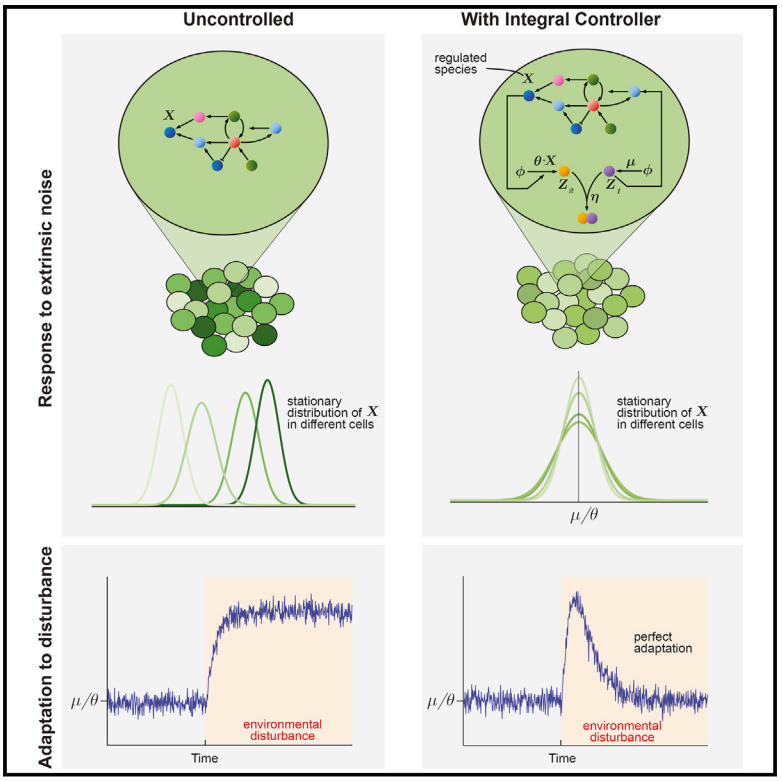
Article

Cell Systems

Antithetic Integral Feedback Ensures Robust Perfect Adaptation in Noisy Biomolecular Networks

Graphical Abstract



Authors

Corentin Briat, Ankit Gupta, Mustafa Khammash

Correspondence

mustafa.khammash@bsse.ethz.ch

In Brief

A novel mechanism for achieving cellular regulation in the presence of biochemical noise and environmental disturbances is proposed and studied.

Highlights

- Homeostatic schemes in noisy cellular conditions remain poorly understood
- A novel integral feedback motif for regulation at the molecular level is presented
- It functions robustly in presence of biochemical noise and environmental disturbances
- It elucidates endogenous regulatory circuits and motivates novel synthetic ones

Briat et al., 2016, Cell Systems 2, 15–26
January 27, 2016
http://dx.doi.org/10.1016/j.cels.2016.01.004

Cell Systems
Article

Antithetic Integral Feedback Ensures Robust Perfect Adaptation in Noisy Biomolecular Networks

Corentin Briat,[1,2] Ankit Gupta,[1,2] and Mustafa Khammash[1,*]
[1]Department of Biosystems Science and Engineering (D-BSSE), ETH–Zürich, Mattenstrasse 26, 4058 Basel, Switzerland
[2]Co-first author
*Correspondence: mustafa.khammash@bsse.ethz.ch
http://dx.doi.org/10.1016/j.cels.2016.01.004

SUMMARY

The ability to adapt to stimuli is a defining feature of many biological systems and critical to maintaining homeostasis. While it is well appreciated that negative feedback can be used to achieve homeostasis when networks behave deterministically, the effect of noise on their regulatory function is not understood. Here, we combine probability and control theory to develop a theory of biological regulation that explicitly takes into account the noisy nature of biochemical reactions. We introduce tools for the analysis and design of robust homeostatic circuits and propose a new regulation motif, which we call antithetic integral feedback. This motif exploits stochastic noise, allowing it to achieve precise regulation in scenarios where similar deterministic regulation fails. Specifically, antithetic integral feedback preserves the stability of the overall network, steers the population of any regulated species to a desired set point, and adapts perfectly. We suggest that this motif may be prevalent in endogenous biological circuits and useful when creating synthetic circuits.

INTRODUCTION

Perfect adaptation is that property of a biological system (e.g., a cell) that enables it to adapt after an external stimulus and maintain responsiveness to further stimuli. This is usually accomplished by returning to a pre-established baseline; a new balance of components or activities is established within the system to maintain that baseline in the presence of the stimulus. Such a balancing tendency is called homeostasis. The ability to maintain homeostasis is a defining feature of many biological systems (e.g., the maintenance of plasma calcium concentration in mammals, the level of excitability in neurons, and internal pressure in various cells). While the mechanisms used to maintain homeostasis are often uncharacterized and likely diverse, they must share certain properties if adaptation is to be perfect. Notably, to be effective, an adaptation mechanism must be robust: that is, it must remain functional over a wide range of stimulus levels and be insensitive to variations in the biochemical parameters of the network.

A central strategy for realizing robust perfect adaptation is to use integral feedback. Roughly speaking, a system that implements integral feedback does two things. First, it measures a quantity that reflects some undesirable imbalance (e.g., deviation from baseline) by integrating that quantity over time. Second, it uses the result of that integration to drive processes that correct the imbalance and drive it to zero (i.e., it engages negative feedback). These behaviors are present in biological systems. For example, in Barkai and Leibler (1997) and Alon et al. (1999), it was observed that the bacterial chemotaxis network achieves adaptation to attractant signals and that the precision of this adaptation is robust to changes in the concentration of the network's proteins. It was subsequently inferred by Yi et al. (2000) that robust perfect adaptation in bacterial chemotaxis is achieved through integral feedback. While bacterial chemotaxis provides one example of a network that can realize integral feedback control, other systems with different network structures have also been shown to utilize integral feedback for achieving robust perfect adaptation. For example, in *Saccharomyces cerevisiae*, nuclear enrichment of the MAP kinase Hog1 adapts perfectly to changes in external osmolarity following an osmotic shock (Muzzey et al., 2009). This adaptation requires Hog1 kinase activity (Muzzey et al., 2009), and together these observations reveal the presence of an uncharacterized integral feedback action that is dependent on the Hog1-containing MAP kinase cascade. At the level of organ systems, we have demonstrated that calcium homeostasis in mammals relies on an integral feedback strategy to achieve robust perfect adaptation following changes in blood plasma calcium clearance or influx (El-Samad et al., 2002). This ability to maintain homeostasis is biologically important: it allows mammals to maintain physiological concentrations of blood plasma calcium within tight tolerances, despite the fact that demands for calcium vary substantially over time. Other motifs ensuring robust perfect adaptation in biochemical networks have also been reported (Ma et al., 2009; Shinar and Feinberg, 2010), and their connection to integral feedback control is currently a topic of research.

In engineering applications, integral feedback is recognized as a principal strategy for homeostatic-type regulation (for a primer, see Box 1). The Proportional-Integral-Derivative (PID) control architecture, which includes integral feedback as an essential element, is a workhorse of industrial control systems and is implemented in the majority of all automatic control applications (Åström and Hägglund, 1995). Undoubtedly, the prevalence of such a control strategy in natural and man-made systems is due to a well-described, inherent property of integral feedback: it can robustly steer a regulated system variable to a desired

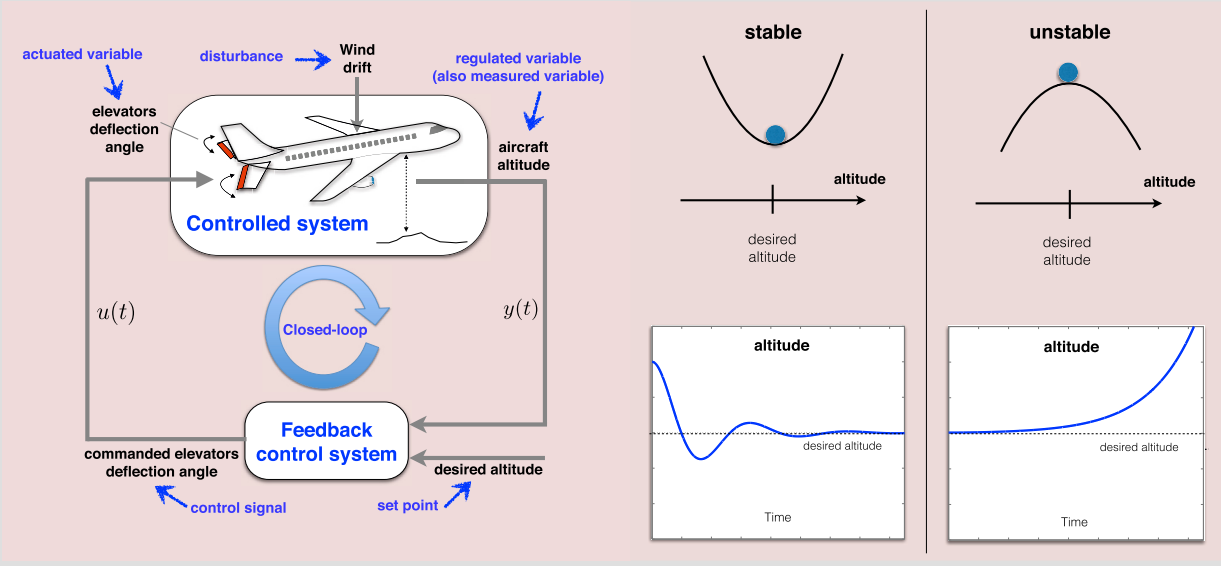
Box 1. Elements of Feedback Control Systems: A Primer

Feedback control systems are workhorses of modern technology. They are also prevalent at every level of biological organization. Control theory deals with the analysis and the design of such systems. We outline the key elements of feedback control systems using a familiar example—a modern airplane. The goal of every feedback control system is to regulate the behavior of a given dynamical system (we refer to this as "the controlled system"). Regulation is achieved by manipulating a variable of the system (we refer to this as "the actuated variable"), which, in turn, affects the system dynamics. The actuated variable is manipulated according to a goal: the system output (which we refer to as "the regulated variable") should follow a desired behavior. In the case of our airplane example (see Box 1 figure, left), the actuated variable is the deflection angle of the elevator, while the regulated variable is the airplane altitude. The control objective is to have the regulated variable reach a desired altitude (we refer to this as the "set point") and to remain there in stable balance over time (we refer to this as "set-point tracking"), in spite of external factors like wind drift (we refer to these as "disturbances") that tend to drive the system into imbalance.

Manipulating the actuated variable to achieve set-point tracking is the task of the "feedback control system." This system can be physically built (in real-world applications) or can be inferred to exist based on systems-level behaviors observed during experiments (in biology). We focus on real-world examples here. The control system takes as its input a continuous measurement of a dynamic variable that gives information about the regulated variable (we refer to this as "the measured variable"; it happens to be equal to the regulated variable in our airplane example). Information from the measured variable is processed in the control system, along with the desired set point, to generate a correction signal (we refer to this signal as "the control signal") that is fed back to the controlled system, thereby "closing the loop." For example, the control signal in our example is the electronic signal that actuates the hydraulics responsible for deflecting the elevators of the airplane. There are various possible strategies that could be employed by the control system. In an integral feedback-based strategy, the control system calculates the integral, with respect to time, of the difference between the set point and the regulated variable and incorporates that into the control signal. This gives the integral feedback control strategy its name.

Importantly, this system actually contains two systems: the control system itself and the controlled system. The control system has its own dynamics, and, together with the dynamics of the controlled system, they form the aggregate dynamics of the closed-loop system, which has new properties that reflect the tight coupling of both. When coupling the dynamics of two systems through feedback, one of the key requirements is the dynamic stability of the new closed-loop system (we term this "closed-loop stability", see Box 1 figure, right). The stability of each system in isolation does not guarantee the stability of the closed loop, and one must carefully design the control system to achieve closed-loop stability. Once stability is assured, one can begin to explore the robust set-point tracking property of the system. Mathematically, set-point tracking is achieved when: $e(t) \rightarrow 0$ as $t \rightarrow \infty$, where $e(t)$ is the error signal, defined as the difference between the desired set point, r, and the controlled output $y(t)$. The above condition assures that the controlled variable tracks the desired set point after some time has elapsed.

Achieving such set-point tracking performance in spite of unknown disturbances and unknown dynamics of the controlled system is referred to as "robust set-point tracking" and is an important goal of many control systems. Control theory stipulates that control systems with certain structural properties, such as integral feedback, will achieve set-point tracking, and that tracking will be maintained regardless of system parameters and constant disturbances (provided stability is preserved). These ideas are explored more fully in Åström and Murray (2008) and Franklin et al. (2014).



set point, while achieving perfect adaptation to disturbances (or stimuli), regardless of the model parameters (Yi et al., 2000; Åström and Murray, 2008; Francis and Wonham, 1976).

Engineered biological circuits that display perfect adaptation are rare, and current synthetic circuits only rely on simpler feedback strategies. For example, several such biochemical networks have been employed to maintain an upward bound on the level of biofuel production in bacteria (Dunlop et al., 2010). These networks were theoretically analyzed by Dunlop and colleagues, revealing that synthetic feedback loops could be used to optimize biofuel production while maintaining a low biofuel toxicity. However, the type of controllers used in Dunlop et al. (2010) did not ensure robust perfect adaptation and remained sensitive to model uncertainties. Similarly, synthetic negative feedback loops have been designed to control the translation rates of specific proteins in mammalian cells, with the goal of tuning the expression levels of multiple proteins in concert (Stapleton et al., 2012). Instead of integral feedback strategies, these engineered circuits relied on simpler feedback strategies: the strength of negative feedback is proportional to the deviation from the desired equilibrium—not the time integral of that deviation, as it is in integral feedback. While systems employing such a strategy will exhibit some amount of adaptation to constant disturbances, this adaptation is neither perfect nor robust. Instead, a constant deviation from the desired equilibrium will persist, and the amount of this deviation will depend on both the system parameters as well as the size and type of the disturbance. Consequently, these synthetic circuits require a cumbersome tuning of parameters to achieve their goals. Such tuning is difficult to realize in a biological setting (Dunlop et al., 2010). This suggests a need for synthetic circuits that are both simple and employ integral feedback in order to achieve robust perfect adaptation.

In the noise-free (deterministic) setting, integral feedback control is well understood (Åström and Hägglund, 1995) and is usually implemented in engineering applications in a configuration similar to that shown in Box 1, using flight control as an example. While the scale and exact implementation of a control strategy may change drastically, all deterministic systems (including biological ones) employing integral feedback in stable feedback loops are guaranteed to achieve robust perfect adaptation. However, in spite of our relatively good understanding of robust perfect adaptation in noise-free contexts and the role of integral feedback in achieving it, determining what constitutes integral feedback in biologically important settings, where the dynamics are described by stochastic processes, remains unclear. Moreover, strategies that are analogous to integral feedback control in intrinsically noisy cellular environments are unknown.

As in the deterministic case, a "stochastic integral feedback" strategy must achieve three goals if effective homeostasis is to be maintained. First, the overall system must have a property called "closed-loop ergodicity," meaning that the joint probability distribution of the random state vector representing the abundances of the species present in the system converges with time to a unique stationary distribution, irrespective of the initial abundances. This is somewhat analogous to the deterministic case where the trajectories of system variables (e.g., concentrations) of a closed-loop stable system converge to a set of stable equilibria regardless of their initial conditions. Second, it must display robust set-point tracking; that is, some statistic that describes a variable of interest (e.g., its mean) reaches a desired value (that is, a "set point") in a manner that is insensitive to system parameters, except those defining the set point itself. Third, it must achieve robust perfect adaptation to external disturbances. Unlike the deterministic setting, however, set-point tracking and adaptation robustness must be maintained not only with respect to model parameters, like biochemical rate constants, but also in the presence of stochastic fluctuations in species abundances.

To better understand how biology may achieve stochastic integral feedback, one can try to engineer biochemical circuits with these goals in mind. For example, one could use statistical moments (such as the mean or the variance) to describe the process to be regulated, and then to design feedback regulation strategies that steer these moments to desired values while achieving perfect adaptation (Milias-Argeitis et al., 2011). While this approach brings the problem back to the deterministic domain (statistical moments evolve according to deterministic dynamics), one is immediately faced with the so-called "moment closure problem," whereby an infinite set of deterministic differential equations is needed to determine even the first two moments (Hespanha, 2008). Similar difficulties arise if one works with the chemical master equation (see, e.g., Gillespie, 1997).

Here, we introduce an integral feedback strategy that we call antithetic integral feedback, which exhibits robust set-point tracking and robust perfect adaptation in the stochastic setting. Rather than dealing with the deterministic moments dynamics, we work with the stochastic chemical reaction network directly, thereby circumventing the moment closure problem. The objective of our control setup, represented in Figure 1A, is to bring the population average of the species $\mathbf{X}_\ell$ involved in an undefined network (the "cloud" in Figure 1A) to a desired set point. To achieve this, a new set of chemical reactions is introduced in a way that effectively implements a stochastic "antithetic integral feedback controller." For clarity, we introduce this strategy twice, once using the language of control theory, and once using more biological language.

Biological Description

The network is sensitive to $\mathbf{X}_1$, which is produced at a rate that is proportional to the amount of $\mathbf{Z}_1$. Through a network that can be arbitrarily large and uncharacterized, $\mathbf{X}_1$ activates $\mathbf{X}_\ell$. $\mathbf{Z}_2$ is produced at a rate that is proportional to the amount $\mathbf{X}_\ell$. $\mathbf{Z}_1$ and $\mathbf{Z}_2$ "annihilate" each other; however, the annihilation reaction need not result in the physical destruction of $\mathbf{Z}_1$ and $\mathbf{Z}_2$. Rather, it can be the biological activity that is abolished. For example, $\mathbf{Z}_1$ and $\mathbf{Z}_2$ may form a high-affinity, biologically inert dimer.

Regulation of *E. coli*'s sigma factor 70, σ^{70}, provides a hypothetical, concrete example of this mechanism (Figure 1B). σ^{70} is an auxiliary component of the RNA polymerase (RNAP) holoenzyme; it acts as a specificity factor and targets RNAP to the promoters of housekeeping genes. σ^{70} is regulated by a corresponding anti-sigma factor (Rsd) (Treviño-Quintanilla et al., 2013), denoted here by $\overline{\sigma}^{70}$. $\overline{\sigma}^{70}$ binds σ^{70} very tightly, sequestering it away from the RNAP core enzyme and changing the transcriptional program of the cell. Importantly, transcription of the Rsd gene is itself controlled by a σ^{70} -dependent promoter (Jishage and Ishihama, 1999); this dependence effectively closes the feedback loop that controls the activity of σ^{70}. We stress that it is yet to be experimentally verified that the σ^{70}

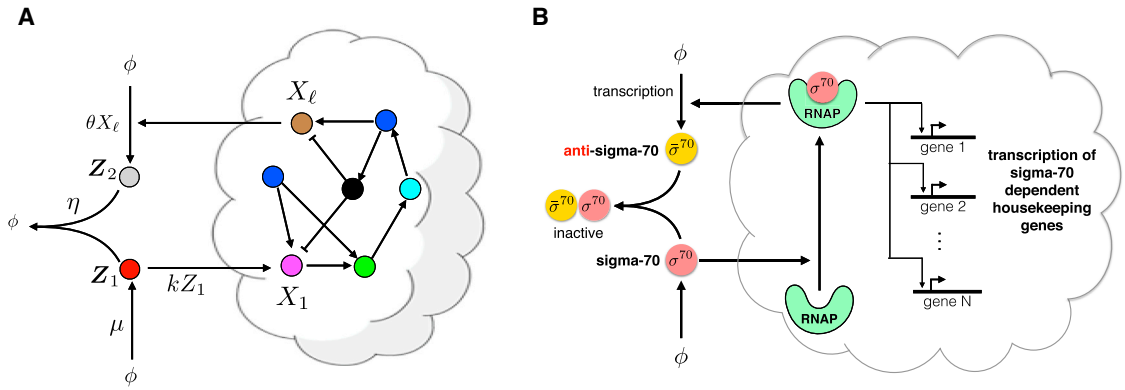


Figure 1. Antithetic Integral Feedback Controller: Generic Realization and a Specific Endogenous Example

(A) Schematic representation showing the constituents of our biomolecular control system. The network on the right (inside the cloud) represents the open-loop network, whose dynamics are to be controlled. Control is achieved by augmenting another network of reactions, referred to as the controller network (outside the cloud). Together the two networks form the closed-loop network whose dynamics are determined by the coupling resulting from the interaction of both networks. The proposed controller network, which we refer to as a "stochastic antithetic integral controller" (defined later on in Equation 1), acts on the open-loop network by influencing the rate of production of the actuated species $\mathbf{X}_1$ by means of the control input species $\mathbf{Z}_1$. The regulated species $\mathbf{X}_\ell$ will be influenced by the increase or decrease of the actuated species $\mathbf{X}_1$ and, in return, will influence the rate of production of the sensing species $\mathbf{Z}_2$, that will, finally, annihilate with the control input species $\mathbf{Z}_1$, thereby implementing a negative feedback control loop. The integral action is encoded in all the reactions of the controller network. (B) Regulation of the housekeeping genes in *E. coli*. An endogenous circuit that may employ the control strategy proposed in this article is the $\boldsymbol{\sigma}^{70}$ regulation circuit. The sigma factor $\boldsymbol{\sigma}^{70}$ binds RNA polymerase core in a complex that controls the expression of housekeeping genes during exponential growth conditions. The anti- $\boldsymbol{\sigma}^{70}$ factor, Rsd, whose gene is controlled by $\boldsymbol{\sigma}^{70}$, binds with a very strong affinity to $\boldsymbol{\sigma}^{70}$, sequestering it away from RNA polymerase. According to the theory put forth in this article, both interactions result in the tight regulation of the concentration of the **RNAP** : $\boldsymbol{\sigma}^{70}$ complex through negative feedback. While the average abundance of RNA polymerase is about 4,600 molecules per cell (Bakshi et al., 2012), the average abundance of the complex **RNAP** : $\boldsymbol{\sigma}^{70}$ (regulated species) is approximately 700 molecules per cell (Jishage and Ishihama, 1999). At the same time, the abundance of *free-* $\boldsymbol{\sigma}^{70}$ is computed to be fewer than one molecule per cell.

system in *E. coli* functions as a stochastic antithetic integral feedback controller. We simply present this as a plausible example here to facilitate understanding. (It is explored more fully in Discussion.)

Control-Theoretic Description

Our controller network consists of four reactions and two additional controller species ($\mathbf{Z}_1$ and $\mathbf{Z}_2$) that can annihilate each other. The species $\mathbf{Z}_1$ actuates the network through $\mathbf{X}_1$ which, in turn, influences the production of $\mathbf{Z}_2$ through the output species $\mathbf{X}_\ell$.

In the analysis of the stochastic antithetic integral feedback controller presented here, we demonstrate that, for a large class of networks, the steady-state value for the population average of $\mathbf{X}_\ell$ depends exclusively on the ratio of two of the controller parameters, and is independent of the network parameters. In this respect, the closed-loop network exhibits stability, robust set-point tracking, and robust perfect adaptation with respect to $\mathbf{X}_\ell$. To analyze such stochastic systems and to guarantee that they achieve these objectives, a new theory is needed. We develop such a theory here and use it to characterize a large class of networks for which we rigorously prove that our proposed controller motif indeed achieves closed-loop ergodicity, robust set-point tracking, and robust perfect adaptation—at the population level and also, in a time-averaged sense, at the single-cell level. Notably, for the class of networks considered our motif requires no specific knowledge of implementation of the biochemical network's structure or parameters. Our "stochastic antithetic integral control motif" can provably achieve all the desired properties mentioned above, even when very low molecular copy numbers exist anywhere in the network. This presents a clear advantage in synthetic biology applications, where synthetic control loops involving large molecular counts can impose a debilitating metabolic load on the cell. Our control scheme can also be shown to possess stabilizing properties that are not found in deterministic implementations of the same circuit. This provides a clear example where the intrinsic stochastic noise is beneficial—it stabilizes a system that would otherwise be unstable. This beneficial effect of noise, in the context of control theory, joins several other recent examples from the literature of noise's constructive functions including stochastic focusing (Paulsson et al., 2000), noise-induced oscillators (Vilar et al., 2002), and noise-induced switches (Tian and Burrage, 2006; Acar et al., 2008).

RESULTS

In what follows, we elaborate on the control problem under consideration, the proposed controller, along with some technical results stating the conditions under which the proposed controller solves the considered control problem. Interestingly, these conditions obtained from probability theory connect to well-known concepts of control theory, such as stability and controllability. Some additional properties, such as robustness and innocuousness, are also discussed.

The Network to Be Controlled: Open-Loop Network

We start by describing the reaction network we aim to control. Consider a reaction network with mass-action kinetics involving d molecular species denoted by $\mathbf{X}_1, \ldots, \mathbf{X}_\mathbf{d}$. Under the

well-mixedness assumption (Anderson and Kurtz, 2011), we can model the dynamics by a Markov process whose state at any given time is simply the vector of molecular counts of the d species. The "Markovian" assumption on the dynamics means that given the current state of the system, the future evolution of the state is independent of the past (memoryless property). The state evolution is influenced by K reaction channels: if the state is x, then the k-th reaction fires at rate $\lambda_k(x)$, and it displaces the state by the "stoichiometric vector" $\zeta_k \in \mathbb{Z}^d$, where $\mathbb{Z}$ denotes the set of integers. Here, λ_k is called the "propensity function" of the k-th reaction and is assumed to satisfy the property that if for any $x \in \mathbb{N}_0^d$ we have $x + \zeta_k \notin \mathbb{N}_0^d$, then $\lambda_k(x) = 0$, where $\mathbb{N}_0$ denotes the nonnegative integers. This property ensures that molecular counts of all the species remain nonnegative throughout the dynamics. In the following, we shall refer to this network as the "open-loop reaction network" and denote it by $(\mathbf{X}, \lambda, \zeta)$.

We now fix a state-space $\mathcal{S}$ for the Markovian reaction dynamics. This set $\mathcal{S}$ is a non-empty subset of $\mathbb{N}_0^d$ which is closed under the reaction dynamics. This means that for any state $x \in \mathcal{S}$ we must also have $(x + \zeta_k) \in \mathcal{S}$ if the the k-th reaction has a positive rate of firing $(\lambda_k(x) > 0)$ at state x. Selecting $\mathcal{S}$ this way allows us to use it as a generic state-space for all Markov processes describing the reaction kinetics and starting at an initial state in $\mathcal{S}$ (Gupta et al., 2014). Henceforth, we denote by $\{X(t) = (X_1(t), \ldots, X_d(t)) : t \geq 0\}$ the continuous time Markov process representing the reaction dynamics with an initial state $x_0 \in \mathcal{S}$.

From a control theoretic point of view, it is necessary to define input and output nodes of the above network. We assume here that species $\mathbf{X}_1$ is the "actuated species" that is the species that the controller can act on. The "regulated species" is $\mathbf{X}_\ell$, for some $\ell \in \{1, \ldots, d\}$, and it is the species we wish to control. The way the controller acts on the actuated species, in order to control the regulated species is depicted in Figure 1A. This will be explained in more detail in the next section.

The Control Objectives

We now state the objectives required from our control system, which includes an open-loop network and a controller network.

Objectives

Find a controller (set of additional reactions and additional species) such that, by suitably acting on the actuated species $\boldsymbol{X}_1$, we have the following properties for the closed-loop network (defined here as the interconnection of the open-loop network $(\mathbf{X}, \lambda, \zeta)$ described above with the controller network):

1. The closed-loop network is ergodic.
2. The first and second-order moments of $X(t)$ exist and are uniformly bounded and globally converging with time to their unique stationary value.
3. We have that $\mathbb{E}[X_\ell(t)] \to \mu^*$ as $t \to \infty$ for some desired set point $\mu^* > 0$.

The first requirement is fairly standard. Indeed, ergodicity is the analog of having a globally attracting fixed point for deterministic dynamics (i.e., global stability) and is required here so that the closed-loop network is well behaved, in the sense that it reaches stationarity starting from any initial distribution. The second requirement is more specific to stochastic processes, because if only the means converge, the variance can still grow unboundedly with time, which would mean that the actual dynamics of the process (its sample paths) is not well behaved, rendering the controller of little practical utility. Finally, the third statement encapsulates our desired objective of perfect adaptation (or set-point tracking), i.e., that the population mean of the regulated species $\mathbf{X}_\ell$ approaches a fixed homeostatic value μ^*.

The Controller Reactions

We propose the following controller network (Figure 1):

$$\underbrace{\varnothing \xrightarrow{\mu} \mathbf{Z}_1}_{\text{reference}}, \quad \underbrace{\varnothing \xrightarrow{\theta X_\ell} \mathbf{Z}_2}_{\text{measurement}}, \quad \underbrace{\mathbf{Z}_1 + \mathbf{Z}_2 \xrightarrow{\eta} \varnothing}_{\text{comparison}}, \quad \underbrace{\varnothing \xrightarrow{kZ_1} \mathbf{X}_1}_{\text{actuation}}, \qquad \text{(Equation 1)}$$

where $\mathbf{X}_\ell$ is the species *measured* by the controller, which is identical to the regulated species in the current setup. The species $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are referred to as the "controller species." Note that the topology of the controller network belongs to a family of four control topologies (see Figure S1 in the Supplemental Information) depending on respective roles of the two controller species.

Although inspired from Oishi and Klavins (2011), the above network has a different philosophy. Besides the fact that the current setting is stochastic, the main difference lies in the way the network interacts with the environment. While the goal of the previously mentioned paper was the biomolecular implementation of linear input-output systems, the goal here is to control a reaction network. In this regard, the birth reactions of $\mathbf{Z}_1$ and $\mathbf{Z}_2$ clearly differ from the way they are defined in the previously cited paper. We now clarify the role and meaning of each of the controller reactions:

1. The first reaction is the "reference reaction" (or set point), which (partially) sets the value of the reference $\mu^* = \mu/\theta$. This value is implemented as the birth rate of species $\mathbf{Z}_1$.
2. The second reaction is the "measurement reaction" and takes the form of a pure-birth reaction with a rate proportional to the current population of the regulated species $\mathbf{X}_\ell$. It is referred to as the measurement reaction as the rate of increase of the population of $\mathbf{Z}_2$ reflects the population size of $\mathbf{X}_\ell$.
3. The third reaction implements the "comparison reaction," which decrements the molecular counts of $\mathbf{Z}_1$ and $\mathbf{Z}_2$ by one each. The rate constant for this reaction is η that can be tuned. The main role of this reaction is to correlate both the populations of $\mathbf{Z}_1$ and $\mathbf{Z}_2$ and to prevent them from growing without bounds. This reaction can be viewed as a "compare and subtract operation" since, when both $\mathbf{Z}_1$ and $\mathbf{Z}_2$ have positive populations (comparison), then this reaction decreases their respective population sizes by one (subtraction), thereby preserving the difference of their population sizes.
4. The last reaction is the "actuation reaction," which implements the way the controller acts on the system, i.e., by acting on the birth rate of the actuated species $\mathbf{X}_1$. The parameter k is also a tuning parameter of the controller.

The above controller has been chosen with an implementability constraint in mind, as it is expressed as plausible reactions that may be implemented in vivo to perform in vivo control. It will be shown later that the proposed controller exhibits strong robustness properties that make its implementation much easier than other types of controllers that require the fine tuning of their reaction rates (see also Section S5.4 of the Supplemental Information). In vitro control is also possible using, for instance, DNA strand displacement (Chen et al., 2013). In silico control (Milias-Argeitis et al., 2011; Uhlendorf et al., 2012) can be considered as well, whenever the population size of regulated species $\mathbf{X}_\ell$ can be measured in real time from outside of the cell(s) using, for instance, microscopy.

Guaranteed Performance Properties of the Controlled Network

We now consider the dynamics of the "closed-loop network," which is formed by interconnecting the open-loop network $(\mathbf{X}, \lambda, \zeta)$ with the controller (Equation 1). We can represent the dynamics by a Markov process $\{(X(t), Z(t)) : t \geq 0\}$, where for each t, $X(t) = (X_1(t), \ldots, X_d(t))$ denotes the molecular counts of the d network species and $Z(t) = (Z_1(t), Z_2(t))$ denotes the molecular counts of the two controller species. Our control objective is to steer the first-order moment $\mathbb{E}[X_\ell(t)]$ corresponding to species $\mathbf{X}_\ell$ to a desired set point μ^*. The dynamics of the first-order moments $\mathbb{E}[X(t)] = (\mathbb{E}[X_1(t)], \ldots, \mathbb{E}[X_d(t)])$ and $\mathbb{E}[Z(t)] = (\mathbb{E}[Z_1(t)], \mathbb{E}[Z_2(t)])$ is described by the following system of ordinary differential equations (ODEs)

$$\begin{aligned} \frac{d\mathbb{E}[X(t)]}{dt} &= \sum_{k=1}^{K} \zeta_k \mathbb{E}\left[\lambda_k(X(t))\right] + k\mathbb{E}[Z_1(t)]e_1, \\ \frac{d\mathbb{E}[Z_1(t)]}{dt} &= \mu - \eta\mathbb{E}[Z_1(t)Z_2(t)], \\ \frac{d\mathbb{E}[Z_2(t)]}{dt} &= \theta\mathbb{E}[X_\ell(t)] - \eta\mathbb{E}[Z_1(t)Z_2(t)], \end{aligned} \qquad \text{(Equation 2)}$$

where e_1 is the d -dimensional vector whose first component is 1 and the rest are zero. Note that this system of ODEs is not *closed* because there is no equation for the dynamics of $\mathbb{E}[Z_1(t)Z_2(t)]$. Moreover, if the propensity functions λ_k's are nonlinear functions of the state variables, additional quantities whose dynamics is not captured by these equations will be encountered. Attempting to "close" this system by adding equations for the dynamics of all these additional quantities and $\mathbb{E}[Z_1(t)Z_2(t)]$ will again lead to another set of quantities with unrepresented dynamics. This is known as the "moment-closure problem" in the literature, a well-known barrier for the direct analysis and simulation of moment equations.

As stated, it is unclear why the proposed controller structure involves integral action. By subtracting the two last equations in Equation 2 and integrating, we get

$$\mathbb{E}[(Z_1 - Z_2)(t)] = \theta \int_0^t \left(\frac{\mu}{\theta} - \mathbb{E}[X_\ell(s)]\right) ds. \qquad \text{(Equation 3)}$$

In other words, the mean difference between the two controller species is proportional to the integral of the mean tracking error. Now the role of integration in achieving perfect steady-state tracking can be understood: if it can be further ensured that as $t \to \infty$, the left-hand side converges to a constant, then the integrand (the mean tracking error) will converge to zero, which is equivalent to having $\mathbb{E}[X_\ell(s)]$ converge to the set point $\mu/\theta = \mu^*$. Ensuring the convergence of the left-hand side to a constant is related to the ergodicity of the network, which we address next.

Our approach in this paper is to find conditions ensuring that the Markov process $\{(X(t), Z(t)) : t \geq 0\}$ describing the dynamics of closed-loop reaction network is ergodic. This means that starting from any initial state, the probability distribution of the state $(X(t), Z(t))$ converges to a unique stationary distribution π as $t \to \infty$. Under fairly general conditions, ergodicity also implies that first- and second-order moments of the state $(X(t), Z(t))$ converge to constant steady-state values, which can be computed by evaluating the expectation $\mathbb{E}_\pi$ with respect to the stationary distribution π. As mentioned above, this guarantees that the integrand in Equation 3 tends to zero as $t \to \infty$ and gives

$$\mathbb{E}_\pi[X_\ell] = \frac{\mu}{\theta} = \mu^*, \qquad \text{(Equation 4)}$$

where $\mathbb{E}_\pi[X_\ell] = \lim_{t\to\infty} \mathbb{E}[X_\ell(t)]$. This shows that if the closed-loop network dynamics is ergodic, then our proposed controller automatically imposes the set-point tracking property, regardless of the initial conditions or model parameters (other than μ/θ). A sudden change in these parameters will result in the same steady-state value after a transient, thereby ensuring perfect adaptation. Achieving steady-state tracking and perfect adaptation is the main rationale behind integral control. Note that these desirable controller properties are demonstrated without solving the first-order moment Equation 2, thereby circumventing the moment-closure problem mentioned earlier.

The following result, proved in Section S4.3 of the Supplemental Information, establishes conditions under which a general stochastic biochemical reaction network can be controlled using the proposed controller network.

Theorem 1: General Network Case

Consider an open-loop reaction network $(\mathbf{X}, \lambda, \zeta)$ and assume that for some given values of its parameters, and the parameters of the controller network, the closed-loop network, formed by augmenting the open-loop network with the controller reactions (Equation 1), is ergodic and has uniformly bounded first- and second-order moments. Then, asymptotic set-point tracking is achieved, i.e., $\mathbb{E}[X_\ell(t)] \to \mu/\theta$ *as* $t \to \infty$.

The main challenge in the above result lies in verifying the ergodicity of a given closed-loop network dynamics. In what follows, we provide simple conditions that allow us to check this property using efficient computational techniques, such as linear programming. This is done for two main classes of networks, namely, those consisting of *unimolecular* reactions and those consisting of *bimolecular* reactions. In the unimolecular case, for each reaction k the propensity function $\lambda_k(x)$ is an affine function of the state variable $x = (x_1, \ldots, x_d)$. Hence, we can express each $\lambda_k(x)$ as

$$\lambda_k(x) = \sum_{i=1}^{d} w_{ki}x_i + w_{k0}$$

where w_{k0} is a nonnegative constant and w_{ki}'s are some real numbers for $i=1,\ldots,d$. Define a $K\times d$ matrix W with entries w_{ki} and let S be the $d\times K$ matrix whose k-th column is the stoichiometry vector ζ_k for reaction k. Also, let w_0 be the d-dimensional vector whose k -th component is w_{k0}. Regarding each vector as a column-vector, for any state-vector x we can write

$$\sum_{k=1}^{K}\lambda_k(x)\zeta_k = SWx + Sw_0,$$

which allows us to express the first equation in Equation 2, with $Z_1(t)\equiv 0$, as

$$\frac{d\mathbb{E}[X(t)]}{dt} = SW\mathbb{E}[X(t)] + Sw_0. \qquad \text{(Equation 5)}$$

This linear system of ODEs describes how the first-order moments of the open-loop network will evolve without the control action. If the matrix SW only has eigenvalues with negative real parts (we say in this case that the matrix SW is Hurwitz stable), then this system is asymptotically stable and the first-order moment vector $\mathbb{E}[X(t)]$ converges as $t\to\infty$. To fulfill our control objective, it is necessary that this system be asymptotically stable. This is because our controller can only act positively on the open-loop network, and hence it cannot stabilize an unstable system.

Recall that $\mathcal{S}\subset\mathbb{N}_0^d$ is the state-space for the Markovian dynamics of the open-loop reaction network. This state-space is irreducible if any state in $\mathcal{S}$ can be reached from any other state in $\mathcal{S}$ by a sequence of reactions having positive propensities at all the intermediate states. A simple example is the single-species birth-death process for which from any state value we can reach any larger state value by a sequence of birth reactions and any smaller state value by a sequence of death reactions. The irreducibility of the state-space is a necessary condition for ergodicity (Meyn and Tweedie, 1993), which holds for many reaction networks in the literature; see e.g., Gupta et al. (2014). For a unimolecular open-loop network with an irreducible state-space $\mathcal{S}$, the asymptotic stability of the linear system (Equation 5) is equivalent to the ergodicity of the open-loop network.

However, in order to ensure the ergodicity of the closed-loop network, we need a couple of other conditions that are both very natural for our control problem. The first condition is that the open-loop system be output controllable (Ogata, 1970), which simply means, in our case, that the molecular count of the regulated species $\mathbf{X}_\ell$ responds to changes in the molecular count of the actuated species $\mathbf{X}_1$. Mathematically, we can express this condition as

$$\left[(SW)^{-1}\right]_{\ell 1}\neq 0 \qquad \text{(Equation 6)}$$

where $[(SW)^{-1}]_{\ell 1}$ is the component at column 1 and row ℓ of the inverse of matrix SW. The second condition that we need is that the set point $\mu^*=\mu/\theta$ should be accessible by the dynamics of $\mathbb{E}[X_\ell(t)]$. Since our controller can only act positively on the system (Equation 5), this accessibility condition can fail if some components of the input vector Sw_0 are too large (see also the discussion below theorem S6.5 in the Supplemental Information). Technically, we can check this accessibility condition by ensuring that there exists a positive constant c and a d-dimensional vector $v=(v_1,\ldots,v_d)$ with positive entries such that each component of the vector $v^T(SW+cI_d)$, where I_d is the $d\times d$ identity matrix, is strictly negative and

$$\mu^* = \frac{\mu}{\theta} > \frac{v^T S w_0}{c v_\ell}. \qquad \text{(Equation 7)}$$

As discussed in Gupta et al. (2014), the above condition can be checked using linear programming techniques.

The following result, proved in Section S6 of the Supplemental Information, provides conditions under which a unimolecular open-loop reaction network can be controlled using the controller network (Equation 1).

Theorem 2: Unimolecular Network Case

Suppose that the open-loop reaction network $(\mathbf{X},\lambda,\zeta)$ is unimolecular and its state-space $\mathcal{S}$ is irreducible. Furthermore, assume that the linear system (Equation 5) of ODEs is asymptotically stable (SW is Hurwitz stable) and output controllable (the condition in [Equation 6] holds), and that the desired set point μ/θ is accessible (the condition in [Equation 7] holds).

Then, for any $k,\eta>0$, the closed-loop reaction network is ergodic and asymptotic set-point tracking is achieved; i.e., $\mathbb{E}[X_\ell(t)]\to\mu/\theta$ *as* $t\to\infty$. Moreover, the steady-state values of the first-order moments $\mathbb{E}[X(t)]=(\mathbb{E}[X_1(t)],\ldots,\mathbb{E}[X_d(t)])$ and $\mathbb{E}[Z_1(t)]$ are

$$\lim_{t\to\infty}\mathbb{E}[X(t)] = \frac{\mu(SW)^{-1}e_1}{\theta\left[(SW)^{-1}\right]_{\ell 1}} \quad \textit{and} \quad \lim_{t\to\infty}\mathbb{E}[Z_1(t)] = \frac{-\mu}{\theta\left[(SW)^{-1}\right]_{\ell 1}}.$$

A more general version of this result and its extension to a class of bimolecular networks are provided in Section S6 and Section S7 of the Supplemental Information. In light of theorem 2, several favorable properties for the controller and the closed-loop network can now be highlighted and expounded.

Ergodicity, Set-Point Tracking, and Bounded First- and Second-Order Moments

These are the main properties sought in our statement of control objectives. Moreover, as stated in theorem 2, the average population sizes of species $\mathbf{X}_1,\ldots,\mathbf{X_d}$ and $\mathbf{Z}_1$, at steady state, are uniquely defined by the set point μ/θ and the parameters of the open-loop network, implying that these quantities are also regulated by our stochastic antithetic integral controller.

Robustness

Robustness is a fundamental requirement that ensures that some properties for the closed-loop network are preserved, even in the presence of model uncertainties. This concept is critical in biology as the environment is fluctuating or noisy and only poorly known models are typically available. The obtained results can automatically guarantee the preservation of all the properties stated in Theorem 2, even in such constraining conditions.

Well-Behaved Single-Cell Tracking Dynamics

Ergodicity ensures that the population average at stationarity is equal to the asymptotic value of the time average of any single-cell trajectory; see, e.g., Gupta et al. (2014). We can therefore conclude that the proposed controller

achieves two goals simultaneously, as it can ensure robust set-point tracking at both the population and single-cell levels. As a consequence, the controller will also ensure single-cell set-point tracking in the presence of cell events such as cell-division when certain conditions are met (see Section S8.5 of the Supplemental Information for more details).

Innocuousness of the Controller

An inaccurate implementation of controller parameters may sometimes lead to an unstable behavior for the closed-loop system. However, the fact that the conditions of Theorem 2 are independent of k and η tells us that the proposed controller will both preserve the ergodicity of the (possibly uncharacterized) open-loop network and ensure set-point tracking/adaptation regardless of the values of its parameters, provided that the open-loop network satisfies the conditions of Theorem 2. This property is crucial in biology, as identifying models and implementing specific reaction rates (even approximately) are difficult tasks. This peculiar and non-standard property is referred here as ''innocuousness,'' and it is illustrated in Figure 4, where the deterministic and stochastic dynamics of the same controlled networks are compared (see also Section S5.4 of the Supplemental Information).

Low Metabolic Load

Even if the controller works in the low copy-number regime, it does not necessarily imply a low metabolic load for the cell. Indeed, if there is fast creation and annihilation of the controller species $\mathbf{Z}_1$ and $\mathbf{Z}_2$, then it will result in many energy consuming futile cycles, which can impose a heavy metabolic burden on the cell, even though the dynamics is still in the low copy-number regime. However, it can be shown (see Section S5.5 of the Supplemental Information) that the power consumption at stationarity of the controller reactions, denoted by $\overline{P}$, can be expressed in the case of unimolecular networks as

$$\overline{P} = \mu(\alpha_1 + \alpha_2 + \alpha_3) + \frac{\mu}{\theta}\frac{\alpha_4}{\left|\left[(SW)^{-1}\right]_{\ell 1}\right|} \qquad \text{(Equation 8)}$$

where $\alpha_1, \alpha_2, \alpha_3$ and α_4 are (positive) weights associated with the reference, measurement, comparison, and actuation reactions, respectively, that represent the energy cost of each reaction. Interestingly, only the first three terms depend on μ while the last term depends on the ratio μ/θ and some network parameters. This last term, however, would also be present in the case of the production of $\mathbf{X}_1$ at a constitutive rate equal to $\mu/(\theta\left|e_\ell^T(SW)^{-1}e_1\right|)$ (which would lead to the same steady-state value for the controlled output, but no adaptation properties). Hence, the effective metabolic load of our controller is equal to the first three terms and is only proportional to μ. A low metabolic load can therefore be easily achieved by first setting μ to a small value and then adjusting the set-point value with θ.

Circumventing Moment Closure Difficulties

Finally, we emphasize that using the proposed approach, the moment closure problem does not arise, as the main conclusions (e.g., ergodicity, set-point tracking and robustness) directly follow from stochastic analysis tools and the structure of the controller, thereby avoiding altogether the framework of the moment equations (see Section S8.6 of the Supplemental Information).

Application to Gene Expression Control: Set-Point Tracking and Perfect Adaptation

The goal of this example is to demonstrate that set-point tracking, and perfect adaptation can be ensured with respect to any change in the parameters of the gene expression network

$$\mathbf{X}_1 \xrightarrow{\gamma_1} \varnothing, \quad \mathbf{X}_1 \xrightarrow{k_2} \mathbf{X}_1 + \mathbf{X}_2, \quad \mathbf{X}_2 \xrightarrow{\gamma_2} \varnothing \qquad \text{(Equation 9)}$$

where $\mathbf{X}_1$ denotes the mRNA and $\mathbf{X}_2$ the measured/regulated species (see Figure 2). The following result is proved in Section S8.2 of the Supplemental Information:

Proposition 3

For any positive values of the parameters $k, k_2, \gamma_1, \gamma_2, \eta, \theta$, and μ, the controlled gene expression network and is ergodic, has bounded and globally converging first- and second-order moments and

$$\mathbb{E}[X_2(t)] \to \frac{\mu}{\theta} \quad as \quad t \to \infty. \qquad \text{(Equation 10)}$$

However, when we consider a Hill-type static control scheme of the form

$$\varnothing \xrightarrow{f(X_2)} \mathbf{X}_1 \quad with \quad f(X_2) = \frac{\alpha K^n}{K^n + X_2^n} \qquad \text{(Equation 11)}$$

where K, α are positive parameters and n is a positive integer, we obtain the results depicted in Figure 3. We can see that, as opposed to the stochastic antithetic integral controller, perfect adaptation is not ensured by the Hill-type static controller. Note that, even though the comparison is only made for the gene expression network and this specific choice for the Hill-type static controller, it is provable that, in general, such controllers can not ensure perfect adaptation; see Section S8.1 of the Supplemental Information for some theoretical arguments and different control schemes.

Noise as a Stabilizing Agent: Stochastic versus Deterministic Population Control

Here, we demonstrate a striking effect of noise as an agent for dynamic stabilization at the population level. We do this by comparing the results that we obtain to those we would have obtained in the deterministic setting (see Figure 4). To this aim, we again consider the gene expression network (Equation 9), and we set $k_2 = \gamma_1 = \gamma_2 = 1$ for simplicity. We then get the deterministic model and stochastic mean model depicted in Figures 4A and 4B, respectively. It is important to emphasize that the deterministic model represents here the evolution of the mean concentration of the species over a population of deterministically behaving cells with identical initial concentrations. On the other hand, the stochastic mean model represents the evolution of the mean number of the species over a population of stochastically behaving cells with identical initial molecular counts. Note that, by virtue of the ergodicity property, the stochastic mean model will always converge to its unique steady-state value regardless of the different initial conditions for the individual cells. This property does not hold in general for a deterministic model representing the average

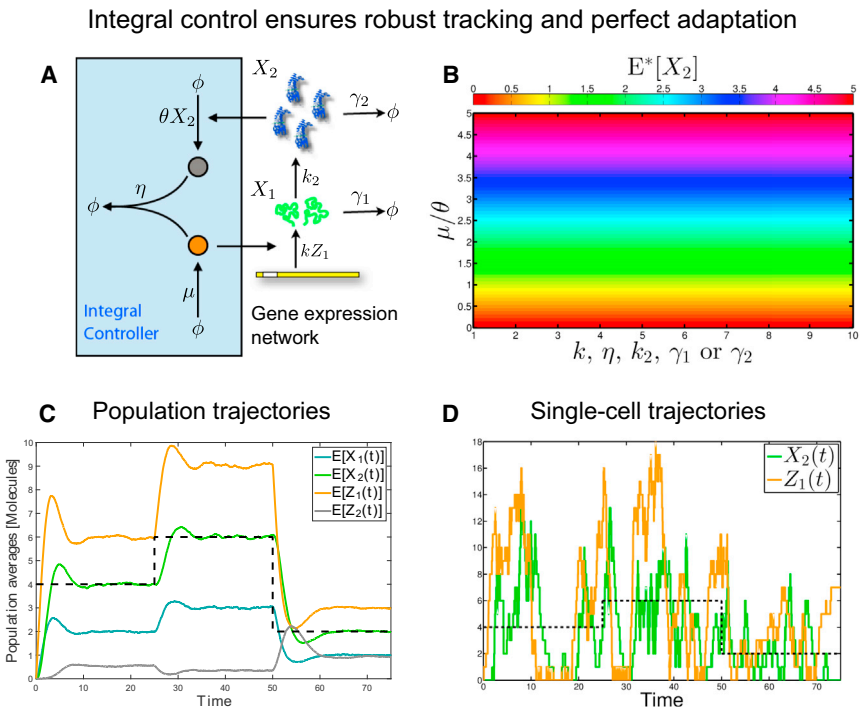


Figure 2. Stochastic Gene Expression Regulated by Antithetic Integral Controller Exhibits Tracking and Robust Perfect Adaptation

(A) The controlled gene expression network (Equation 9) with the proposed stochastic antithetic integral controller (Equation 1).

(B) The closed-loop reaction network shows perfect adaptation (at stationarity) with respect to any changes in the parameters of the network as we have that $\mathbb{E}_\pi[X_2] = \mu/\theta$ for any values of the parameters k, η, k_2, γ_1, and γ_2 where $\mathbb{E}_\pi[X_2]$ denotes the mean number of molecules of $\mathbf{X}_2$ at stationarity.

(C) The controlled-output $\mathbb{E}[X_2(t)]$ of the closed-loop network tracks the reference value (in black-dash). The mean population of input species $\mathbb{E}[Z_1(t)]$ adapts automatically to changes in the reference value $\mu^* = \mu/\theta$ without requiring re-implementation.

(D) Single-cell trajectories, although strongly affected by noise, still have an underlying regularity ensuring the convergence of the moments at the population level. All simulations have been performed using Gillespie's stochastic simulation algorithm with the parameters $k = 1$, $\gamma_1 = 3$, $k_2 = 2$, $\gamma_2 = 1$, $\theta = 1$, and $\eta = 50$.

concentrations of a population of deterministically oscillating cells. The stochastic mean model has been obtained using the identity $\mathbb{E}[Z_1 Z_2] = \mathbb{E}[Z_1]\mathbb{E}[Z_2] + Cov(Z_1, Z_2)$ where the covariance term is nonzero as the random variables are not independent. If such a term would be zero, then we would recover the deterministic dynamics, but, due to noise, we can see in Figure 4C that while the deterministic dynamics may exhibit oscillations, the dynamics of the first-order moment is always globally converging to the desired steady-state value. As a final comment, we note that if we were closing the moment equations in Figure 4C by neglecting the second-order cumulant, then we would fail in predicting the correct behavior of the first-order moments. This demonstrates the central role of the noise in the stabilizing properties of the proposed stochastic antithetic integral controller. The noise can hence be viewed as here a stabilizing agent through the randomness it adds to the dynamics, allowing then for their systematic compensation at the population level, regardless of the initial conditions and the system parameters. This phenomenon is entirely due to stochasticity, and it does not generically occur in the deterministic setting.

DISCUSSION

A general control theory for stochastic biochemical reaction networks with tailored mathematical concepts and tools has been missing. We believe that a well-grounded biomolecular control theory would enable a deeper understanding of biological regulation at the molecular level. Moreover, it could pave the way for an efficient and systematic rational design of synthetic genetic circuits that steer cellular dynamics in understood ways. For such circuits, we propose the term "cybergenetic," which combines the genetic nature of the system with its cybernetic function of dynamic steering and control ("cyber" means "to steer" in Greek). In this article, we take a first step in the development of such a theory by addressing one of the central dynamic control motifs: integral feedback. The methods we developed are the product of a synthesis of ideas from control theory, probability theory, linear algebra, and optimization theory. Even though our findings are specific to the class of integral controllers we consider, they may serve as the foundation on which more general biomolecular control theory can be developed—one that deals with a larger class of stochastic dynamic controllers and networks. Indeed, numerical experiments performed on more general networks lying outside the scope of the developed theory tend to support this claim (see Section S8.7 of the Supplemental Information).

Although the proposed control motif served as a frame around which the theoretical ideas of molecular stochastic control were developed, a cybergenetic circuit implementing such a motif may be of biological significance in its own right, a possibility that we explore below for both endogenous and synthetic regulation. Indeed, the simplicity of the mechanism and the regulation properties that it confers raises the question whether such a motif could have evolved for the purpose of endogenous regulation. The species $\mathbf{Z}_1$ and $\mathbf{Z}_2$ may be RNA, protein, metabolite, or etc., but they must act to effectively "annihilate" each other. As discussed in the Introduction, this annihilation need not be physical, as long as the two species act to render each other functionless. This could occur, for example, as a result of irreversible binding of $\mathbf{Z}_1$ and $\mathbf{Z}_2$, where the new complex effectively sequesters both from performing their function.

Returning to the example of σ^{70}, $\overline{\sigma}^{70}$, and the regulation of *E. coli*'s housekeeping genes (Figure 1B), what is understood about this system's regulation fits well our antithetic integral feedback motif and suggests that the RNAP core enzyme complex with σ^{70} is the object of tight regulation. This, in turn,

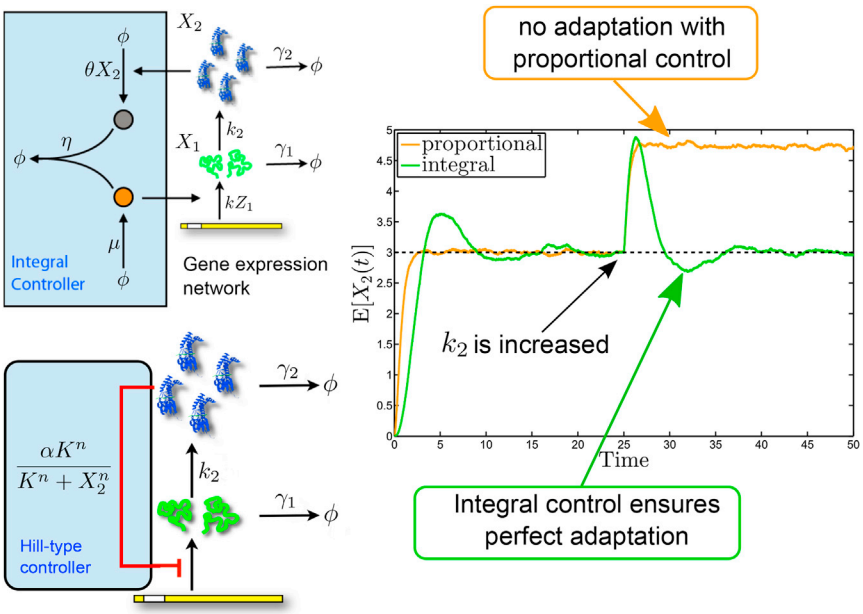


Figure 3. Comparison between Proposed Antithetic Integral Controller and Hill-type Controller

The simulation is performed using parameters initialized to $\mu = 3$, $\theta = 1$, $k = 1$, $\gamma_1 = 3$, $k_2 = 3$, $\gamma_2 = 1$ and $\eta = 50$ for the stochastic antithetic integral controller (Equation 1) and $n = 1$, $\alpha = 8.22$ and $K = 3$ for the Hill-type static controller (Equation 11). The averaging is performed over 8,000 cells simulated with Gillespie's stochastic simulation algorithm. At $t = 25$ s, the value of k_2 jumps from 3 to 6. While the proposed stochastic antithetic integral controller shows perfect adaptation, the Hill-type static controller is unable to return to the mean value of the population of $\mathbf{X}_2$ before the stimulus. This demonstrates the advantage of the integral feedback strategy over the Hill-type strategy.

suggests that the expression of all the housekeeping genes whose promoters are recognized by $\boldsymbol{\sigma}^{70}$ will be regulated in a correspondingly tight way. When binding reaction rates in the literature (Maeda et al., 2000; Bakshi et al., 2012) are used for $\boldsymbol{\sigma}^{70}$ and RNAP along with estimated average total numbers, one comes up with a very low (<1) average copy number of *free*- $\boldsymbol{\sigma}^{70}$ per cell, which would indicate that the system operates in the stochastic regime. Intriguingly, the anti- $\boldsymbol{\sigma}^{70}$ factor $\overline{\boldsymbol{\sigma}}^{70}$ is known to be involved in the transition from the log to the stationary phase of growth, during which time its concentrations are significantly elevated, leading to the downregulation of the housekeeping genes. One way to achieve this is through the increase of the rate of transcription of anti-sigma factor Rsd during transition to stationary phase, leading to the decrease in the set point for $\mathbf{RNAP} : \boldsymbol{\sigma}^{70}$ (in our analysis, this corresponds to the quantity μ/θ). Supporting this hypothesis is the fact that transcription of the $\overline{\boldsymbol{\sigma}}^{70}$ gene is known to be controlled by a $\boldsymbol{\sigma}^{38}$ -dependent promoter (Jishage and Ishihama, 1999), where $\boldsymbol{\sigma}^{38}$ is recognized as the master regulator for adaptation to stationary phase transcription. The full regulatory details of $\boldsymbol{\sigma}^{70}$ continue to unknown, and, while it remains to be experimentally verified that the $\boldsymbol{\sigma}^{70}$ and $\overline{\boldsymbol{\sigma}}^{70}$ interactions in *E. coli* function as a stochastic antithetic integral feedback controller, as put forth here, it is quite likely that these interactions play a central part in the regulation of the $\boldsymbol{\sigma}^{70}$ system, even if they do not fit the scheme proposed exactly. For example, if the expression rate of $\boldsymbol{\sigma}^{70}$, assumed here as constant, turns out to be dependent on $\mathbf{RNAP} : \boldsymbol{\sigma}^{70}$, then our analysis can be extended to show that similar regulation properties still hold. Finally, we speculate that the sigma/anti-sigma regulatory motif similar to that presented here is not uncommon in biology, as several anti-sigma factors have been found in a number of bacteria, including *E. coli* and Salmonella, as well as in the T4 bacteriophage.

Beyond endogenous circuits, our cybergenetic motif presents opportunities for applications in synthetic biology. Until now, most of the synthetic regulatory circuits have relied on proportional action—a control scheme that fails to ensure perfect adaptation in many practical situations. Moreover, existing theoretical studies of synthetic biological circuits mainly considered the deterministic setting, and hence they implicitly assumed large molecular abundances. However, the implementation of control circuits that rely on high component abundances severely impinges on the host circuit's material and energy resources, leading to increased metabolic burden that can affect both function and viability. Fortunately, this is largely avoidable, as effective control involves mostly information processing, which in principle requires little energy and material resource consumption.

The regulatory motif that we propose exhibits characteristics that provably ensure robust stability, robust set-point tracking, and robust perfect adaptation for the controlled network and is achieved with a low metabolic cost and with molecular species that can have very low abundances. It can be used for both single-cell set-point tracking (on average) and for population control. Notably, perfect adaptation will also hold in presence of static extrinsic noise (cell-to-cell variability) when understood as network parameter discrepancies. Thanks to the innocuousness of the controller, it does not need to be fine tuned and can therefore be used in many practical situations, e.g., when the controlled network is very poorly known. In this regard, the proposed controller maintains clear implementability advantages over controllers requiring parameter tuning. This latter property emerges from the random nature of the reactions, as its deterministic counterpart leads to oscillating trajectories when the controller parameters are located in a certain instability region. In spite of this, this controller can still be used in a deterministic setting even though some of the properties, such the innocuousness property, are lost. With this in mind, the proposed controller may find several applications within synthetic biology. An immediate one is the optimization of drug or fuel production in bioreactors; see, e.g., Dunlop et al. (2010). Currently, simple control strategies, such as proportional feedback or constitutive production, are used in these applications. By utilizing slightly more complex controllers, such as the one proposed here, dramatic improvements in the production process can be

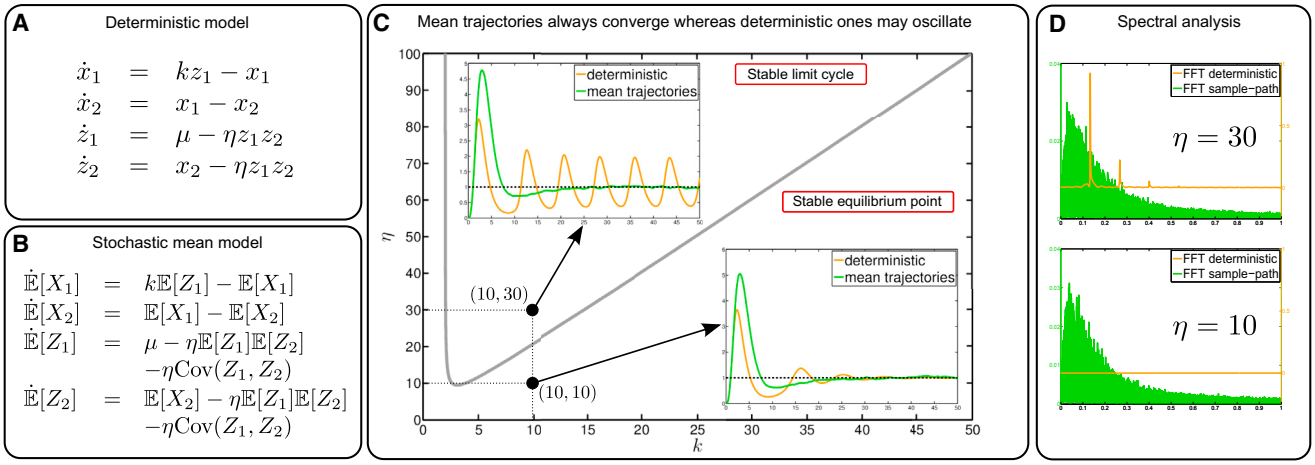


Figure 4. Antithetic Integral Feedback Controller Exploits Stochastic Noise to Achieve Robust Regulation

(A) Deterministic model for the gene expression network (Equation 9) with $k_2 = \theta = \gamma_1 = \gamma_2 = 1$.

(B) Mean model for the gene expression network (Equation 9) with $k_2 = \theta = \gamma_1 = \gamma_2 = 1$.

(C) The deterministic dynamics bifurcates from a unique stable equilibrium point when the controller parameters (k, η) are chosen below the bifurcation curve into a stable limit cycle when the controller parameters are chosen above. The first-order moments, however, always converge to the desired steady-state value for the regulated species, here $\mu = 1$, regardless of the values of the controller parameters. This can be explained by the presence of the stabilizing covariance term in the model for the stochastic means.

(D) While the frequency content at stationarity of the deterministic dynamics dramatically changes when crossing the bifurcation curve, the frequency content of the sample paths remains qualitatively the same. In this regard, the controller can be considered to perform the same way in both cases. This demonstrates the superiority and the central role of the noise in the stabilizing properties of the proposed stochastic integral controller.

expected, thanks to their enhanced robustness properties. Another important application is the design of signal insulators; see, e.g., Mishra et al. (2014), where our motif may also be applied. For example, it can be used as a buffering element in order to drive the output of a module to the input of another one. It can also be used as a constant signal generator that can be used to act on a network to be analyzed. The amplitude can be tuned by acting on the reference, which can be modified from outside the cell using light-induced techniques (Milias-Argeitis et al., 2011).

The proposed controller, however, may have some drawbacks, as it seems to introduce some additional variance to the controlled process. Even though this extra variance is not detrimental to the current control objectives, it may be a problem if the goal is to reduce the variance over a cell population. Whether the variance can be reduced via a more optimal choice of parameters or through some additional controller reactions remains a question for further research. In the end, some "extra" variance due to the controller may be unavoidable, as fundamental limitations to variance reduction with feedback (Lestas et al., 2010) are likely to hold for any molecular control circuit. Even so, controller noise should not detract from the tremendous promise of designing novel stochastic control circuits at the molecular level, where the dynamic properties, benefits, and limitations seem to be exquisitely different from those at the macroscopic scale.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and seven figures and can be found with this article online at http://dx.doi.org/10.1016/j.cels.2016.01.004.

AUTHOR CONTRIBUTIONS

Conceptualization and Methodology, C.B., A.G., and M.K.; Formal Analysis, C.B. and A.G.; Software, C.B.; Writing, C.B., A.G., and M.K.; Supervision and Funding, M.K.

ACKNOWLEDGMENTS

The authors are grateful to Adam Arkin for pointing out the similarity of the tight sigma/anti-sigma factor binding reactions to our feedback controller reactions. The authors acknowledge funding support from the Human Frontier Science Program Grant RGP0061/2011 and the Swiss National Science Foundation grant 200021-157129.

Received: June 24, 2015
Revised: November 10, 2015
Accepted: January 6, 2016
Published: January 27, 2016

REFERENCES

Acar, M., Mettetal, J.T., and van Oudenaarden, A. (2008). Stochastic switching as a survival strategy in fluctuating environments. Nat. Genet. *40*, 471–475.

Alon, U., Surette, M.G., Barkai, N., and Leibler, S. (1999). Robustness in bacterial chemotaxis. Nature *397*, 168–171.

Anderson, D., and Kurtz, T.G. (2011). Continuous time Markov chain models for chemical reaction networks. In Design and Analysis of Biomolecular Circuits: Engineering Approaches to Systems and Synthetic Biology, H. Koeppl, D. Densmore, G. Setti, and M. di Bernardo, eds. (Springer), pp. 3–42.

Åström, K.J., and Hägglund, T. (1995). PID Controllers: Theory, Design, and Tuning (Instrument Society of America).

Åström, K.J., and Murray, R.M. (2008). Feedback Systems: An Introduction for Scientists and Engineers (Princeton University Press).

Bakshi, S., Siryaporn, A., Goulian, M., and Weisshaar, J.C. (2012). Superresolution imaging of ribosomes and RNA polymerase in live Escherichia coli cells. Mol. Microbiol. *85*, 21–38.

Barkai, N., and Leibler, S. (1997). Robustness in simple biochemical networks. Nature *387*, 913–917.

Chen, Y.-J., Dalchau, N., Srinivas, N., Phillips, A., Cardelli, L., Soloveichik, D., and Seelig, G. (2013). Programmable chemical controllers made from DNA. Nat. Nanotechnol. *8*, 755–762.

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.

Francis, B.A., and Wonham, W.M. (1976). The internal model principle of control theory. Automatica *12*, 457–465.

Franklin, G., Powel, J.D., and Emami-Naeinin, A. (2014). Feedback Control of Dynamic Systems, Seventh Edition (Prentice Hall).

Gillespie, D.T. (1997). A rigorous derivation of the chemical master equation. Physica A *188*, 404–425.

Gupta, A., Briat, C., and Khammash, M. (2014). A scalable computational framework for establishing long-term behavior of stochastic reaction networks. PLoS Comput. Biol. *10*, e1003669.

Hespanha, J.P. (2008). Moment closure for biochemical networks. In 3rd International Symposium on Communications, Control and Signal Processing (St. Julian's), pp. 142–147.

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.

Lestas, I., Vinnicombe, G., and Paulsson, J. (2010). Fundamental limits on the suppression of molecular fluctuations. Nature *467*, 174–178.

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.

Maeda, H., Fujita, N., and Ishihama, A. (2000). Competition among sevn Escherichia coliσ. Nucleic Acids Res. *28*, 3497–3503.

Meyn, S.P., and Tweedie, R.L. (1993). Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes. Adv. Appl. Probab. *25*, 518–548.

Milias-Argeitis, A., Summers, S., Stewart-Ornstein, J., Zuleta, I., Pincus, D., El-Samad, H., Khammash, M., and Lygeros, J. (2011). In silico feedback for in vivo regulation of a gene expression circuit. Nat. Biotechnol. *29*, 1114–1116.

Mishra, D., Rivera, P.M., Lin, A., Del Vecchio, D., and Weiss, R. (2014). A load driver device for engineering modularity in biological networks. Nat. Biotechnol. *32*, 1268–1275.

Muzzey, D., Gómez-Uribe, C.A., Mettetal, J.T., and van Oudenaarden, A. (2009). A systems-level analysis of perfect adaptation in yeast osmoregulation. Cell *138*, 160–171.

Ogata, K. (1970). Modern Control Engineering (Prentice Hall).

Oishi, K., and Klavins, E. (2011). Biomolecular implementation of linear I/O systems. IET Syst. Biol. *5*, 252–260.

Paulsson, J., Berg, O.G., and Ehrenberg, M. (2000). Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. Proc. Natl. Acad. Sci. USA *97*, 7148–7153.

Shinar, G., and Feinberg, M. (2010). Structural sources of robustness in biochemical reaction networks. Science *327*, 1389–1391.

Stapleton, J.A., Endo, K., Fujita, Y., Hayashi, K., Takinoue, M., Saito, H., and Inoue, T. (2012). Feedback control of protein expression in mammalian cells by tunable synthetic translational inhibition. ACS Synth. Biol. *1*, 83–88.

Tian, T., and Burrage, K. (2006). Stochastic models for regulatory networks of the genetic toggle switch. Proc. Natl. Acad. Sci. USA *103*, 8372–8377.

Treviño-Quintanilla, L.G., Freyre-González, J.A., and Martínez-Flores, I. (2013). Anti-sigma factors in e. coli: Common regulatory mechanisms controlling sigma factors availability. Curr. Genomics *14*, 378–387.

Uhlendorf, J., Miermont, A., Delaveau, T., Charvin, G., Fages, F., Bottani, S., Batt, G., Hersen, P., Batt, G., and Hersen, P. (2012). Long-term model predictive control of gene expression at the population and single-cell levels. Proc. Natl. Acad. Sci. USA *109*, 14271–14276.

Vilar, J.M.G., Kueh, H.Y., Barkai, N., and Leibler, S. (2002). Mechanisms of noise-resistance in genetic oscillators. Proc. Natl. Acad. Sci. USA *99*, 5988–5992.

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.