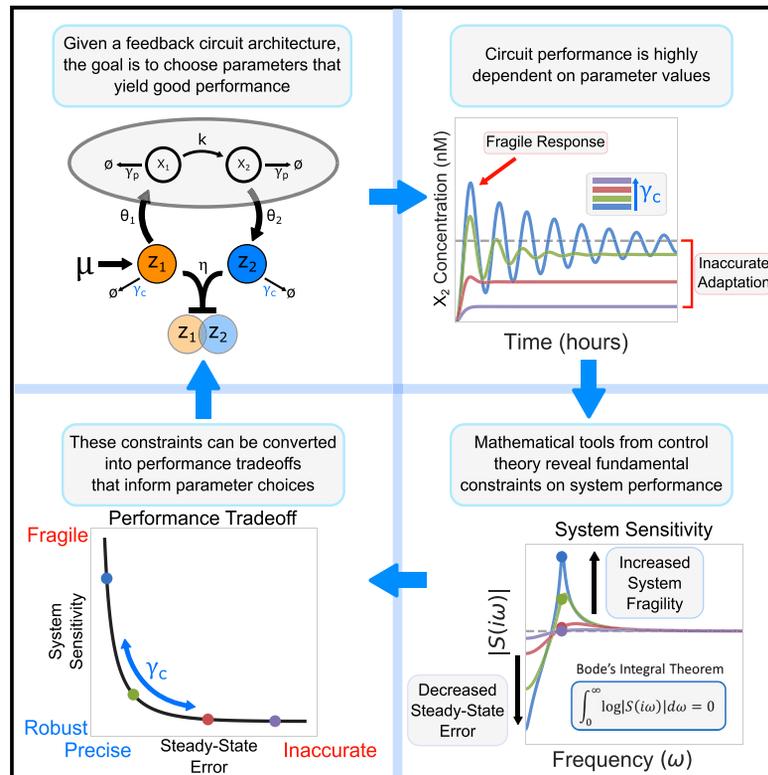


Hard Limits and Performance Tradeoffs in a Class of Antithetic Integral Feedback Networks

Graphical Abstract



Highlights

- Feedback control is an essential component of biomolecular systems
- The design of feedback systems necessarily imposes performance tradeoffs
- We use control theory to study an important class of molecular feedback motifs
- Our work provides a map between biochemical parameters and circuit performance

Authors

Noah Olsman, Ania-Ariadna Baetica, Fangzhou Xiao, Yoke Peng Leong, Richard M. Murray, John C. Doyle

Correspondence

noah_olsman@hms.harvard.edu

In Brief

While feedback regulation is pervasive at every level of biology, it has proven difficult to design synthetic biomolecular feedback systems that match the performance found in nature. The recently developed antithetic integral feedback motif provides a promising mechanism for the implementation of robust control of molecular processes. Our work applies mathematical tools from control theory to this motif, with the goal of taking steps towards the development of a coherent theoretical framework to guide the design of synthetic feedback networks. We characterize the stability and performance tradeoffs of the network, clarifying the relationship between low-level biomolecular rate parameters and high-level system performance (e.g., speed, robustness, tracking error). While these observations can be taken separately, we highlight that a mathematical result known as Bode's integral theorem provides a unifying framework for considering the fundamental constraints on feedback control systems.



Hard Limits and Performance Tradeoffs in a Class of Antithetic Integral Feedback Networks

Noah Olsman,^{1,2,5,*} Ania-Ariadna Baetica,^{1,3} Fangzhou Xiao,⁴ Yoke Peng Leong,¹ Richard M. Murray,^{1,4} and John C. Doyle^{1,4}

¹Department of Control and Dynamical Systems, California Institute of Technology, 1200 E, Pasadena, CA 91125, USA

²Department of Systems Biology, Harvard Medical School, Boston, MA 02215, USA

³Department of Biochemistry and Biophysics, University of California, San Francisco, Box 2542, San Francisco, CA 94158, USA

⁴Division of Biology and Biological Engineering, California Institute of Technology, 1200 E, Pasadena, CA 91125, USA

⁵Lead Contact

*Correspondence: noah_olsman@hms.harvard.edu

<https://doi.org/10.1016/j.cels.2019.06.001>

SUMMARY

Feedback regulation is pervasive in biology at both the organismal and cellular level. In this article, we explore the properties of a particular biomolecular feedback mechanism called antithetic integral feedback, which can be implemented using the binding of two molecules. Our work develops an analytic framework for understanding the hard limits, performance tradeoffs, and architectural properties of this simple model of biological feedback control. Using tools from control theory, we show that there are simple parametric relationships that determine both the stability and the performance of these systems in terms of speed, robustness, steady-state error, and leakiness. These findings yield a holistic understanding of the behavior of antithetic integral feedback and contribute to a more general theory of biological control systems.

INTRODUCTION

One of the central goals of systems biology is to gain insight into the design, function, and architecture of biomolecular circuits. When systems biology emerged as a field, there was a focus on the precise measurement of parameters in canonical pathways, for example, those that govern glucose metabolism [Rizzi et al. \(1997\)](#) and developmental signaling [Lee et al. \(2003\)](#); [Schoeberl et al. \(2002\)](#). As both our understanding of these pathways and our quantitative measurements improved, it became apparent that many of the underlying circuit parameters are subject to large amounts of variability, despite the circuit's overall performance being robust [Goentoro and Kirschner \(2009\)](#); [Shinar et al. \(2007\)](#); [El-Samad et al. \(2005\)](#); [Barkai and Leibler \(1997\)](#). These observations led to the important insight that biological networks have evolved sophisticated feedback control mechanisms that confer robustness, similar to those developed for classical engineering systems [Chandra et al. \(2011\)](#); [Yi et al.](#)

[\(2000\)](#); [Shimizu et al. \(2010\)](#); [Muzzey et al. \(2009\)](#); [Hancock et al. \(2017\)](#). To this end, understanding the architecture and constraints of these regulatory processes is essential both to assessing the range of biological functions that they can implement and to building functional synthetic systems [Szekely et al. \(2013\)](#); [Stelling et al. \(2004\)](#); [Adler et al. \(2017\)](#).

For many systems, the key to achieving robust performance is feedback control, which can provide robustness to both external noise and disturbances and to internal system variability [Ferrell \(2016\)](#); [El-Samad et al. \(2002\)](#); [Yi et al. \(2000\)](#); [Aström and Murray \(2008\)](#). When the system undergoes a change, such as an external disturbance or a variation in parameters, feedback can ensure that the system returns to its desired steady state with a small error [Aström and Murray \(2008\)](#). Additionally, feedback control can stabilize and speed up unstable or slow processes [Ma et al. \(2009\)](#); [Chandra et al. \(2011\)](#); [Stein \(2003\)](#). However, feedback must be correctly designed and tuned, as it can inadvertently amplify noise and exacerbate instability [Del Vecchio and Murray \(2015\)](#); [Aström and Murray \(2008\)](#). Despite some limitations, feedback control is ubiquitous in natural biological systems, where it serves to regulate diverse processes such as body temperature, circadian rhythms, calcium dynamics, and glycolysis [Werner \(2010\)](#); [Rust et al. \(2007\)](#); [El-Samad et al. \(2002\)](#); [Chandra et al. \(2011\)](#); [Ferrell \(2016\)](#).

There are a variety of circuit architectures capable of implementing feedback control in a biomolecular network. However, the time scale and dynamic range of their response can vary greatly depending on implementation details, such as whether the circuit relies on gene regulation [Cohen-Saidon et al. \(2009\)](#), post-translational modification [Barkai and Leibler \(1997\)](#), or RNA interactions [Agrawal et al. \(2018\)](#). Similarly, some circuits are robust over a broad range of inputs [Sourjik and Wingreen \(2012\)](#), while others may have a more modest functional range of response [Goentoro and Kirschner \(2009\)](#).

A particularly interesting class of biological control circuits was recently proposed by [Briat et al. \(2016\)](#). The authors showed that feedback implemented with an antithetic bimolecular mechanism is equivalent to integral feedback control [Aström and Murray \(2008\)](#), which guarantees perfect steady-state adaptation of the output of a network to an input signal [Ferrell \(2016\)](#). For example, an endogenous biological system that uses antithetic



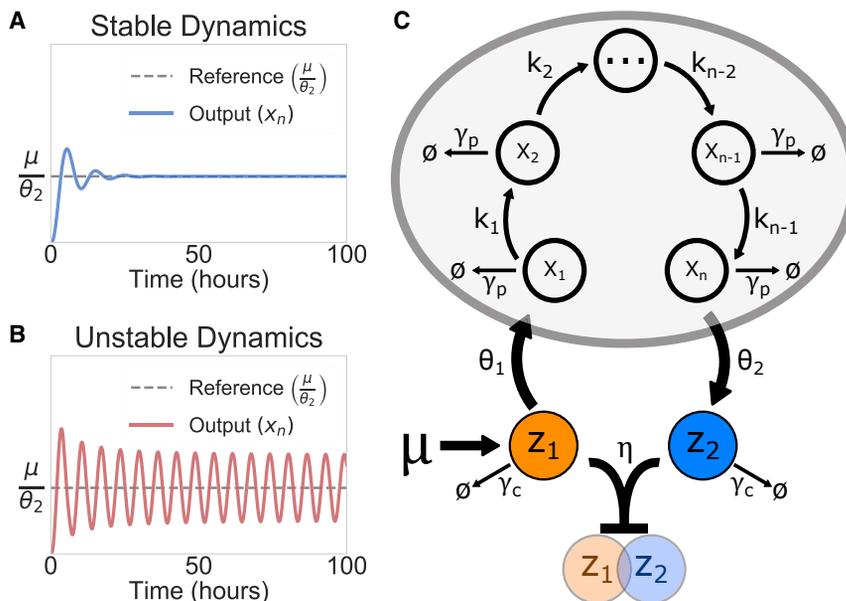


Figure 1. The Antithetic Integral Feedback Network

(A–C) (A) Stable dynamics of an antithetic integral feedback system, where the output (solid line) precisely adapts to a reference signal (dashed line). (B) Unstable dynamics of the same circuit, where the system is now in a parameter regime that results in sustained oscillations that never converge to the desired reference point μ/θ_2 . (C) A class of antithetic integral feedback networks. This general model has two control species, Z_1 and Z_2 , and n process species. The two controller species are subject to a reaction with binding rate η . Additionally, we assume that the binding of the two controller species is much faster than their unbinding. The process species production rates are denoted as $\theta_1, \theta_2, k_1, \dots, k_{n-1}$. For simplicity, the process species degradation rate γ_p is assumed to be equal for each X_i , as is the controller species degradation rate γ_c . This class of networks is defined by a simple set of possible processes where each species is only involved in the production of the next species.

integral feedback involves the binding of sigma factor σ^{70} to anti-sigma factor Rsd (Jishage and Ishihama (1999)). Examples of synthetic biological systems that employ antithetic integral feedback include a concentration tracker (Hsiao et al. (2015); Franco et al. (2014)), two bacterial cell growth controllers (McCardell et al. (2017)), and a gene expression controller (Annunziata et al. (2017)). Recent work has shown that antithetic integral controllers can be implemented in a variety of *in vivo* circuits, using both RNA and protein-based mechanisms (Huang et al. (2018); Kelly et al. (2018); Lillacci et al. (2017)).

While integral control is a powerful tool, its stability and performance are not guaranteed to be well-behaved. Even if both the controller and the network being controlled are stable, their closed-loop dynamics may be either stable (Figure 1A) or unstable (Figure 1B). If the closed-loop system is stable, performance can be characterized by metrics such as tracking error, response time, leakiness, and sensitivity to disturbances. Although these metrics can be optimized individually, they can rarely all yield good results simultaneously because of the constraints imposed by performance tradeoffs. These hard limits have been studied in a variety of contexts, for example, in general stochastic biological control systems (Lestas et al. (2010)) and in the particular context of metabolic control in the yeast glycolysis system (Chandra et al. (2011); Hancock et al. (2017)). Though the original work from Briat et al. (2016) proves a variety of general properties of the antithetic integral feedback circuit, such as ergodicity and the existence of both stable and unstable dynamics, the precise relationship between low-level circuit parameters and high-level system performance remains an open question.

We present here an analytical treatment of stability, performance tradeoffs, and hard limits in a class of antithetic integral feedback circuits, graphically represented in Figure 1C. Specifically, we analyze circuits where the process being controlled consists of a chain of reactions with linear rates such that the production X_i is proportional to the concentration of X_{i-1} and each X_i undergoes degradation at a rate that is proportional to

its own concentration. The antithetic feedback interaction is an irreversible bimolecular reaction which occurs at a rate proportional to the concentrations of both an actuator species Z_1 and a measurement species Z_2 , each of which may also undergo degradation independent of their pairwise binding. The analysis we present here is of a linearized model of the system, which reflects a conscious decision to favor tractability over generality. Notably, however, this simplification also focuses our analysis on the control problem most relevant to biological systems. Though, most systems we desire to control are not truly linear, the most basic nature of feedback control is to keep a specific property of a system near an equilibrium point. Consequently, these systems often behave as if they were linear because the dynamics of interest are those of the system responding to disturbances around their set point, which is the physical realization of the core mathematical assumption underlying the linearized analysis of nonlinear systems.

While many biomolecular circuits of interest are too complex to yield clear theoretical results that describe system-level dynamics and performance, we show in Linear Stability Analysis that a class of antithetic integral feedback networks is amenable to theoretical analysis using techniques from control theory. In particular, we find that there exists an analytic stability criterion for a class of antithetic integral feedback systems (described in Figure 1C). This stability criterion gives rise to performance tradeoffs, for example, between speed and sensitivity, since fast-responding controllers are intrinsically less robust. We prove these results both in the case where there is no controller degradation (Performance Tradeoffs and Hard Limits), as in the model from Briat et al. (2016), and in the more biologically realistic context where there may be such degradation (The Effects of Controller Species Degradation) Qian and Del Vecchio (2018). Though we determine many different classes of tradeoffs for the circuit, we find that they can all be viewed as different aspects of Bode's integral theorem, which states a conservation law for the sensitivity of feedback control systems (Åström and Murray

(2008). We also provide a less technical description of these results, as well as an analysis of noise in the system and simulations of synthetic circuit performance, in a companion piece [Olsman et al. \(2019\)](#).

These theoretical tools provide novel insight into both the analysis of endogenous biological systems and the design of synthetic systems, which we demonstrate by applying our results to a synthetic bacterial growth control circuit in [A Synthetic Growth Control Circuit](#). Finally, we demonstrate in [Controlling Autocatalytic Processes](#) that it is possible to develop control architectures that will stabilize an otherwise unstable chemical reaction process. This result points toward new application domains for antithetic integral feedback controllers, such as autocatalytic metabolic networks, that have yet to be explored in detail.

RESULTS

Our goal here will be to develop a mathematical framework to investigate the general constraints that shape the structure of the closed-loop antithetic integral feedback network. For the sake of clarity, we focus the results on the simplest examples of a network regulated by antithetic integral feedback, however, many of the results presented in this section generalize to a broader class of systems (e.g. the case with more network species and the case with controller degradation). While our analysis is by no means completely generalizable, it does give insight into a class of linear processes and corresponding nonlinear processes for which linearized dynamics do well to capture the circuit's behavior near equilibrium. We discuss in the [Method Details](#) section "[Motivation for linearized analysis of nonlinear systems](#)" both the benefits and the limitations of our linearized approach to modeling and analysis.

Model Description

We first describe the simple antithetic integral feedback model proposed by [Briat et al. \(2016\)](#) with two control species (Z_1 and Z_2) and two species in the open-loop network (X_1 and X_2), which corresponds to the case of $n=2$ in the general circuit diagram presented in [Figure 1C](#) with $\gamma_c = 0$. In the control theory literature the network being controlled is often referred to as the process, a convention we will use in the rest of the paper.

We model the full closed-loop network using the following system of ordinary differential equations:

$$\dot{x}_1 = \theta_1 z_1 - \gamma_p x_1, \quad (\text{Equation 1a})$$

$$\dot{x}_2 = k x_1 - \gamma_p x_2, \quad (\text{Equation 1b})$$

$$\dot{z}_1 = \mu - \eta z_1 z_2, \quad (\text{Equation 1c})$$

$$\dot{z}_2 = \theta_2 x_2 - \eta z_1 z_2. \quad (\text{Equation 1d})$$

As a convention, we will use uppercase letters to denote species in the circuit and lowercase letters to denote the corresponding variables in the models throughout the paper. The rates k and γ_p are production and degradation rates that are in-

ternal to the process. The parameters θ_1 and θ_2 are production rates that provide an interface between the process and the controller. An external reference inducer μ determines production rate of Z_1 , and the two control species Z_1 and Z_2 interact with each other at the rate η .

While realistic models of biological circuits often have both more complex interactions and many more states, this model captures much of the important structural information about the antithetic integral feedback system. In particular, [Briat et al. \(2016\)](#) found that the network defined by [Equation 1](#) implements the precise adaptation of X_2 via integral feedback, as shown by the following relationship:

$$\dot{z}_1 - \dot{z}_2 = \mu - \theta_2 x_2 \Rightarrow (z_1 - z_2)(t) = \theta_2 \int_0^t \left(\frac{\mu}{\theta_2} - x_2(t') \right) dt'. \quad (\text{Equation 2})$$

This ensures that, if the system is stable (i.e., $\dot{z}_1 - \dot{z}_2 \rightarrow 0$), then at steady state (denoted with a *) $x_2^* = \mu/\theta_2$. The parametric conditions that guarantee stability are not, however, obvious at first glance. [Briat et al. \(2016\)](#) showed general algebraic conditions that prove the existence of both stable and unstable dynamics of the linearized antithetic integral feedback system (using Descartes' rule of sign), however, it is not trivial to use their methods to explicitly describe stability in general.

We find that, in the limit of strong binding (large η), there is a simple closed-form criterion for system-level stability. Later, we will show that a one-state network is intrinsically stable for all parameters, and that there exists a simple stability criterion for the general class of networks with many states represented in [Figure 1C](#). For the analysis, we assume both that a set of process parameters (k and γ_p) and a desired set point (determined by μ and θ_2) are given, and we study how stability and performance relate to the rest of the control parameters (θ_1 and η).

Linear Stability Analysis

In this section, we derive an analytic criterion for the stability of the particular class of antithetic integral feedback networks described in [Figure 1C](#). For simplicity, we assume strong binding of the controller species (which we define mathematically later in the section). A key difficulty in studying antithetic integral feedback is the nonlinear term $\eta z_1 z_2$ that mediates feedback in [Equations 1c](#) and [1d](#). Though there exist techniques to study nonlinear feedback systems, there are far more general tools available to study linear ones. While analysis of the linear system does not give guarantees about global behavior, it does allow us to characterize the local stability of the steady state to which we would like X_2 to adapt. We present further discussion of the broader role of linearized analysis in studying feedback control systems and the benefits and limitations that come along with it in the [Method Details](#) section "[Motivation for linearized analysis of nonlinear systems](#)". Here, we linearize the antithetic integral feedback network around the point

$$x_1^* = \frac{\mu \gamma_p}{\theta_2 k}, x_2^* = \frac{\mu}{\theta_2}, z_1^* = \frac{\mu \gamma_p^2}{\theta_1 \theta_2 k}, z_2^* = \frac{\theta_1 \theta_2 k}{\eta \gamma_p^2},$$

the nonzero steady-state values derived from, Equation 1 to derive the following dynamical systems:

$$\dot{\mathbf{x}} = M\mathbf{x},$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ z_1 \\ z_2 \end{bmatrix}, M = \begin{bmatrix} -\gamma_p & 0 & \theta_1 & 0 \\ k & -\gamma_p & 0 & 0 \\ 0 & 0 & -\alpha & -\beta/\alpha \\ 0 & \theta_2 & -\alpha & -\beta/\alpha \end{bmatrix}, \quad (\text{Equation 3})$$

where $\alpha = \theta_1\theta_2k/\gamma_p^2$ and $\beta = \eta\mu$. We can think of α as representing the open-loop rate at which Z_2 molecules are produced from each Z_1 molecule, and β as representing the linearized feedback strength, as it quantifies the linearized rate at which Z_1 is annihilated by Z_2 .

In general, stability of linear systems is determined by the sign of the real parts of its eigenvalues. If they are all strictly negative, then the dynamical system is stable and the system will converge to the equilibrium point. Ideally, we would be able to directly compute the eigenvalues of M ; however, this computation corresponds to finding the roots of a fourth-order polynomial $p(s) = \det(sI - M)$. While this is difficult to do in general, it is possible to study stability by finding what has to be true of the parameters for the system to have a pair of purely imaginary eigenvalues, which characterizes the boundary between stable and unstable behavior. We find that, in the limit of strong binding (specifically $\eta \gg \max(\alpha, \gamma_p) \cdot \alpha/\mu$), M will have the purely imaginary eigenvalues $\lambda = \pm i\omega$ when $\omega = \gamma_p = \sqrt[3]{\frac{\theta_1\theta_2k}{2}}$. More generally, we find that the criterion for stability of the closed-loop antithetic integral feedback circuit described in Equation 3 is

$$\sqrt[3]{\frac{\theta_1\theta_2k}{2}} < \gamma_p, \quad (\text{Equation 4})$$

a relationship we refer to as the production-degradation inequality (proved in The Stability Criterion). In Olsman et al. (2019), we expand on the role of η and how it may affect design decisions.

This implies that the closed-loop (i.e., all parameters in Equation 1. have positive values) system will be stable so long as the degradation rate is larger than a constant that is proportional to the geometric mean of the production rates ($\sqrt[3]{\theta_1\theta_2k}$). We note that, in this strong binding limit, Equation 4 is independent of the controller variables μ and η . Thus, this relationship tells us that stability is purely a function of the parameters describing the process and its connection to the controller, and is independent of the controller itself. Intuitively, the degradation rate sets the rate of adaptation of X_1 and X_2 , so Equation 4 tells us that, so long as the species have a rate of adaptation that is faster than the rate of change in production, the system will be stable.

Through a more technical argument (also in The Stability Criterion), we find that a generalized system with a chain of n process species has a production-degradation inequality of the form

$$\sqrt[n+1]{\frac{\theta_1\theta_2\prod_{i=1}^{n-1}k_i}{\Omega_n}} < \gamma_p, \quad (\text{Equation 5})$$

where Ω_n is a constant that is only a function of the number of process species. When the system has purely imaginary eigenvalues, each species will oscillate at the frequency

$$\omega = \tan\left(\frac{\pi}{2n}\right)\gamma_p.$$

For $n=1$ we get $\omega = \tan(\pi/2)\gamma_p = \infty$, corresponding to an intrinsically stable system (i.e., it cannot oscillate or become otherwise unstable). At $n=2$ we find $\omega = \gamma_p$, so the frequency of oscillation is equal to the process degradation rate. Since $\tan(\pi/(2n))$ is a decreasing function of n , the frequency of oscillation will monotonically decrease as the system grows (assuming a fixed value of γ_p).

Rearranging Equation 5, we get the inequality

$$\alpha < \Omega_n\gamma_p,$$

which says that the degradation rate γ_p sets a bound on how large α can be while still maintaining stability.

For simplicity, the results so far focus only on the strong binding regime. However, we show in The Stability Criterion that there are also tractable and interesting results in the regime of weak feedback (η small). The results have a form similar to that of the strong binding limit, however, the direction of the inequality is reversed. The stability condition for weak feedback is

$$\sqrt[n-1]{\frac{\Omega_n\theta_1\theta_2\prod_{i=1}^{n-1}k_i}{\beta}} > \gamma_p.$$

One interpretation of these results as a whole is that stability is achieved when either feedback or open-loop process production are sufficiently fast, but not when both are.

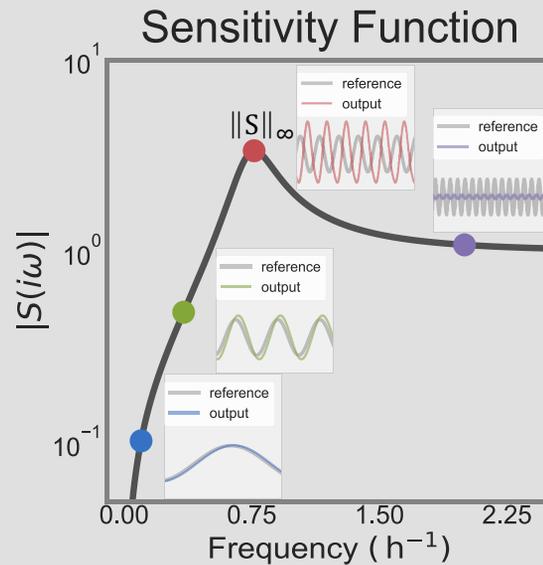
Performance Tradeoffs and Hard Limits

While Equation 4 gives us a binary condition that determines stability, it does not directly tell us about the overall performance of the system. We know when the system becomes unstable, but it is unclear how the system behaves as it approaches instability. Let

$$\mathcal{M} = 1 - \frac{\theta_1\theta_2k}{2\gamma_p^3},$$

which quantifies how far the system is from becoming unstable and which we will refer to as a *stability measure* of the system. For simplicity the analysis here will focus on the $n=2$ species case; however, the results naturally generalize for arbitrary finite n . From Equation 4, $\mathcal{M} = 0$ implies instability, and the larger \mathcal{M} is, the further the system is from becoming unstable. Intuitively it seems that the system should become increasingly fragile as \mathcal{M} approaches zero. Conversely, we can increase \mathcal{M} by decreasing the production rates θ_1 , θ_2 , and k , but this will slow down the dynamics of the system and could potentially hurt performance.

BODE'S INTEGRAL THEOREM AND THE ANATOMY OF A SENSITIVITY FUNCTION



The Sensitivity Function. The sensitivity function for a system, with simulations of reference tracking dynamics for various inputs. We see that when $|S(i\omega)| < 1$, the system has small error and performs well (blue and green). At the peak $|S(i\omega)| = \|S\|_\infty$ (red), we see that the output magnitude is not only amplified, but also phase shifted such that it is almost exactly out of sync with the reference. At high frequencies (purple), the reference is changing so quickly that the system can barely track it.

The primary goal of any control system is to ensure that a process has a desirable response to an input signal, while minimizing the effect of external disturbances (such as noise and systematic modeling errors). While we often think of the time evolution of the full state of a dynamical system $x(t)$, it is often useful to study the input-output relationship of a dynamical system using the (one-sided) Laplace transform

$$X(s) = \int_0^\infty x(t)e^{-st} dt,$$

where it becomes straightforward to mathematically analyze the input-output relationship of a given process.

We call functions that describe the input-output response of a system in the Laplace domain **transfer functions**, and in particular the transfer function between a reference and the output error of a system is the **sensitivity function** of a system $S(s)$. If we take $y(t)$ as the output state of the system (in the antithetic integral feedback circuit $y(t) = x_2(t)$), we denote the Laplace transform of the output $Y(s)$. We can similarly define an input or reference signal $r(t)$ (corresponding to μ) with a corresponding transformed signal $R(s)$. We then define the error of the closed-loop system as $E(s) = R(s) - Y(s)$, and ask how large the error of the system will be when tracking a given reference. This is given by the function

$$S(s) = \frac{E(s)}{R(s)} = \frac{1}{1 + P(s)C(s)},$$

where $P(s)$ and $C(s)$ are the transfer functions for the open-loop process and controller, respectively. The loop transfer function is then the product PC , which describes the open-loop behavior of the full control system. If there exist right-half plane poles of PC , which is to say values of $s = p_k$ such that $\text{Re}(p_k) > 0$ and $PC(p_k) = \infty$, then those poles correspond to unstable eigenvalues of the open-loop system.

Intuitively, when the magnitude of this function $|S(s)|$ is small, then there will be a small tracking error between the reference signal $r(t)$ and the output $y(t)$. Conversely when $|S(s)|$ is large, then there is a large tracking error. If we assume $r(t) = A\sin(\omega t)$, then we can study the frequency response of the system $|S(i\omega)|$ to a sinusoidal input with frequency ω .

$|S(i\omega)|$ provides a way to measure system robustness, by quantifying how well a system attenuates errors to a given input. The worst-case robustness can be described by the maximum value of $|S(i\omega)|$, denoted $\|S\|_\infty$. Ideally we would have $|S(i\omega)| \ll 1$ for all ω . However, a deep result known as Bode's integral theorem (proved in [Bode \(1945\)](#)) states that, if S is a stable rational transfer function with relative degree ≥ 2 (the denominator polynomial has degree that is at least two more than that of the numerator), then the following is true of the closed-loop response:

$$\int_0^\infty \log(|S(i\omega)|)d\omega = \pi \sum_k \text{Re}(\rho_k). \tag{Equation 6}$$

This implies that to reduce error in one frequency range, it *must* be increased elsewhere. This phenomenon is known as the waterbed effect, and sets a fundamental limitation on the performance of any feedback control system. Bode's result was an important early development in control theory, as it helped to formalize the general notion that feedback control problems can be viewed from the perspective of shaping of a system's sensitivity. In the special case where PC has no right-half plane poles (i.e., the open-loop system is stable), we get the simpler relationship

$$\int_0^\infty \log(|S(i\omega)|)d\omega = 0. \tag{Equation 7}$$

To analyze this problem, we will study the sensitivity function $S(s)$, which is the transfer function between the reference signal and the output error of the system [Aström and Murray \(2008\)](#). The sensitivity function is described in greater detail in the box above.

Though there are many different ways to characterize robustness, generally we consider a system to be robust if there are no small changes in parameters that would cause it to become unstable. A mathematically equivalent interpretation is that a system is robust when its worst-case error while tracking references (i.e., the maximum value of S) is small [Aström and Murray \(2008\)](#). For the $n=2$ case of the circuit in [Figure 1](#), we have (see [The Sensitivity Function](#)):

$$|S(i\omega)| = \frac{\gamma_p^2 + \omega^2}{\sqrt{\left(\frac{1}{\omega}\theta_1\theta_2k - 2\omega\gamma_p\right)^2 + \left(\gamma_p^2 - \omega^2\right)^2}}. \tag{Equation 8}$$

The robustness of a system can be formally quantified by $\|S\|_\infty = \max_\omega |S(i\omega)|$, the maximum magnitude of the sensitivity function across all frequencies. (In mathematics, the quantity $\|\cdot\|_\infty$ is referred to as the infinity norm of a function.) The quantity $\|S\|_\infty$ describes the worst-case disturbance amplification for the system to an oscillatory input. If $\|S\|_\infty$ is in some sense small enough to be manageable, then values of $|S|$ across all frequencies are also small and the system is robust to any disturbance. If $\|S\|_\infty$ is large enough to be problematic, there is at least one disturbance against which the system is fragile. How much fragility the system can endure is inherently an application-specific property; for example, a process that is essential to survival (such as ATP synthesis) may need to be much more robust than one that is nonessential to the cell.

Directly computing $\|S\|_\infty$ in terms of the parameters of a system is difficult in general, but it is sometimes possible to compute

good lower bounds that yield insight into a system's robustness. To this end, we find that (see [The Sensitivity Function](#) for a detailed proof):

$$\|S\|_\infty \geq \mathcal{F} = \frac{1 + \frac{\alpha}{2\gamma_p}}{1 - \frac{\alpha}{2\gamma_p}} = \frac{2\gamma_p^3 + \theta_1\theta_2k}{2\gamma_p^3 - \theta_1\theta_2k}, \tag{Equation 9}$$

with equality when,

$$\mathcal{M} = 0 \Leftrightarrow \gamma_p = \sqrt[3]{\frac{\theta_1\theta_2k}{2}} \Leftrightarrow \|S\|_\infty = \mathcal{F} = \infty.$$

The fragility bound \mathcal{F} is constructive, in that we can write down the frequency ω that achieves it:

$$|S(i\omega)| = \mathcal{F} \Leftrightarrow \omega = \sqrt{\frac{\alpha\gamma_p}{2}}.$$

For a given constant reference μ/θ_2 we use [Equation 7](#) to derive a tradeoff between fragility and response time, which we quantify with $1/\theta_1$. While the response time will in principle depend on the interactions of many different parameters, $1/\theta_1$ serves as a good proxy in relative terms (a network with large θ_1 will respond faster than an equivalent network that has a small θ_1). Since we assume the rest of the parameters in the network are fixed, this relative quantification does well to capture the dynamics of the system. It is worth noting that there are limitations to this characterization, as will be seen [The Effects of Controller Species Degradation](#) varying θ_1 has diminishing effects on response time due to other rate limiting reactions downstream of X_1 .

[Figure 2A](#) shows this tradeoff curve for a particular set of parameters as θ_1 varies, with the corresponding dynamics shown in [Figure 2B](#). We see from the latter plots that the response time ($1/\theta_1$) and fragility (\mathcal{F}) correspond directly to the rise times

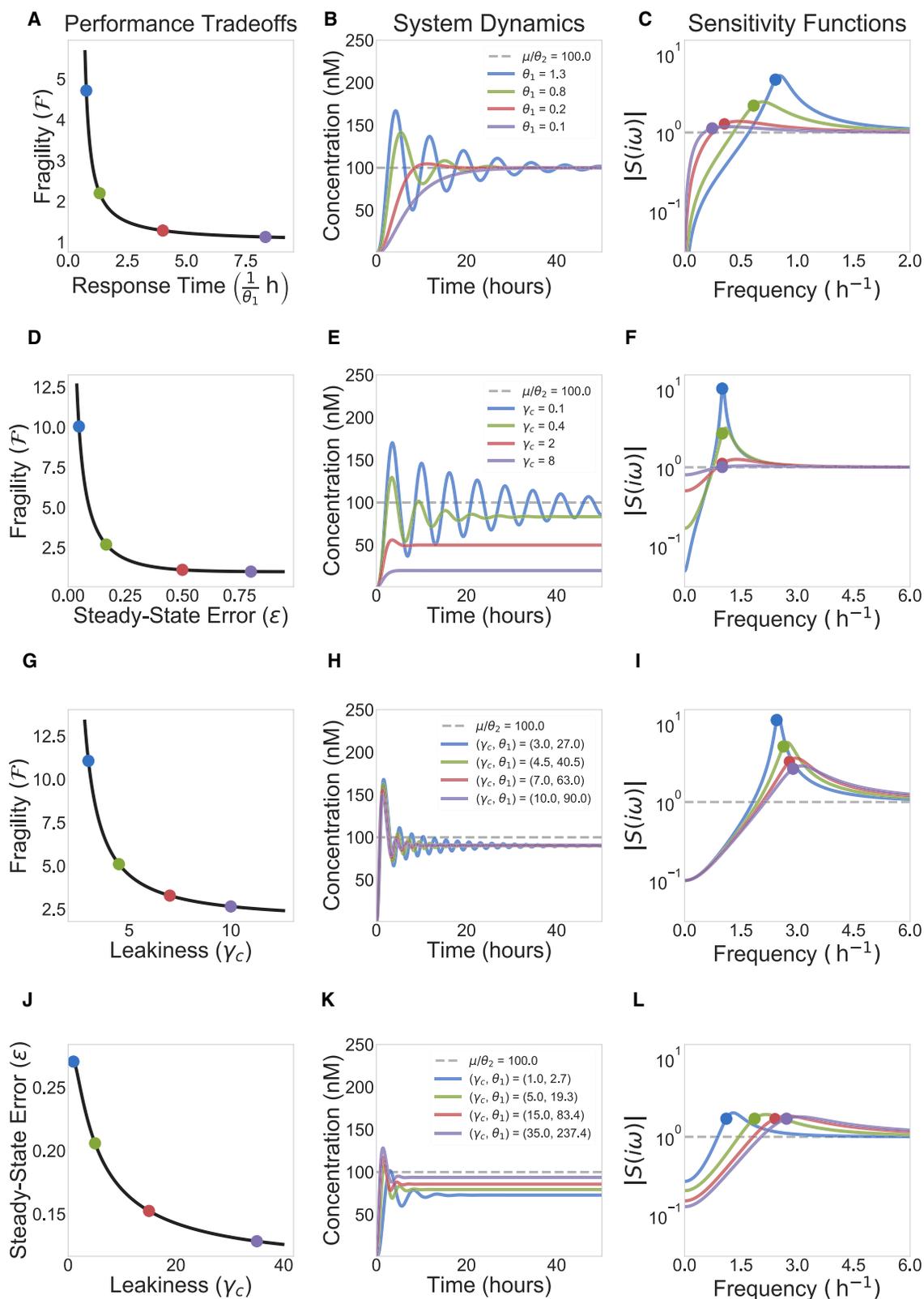


Figure 2. Hard Limits and Performance Tradeoffs in Antithetic Integral Feedback Circuits

(A–L) (A) We see the relationship between speed and fragility in the antithetic integral feedback system. Speed can be characterized in terms of any of the production rates of the system (here we vary θ_1^{-1}), where higher production rates lead to a faster response. Fragility is defined as a lower bound on the maximum

(legend continued on next page)

and oscillatory behavior of simulations in Figure 2B. Figure 2C shows the corresponding sensitivity functions, with colored dots corresponding to values of \mathcal{F} . Here we can clearly see Bode's integral theorem (Equation 7) at work, in that the area above and below the dashed line (corresponding to $\log|S(i\omega)| = 0$) is always equal. We can observe that, as dynamics become more oscillatory, $\|S\|_\infty$ becomes large. While \mathcal{F} is, strictly speaking, only a lower bound of $\|S\|_\infty$, simulation results imply that \mathcal{F} and $\|S\|_\infty$ are both closely matched and respond similarly to changes in parameters.

Because we have fixed μ/θ_2 and assumed that η is large, the only remaining control parameter to vary is θ_1 , so there will only ever be one meaningful tradeoff dimension to study for this system. In the next section, we present results for the case with nonzero controller degradation rates. This model both captures a broader range of biological processes and provides a richer tradeoff space to analyze.

The Effects of Controller Species Degradation

In the previous sections, we assumed that the controller species does not degrade, and we derived an analytic stability criterion for the closed-loop antithetic integral feedback networks. Fulfilling the stability criterion ensures that the antithetic integral feedback network precisely adapts. As discussed, perfect adaptation is a desirable property because it facilitates disturbance rejection and robustness despite variability in process dynamics. However, the literature suggests that implementing antithetic integral feedback with no controller species degradation can potentially be challenging, depending on the biological context of the circuit Ang et al. (2010); Qian et al. (2017). This degradation can act like a leakiness in the system, in that controller species are lost before they have the chance to play a role in the feedback loop. Even if the controller species are not actively degraded, it may be the case that dilution due to cell division has a significant effect on the circuit's behavior. Because of this, we will now extend our analysis of stability, performance, and tradeoffs to antithetic integral feedback networks with nonzero controller species degradation rates.

To model the effects of controller species degradation, we modify Equations 1c and 1d such that,

$$\dot{z}_1 = \mu - \eta z_1 z_2 - \gamma_c z_1, \quad (\text{Equation 10a})$$

$$\dot{z}_2 = \theta_2 x_2 - \eta z_1 z_2 - \gamma_c z_2, \quad (\text{Equation 10b})$$

where γ_c is the degradation rate of the control species z_1 and z_2 .

Including the controller species degradation rate in the antithetic integral feedback network model changes its properties of stability and performance. In particular, the closed-loop antithetic integral feedback network has a zero steady-state error for $\gamma_c = 0$, whereas if $\gamma_c > 0$, there will generally be some nonzero error in X_2 .

In the limit of strong binding, we can analytically compute the steady-state values of the system species and bound its sensitivity function. While it is somewhat more complicated to compute even the steady-state values of each species for this system, we show (see Steady-State Error) that, in the limit of large η , it is possible to derive a simple approximate formula for x_2^* :

$$x_2^* \approx \frac{\mu}{\theta_2} \frac{1}{1 + \frac{\gamma_c}{\alpha}}, \quad (\text{Equation 11})$$

from which all other steady-state values can be derived. Under the strong binding assumption, X_2 no longer precisely adapts to the set point μ/θ_2 , but rather will have some amount of steady-state error determined by the ratio γ_c/α . The relative error in x_2^* can be quantified by the relationship

$$\varepsilon = \frac{\mu/\theta_2 - x_2^*}{\mu/\theta_2} = \frac{1}{1 + \frac{\gamma_c}{\alpha}}. \quad (\text{Equation 12})$$

We see that $\gamma_c = 0 \Rightarrow \varepsilon = 0$, corresponding to our previous results that precise adaptation is achieved when there is no controller degradation. Using this simplified expression, the relative steady-state error function can be bounded. For example, if we are interested in obtaining $\varepsilon < .1$, then we can choose a controller degradation rate such that $\gamma_c < \frac{\alpha}{9}$.

Moreover, we can also derive the corresponding stability criterion (see Stability Analysis with Controller Degradation). Here we present the stability criterion for the two process species case:

$$\frac{\theta_1 \theta_2 k}{2} < \gamma_p (\gamma_c + \gamma_p)^2. \quad (\text{Equation 13})$$

This reduces to Equation 4 when $\gamma_c = 0$, and shows that $\gamma_c > 0$ leads to an increased stability measure. If we only consider variations in γ_c , then the combination of Equation 10 and Equation 11 yields yet another tradeoff. As γ_c increases, the system becomes increasingly stable at the cost of worse steady-state error (see Figures 2D and 2E). In Antithetic Integral Feedback with Controller Species Degradation, we derive a general stability criterion that depends on comparing the controller and the process species degradation rates for $n > 2$ process species. When

value of the sensitivity function $\|S\|_\infty$ as defined in Equation 7. (B) Time-domain simulations corresponding to different points on the tradeoff curve in A. We see that speed and fragility naturally relate to the rise time and settling time of the system. (C) Sensitivity functions for various parameter values. We see what is known in control theory as a waterbed effect, where better attenuation of disturbances at low frequencies necessarily implies worse amplification of disturbances at higher frequencies because of Equation 7. (D) Here we set $\gamma_c > 0$ and observe the effects of controller degradation being varied on its own. We set $\theta_1 = 2 \text{ h}^{-1}$ so that, if $\gamma_c = 0$, the system would be unstable. We see that increasing γ_c decreases fragility, at the cost of introducing steady-state error, which is illustrated in the dynamics shown in panel. (E and F) The corresponding sensitivity functions also illustrate the tradeoff, where the peak of $|S(i\omega)|$ (fragility) decreases as the value of $|S(0)|$ (steady-state error) increases. \mathcal{F} is now computed using Equation 12. (G, H, I) In these plots we vary both γ_c and θ_1 such that $\theta_1/\gamma_c = 9$, corresponding to $\varepsilon = .1$ in Equation 10. We now observe a tradeoff between fragility and leakiness, the latter being captured by how much turnover of Z_1 and Z_2 is introduced by γ_c . (J, K, L) Finally, we can instead hold \mathcal{F} constant and numerically solve for θ_1 given a value of γ_c . This introduces a tradeoff between steady-state error and leakiness. In all simulations $k = \theta_2 = \gamma_p = 1 \text{ h}^{-1}$, $\eta = 1000 \text{ nM}^{-1} \text{ h}^{-1}$, and $\mu = 100 \text{ nM h}^{-1}$.

$\gamma_p \gg \gamma_c$ or $\gamma_p \approx \gamma_c$, the stability criterion is essentially the same as the production-degradation inequality in Equation 4. However, when $\gamma_p \ll \gamma_c$, stability depends primarily on the value of γ_c , rather than γ_p as in the previous cases.

We now focus on analyzing the properties of the sensitivity function and the tradeoff it introduces. Figure 2F shows the corresponding sensitivity function for this system. One major difference between these sensitivity functions and those in Figure 2C is that we now have $|S(0)| > 0$. This is directly related to the steady-state error in Equation 9, as we can think of a signal with frequency $\omega = 0$ as a constant reference. A convenient property of the sensitivity function is that $|S(0)| = \varepsilon$, so the previously mentioned tradeoff between robustness and steady-state error can be recast as a tradeoff between $|S(0)|$ and $\|S\|_\infty$. In Figure 2C we see that $\log|S(0)| = -\infty$, corresponding to $|S(0)| = 0$, implying $\varepsilon = 0$ steady-state error. Because of the waterbed effect, increasing $|S(0)|$ tends to reduce $\|S\|_\infty$. This can be seen directly by deriving a bound similar to the one in Equation 7 for the case $\gamma_c > 0$:

$$\|S\|_\infty > \mathcal{F} = \frac{(\gamma_p^2 + \omega^2) \sqrt{1 + \left(\frac{\gamma_c}{\omega}\right)^2}}{(\gamma_p^2 - \omega^2) + 2\gamma_c\gamma_p}, \quad (\text{Equation 14})$$

$$\omega = \gamma_p \sqrt{\frac{\alpha + \gamma_c}{2\gamma_p + \gamma_c}}. \quad (\text{Equation 15})$$

Though \mathcal{F} is now more complicated, we can see that it will scale as $\mathcal{O}(1/\gamma_c)$ (i.e. \mathcal{F} is asymptotically bounded by some constant multiple $1/\gamma_c$ for small γ_c). This tells us that increasingly γ_c has the potential to reduce \mathcal{F} . In Figure 2D we see this effect, where \mathcal{F} asymptotically decreases to 1 as γ_c (and consequently ε) increases. It is also straightforward to check that \mathcal{F} reduces to Equation 7 when $\gamma_c = 0$.

So far we have shown what happens when the control parameters θ_1 and γ_c are varied individually; it is also interesting, however, to study what happens when they are varied such that a particular performance characteristic is held constant. Figures 2G–2I demonstrate the system's response when we vary θ_1 and γ_c such that the steady-state error ε is fixed. This sort of variation can be interpreted as changing the turnover rate, and consequently the leakiness (described in more depth in Olsman et al. (2019)), of the controller. This leakiness can also be thought of decreasing efficiency, as it means that control species are degraded before ever being involved in feedback. By increasing γ_c , we make the system less efficient because the controller spends resources producing and then degrading molecules of z_1 and z_2 . Figure 2G shows that highly efficient controllers are more fragile than less efficient ones. We can also see this in Figure 2I, where the integrated area of $|S(i\omega)|$ gets spread out over high frequencies, rather than having large and narrow peaks. This leads to a lower value of $\|S\|_\infty$ and a corresponding increase in robustness to a worst-case input. This represents a tradeoff between having highly localized sensitivity (i.e., there is a small range of frequencies for which the system responds very badly) and distributed sensitivity (i.e., there is a wide range of frequencies, but the performance is not particularly bad for any single frequency).

Conversely, we can fix \mathcal{F} and see how ε changes with leakiness. In Figures 2J and 2K we see that highly efficient controllers have worse steady-state error, and as the controller becomes less efficient, ε improves. This can be observed in Figure 2L, where $|S(0)|$ is reduced as γ_c increases. Because $|S(0)|$ is decreasing and $\|S\|_\infty$ is fixed, we see that $|S(i\omega)|$ stays large at higher frequencies rather than falling off quickly after its peak. These latter two tradeoffs demonstrate how some properties of a control system may shape sensitivity in a relatively nuanced way that is not easily quantified by a particular number; however the overall effect becomes clear when looking at the sensitivity function as a whole.

While any of these tradeoffs could be studied in its own right, the important conceptual takeaway is that what underlies all of them is Bode's integral theorem. In the same way that conservation laws provide a broad understanding of the constraints on physical quantities (like momentum and energy), Equation 7 gives us a unifying framework for understanding the fundamental performance limitations of control systems. With this result in hand, we see that the performance tradeoffs shown here are simply different ways of tuning parameters to shape the function $|S(i\omega)|$.

In this section, we focused primarily on studying one-dimensional curves that relate parameters and performance. This type of analysis inherently makes simplifications, as all parametric relationships are fundamentally slices of a much higher-dimensional tradeoff space. Baetica et al. (2018) analyze these higher-dimensional tradeoffs in a more general context, exploring tradeoff surfaces that combine multiple parameters simultaneously. In the next section, we will apply some of these theoretical concepts to a particular biological circuit model. Though this model is more complex and nonlinear than those we have discussed so far, we will see that the same essential theoretical approach applies.

A Synthetic Growth Control Circuit

Here we will use the results from previous sections to study a simple model of a synthetic antithetic integral feedback circuit based on the work in You et al. (2004); McCardell et al. (2017), illustrated in Figure 3A. The intended function of this circuit is to regulate the population level of a colony of *E. coli* via an external reference signal such as an inducer. We model the circuit with the following set of differential equations:

$$\frac{d}{dt}[\text{CcdB}] = k_p[\text{mRNA}] - \gamma_p[\text{CcdB}] \quad (\text{Equation 16a})$$

$$\frac{d}{dt}N = rN \left(1 - \frac{N}{N_m}\right) - \tau[\text{CcdB}]N \quad (\text{Equation 16b})$$

$$\frac{d}{dt}[\text{mRNA}] = k_R G_a N - \eta[\text{mRNA}][\text{asRNA}] - \gamma_R[\text{mRNA}] \quad (\text{Equation 16c})$$

$$\frac{d}{dt}[\text{asRNA}] = \mu - \eta[\text{mRNA}][\text{asRNA}] - \gamma_R[\text{asRNA}]. \quad (\text{Equation 16d})$$

Quantities of the form $[\cdot]$ represent intracellular concentrations for each cell, and N represents the total number of cells.

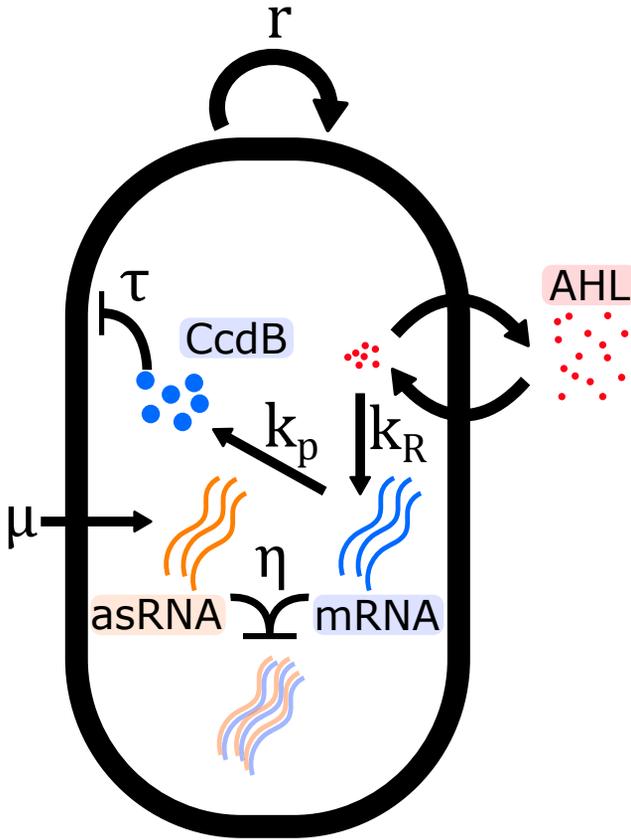


Figure 3. A Synthetic Growth Control Circuit

The circuit diagram for the dynamics described in Equation 16. This circuit controls the growth of a bacterial population via the toxin CcdB. The concentration of CcdB is in turn regulated by a quorum-sensing molecule AHL, whose mRNA can be sequestered by an antisense Agrawal et al. (2018). This circuit is inspired by the work in You et al. (2004) and has been implemented in McCardell et al. (2017). This figure is adapted from Olsman et al. (2019).

N follows logistic dynamics with an additional death rate due to toxicity τ 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 a short antisense RNA that has a complementary sequence to the CcdB mRNA, thus acting as a binding partner. The term $G_a = 10^{-6}$ nM captures the gain between N and mRNA induction mediated by the quorum-sensing molecule AHL.

As before, we will analyze a linearized version of this circuit. To do this we must first compute the steady-state values, shown in Table 1. The linearized dynamics can now be written as

$$\dot{\mathbf{x}} = M\mathbf{x},$$

where

$$\mathbf{x} = \begin{bmatrix} [\text{CcdB}] \\ N \\ [\text{mRNA}] \\ [\text{asRNA}] \end{bmatrix}, M = \begin{bmatrix} -\gamma_p & 0 & k_p & 0 \\ -T & -\gamma_N & 0 & 0 \\ 0 & \hat{k}_R & -\nu - \gamma_R & -\rho \\ 0 & 0 & -\nu & -\rho - \gamma_R \end{bmatrix},$$

and $\hat{k}_R = k_R G_a$, $T = \tau N^*$, $\gamma_N = r N^* / N_m$, $\alpha = (\hat{k}_R \tau N_m) / (\gamma_R r)$, $\nu = [\text{asRNA}]^*$, and $\rho = [\text{mRNA}]^*$. From this, we can again derive stability results in the limit of large η . In terms of the parameters in M , we get a similar relationship to that of Equation 11, with the introduction of heterogeneous degradation rates (discussed in more generality in Baetica et al. (2018)):

$$k_p \hat{k}_R T < (\gamma_p + \gamma_R)(\gamma_N + \gamma_R)(\gamma_p + \gamma_N), \quad (\text{Equation 17})$$

and the corresponding stability measure

$$\mathcal{M} = 1 - \frac{k_p \hat{k}_R T}{(\gamma_p + \gamma_R)(\gamma_N + \gamma_R)(\gamma_p + \gamma_N)}.$$

A notable difference about this circuit is that stability is implicitly dependent on μ . This is because μ appears in N^* , which determines the values of γ_N and T . Given that the function of this circuit is to control cell proliferation, it is natural to ask what steady-state levels of N^* are achievable for a given set of parameters. Because the scale of N^* is set by N_m , we can nondimensionalize the population size with the term N^* / N_m . In the case of $\gamma_R = 0$, we can recast Equation 17 as

$$\frac{N^*}{N_m} = \frac{\mu}{\hat{k}_R N_m} > \frac{k_p}{\gamma_p} \frac{\tau \hat{k}_R N_m}{r^2} - \frac{\gamma_p}{r}. \quad (\text{Equation 18})$$

One immediate result of Equation 18 is that, if the following holds:

$$\frac{k_p}{\gamma_p} \frac{\tau \hat{k}_R N_m}{r^2} < \frac{\gamma_p}{r} \Rightarrow \frac{\tau \hat{k}_R N_m}{r} < \frac{\gamma_p^2}{k_p},$$

then the steady-state N^* is stable for any μ such that $\frac{\mu}{\hat{k}_R} < N_m$ (the steady-state value of N^* cannot exceed the carrying capacity N_m in Equation 16b from the nonlinear model). This constraint is also implicit in the steady-state value $[\text{asRNA}]^*$, which is infinite if $\frac{\mu}{\hat{k}_R} = N_m$. Because the right-hand side of the inequality has a factor of γ_p^2 / k_p , it is possible to improve performance without changing the steady-state concentration of [CcdB] by increasing both k_p and γ_p simultaneously, effectively increasing the protein's turnover rate. If the right-hand side of Equation 18 is positive, then we see that the system's performance is constrained, in that there is a certain population threshold below which N^* cannot be set. Just as in the previous section, as the system approaches this threshold it will become increasingly oscillatory. These effects were observed experimentally in Balagaddé et al. (2005), which uses the same general growth control architecture as in You et al. (2004).

This section illustrates two key points, the first being that the general theoretical results from our initial analysis can be adapted to specific biologically motivated models of control. The second more general takeaway is that systems that look on the surface to be both biologically and mathematically distinct (e.g. a linear model of a chemical reaction network and a nonlinear population-level growth control circuit) have the same underlying structure. We often think of linearization simply as a method of approximation, but its real power often lies in showing us

Table 1. Steady-State Parameter Values Derived from Equation 16

Species	Steady State	Exact Solutions, $\gamma_R = 0$	Approximate Solutions, $\gamma_R > 0$
N/N_m	N^*/N_m	$\frac{\mu}{N_m \hat{k}_R}$	$\frac{\frac{\mu}{N_m \hat{k}_R} + \frac{\gamma_R}{\alpha}}{1 + \frac{\gamma_R}{\alpha}}$
[CcdB]	$\frac{r}{\tau} \left(1 - \frac{N^*}{N_m}\right)$	$\frac{r}{\tau} \left(1 - \frac{\mu}{N_m \hat{k}_R}\right)$	$\frac{r}{\tau} \frac{1}{1 + \frac{\gamma_c}{\alpha}} \left(1 - \frac{\mu}{N_m \hat{k}_R}\right)$
[mRNA]	$\frac{\gamma_p}{k_p} [\text{CcdB}]^*$	$\frac{\gamma_p r}{k_p \tau} \left(1 - \frac{\mu}{N_m \hat{k}_R}\right)$	$\frac{\gamma_p r}{k_p \tau} \frac{1}{1 + \frac{\gamma_c}{\alpha}} \left(1 - \frac{\mu}{N_m \hat{k}_R}\right)$
[asRNA]	$\frac{\mu}{\eta} \frac{1}{[\text{mRNA}]^*}$	$\frac{\mu k_p \tau}{\eta \gamma_p r} \left(1 - \frac{\mu}{N_m \hat{k}_R}\right)^{-1}$	$\frac{\mu k_p \tau}{\eta \gamma_p r} \left(1 + \frac{\gamma_c}{\alpha}\right) \left(1 - \frac{\mu}{N_m \hat{k}_R}\right)^{-1}$

For the case $\gamma_R = 0$ these solutions are exact, while they are approximated (assuming η large) for $\gamma_R > 0$.

the connection between seemingly different mathematical models. In this case, it becomes clear what the analogous production and degradation rates are in Equations 1 and 16.

This type of system-level theory allows us to abstract away details to see that seemingly different problems can be tackled with the same class of tools. In Olsman et al. (2019), we delve into simulations using biologically plausible parameter values and demonstrate that controller degradation can dramatically improve the circuit's performance at relatively little cost. In *Limitations of Linear Analysis* we present another nonlinear circuit with a linearization which structurally matches that of Equation 1, but whose dynamics are poorly described by the linear theory. This demonstrates some of the limitations of the linearized approach presented in this paper. For circuits like these, there may be parameter regimes where more specialized tools from nonlinear dynamical systems are required to accurately describe global behavior.

Controlling Autocatalytic Processes

The general approach of the results presented so far has been to analyze in detail the simplest classes of networks that can be controlled by antithetic integral feedback. Going forward, it will be important to study networks where both the process and controller have more complex architecture. At the controller level, the antithetic mechanism alone only implements integral feedback. It will be useful to investigate mechanisms that could robustly implement proportional and derivative control mechanisms with the ultimate goal of synthesizing full proportional-integral-derivative (PID) controller Aström and Murray (2008); Briat et al. (2018) in synthetic circuits.

Our results thus far have focused on the application of antithetic integral feedback to processes that are open-loop stable. It will likely be important to study the case of unstable processes, which can occur in autocatalytic networks such as the one involved in glycolysis and other metabolic processes. In control theory, unstable processes require the more general version of Bode's integral theorem from Equation 6:

$$\int_0^{\infty} \ln(|S(i\omega)|) d\omega = \pi \sum_k \text{Re}(\rho_k), \quad (\text{Equation 19})$$

where $\text{Re}(\rho_k)$ is the real part of the right-half plane poles of the system's loop transfer function Aström and Murray (2008). These poles correspond to unstable eigenvalues of the open-loop system, and serve an important role in determining the dynamics of a given feedback system. Larger values of $\pi \sum_k \text{Re}(\rho_k)$ correspond to more global sensitivity to disturbances and harsher performance tradeoffs. Intuitively, this is because the open-loop system is more unstable, making closed-loop stabilization more difficult to achieve.

To demonstrate the nuance and complexity added by unstable processes, we demonstrate two seemingly similar control architectures that yield diametrically opposed behavior. The specific sense in which we use the term 'architecture' here refers to the pattern of connectivity in the network. In contrast we would refer to the circuits analyzed in *Performance Tradeoffs and Hard Limits* and *The Effects of Controller Species Degradation*, which have different parameter values but the same network topology, as having the same architecture. As a simple model of an unstable process, we will use the process described in Figure 4, which has the following dynamics:

$$\dot{x}_1 = k_2 x_2 - \gamma_p x_1,$$

$$\dot{x}_2 = k_1 x_1 - \gamma_p x_2.$$

Since the system is linear, it is straightforward to check that the system is unstable when $k_1 k_2 > \gamma_p^2$. Because of the instability of the process, our controller will need to be repressive rather than activating, as it has been throughout the paper. The left panel of Figure 4 describes a plausible control architecture for such a system:

$$\dot{x}_1 = \frac{k_2 x_2}{1 + \theta_1 z_1} - \gamma_p x_1,$$

$$\dot{x}_2 = k_1 x_1 - \gamma_p x_2,$$

$$\dot{z}_1 = \mu_1 - \eta z_1 z_2,$$

$$\dot{z}_2 = \frac{\mu_2}{1 + \theta_2 x_2} - \eta z_1 z_2.$$

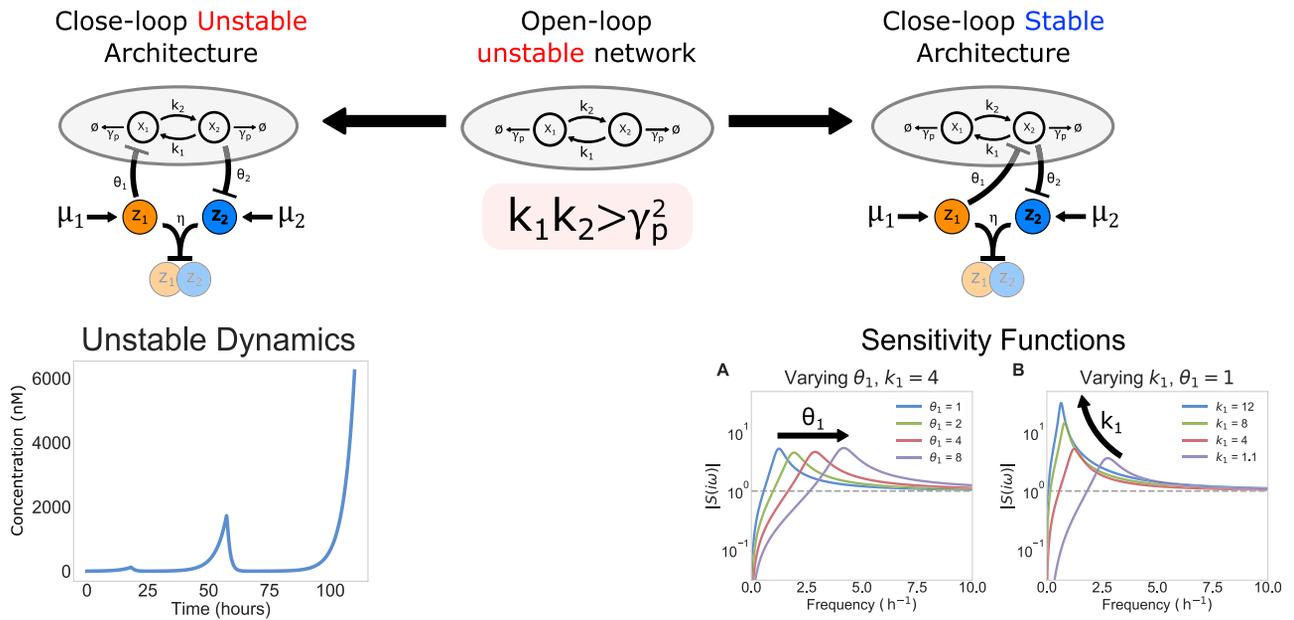


Figure 4. Using Antithetic Integral Feedback to Control an Unstable Network

(A and B) Here we take an unstable process (center) and study two different antithetic integral control architectures (left and right). This network is unstable so long as $k_1 k_2 > \gamma_p^2$. The repressive architecture (represented by flat-headed arrows) on the left is *intrinsically* unstable, in that there are no values of the control parameters that lead to the system's reaching a stable steady state. A representative simulation of the unstable dynamics is presented below the architecture diagram. In contrast, the repressive architecture on the right is not only stabilizing, but *intrinsically* stabilizing. Any nonzero parameter values that result in positive steady-state concentrations of species will yield a stable closed-loop network. (A) shows the sensitivity function as θ_1 varies for a fixed value of $k_1 = 4 \text{ h}^{-1}$. In this case, Equation 19 tells us that the integrated area of the $|S(i\omega)|$ will be constant as θ_1 varies, because θ_1 does not effect the location of right-half plane poles of the loop transfer function. In (B), $\theta_1 = 1 \text{ h}^{-1}$ is fixed and k_1 varies. This will change the location of the right-half plane pole, and we see a consequent change in the integrated area of $|S(i\omega)|$, with large values of k_1 leading to higher overall sensitivity of the system. In all simulations we take $\theta_2 = k_2 = \gamma_p = 1 \text{ h}^{-1}$, $\eta = 1000 \text{ nM}^{-1} \text{ h}^{-1}$, $\mu_1 = 10 \text{ nM h}^{-1}$, and $\mu_2 = 110 \text{ nM h}^{-1}$.

Here Z_1 represses X_1 and X_2 represses Z_2 . Intuitively, if x_2 is large, then Z_1 will be reduced, increasing the amount of Z_1 which in turn reduces the amount of X_1 and X_2 . We prove that this controller is actually *incapable* of stabilizing an unstable process, in that there are no parameters for which the closed-loop system is stable (Controlling an Unstable Process). If, however, we instead have Z_1 directly repress X_2 (Figure 4, right), the model is as follows:

$$\begin{aligned} \dot{x}_1 &= k_2 x_2 - \gamma_p x_1, \\ \dot{x}_2 &= \frac{k_1 x_1}{1 + \theta_1 z_1} - \gamma_p x_2, \\ \dot{z}_1 &= \mu_1 - \eta z_1 z_2, \\ \dot{z}_2 &= \frac{\mu_2}{1 + \theta_2 x_2} - \eta z_1 z_2. \end{aligned}$$

It is not only possible to stabilize the closed-loop system, but the system is *intrinsically* stable. So long as the system has positive parameter values and steady-state concentrations, we recover robust precise adaptation as presented in the earlier sections (Controlling an Unstable Process). While the stable process architecture could either be stable or unstable in closed-loop, this unstable process architecture confers a sort of inherent closed-loop stability that is quite surprising. If the system were linear, this would not be possible. Stability is a direct result of the nonlinearity introduced by

repression. There is, however, a limitation: Equation 19 tells us that a very unstable process ($k_1 k_2 \gg \gamma_p^2$) must exhibit extreme disturbance amplification. In terms of reference tracking, this implies that even the intrinsically stable controller will potentially have very bad transient behavior (e.g. extreme overshoot and ringing as the system stabilizes). While we can use the techniques developed in this paper to mathematically prove why these two architectures behave so differently, we have little biological insight into the architectural requirements for a stabilizing antithetic integral feedback controller. In the future we hope to develop a more general theoretical understanding of which architectures can confer stability to unstable networks.

DISCUSSION

The development of synthetic biomolecular controllers could enable bioengineering to yield new solutions to problems in drug synthesis, immune system engineering, waste management, and industrial fermentation Narcross et al. (2016); Dunlop et al. (2010); Csete and Doyle (2002). In their current state, however, most current synthetic circuits lack the requisite robustness and scalability required of industrial technologies. The application of control theory to synthetic biological controllers aims to ensure that they function robustly in different host organisms and signaling contexts, despite perturbations from an

uncertain environments. Concurrently, being able to understand the design principles of endogenous controllers will both yield insight into natural biological function and give guidance on how to design synthetic systems.

The recent development of antithetic integral feedback controllers represents a promising step towards a more general framework for implementing control in biological networks. This is best demonstrated by the rapid experimental progress toward implementing these controllers in a variety of contexts and with different control species interaction mechanisms [McCardell et al. \(2017\)](#); [Hsiao et al. \(2015\)](#); [Folliard et al. \(2017\)](#); [Lillacci et al. \(2018, 2017\)](#). Our work extends the ideas developed in [Briat et al. \(2016\)](#) and demonstrates a first step toward clarifying the connection between the high-level behavior of the circuit with its low-level rate parameters. We first develop a linearized theory that characterizes the local stability of this class of circuits near the reference point to which we want the system to adapt. We then derive theoretical results that characterize how the parameters of the circuit create relate to performance tradeoffs, which can be understood in terms of classical frequency-domain tools from control theory. Finally we extend the existing model in a variety of ways to probe different aspects of its behavior, for example, by adding degradation of control species and studying various circuits that exhibit interesting nonlinear behavior. We show that, though there are limitations to linear analysis, the linearized theory can capture a large amount of both quantitative and qualitative detail about the system's dynamics. These nonlinear circuits have the common property that their linearization is amenable to the mathematical tools used to analyze the simpler circuit in [Equation 1](#).

As antithetic integral feedback controllers become widely used, we believe that the theoretical results in this paper will not only provide a broad perspective on how the parameters of these networks interact to determine circuit performance, but also provide practical design rules that will tune circuit behavior to meet performance requirements. We begin to investigate these rules in [Olsman et al. \(2019\)](#), where we recast some of the results presented here (as well as some standalone results) as high-level architectural principles for understanding the performance of antithetic integral feedback circuits. We hope that the results here and in related work (e.g., [Briat et al. \(2018\)](#); [Qian and Del Vecchio \(2018\)](#)) will serve as a starting point for broader theoretical exploration of novel biological feedback control systems.

The work presented here provides a link between the tools from classical control theory and contemporary problems in systems and synthetic biology. In particular, we showed that it is possible to explicitly describe parametric conditions that determine stability, performance tradeoffs, and hard limits for a class of antithetic integral feedback controllers. While these limits can each be evaluated on their own, we observe that they can all be interpreted as different aspects of Bode's integral theorem. This result acts like a fundamental conservation law for the performance of feedback control systems. By understanding these general theoretical constraints, we can gain a broad understanding of what is and is not achievable with a given control architecture.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [CONTACT FOR REAGENT AND RESOURCE SHARING](#)
- [METHOD DETAILS](#)
 - Motivation for Linearized Analysis of Nonlinear Systems
 - The Stability Criterion
 - The Sensitivity Function
 - Antithetic Integral Feedback with Controller Species Degradation
 - Limitations of Linear Analysis
 - Controlling an Unstable Process
- [DATA AND CODE AVAILABILITY](#)

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cels.2019.06.001>.

ACKNOWLEDGMENTS

The authors would like to thank Reed McCardell for providing insight into the synthetic growth circuit, James Anderson for useful discussions of tools for analyzing nonlinear systems, and Hana El-Samad for providing feedback on the manuscript. 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.

AUTHOR CONTRIBUTIONS

Conceptualization and Methodology, NO, AAB, FX, YPL, JCD, and RMM; formal analysis, NO, AAB, FX, and YPL; software, NO, AAB, and YPL; writing, NO, AAB, and FX; supervision and funding, JCD and RMM.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 21, 2018

Revised: February 28, 2019

Accepted: May 30, 2019

Published: July 3, 2019

REFERENCES

- [Adler, M., Szekely, P., Mayo, A., and Alon, U. \(2017\). Optimal regulatory circuit topologies for fold-change detection. *Cell Syst* 4, 171–181.](#)
- [Agrawal, D.K., Tang, X., Westbrook, A., Marshall, R., Maxwell, C.S., Lucks, J., Noireaux, V., Beisel, C.L., Dunlop, M.J., and Franco, E. \(2018\). Mathematical modeling of rna-based architectures for closed loop control of gene expression. *ACS Synth. Biol.* 7, 1219–1228.](#)
- [Anderson, J., Chang, Y.-C., and Papachristodoulou, A. \(2011\). Model decomposition and reduction tools for large-scale networks in systems biology. *Automatica* 47, 1165–1174.](#)
- [Anderson, J., and Papachristodoulou, A. \(2012\). A decomposition technique for nonlinear dynamical system analysis. *IEEE Trans. Automat. Contr.* 57, 1516–1521.](#)
- [Ang, J., Bagh, S., Ingalls, B.P., and McMillen, D.R. \(2010\). Considerations for using integral feedback control to construct a perfectly adapting synthetic gene network. *J. Theor. Biol.* 266, 723–738.](#)

- Annunziata, F., Matyjaszkiewicz, A., Fiore, G., Grierson, C.S., Marucci, L., di Bernardo, M., and Savery, N.J. (2017). An orthogonal multi-input integration system to control gene expression in *Escherichia coli*. *ACS Synth. Biol.* **6**, 1816–1824.
- Aström, K.J., and Murray, R.M. (2008). *Feedback Systems: an Introduction for Scientists and Engineers* (Princeton University Press).
- Baetica, A.-A., Leong, Y.P., Olsman, N., and Murray, R.M. (2018). Design guidelines for sequestration feedback networks. *bioRxiv* <https://www.biorxiv.org/content/early/2018/10/30/455493>.
- Balagaddé, F.K., You, L., Hansen, C.L., Arnold, F.H., and Quake, S.R. (2005). Long-term monitoring of bacteria undergoing programmed population control in a microchemostat. *Science* **309**, 137–140.
- Barkai, N., and Leibler, S. (1997). Robustness in simple biochemical networks. *Nature* **387**, 913–917.
- Bode, H.W. (1945). *Network Analysis and Feedback Amplifier Design* (Van Nostrand).
- Briat, C., Gupta, A., and Khammash, M. (2016). Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell Syst.* **2**, 15–26.
- Briat, C., Gupta, A., and Khammash, M. (2018). Antithetic proportional-integral feedback for reduced variance and improved control performance of stochastic reaction networks. *J. R. Soc. Interface* **15**.
- Chandra, F.A., Buzi, G., and Doyle, J.C. (2011). Glycolytic oscillations and limits on robust efficiency. *Science* **333**, 187–192.
- Chandra, F.A., Buzi, G., and Doyle, J.C. (2009). Linear control analysis of the autocatalytic glycolysis system. In *American Control Conference, 2009 (Australia-China Council) '09 (IEEE)*, pp. 319–324.
- Cohen-Saidon, C., Cohen, A.A., Sigal, A., Liron, Y., and Alon, U. (2009). Dynamics and variability of erk2 response to egf in individual living cells. *Mol. Cell* **36**, 885–893.
- Csete, M.E., and Doyle, J.C. (2002). Reverse engineering of biological complexity. *Science* **295**, 1664–1669.
- Del Vecchio, D., and Murray, R.M. (2015). *Biomolecular Feedback Systems* (Princeton University Press).
- Dunlop, M.J., Keasling, J.D., and Mukhopadhyay, A. (2010). A model for improving microbial biofuel production using a synthetic feedback loop. *Syst. Synth. Biol.* **4**, 95–104.
- El-Samad, H., Goff, J.P., and Khammash, M. (2002). Calcium homeostasis and parturient hypocalcemia: an integral feedback perspective. *J. Theor. Biol.* **214**, 17–29.
- El-Samad, H., Kurata, H., Doyle, J.C., Gross, C.A., and Khammash, M. (2005). Surviving heat shock: control strategies for robustness and performance. *Proc. Natl. Acad. Sci. USA* **102**, 2736–2741.
- Ferrell, J.E. (2016). Perfect and near-perfect adaptation in cell signaling. *Cell Syst.* **2**, 62–67.
- Folliard, T., Steel, H., Prescott, T.P., Wadhams, G., Rothschild, L.J., and Papachristodoulou, A. (2017). A synthetic viral feedback loop results in robust expression. *ACS Synth. Biol.* **6**, 1663–1671.
- Franco, E., Giordano, G., Forsberg, P.O., and Murray, R.M. (2014). Negative autoregulation matches production and demand in synthetic transcriptional networks. *ACS Synth. Biol.* **3**, 589–599.
- Gao, X.J., Chong, L.S., Kim, M.S., and Elowitz, M.B. (2018). Programmable protein circuits in living cells. *Science* **361**, 1252–1258.
- Goentoro, L., and Kirschner, M.W. (2009). Evidence that fold-change, and not absolute level, of β -catenin dictates wnt signaling. *Mol. Cell* **36**, 872–884.
- Hancock, E.J., Ang, J., Papachristodoulou, A., and Stan, G.B. (2017). The interplay between feedback and buffering in cellular homeostasis. *Cell Syst.* **5**, 498–508.
- Hsiao, V., de los Santos, E.L., Whitaker, W.R., Dueber, J.E., and Murray, R.M. (2015). Design and implementation of a biomolecular concentration tracker. *ACS Synth. Biol.* **4**, 150–161.
- Huang, H.-H., Qian, Y., and Del Vecchio, D. (2018). A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *bioRxiv*.
- Jishage, M., and Ishihama, A. (1999). Transcriptional organization and in vivo role of the *Escherichia coli* *rsd* gene, encoding the regulator of RNA polymerase sigma d. *J. Bacteriol.* **181**, 3768–3776.
- Kelly, C.L., Harris, A.W.K., Steel, H., Hancock, E.J., Heap, J.T., and Papachristodoulou, A. (2018). Synthetic negative feedback circuits using engineered small rnas. *Nucleic Acids Res.* **46**, 9875–9889.
- Khalil, H.K., and Grizzle, J. (2002). *Nonlinear Systems*, 3 (Prentice Hall).
- Lee, E., Salic, A., Krüger, R., Heinrich, R., and Kirschner, M.W. (2003). The roles of *apc* and *axin* derived from experimental and theoretical analysis of the wnt pathway. *PLoS Biol.* **1**, E10.
- Lestas, I., Vinnicombe, G., and Paulsson, J. (2010). Fundamental limits on the suppression of molecular fluctuations. *Nature* **467**, 174–178.
- Lillacci, G., Aoki, S.K., Schweingruber, D., and Khammash, M. (2017). A synthetic integral feedback controller for robust tunable regulation in bacteria. *bioRxiv*.
- Lillacci, G., Benenson, Y., and Khammash, M. (2018). Synthetic control systems for high performance gene expression in mammalian cells. *Nucleic Acids Res.* **46**, 9855–9863.
- Ma, W., Trusina, A., El-Samad, H., Lim, W.A., and Tang, C. (2009). Defining network topologies that can achieve biochemical adaptation. *Cell* **138**, 760–773.
- Malleshaiah, M., and Gunawardena, J. (2016). Cybernetics, redux: an outside-in strategy for unraveling cellular function. *Dev. Cell* **36**, 2–4.
- McCardell, R.D., Huang, S., Green, L.N., and Murray, R.M. (2017). Control of bacterial population density with population feedback and molecular sequestration. *bioRxiv*.
- Muzzey, D., Gómez-Urbe, C.A., Mettetal, J.T., and van Oudenaarden, A. (2009). A systems-level analysis of perfect adaptation in yeast osmoregulation. *Cell* **138**, 160–171.
- Narcross, L., Fossati, E., Bourgeois, L., Dueber, J.E., and Martin, V.J.J. (2016). Microbial factories for the production of benzylisoquinoline alkaloids. *Trends Biotechnol.* **34**, 228–241.
- Olsman, N., Xiao, F., and Doyle, J.C. (2019). Architectural principles for characterizing the performance of antithetic integral feedback networks. *iScience* **14**, 277.
- Qian, Y., and Del Vecchio, D. (2018). Realizing ‘integral control’ in living cells: how to overcome leaky integration due to dilution? *J. R. Soc. Interface* **15**.
- Qian, Y., McBride, C., and Del Vecchio, D. (2017). Programming cells to work for us. *Annu. Rev. Control Rob. Auton. Syst.*
- Rizzi, M., Baltés, M., Theobald, U., and Reuss, M. (1997). In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol. Bioeng.* **55**, 592–608.
- Rust, M.J., Markson, J.S., Lane, W.S., Fisher, D.S., and O’Shea, E.K. (2007). Ordered phosphorylation governs oscillation of a three-protein circadian clock. *Science* **318**, 809–812.
- Schoeberl, B., Eichler-Jonsson, C., Gilles, E.D., and Müller, G. (2002). Computational modeling of the dynamics of the map kinase cascade activated by surface and internalized egf receptors. *Nat. Biotechnol.* **20**, 370–375.
- Shimizu, T.S., Tu, Y., and Berg, H.C. (2010). A modular gradient-sensing network for chemotaxis in *Escherichia coli* revealed by responses to time-varying stimuli. *Mol. Syst. Biol.* **6**, 382.
- Shinar, G., Milo, R., Martínez, M.R., and Alon, U. (2007). Input-output robustness in simple bacterial signaling systems. *Proc. Natl. Acad. Sci. USA* **104**, 19931–19935.
- Slotine, J.-J.E., Li, W., et al. (1991). *Applied Nonlinear Control*, 199 (Prentice Hall).
- Sootla, A., and Anderson, J. (2017). Structured projection-based model reduction with application to stochastic biochemical networks. *IEEE Trans. Automat. Contr.* **62**, 5554–5566.
- Sourjik, V., and Wingreen, N.S. (2012). Responding to chemical gradients: bacterial chemotaxis. *Curr. Opin. Cell Biol.* **24**, 262–268.
- Stein, G. (2003). Respect the unstable. *IEEE Control Syst.* **23**, 12–25.

- Stelling, J., Sauer, U., Szallasi, Z., Doyle, F.J., III, and Doyle, J. (2004). Robustness of cellular functions. *Cell* 118, 675–685.
- Szekely, P., Sheftel, H., Mayo, A., and Alon, U. (2013). Evolutionary tradeoffs between economy and effectiveness in biological homeostasis systems. *PLoS Comput. Biol.* 9, e1003163.
- Werner, J. (2010). System properties, feedback control and effector coordination of human temperature regulation. *Eur. J. Appl. Physiol.* 109, 13–25.
- Wikipedia contributors. (2018). Hartman–Grobman Theorem — Wikipedia, the Free Encyclopedia, [Online; accessed 3-January-2019] https://en.wikipedia.org/w/index.php?title=Hartman%E2%80%93Grobman_theorem&oldid=861017253.
- Yi, T.M., Huang, Y., Simon, M.I., and Doyle, J. (2000). Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl. Acad. Sci. USA* 97, 4649–4653.
- You, L., Cox, R.S., III, Weiss, R., and Arnold, F.H. (2004). Programmed population control by cell-cell communication and regulated killing. *Nature* 428, 868–871.

STAR★METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Noah Olsman (noah_olsman@hms.harvard.edu).

METHOD DETAILS

Motivation for Linearized Analysis of Nonlinear Systems

For the purposes of this paper, most of the core theoretical results assume a particular class of circuits. Specifically, we examine the architecture presented in [Figure 1](#) with a process containing only linear, unimolecular reactions and a single antithetic integral feedback loop mediated by the interaction of Z_1 and Z_2 . While this architecture is admittedly simple, it has two important features that make it amenable to analytic treatment.

The first is that it allows us to derive clean theoretical results, avoiding as much as possible the complicated parametric relationship we would see in more complex models (e.g., [A Synthetic Growth Control Circuit](#)). We made the conscious decision to favor analytic tractability over generality, as we believe the direct analytic results provide a great deal of insight into the important structure of the problem at hand that would likely be obscured were it to be treated in a more general setting. As such, the goal of this paper is to be constructive whenever possible, rather than focusing primarily on existence proofs and more abstract formulations. For example, if we were to perform the analysis in [The Stability Criterion](#) for a general linear process, it would not be possible to explicitly derive the parametric stability criterion in [Equation 5](#). Instead, the inequality would involve generic statements about the roots of equation [The Stability Criterion](#). By focusing on the chain topology (species X_i being produced by X_{i-1} and producing X_{i+1}), we are able to see directly how the parameters of the model relate to the system's stability.

A second motivation for our analysis is that many of the core results in control theory focus on linear systems. While the linearization of a general physical system will not necessarily accurately capture its full dynamics, the particular case of engineering and biological systems which are designed or evolved to exhibit robust stability around a fixed point are often well characterized by their linearization around that point. This idea is formalized by the Hartman-Grobman Theorem in dynamical systems, which proves the existence of neighborhoods near an equilibrium point for which a nonlinear system will behave indistinguishably from its linearization [Wikipedia contributors \(2018\)](#). A compelling argument motivating the use of simplified and linearized models for studying complex biological phenomena was published in a commentary by [Malleshaiah and Gunawardena \(2016\)](#), which focuses on some recent practical successes of such models in explaining experimental data. Because the purpose of the antithetic integral feedback circuit is to drive the system to a given set of steady-state values, linearized analysis can be particularly well suited to describe the behavior of the system near the desired set point. While these results may break down when the system is driven far from equilibrium, we can often learn a great deal about the nominal operating conditions of a system by developing a thorough understanding of its local behavior.

This is not to say that linearization is the be-all and end-all of control theory. There are many applications for which nonlinearities play a fundamental role and where linearizations fail to capture important aspects of a system's dynamics, which is the case for many contemporary problems in robotics [Slotine et al. \(1991\)](#). Many tools for analyzing the stability of nonlinear control problems involve Lyapunov functions, which serve as a certificate of stability [Khalil and Grizzle \(2002\)](#). Since Lyapunov functions are difficult to construct and existing methods scale poorly with state dimension, a common strategy is to find a decomposition of a given system into well-behaved subsystems which are easier to analyze [Anderson and Papachristodoulou \(2012\)](#). These tools have successfully been applied in the context of both deterministic and stochastic biological systems [Anderson et al. \(2011\)](#); [Sootia and Anderson \(2017\)](#). While these tools are becoming increasingly applicable to problems in biology, they tend to serve more as a way of computationally verifying stability. If the goal is to gain some analytic insight the system, it is often the case that classical tools from linear systems theory are the best option.

Be that as it may, it is often a best practice to start from the linearized theory, where the mathematical tools are both easier to use and more general, and only worry about nonlinearities when there is reason to believe that linearization misses an important aspect of the problem. Mathematically, this comes from the fact that deriving a linearization around a stable fixed point is straightforward, however assessing the size of the fixed point's basin attraction is a more complex problem and is generally not amenable to analytical derivations. As an example of where linearization can potentially do a poor job of describing locally stable behavior, we present in [Limitations of Linear Analysis](#) a nonlinear circuit for which there exist parameters such that the linearized theory predicts local stability of the desired set-point, yet nonlinear simulations of the circuit shows that the basin of attraction for the stable fixed point can be made extremely small.

Be that as it may, the linearized theory in [Results](#) does a very good job of predicting the performance of the various nonlinear simulations we show alongside our analytical results. For example the theoretical tradeoff curves derived in [Figure 2](#), which come from the linearized analysis, do a very good job of predicting the behavior of the accompanying simulations. However, the linearized theory breaks down when the system becomes locally unstable near the desired equilibrium point. Where the linear theory tells us that the dynamics of X_2 will be exponentially unstable near the set point, it is incapable of predicting that X_2 will enter a limit cycle and oscillate indefinitely, as seen in [Figure 1B](#). This is a fundamentally nonlinear phenomenon, and is outside the scope of the methods we use

here. In fact we have made little progress towards proving even the existence of these limit cycles, despite their predictable appearance in simulations.

A middle ground between these two extremes can be seen in A synthetic Growth Control Circuit, where we study a synthetic growth control circuit that is fundamentally nonlinear. Fortunately, the linearization of the model has a structure that is almost identical to that of the simpler architecture presented in Model Description. While the underlying parametric relationships in Equation 17 are more complicated than those of Equation 4, we show in Olsman et al. (2019) that the linearized theory does well to characterize the stability and performance of the nonlinear circuit around the ultimate steady-state population. While our analysis does not capture the initial transients due to the logistic growth expression in Equation 14, the theory does a good job of predicting behavior around the equilibrium point to which we have designed the system to adapt.

The Stability Criterion

We consider the mathematical model of the antithetic integral feedback network described in Equation 1. This mathematical model has a nonlinear term introduced by the bimolecular interaction between Z_1 and Z_2 . To evaluate its properties of stability and performance, we first linearize its dynamics. We can then describe the block structure of the linearized system in terms of the following matrices:

$$A = \begin{bmatrix} -\gamma_p & 0 & \cdots & 0 \\ k_1 & -\gamma_p & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & k_{n-1} & -\gamma_p \end{bmatrix}, B = \begin{bmatrix} \theta_1 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & \theta_2 \end{bmatrix}, D = \begin{bmatrix} -\alpha & -\beta/\alpha \\ -\alpha & -\beta/\alpha \end{bmatrix},$$

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

where $\alpha = (\theta_1 \theta_2 \prod_{i=1}^{n-1} k_i) / \gamma_p^n$ and $\beta = \eta \mu$. The linearized dynamics will then be of the form

$$\dot{\mathbf{x}} = M\mathbf{x},$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ z_1 \\ z_2 \end{bmatrix}$$

To prove our main stability result, we will analyze the characteristic polynomial of M , $p(s)$. The roots of $p(s)$ correspond to eigenvalues of M . In general it is difficult to analyze these roots, however we will see that the $p(s)$ has a great deal of useful structure which we can exploit. First, we have to write down what $p(s)$ actually is.

Lemma S1

The characteristic polynomial of M is

$$p(s) = \det(sI - M) = (s + \gamma_p)^n \left[s^2 + \left(\alpha + \frac{\beta}{\alpha} \right) s \right] + \gamma_p^n \beta.$$

Proof. We start by using the result that, for a block matrix such as M , we can use the classical result from linear algebra

$$\begin{aligned} p(s) &= \det(sI - M) \\ &= \det \begin{bmatrix} sI - A & -B \\ -C & sI - D \end{bmatrix} \\ &= \det(sI - A) \det \left[(sI - D) - C(sI - A)^{-1} B \right]. \end{aligned}$$

Since A is lower-triangular, we see immediately that the first term is

$$\det(sI - A) = (s + \gamma_p)^n.$$

To analyze the second term, we first focus on computing $C(sI - A)^{-1}B$. Because of the sparse structure of B and C , we have

$$C(sI - A)^{-1}B = \begin{bmatrix} 0 & 0 \\ \theta_1\theta_2(sI - A)^{-1}_{n1} & 0 \end{bmatrix},$$

where $(sI - A)^{-1}_{n1}$ is the bottom-left most entry of $(sI - A)^{-1}$. Using Cramer's rule, we can compute

$$\begin{aligned} (sI - A)^{-1}_{n1} &= \frac{1}{(s + \gamma_p)^n} (-1)^{n+1} \det \begin{bmatrix} -k_1 & s + \gamma_p & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -k_{n-2} & s + \gamma_p \\ 0 & 0 & \cdots & -k_{n-1} \end{bmatrix} \\ &= \frac{1}{(s + \gamma_p)^n} (-1)^{n+1} (-1)^{n-1} \prod_{i=1}^{n-1} k_i \\ &= \frac{\prod_{i=1}^{n-1} k_i}{(s + \gamma_p)^n}. \end{aligned}$$

Combing these results, we see that

$$\begin{aligned} \rho(s) &= (s + \gamma_p)^n \det \begin{bmatrix} s + \alpha & \frac{\beta}{\alpha} \\ \alpha - \frac{\theta_1\theta_2 \prod_{i=1}^{n-1} k_i}{(s + \gamma_p)^n} & s + \frac{\beta}{\alpha} \end{bmatrix} \\ &= (s + \gamma_p)^n \left[(s + \alpha) \left(s + \frac{\beta}{\alpha} \right) - \beta + \frac{\gamma_p^n \beta}{(s + \gamma_p)^n} \right] \\ &= (s + \gamma_p)^n \left[s^2 + \left(\alpha + \frac{\beta}{\alpha} \right) s + \gamma_p^n \beta \right]. \end{aligned} \tag{Equation S1}$$

We can now use this result about $\rho(s)$ to analyze the stability of the linearized antithetic integral feedback system.

Theorem S2 (Eigenvalue Classification Theorem)

For a given n and $\beta \gg \max\{\alpha^2, \alpha\gamma_p\}$, the eigenvalues λ of M has a parameter-independent classification of the form $|\arg(\lambda/\gamma_p) + \arg(\lambda/\gamma_p + 1)| = m\pi$, for an integer m .

Proof. To study the eigenvalues of M , we will analyze the roots of $\rho(s)$. We begin by substituting $s = \gamma_p\sigma$ in Equation S1 and setting $\rho(s = \gamma_p\sigma) = 0$:

$$\gamma_p^2 \sigma (1 + \sigma)^n \left[\sigma + \frac{\alpha^2 + \beta}{\alpha\gamma_p} \right] = -\beta.$$

Taking the limit of strong binding ($\beta \gg \alpha^2, \alpha\gamma_p$), this equation reduces to

$$\sigma (1 + \sigma)^n \left[1 + \sigma \frac{\alpha\gamma_p}{\beta} \right] = -\frac{\alpha}{\gamma_p}.$$

From this relationship we see that $\rho(s = \gamma_p\sigma)$ has one large real root at $\sigma \approx -\beta/\alpha\gamma_p$. If we plug this into the phase constraint equation, this gives a phase of $(n + 1)\pi$. We will say the index of this root is $n + 1$. If $|\sigma| \ll \beta/\alpha\gamma_p$, we get the simplified magnitude constraint

$$|\sigma| |1 + \sigma|^n = \frac{\alpha}{\gamma_p}$$

and the phase constraint

$$\arg(\sigma) + n\arg(1 + \sigma) = \pi + 2k\pi = (2k + 1)\pi.$$

We can see from this that the maximum phase possible is $n + 1$ and that each of the indices will be of the form $2k + 1$ (i.e., odd integers). Because the magnitude constraint is independent of k , fundamentally we can have phase indices for any odd integer m such that $|m| \leq n + 1$.

First we will see what conditions can produce purely real roots. If σ is real and $\sigma > 0$, then

$$\arg(\sigma) + n \arg(1 + \sigma) = 0,$$

Which violates the phase constraint. This implies that, if there are unstable roots, they are not purely real. If $-1 < \sigma < 0$, then

$$\arg(\sigma) + n \arg(1 + \sigma) = \pi,$$

and we can have stable real roots with index 1. The magnitude constraint tells us that we will have a pair of these real roots if $(\alpha/\gamma_p) < (n^n/(n+1)^{n+1})$ (which have index 1) with a bifurcation that generates conjugate pairs of roots when $(\alpha/\gamma_p) \geq (n^n/(n+1)^{n+1})$. These conjugate roots will have indices ± 1 .

An immediate result of these observations is that, for any positive odd integer m such that $1 < m < n + 1$, roots cannot be purely real and must come in conjugate pairs $\pm m$. If n is odd, then we will have a conjugate pair of roots for each $m \in [3, n - 1]$, either a pair of small real roots or a conjugate pair for $m = 1$, and a single large negative real root for $m = n + 1$.

If n is even, then the situation will be almost the same except for the fact that there will be a second real root with index $n + 1$. By some simple accounting, this analysis accounts for all $n + 1$ roots of $\rho(s = \gamma_p \sigma)$, which correspond to roots of $\rho(s)$ by a simple re-scaling by $1/\gamma_p$.

Theorem S3 (Production-Degradation Inequality)

Let M be the matrix corresponding to a linearization of the antithetic integral feedback system with two control molecules (z_1 and z_2) and n process species. In the limit of strong binding ($\beta \gg \alpha^2, \alpha\gamma_p$), the system is stable if and only if $\sqrt[n+1]{(\theta_1 \theta_2 \prod_{i=1}^{n-1} k_i)/\Omega_n} < \gamma_p$, where Ω_n is a constant that only depends on n . Further, when the system has purely imaginary eigenvalues the frequency of oscillation will be $\omega = \tan(\pi/2n)\gamma_p$.

Proof. We will prove the results by finding parametric conditions that will result in purely imaginary eigenvalues, and then study what happens to the stability of the system when those parametric conditions do not hold (i.e., equalities become inequalities). To do this, we generalize a technique from Briat et al. (2016), where we evaluate $\rho(s) = 0$ on the imaginary axis. In particular, we pick the change of variable $s = i\omega^* \gamma_p$, where ω^* is a positive real number (which we can assume without loss of generality because complex roots come in conjugate pairs), and evaluate $\rho(s = i\omega^*)$. This yields the equations

$$\rho(s = i\omega^*) = 0 \Rightarrow \gamma_p^2 i \omega^* (1 + i\omega^*)^n (\phi + i\omega^*) = -\beta, \quad (\text{Equation S2})$$

where $\phi = (\alpha^2 + \beta)/\alpha\gamma_p$. If we take the magnitude and phase of the left-hand side of Equation S2, we get the constraints

$$\gamma_p^2 \omega^* (1 + \omega^{*2})^{\frac{n}{2}} \sqrt{\phi^2 + \omega^{*2}} = \beta \quad (\text{Equation S3})$$

$$n \tan^{-1}(\omega^*) + \tan^{-1}\left(\frac{\omega^*}{\phi}\right) = \frac{\pi}{2} + 2k\pi. \quad (\text{Equation S4})$$

From Theorem S2, we know that, in the limit of strong binding, all complex eigenvalues have magnitude much less than ϕ , therefore $\tan^{-1}(\omega^*/\phi) \rightarrow 0$. From these observations, we get the simplified relationship

$$n \tan^{-1}(\omega^*) = \frac{\pi}{2} + 2k\pi \Rightarrow \omega^* = \tan\left(\frac{\pi}{2n} + \frac{2k}{n}\pi\right),$$

and Equation S3 becomes

$$\omega^* (1 + \omega^{*2})^{\frac{n}{2}} \frac{\gamma_p}{\alpha} = 1$$

$$\Rightarrow \gamma_p = \frac{\omega^* (1 + \omega^{*2})^{\frac{n}{2}}}{\Omega_n}, \quad (\text{Equation S5})$$

where $\Omega_n = \omega^* (1 + \omega^{*2})^{\frac{n}{2}}$. We can think of the parametric constraint from Equation S5 as the boundary between stable and unstable behavior in the system. Because the left-hand side of Equation S3 is monotone in ω^* , we can infer that ω^* is unique and consequently there can only be one point in parameter space where there exist purely imaginary eigenvalues.

The final step is to study what happens when Equation S5 does not hold. First we look at the regime $\sqrt[n+1]{(\theta_1 \theta_2 \prod_{i=1}^{n-1} k_i)/\Omega_n} < \gamma_p$. Again using the uniqueness of ω^* , if we understand the stability behavior of the system for a particular value of γ_p in this regime, the same stability behavior must hold for all γ_p in this range. Because of this, we can first examine the range where γ_p is large.

Intuitively, if degradation is sufficiently stronger than production then all species subject to degradation should converge to 0. To prove this rigorously, we will first search for roots with a large magnitude. If we apply the strong binding limit to the characteristic equation from Equation S1, we get

$$p(s) = s(s + \gamma_p)^n \left(s + \frac{\beta}{\alpha} \right) + \gamma_p^n \beta = 0$$

$$\Rightarrow s \left(\frac{s}{\gamma_p} + 1 \right)^n \left(s \frac{\gamma_p \alpha}{\beta} + 1 \right) + \gamma_p \alpha = 0.$$

When $|s| \gg \gamma_p \alpha$, the characteristic equation will have the approximate form

$$\left(\frac{s}{\gamma_p} + 1 \right)^n \left(s \frac{\gamma_p \alpha}{\beta} + 1 \right) = 0,$$

which gives us n roots at $-\gamma_p$ and one root at $-\beta/(\gamma_p \alpha)$. Since Equation S1 is order $n + 2$, we know there is one remaining root outside of this regime. Next, we search for the final small root ($|s| \ll \min(\gamma_p, \beta/(\gamma_p \alpha))$), which gives the relationship

$$s + \gamma_p \alpha = 0,$$

which gives a final small root at $-\gamma_p \alpha$. Since each of the $n + 2$ roots is negative, the system is stable for all $\sqrt[n+1]{(\theta_1 \theta_2 \prod_{i=1}^{n-1} k_i) / \Omega_n} < \gamma_p$.

Now we examine the regime $\sqrt[n+1]{(\theta_1 \theta_2 \prod_{i=1}^{n-1} k_i) / \Omega_n} > \gamma_p$. Here we will use a different technique, as taking the analogous limit of very small γ_p is less straight-forward to analyze. To start, we will perform a change of variable $s = \gamma_p \sigma$, where σ is a complex number. We will again use the strong binding limit, and study roots near the stability boundary, such that the characteristic equation still has the general form

$$\sigma(1 + \sigma)^n = -\frac{\alpha}{\gamma_p}. \quad (\text{Equation S6})$$

If we write $\sigma = a + ib$, we have the magnitude constraint

$$(a^2 + b^2) \left[(1 + a)^2 + b^2 \right]^n = \left(\frac{\alpha}{\gamma_p} \right)^2 > \Omega_n^2.$$

We also get the phase relationship

$$\tan^{-1} \left(\frac{b}{a} \right) + n \tan^{-1} \left(\frac{b}{1 + a} \right) = \pi$$

$$\Rightarrow \frac{b}{1 + a} < \omega^*.$$

Combining these relationships, we get

$$\left[a^2 + \omega^{*2} (1 + a)^2 \right] (1 + \omega^{*2})^n (1 + a)^{2n} > \Omega_n^2.$$

$$\Rightarrow \left[\left(\frac{a}{\omega^*} \right)^2 + (1 + a)^2 \right] (1 + a)^{2n} > 1.$$

Since $a = 0$ at the stability boundary, there must be a regime of parameters sufficiently close to the boundary such that $|a| \ll \omega^*$, for which we have the relationship

$$(1 + a)^{2(n+1)} > 1 \Rightarrow a > 0.$$

This proves the existence of an unstable point when $\sqrt[n+1]{(\theta_1 \theta_2 \prod_{i=1}^{n-1} k_i) / \Omega_n} > \gamma_p$, which implies that all parameters in this regime will yield unstable dynamics (so long as the strong binding assumption still holds).

We note that, though previous results studied the regime of strong binding (β large), the core assumption that was made is that the quantity

$$\frac{\alpha^2 + \beta}{\alpha \gamma_p} \gg 1.$$

We note that there is an entirely different way to achieve this, by making $\alpha^2 \gg \beta, \alpha\gamma_p$. In this regime, all of the previous results follow in almost exactly the same way, except for changes to the constants involved. It is relatively straightforward to show that the characteristic for the system reduces to

$$\sigma(1 + \sigma)^n = -\frac{\beta\gamma_p}{\alpha}.$$

Following the same steps from the previous proofs, we can find that instability now occurs when

$$\sqrt[n-1]{\frac{\Omega_n \theta_1 \theta_2 \prod_{i=1}^{n-1} k_i}{\beta}} = \gamma_p.$$

Interestingly, the stable regime is now

$$\sqrt[n-1]{\frac{\Omega_n \theta_1 \theta_2 \prod_{i=1}^{n-1} k_i}{\beta}} > \gamma_p,$$

the opposite of what occurs in the strong binding limit. One interpretation of these results as a whole is that stability is achievable when either controller binding or process production are individually large, but not when both are large simultaneously.

The Sensitivity Function

The sensitivity function $S(s)$, where s is a complex number, is the transfer function between an input reference to a system and output error [Aström and Murray \(2008\)](#). It is particularly useful to examine $|S(i\omega)|$, which corresponds to the magnitude of S given a purely oscillatory disturbance. If $|S(i\omega)| > 1$, then the system will amplify disturbances at a frequency ω . Conversely, if $|S(i\omega)| < 1$ then the system will attenuate disturbances at frequency ω .

Define $P(s)$ and $C(s)$ to be the transfer function between inputs and outputs of the process and controller, respectively. It is a standard result in control theory that

$$S = \frac{1}{1 + PC}.$$

In general, for a linear system

$$\dot{x} = Ax + Bu$$

$$y = Cx,$$

The transfer function has the form $H(s) = C(sI - A)^{-1}B$. For the antithetic integral feedback system, we have that

$$P(s) = [0, \dots, 1](sI - A)^{-1} \begin{bmatrix} 0 \\ \vdots \\ \theta_1 \end{bmatrix} = \frac{\theta_1 \prod_{i=1}^{n-1} k_i}{(s + \gamma_p)^n},$$

where just as before we use

$$A = \begin{bmatrix} -\gamma_p & 0 & \dots & 0 \\ k_1 & -\gamma_p & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \dots & k_{n-1} & -\gamma_p \end{bmatrix}.$$

Similarly, we have that

$$C(s) = [1, 0](sI - D)^{-1} \begin{bmatrix} 0 \\ \theta_2 \end{bmatrix} = \frac{1}{s} \frac{\theta_2 \frac{\beta}{\alpha}}{[s + (\alpha + \frac{\beta}{\alpha})]},$$

where

$$D = \begin{bmatrix} -\alpha & -\beta/\alpha \\ -\alpha & -\beta/\alpha \end{bmatrix}.$$

Note that $C(s)$ has a factor of $1/s$, indicating that it corresponds to an integrator. From P and C , we see that

$$S(s) = \frac{1}{1 + \frac{\frac{\beta}{\alpha} \theta_1 \theta_2 \prod_{i=1}^{n-1} k_i}{s(s + \gamma_p)^n \left[s + \left(\alpha + \frac{\beta}{\alpha} \right) \right]}} = \frac{s(s + \gamma_p)^n \left[s + \left(\alpha + \frac{\beta}{\alpha} \right) \right]}{s(s + \gamma_p)^n \left[s + \left(\alpha + \frac{\beta}{\alpha} \right) \right] + \beta \gamma_p^n}$$

If we again take the limit $(\alpha + (\beta/\alpha))/\gamma_p \gg 1$ and substitute $s = \gamma_p \sigma$ we get the approximation

$$S(\sigma) \approx \frac{\sigma(1 + \sigma)^n}{\sigma(1 + \sigma)^n + \frac{\alpha}{\gamma_p}}$$

Ideally we would like to analyze $\|S(i\omega)\|_\infty = \max_\omega |S(i\omega)|$, however this is difficult to compute in general. A lower bound for this term can, however, be easily computed by evaluating a particular value of ω close to the maximum. Specifically, we will use $\omega = \tan(\pi/2n)\gamma_p = \omega^* \gamma_p$. At $\sigma = i\omega^*$, we get

$$|S(i\omega^*)| \approx \frac{\omega^*(1 + \omega^{*2})^{\frac{n}{2}}}{\omega^*(1 + \omega^{*2})^{\frac{n}{2}} - \frac{\alpha}{\gamma_p}} = \frac{\Omega_n}{\Omega_n - \frac{\alpha}{\gamma_p}}$$

From our previous results, we know that the system is purely oscillatory when $\Omega_n = \alpha/\gamma_p$, which corresponds to $|S(i\omega^*)| = \|S(i\omega)\|_\infty = \infty$. This gives the intuitive result that the system is infinitely sensitive to a periodic disturbance at $\omega = \omega^* \gamma_p$ when $\Omega_n = \alpha/\gamma_p$. In general, we will have that

$$\|S(i\omega)\|_\infty \geq \frac{\Omega_n}{\Omega_n - \frac{\alpha}{\gamma_p}} \tag{Equation S7}$$

For the special case of $n = 2$, we can explicitly derive an even tighter bound than the one in Equation S7. First, we can explicitly compute

$$\begin{aligned} |S(i\omega^*)| &\approx \frac{|i\omega^*(1 + i\omega^*)^2|}{\left| i\omega^*(1 + i\omega^*)^2 + \frac{\alpha}{\gamma_p} \right|} \\ &= \frac{\omega^*(1 + \omega^{*2})}{\left| \left(\frac{\alpha}{\gamma_p} - 2\omega^{*2} \right) + i\omega^*(1 - \omega^{*2}) \right|} \\ &= \frac{\omega^*(1 + \omega^{*2})}{\sqrt{\left(\frac{\alpha}{\gamma_p} - 2\omega^{*2} \right)^2 + (\omega^*(1 - \omega^{*2}))^2}} \end{aligned}$$

Much of the complexity in this equation comes from the denominator, which can be simplified if we pick ω^* such that either the real or imaginary part is 0. If we plug in $\omega^* = \tan(\pi/4) = 1$, the complex part of the denominator vanishes and we recover the original bound:

$$\|S\|_\infty \geq |S(i)| = \frac{2}{2 - \frac{\alpha}{\gamma_p}} = \frac{1}{1 - \frac{\alpha}{2\gamma_p}}$$

To set the real part to zero, it must be the case that

$$\frac{\alpha}{\gamma_p} - 2\omega^{*2} = 0 \Rightarrow \omega^* = \sqrt{\frac{\alpha}{2\gamma_p}}$$

Plugging this in, we get that

$$\|S\|_\infty \geq \left| S\left(i\sqrt{\frac{\alpha}{2\gamma_p}} \right) \right| = \frac{1 + \frac{\alpha}{2\gamma_p}}{1 - \frac{\alpha}{2\gamma_p}} > \frac{1}{1 - \frac{\alpha}{2\gamma_p}} \tag{Equation S8}$$

We see that this new bound is strictly greater than the one derived in Equation S7, and therefore is a better approximation of $\|S\|_\infty$. While Equation S7 generalizes to all value of n , the latter bound unfortunately requires us to find real roots of order n polynomials, which scales poorly for this problem. We note that these bounds are chosen purely because they are computationally tractable. We were unable to rigorously demonstrate the conditions under which they will be close to $\|S\|_\infty$, however Figure 2 suggest that bound does well, at least in the range of parameters used in our simulations. Empirically, we never observed conditions while developing this paper for which the matching between \mathcal{F} and $\|S\|_\infty$ was particularly worse than what is shown in Figure 2.

Antithetic Integral Feedback with Controller Species Degradation Steady-State Error

Following the same notation as the previous sections, we can model the role of controller degradation as

$$\dot{x}_1 = \theta_1 z_1 - \gamma_p x_1 \quad (\text{Equation S9a})$$

$$\dot{x}_2 = k_1 x_1 - \gamma_p x_2 \quad (\text{Equation S9b})$$

⋮

$$\dot{x}_n = k_{n-1} x_{n-1} - \gamma_p x_n \quad (\text{Equation S9c})$$

$$\dot{z}_1 = \mu - \eta z_1 z_2 - \gamma_c z_1 \quad (\text{Equation S9d})$$

$$\dot{z}_2 = \theta_2 x_n - \eta z_1 z_2 - \gamma_c z_2, \quad (\text{Equation S9e})$$

where γ_p is the degradation of the process species x_i and γ_c is the degradation rate of the control species z_1 and z_2 . At a high level we will proceed much in the same way as we did previously, however we will see that nonzero controller degradation leads to several technical challenges that do not appear when $\gamma_c = 0$. The first of these arises from simply solving for the steady values around which we will linearize the model. Where previously we used the fact that $\dot{z}_1 - \dot{z}_2 = 0 \Rightarrow x_n^* = \mu / \theta_2$, where $*$ denotes a steady-state value. To subsequently solve for all other steady-state concentrations, we are now left with the messier relationship

$$\dot{z}_1 - \dot{z}_2 = 0 \Rightarrow x_n^* = \frac{\mu}{\theta_2} - \frac{\gamma_c}{\theta_2} (z_1^* - z_2^*).$$

This implies that, for $\gamma_c > 0$, we expect x_n to differ from the desired steady-state μ / θ by some error that depends on the values of z_1^* and z_2^* . Since this error is almost surely a function of many other parameters, we essentially lose the robust precise adaptation property where x_n^* is completely independent of the network's parameters. We will first calculate a general form for x_n^* , then derive a large η limit that makes further calculations tractable.

To begin, we use Equation S9d and S9e to derive the relationships

$$\mu = z_1^* (\eta z_2^* + \gamma_c) \Rightarrow z_2^* = \frac{1}{\eta} \left(\frac{\mu}{z_1^*} - \gamma_c \right)$$

$$x_n^* = \frac{1}{\theta_2} z_2^* (\eta z_1^* + \gamma_c).$$

Combining these equations, we find that

$$x_n^* = \frac{\mu}{\theta_2} + \frac{\gamma_c \mu}{\eta \theta_2} \frac{1}{z_1^*} - \frac{\gamma_c}{\theta_2} z_1^* - \frac{\gamma_c^2}{\eta \theta_2}.$$

Finally, we observe that

$$z_1^* = \frac{\gamma_p^n}{\theta_1 \prod_i k_i} x_n^* = \frac{\theta_2}{\alpha} x_n^*,$$

which yields the relationship

$$x_n^* = \frac{\mu}{\theta_2} + \frac{\gamma_c \mu}{\eta \alpha} \frac{1}{x_n^*} - \frac{\gamma_c x_n^*}{\alpha} - \frac{\gamma_c^2}{\eta \theta_2},$$

$$\Rightarrow \left(1 + \frac{\gamma_c}{\alpha}\right) x_n^{*2} = \left(\frac{\mu}{\theta_2} - \frac{\gamma_c^2}{\eta \theta_2}\right) x_n^* + \frac{\gamma_c \mu}{\eta \alpha}. \quad (\text{Equation S10})$$

While this quadratic can be solved explicitly, the result can be greatly simplified by again taking the limit of large η . Here the sense in which we take this limit is such that $\mu/\theta_2 \gg \gamma_c^2/\eta\theta_2$ and $1 \gg \gamma_c\mu/\eta\alpha$. These reduce to the condition

$$\eta \gg \frac{\gamma_c^2}{\mu}, \frac{\gamma_c \mu}{\alpha}.$$

Combined with the previous assumption about the size of η we now have a large number of conditions to fulfill, however we find that in practice we rarely are in parameter regimes where a great deal of tuning needs to be done to satisfy everything. That being said, we can use this limit to reduce [Equation S10](#) to

$$\left(1 + \frac{\gamma_c}{\alpha}\right) x_n^{*2} = \frac{\mu}{\theta_2} x_n^* \Rightarrow x_n^* \approx \frac{\mu}{\theta_2} \frac{1}{1 + \frac{\gamma_c}{\alpha}}. \quad (\text{Equation S11})$$

Using the same approximation, we can also compute

$$z_1^* = \frac{\theta_2}{\alpha} x_n^* \approx \frac{\mu}{\alpha + \gamma_c}, \quad (\text{Equation S12})$$

$$z_2^* = \frac{\theta_2}{\eta z_1^* + \gamma_c} x_n^* \approx \frac{\alpha}{\eta}. \quad (\text{Equation S13})$$

These will be useful for computing the linearized dynamics of the system in the next section.

As a sanity check, we can immediately see that $x_n^* = \mu/\theta_2$ when $\gamma_c = 0$, as expected. For $\gamma_c > 0$, [Equation S11](#) captures the steady-state error relative to the set point μ/θ induced by nonzero controller degradation. We see that, so long as the ratio $\gamma_c/\alpha \ll 1$, error will be negligible. What is unclear at this point is under what conditions this can be achieved while still ensuring stability of the overall system. To this end, we will now characterize stability and performance for $\gamma_c > 0$.

Linearized Dynamics with Controller Degradation

Here we present results analogous to those in [The Stability Criterion](#), omitting detailed proofs since the structure of the argument from this point on is essentially identical to what was shown in the previous section. Because the only nonlinear terms in our system are in [Equation S9d](#) and [S9e](#), the only matrix to change in our linearization from [The Stability Criterion](#) is

$$D = \begin{bmatrix} -\eta z_2^* - \gamma_c & -\eta z_1^* \\ -\eta z_2^* & -\eta z_1^* - \gamma_c \end{bmatrix} \approx \begin{bmatrix} -\alpha - \gamma_c & -\beta/(\alpha + \gamma_c) \\ -\alpha & -\beta/(\alpha + \gamma_c) - \gamma_c \end{bmatrix}.$$

Using this D matrix and proceeding with precisely the same calculation as before, we can derive the characteristic equation for the system:

$$(s + \gamma_p)^n (s + \gamma_c) \left[s + \gamma_c + \alpha + \frac{\beta}{\alpha + \gamma_c} \right] + \gamma_p^n \beta \frac{\alpha}{\alpha + \gamma_c} = 0. \quad (\text{Equation S14})$$

We again take the appropriate limit of $\beta \gg (\gamma_c + \alpha) \gamma_p$, $(\gamma_c + \alpha)^2$ to follow the same argument as in [The Stability Criterion](#) to get the simplified expression in terms of $\sigma = s/\gamma_p$:

$$(1 + \sigma)^n \left(\frac{\gamma_c}{\gamma_p} + \sigma \right) = -\frac{\alpha}{\gamma_p}. \quad (\text{Equation S15})$$

First we note that, when $\gamma_c = 0$, we recover [Equation S6](#) as expected. Proceeding as before, we can write the characteristic polynomial in terms of phase and magnitude constraints for $\sigma = i\omega^*$:

$$(1 + \omega^{*2})^{\frac{n}{2}} \sqrt{\frac{\gamma_c^2}{\gamma_p^2} + \omega^{*2}} = \frac{\alpha}{\gamma_p} \quad (\text{Equation S16})$$

$$n \tan^{-1}(\omega^*) + \tan^{-1}\left(\frac{\gamma_p \omega^*}{\gamma_c}\right) = \pi. \quad (\text{Equation S17})$$

Stability Analysis with Controller Degradation

Unfortunately, the additional complexity in Equation S17 makes it challenging to write down the sort of explicit closed-form expressions for stability seen in Theorem S3. While we can write out explicit stability conditions for $n = 2$, we will need to study particular parameter regimes for $n > 2$ as the summation relationship for \tan^{-1} scales poorly.

To solve for ω^* in Equation S17 we make use of the inverse trigonometric identity

$$\tan^{-1}(a) + \tan^{-1}(b) = \tan^{-1}\left(\frac{a+b}{1-ab}\right) \pmod{\pi}.$$

Applying this identity twice yields the relationship

$$\begin{aligned} 2 \tan^{-1}(\omega^*) + \tan^{-1}\left(\frac{\gamma_p \omega^*}{\gamma_c}\right) &= \pi \\ \Rightarrow \tan^{-1}\left(\frac{2\omega^*}{1-\omega^{*2}}\right) + \tan^{-1}\left(\frac{\gamma_p \omega^*}{\gamma_c}\right) &= 0 \pmod{\pi} \\ \Rightarrow \tan^{-1}\left(\frac{2\omega^* + \frac{\gamma_p \omega^*}{\gamma_c}(1-\omega^{*2})}{1 - \left(1 + \frac{\gamma_p}{\gamma_c}\right)\omega^{*2}}\right) &= 0 \pmod{\pi}. \end{aligned}$$

Since the only value for which $\tan^{-1}(x) = 0$ is $x = 0$, the problem reduces to solving the equation

$$\begin{aligned} 2\omega^* + \frac{\gamma_p}{\gamma_c}\omega^*(1-\omega^{*2}) &= 0 \\ \Rightarrow 2 + \frac{\gamma_p}{\gamma_c}(1-\omega^{*2}) &= 0 \\ \Rightarrow \omega^* &= \sqrt{2\frac{\gamma_c}{\gamma_p} + 1}. \end{aligned}$$

Combining this with Equation S16 yields the stability criterion

$$\frac{\theta_1 \theta_2 k}{2} < \gamma_p (\gamma_c + \gamma_p)^2. \quad (\text{Equation S18})$$

If we assume that we have full freedom to set control parameters, then Equation S18 that it is possible to make the production rates θ_1 and θ_2 large, so long as there is a compensatory increase in γ_c . This implies that we can, in a sense, sidestep the performance tradeoffs between speed and stability if we are willing pay a price in terms of *efficiency*, measured by the turnover rates of z_1 and z_2 .

Next we will study what happens when $n > 2$. We note that there is an interesting topological distinction going from $n = 2$ to $n > 2$ which yields qualitatively different stability results. To see why this is the case, we return to Equation S17:

$$n \tan^{-1}(\omega^*) + \tan^{-1}\left(\frac{\gamma_p \omega^*}{\gamma_c}\right) = \pi.$$

Recall that $\tan^{-1}(x) < \pi/2$ for all x . Because this is the case, when $n = 2$ it is always the case that $2 \tan^{-1}(\omega^*) < \pi$, implying that satisfying the phase condition is strictly contingent of the value of the term $\tan^{-1}((\gamma_p/\gamma_c)\omega^*)$. On the other hand, for $n > 2$, there exist values of ω^* such that $n \tan^{-1}(\omega^*) \geq \pi$, so depending on the relative magnitude of the ratio γ_p/γ_c satisfying the phase condition may or may not depend strongly on γ_c .

If we look again at Equation S15:

$$(1 + \sigma)^n \left(\frac{\gamma_c + \sigma}{\gamma_p}\right) = -\frac{\alpha}{\gamma_p},$$

we notice that the only place in which γ_c appears is in the ratio γ_c/γ_p . One natural approach to studying the solutions to this equation is to examine what happens at various limits, namely $\gamma_c \ll \gamma_p$, $\gamma_c = \gamma_p$, and $\gamma_c \gg \gamma_p$. Here we will present results without going into formal detail, however the analysis can be made rigorous by analyzing Equation S17.

Case 1 $\gamma_c \ll \gamma_p$. This case is fairly straightforward, as it is it reduces to the case of no controller degradation. We recover the characteristic polynomial

$$\sigma(1 + \sigma)^n = -\frac{\alpha}{\gamma_p},$$

which has the same exact stability condition as in Theorem S3.

Case II $\gamma_c = \gamma_p$. This case is representative of what happens when controller and process degradation have the same order of magnitude. We use $\gamma_p/\gamma_c = 1$ in Equation S17 to find that the stability boundary is characterized by

$$\begin{aligned} n \tan^{-1}(\omega^*) + \tan^{-1}\left(\frac{\gamma_p \omega^*}{\gamma_c}\right) &= \pi \\ \xrightarrow{\gamma_c = \gamma_p} (n + 1) \tan^{-1}(\omega^*) &= \pi \\ \Rightarrow \omega^* &= \tan\left(\frac{\pi}{n + 1}\right). \end{aligned}$$

Here it is useful to define the quantity

$$\tilde{\Omega}_n = \left(1 + \tan\left(\frac{\pi}{n}\right)^2\right)^{\frac{n}{2}},$$

where $\tilde{\Omega}_n$ differs from the previously defined Ω_n by a factor of 1/2 in the argument of the tangent term. Using this expression, we can use Equation S16 to derive the stability criterion

$$\frac{\alpha}{\gamma_p} < \tilde{\Omega}_{n+1} \Rightarrow \sqrt[n+1]{\frac{\theta_1 \theta_2 \prod_i k_i}{\tilde{\Omega}_{n+1}}} < \gamma_p.$$

This condition is qualitatively the same as the one in Theorem S3 up to a constant difference accounted for by the $\tilde{\Omega}_{n+1}$ term.

Case III $\gamma_c \gg \gamma_p$. Following a similar line of reasoning as in the previous case, taking the limit $\gamma_p/\gamma_c \ll 1$ in Equation S17 to show that

$$\begin{aligned} n \tan^{-1}(\omega^*) + \tan^{-1}\left(\frac{\gamma_p \omega^*}{\gamma_c}\right) &= \pi \\ \xrightarrow{\gamma_c \gg \gamma_p} n \tan^{-1}(\omega^*) &= \pi \\ \Rightarrow \omega^* &= \tan\left(\frac{\pi}{n}\right). \end{aligned}$$

We again use Equation S16 to find that the stability boundary is set by the following relationship:

$$\begin{aligned} (1 + \omega^{*2})^{\frac{n}{2}} \sqrt{\frac{\gamma_c^2}{\gamma_p^2} + \omega^{*2}} &= \frac{\alpha}{\gamma_p} \\ \xrightarrow{\gamma_c \gg \gamma_p} \tilde{\Omega}_n \frac{\gamma_c}{\gamma_p} &= \frac{\alpha}{\gamma_p} \\ \Rightarrow \tilde{\Omega}_n &= \frac{\alpha}{\gamma_c}. \end{aligned}$$

This implies that the stability criterion for this case is

$$\frac{\alpha}{\gamma_c} < \tilde{\Omega}_n \Rightarrow \sqrt[n]{\frac{\theta_1 \theta_2 \prod_i k_i}{\gamma_c \tilde{\Omega}_n}} < \gamma_p.$$

Notice that in the $n = 2$ case, ω^* is a function of γ_c and the subsequent stability criterion depends on the term $(\gamma_p + \gamma_c)^2$. This is quite different from the $n > 2$ cases where in each regime, the γ_c dependence in ω^* disappears. Similarly, in the stability criterion we see a linear (rather than quadratic) dependence on γ_c . This is a direct result of the previously mentioned topological difference between the $n = 2$ and $n > 2$ cases.

One interesting side effect of this results is that, when the system is purely oscillatory (on the stability boundary), the frequencies of oscillation may be dramatically different depending on n . Consider the case where $\gamma_c \gg \gamma_p$. If $n = 2$, this frequency will be

$$\omega = \gamma_p \omega^* = \gamma_p \sqrt{2 \frac{\gamma_c}{\gamma_p} + 1} \approx \sqrt{2 \gamma_p \gamma_c}.$$

If $n > 2$, we use the results from Case III above to find

$$\omega = \gamma_p \omega^* \approx \gamma_p \tan\left(\frac{\pi}{n}\right).$$

In the former case, ω scales with $\sqrt{\gamma_p}$, whereas in the latter case ω is independent of γ_c . This implies that for large controller degradation rates we would expect much faster oscillatory modes for $n=2$ than for $n>2$.

The Effects of Degradation on Sensitivity and Performance

Just as in [The Sensitivity Function](#), we can write the generic sensitivity function for the linearized antithetic integral feedback system with degradation in terms of the variable $\sigma = \gamma_p s$ as

$$S(\sigma) = \frac{(1 + \sigma)^n \left(\frac{\gamma_c}{\gamma_p} + \sigma \right)}{(1 + \sigma)^n \left(\frac{\gamma_c}{\gamma_p} + \sigma \right) + \frac{\alpha}{\gamma_p}}. \quad (\text{Equation S19})$$

For the case $n = 2$, we can again derive an explicit lower bound for $\|S(i\omega^*)\|_\infty$:

$$\begin{aligned} |S(i\omega^*)| &= \frac{\left| (1 + i\omega^*)^2 \left(\frac{\gamma_c}{\gamma_p} + i\omega^* \right) \right|}{\left| (1 + i\omega^*)^2 \left(\frac{\gamma_c}{\gamma_p} + i\omega^* \right) + \frac{\alpha}{\gamma_p} \right|}, \\ &= \frac{(1 + \omega^{*2}) \sqrt{\frac{\gamma_c^2}{\gamma_p^2} + \omega^{*2}}}{\left| \left(\frac{\gamma_c}{\gamma_p} (1 - \omega^{*2}) - 2\omega^{*2} \right) + i \left(2\omega^* \frac{\gamma_c}{\gamma_p} + \omega^* (1 - \omega^{*2}) \right) \right|}. \end{aligned} \quad (\text{Equation S20})$$

We can again solve for ω^* such that the real part of the denominator is zero:

$$\frac{\gamma_c}{\gamma_p} (1 - \omega^{*2}) - 2\omega^{*2} = 0 \Rightarrow \omega^* = \sqrt{\frac{\alpha + \gamma_c}{2\gamma_p + \gamma_c}}.$$

If we evaluate [Equation S20](#) at ω^* , we can write the bound

$$\|S\|_\infty > |S(i\omega^*)| = \mathcal{F} = \frac{(1 + \omega^{*2}) \sqrt{1 + \left(\frac{\gamma_c}{\omega^* \gamma_p} \right)^2}}{(1 - \omega^{*2}) + 2 \frac{\gamma_c}{\gamma_p}}. \quad (\text{Equation S21})$$

It is easy to check that, for $\gamma_c = 0$, we recover the bound from [Equation S8](#). As ω^* approaches $1 + 2\gamma_c/\gamma_p$, $\|S\|_\infty$ will asymptotically increase to ∞ . Alternatively, increasing γ_c will decrease sensitivity, and consequently improve robustness. We can think of γ_c as capturing the inefficiency of our controller (higher degradation means the control species are degraded before being used in an antithetic feedback reaction). In these terms, we see that increasing γ_c will reduce \mathcal{F} at the cost of increased steady-state error (see [Figures 2D–2F](#)). If we hold ε constant by varying both γ_c and θ_1 , we can decrease \mathcal{F} at the cost of on decreasing efficiency of the controller (see [Figures 2G–2I](#)). Finally, we can vary γ_c and θ_1 such that \mathcal{F} is constant, which leads to a tradeoff between steady-state error and efficiency (see [Figures 2J–2L](#)).

Limitations of Linear Analysis

To demonstrate the limits of the linearized theory presented throughout the paper, we will briefly describe a circuit architecture that uses antithetic integral feedback, yet for which there exist parameters such that linearization does not yield practically useful predictions of stability. This circuit still has two process species X_1 and X_2 and two control species Z_1 and Z_2 , however the process will follow a new set of dynamics:

$$\dot{X}_1 = \phi_1 - \theta_1 Z_1 X_1, \quad (\text{Equation S22a})$$

$$\dot{X}_2 = \phi_2 - k X_1 X_2, \quad (\text{Equation S22b})$$

$$\dot{Z}_1 = \mu - \eta Z_1 Z_2, \quad (\text{Equation S22c})$$

$$\dot{Z}_2 = \theta_2 X_2 - \eta Z_1 Z_2. \quad (\text{Equation S22d})$$

We see that now, rather than X_1 and X_2 having linear dynamics as in Equation 1, they interact via bimolecular degradation, as would be the case if Z_1 and X_1 were proteases respectively targeting X_1 and X_2 (possibly using the mechanism described in Gao et al. (2018)). Solving for steady-state values yields:

$$x_1^* = \frac{\phi_2 \theta_2}{\mu k}, x_2^* = \frac{\mu}{\theta_2}, z_1^* = \frac{\phi_1 \mu k}{\phi_2 \theta_1 \theta_2}, z_2^* = \frac{\phi_2 \theta_1 \theta_2}{\eta \phi_1 k}, \quad (\text{Equation S23})$$

and the corresponding linearized dynamics are

$$\dot{\mathbf{x}} = M\mathbf{x},$$

$$M = \begin{bmatrix} -\theta_1 z_1^* & 0 & -\theta_1 x_1^* & 0 \\ -k x_2^* & -k x_1^* & 0 & 0 \\ 0 & 0 & -\eta z_2^* & -\eta z_1^* \\ 0 & \theta_2 & -\eta z_2^* & -\eta z_1^* \end{bmatrix}.$$

since these linearized dynamics have the same essential structure as those of Equation 3 with the addition of heterogeneous degradation rates $\theta_1 z_1^*$ and $k x_1^*$, we can use the same methods as in The Stability Criterion to find that the associated linearized stability criterion for Equation S22 is

$$\begin{aligned} \theta_2 (\theta_1 x_1^*) (k x_2^*) &< (\theta_1 z_1^*) (k x_1^*) [\theta_1 z_1^* + k x_1^*] \\ \Rightarrow \frac{\phi_2}{\phi_1} \frac{\theta_1 \theta_2}{k} &< \frac{\phi_1}{\phi_2} \frac{\mu k}{\theta_2} + \frac{\phi_2 \theta_2}{\mu}. \end{aligned} \quad (\text{Equation S24})$$

This local stability relationship is harder to interpret than the one presented in Equation 4, and does not yield the same clean separation of parameters. This is less due to the structure of the dynamics, and more because of the complex way that the steady-state values depend on parameters. Figure S1 shows that, for a range of parameters, the behavior of the nonlinear circuit in Equation S22 is well characterized by Equation S24 when the initial conditions are near the equilibrium determined by Equation S23.

What is more interesting is that we find that there exist parameters for which Equation S24 fails to capture global stability. Equation S24 successfully predicts stability in Figure S2A–S2C, where X_2 converges to its desired reference point. There are, however, parameters for which Equation S24 is satisfied, but simulations show clear limit cycle behavior (see Figure S2D) if the initial conditions are sufficiently far from equilibrium. This implies that, for these parameters and initial conditions, the limit cycle is a strong attractor and the local point to which we want the system to adapt, $x_2^* = \mu/\theta_2$, has a sufficiently small basin of attraction that initial transients are enough to push the system into the regime where the limit cycle dominates. More concisely, it appears that for the system described in Equation S22, Equation S24 is a necessary but not sufficient condition for global stability.

While we were able to get some sense via simulation for when Equation S24 is a good predictor of global stability, we hope that future work in this area will be able to produce a constructive, unifying theory that is sufficient to explain these phenomena precisely.

Controlling an Unstable Process

In all prior sections, we have assumed that the underlying process being controlled is open-loop stable. Here we will examine a simple model of an open-loop unstable process and describe which control architectures are capable of stabilizing the closed-loop system. To start, we will use a simple linear system as our process:

$$\dot{x}_1 = k_2 x_2 - \gamma_p x_1,$$

$$\dot{x}_2 = k_1 x_1 - \gamma_p x_2.$$

This system will be unstable when at least one eigenvalue of the matrix:

$$A = \begin{bmatrix} -\gamma_p & k_2 \\ k_1 & -\gamma_p \end{bmatrix}$$

has positive real part. With some straightforward linear algebra we can find that the eigenvalues of A are

$$\lambda_{\pm} = -\gamma_p \pm \sqrt{k_1 k_2}.$$

Because $k_1, k_2, \gamma_p > 0$, we know that $\lambda_- < 0$ for all parameters. λ_+ , however, can be either positive or negative. In particular,

$$\sqrt{k_1 k_2} > \gamma_p \Leftrightarrow \lambda_+ > 0.$$

To facilitate our study of unstable processes, we will assume $\sqrt{k_1 k_2} > \gamma_p$ for the rest of the section. One immediate difference is that, due to the unstable process, any controller must now interact with the process via repression rather than activation. To model this, we will first study the following architecture (described in Equation 18 and the left panels of Figure 4):

$$\dot{x}_1 = \frac{k_2 x_2}{1 + \theta_1 z_1} - \gamma_p x_1, \quad (\text{Equation S25a})$$

$$\dot{x}_2 = k_1 x_1 - \gamma_p x_2, \quad (\text{Equation S25b})$$

$$\dot{z}_1 = \mu_1 - \eta z_1 z_2, \quad (\text{Equation S25c})$$

$$\dot{z}_2 = \frac{\mu_2}{1 + \theta_2 x_2} - \eta z_1 z_2. \quad (\text{Equation S25d})$$

If $\theta_1 = \theta_2 = 0$, then this architecture reduces to the open-loop system described above. The controller topology is essentially the same as in the stable case, with the core difference that z_1 represses x_1 and x_2 represses z_2 , where before these interactions were activating. Since now there is no reaction synthesizing z_2 , we must add in some external production rate μ_2 . We will again proceed by solving for the steady-state concentrations of each species and linearizing around these values. The steady-state concentrations are as follows:

$$\begin{bmatrix} x_1^* \\ x_2^* \\ z_1^* \\ z_2^* \end{bmatrix} = \begin{bmatrix} \gamma_p & \mu_2 - \mu_1 & 1 & \mu_2 - \mu_1 & k_1 k_2 - \gamma_p^2 & \mu_1 & \theta_1 \gamma_p^2 \\ k_1 \theta_2 & \mu_1 & \theta_2 & \mu_1 & \theta_1 \gamma_p^2 & \eta & k_1 k_2 - \gamma_p^2 \end{bmatrix}^T. \quad (\text{Equation S26})$$

If we now linearize around this fixed point, we can define a new set of parameters:

$$\begin{aligned} \hat{k}_2 &= \frac{d}{dx_2} \left(\frac{k_2 x_2}{1 + \theta_1 z_1} \right)_{z_1^*} = \frac{\gamma_p^2}{k_1}, \\ \hat{\theta}_1 &= \left| \frac{d}{dz_1} \left(\frac{k_2 x_2}{1 + \theta_1 z_1} \right)_{x_2^*, z_1^*} \right| = \frac{\theta_1 \gamma_p^4}{\theta_2 k_1^2 k_2} \frac{\mu_2 - \mu_1}{\mu_1}, \\ \hat{\theta}_2 &= \left| \frac{d}{dx_2} \left(\frac{\mu_2}{1 + \theta_2 x_2} \right)_{x_2^*} \right| = \theta_2 \frac{\mu_1^2}{\mu_2}, \\ \alpha &= \eta z_2^* = \frac{\mu_1 \theta_1 \gamma_p^2}{k_1 k_2 - \gamma_p^2}, \\ \beta &= \eta \mu_1, \end{aligned}$$

which characterize the linearized set of dynamics:

$$\dot{\mathbf{x}} = M \mathbf{x},$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ z_1 \\ z_2 \end{bmatrix}, M = \begin{bmatrix} -\gamma_p & \hat{k}_2 & -\hat{\theta}_1 & 0 \\ k_1 & -\gamma_p & 0 & 0 \\ 0 & 0 & -\alpha & -\beta/\alpha \\ 0 & -\hat{\theta}_2 & -\alpha & -\beta/\alpha \end{bmatrix}.$$

Following the same methods in [The Stability Criterion](#), we can derive the characteristic polynomial for M ,

$$p(s) = s(s - \lambda_+)(s - \lambda_-) \left(s + \alpha + \frac{\beta}{\alpha} \right) + \frac{\beta}{\alpha} \hat{\theta}_1 \hat{\theta}_2 k_1,$$

where we now use $\lambda_{\pm} = -\gamma_p \pm \sqrt{k_1 \hat{k}_2}$. If we plug in \hat{k}_2 , we see that $\lambda_+ = 0$ and $\lambda_- = -2\gamma_p$. The fact that the process's eigenvalues change when comparing the open- and closed-loop systems is a byproduct of the fact that our original model was nonlinear, and is something that would not occur for a strictly linear system.

Again taking the limit of strong binding, which here takes the form $\beta \gg \alpha^2, 2\alpha\gamma_p$, and setting $p(s) = 0$, we get the equation

$$s^2(s + 2\gamma_p) = -\hat{\theta}_1\hat{\theta}_2k_1. \quad (\text{Equation S27})$$

Theorem S4

Equation S25 is intrinsically unstable, so long as $k_1k_2 > \gamma_p$.

Proof. The corresponding phase constraint for Equation S27 when $s = i\omega$ is:

$$\tan^{-1}\left(\frac{\omega}{2\gamma_p}\right) = 0 \Rightarrow \omega = 0.$$

First, we note that, because the right-hand side of Equation S27 has all positive coefficients, it is not possible for the equation to be satisfied by a value of s that is both real and positive. Further, because the equation is monotone, there can be at most most real and negative solutions. This means that there are no parameters such that the linearized system transitions from being stable to unstable or vice versa, since the system's complex conjugate solutions would need to pass through a purely oscillatory mode with $\omega > 0$ for this to occur. This implies that the system is either always stable or always unstable, independent of the particular values of parameters.

From this, we can prove that there are no solutions to Equation S27 with $\text{Re}(s) < 0$ by simply proving the existence of a single such solution. If we take $s = a + ib$, $a < 0$, $b > 0$ (we can assume $b > 0$ without loss of generality because solutions come in conjugate pairs), the phase constraint for Equation S27 has the form

$$2\arg(s) + \arg(s + 2\gamma_p) = \pi$$

Since $\pi > \arg(s) > \pi/2$ and $\pi > \arg(s + 2\gamma_p) > 0$, it must be the case that

$$3\pi > 2\arg(s) + \arg(s + 2\gamma_p) > \pi.$$

Consequently, there is no value of γ_p for which the phase constraint can be satisfied by s with $a < 0$. By contradiction, it must then be the case that all complex conjugate solutions to Equation S27 have $\text{Re}(s) > 0$, and the system is therefore the system is intrinsically unstable. It is also relatively easy to check that, if $a > 0$, the corresponding inequalities are satisfiable (though explicitly constructing the solutions is not straightforward).

Next, we consider an alternative architecture, shown in Figure 4 and Equation 19. This system is described by the same dynamics as before, except we now have z_1 directly repressing x_2 (rather than indirectly doing so via x_1):

$$\dot{x}_1 = k_2x_2 - \gamma_px_1, \quad (\text{Equation S28a})$$

$$\dot{x}_2 = \frac{k_1x_1}{1 + \theta_1z_1} - \gamma_px_2, \quad (\text{Equation S28b})$$

$$\dot{z}_1 = \mu_1 - \eta z_1z_2, \quad (\text{Equation S28c})$$

$$\dot{z}_2 = \frac{\mu_2}{1 + \theta_2x_2} - \eta z_1z_2. \quad (\text{Equation S28d})$$

The steady-state values are almost identical to those of Equation S26, except we now have that

$$x_1^* = \frac{k_2}{\gamma_p\theta_2} \frac{\mu_2 - \mu_1}{\mu_1}.$$

We can define another set of linearized parameters,

$$\hat{k}_1 = \frac{d}{dx_1} \left(\frac{k_1x_1}{1 + \theta_1z_1} \right)_{z_1^*} = \frac{\gamma_p^2}{k_2},$$

$$\hat{\theta}_1 = \left| \frac{d}{dz_1} \left(\frac{k_1x_1}{1 + \theta_1z_1} \right)_{x_1^*, z_1^*} \right| = \frac{\theta_1}{\theta_2} \frac{\gamma_p^3}{k_1k_2} \frac{\mu_2 - \mu_1}{\mu_1},$$

with $\hat{\theta}_2$, α , and β the same as before. Our linearized dynamics are now described by the matrix

$$M = \begin{bmatrix} -\gamma_p & k_2 & 0 & 0 \\ \widehat{k}_1 & -\gamma_p & -\widehat{\theta}_1 & 0 \\ 0 & 0 & -\alpha & -\beta/\alpha \\ 0 & -\widehat{\theta}_2 & -\alpha & -\beta/\alpha \end{bmatrix},$$

with a corresponding characteristic polynomial

$$\rho(s) = s(s - \lambda_+)(s - \lambda_-) \left(s + \alpha + \frac{\beta}{\alpha} \right) + \frac{\beta}{\alpha} \widehat{\theta}_1 \widehat{\theta}_2 (s + \gamma_p),$$

with $\lambda_{\pm} = -\gamma_p \pm \sqrt{\widehat{k}_1 k_2}$. The limiting form of the characteristic equation is now

$$s^2(s + 2\gamma_p) = -\widehat{\theta}_1 \widehat{\theta}_2 (s + \gamma_p). \quad (\text{Equation S29})$$

Theorem S5

Equation S28 is intrinsically stable, so long as $k_1 k_2 > \gamma_p$ and non-negative steady-state values exist for each species.

Proof. The structure of this proof will follow the same structure as Theorem S4, except that the conclusion will be reversed because of the architectural differences between this circuit and the previous one. In particular, if we examine the phase constraint corresponding to Equation S29, we get

$$2\arg(s) + \arg(s + 2\gamma_p) = \pi + \arg(s + \gamma_p).$$

Assume that $s = a + ib, a, b > 0$, corresponding to the existence of unstable dynamics. In this case, this yields the constraint

$$2 \tan^{-1}\left(\frac{b}{a}\right) + \tan^{-1}\left(\frac{b}{a + 2\gamma_p}\right) = \pi + \tan^{-1}\left(\frac{b}{a + \gamma_p}\right).$$

However, we can also observe both that $2 \tan^{-1}(b/a) < \pi$ and $\tan^{-1}(b/(a + 2\gamma_p)) < \tan^{-1}(b/(a + \gamma_p))$ for all $a, b > 0$. Because of this, there is no way for the sum of the left-hand side terms to ever equal the sum of the right-hand terms, therefore there cannot be a solution to Equation S29 with $\text{Re}(s) > 0$. This implies that Equation S28 is intrinsically stable whenever non-negative steady-state values exist.

Finally, we can find the sensitivity function for the stabilizing architecture. This is somewhat complicated by the fact that the process transfer function varies with control parameters, so it is difficult to separate the process and the controller transfer functions. However we can use a convenient form

$$S(s) = \frac{p_{ol}(s)}{p_{cl}(s)},$$

where $p_{cl}(s)$ is the characteristic equations for the closed-loop systems, and $p_{ol}(s) = \lim_{\theta_i \rightarrow 0} p_{cl}(s)$ (Chandra et al. (2009, 2011)). Using

$$p_{cl}(s) = s^2(s + 2\gamma_p) \left(s + \alpha + \frac{\beta}{\alpha} \right) + \frac{\beta}{\alpha} \widehat{\theta}_1 \widehat{\theta}_2 (s + \gamma_p),$$

we get the sensitivity function (assuming large β):

$$S(s) = \frac{s(s + \gamma_p - \sqrt{k_1 k_2})(s + \gamma_p + \sqrt{k_1 k_2})}{s^2(s + 2\gamma_p) + \widehat{\theta}_1 \widehat{\theta}_2 (s + \gamma_p)}.$$

Note that it is important that we take care with the limits, as the roots of $p_{ol}(s)$ should reflect the eigenvalues of the unstable open-loop system. This is used to generate the right-hand plots in Figures 4A and 4B.

DATA AND CODE AVAILABILITY

All figures and simulations for this paper were generated in Python. A Jupyter notebook has been provided as Data S1, each of which produces the elements of a corresponding figure in our paper. While some annotations and cartoons were added using image editing software, wherever possible we tried to make the generation of plots entirely self-contained. The only Python modules that need to be imported are matplotlib, scipy, and numpy. An optional module, jupyterthemes, is used purely for stylistic purposes and can be commented out if desired. The only difference will be a change in the color scheme for plots.