bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Realizing 'integral control' in living cells: How to overcome leaky integration due to dilution?

Yili Qian[1] and Domitilla Del Vecchio[1,2,*]

[1] Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA02139, USA
[2] Synthetic Biology Center, Massachusetts Institute of Technology, 500 Technology Square, Cambridge, MA02139, USA

∗ Correspondence:
Address: Room 3-469, Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA02139, USA.
Phone: +1-617-324-6108
Email: ddv@mit.edu

Abstract

A major problem in the design of synthetic genetic circuits is robustness to perturbations and uncertainty. Because of this, there have been significant efforts in recent years in finding approaches to implement integral control in genetic circuits. Integral controllers have the unique ability to make the output of a process adapt perfectly to disturbances. However, implementing an integral controller is challenging in living cells. This is because a key aspect of any integral controller is a "memory" element that stores the accumulation (integral) of the error between the output and its desired set-point. The ability to realize such a memory element in living cells is fundamentally challenged by the fact that all biomolecules dilute as cells grow, resulting in a "leaky" memory that gradually fades away. As a consequence, the adaptation property is lost. Here, we propose a general principle for designing integral controllers such that the performance is practically unaffected by dilution. In particular, we mathematically prove that if the reactions implementing the integral controller are all much faster than dilution, then the adaptation error due to integration leakiness becomes negligible. We exemplify this design principle with two synthetic genetic circuits aimed at reaching adaptation of gene expression to fluctuations in cellular resources. Our results provide concrete guidance on the biomolecular processes that are most appropriate for implementing integral controllers in living cells.

1 Introduction

A long-standing challenge in synthetic biology is the creation of genetic circuits that perform predictably and reliably in living cells [1, 2]. Variations in the environment [3], unforeseen interactions between a circuit and the host machinery [4], and unknown interference between circuit components [5] can all significantly alter the intended behavior of a genetic circuit. Integral control, which is ubiquitous in traditional engineering systems, is a promising approach to mitigate the impact of such unknowns on genetic circuits. At the most basic level, an integral controller ensures that the output of a system can reach a desired set-point while perfectly rejecting (i.e., adapting to) constant disturbances [6–8]. Due to this unique ability, synthetic biology research has witnessed increasing efforts toward establishing procedures to realize integral control in living cells [9–17].

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

At the core of an integral controller is an indispensable "memory" element that accumulates (i.e., integrates) the error between the desired set-point and the measured output over time. While computing this integral is almost trivial using electronic components, realizing it through biomolecular reactions in living cells is problematic. This is because information about the error is usually stored in the form of molecular concentrations, and as host cell grows and divides, biomolecules in every single cell dilute [18, 19], leading to a "leaky" memory that fades away in time. Since cell growth is unavoidable and even beneficial in a number of applications, such as in biofuel production [20], leaky integration is a fundamental physical limitation for the implementation of integral control through genetic circuits in living cells.

Prior theoretical studies have proposed integral control motifs (ICMs), but neglected the presence of molecule dilution in the analysis [10, 12, 15, 17]. Consequently, while these motifs have theoretically guaranteed performance in cell-free systems, where dilution is non-existent, their performance in the presence of molecule dilution is not preserved. In fact, the deteriorating effect of leaky integration can be significant, as shown in this paper and in other recent studies (e.g., [21, 22]). Therefore, possible solutions to leaky integration have been proposed [10, 21]. Specifically, the approach proposed in [10] hinges on "canceling" the effect of dilution by producing the "memory species" at the same rate as its dilution. This approach, however, relies on exact parameter matching and is hence difficult to realize in practice. A different approach is to use exhaustive numerical simulation to extract circuit parameters that minimize the effects of leaky integration [21], leading to a lengthy and *ad hoc* design process.

In this report, we propose general design principles for ICMs that perform well despite the presence of dilution. In particular, we mathematically demonstrate that the undesirable effect of dilution can be arbitrarily suppressed by increasing the rate of all controller reactions. This is true under mild conditions that are largely independent of specific parameter values. This design principle guides the choice of core biomolecular processes and circuit parameters that are most suitable for realizing an ICM in living cells. We illustrate this guided choice on two circuits that are designed to mitigate the effects of transcription/translation resource fluctuations on gene expression, a problem that has gained significant attention recently [4, 23–25].

2 Quasi-integral control

The approach that we take in this paper is as follows. We first describe two types of ideal ICMs in Section 2.1, which were previously proposed abstract circuit motifs for adapting to constant disturbances in the absence of dilution. We then introduce leaky ICMs, which add dilution to these ideal ICMs, and demonstrate that the adaptation property is lost. Finally, we describe quasi-ICMs in Section 2.2, the main novelty of this paper, in which the controller reactions are engineered to be much faster than dilution, enabling the circuit to restore almost perfect adaptation to constant disturbances.

2.1 Ideal ICMs and leaky ICMs

We illustrate in Figure 1A two different types of ideal ICMs that abstract the two main mechanisms for biomolecular integral control proposed in the literature. In both types of motifs, we denote by x the species whose concentration needs to be kept at a set-point u, despite that its production rate is affected by a constant disturbance d. Therefore, the output of the process to be controlled is the concentration of species x, which is denoted by x (*italic*). The purpose of implementing an integral controller is thus to have the steady state of x, denoted by $\bar{x}$, to satisfy $\bar{x} = u$ regardless of d (i.e., x adapts perfectly to d).

In these ideal ICMs, selected controller species realize the memory function essential for integral control. Specifically, a type I ideal ICM (Figure 1A-I) regulates the expression of x using a single controller species z. This motif arises from saturating certain Hill-type or Michaelis-Menten-type kinetics such that the production and the removal rates of z become roughly proportional to u and

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

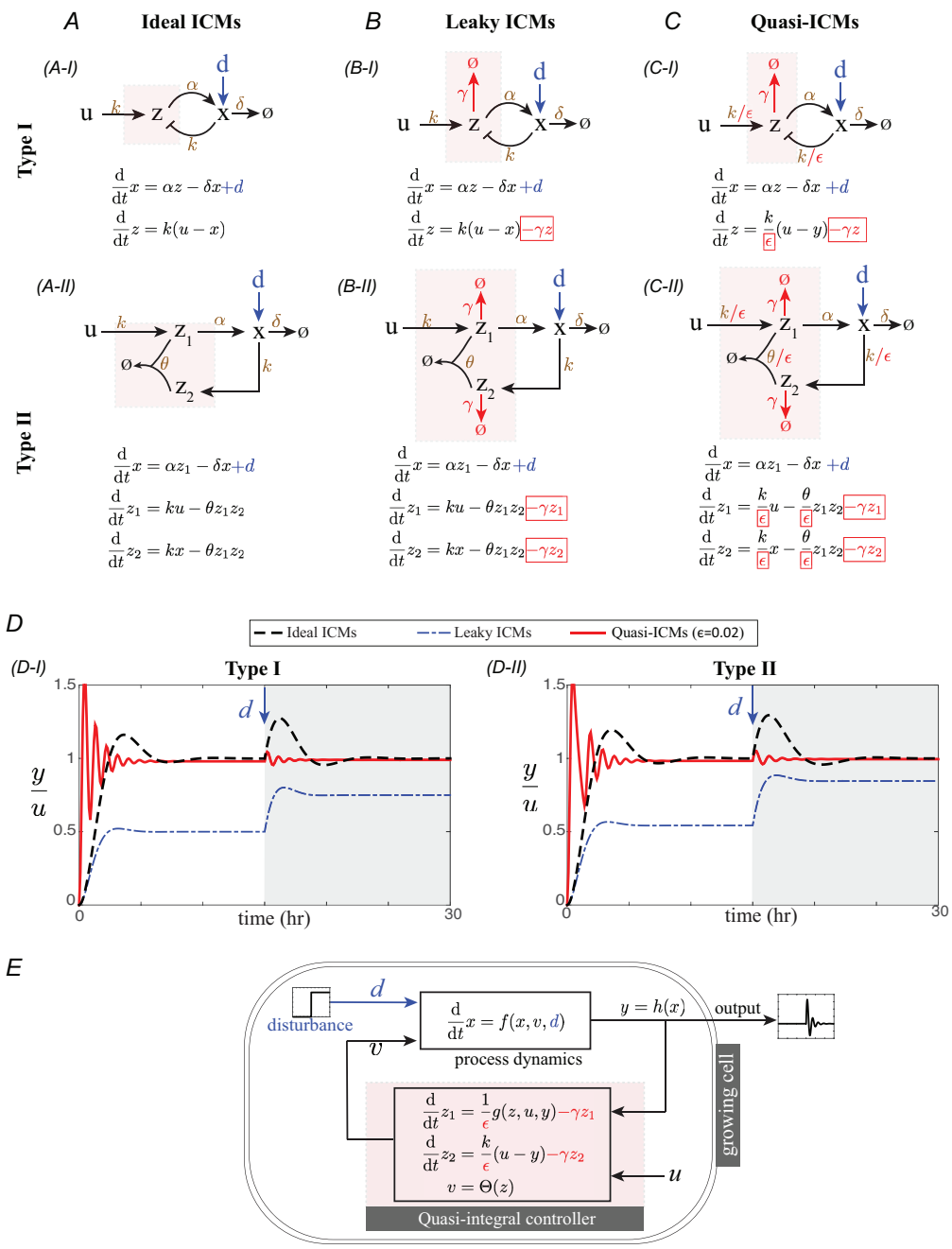

Figure 1: **Quasi-integral control mitigates the effect of leaky integration due to dilution.** (A) Two types of ideal ICMs. The controller reactions are boxed in pink, and the rest of the circuit belongs to the process to be controlled. Dilution of the controller species are neglected in ideal ICMs. (B) When dilution of the controller species are considered, ideal ICMs become leaky ICMs that cannot carry out integration. The adaptation error, defined as the extent to which the output is affected by the disturbance, can be arbitrarily large. (C) All controller reactions in a quasi-ICM are $1/\epsilon$ times faster than those in the corresponding leaky ICM, with $\epsilon \ll 1$. (D) Simulations of the type I and type II ideal, leaky and quasi-ICMs. Simulation parameters for both motifs: $\alpha = \gamma = \delta = k = 1$ hr^{-1}, $\theta = 1$ nM^{-1}·hr^{-1}, $\epsilon = 0.02$, $u = 10$ nM and $d = 5$ nM·hr^{-1}. Disturbance input d is applied at 15 hr. (E) A general ϵ-quasi-integral control system. Output of the process y becomes an input to the quasi-integral controller. Variable z_2 is the leaky memory variable, and z_1 represents the remaining controller states, if any. All controller reactions are much faster than dilution, as characterized by parameter ϵ. Output of the controller v drives the process to set-point u and adapt to d.

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

x, respectively, resulting in the following approximate mass action kinetics model for the controller species z [10, 15, 17]:

$$\frac{\mathrm{d}}{\mathrm{d}t}z = k(u - x), \tag{1}$$

where k is a positive constant called integral *gain*. Since $z = \int k(u - x)\mathrm{d}t$, z is called a *memory variable* that represents the integral of the error $(u - x)$ between the set-point and the output over time. Independent of the exact reactants implementing this ideal ICM, it is immediate from (1) that by setting the time derivative to 0, the steady state output satisfies $\bar{x} = u$ regardless of d.

With reference to Figure 1A-II, a type II ideal ICM, arising from what was called the antithetic integral controller in [12], realizes integral action using two controller species z_1 and z_2. Their production rates are engineered to be proportional to the set-point (u) and to the concentration of the output species (x), respectively. The two controller species can bind together to form a complex that is removed with rate constant θ: $\mathrm{z}_1 + \mathrm{z}_2 \xrightarrow{\theta} \emptyset$. These biomolecular processes can be described by the following mass action kinetics model:

$$\frac{\mathrm{d}}{\mathrm{d}t}z_1 = ku - \theta z_1 z_2, \qquad \frac{\mathrm{d}}{\mathrm{d}t}z_2 = kx - \theta z_1 z_2. \tag{2}$$

The memory function is carried out by the hidden memory variable $z := z_1 - z_2$, which satisfies $\mathrm{d}z/\mathrm{d}t = \mathrm{d}z_1/\mathrm{d}t - \mathrm{d}z_2/\mathrm{d}t = k(u - x)$. As a consequence, independent of the exact choice of reactants, parameters and the magnitude of disturbance d, at steady state, we must have $\bar{x} = u$. With reference to the simulation results in Figure 1D (black dashed lines), the output x of both types of ideal ICMs reaches the set-point u and adapts to disturbance d perfectly at steady state.

When dilution of the controller species due to host cell growth is taken into account, the key integral structure of the memory variables is disrupted (Figure 1B). In fact, in a standard mass action kinetics model describing reactions in exponentially growing cells, the average effect of cell growth and division on the dynamics of any species z can be modeled by a first-order additive term $-\gamma z$, where γ is the specific growth rate of the host cell [18, 19, 26, 27] (see SI Section S1.1 for details). Hence, with reference to Figure 1B-I, the dynamics of the memory variable z in a type I ICM become

$$\frac{\mathrm{d}}{\mathrm{d}t}z = k(u - x) - \gamma z. \tag{3}$$

Similarly, as shown in Figure 1B-II, the dynamics of the two controller species z_1 and z_2 in the type II ICM become

$$\frac{\mathrm{d}}{\mathrm{d}t}z_1 = ku - \theta z_1 z_2 - \gamma z_1, \qquad \frac{\mathrm{d}}{\mathrm{d}t}z_2 = kx - \theta z_1 z_2 - \gamma z_2, \tag{4}$$

resulting in the dynamics of the hidden memory variable to become

$$\frac{\mathrm{d}}{\mathrm{d}t}z = \frac{\mathrm{d}}{\mathrm{d}t}z_1 - \frac{\mathrm{d}}{\mathrm{d}t}z_2 = k(u - x) - \gamma z. \tag{5}$$

From equations (3) and (5), we observe that in both types of motifs, the memory variable z is no longer integrating the error between the set-point and the output, but rather carries out leaky integration. We therefore call the motifs in Figure 1B *leaky ICMs*. In simulations of Figure 1D (blue dash-dot lines), we demonstrate that including dilution significantly hinders the ability of leaky ICMs to achieve adaptation to disturbance d, resulting in nearly 100% relative adaptation error. More generally, as we exemplify in SI Section S1.5, for any fixed dilution rate constant (γ), depending on the process to be controlled and the parameters in the leaky ICM, the adaptation error can be arbitrarily large.

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2.2 Quasi-ICMs: almost perfect adaptation to disturbances

In this section, we introduce quasi-ICMs, the main contribution of this paper. While leaky ICMs cannot achieve perfect adaptation to disturbances, and their performance can be arbitrarily deteriorated by the presence of dilution, we propose that one can achieve almost perfect adaptation to disturbances despite dilution by increasing the rate of all controller reactions in the leaky ICMs. Therefore, we call a leaky ICM whose controller reactions are much faster than the slower dilution process a *quasi-ICM*. These motifs are shown in Figures 1C-I and C-II respectively for the two types of motifs. In particular, with reference to Figure 1C-I, we use a small dimensionless positive parameter $0 < \epsilon \ll 1$ to capture the fact that the controller reaction rates in a type I quasi-ICM are $1/\epsilon$ times faster than those in a type I leaky ICM (Figure 1B-I), resulting in the memory variable dynamics to become

$$\frac{\mathrm{d}}{\mathrm{d}t}z = \frac{k}{\epsilon}(u - x) - \gamma z. \tag{6}$$

Similarly, when all the controller reaction rates in a type II leaky ICM are increased by a factor of $1/\epsilon$ (see Figure 1C-II), the controller species dynamics become

$$\frac{\mathrm{d}}{\mathrm{d}t}z_1 = \frac{k}{\epsilon}u - \frac{\theta}{\epsilon}z_1 z_2 - \gamma z_1, \qquad \frac{\mathrm{d}}{\mathrm{d}t}z_2 = \frac{k}{\epsilon}x - \frac{\theta}{\epsilon}z_1 z_2 - \gamma z_2, \tag{7}$$

resulting in the following hidden memory variable dynamics:

$$\frac{\mathrm{d}}{\mathrm{d}t}z = \frac{\mathrm{d}}{\mathrm{d}t}z_1 - \frac{\mathrm{d}}{\mathrm{d}t}z_2 = \frac{k}{\epsilon}(u - x) - \gamma z. \tag{8}$$

As shown in Figure 1D (red solid lines), for small ϵ, both types of quasi-ICMs restore the ability to drive the output x to the set-point u, and to adapt to d almost perfectly despite dilution. These qualitative observations are reflected mathematically in equations (6) and (8), where for both motifs the steady state adaptation error e can be computed as $e = u - \bar{x} = \epsilon\gamma\bar{z}/k$, whose magnitude can be arbitrarily decreased by detuning ϵ (i.e., increasing controller reaction rates). Although this reasoning is intuitive, it is based on the implicit assumption that the steady state concentration of controller species ($\bar{z}$) stays roughly unchanged as ϵ decreases, which is, unfortunately, not always true and hard to verify in general (see SI Section S1.6 for such an example). Therefore, in the rest of this section, we provide precise and general mathematical conditions under which performance deterioration of a leaky integral controller can be arbitrarily suppressed by faster controller reactions. These results establish a general circuit design principle applicable to any biomolecular process to be controlled and to any leaky integral controller, including, but not limited to those arising from the type I and type II ICMs.

We consider a biomolecular process to be controlled, whose dynamics can be written as:

$$\frac{\mathrm{d}}{\mathrm{d}t}x = f(x, v, d), \qquad y = h(x), \tag{9}$$

where x represents process states (i.e., the concentrations of species x forming the biomoelcular process). The process takes two inputs: v is the control input produced by the controller, and d is a disturbance input, which could also represent an uncertainty in process parameters. The output of the process y is determined by a function $h(x)$. The process (9) is connected to an ϵ-parameterized biomolecular controller, which contains a leaky integral controller. We study the case where all controller reactions have an ϵ-parameterized time-scale separation from dilution. To this end, we write the controller in the general form:

$$\frac{\mathrm{d}}{\mathrm{d}t}z_1 = \frac{1}{\epsilon}g(z, u, y) - \gamma z_1, \qquad \frac{\mathrm{d}}{\mathrm{d}t}z_2 = \frac{k}{\epsilon}(u - y) - \gamma z_2, \qquad v = \Theta(z), \tag{10}$$

where $z := [z_1, z_2]^T$ represents the controller states (e.g., concentrations of controller species) and ϵ is a small positive parameter. Specifically, z_2 is the intended memory variable that carries out the

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

leaky integration with gain k, and z_1 represents concentrations of additional controller species, if any. The controller takes the output of the process y as input, and compares it with a set-point u. The *closed loop system* is therefore a feedback interconnection of process (9) and controller (10) (see Figure 1E). Assuming that the closed loop system has a *unique locally asymptotically stable steady state* $(\bar{x}, \bar{z})$, we formally define ϵ-quasi-integral control as follows.

Definition 1. *The closed loop system (9)-(10) realizes ϵ-quasi-integral control in an admissible input set $\mathbb{U} \times \mathbb{D}$ if for all $(u, d) \in \mathbb{U} \times \mathbb{D}$, the system's steady state output $\bar{y}$ satisfies*

$$\lim_{\epsilon \to 0} \bar{y}(u, d, \epsilon) = u. \tag{11}$$

Equation (11) implies that, by increasing the rate of controller reactions (i.e., decreasing ϵ), the steady state output $\bar{y}$ of an ϵ-quasi-integral control system can be made arbitrarily close to the set-point u in the presence of disturbance d. The following claim provides an easy-to-check condition for ϵ-quasi-integral control.

Claim 1. *The closed loop system (9)-(10) realizes ϵ-quasi-integral control if the auxiliary ideal integral control system consisting of the process (9) and the ideal integral controller*

$$\frac{\mathrm{d}}{\mathrm{d}t} z_1 = g(z, u, y), \qquad \frac{\mathrm{d}}{\mathrm{d}t} z_2 = k(u - y), \qquad v = \Theta(z). \tag{12}$$

has a unique locally exponentially stable steady state for all $(u, d) \in \mathbb{U} \times \mathbb{D}$.

The proof of this Claim can be found in SI Section S1.2. It establishes that if one can use the ideal integral controller (12) to obtain perfect adaptation in the absence of dilution, then when dilution is taken into account, one can always speed up all controller reactions to restore adaptation. From an engineering perspective, one should therefore select controller biomolecular processes whose rates are significantly faster with respect to dilution rates. Such processes include, for example, enzymatic reactions (i.e., phosphorylation, methylation, and ubiquitination), protein-protein interactions, and RNA-based interactions (i.e., sRNA or microRNA-enabled degradation or sequestration). With these processes, the designer can then realize type I or type II ideal ICMs as well as additional ideal ICMs that are in the form of system (12). It is worth to note that even in cases where only part of the controller reactions can be sped up, it may still be possible to reach ϵ-quasi-integral control, although the algebraic conditions to check become slightly more involved (see SI Section S1.7). In SI Section S4.3, we further demonstrate that when dilution rate constant varies slowly in time, the adaptation error is still guaranteed to decrease as ϵ is decreased.

Claim 1 provides the theoretical underpinning to the type I and type II quasi-ICM designs of Figure 1C. Since the corresponding type I and II ideal ICMs are in the form of (12) and have unique locally exponentially stable steady states (see proof in SI Section S1.3-1.4), increasing all controller reaction rates will quench the steady state effect of leaky integration. Furthermore, we also show in SI Section S1.3-1.4 that the steady states of the type I and II quasi-ICMs are unique and stable.

3 Two biomolecular realizations of quasi-ICMs

In this section, we propose physical implementations of a phosphorylation-based quasi-integral controller and a small RNA-based quasi-integral controller. They are realizations of the type I and type II quasi-ICMs, respectively, which satisfy the general design principle of fast controller reactions. For illustration purposes, we use these controllers to mitigate the effect of transcriptional and translational resource fluctuations on gene expression, a problem that has received considerable attention in synthetic biology in recent years [4, 23–25].

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3.1 Type I quasi-ICM: phosphorylation cycle

A diagram of the phosphorylation-based quasi-integral control system is shown in Figure 2A. The controller is intended to regulate the production of protein p to adapt to a disturbance d, which models a reduction in protein production rate due to, for example, depletion of transcriptional and/or translational resources in the host cell [4, 23]. In this system, the intended integral control is accomplished by a phosphorylation cycle. The regulated protein p is co-expressed with a phosphatase and the concentration of the kinase acts as the set-point u. The substrate b is expressed constitutively. When b is phosphorylated by kinase u to become active substrate b*, it transcriptionally activates production of protein p. A simplified mathematical model of this system is (see SI Section S2.1 for detailed derivation from chemical reactions):

$$\frac{\mathrm{d}}{\mathrm{d}t}b^* = k_1\frac{ub}{b+K_1} - k_2\frac{pb^*}{b^*+K_2} - \gamma b^*, \qquad \frac{\mathrm{d}}{\mathrm{d}t}p = R(1-d)\frac{b^*}{\lambda+b^*} - \gamma p, \tag{13}$$

where k_1 and k_2 are the catalytic rate constants of the phosphorylation and dephosphorylation reactions, respectively, K_1 and K_2 are the Michaelis-Menten constants, R is the protein production rate constant, and λ is the dissociation constant between b* and the promoter of the regulated gene.

System (13) can be taken to the form of a type I leaky-ICM (3) if both the kinase and the phosphatase are saturated (i.e., $b \gg K_1$ and $b^* \gg K_2$). These design constraints are present in any type I ICM, which relies on saturation of Hill-type or Michaelis-Menten kinetics [10, 15, 17], and can be satisfied in practice when the substrate is over-expressed and the kinase concentration is not too small (see SI Section S2.3 for details). Under these assumptions and by setting $x := p$, $z := b^*$ and $\sigma := k_2/k_1$, we can approximate (13) by

$$\frac{\mathrm{d}}{\mathrm{d}t}x = R(1-d)\frac{z}{\lambda+z} - \gamma x, \qquad \frac{\mathrm{d}}{\mathrm{d}t}z = k_1(u-y) - \gamma z, \qquad y = \sigma x, \tag{14}$$

which is in the form of a type I leaky-ICM. Since phosphorylation and dephosphorayltion reactions are typically much faster than dilution, by choosing a phosphorylation cycle to realize the type I motif, the effect of leaky integration is naturally mitigated. In fact, by setting $\epsilon := \gamma/k_1$ ($\epsilon \sim 10^{-3}$ in bacteria [19]), system (14) is in the form of a type I quasi-ICM:

$$\frac{\mathrm{d}}{\mathrm{d}t}x = R(1-d)\frac{z}{\lambda+z} - \gamma x, \qquad \frac{\mathrm{d}}{\mathrm{d}t}z = \frac{\gamma}{\epsilon}(u-y) - \gamma z, \qquad y = \sigma x. \tag{15}$$

As demonstrated in the simulation results in Figure 2B and Figure S8 in SI, the steady state adaptation error due to leaky integration becomes smaller as ϵ decreases (i.e., larger time-scale separation between phosphorylation/dephosphorylation and dilution). We refer the readers to SI Section S2.2 for mathematical details on the uniqueness and stability of the steady state.

3.2 Type II quasi-ICM: small RNA (sRNA) interference

The sRNA-based quasi-integral controller is a realization of the type II quasi-ICM. It is intended to regulate translation of protein p to adapt to a disturbance d, which models uncertainty in translation rate constant due to fluctuation in the amount of available ribosomes [4, 23, 24]. With reference to Figure 2C, the controller consists of protein p transcriptionally activating production of an sRNA (s) that is complementary to the mRNA (m) of the regulated protein p. The sRNA and mRNA can bind and degrade together rapidly [28–30]. The mRNA concentration m is the control input to the translation process. A constant upstream transcription factor regulates mRNA production as a set-point input u to the controller. Based on the chemical reactions in SI Section S3.1, the ODE model of this system is:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}m &= Tu - \theta ms/\beta - \delta m,\\
\frac{\mathrm{d}}{\mathrm{d}t}s &= T_s\frac{p/k_s}{1+p/k_s} - \theta ms/\beta - \delta s,\\
\frac{\mathrm{d}}{\mathrm{d}t}p &= R(1-d)m - \gamma p,
\end{aligned} \tag{16}$$

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

where T and T_s characterize the production rates of mRNA and sRNA, respectively. They are proportional to the copy numbers of the regulated gene and the sRNA-expression DNA. Parameter k_s is the dissociation constant of the binding between protein p and the sRNA promoter, β is the dissociation constant of mRNA-sRNA binding, θ is the degradation rate constant of the mRNA-sRNA complexes, and R is the translation rate constant. In addition to dilution due to cell growth, characterized by rate constant γ, uncoupled mRNA and sRNA are degraded by RNAse [31, 32]. Therefore, we model decay (i.e., dilution and degradation) of uncoupled RNAs by a lumped rate constant δ such that $\delta \geq \gamma$, and assume that this rate constant is the same for mRNA and sRNA without loss of generality. Our model (16) is similar to established sRNA-based translation regulation models [29, 33, 34], except that expression of sRNA is activated by the regulated protein (p) to constitute a feedback loop. These additional reactions are modeled using a standard Hill-type function in (16), and we refer the readers to SI Section S3.1 for its derivation.

Setting $z_1 := m$, $z_2 := s$ and $x := p$, system (16) can be taken to the form of a type II leaky-ICM:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}x &= R(1-d)z_1 - \gamma x, \\
\frac{\mathrm{d}}{\mathrm{d}t}z_1 &= Tu - \theta z_1 z_2/\beta - \delta z_1, \\
\frac{\mathrm{d}}{\mathrm{d}t}z_2 &= Ty - \theta z_1 z_2/\beta - \delta z_2, \\
y &= \frac{T_s}{T}\frac{x/k_s}{1+x/k_s}.
\end{aligned} \tag{17}$$

In order to achieve negligible adaptation error, we need to ensure that the controller reactions (i.e., z_1 and z_2 dynamics) are much faster than uncoupled RNA decay. On the one hand, coupled degradation of mRNA-sRNA complexes are by nature much more rapid than uncoupled RNA decay [30]. We therefore use $\epsilon := \delta/\theta \ll 1$ to characterize this time-scale separation. Additionally, we can simultaneously increase DNA copy numbers of the controller species m and s by a factor of $1/\epsilon$ to increase their production rates. Consequently, the production rate constants of mRNA and sRNA become T/ϵ and T_s/ϵ, respectively. Under these assumptions, system (17) is in the form of a type II quasi-ICM:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}x &= R(1-d)z_1 - \gamma x, \\
\frac{\mathrm{d}}{\mathrm{d}t}z_1 &= \frac{T}{\epsilon}u - \frac{\theta}{\epsilon\beta}z_1 z_2 - \delta z_1, \\
\frac{\mathrm{d}}{\mathrm{d}t}z_2 &= \frac{T}{\epsilon}y - \frac{\theta}{\epsilon\beta}z_1 z_2 - \delta z_2, \\
y &= \frac{T_s}{T}\frac{x/k_s}{1+x/k_s}.
\end{aligned} \tag{18}$$

The adaptation error can be arbitrarily decreased by detuning ϵ. In practice, this can be accomplished by simultaneously (i) increasing the mRNA-sRNA complex degradation rate constant θ, and (ii) increasing the copy numbers of the regulated gene and the sRNA-expression DNA. While directly increasing θ may be difficult to implement in practice, since the parameters θ and β are clustered together in model (16), we can achieve the same effect by increasing the affinity between sRNA and mRNA ($1/\beta$) [35]. The above results are confirmed by simulations in Figure 2D and Figure S9 in SI using biologically relevant parameters from bacteria *E. coli*. In SI Section S3.2, we analytically demonstrate that the conditions in Claim 1 are indeed satisfied for system (18).

4 Discussion

Due to integral controllers' unique ability to drive a process to reach a set-point regardless of disturbances/uncertainties, they are especially appealing to synthetic biology, where various forms

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

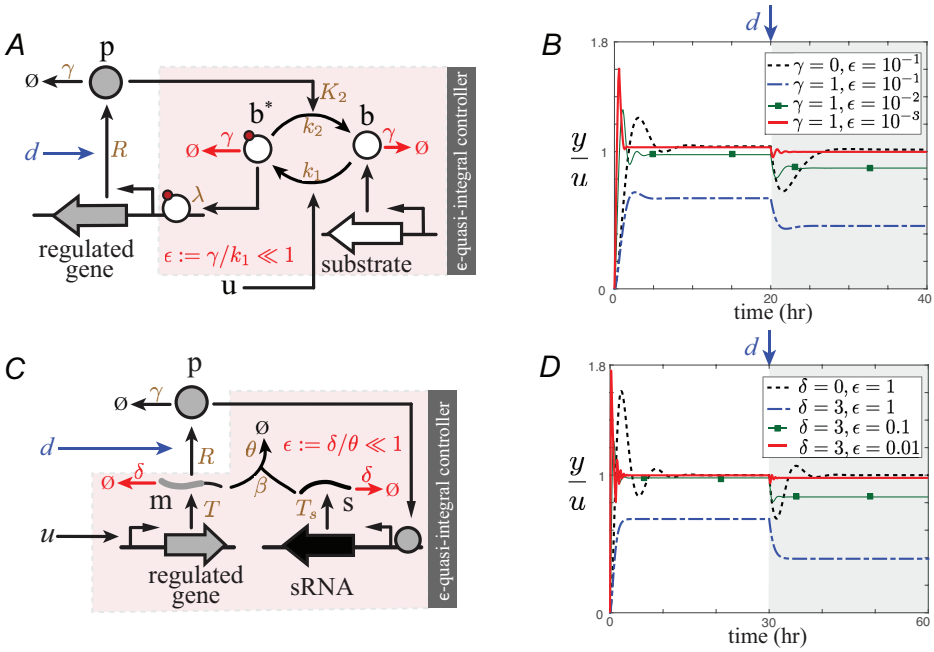

Figure 2: **Two physical realizations of quasi-ICMs.** (A) Genetic circuit diagram of the phosphorylation-based quasi-integral controller. Chemical reactions realizing the controller are boxed in pink. (B) Simulation of the circuit's response according to (13). A set-point input $u = 20$ nM is applied at time 0 and a disturbance input $d = 0.5$ is applied at 20 hr. The vertical axis represents the ratio between output y, defined to be proportional to p ($y = \sigma p$), and set-point u. The dashed black line is the response of the phosphorylation-based control system assuming no dilution of the active substrate b^*. The dotted blue line, the thin green line with square markers and the solid red line represent circuit's response in the presence of nonzero substrate dilution ($\gamma = 1$ hr^{-1}) and decreasing ϵ, which corresponds to increasing catalytic rates (k_i, $i = 1, 2$). (C) Genetic circuit diagram of the sRNA-based quasi-integral controller. (D) Simulation of the circuit's response according to (16). A set-point input $u = 1$ is applied at time 0 and a disturbance input $d = 0.5$ is applied at 30 hr. The vertical axis represents the ratio between output y, defined in (18), and set-point input u. The dashed black line represents response of an ideal integral control system, where RNA decay rate δ is set to 0. The dotted blue line, the thin green line with square markers and the solid red line represent circuit's responses in the presence of nonzero RNA decay rate ($\delta = 3$ hr^{-1}, corresponding to half-life of about 13 mins) and decreasing ϵ. Parameter ϵ is decreased by increasing the mRNA-sRNA removal rate (θ/β). The DNA copy numbers of the regulated gene and the sRNA are increased simultaneously by a factor of $1/\epsilon$ as ϵ decreases. Simulation parameters are listed in SI Section S5, Table 2.

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of disturbances and uncertainties limit the robustness and predictability of synthetic circuits. A unique and fundamental challenge to realize integral control in living cells is that the concentration of biomolecules dilute as cells grow, resulting in the impossibility to create a non-leaky biomolecular memory element, which is indispensable to implement integral control.

In this report, we propose a general design principle to overcome this obstacle. We mathematically demonstrate that one can design an ideal integral controller to obtain perfect adaptation in the absence of dilution and then speed up all controller reactions with respect to dilution to restore adaptation in the presence of dilution. This implies that the designer should select controller biomolecular processes whose rates are significantly faster than dilution rates. Our result is equally applicable to situations where controller species are degraded in addition to being diluted. Due to the asymptotic nature of our results, they are largely insensitive to system parameters. To solidify this point, we simulated the phosphorylation-based and the sRNA-based quasi-integral controllers with randomly generated parameters, and demonstrated that, regardless of the rest of the parameters, the adaptation error always decreases as the rates of controller reactions are increased (SI Section S5.2).

While our mathematical analyses are carried out using ordinary differential equation models derived from mass action kinetics, we numerically exemplify their effectiveness using a time-varying Gillespie simulation [36], in which the stochasticity of biomolecular reactions as well as cell volume expansion and division are explicitly taken into account (see SI Section S5.3 for details). In addition, we have assumed that the host cells grow at constant rate (constant γ). In practice, however, their growth rate may change with, for example, temperature, nutrients and metabolic load of the circuit [37]. We have therefore studied the performance of quasi-integral controllers in changing growth conditions (see SI Section S4.3), and found that the circuit's output can still adapt to disturbances as long as growth rate changes slowly.

Nature may have already utilized the time-scale separation strategy between process and controller that we propose here for realizing adaptation and homeostasis. For example, it has been suggested that the bacterial chemotaxis system [38] and the yeast osmoregulation system [39] contain integral controllers realized through methylation and phosphorylation, respectively. Since these processes are much faster than dilution in their respective host cells, it is justifiable to neglect dilution in these systems, and almost perfect adaptation can be achieved. Therefore, experimental study of synthetic realizations of the quasi-integral controllers proposed here may enhance our understanding of their natural counterparts, which are often embedded in more complex biomolecular networks and are therefore more difficult to analyze.

From a classical control-theoretic perspective, integral controllers often operate with low integral gains (i.e., slow integration). This is because with a perfect integrator, steady state adaptation can be achieved regardless of the magnitude of the integral gain, and a high integral gain often leads to undesirable effects such as higher energy consumption and tendency to instability [6, 22]. However, a potential advantage of increasing the leaky integral gain is that it can improve a system's ability to track (reject) time-varying references (disturbances). While we demonstrate this capability for the two example systems through linearization and simulations in SI Section S4.1-4.2, further theoretical assessment of this property requires deeper study [40, 41].

Data accessibility

All simulation codes are provided as Electronic Supplementary Material.

Author contributions

Y.Q derived the mathematical model and performed analyses and simulations; Y.Q and D.D.V wrote the manuscript.

Competing interests

We have no competing interests.

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Funding

This work is supported by AFOSR grant FA9550-14-1-0060.

Acknowledgement

We thank Cameron McBride for proof-reading the manuscript; Prof. Mustafa Khammash, Dr. Corentin Briat, Dr. Hsin-Ho Huang, Ross Jones and Theodore W. Grunberg for helpful discussions and suggestions; and Prof. Elisa Franco and Dr. Christian Cuba Samaniego for comments on earlier versions of the manuscript. We thank the anonymous reviewers for their constructive comments.

References

[1] Priscilla E. M. Purnick and Ron Weiss. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.*, 10:410–422, 2009.

[2] Stefano Cardinale and Adam Paul Arkin. Contextualizing context for synthetic biology- identifying causes of failure of synthetic biological systems. *Biotechnol. J.*, 7:856–866, 2012.

[3] H. Giladi, D. Goldenberg, S. Koby, and A. B. Oppenheim. Enhanced activity of the bacteriophage lambda PL promoter at low temperature. *FEMS Microbiol. Rev.*, 17(1-2):135–140, 1995.

[4] Yili Qian, Hsin-Ho Huang, José I. Jiménez, and Domitilla Del Vecchio. Resource competition shapes the response of genetic circuits. *ACS Synth. Biol.*, 6(7):1263–1272, 2017.

[5] Enoch Yeung, Aaron J. Dy, Kyle B. Martin, Andrew H. Ng, Domitilla Del Vecchio, James L. Beck, James J. Collins, and Richard M. Murray. Biophysical constraints arising from compositional context in synthetic gene networks. *Cell Systems*, 5(1):11–24.e12, 2017.

[6] K. J. Ăström and R. M. Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, 2008.

[7] B.A. Francis and W.M. Wonham. The Internal Model Principle of Control Theory The Internal Model Principle of Control Theory. *Automatica*, 12:457–465, 1976.

[8] Eduardo D. Sontag. Adaptation and regulation with signal detection implies internal model. *Syst. Control Lett.*, 50(2):119–126, 2003.

[9] Domitilla Del Vecchio, Aaron J. Dy, and Yili Qian. Control theory meets synthetic biology. *J. R. Soc. Interface*, 13(120):20160380, 2016.

[10] Jordan Ang, Sangram Bagh, Brian P. Ingalls, and David R. McMillen. Considerations for using integral feedback control to construct a perfectly adapting synthetic gene network. *J. Theor. Biol.*, 266(4):723–738, 2010.

[11] Jordan Ang and David R. McMillen. Physical constraints on biological integral control design for homeostasis and sensory adaptation. *Biophys. J.*, 104(2):505–515, 2013.

[12] Corentin Briat, Ankit Gupta, and Mustafa Khammash. Antithetic integral feedback ensures robust perfect adaptation in noisy bimolecular networks. *Cell Systems*, 2(1):15–26, 2016.

[13] Corentin Briat, Christoph Zechner, and Mustafa Khammash. Design of a synthetic integral feedback circuit: Dynamic analysis and DNA implementation. *ACS Syn. Biol.*, 5(10):1108–1116, 2016.

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[14] Xinying Ren, Ania-Ariadna Baetica, Anandh Swaminathan, and Richard M Murray. Population regulation in microbial consortia using dual feedback control. In *Proceedings of the 56th IEEE Conference on Decision and Control*, Melbourne, Australia, 2017.

[15] T. Drengstig, X. Y. Ni, K. Thorsen, I. W. Jolma, and P. Ruoff. Robust adaptation and homeostasis by autocatalysis. *J. Phys. Chem. B*, 116(18):5355–5363, 2012.

[16] Zhe F. Tang and David R. McMillen. Design principles for the analysis and construction of robustly homeostatic biological networks. *J. Theor. Biol.*, 408:274–289, 2016.

[17] E. Klavins. Proportional-integral control of stochastic gene regulatory networks. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 2547–2553, Atlanta, GA, 2010.

[18] Uri Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC Press, 2006.

[19] Domitilla Del Vecchio and Richard M. Murray. *Biomolecular Feedback Systems*. Princeton University Press, Princeton, 2014.

[20] P. P. Peralta-Yahya, F. Zhang, S. B. del Cardayre, and J. D. Keasling. Microbial engineering for the production of advanced biofuels. *Nature*, 488:320–328, 2012.

[21] Christian Cuba Samaniego and Elisa Franco. An ultrasensitive biomolecular network for robust feedback control. In *Proceedings of the 20th World Congress of International Federation of Automatic Control (IFAC)*, pages 11438–11443, July 2017.

[22] Noah Olsman, Ania-Ariadna Baetica, Fangzhou Xiao, Yoke Peng Leong, John Doyle, and Richard Murray. Hard limits and performance tradeoffs in a class of sequestration feedback systems. *bioRxiv*, 2017.

[23] Andras Gyorgy, José I. Jiménez, John Yazbek, Hsin-Ho Huang, Hattie Chung, Ron Weiss, and Domitilla Del Vecchio. Isocost lines describe the cellular economy of gene circuits. *Biophys. J.*, 109(3):639–646, 2015.

[24] Alon Raveh, Michael Margaliot, Eduardo D Sontag, and Tamir Tuller. A model for competition for ribosomes in the cell. *J. R. Soc. Interface*, 13:20151062, 2015.

[25] Tatenda Shopera, Lian He, Tolutola Oyetunde, Yinjie J. Tang, and Tae Seok Moon. Decoupling resource-coupled gene expression in living cells. *ACS Synth. Biol.*, 6(8):1596–1604, 2017.

[26] J Monod, A.M Pappenheimer, and G Cohen-Bazire. La cinétique de la biosynthèse de la β-galactosidase chez e. coli considérée comme fonction de la croissance. *Biochimica et Biophysica Acta*, 9:648–660, 1952.

[27] Nitzan Rosenfeld, Michael B Elowitz, and Uri Alon. Negative autoregulation speeds the response times of transcription networks. *J. Mol. Biol.*, 323(5):785–793, 2002.

[28] Seung Min Yoo, Dokyun Na, and Sang Yup Lee. Design and use of synthetic regulatory small RNAs to control gene expression in Escherichia coli. *Nat. Protocols*, 8(9):1694–1707, 2013.

[29] Erel Levine, Zhongge Zhang, Thomas Kuhlman, and Terence Hwa. Quantitative characteristics of gene regulation by small RNA. *PLoS Biol.*, 5(9):e229, 2007.

[30] Eric Massé, Freddy E. Escorcia, and Susan Gottesman. Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev.*, 17(19):2374–2383, 2003.

[31] George A. Mackie. RNase E: at the interface of bacterial RNA processing and decay. *Nat. Rev. Micro.*, 11(1):45–57, 2013.

bioRxiv preprint doi: https://doi.org/10.1101/141051; this version posted February 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[32] Razika Hussein and Han N. Lim. Direct comparison of small RNA and transcription factor signaling. *Nucleic. Acids. Res.*, 40(15):7269–7279, 2012.

[33] Pankaj Mehta, Sidhartha Goyal, and Ned S. Wingreen. A quantitative comparison of sRNA-based and protein-based gene regulation. *Mol. Syst. Biol.*, 4:221, 2008.

[34] Yishai Shimoni, Gilgi Friedlander, Guy Hetzroni, Gali Niv, Shoshy Altuvia, Ofer Biham, and Hanah Margalit. Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol. Syst. Biol.*, 3, 2007.

[35] Dokyun Na, Seung Min Yoo, Hannah Chung, Hyegwon Park, Jin Hwan Park, and Sang Yup Lee. Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. *Nat. Biotechnol.*, 31:170–174, 2013.

[36] T. Lu, D. Volfson, L. Tsimring, and J. Hasty. Cellular growth and division in the gillespie algorithm. *Systems Biology*, 1(1):121–128, 2004.

[37] Stefan Klumpp, Zhongge Zhang, and Terence Hwa. Growth-rate dependent global effect on gene expression in bacteria. *Cell*, 139:1366–1375, 2009.

[38] Tau-Mu Yi, Yun Huang, Melvin I. Simon, and John Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl. Acad. Sci.*, 97(9):4649–4653, 2000.

[39] Dale Muzzey, Carlos A. Gómez-Uribe, Jerome T. Mettetal, and Alexander van Oudenaarden. A Systems-Level Analysis of Perfect Adaptation in Yeast Osmoregulation. *Cell*, 138(1):160–171, 2009.

[40] Phillip Rivera-Ortiz and Domitilla Del Vecchio. Integral action with time scale separation: A mechanism for modularity in biological systems. In *Proceedings of the 53rd IEEE Conference on Decision and Control*, pages 49–55, Los Angeles, CA, 2014.

[41] Petar Kokotović, Hassan K. Khalil, and John O'Reilly. *Singular Perturbation Methods in Control: Analysis and Design*. SIAM, Philadelphia, PA.