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

Feedback control systems with integral action have the unique ability to perfectly reach constant set-points and to reject constant disturbances. Due to these properties, they are ubiquitous in engineering systems and frequently found in natural biological systems that achieve homeostasis and adaptation. Recently, there has been increasing interest in realizing integral controllers in the synthetic biology community as a way to improve robustness of circuits and tightly regulate gene expression. Although a number of circuit designs have been proposed, their functionality often hinges on the critical assumption that species in the controller do not dilute, since dilution breaks the controller integration structure and leads to "leaky integration". While this assumption is satisfied in cell-free systems, it is not met in implementations in living cells, where cell growth and division dictates species dilution. In this report, we abstract previously proposed designs of biomolecular integral controllers into two ideal integral control motifs (IICM), type I and type II. Based on these, we mathematically demonstrate how engineering a time-scale separation between the controller reactions and dilution, we can obtain quasi-integral control motifs (qICM) that recover almost perfect adaptation even in the presence of dilution. Contrary to previous hypothesis, our results show that implementing a fast integral action alone may not be sufficient to recover the adaptation property. Our results are based on a general dynamical model of a "plant" and a "controller" for which we provide easy-to-check algebraic conditions for almost perfect adaptation. These conditions can be used as guidance for design. Here, we propose two bimolecular implementations of qICMs both designed to reach adaptation of gene expression to fluctuations in cellular resources.

1 Introduction

Integral feedback control is a widely applied engineering principle that ensures the output of a system can reach a constant set-point in the presence of constant external disturbances or parameter uncertainties [1, 2]. This principle is implemented by many biological systems that demonstrate homeostasis and adaptation, key properties that ensure robustness to changing or uncertain environments. For instance, integral feedback motifs have been identified in models of bacterial chemotaxis [3], calcium homeostasis [4], and yeast osmoregulation [5] (see [6] for a comprehensive review). In fact, according to the Internal Model Principle of control theory [7], a linear system must contain an

integrator to achieve perfect homeostasis/adaptation. This principle has been extended to classes of nonlinear systems under mild technical conditions [2].

In a standard integral feedback control setup, a "memory variable" z integrates the error between the constant set-point u and the process output y to generate a control input v to the process (Figure 1A). The process dynamics are subject to constant external disturbances (or parameter uncertainties) d. This simple integral feedback control mechanism can be modeled by the following ordinary differential equations (ODEs):

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}x &= f(x,v,d), \quad y = h(x), \\ \frac{\mathrm{d}}{\mathrm{d}t}z &= k(u-y), \quad v = \Theta(z), \end{aligned} \tag{1}$$

where k is the *integral gain*, and $f(\cdot)$, $h(\cdot)$ and $\Theta(\cdot)$ represent the process dynamics, the process output function and the controller output function, respectively. Assuming the existence of a *unique stable* steady state, it is immediate from (1) that by setting the time derivatives to 0, the steady state output $\bar{y}$ must satisfy $\bar{y} = u$ regardless of d. This implies that the steady state output *reaches* set-point u and *adapts* perfectly to disturbance d.

Due to this robustness property, integral control has raised much attention in the synthetic biology community in recent years [8, 9, 10, 11, 12, 13, 14, 15]. In particular, a physical constraint that has been widely acknowledged to be a challenge for biomolecular realizations of integral controllers in living cells is that concentrations of all species dilute due to cell growth [6, 9, 10, 13, 16]. In fact, with reference to (1), if the memory variable z represents the concentration of a biochemical species, then the integration dynamics $\mathrm{d}z/\mathrm{d}t = k(u-y)$ become

$$\frac{\mathrm{d}}{\mathrm{d}t}z = k(u-y) - \gamma z, \tag{2}$$

where γ represents the dilution rate constant, determined by the specific growth rate of the host cell. Therefore, z is no longer integrating the error $(u-y)$, instead, it implements a *leaky integration*.

Similar problems exist in electronics. Leaky capacitors may cause catastrophic system level malfunctions, which are often solved by replacing leaky capacitors with higher quality ones that have less leakiness (smaller γ) [17]. However, in most biological contexts, it is generally undesirable to reduce γ because this requires inhibiting host cell growth. This is especially the case in certain synthetic biology applications, such as in biofuel production, where fast host cell growth is advantageous [18]. Therefore, how to overcome the integration leakiness arising from host cell growth is a fundamental issue facing the implementation of integral control in living cells.

Previous studies have proposed to overcome this leaky integration problem in a few directions [9, 10, 16]. One direction, as proposed in [10] and [16], is motivated by the fact that if dilution rate $(-\gamma z)$ remains sufficiently small over all time, then the trajectory of a system with leaky integral feedback (2) would be close to that with ideal integral feedback (1). However, this condition would require z, which represents the integral of the error $(u-y)$, to be sufficiently small over all time, resulting in severe restrictions on the range of set-point/disturbance input that the system can reach/adapt to [10, 16]. Another approach proposed in [9] is to counteract dilution $(-\gamma z)$ by engineering an additional first-order production dynamics $(k_a z)$ in the controller that effectively cancels dilution. While this approach works in theory, it relies on exact parameter matching $k_a = \gamma$, which is often impractical to realize experimentally.

It has been also suggested that dilution may be neglected if the integral action is implemented by reactions that are much faster than dilution [6, 9, 10]. However, no mathematical analysis or proposed implementation has appeared leveraging this time-scale separation. Additionally, as we demonstrate here, fast integral action alone may not be sufficient to mitigate the effect of leaky integration.

In this paper, we take this time-scale separation approach. Specifically, we mathematically demonstrate how near perfect adaptation can be achieved utilizing the time-scale separation between

dilution and the reactions that implement the controller, given that a set of easy-to-check algebraic conditions on system dynamics are satisfied. These results provide guidance for the selection of core biomolecular processes that implement the controller in order to overcome leaky integration problems in previously proposed *ideal integral control motifs* (IICMs). We propose two physical realizations of quasi-integral controllers (Section 3), which can be applied to mitigate the effect of cellular resource fluctuation on the expression of the regulated gene. We briefly discuss the practical implications of our results from both the control-theoretic and the biological perspective (Section 4).

2 Quasi-integral control with a leaky integrator

In Section 2.1, we first introduce two types of IICMs, which implement ideal integral feedback actions (Figure 1A) and are guaranteed to reach perfect adaptation in the absence of dilution. For both types of IICMs, we then introduce their corresponding *quasi-integral control motifs* (qICMs), which include dilution (Figure 1B) but guarantee almost perfect adaptation. These qICMs fit into a general system structure that we analyze in Section 2.2, providing sufficient mathematical conditions to realize quasi-integral control.

2.1 Two types of integral control motifs

As a motivation to the general controller structure we propose in this paper, we first consider two different types of integral control motifs (Figure 1C). We refer to the motifs that assume no dilution of the controller species as IICMs. They are shown in the left column in Figure 1C. Type I and type II IICMs represent abstractions of previously proposed ideal biomolecular integral control systems [9, 11, 15]. In these motifs, we have assumed the process dynamics to be the production and dilution of species x subject to an additive external disturbance d. This simplification is for illustration purposes only and our conclusions are independent of it. A type I IICM regulates the dynamics of x using a single controller species z, whose concentration z is the memory variable. This motif arises from saturating certain Hill-type or Michaelis-Menten-type kinetics such that the production rate of z becomes approximately proportional to the error $(u - y)$ [9, 15]. A type II IICM, arising from what was called the antithetic integral controller [11], realizes integral action using two controller species z_1 and z_2, whose production rates are proportional to the set-point (u) and the concentration of the output species (x), respectively. The two controller species then annihilate when bound together. The integral action is carried out by the "hidden" memory variable $\tilde{z} := z_1 - z_2$, which satisfies $\mathrm{d}\tilde{z}/\mathrm{d}t = k(u - x)$. With reference to Figure 1D (black dashed lines), in the absence of dilution of the controller species, both types of IICMs ensure set-point regulation and adaptation (to disturbances).

However, when dilution of the controller species is included in these motifs as a consequence of cell growth, the integration dynamics disappear (central column in Figure 1C). We call the corresponding motifs *leaky integral control motifs* (LICMs), since for both types of motifs, the dynamics of their memory variables take the leaky integration form $\mathrm{d}z/\mathrm{d}t = k(u-x)-\gamma z$. Dilution of controller species significantly hinders the capabilities of these motifs to achieve set-point regulation and adaptation (Figure 1D, blue dash-dot lines).

While LICM do not have the ability to achieve set-point regulation and adaptation, an ϵ-parameterized time-scale separation between dilution and controller reactions can suppress the effect of leaky integration. We call motifs with such a property ϵ-qICMs. These motifs are shown in the right column in Figure 1C. By picking a small ϵ, the time-scale of the controller dynamics, captured by the z variables, becomes much faster than dilution dynamics. As shown in Figure 1D (red solid line), for ϵ small, the ϵ-qICMs recover the ability to achieve set-point regulation and adaptation. The theoretical underpinning of type I and II qICMs is based on a general fast controller structure that satisfies mild well-posedness and easy-to-check conditions, which we introduce in the next section.

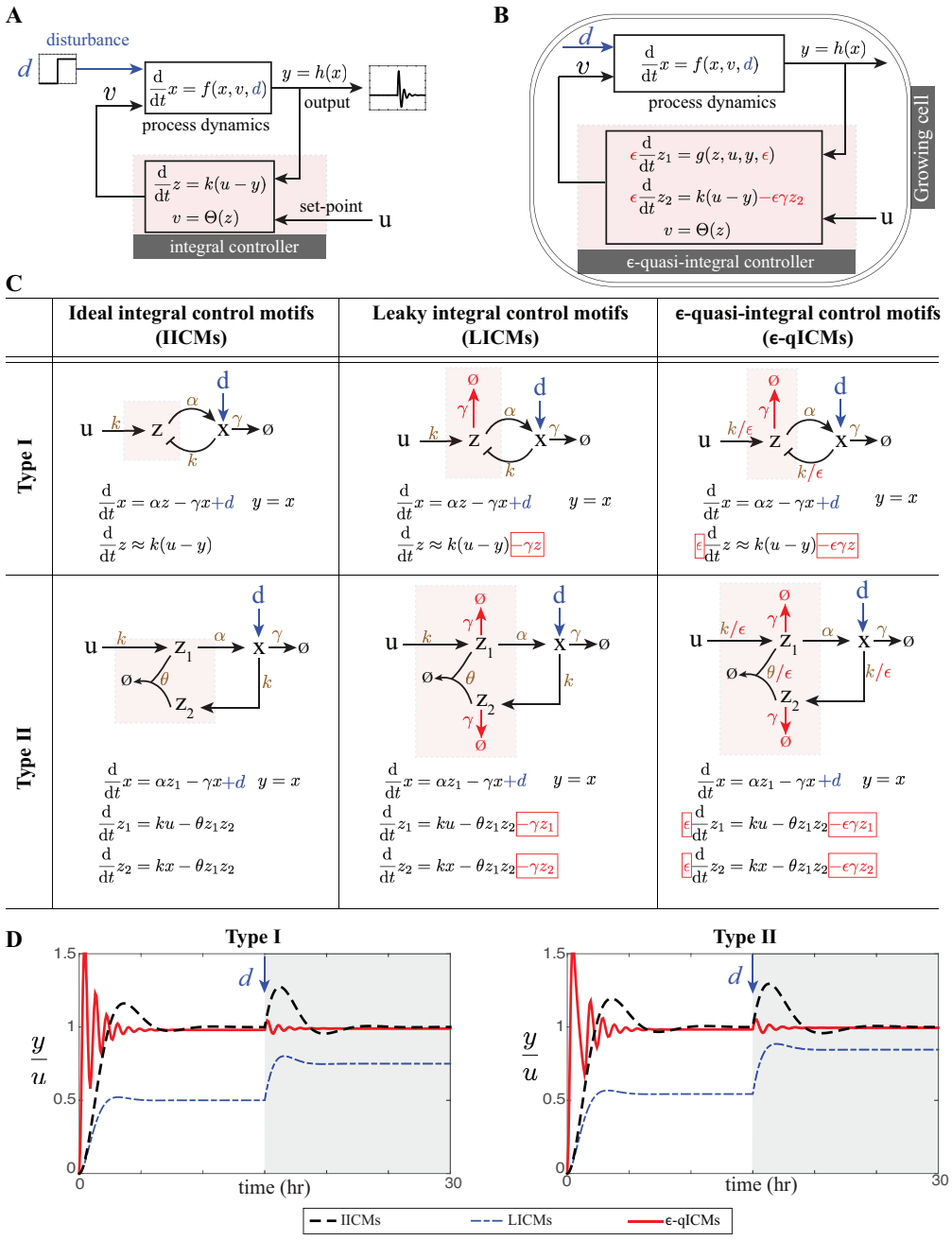

Figure 1: **Quasi-integral control mitigates the effect of leaky integration due to dilution.** (A) Schematic of an integral control system. With reference to (1), the controller integrates the error between set-point u and output y with gain k. The steady state output is independent of disturbance input d. (B) Schematic diagram of an ϵ-quasi-integral control system, where the controller species dilute due to cell growth. Variable z_2 represents the memory variable that integrates the error in the controller, and z_1 represents the remaining controller states, if any (see equations (3)-(4)). There is an ϵ-parameterized time-scale separation between dilution and the rest of the controller dynamics. (C) Two types of integral control motifs. The IICMs neglect dilution of the controller species, which are boxed in pink. When dilution of controller species is considered, the integral structure breaks and the motifs become LICMs. The ϵ-qICMs mitigate the leaky integration effect by introducing an ϵ-parameterized time-scale separation between dilution and all controller reactions. (D) Simulations of type I and type II IICMs, LICMs and ϵ-qICMs. Simulation parameters for both motifs: $\alpha = \gamma = k = 1$ hr^{-1}, $\theta = 1$ nM^{-1}·hr^{-1}, $\epsilon = 0.02$, $u = 10$ nM and $d = 5$ nM·hr^{-1}. Set-point input u is applied at time 0 and disturbance input d is applied at 15 hr.

2.2 Sufficient conditions for ϵ-quasi-integral control

In this section, we show that a general controller structure can capture the key design principles of type I and type II ϵ-qICMs, which offers a broad framework useful to uncover additional ϵ-qICMs.

We consider a biomolecular process to be controlled Σ_p, with dynamics that can be written as:

$$\Sigma_p: \qquad \frac{\mathrm{d}}{\mathrm{d}t}x = f(x,v,d), \qquad y = h(x), \tag{3}$$

where x represents process states (i.e. the concentration of species x forming the biomoelcular process). Their dynamics are described by $f(\cdot)$. The process takes two inputs: v is the control input, and d is a constant external disturbance input, which could also represent a constant uncertainty in process parameters. The output of the process y is determined by a function $h(x)$.

The process Σ_p is connected to an ϵ-parameterized biomolecular controller Σ_c^ϵ, which contains a leaky integral action due to dilution. We study the case where all controller reactions have an ϵ-parameterized time-scale separation with dilution. To this end, we write the controller in the general form:

$$\Sigma_c^\epsilon: \qquad \epsilon\frac{\mathrm{d}}{\mathrm{d}t}z_1 = g(z,u,y,\epsilon) \qquad \epsilon\frac{\mathrm{d}}{\mathrm{d}t}z_2 = k(u-y) - \epsilon\gamma z_2, \qquad v = \Theta(z), \tag{4}$$

where $z := [z_1, z_2]^T$ represents the controller states and ϵ is a small positive parameter. Specifically, z_2 is the memory variable that carries out the leaky integration, z_1 represents concentrations of additional controller species, if any. In Figure 1C, for the type I ϵ-qICM, $z_2 := z$ and there is no z_1, while for the type II ϵ-qICM, we set $z_2 := \tilde{z}$. The controller takes the output of the process y as input and compares it with a constant external set-point input u. The error is amplified by gain k/ϵ. The *closed loop system* Σ^ϵ is therefore a feedback interconnection of Σ_p and Σ_c^ϵ (Figure 1B).

We would like the steady state output $\bar{y}$ of Σ^ϵ to approach the set-point u as parameter ϵ decreases, regardless of d. This formulation has similarities with existing mathematical formulation of approximate adaptation developed in the control theory literature [19]. However, since the results in [19] require an integrator to appear in the controller dynamics when $\epsilon = 0$, they are inapplicable to our model (4). Assuming that Σ^ϵ has a *unique locally asymptotically stable* steady state $(\bar{x}, \bar{z})$, we define ϵ-quasi-integral control as follows.

Definition 1. *System Σ^ϵ 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 $\bar{y}$ is such that*

$$\lim_{\epsilon\to 0}\bar{y}(u,d,\epsilon) = u. \tag{5}$$

The unique steady state $(\bar{x},\bar{z})$ of Σ^ϵ can be computed from the following algebraic equations:

$$F(\bar{x},\bar{z},d) := f(\bar{x},\Theta(\bar{z}),d) = 0, \quad G(\bar{x},\bar{z},u,\epsilon) := g(\bar{z},u,h(\bar{x}),\epsilon) = 0, \quad k[u-h(\bar{x})] = \epsilon\gamma\bar{z}_2, \tag{6}$$

in which we assume that functions $F(\cdot)$, $G(\cdot)$ and $h(\cdot)$ are sufficiently smooth in their arguments (see SI Section S1 for precise mathematical requirements). The following claim provides sufficient conditions for the steady state of Σ^ϵ to satisfy (5).

Claim 1. *Let $(\bar{x},\bar{z})$ be the unique locally asymptotically stable steady state of Σ^ϵ, system Σ^ϵ realizes ϵ-quasi-integral control in $\mathbb{U}\times\mathbb{D}$ if the following conditions are satisfied for all $(u,d)\in\mathbb{U}\times\mathbb{D}$:*

(C1) There exists a unique solution (x^*,z^*) to (6) when $\epsilon = 0$, that is,

$$F(x^*,z^*,d) = 0, \qquad G(x^*,z^*,u,0) = 0, \qquad u - h(x^*) = 0.$$

(C2) The matrix

$$\left.\begin{bmatrix} \partial F/\partial x & \partial F/\partial z_1 & \partial F/\partial z_2 \\ \partial G/\partial x & \partial G/\partial z_1 & \partial G/\partial z_2 \\ -k\cdot\partial h/\partial x & 0 & 0 \end{bmatrix}\right|_{(x=x^*,z=z^*,\epsilon=0)}$$

is nonsingular.

The proof of this claim is a direct application of the Implicit Function Theorem [20] (see SI Section S1). Condition (C1) ensures the existence of a unique closed loop steady state (x^*, z^*) that allows the output to reach u and adapt to d in the limit $\epsilon = 0$. Condition (C2) further guarantees that for nonzero ϵ, the steady state $(\bar{x}, \bar{z})$ is close to (x^*, z^*).

We can apply Claim 1 to the ϵ-qICMs in Figure 1C and verify that (C1) and (C2) are indeed satisfied for both motifs (see SI Section S1.2). From an engineering perspective, as long as (C1) and (C2) are satisfied, one may consider increasing the rate constants of all controller reactions in any IICM by a factor of $1/\epsilon$ to realize an ϵ-qICM in the presence of dilution. In the next section, we propose two physical implementations of type I and II ϵ-qICMs.

3 Two biomolecular implementations

In this section, we propose physical implementations of an sRNA-based quasi-integral controller and a phosphorylation-based quasi-integral controller. They can be viewed as realizations of the type II and type I qICMs, respectively. We intend to use these controllers to mitigate the effect of cellular resource competition on gene expression, which can create significant unintended interactions among genes [21, 22].

3.1 sRNA-based quasi-integral controller

With reference to Figure 2A, the sRNA-based quasi-integral controller regulates translation of protein p to adapt to a disturbance input d. The disturbance input $d \in [0, 1)$ models a reduction in the translation rate constant arising from, for example, reduced amount of ribosomes translating p as more ribosomes are sequestered to produce other proteins in the host cell [21, 22]. The controller consists of 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 [23, 24, 25]. 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 S2, a simplified ODE model of this system is:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}m &= Tu - \delta m - \theta m s/\beta, \\ \frac{\mathrm{d}}{\mathrm{d}t}s &= T_s \frac{p/k_s}{1 + p/k_s} - \delta s - \theta m s/\beta, \\ \frac{\mathrm{d}}{\mathrm{d}t}p &= R(1-d)m - \gamma p, \end{aligned} \tag{7}$$

where T and T_s are the transcription rate constants to produce mRNA and sRNA, respectively. They are respectively proportional to the copy numbers of the regulated gene and the sRNA. Parameter k_s is the binding dissociation constant 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 per mRNA copy. In addition to dilution due to cell growth, characterized by rate constant γ, uncoupled mRNA and sRNA are degraded by RNAse [26, 27]. 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.

When decay of uncoupled RNAs is neglected ($\delta \equiv 0$), dynamics of the memory variable $z_2 := m - s$ is an integration of the error $(u - y)$. In particular, we can find $\mathrm{d}z_2/\mathrm{d}t = \mathrm{d}m/\mathrm{d}t - \mathrm{d}s/\mathrm{d}t = T(u - y)$, in which we take

$$y := \frac{T_s}{T} \frac{p/k_s}{1 + p/k_s}, \tag{8}$$

to be the output of the system. In this case, system (7) is an ideal integral control system that guarantees the steady state output to satisfy $\bar{y} = u$ regardless of d. According to (8), since protein concentration p is uniquely determined by y according to $p = Tyk_s/(T_s - Ty)$, p adapts to d. However, this integration structure breaks in the presence of nonzero δ, in which case, the z_2 dynamics become

$$\frac{\mathrm{d}}{\mathrm{d}t}z_2 = T\left(u - \frac{T_s}{T}\frac{p/k_s}{1+p/k_2}\right) - \delta z_2 = T(u-y) - \delta z_2. \tag{9}$$

Due to the fact that mRNA-sRNA complexes are degraded much more rapidly than uncoupled RNAs [25], the effect of leaky integration in this controller can be attenuated. In particular, we leverage this time-scale separation property to put (7) into the general form of ϵ-quasi-integral control in (3)-(4). We use $\epsilon := \delta/\theta \ll 1$ to characterize the rate difference between mRNA-sRNA complex degradation and uncoupled RNA decay, and 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 transcription rate constants of mRNA and sRNA become T/ϵ and T_s/ϵ, respectively. Under the coordinate transformation $x := p$, $z_1 := m$ and $z_2 := m - s$, system (7) can be re-written as:

$$\frac{\mathrm{d}}{\mathrm{d}t}x = R(1-d)z_1 - \gamma x, \tag{10a}$$

$$\epsilon\frac{\mathrm{d}}{\mathrm{d}t}z_1 = Tu - \epsilon\delta z_1 - \delta z_1(z_1 - z_2)/\beta, \tag{10b}$$

$$\epsilon\frac{\mathrm{d}}{\mathrm{d}t}z_2 = T\left(u - \frac{T_s}{T}\frac{x/k_s}{1+x/k_2}\right) - \epsilon\delta z_2 = T\left(u-y\right) - \epsilon\delta z_2. \tag{10c}$$

Since system (10) takes the form of (3)-(4), we apply Claim 1 to determine whether the steady state output $\bar{y}$ is close to u when ϵ is small. We find that when $u \in \mathbb{U} := (0, T_s/T)$, conditions (C1) and (C2) are satisfied, and the steady state is unique and locally asymptotically stable for all positive ϵ (see SI Section S2). Therefore, the sRNA-based controller (10) can achieve quasi-integral control for all $u \in \mathbb{U}$ and $d \in [0, 1)$. Condition (C1) breaks when $u \geq T_s/T$, corresponding to a situation where the set-point input is larger than the maximum output the system can reach, due to the saturation of the Hill function in equation (8).

We can decrease the effect of leaky integration by simultaneously (i) increasing the mRNA-sRNA degradation rate constant θ to decrease ϵ, and (ii) increasing the DNA copy numbers of the regulated gene and the sRNA such that their transcription rate constants become T/ϵ and T_s/ϵ. While directly increasing θ may be difficult to implement in practice, since the parameters θ and β are clustered together in model (7), we can achieve the same effect by increasing the affinity between sRNA and mRNA ($1/\beta$) [28]. We confirm the above results by simulations in Figure 2B and Figure S4 in SI using biologically relevant parameters from bacteria *E. coli*.

As a final remark, note that according to the leaky integration dynamics (9), the integral action can be made faster by simultaneously increasing the DNA copy numbers of the gene and the sRNA to increase T and T_s. However, as we demonstrate numerically in SI Figure S1, unless removal rate (θ/β) of the mRNA-sRNA complex is also increased simultaneously, increasing DNA copy numbers alone does not decrease the effect of leaky integration. This is in contrary to previous hypothesis that increasing the speed of integral action is sufficient to suppress the leaky integration effect [6, 9, 10]. In fact, conditions (C1) and (C2) are violated if the complex removal rate is small (see discussion in SI Section 1.3).

3.2 Physphorylation-based quasi-integral controller

A diagram of the phosphorylation-based quasi-integral controller is shown in Figure 2C. The controller is intended to regulate the production (i.e. transcription and translation) of protein p to adapt to a disturbance $d \in [0, 1)$, which models a reduction in protein production rate due to, for example, depletion of transcriptional and/or translational resources in the host cell by other genes [21, 22]. In this system, the integral action is accomplished by a phosphorylation cycle, in which

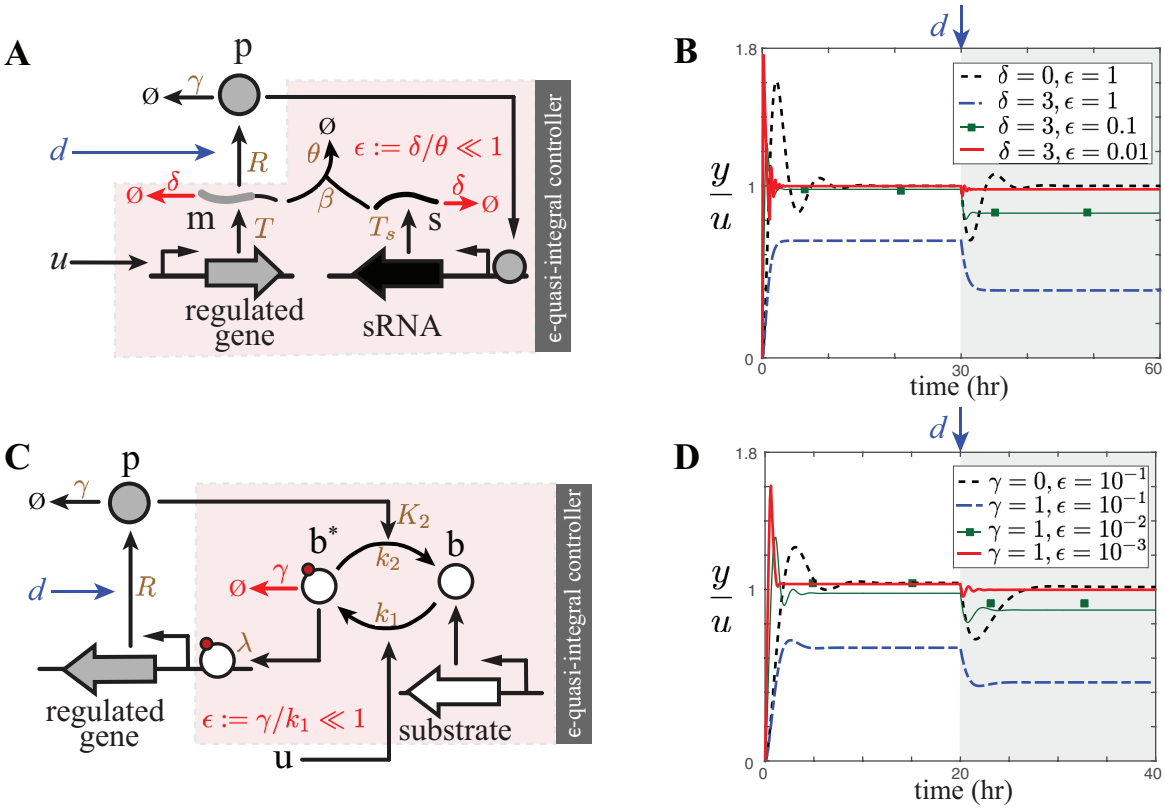

Figure 2: **Two physical realizations of ϵ-qICMs.** (A) Genetic circuit diagram of the sRNA-based quasi-integral controller. Chemical reactions realizing the controller are boxed in pink. (B) Simulation of the circuit's response according to (7). A set-point input $u = 1$ is applied at time 0 and a disturbance input $d = 0.5$ is applied at 30 hr. The vertical axis represents the ratio between output y, defined in (8), 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. (C) Genetic circuit diagram of the phosphorylation-based quasi-integral controller. (D) Simulation of the circuit's response according to (11). 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 input 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, realized by increasing the catalytic rates (k_i, $i = 1, 2$). Simulation parameters are listed in SI Section S4.

both the kinase and the phosphatase are saturated by their substrates. The regulated protein p is co-expressed with a phosphatase and the set-point input u represents the concentration of a kinase. The substrate b is expressed constitutively. When b is phosphorylated by kinase u to become active substrate b*, it transcriptionally activates production of p. A simplified mathematical model of this system is (see SI Section S3 for derivation):

$$\frac{\mathrm{d}}{\mathrm{d}t}b^* = k_1\frac{ub}{b+K_1} - k_2\frac{pb^*}{b^*+K_2} - \gamma b^*, \qquad \frac{\mathrm{d}}{\mathrm{d}t}p = R(1-d)\frac{b^*}{\lambda+b^*} - \gamma p, \tag{11}$$

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 that characterize the binding affinity between u and b, and between p and b*, respectively, R is the protein production rate constant, and λ is the dissociation constant between b* and the promoter of the regulated gene.

Phosphorylation and dephosphorayltion reactions are often much faster than dilution. For example, $k_i/\gamma \approx 10^3$ $(i = 1, 2)$ in bacteria [29]. Therefore, a natural choice of ϵ is to let $\epsilon := \gamma/k_1$. System (11) can be taken to the form of an ϵ-quasi-integral control system (3)-(4) if both the kinase and the phosphatase are saturated (i.e. $b \gg K_1$ and $b^* \gg K_2$). In particular, in SI Section S3.3, we find that when (I) the Michaelis-Menten constant between p and b* (K_2) is small compared to the dissociation constant λ between b* and the regulated gene, (II) the set-point input u is not too small, and (III) the production rate of b is sufficiently large, both the kinase and the phosphatase can be saturated by their substrates at steady state. These design constraints are inherited from the mechanism of type I IICMs and can be satisfied through proper engineering of binding affinities (in particular, λ) and DNA copy numbers (see SI Section S3.3 for a detailed discussion).

Assuming that both the kinase and the phosphatase are saturated, by setting $x := p$, $z_2 := b^*$ and $\sigma := k_2/k_1$, we can approximate (11) to become

$$\frac{\mathrm{d}}{\mathrm{d}t}x = R(1-d)\frac{z_2}{\lambda+z_2} - \gamma x, \qquad \epsilon\frac{\mathrm{d}}{\mathrm{d}t}z_2 = \gamma(u-\sigma x) - \epsilon\gamma z_2 = \gamma(u-y) - \epsilon\gamma z_2, \tag{12}$$

where we defined the output to be $y := \sigma x = \sigma p$. System (12) is in the form of (3)-(4). When $u < \sigma R(1-d)/\gamma$, conditions (C1) and (C2) in Claim 1 are satisfied, and the steady state is unique and locally asymptotically stable for all positive ϵ (see SI Section S3). We therefore conclude that system (12) can achieve ϵ-quasi-integral control with steady state output $\bar{y} = u$. This is supported by simulation results in Figure 2D and Figure S5 in SI, in which the steady state error due to leaky integration can be reduced as we decrease ϵ by increasing the time-scale separation between dilution and phosphorylation/dephosphorylation.

4 Discussion

Due to their appealing abilities to control a process to achieve set-point regulation regardless of disturbances/uncertainties, integral controllers have many potential applications in the field of synthetic biology. A well-known challenge to realize integral control in living cells is that the concentration of controller species dilute as cells grow. In particular, IICMs lose the integral structure once dilution is taken into account (Figure 1C). In this report, we propose an approach based on time-scale separation to overcome this obstacle to realize integral control in living cells. We establish a general mathematical structure of ϵ-quasi-integral controllers and provide easy-to-check algebraic conditions, under which the effect of leaky integration due to dilution can be made arbitrarily small as we increase the time-scale separation between dilution and all controller reactions. (A more general situation where only part of the controller reactions are much faster than dilution is considered in SI Section S1.) In addition to the two types of ϵ-qICMs presented here, these mathematical results can facilitate researchers to uncover/design additional ϵ-qICMs in the future.

The physical implementations of quasi-integral controllers proposed in Section 3 share common elements with some well-known adaptive/homeostatic networks in nature. In bacterial chemotaxis

[3] and yeast osmoregulation [5], integral actions are carried out by the fast methylation and phosphorylation processes. Similarly, negative feedback systems involving sRNA have been identified in iron homeostasis [30], in quorum sensing [31] and in sugar metabolism [32]. Therefore, further experimental study of the simple synthetic 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 integral actions). This is because with a perfect integrator, steady state set-point regulation and disturbance 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 [1]. For example, due to the high DNA copy number requirement in the sRNA-based controller, circuit operation sequesters a significant amount of RNA polymerase for transcription, which may affect the expression of other genes in the cell [22]. A potential advantage of increasing 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 S5, further theoretical assessment of this property may require us to leverage existing control theoretic results to new contexts [33, 34].

Acknowledgement

We thank Prof. Mustafa Khammash, Dr. Corentin Briat and Theodore Grunberg for helpful discussions and suggestions. We thank Cameron McBride who assisted proof-reading the manuscripts. This work is supported by AFOSR grant FA9550-14-1-0060.

Author contributions

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

Conflicts of interest

The authors declare no conflicts of interest.

References

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

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

[3] 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.

[4] H. El-Samd, J. P. Goff, and M. Khammash. Calcium Homeostasis and Parturient Hypocalcemia: An Integral Feedback Perspective. *J. Theor. Biol*, 214:17–29, 2002.

[5] 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.

[6] Pramod R. Somvanshi, Anilkumar K. Patel, Sharad Bhartiya, and K. V. Venkatesh. Implementation of integral feedback control in biological systems. *Wiley Interdiscip. Rev. Syst. Biol. Med.*, 7(5):301–316, 2015.

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

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] Xinying Ren, Ania-Ariadna Baetica, Anandh Swaminathan, and Richard M Murray. Population regulation in microbial consortia using dual feedback control. *bioRxiv*, 2017.

[14] 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.

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

[16] E D Sontag. Remarks on feedforward circuits, adaptation, and pulse memory. *IET Syst Biol*, 4(1):39–51, 2010.

[17] Yu-Tzu Chiu and S.K. Moore. Leaking capacitors muck up motherboards. *IEEE Spectr.*, 40(2):16–17, 2003.

[18] 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.

[19] Burton W. Andrews, Eduardo D. Sontag, and Pablo A. Iglesias. An approximate internal model principle: Applications to nonlinear models of biological systems. *IFAC Proceedings Volumes*, 41(2):15873 – 15878, 2008.

[20] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, 3rd edition, 1976.

[21] 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.

[22] Yili Qian, Hsin-Ho Huang, José I. Jiménez, and Domitilla Del Vecchio. Resource competition shapes the response of genetic circuits. *ACS Synth. Biol.*, 2017.

[23] 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.

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

[25] 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.

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

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

[28] 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.

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

[30] Eric Massé and Susan Gottesman. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.*, 99(7):4620–4625, 2002.

[31] Kimberly C. Tu, Tao Long, Sine L. Svenningsen, Ned S. Wingreen, and Bonnie L. Bassler. Negative feedback loops involving small regulatory RNAs precisely control the Vibrio harveyi quorum-sensing response. *Mol. Cell*, 37(4):567–579, 2010.

[32] Falk Kalamorz, Birte Reichenbach, Walter März, Bodo Rak, and Boris Görke. Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in Escherichia coli. *Mol. Microbiol.*, 65(6):1518–1533, 2007.

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

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