0$ if $x_2^* = \mu/\theta_2$ (no error), and $\epsilon = 1$ when $x_2^* = 0$ (maximum error).

**Figure 5: Controller Degradation Introduces Steady-State Error.** A) We see that controller degradation $\gamma_c > 0$ improves stability at the cost of introducing steady-state error. This tradeoff curve is a parametric plot where $\gamma_c$ is varied and equation (6) is compared to a generalization of equation (5) that incorporates the effects of $\gamma_c$. B) Here we show the effects of the tradeoff in A on the dynamics of $x_2$. The parameters are chosen such that, if $\gamma_c = 0$, the system would be unstable. The trajectory with small $\gamma_c$ (blue) is stabilized, but still has long-term oscillations indicative of fragility, but has little steady-state error. The trajectory with large $\gamma_c$ (purple) is extremely robust, but with large steady-state error. We vary $\gamma_c$ and use $k = \theta_2 = \gamma_p = 1 \text{ h}^{-1}$, $\theta_1 = 2 \text{ h}^{-1}$, $\mu = 100 \text{nM h}^{-1}$, and $\eta = 300 \text{nM}^{-1} \text{h}^{-1}$. C) In this tradeoff curve, we hold the error $\epsilon = 0.1$ constant by varying both $\theta_1$ and $\gamma_c$ as a constant ratio. We now observe that there is a tradeoff between fragility and leakiness, the latter parametrized by $\gamma_c$. Intuitively, if $\gamma_c$ is large, then many copies of $z_1$ and $z_2$ are being degraded without ever being involved in the feedback process. D) We see that simulated trajectories display constant steady-state error, with less oscillatory behavior when $\gamma_c$ is large. Panels C and D use the same parameters as A and B, with the exception that $\theta_1$ is no longer fixed. This figure is adapted from one presented in [31].

It is also possible to describe the fragility $F$ for the case where $\gamma_c > 0$. This expression is much more complex than the one presented in equation (5), so we refer the reader to our analysis in [31] for details. The qualitative behavior of $F$ is shown in figure 6A, where increasing $\gamma_c$ decreases fragility. Combining these observations, we see that varying controller degradation introduces a tradeoff between steady-state error and robustness. Increasing $\gamma_c$ reduces $F$ at the cost of increasing $\epsilon$. Figure 5B demonstrates this effect via simulations of equation (1) with increasing values of $\gamma_c$. We see that the trajectory with
large controller degradation (purple) is far more stable than the trajectory with small degradation (blue), however the latter differs substantially from the original set point \( \mu/\theta_2 \).

If it were the case that this error were simply a constant offset, then it could, in principle, be corrected for after the fact. The issue, however, is that the error is highly parameter dependent, as can be seen from equation (6). If the goal is to ensure that the steady-state value of \( x_2 \) is robust to variation in parameters, this error term may undermine the whole purpose of the circuit. Ideally, we would have some way to preserve the increased stability from controller degradation without suffering the consequences of large error. Fortunately, this is possible if we think not only in terms of varying \( \gamma_p \), but also the production rate \( \theta_1 \).

Equation (2), tells us that \( \alpha \) is proportional to \( \theta_1 \). This means that, if we increase \( \theta_1 \) to match an increase in \( \gamma_c \), it would be possible to hold the error in equation (6) constant. This would not necessarily be very interesting if this increase in \( \theta_1 \) increased \( F \) by a corresponding amount, however it turns out that we are still about to decrease \( F \) while keeping the ratio \( \theta_1/\gamma_c \) constant, as seen in figure 5C. We see in figure 5D this effect in simulations, where each trajectory has a constant steady-state error of \( \varepsilon = 0.1 \), however the trajectory with larger \( \gamma_c \) values are significantly less oscillatory. This tells us that, if high turnover of \( z_1 \) and \( z_2 \) is not too costly, it is possible to mitigate the downside of degradation while preserving its benefits.

### 2.5 Sequestration Feedback in a Synthetic Bacterial Growth Control Circuit

The results presented so far have focused on the simple model of sequestration feedback presented in figure 1. This approach has facilitated our development of theoretical results that characterize some of the important features of sequestration as a mechanism for biological feedback control. We will now make use of the insights gained from the simplified model to study a particular synthetic bacterial growth control circuit. We show that the conceptual guidelines developed so far yield practical insight into the design of this circuit. Specifically, we find that the incorporation of controller degradation can lead to dramatically improved performance. The mathematical details of this analysis are presented in [31], here will show simulation results and explain at a high level how the theory can help guide circuit design.

A diagram of the circuit architecture is presented in figure 6A, where growth control is achieved by regulating the production of the toxin CcdB. Conceptually, if the intracellular concentration of toxin is proportional to the total number of cells, then the population as a whole will converge to a steady-state size that is less than the carrying capacity of the environment. The circuit uses a quorum sensing mechanism (involving the autoinducer AHL) to implement the coupling between population size and CcdB expression. The circuit described so far is capable of constant regulation, but lacks an extracellular mechanism through which we can control the population size (assuming that we do not want to be directly tuning protein expression, for example by altering the strength of ribosome binding sites).

The desired control can be implemented with molecular sequestration. For simplicity we will focus our modeling efforts on the particular mechanism of sense-antisense RNA pairing, however it would also be feasible to implement sequestration with sigma factor/anti-sigma factor binding or toxin/anti-toxin binding. An antisense RNA (asRNA) is one that has a complimentary sequence to a messenger RNA (mRNA) strand. This complimentarity allows an asRNA to hybridize with an mRNA and block translation. By controlling the expression of asRNA, it is possible to modulate how responsive CcdB expression is to changes in AHL concentration and thus control the total size of the population. This architecture was originally proposed for experimental purposes in [32], and a functionally similar circuit
Figure 6: A Synthetic Growth Control Circuit. A) The circuit diagram for the dynamics described in equation \((7)\). B) Simulations of the growth control circuit without RNA degradation (solid lines) for various set points \(\mu\) (dashed lines). This architecture exhibits the precise adaptation property, though the response is relatively slow and oscillatory. C) Here we see the same circuit simulated with RNA degradation. The response is much faster and more robust, however there is non-zero steady-state error for each trajectory. D) Here we again simulate the circuit without degradation, but now vary \(k_R\). We see qualitatively similar performance tradeoffs to those in figure 4. E) As before, we see that adding controller degradation yields a very fast and consistent response. For these particular parameters, the circuit can achieve this performance with relatively little steady-state error. For all circuits we use the parameters \(N_m = 10^9\), \(r = 1\ h^{-1}\), \(\eta = 20\ nM^{-1}\ h^{-1}\), \(k_p = 10\ h^{-1}\), \(\gamma_p = 3\ h^{-1}\), \(G_a = 10^{-6}\ nM\), and \(\tau = 4 \times 10^{-3}\ nM^{-1}\ h^{-1}\). Panel B uses \(k_R = 0.1\ h^{-1}\) and C uses \(k_R = 10\ h^{-1}\) and \(\gamma_R = 20\ h^{-1}\).

was tested in [33]. We model this circuit with the following set of differential equations:

\[
\frac{d}{dt}[\text{CcdB}] = k_p[mRNA] - \gamma_p[\text{CcdB}] \tag{7a}
\]

\[
\frac{d}{dt}N = rN \left(1 - \frac{N}{N_m}\right) - \tau[\text{CcdB}]N \tag{7b}
\]

\[
\frac{d}{dt}[\text{mRNA}] = k_RG_aN - \eta[mRNA][\text{asRNA}] - \gamma_R[\text{mRNA}] \tag{7c}
\]

\[
\frac{d}{dt}[\text{asRNA}] = \mu - \eta[mRNA][\text{asRNA}] - \gamma_R[\text{asRNA}] \tag{7d}
\]

Quantities of the form \([\cdot]\) represent intracellular concentrations for each cell, and \(N\) represents the total number of cells. \(N\) follows logistic dynamics with an additional death rate due to toxicity \(\tau\) proportional to the concentration of [CcdB] per cell. [CcdB] is a protein that is toxic to the cell, [mRNA] is the corresponding messenger RNA, the transcription of which we model as being induced by a quorum sensing ligand that is produced at a rate proportional to \(N\), and [asRNA] is an antisense RNA that has a complementary sequence to the CcdB mRNA, thus acting as a sequestering partner. The term \(G_a\) captures the gain between \(N\) and mRNA induction mediated by the quorum-sensing molecule AHL. We can think of [asRNA] and [mRNA] as representing \(z_1\) and \(z_2\), and the quantities
[CedB] and N as representing \( x_1 \) and \( x_2 \). This highlights the generality of the modeling framework in figure 1A: because we did not make any assumptions about the particular nature of the underlying variables, we are able to analyze a circuit with extremely heterogeneous underlying quantities (i.e. RNA, proteins, and cell population).

Figure 6B demonstrates how the growth control circuit adapts to various steady-state population levels when there is no controller degradation (\( \gamma_R = 0 \)). The steady state is set by varying \( \mu \). When possible, parameters for this model are taken from [34]. What is clear across all set points is that the population first grows to carrying capacity before the circuit is activated. Intuitively, the blue curve in Figure 6B has a large amount of asRNA that sequesters mRNA. Because of this, it takes longer to accumulate enough mRNA to make CedB and lower the population level. In contrast, the purple curve has comparatively little asRNA, effectively increasing the rate at which CedB can be produced. Qualitatively similar long-term oscillatory behavior in a CedB-based growth control circuit was observed in [35].

Since [34] does not explicitly model transcription, we would ideally pick realistic transcription and translation timescales for bacteria. If we were to naively assume that we could model asRNA and mRNA as if they were like \( z_1 \) and \( z_2 \) in section 2.3, i.e. neglecting controller degradation and assuming they are only removed via sequestration, then we run into an issue. Because the sequestration mechanism modeled in section 2.3 assumes that controller degradation is negligible, we must use a very small mRNA synthesis rate to achieve stable dynamics, assuming all other parameters are fixed in a biologically plausible regime, even if sequestration is fast. This leads to a slow circuit response and a large transient overshoot. This is demonstrated in figure 6B, where CedB production is so slow that the population reaches carrying capacity before the circuit can become active. In Figure 6D we see similar dynamics to those in figure 4, where the circuit faces harsh tradeoffs between speed and robustness.

We see from figure 6C and E that good performance requires not only that RNA is removed via sequestration, but also that it is degraded at a nontrivial rate. At the cost of a lack of precise adaptation, these circuits display dramatically improved performance. In figure 6C and E, the transient overshoot from figure 6B and D has almost entirely disappeared, and each system adapts on a nearly identical time scale, independent of \( \mu \). Figure 6D and E compare performance for various values of \( k_R \) in each system. Figure 6E shows that the introduction of \( \gamma_R \) makes the system’s dynamics extremely robust to variations in \( k_R \) over a wide range of values. We can interpret \( \gamma_R \) as introducing a third tradeoff dimension, namely steady-state error. By allowing the system flexibility along this axis, its speed and robustness are greatly improved.

### 2.6 Noise and Fragility Are Two Sides of the Same Coin

Our analysis so far has assumed that the underlying circuit is perfectly deterministic, i.e. that its dynamics can be modeled by systems of ordinary differential equations. While these models serve as a good starting point for studying many biomolecular systems, they do not capture the effects of noise on the system. While noise is not always an important feature of biological processes, it can sometimes drastically alter the actual behavior of a circuit in a cell (for example, when certain molecules are at a low copy number) [5,36].

Here we will examine the steady-state variance of the output species \( X_2 \) of the sequestration feedback system when its dynamics follow a stochastic chemical reaction model and relate it to the performance of the deterministic model equation (1). The capitalization in \( X_2 \) reflects that this it is now a random variable, whereas \( x_2 \) in equation (1b) is deterministic. To simplify our analysis, we will return to the assumption that there is no controller degradation (\( \gamma_c = 0 \)). In [20], the authors observed that there exist parameter values such that the deterministic model is unstable, but the average behavior of the stochastic model is stable (i.e. each species, on average, adapts to the desired steady-state concentration). We might think of each stochastic simulation representing a single cell’s dynamics, and the average representing the population-level behavior.

We find that, while this result regarding the mean behavior is correct, it does not tell the whole story. In particular we show that, while the population average is generally well behaved, the population variance can become large. In other words the population as a whole is predictable, however there is
Figure 7: The Relationship Between Noise and Robustness. A) Here we see a general tradeoff between response time and noise (quantified by the Fano factor) of the sequestration feedback network. This is analogous to the tradeoff in figure 4A. The plot demonstrates the approximate behavior of equation (8) (black), simulation results for the same approximate model (blue), and simulation results for the fully nonlinear model without approximations (green). B) For $\theta_1 = 0.75$ h$^{-1}$, the deterministic and stochastic mean converge with good performance; individual stochastic trajectories are not very noisy. C) For $\theta_1 = 1.3$ h$^{-1}$, the deterministic and stochastic mean have damped oscillations; individual stochastic trajectories are noisy. D) For $\theta_1 = 3.5$ h$^{-1}$, the deterministic model is unstable and oscillates while the stochastic mean is stable, as demonstrated in [20]. We see however, that the individual trajectories oscillate with randomized phase. In all simulations $k = \theta_2 = \gamma_p = 1$ h$^{-1}$, $\eta = 10$ h$^{-1}$ nM$^{-1}$, $\mu = 100$ nM h$^{-1}$. Mean trajectories and standard deviations are computed using $N = 1000$ trajectories.
a large amount of cell-to-cell variability at any given time. In fact, it is the case that the noise scales in approximately the same way as the fragility of the system in equation (5) (as shown in figure 7A).

Formally, we derive an approximate expression for the steady-state Fano factor (the variance divided by the mean) of $X_2$ in the limit of fast sequestration:

$$\text{Fano}[X_2] = \frac{\text{Var}[X_2]}{\mathbb{E}[X_2]} \approx \frac{\gamma_p (2\theta_1 k + k\gamma_p + 2\gamma_p^2)}{2\gamma_p^3 - \theta_1 \theta_2 k}$$  

We see in figure 7A that the variability of $X_2$ increases as the deterministic system approaches instability. This illustrates that there is a fundamental tradeoff such that the system can be either fast and noisy or slow and accurate, mirroring the deterministic tradeoff described in figure 4. We can get a sense for why this happens by observing that the denominator in equation (8) is the same as the denominator of equation (5). This tells us that we can expect each expression to grow in the same way as the respective denominators approaches 0. Thus, there is an intimate connection between the sensitivity of the deterministic sequestration feedback system, which corresponds to oscillatory behavior, and the sensitivity of the stochastic sequestration feedback system, which corresponds to increased noise. To give a more concrete sense for this relationship, we present representative simulation results that demonstrate this behavior.

In figure 7B, we see that a slow and robust deterministic performance (in the sense described in section 2.3) corresponds to a stochastic model with low noise. The left panel shows the mean behavior matching closely to the deterministic trajectory, with a fairly small amount of noise throughout the simulations. The right panel displays some sample individual trajectories, which essentially look like we would expect: closely following the mean with small deviations. If we look at figure 7C, we see that the deterministic model and the mean of the stochastic model converge to the reference quickly, with damped oscillations. Just as the fragility of the deterministic system is larger in the left panel, we see that the corresponding noise in the stochastic system is much larger than in figure 7B. Just as speed increased fragility in section 2.3, it appears to increase variability here. We note that these results assume that the sequestration reactions constitutes the only feedback in the system, recent work has shown that additional feedback loops can potentially serve to reduce noise [37].

Finally, figure 7D demonstrates a parameter regime where the deterministic model becomes unstable. In the left panel we see precisely the type of behavior described by Briat et al. in [20], where the stochastic mean appears to converge despite deterministic instability. The right panel, however, demonstrates that each individual trajectory is in fact exhibiting noisy oscillations, but with phases that are randomized relative to one another. Each individual cell is unstable, but this instability averages out at the population level. This highlights the importance of distinguishing between average and individual behavior.

### 3 Discussion

Though we could have, in principle, made some of the qualitative observation presented in this work from simulations alone, it is important to emphasize the fact that these theoretical results not only formalize numerical observations, but also force us to state exactly what it is we are measuring. An important contribution from [20] was not just that the authors proposed a clever mechanism to implement feedback control in biological contexts. They also went to great effort to clearly state and prove the existence of that the circuit is capable of achieving what they claimed. Our work here and in [31] is an attempt to pursue this line of reasoning and further characterize the qualitative and quantitative behavior of this circuit architecture.

More generally, this theoretical perspective sheds light on a variety of non-trivial parameter relationship which we hope will allow researchers avoid the need for brute-force parameter tuning when designing future control circuits. Where this article was intended to provide a relatively non-technical description of the work, the interested reader can find a great deal more mathematical depth and generality in [31]. We believe that this an exciting time for biology, where theory and experiment can productively guide each other towards new and interesting directions of inquiry.
Author Contributions

Conceptualization and Methodology, NO, FX, and JCD; Formal Analysis, NO and FX; Software, NO, FX; Writing, NO; Supervision and Funding, JCD.

Acknowledgements

The authors would like to thank Harry Nunns for providing feedback on the manuscript, Anandh Swaminathan for helping with stochastic simulations, and Reed McCardell for providing insight into the synthetic growth circuit. The project was sponsored by the Defense Advanced Research Projects Agency (Agreement HR0011-17-2-0008). The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

Declaration of Interests

The authors declare no competing interests.

4 Supplemental Information

4.1 Analysis of Stochastic Systems

In this section we derive the approximate expression for the Fano factor of the output species in a first order system size expansion [38–40], write down the chemical master equation for the stochastic system, and perform approximations with justifications to obtain an expression for the Fano factor of the output species.

We describe the biochemical reactions of sequestration feedback system with two process species: With this simplification, the dynamics of the stochastic sequestration feedback system can be described by a continuous-time Markov chain (CTMC) over the counts of species $Z$ at any time, only one of species $Z_1$ and $Z_2$ can be non-zero. If both species are non-zero, then they sequester each other infinitely fast through reaction $Z_1 + Z_2 \rightarrow \emptyset$ until one of them becomes zero. Therefore, we can define variable $Z = Z_1 - Z_2$, which has a one-to-one correspondence to species $Z_1$ and $Z_2$ counts, where positive $Z$ indicates counts of $Z_1$, and negative $Z$ indicates counts of $Z_2$.

With this simplification, the dynamics of the stochastic sequestration feedback system can be described by a continuous-time Markov chain (CTMC) over the counts of species $Z$, $X_1$ and $X_2$ using the following master equation dynamics [9]:

$$p(x_1, x_2, z) = \mu(p(x_1, x_2, z - 1) - p(x_1, x_2, z)) + \theta_1 \max\{z, 0\}[p(x_1 - 1, x_2, z) - p(x_1, x_2, z)] + k_{x_1}[p(x_1, x_2 - 1, z) - p(x_1, x_2, z)] + \theta_2 x_2[p(x_1, x_2, z + 1) - p(x_1, x_2, z)] + \gamma_p[x_1 + 1)p(x_1 + 1, x_2, z) - x_1p(x_1, x_2, z)] + \gamma_p[(x_2 + 1)p(x_1, x_2 + 1, z) - x_2p(x_1, x_2, z)],$$

where $p(x_1, x_2, z; t)$ denotes the probability for the system to have $Z = z$, $X_1 = x_1$, and $X_2 = x_2$ at time $t$. 

\[\text{(9)}\]
We observe that all the terms on the right hand side of Equation (equation (9)) are linear, except for the max\{z, 0\} term. We can see this more clearly if we consider the first moment equation.

If we consider the steady state master equation, we set the left hand side to 0 and we apply \(\sum_{x_1, x_2, z} x_1\) with the sum over all \(x_1, x_2 \in \mathbb{N}, z \in \mathbb{Z}\), then we obtain that

\[
\theta_1 E(Z | Z \geq 0)P(Z \geq 0) = \gamma_p E(X_1)
\]

Similarly, if we apply \(\sum_{x_1, x_2, z} x_2\), and \(\sum_{x_1, x_2, z} z\), we get

\[
k E(X_1) = \gamma_p E(X_2) \quad \mu = \theta_2 E(X_2).
\]

The term that prevents us from solving this set of linear equations for the first moments is the max\{z, 0\} term, which results in the probability for \(Z\) to be non-negative in the moment equations.

Therefore, we make a second assumption that \(Z \geq 0\) with probability 1 at steady state. This means \(Z_2\) is zero with probability 1 and this represents a good approximation if the system is stable, without \(Z_1\) oscillating to a very low count.

Under this assumption, we then obtain the linear equation:

\[
\theta_1 E(Z) = \gamma_p E(X_1).
\]

Similarly, if we apply sum \(\sum_{x_1, x_2, z} x_1 z\) to the master equation, we obtain a system of linear equations for steady-state moments of both the first and the second order terms. As the system of equations becomes cumbersome to solve by hand, a Mathematica script was written to automatically derive and solve the moment equations. Solution gives the Fano factor of \(x_2\) as the following:

\[
\frac{\text{Var} X_2}{E X_2} = \frac{\gamma_p (2\theta_1 k + k\gamma_p + 2\gamma_p^2)}{2\gamma_p^3 - \theta_1 \theta_2 k}.
\]

As \(\gamma_p \to \infty\), we obtain an additional simplification

\[
\frac{\text{Var} X_2}{E X_2} \sim 1 + \frac{k}{2\gamma_p}.
\]

References


