Design and Analysis of Genetic Feedback Architectures for Synthetic Biology

Emmanouil Alexis

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2023

Ἀριστοτέλης ἐρωτηθεὶς ποῦ κατοικοῦσιν αἱ Μοῦσαι, ἔφη:

" ἐν ταῖς ψυχαῖς τῶν φιλοπόνων."

When asked where the Muses dwell, Aristotle replied:

"In the souls of the diligent."

Acknowlegements

My doctoral journey would not have been nearly as fruitful – if possible at all – without the guidance of my advisors, Professors Antonis Papachristodoulou and Luca Cardelli. I am deeply grateful to them for providing me the freedom to follow my own research paths, constantly challenging me to acquire new skills and supporting all my academic endeavours, even the unsuccessful ones. Throughout this journey, they have been a unique source of inspiration, encouragement and scientific passion or, to put it simply, the best mentors one could ask for.

My sincerest thanks go to Professor Harrison Steel for his thoughtful advice and academic support at every stage of my doctorate. My gratitude also extends to Professor Kostas Margellos for his kind presence and contribution in all the key milestones of my DPhil programme. Moreover, I would like to thank Professor Lucia Marucci from the University of Bristol for serving in my DPhil viva examination.

I am very grateful to the Director of Stanford House (Montag Centre for Overseas Studies), Dr. Stephanie Solywoda, for offering me the opportunity to join the academic community of Stanford University as a tutor. I would also like to thank Professor Richard Murray for hosting my visit to California Institute of Technology (Caltech). In addition, I have immensely benefited from the insightful conversations with Professor James Forbes from McGill University.

I would like to express my profound appreciation to all the (present and past) members of Control Group at Engineering Science Department and in particular to the SySOS peo-

Abstract

Synthetic Biology seeks to design and assemble novel biological systems with favourable properties. It allows us to comprehend and modify the fundamental mechanisms of life and holds significant promise in revolutionizing current technologies ranging from medicine and biomanufacturing to energy and environmental protection. Biological processes constitute remarkably complex dynamical systems operating impeccably well in messy and constantly changing environments. Their ability to do so is rooted in sophisticated molecular control architectures crafted by natural evolutionary innovation over billions of years. Such control architectures, often blended with human- engineering approaches, are the key to realizing efficient and reliable synthetic biological systems. Aiming to accelerate the development of the latter, the present thesis addresses some fundamental challenges in biomolecular systems and control design.

We begin by elucidating biological mechanisms of temporal gradient computation, enabling cells to adjust their behaviour in response to anticipated environmental changes. Specifically, we introduce biomolecular motifs capable of functioning as highly tunable and accurate signal differentiators to input molecular signals around their nominal operation. We investigate strategies to deal with high-frequency input signal components which can be detrimental to the performance of most differentiators. We ascertain the occurrence of such motifs in natural regulatory networks and demonstrate the potential of synthetic experimental realizations. Our motifs can serve as reliable speed biosensors and can form the basis for derivative feedback control.

Motivated by the pervasiveness of Proportional-Integral-Derivative (PID) controllers in modern technological applications, we present the realization of a PID controller via biomolecular reactions employing, among others, our differentiator motifs. This biomolecular architecture represents a PID control law with set point weighting and filtered derivative action, offering robust regulation of a single-output biological process with enhanced dynamic performance and low levels of stochastic noise. It is characterized by significant ease of tuning and can be of particular experimental interest in molecular programming applications.

Finally, we investigate efficient regulation strategies for multi-output biological processes with internal coupling interactions, expanding previously established single-output control approaches. More specifically, we propose control schemes allowing for robust manipulation of the outputs in various ways, namely manipulation of their product/ratio, linear combinations of them as well as manipulation of each of the outputs independently. Our analysis is centered around two-output biological processes, yet the scalability of the proposed regulation strategies to processes with a higher number of outputs is highlighted. In parallel, their experimental implementability is explored in both *in vivo* and *in vitro* settings.

Contents

Chapter 1

Introduction

1.1 Motivation

Synthetic Biology undoubtedly constitutes one of the most promising and rapidly developing fields of research in the 21st century, lying at the crossroads of Engineering and Biology. It comprises the concepts, methods and technologies enabling the design and engineering of novel biological systems or modification of existing, natural ones in order to carry out predefined tasks. The well-known Richard Feynman's dictum, "What I cannot create, I do not understand", is often cited by synthetic biologists to eloquently summarize the field's promise: only via the process of creating artificial forms of life can the mysteries of life be fully illuminated.

Recent advancements have illustrated the immense potential for Synthetic Biology applications to revolutionize the current technology landscape and address major humanitarian problems across areas as diverse as healthcare, manufacturing, energy, the environment and others. Examples include: bacterial and mammalian cell-based systems for diagnostic and drug delivery purposes targeting a variety of human diseases such as metabolic and autoimmune disorders [59, 15]; programmable living materials created by genetically engineered microorganisms [28]; fuels and chemical synthesis through cell factories em-
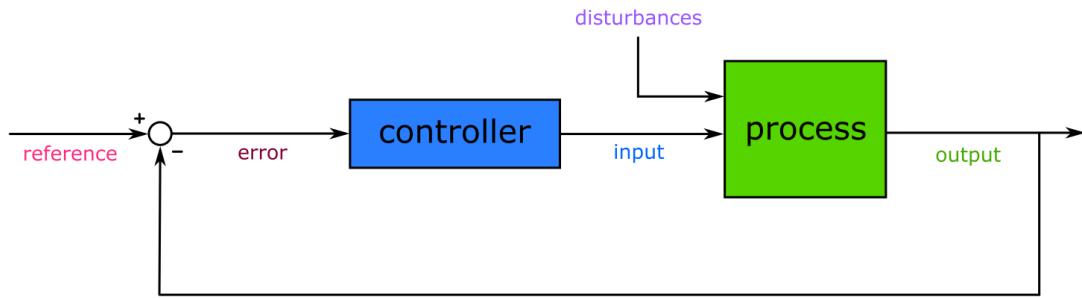
Figure 1.1: General structure of a feedback control closed loop

ploying renewable resources and atmospheric CO_2 [52, 50]; fast microbial production of high-quality food ingredients [78] and cellular biosensors for environmental pollution monitoring [30].

Control Theory provides a mathematical framework encompassing powerful techniques for the analysis, design and manipulation of physical systems. A central notion in this theory is feedback control referring to the ability of a system to fine-tune its behaviour by measuring one or more of its outputs. Feedback can be found almost in every aspect of modern life including power generation and transmission, transportation, internet networks, electronic devices, industrial robotics and economics. Negative feedback is probably the most prevalent form of feedback in engineering applications. The general idea is the creation of a closed loop where the output of a process is compared with a reference input (**Figure 1**). As a result, an error signal is generated which is then processed by a controller. The latter, in turn, produces a new input to the initial process, which steers the output towards the desired value even in the presence of disturbances.

Intriguingly, feedback control architectures are ubiquitous in nature at various organizational levels and timescales, regulating essential processes for the survival of living systems. Numerous natural homeostatic mechanisms exploit integral feedback control [8], one of the fundamental regulation schemes in traditional control engineering, as happens, for instance, with bacterial chemotaxis [88], yeast osmoregulation [62] or the regulation of calcium concentration in mammals [20]. Other mechanisms, such as the one responsible

for bacterial persistence [73], take advantage of positive feedback strategies while in more complex cases, such as the regulatory network of galactose metabolism in yeast [85], a combination of positive and negative feedback loops is used.

Drawing inspiration from the natural evolutionary innovation blended with concepts from modern technology has led to several success stories in the front of biological feedback control systems. In fact, the first two synthetic genetic circuits built, namely a bistable toggle switch [27] and an oscillatory network termed the repressilator [21], which marked the initiation of Synthetic Biology as a field at the beginning of the current millennium, both used feedback to design dynamics. In the recent years, there have been various successful synthetic implementations of feedback regulation strategies within cells (*in vivo*) such as integral controllers [42, 38, 7, 26], feedforward control topologies [9, 41, 25, 31], paradoxical [55], layered [37] or burden-driven [14] feedback motifs and others. In parallel, similar approaches have been studied in cell-free or *in vitro* environments [1, 65, 71]. Additionally, substantial research has focused on *in silico* control of biological processes by interfacing the latter with a computer executing the control algorithm [36, 69, 82, 48, 70, 68, 47].

The design of well-regulated biomolecular devices with reliable performance is a key prerequisite for synthetic biology to fulfill its potential, allowing current and future developments to be directly transferable to out-of-the-lab contexts. Despite recent progress, our capacity to attain this aim is still hampered by a number of ongoing system-level challenges resulting to a tremendous lack of robustness, modularity and predictability. More specifically, the performance of artificial biomolecular networks is severely affected by the presence of undesirable crosstalk interactions between different genetic modules as well as by environmental variations in terms of, for instance, temperature, growing media and cell cycle stage [61, 13]. Unexpected or even catastrophic changes in a design's behaviour may also occur due to loading effects stemming from the interconnection of biomolecular components [19] or due to competition for accessing the limited resources of the host cell [72]. Another important contributing factor is the stochastic noise which is generated

from randomizing effects of the cellular environment (extrinsic noise) and the probabilistic nature of chemical reactions (intrinsic noise) [83, 67]. Moreover, synthetic circuit failure might be caused by the emergence of random genetic mutants with a fitness advantage, capable of dominating the cell population (mutational escape) [81]. These challenges, among others, highlight the fact that biological systems are quite distinct, exhibiting a plethora of features and requirements that are absent in traditional engineering contexts. As a consequence, our ability to mimic conventional control schemes found in the latter is severely restricted. This, by extension, necessitates the construction of new systematic design tools and methods specifically tailored to biomolecular environments.

This thesis constitutes a step forward in addressing the problems and limitations discussed above, with the ultimate goal of accelerating the realization of dependable synthetic bio-devices. This work provides novel theory-grounded frameworks and architectures for efficient biomolecular system design and control by leveraging principles and techniques from the broad area of Dynamical Systems, and especially Control Theory. Special emphasis has been placed on establishing stability, disturbance rejection, and predictable performance while guaranteeing biological implementability. Furthermore, new light is shed on the underlying mechanisms of naturally occuring motifs.

1.2 Thesis Outline

An overview of the thesis is provided below, briefly summarizing the content of each chapter. The contributions and novelties of this work are organized into research papers which have been published or submitted for publication.

- **Chapter 2**

  This chapter presents a high-level review of fundamental ideas necessary for the analysis and design of biomolecular systems which are not covered in later chapters. It begins with a concise introduction to basic mathematical tools, followed by a brief discussion of some core concepts in molecular biology.

- **Chapter 3**

  This chapter introduces three modular and tunable biomolecular devices capable of operating as signal differentiators of high accuracy. Tuning strategies and structural additions for enhanced performance and high-frequency input noise insensitivity are also investigated. Finally, natural regulatory networks whose structure resemble the core of these architectures are studied and guidelines for potential synthetic implementations are provided. The corresponding manuscript has been published [5].

- **Chapter 4**

  This chapter describes a chemical reaction network implementation of a highly tunable Proportional-Integral-Derivative (PID) controller. It shows two special characteristics, namely set-point weighting and high-frequency noise filtering regarding derivative control. It is demonstrated that the feedback bio-controller in question is able to achieve robust regulation of a biological process with improved transient dynamics and mitigates stochastic noise. The corresponding manuscript has been published [4].

- **Chapter 5**

This chapter investigates the problem of regulating biomolecular systems of multiple outputs interacting with each other in the presence of disturbances from the external environment. In particular, feedback architectures for processes with two outputs of interest are introduced, capable of robustly controlling the ratio/product of the latter, a linear combination of them and each of them independently. Potential synthetic experimental realizations are also discussed. The corresponding manuscript has been submitted for publication and is currently under revision [6].

- **Chapter 6**

  This chapter provides a synopsis of the thesis' findings and a discussion on future research avenues.

1.3 Thesis contributions and related works

My doctoral research has led to the following publications which are part of the present thesis:

- Alexis, E., Schulte, C. C., Cardelli, L., Papachristodoulou, A., *Biomolecular mechanisms for signal differentiation*, Iscience, 24(12), 103462, 2021

- E. Alexis, L. Cardelli and A. Papachristodoulou, *On the Design of a PID Bio-Controller With Set Point Weighting and Filtered Derivative Action*, IEEE Control Systems Letters, vol. 6, pp. 3134-3139, 2022 [*]

- Alexis, E., Schulte, C. C., Cardelli, L., Papachristodoulou, A., *Regulation strategies for two-output biomolecular networks*, bioRxiv, 2022

  [*] This paper was presented as an invited paper in 61st IEEE Conference on Decision and Control (CDC) 2022.

I have also contributed to the following study which is not covered herein:

- Sootla, A.[#], Delalez, N.[#], Alexis, E., Norman, A., Steel, H., Wadhams, G. H., Papachristodoulou, A., *Dichotomous feedback: a signal sequestration-based feedback mechanism for biocontroller design*, Journal of the Royal Society Interface, 19(189), 20210737, 2022

  [#] joint first authors

There have been several recent efforts towards the three primary research directions of this thesis, namely biomolecular realizations of signal differentiators, PID controllers, and multi-output control strategies. Subsequently, a concise comparative discussion on related works is presented, highlighting the main novelties of our results.

Biomolecular signal differentiators

<u>Literature review</u>

The main research attempts in this area are summarized below:

- In the design of biological topologies, dual rail encoding [63] can be employed to represent a signal as the non-physical difference between two molecular species components obeying mass-action kinetics. Exploiting this technique and incorporating time delays, two structurally distinct topologies are presented by the authors in [87] and [66], capable of approximating time derivatives of input signals. While these topologies are synthesizable *in vitro* via nucleic acid-based chemistry, their realization *in vivo* poses considerable challenges.

- A feedback architecture containing integral action and a high gain is proposed in [35, 34] to estimate temporal gradients of input signals. Furthermore, a potential experimental implementation through a two-gene circuit in a cell-free environment is discussed. The design approach used therein is similar to [63] and, thus, the output temporal derivative is given by the difference of two biomolecular concentrations.

- A two-node circuit based on a bacterial chemotaxis-like mechanism is adopted by the authors in [17] to design a signal differentiator. This circuit involves active degradation processes following Michaelis-Menten kinetics and operating close to saturation.

- Three feedback motifs with ultrasensitive and quasi-integral components are introduced in [75], which are able to function as practical differentiators. Their structure includes reactions related to enzymatic activation/deactivation cycles and molecular sequestration.

- Constructing a derivative operator via an incoherent feedforward loop and time delays is investigated in [76]. Specifically, architectures based on molecular sequestration, a push-pull system, and a hybrid promoter are examined. The output of the

8

resulted operator does not estimate the "full derivative" of an input signal but rather (only) the positive or negative gradients of it.

- The authors in [23] explore different ways of realizing derivative action for biomolecular feedback control applications. In particular, they accomplish this by utilizing incoherent feedforward loops. In parallel, by placing various integral motifs within feedback loops, they construct different types of differentiator modules, namely "antithetic" differentiators (based on molecular sequestration), inflow/outflow zero-order differentiators, and auto-catalytic differentiators. An important characteristic of these topologies is that the computed derivative is represented by a reaction rate rather than a biomolecular species. While this can be convenient in biocontroller design, it might not be favourable in situations where the computed derivative needs to be (experimentally) measured, for example, in the case of speed sensing.

- A derivative-based controller tailored to gene expression processes is proposed in [60]. The controller is implemented through a coupled feedforward-feedback biochemical circuit taking advantage of time delays.

Thesis contribution

We introduce three architectures employing production-inhibition loops and molecular sequestration. The biomolecular reactions involved following mass-action kinetics while, in one of the architectures, an enzyme-catalyzed degradation based on Michaelis-Menten kinetics is used. Our circuits can function as general-purpose signal differentiators of high accuracy where the derivative of an input signal is encoded as a biomolecular species. They are suitable for (*in vivo* and/or *in vitro*) synthetic experimental implementations and can be used as lenses to investigate natural regulatory mechanisms.

Biomolecular PID controllers

<u>Literature review</u>

In some of the above works, besides the development of differentiator modules, biomolecular realizations regarding the remaining constituent components of a PID controller are also presented. More specifically:

- Two alternative PID control strategies, especially suitable for molecular programming applications, are reported in [87] and [66]. Using dual rail encoding, each of the controller parts is designed independently and, then, their control actions are added up. The improved behaviour of the resulting closed-loop systems is demonstrated in the deterministic setting.

- Similarly to the aforementioned concept, the authors in [17] design the P, I and D components independently to build a controller with enhanced dynamic performance in the deterministic setting. This PID architecture heavily depends on processes that follow Michaelis-Menten kinetics.

- Adopting Hill and mass-action kinetics, the study in [23] introduces a set of PID topologies of varying structural complexity, where the P, I and D parts are not explicitly separable. Higher complexity results in additional degrees of freedom, increased structural separability, and improved performance capabilities. Compared to PI control, the deterministic closed-loop behaviour of the PID topologies is shown to be improved in terms of stability and transient dynamics while some of these topologies are able to suppress stochastic fluctuations. A noteworthy feature is the incorporation of a first-order low-pass filter into some or all the control terms of these designs.

- Focusing on the stochastic nature of gene expression, the work in [60] studies the efficacy of proportional, integral, and derivative control action in mitigating protein

count fluctuations. Note that each type of control is analyzed separately and, therefore, a full PID architecture is not discussed.

Thesis contribution

We propose a chemical reaction network based on mass-action kinetics capable of operating as a PID controller with set point weights and filtered derivative action. The latter is achieved through a second-order low-pass filter. The effectiveness of our strategy is demonstrated in both the deterministic and stochastic setting and compared with PI regulation. Our PID scheme can offer significant tunability and can enhance the closed-loop output behavior by eliminating overshoots and oscillations as well as reducing stochastic noise. Note that our design approach is relatively similar to [23] . Nevertheless, our architecture implements a different type of PID control law which, in turn, results in distinct structural differences.

Biomolecular multi-output control strategies

Literature review

Multi-output control of biological systems has been previously studied exclusively in contexts with either external computer-based control or via genetic logic gates. In particular:

- The genetic toggle switch [27] is a two-output, bistable circuit composed of two mutually repressing genes. The authors in [54] examine different *in-silico* control approaches to balance the toggle switch, namely PI control, Bang Bang control, and open-loop periodic forcing based on mutually exclusive pulse waves. At the same time, two alternative *in-silico* methods are reported in [32]. This study exploits a model-based hybrid strategy of PI control and pulse-width modulation (PWM) as well as a PWM-based strategy using Zero Average Dynamics (ZAD) control.

11

- Combinatorial logic circuits with multiple inputs and multiple outputs have been successfully engineered in bacterial [10] and mammalian [86] cells. While these advancements hold great potential for biomolecular computation, their applicability in feedback control applications may be limited.

Thesis contribution

We present the first chemical reaction network implementation of multi-output feedback control strategies. Using mass-action kinetics and leveraging the concept of antithetic integral feedback [11], our strategies are able to achieve robust steady-state tracking with respect to each of the outputs independently or a desired combination of them. Notably, our regulatory topologies are suitable for fully *in vivo* or *in vitro* experimental settings without the need for external computer-based control.

Finally, it is important to emphasize that the mathematical modelling and analysis in all three main parts of the thesis follow a deterministic approach based on Ordinary Differential Equations (ODEs). Additionally, in the second part, Van Kampen's Linear Noise Approximation (LNA) of the Chemical Master Equation (CME) is adopted to study the stochastic behaviour of the topology under consideration.

Chapter 2

Background

The aim of the present chapter is to provide the reader with a succinct overview of important mathematical and biological concepts that are used in the ensuing chapters of this thesis, yet are not expounded upon therein.

2.1 Mathematical Background

We commence with an exposition of some mathematical ideas that play an instrumental role in modelling and analyzing the dynamics of the biological systems investigated in this thesis.

2.1.1 Modelling chemical reaction networks

Consider q biochemical species $X_1, \ldots, X_q$, where $q \in \mathbb{N}$, which undergo m chemical reactions, where $m \in \mathbb{N}$, composing the following chemical reaction network (CRN)

$$\sum_{j=1}^{q} A_{ij} X_j \xrightarrow{k_i} \sum_{j=1}^{q} B_{ij} X_j, \ i = 1, \ldots, m, \tag{2.1}$$

14
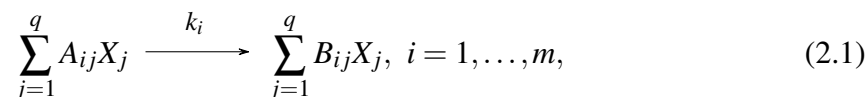
The terms $\sum_{j=1}^{q} A_{ij}X_j$ and $\sum_{j=1}^{q} B_{ij}X_j$ represent the reactant and the product of the ith reaction, respectively, where A_{ij}, B_{ij}, also known as stoichiometric coefficients, are non-negative integers. In addition, k_i refers to the reaction rate constant of the ith reaction and is a positive real number.

Each of the reactions participating in CRN (2.1) is assumed to be irreversible - the transformation of reactants into products is carried out only in the direction of the arrow ($\rightarrow$). In case of a reversible reaction where the converse transformation also takes place the notation ($\leftrightarrows$) or ($\leftrightarrow$) is often used instead. To be aligned with the formalism above (CRN (2.1)), the reverse reaction can be introduced as a separate reaction. Furthermore, CRN (2.1) is considered a closed system since there is no material exchange with the external environment. In other words, the reactants and the products of the reactions involved lie within the network in question. Another commonly encountered symbol is ($\varnothing$) which may appear at the left-hand side or the right-hand side of a reaction representing an unspecified process which is not important for the problem under consideration. In this scenario A_{ij} or B_{ij} is 0, respectively, corresponding to an open system where mass addition/removal is permitted.

The dynamic behaviour of a CRN is usually studied under the following assumptions :

- *Spatial homogeneity* : This is also known as the *well-mixed assumption* and entails equal distribution of the reactants throughout the reaction volume - no spatial structure exists. The reaction rates are therefore spatially independent. This implies that the time scale governing the evolution of the process of interest is longer than the one corresponding to diffusion of its molecular components.

- *Continuum hypothesis* : This refers to the description of molecular abundance by a continuously changing (real-valued) concentration instead of using a discrete (integer-valued) measure. This approximation is valid provided that the number of molecules of the species involved is sufficiently large.

In a fixed volume where the above two assumptions hold, the dynamics of a CRN can be derived by invoking the *Law of Mass Action*, according to which the rate of a chemical reaction is proportional to the product of the concentrations of its reactants. Thus, the dynamics of CRN (2.1) can be given by the following ordinary differential equations which are referred to as the mass action kinetics or the reaction rate equations of the network:

$$\dot{x}(t) = (B-A)^T[k \circ x^A(t)], \ x(0) = x_0, \ t \geq 0, \tag{2.2}$$

where $x_j(t)$ represents the concentration of species X_j at time t, $x(t) = [x_1(t),\ldots,x_q(t)]^T$, $k = [k_1,\ldots,k_m]^T$ and $A = [A_{ij}]$, $B = [B_{ij}]$ correspond to $m \times q$ non-negative matrices. Moreover, the notations $\circ$ and $x^A(t)$ denote component-wise multiplication and vector-matrix exponentiation, respectively.

Equation (2.2) can be rewritten as:

$$\dot{x}(t) = NRx^A(t), \ x(0) = x_0, \ t \geq 0, \tag{2.3}$$

where $N = (B-A)^T$, commonly known as the stoichiometry matrix, and $R = diag(k_1,\ldots,k_m)$.

As expected from their physical interpretation, an important property of the above kinetic equations is that their states remain non-negative provided that the initial conditions are also non-negative and that the solution exists.

Enzymes (a specialized class of proteins) are responsible for catalyzing the vast majority of chemical reactions taking place within the cell. To achieve that, enzymes are able to bind the reactants - termed (enzyme) substrates - and aid their transformation into the products. A general enzymatic reaction can be described as:

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C \overset{k_3}{\longrightarrow} E + P \tag{2.4}$$
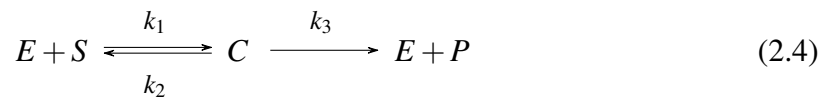
in which E, S, C represent the (free) enzyme, the substrate and the complex formed by the

former two, respectively.

By applying mass action kinetics and assuming that the first, reversible reaction is much faster than the second, reversible one in CRN (2.4), meaning $k_1, k_2 \gg k_3$, and that the initial concentration of S is sufficiently high, the so-called Michaelis-Menten kinetics can be derived:

$$\dot{p}(t) = k_3 \frac{e_{tot}s}{s + K_m} = V_{max}\frac{s}{s + K_m} \tag{2.5}$$

where:

- $e_{tot} = e + c$ is the total enzyme concentration which remains constant (the enzyme is not consumed)

- $V_{max} = k_3 e_{tot}$ is known as the maximal flow or maximal velocity.

- $K_m = \dfrac{k_2 + k_3}{k_1}$ is known as the Michaelis-Menten constant

Equation (2.5) can be further simplified in the following two scenarios:

- For $s \gg K_m$ (the enzyme is saturated by the substrate), Equation (2.5) becomes:

$$\dot{p}(t) \approx V_{max}$$

which states that the reaction practically reaches its maximal speed. The production rate does not depend on the concentration of the substrate and it is often referred to as zero-order kinetics.

- For $s \ll K_m$, Equation (2.5) becomes:

$$\dot{p}(t) \approx \frac{V_{max}}{K_m}s$$

which states that the production rate varies almost linearly with the concentration of the substrate and it is often referred to as first-order kinetics.

Further details on the above can be found in [16, 39, 18, 40].

2.1.2 State-space representation

The dynamical systems studied in this thesis can be modelled via a finite number of coupled first-order ordinary differential equations [43, 2]. Using vector notation these equations can be represented in a compact form as :

$$\dot{x} = f(t, x, u) \tag{2.6}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ is the input vector and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a smooth (possibly time-varying and/or non-linear) mapping. Equation (2.6) is called the state equation.

Another equation, called the output equation, is often associated with Equation (2.6):

$$y = h(t, x, u) \tag{2.7}$$

where $y \in \mathbb{R}^q$ and $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^q$ is a smooth (possibly time-varying and/or non-linear) mapping.

The state variables composing x represent the memory of a dynamical system in terms of its past whereas the output variables composing y represent the variables of interest which must behave in a certain way and/or can be physically measured. Equations (2.6) and (2.7) together constitute the so-called (normalized) state-space representation or system realization.

Systems with multiple inputs and outputs, meaning m, $q > 1$, are referred to as multi-input, multi-output (MIMO) systems. In the special case where $m = q = 1$, systems are referred to as single-input, single-output (SISO).

2.1.3 Linear, time-invariant systems and linearisation

In many practical situations, modelling is based on finite-dimensional linear, time invariant (LTI) systems. Equations (2.6) and (2.7) can therefore be simplified as:

$$\dot{x} = Ax + Bu \tag{2.8}$$

$$y = Cx + Du \tag{2.9}$$

where A, B, C, D are real, constant matrices of appropriate dimensions known as the dynamics matrix, the control matrix, the sensor matrix, the direct term, respectively.

A very common approach of studying the behaviour of a non-linear system is to analyze its dynamics in a regime of interest where this system can be approximated by an LTI system of the form (2.8)-(2.9) (local behaviour). An important characteristic of a system that plays a central role here is that of equilibrium point describing a stationary condition for its dynamics. Considering again the non-linear system (2.6), (2.7), the fixed point (x^*, u^*) constitutes an equilibrium point if $f(x^*, u^*) = 0$. Assuming now that the functions f and h are continuously differentiable, Jacobian linearization based on a truncated Taylor series expansion can be employed. More specifically, the following linearised model can be obtained:

$$\dot{\bar{x}} = A\bar{x} + B\bar{u}$$

$$\bar{y} = C\bar{x} + D\bar{u}$$

where $\bar{x} = x - x^*$, $\bar{u} = u - u^*$ represent small perturbations around (x^*, u^*) and:

$$A = \left. \frac{\partial f(x,u)}{\partial x} \right|_{(x,u)=(x^*,u^*)}$$

$$B = \left. \frac{\partial f(x,u)}{\partial u} \right|_{(x,u)=(x^*,u^*)}$$

$$C = \left. \frac{\partial h(x,u)}{\partial x} \right|_{(x,u)=(x^*,u^*)}$$

$$D = \left. \frac{\partial h(x,u)}{\partial u} \right|_{(x,u)=(x^*,u^*)}$$

2.1.4 State controllability and state observability

Here, we introduce the concepts of state controllability and state observability with respect to systems of the form (2.8)-(2.9).

Definition 2.1.1 (Definition 4.1 in [79]). *The dynamical system (2.8), or equivalently the pair (A,B), is said to be state controllable if, for any initial state $x(0) = x_0$, any time $t_1 > 0$ and any final state x_1, there exists an input $u(t)$ such that $x(t_1) = x_1$. Otherwise the system, or (A,B), is said to be state uncontrollable.*

Definition 2.1.2 (Definition 4.2 in [79]). *The dynamical system (2.8)-(2.9) (or the pair (A,C)), is said to be state observable if, for any time $t_1 > 0$, the initial state $x(0) = x_0$ can be determined from the time history of the input $u(t)$ and the output $y(t)$ in the interval $[0,t_1]$. Otherwise the system, or (A,C), is said to be state unobservable.*

To check if a system is state controllable and state observable, one can use the following tests: [79]:

- Test for state controllability: The system (A,B) is state controllable if and only if the controllability matrix

$$\begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix}$$

  has rank n (full row rank) . Note that n corresponds to the number of states.

- Test for state observability: The system (A,C) is state observable if and only if the

observability matrix

$$
\begin{bmatrix}
C \\
CA \\
\vdots \\
CA^{n-1}
\end{bmatrix}
$$

has rank n (full column rank).

2.1.5   Stability of autonomous systems

The stability analysis presented in this thesis mainly focuses on non-linear systems that take the form [57]:

$$\dot{x} = f(x) \tag{2.10}$$

System (2.10) does not explicitly depend on time and is called autonomous. Without loss of generality, $x^* = 0$ is assumed to be an equilibrium point of this system.

Definition 2.1.3 (Definition A.2 in [12]). *The equilibrium point $x^* = 0$ is said to be stable if, for any $\rho > 0$, there exists $r > 0$ such that if $\| x(0) \| < r$, then $\| x(t) \| < \rho$ for all $t \geq 0$.*

Definition 2.1.4 (Definition A.3 in [12]). *The equilibrium point $x^* = 0$ is asymptotically stable if it is stable, and if in addition there exists some $r > 0$ such that $\| x(0) \| < r$ implies that $x(t) \to 0$ as $t \to \infty$.*

Definition 2.1.5 (Definition A.5 in [12]). *The equilibrium point $x^* = 0$ is exponentially stable if there exist two strictly positive numbers α and λ, independent of time, and initial conditions such that*

$$\| x(t) \| \leq \alpha \| x(0) \| e^{-\lambda t}, \quad \text{for all } t > 0$$

in some ball around the origin.

In the above definitions, the notation $\| \cdot \|$ refers to the Euclidean (L^2) norm in $\mathbb{R}^n$, i.e. $\| x \| = \sqrt{x^T x}$ for all $x \in \mathbb{R}^n$.

It is apparent that exponential stability is the strongest type of stability from the concepts defined above and that it implies asymptotic stability (whereas the converse is not true). Moreover, these stability concepts refer to the local behaviour of system (2.10) around x^*. If their corresponding conditions are satisfied for any initial state, then they become global.

Theorem 2.1.6 (Theorem 3.11 in [57]). *Let $x^* = 0$ be an equilibrium point for system (2.10). Assume that f is continuously differentiable in D ($f : D \rightarrow \mathbb{R}^n$) and $A = \left. \frac{\partial f(x)}{\partial x} \right|_{(x)=(x^*)}$. Then if the eigenvalues λ_i of the matrix A satisfy $\mathscr{R}e(\lambda_i) < 0$, the origin is a (locally) exponentially stable equilibrium point for system (2.10).*

The eigenvalues of a given matrix A are the roots of the characteristic equation:

$$\Delta(s) = det(sI - A) = \alpha_n s^n + \alpha_{n-1} s^{n-1} + \cdots + \alpha_1 s + \alpha_0 = 0 \tag{2.11}$$

where I denotes the identity matrix.

Linear stability of A requires that the real parts of these roots are strictly negative. Nevertheless, explicit calculation of the eigenvalues can often be challenging, especially for high-order polynomials. Instead, the *Routh-Hurwitz criterion* can be exploited in order to determine if the position of the eigenvalues is in the open-half plane. This constitutes a necessary and sufficient criterion for linear stability. More specifically, given polynomial (2.11), this method uses an array based on the ordering of the coefficients regarding the

polynomial as follows [56]:

$$
\begin{bmatrix}
s^n & \alpha_n & \alpha_{n-2} & \alpha_{n-4} & \cdots \\
s^{n-1} & \alpha_{n-1} & \alpha_{n-3} & \alpha_{n-5} & \cdots \\
s^{n-2} & b_{n-1} & b_{n-3} & b_{n-5} & \cdots \\
s^{n-3} & c_{n-1} & c_{n-3} & c_{n-5} & \cdots \\
s^{n-4} & d_{n-1} & d_{n-3} & d_{n-5} & \cdots \\
\vdots
\end{bmatrix}
$$

where

$$
b_{n-1} = -\frac{1}{\alpha_{n-1}} \begin{vmatrix} \alpha_n & \alpha_{n-2} \\ \alpha_{n-1} & \alpha_{n-3} \end{vmatrix}, \quad
b_{n-3} = -\frac{1}{\alpha_{n-1}} \begin{vmatrix} \alpha_n & \alpha_{n-4} \\ \alpha_{n-1} & \alpha_{n-5} \end{vmatrix}
$$

$$
c_{n-1} = -\frac{1}{b_{n-1}} \begin{vmatrix} \alpha_{n-1} & \alpha_{n-3} \\ b_{n-1} & b_{n-3} \end{vmatrix}, \quad
c_{n-3} = -\frac{1}{b_{n-1}} \begin{vmatrix} \alpha_{n-1} & \alpha_{n-5} \\ b_{n-1} & b_{n-5} \end{vmatrix}
$$

$$
d_{n-1} = -\frac{1}{c_{n-1}} \begin{vmatrix} b_{n-1} & b_{n-3} \\ c_{n-1} & c_{n-3} \end{vmatrix}, \ldots
$$

According to *Routh-Hurwitz criterion*, all the elements of the first column are required to be nonzero and have the same sign (necessary and sufficient condition).

A thorough treatment of stability of dynamical systems can be found in [12, 43, 57, 80].

2.1.6 Transfer function representation and frequency response

It is often convenient to study the input-output relationships in LTI systems by using transfer functions. In particular, applying the Laplace transform to system (2.8)-(2.9), under the assumption that all initial conditions are zero, yields:

$$
sX(s) = AX(s) + BU(s)
$$

$$Y(s) = CX(s) + DU(s)$$

where $s \in \mathbb{C}$ is the Laplace variable.

Thus, the following expression is obtained:

$$Y(s) = G(s)U(s) \tag{2.12}$$

in which

$$G(s) = \left[C(sI - A)^{-1}B + D \right]$$

represents, for a general MIMO system, a transfer function matrix whose elements are (SISO) rational transfer functions between a specific input and output. Two important concepts in transfer function representation are those of poles and zeros which are defined below.

Definition 2.1.7 (Definition 4.6 in [79]). *The poles p_i of a system with state-space description (2.8)-(2.9) are the eigenvalues $\lambda_i(A)$, $i = 1, \ldots, n$ of the matrix A. The pole or characteristic polynomial $\phi(s)$ is defined as $\phi(s) \triangleq det(sI - A) = \prod_{i=1}^{n}(s - p_i)$. Thus the poles are the roots of the characteristic equation $\phi(s) \triangleq det(sI - A) = 0$*

Definition 2.1.8 (Definition 4.7 in [79]). *z_i is a zero of $G(s)$ if the rank of $G(z_i)$ is less than the normal rank of $G(s)$. The zero polynomial is defined as $z(s) = \prod_{i=1}^{n_z}(s - z_i)$ where n_z is the number of finite zeros of $G(s)$.*

Finally, substituting $s = j\omega$ in Equation (2.12) results in:

$$Y(j\omega) = G(j\omega)U(j\omega) \tag{2.13}$$

which describes the input-output relationship of the system via Fourier transform. $G(j\omega)$ is known as the frequency response of the system and can provide significant insight into its behaviour such as revealing its response to sinusoids of varying frequency.

Note that switching between Laplace and Fourier representation via the aforementioned substitution is not always possible. Nevertheless, this process is valid when followed in this work since the systems under consideration are asymptotically stable.

Further information on the above topics is provided in [2, 79, 49, 64].

2.1.7 Positive realness

Here we discuss the notion of positive realness with respect to transfer functions as well as its application to stability of interconnected systems within a negative feedback loop by presenting some key definitions and theorems.

About the notations used below: ^{H} and ^{T} indicate the conjugate transpose and the transpose of a matrix, respectively while $\succ$ ($\succeq$) indicates a positive definite (positive-semidefinite) matrix.

Definition 2.1.9 (Definition 2.34 in [12])**.** *The transfer matrix $H(s) \in \mathbb{C}^{m \times m}$ is positive real (PR) if:*

- *$H(s)$ has no pole in $\boldsymbol{Re}[s] > 0$.*

- *$H(s)$ is real for all positive real s.*

- *$H(s) + H^H(s) \succeq 0$ for all $\boldsymbol{Re}[s] > 0$.*

Theorem 2.1.10 (Theorem 2.48 in [12])**.** *The rational function $H(s) \in \mathbb{C}^{m \times m}$ is positive real (PR) if and only if:*

- *$H(s)$ has no pole in $\boldsymbol{Re}[s] > 0$.*

- *$H(j\omega) + H^H(j\omega) \succeq 0$ for all positive real ω such that $j\omega$ is not a pole of $H(\cdot)$.*

- *If $j\omega_0$, finite or infinite, is a pole of $H(\cdot)$, it is a simple pole and the corresponding residual $K_0 = \lim_{s \to j\omega_0}(s - \omega_0)H(s)$ if $\omega_0 < +\infty$, or $K_\infty = \lim_{\omega \to \infty} \dfrac{H(j\omega)}{j\omega}$ if $\omega_0 = \infty$, is a positive semi-definite Hermitian matrix.*

Definition 2.1.11 (Definition 2.58 in [12])**.** *A rational transfer function matrix $H(s) \in \mathbb{C}^{m \times m}$ that is not identically zero for all s is strictly positive real (SPR) if $H(s - \varepsilon)$ is PR for some $\varepsilon > 0$.*

Definition 2.1.12 (Definition 2.77 in [12])**.** *A rational transfer function matrix $H(s) \in \mathbb{C}^{m \times m}$ is weakly strictly positive real (WSPR) if:*

- *$H(s)$ is analytic in $\boldsymbol{Re}[s] \geq 0$.*

- *$H(j\omega) + H^T(-j\omega) \succ 0$ for all $\omega \in \mathbb{R}$.*

Theorem 2.1.13 (Lemma 3.67 in [12])**.** *Consider a system $H_1 : u_1 \rightarrow y_1$ in negative feedback with a system $H_2 : u_2 \rightarrow y_2$, i.e. $u_1 = -y_2$ and $u_2 = y_1$, where H_1 is PR and H_2 is WSPR. Under those conditions u_1, u_2, y_1 and y_2 all converge to zero exponentially.*

A detailed discussion on topics associated with positive realness can be found in [12, 43, 45].

2.1.8 Singular perturbations and model reduction

Due to their multi-time-scale behaviour, many biological processes can take the form of the singular perturbation model [43]:

$$\dot{x} = f(t, x, z, \varepsilon) \tag{2.14}$$

$$\varepsilon \dot{y} = g(t, x, z, \varepsilon) \tag{2.15}$$

where ε is considered a small positive parameter. In additon, f and g are continuously differentiable in their arguments for $(t, x, z, \varepsilon) \in [0, t_1] \times D_x \times D_z \times [0, \varepsilon_0]$ and $D_x \subset \mathbb{R}^n$, $D_z \subset \mathbb{R}^m$ are open connected sets.

Setting $\varepsilon = 0$ results in degeneration of differential Equation (2.15) into the algebraic or transcendental equation:

$$0 = g(t,x,z,0) \tag{2.16}$$

thus reducing the dimension of state-space model (2.14), (2.15) from $n+m$ to n. The model is said to be in standard form if Equation (2.16) has $k \geq 1$ isolated real roots

$$z = h_i(t,x), \quad i = 1,2,\ldots,k \tag{2.17}$$

for each $(t,x) \in [0,t_1] \times D_x$. Assuming $\varepsilon = 0$ and substituting Equation (2.17) into Equation (2.14) gives the following reduced model:

$$\dot{x} = f(t,x,h(t,x),0) \tag{2.18}$$

where the subscript i has been removed from h. Equation (2.18) is also known as the slow model or the quasi-steady state model. Note that when ε is small and $g \neq 0$, $\dot{z} = \dfrac{g}{\varepsilon}$ can be large and z may quickly converge to a root of Equation (2.16).

To get a deeper insight, we focus on the problem of solving the system :

$$\dot{x} = f(t,x,z,\varepsilon), \quad x(t_0) = \xi(\varepsilon) \tag{2.19}$$

$$\varepsilon \dot{y} = g(t,x,z,\varepsilon), \quad z(t_0) = \eta(\varepsilon) \tag{2.20}$$

where $\xi(\varepsilon)$ and $\eta(\varepsilon)$ depend smoothly on ε and $t_0 \in [0,t_1)$. Additionally, let $x(t,\varepsilon)$, $z(t,\varepsilon)$ represent the solution of the system.

The reduced model (see Equation (2.18)) is given by:

$$\dot{x} = f(t,x,h(t,x),0), \quad x(t_0) = \xi_0 \triangleq \xi(0) \tag{2.21}$$

Let $\bar{x}(t)$ be the solution of Equation (2.21). The quasi-steady state behaviour of z when

27

$x = \bar{x}$ can be obtained as:

$$\bar{z} \triangleq h(t, \bar{x}(t)) \tag{2.22}$$

Performing the change of variables $y = z - h(t,x)$ and assuming $\varepsilon = 0$, Equation (2.20) becomes:

$$\frac{dy}{d\tau} = g(t_0, \xi_0, y + h(t_0, \xi_0), 0), \quad y(0) = \eta(0) - h(t_0, \xi_0) \triangleq \eta_0 - h(t_0, \xi_0) \tag{2.23}$$

where $\tau = \dfrac{t - t_0}{\varepsilon}$ is the new time variable (fast time-scale).

Assuming that $\bar{x}(t)$ is defined for $t \in [0, t_1]$ and $\bar{x}(t) \in D_x \subset \mathbb{R}^n$, for some domain D_x, Equation (2.23) can be rewritten as:

$$\frac{dy}{d\tau} = g(t, x, y + h(t,x), 0) \tag{2.24}$$

where the slowly varying parameters $(t, x) \in [0, t_1] \times D_x$ are treated as fixed parameters.

Equation (2.24) is referred to as the boundary-layer model/system. Note that the same terminology is often used for Equation (2.23) which constitutes an evaluation of Equation (2.24) for some given initial time and state.

Theorem 2.1.14 (Theorem 11.1 in [43])**.** *Consider the singular perturbation problem of Equations (2.19) and (2.20) and let $z = h(t,x)$ be an isolated root of Equation (2.16). Assume that the following conditions are satisfied for all*

$$[t, x, z - h(t,x), \varepsilon] \in [0, t_1] \times D_x \times D_y \times [0, \varepsilon_0]$$

for some domains $D_x \subset \mathbb{R}^n$ and $D_y \subset \mathbb{R}^m$, in which D_x is convex and D_y contains the origin:

- *The functions f and g, their first partial derivatives with respect to (x, z, ε), and the first partial derivative of g with respect to t are continuous; the function $h(t,x)$ and the Jacobian $\left[\dfrac{\partial g(t, x, z, 0)}{\partial z} \right]$ have continuous first partial derivatives with respect to*

28

their arguments; the initial date $\xi(\varepsilon)$ and $\eta(\varepsilon)$ are smooth functions of ε.

- *The reduced problem (2.21) has a unique solution $\bar{x}(t) \in S$, for $t \in [t_0, t_1]$, where S is a compact subset of D_x.*

- *The origin is an exponentially stable equilibrium point of the boundary-layer model (2.24), uniformly in (t,x); let $R_y \subset D_y$ be the region of attraction of Equation (2.23) and Ω_y be a compact subset of R_y.*

Then, there exists a positive constant ε^* such that for all $\eta_0 - h(t_0, \xi_0) \in \Omega_y$ and $0 < \varepsilon < \varepsilon^*$, the singular perturbation problem of Equations (2.19) and (2.20) has a unique solution $x(t, \varepsilon)$, $z(t, \varepsilon)$ on $[t_0, t_1]$ and

$$x(t, \varepsilon) - \bar{x}(t) = \mathscr{O}(\varepsilon)$$

$$z(t, \varepsilon) - h(t, \bar{x}(t)) - \hat{y}(t/\varepsilon) = \mathscr{O}(\varepsilon)$$

hold uniformly for $t \in [t_0, t_1]$, where $\hat{y}(\tau)$ is the solution of the boundary-layer model (2.23). Moreover, given any $t_b > t_0$, there is $\varepsilon^{**} \leq \varepsilon^*$ such that

$$z(t, \varepsilon) - h(t, \bar{x}(t)) = \mathscr{O}(\varepsilon)$$

holds uniformly for $t \in [t_b, t_1]$ whenever $\varepsilon < \varepsilon^{**}$.

Note that the region of attraction mentioned above refers to the set of all initial conditions that converge to a given asymptotically stable equilibrium point [8].

Theorem 2.1.14 is also known as Tikhonov's theorem.

A comprehensive treatment of singular perturbation methods can be found in [77, 44, 43].

2.2 Biological Background

To facilitate the comprehension of the experimental implementations discussed in this thesis, here a number of fundamental concepts in molecular biology are briefly introduced. The definitions presented in this section are adopted from [53] and, for the reader's convenience, are listed alphabetically.

Activator Specific transcription factor that stimulates transcription.

ATP (adenosine triphosphate) A nucleotide that is the most important molecule for capturing and transferring free energy in cells. Hydrolysis of each of the two phospho-anhydride bonds in ATP releases a large amount of free energy that can be used to drive energy-requiring cellular processes.

Amino acid An organic compound containing at least one amino group and one carboxyl group. In the amino acids that are the monomers for building proteins, an amino group and carboxyl group are linked to a central carbon atom, the α carbon, to which a variable side chain is attached.

Antibody A protein (immunoglobulin), normally produced in response to an antigen, that interacts with a particular site (epitope) on the same antigen and facilitates its clearance from the body.

Antigen Any material (usually foreign) that elicits an immune response.

Bacteria Class of prokaryotes that constitutes one of the three distinct evolutionary lineages of modern-day organisms; also called eubacteria. Phylogenetically distinct from archaea and eukaryotes.

Bacteriophage (phage) Any virus that infects bacterial cells. Some phages are widely used as vectors in DNA cloning.

Base Any compound, often containing nitrogen, that can accept a proton (H^+) from an acid. Also, commonly used to denote the purines and pyrimidines in DNA and RNA.

Base pair Association of two complementary nucleotides in a DNA or RNA molecule stabilized by hydrogen bonding between their base components. Adenine pairs with thymine or uracil ($A \cdot T$, $A \cdot U$) and guanine pairs with cytosine ($G \cdot C$).

Catalyst A substance that increases the rate of a chemical reaction without undergoing a permanent change in its structure. Enzymes are proteins with catalytic activity, and ribozymes are RNAs that can function as catalysts.

Cell cycle Ordered sequence of events in which a eukaryotic cell duplicates its chromosomes and divides into two. The cell cycle normally consists of four phases: G1 before DNA synthesis occurs; S when DNA replication occurs; G2 after DNA synthesis; and M when cell division occurs, yielding two daughter cells. Under certain conditions, cells exit the cell cycle during G1 and remain in the G0 state as nondividing cells.

Cell division Separation of a cell into two daughter cells. In higher eukaryotes, it involves division of the nucleus (mitosis) and of the cytoplasm (cytokinesis); mitosis often is used to refer to both nuclear and cytoplasmic division.

Cell strain A population of cultured cells, of plant or animal origin, that has a finite life span and eventually dies, commonly after 25–50 generations.

Chaperone Collective term for two types of proteins — molecular chaperones and chaperonins — that prevent misfolding of a target protein or actively facilitate proper folding of an incompletely folded target protein, respectively.

Chemotaxis Movement of a cell or organism toward or away from certain chemicals.

Constitutive Referring to the continuous production or activity of a cellular molecule or the continuous operation of a cellular process (e.g., constitutive secretion) that is not regulated by internal or external signals.

Cytoplasm Viscous contents of a cell that are contained within the plasma membrane but, in eukaryotic cells, outside the nucleus.

DNA (deoxyribonucleic acid) Long linear polymer, composed of four kinds of deoxyribose nucleotides, that is the carrier of genetic information.

DNA-binding domain The domain of a transcription factor that binds specific, closely related DNA sequences.

DNA polymerase An enzyme that copies one strand of DNA (the template strand) to make the complementary strand, forming a new double-stranded DNA molecule.

Double helix, DNA The most common three-dimensional structure for cellular DNA in which the two polynucleotide strands are antiparallel and wound around each other with complementary bases hydrogen-bonded.

Enzyme A protein that catalyzes a particular chemical reaction involving a specific substrate or small number of related substrates.

Gene Physical and functional unit of heredity, which carries information from one generation to the next. In molecular terms, it is the entire DNA sequence — including exons, introns, and transcription-control regions — necessary for production of a functional polypeptide or RNA.

Gene control All of the mechanisms involved in regulating gene expression. Most common is regulation of transcription, although mechanisms influencing the processing, stabilization, and translation of mRNAs help control expression of some genes.

Gene expression Overall process by which the information encoded in a gene is converted into an observable phenotype (most commonly production of a protein).

Genetic code The set of rules whereby nucleotide triplets (codons) in DNA or RNA specify amino acids in proteins.

In vitro Referring to experiments or manipulations performed outside a cell (including cell fragments, lysates, or purified molecules) or to cells placed in an artificial environment such as in a petri dish or test tube; literally, *in glass*.

In vivo Referring to experiments or manipulations performed in the context of an intact organism or intact cell, in contrast to experiments using cell fragments, lysates, or purified molecules; literally, *in the living*.

Monomer Any small molecule that can be linked chemically with others of the same

type to form a polymer. Examples include amino acids, nucleotides, and monosaccharides.

mRNA (messenger RNA) Any RNA that specifies the order of amino acids in a protein (i.e., the primary structure). It is produced by transcription of DNA by RNA polymerase. In eukaryotes, the initial RNA product (primary transcript) undergoes processing to yield functional mRNA.

Nucleic acid A polymer of nucleotides linked by phosphodiester bonds. DNA and RNA are the primary nucleic acids in cells.

Nucleotide A nucleoside with one or more phosphate groups linked via an ester bond to the sugar moiety, generally to the $5'$ carbon atom. DNA and RNA are polymers of nucleotides containing deoxyribose and ribose, respectively.

Peptide A small linear polymer composed of amino acids connected by peptide bonds. The terms peptide and oligopeptide are often used interchangeably.

Plasmid Small, circular extrachromosomal DNA molecule capable of autonomous replication in a cell.

Polymer Any large molecule composed of multiple identical or similar units (monomers) linked by covalent bonds.

Polypeptide Linear polymer of amino acids connected by peptide bonds, usually containing 20 or more residues.

Promoter DNA sequence that determines the site of transcription initiation for an RNA

polymerase.

Protease Any enzyme that cleaves one or more peptide bonds in target proteins.

Protein A macromolecule composed of one or more linear polypeptide chains and folded into a characteristic three-dimensional shape (conformation) in its native, biologically active state.

Repressor Specific transcription factor that inhibits transcription.

Residue General term for the repeating units in a polymer that remain after covalent linkage of the monomeric precursors.

Ribosome A large complex comprising several different rRNA molecules and as many as 83 proteins, organized into a large subunit and small subunit; the engine of translation (protein synthesis).

RNA (ribonucleic acid) Linear, single-stranded polymer, composed of ribose nucleotides. mRNA, rRNA, and tRNA play different roles in protein synthesis; a variety of small RNAs play roles in controlling the stability and translation of mRNAs and in controlling chromatin structure and transcription.

RNA polymerase An enzyme that copies one strand of DNA (the *template* strand) to make the complementary RNA strand using as substrates ribonucleoside triphosphates.

rRNA (ribosomal RNA) Any one of several large RNA molecules that are structural and functional components of ribosomes.

Substrate Molecule that undergoes a charge in a reaction catalyzed by an enzyme.

Transcription Process in which one strand of a DNA molecule is used as a template for synthesis of a complementary RNA by RNA polymerase.

Transcription factor (TF) General term for any protein, other than RNA polymerase, required to initiate or regulate transcription in eukaryotic cells. *General* factors, required for transcription of all genes, participate in formation of the transcription-preinitiation complex near the start site. Specific factors stimulate (activators) or inhibit (repressors) transcription of particular genes by binding to their regulatory sequences.

Translation The ribosome-mediated assembly of a polypeptide whose amino acid sequence is specified by the nucleotide sequence in an mRNA.

tRNA (transfer RNA) A group of small RNA molecules that function as amino acid donors during protein synthesis. Each tRNA becomes covalently linked to a particular amino acid, forming an aminoacyl-tRNA.

Virus A small intracellular parasite, consisting of nucleic acid (RNA or DNA) enclosed in a protein coat, that can replicate only in a susceptible host cell; widely used in cell biology research.

For a full treatment of the principles and phenomena encountered in the biological parts of this work, the reader is referred to [53, 46, 3].

Chapter 3

Biomolecular mechanisms for signal differentiation

Article

Biomolecular mechanisms for signal differentiation
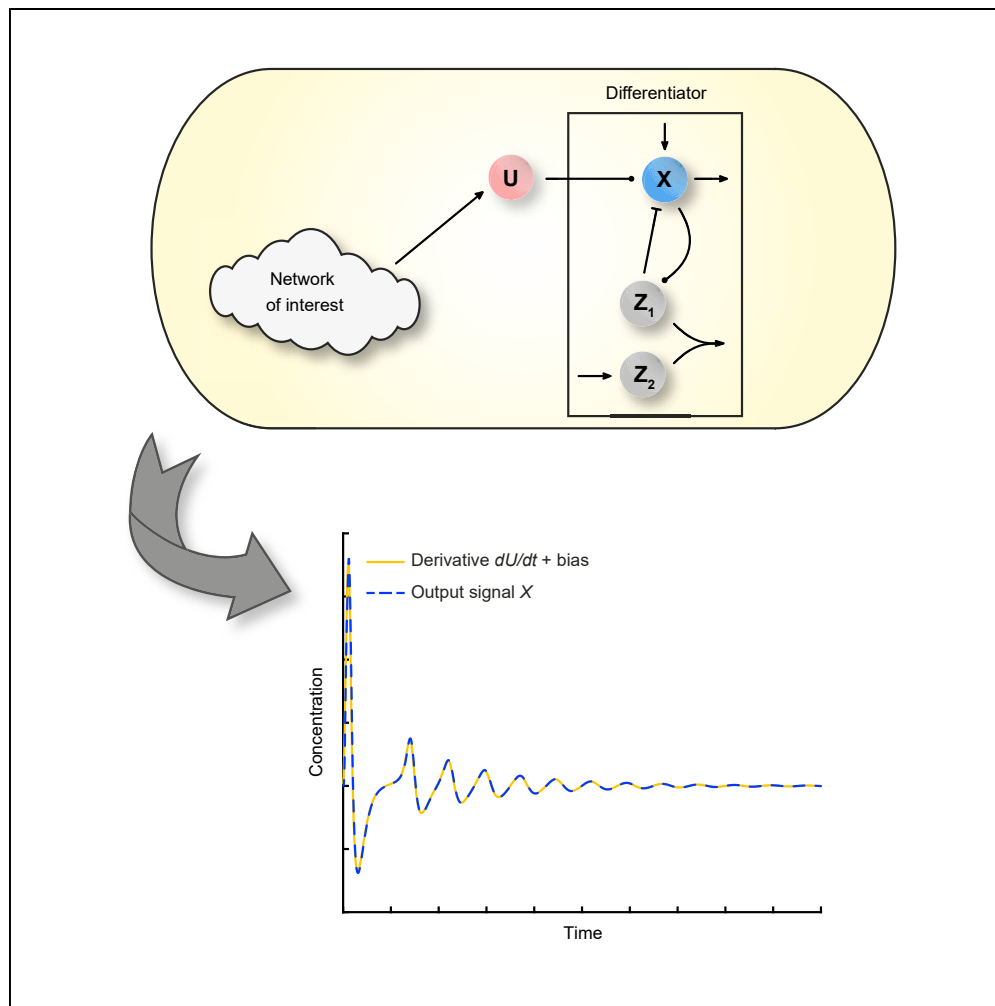
Emmanouil Alexis,
Carolin C.M.
Schulte, Luca
Cardelli, Antonis
Papachristodoulou

antonis@eng.ox.ac.uk

Highlights

Calculating the speed or
higher derivatives of
molecular signals has
been a challenge

Biomolecular topologies
acting as signal
differentiators are studied

High performance and
insensitivity to high-
frequency input signals
can be achieved

Natural circuits and
synthetic realizations for
signal differentiation are
described

Article

Biomolecular mechanisms for signal differentiation

Emmanouil Alexis,[1] Carolin C.M. Schulte,[1,2] Luca Cardelli,[3] and Antonis Papachristodoulou[1,4,*]

SUMMARY

Cells can sense temporal changes of molecular signals, allowing them to predict environmental variations and modulate their behavior. This paper elucidates biomolecular mechanisms of time derivative computation, facilitating the design of reliable synthetic differentiator devices for a variety of applications, ultimately expanding our understanding of cell behavior. In particular, we describe and analyze three alternative biomolecular topologies that are able to work as signal differentiators to input signals around their nominal operation. We propose strategies to preserve their performance even in the presence of high-frequency input signal components which are detrimental to the performance of most differentiators. We find that the core of the proposed topologies appears in natural regulatory networks and we further discuss their biological relevance. The simple structure of our designs makes them promising tools for realizing derivative control action in synthetic biology.

INTRODUCTION

Measuring the speed at which a physical process evolves over time is of central importance to science and engineering. This can be done by computing the time derivative of the function describing the process. Several examples of cellular systems exhibiting derivative action indicate that calculating the rate of change of biological processes is essential in nature. The retina of our eyes, for instance, is one of the best-studied neural networks of the brain. Its response to changes in light intensity reveals typical characteristics of derivative action which stem from the interaction between cone and horizontal cells (Wilson, 1999; Åström and Murray, 2021). In microbiology, the chemotaxis signaling pathway in bacteria such as *Escherichia coli* involves computation of time derivatives: To navigate toward nutrients and away from toxins, bacteria are able to sample their environment as they move and convert spatial gradients into temporal ones (Alon, 2019; Shimizu et al., 2010; Iglesias and Devreotes, 2008; Barkai and Leibler, 1997; Block et al., 1983; Macnab and Koshland, 1972). Furthermore, in the context of cellular energy metabolism, *in silico* studies have revealed the role of creatine phosphate as a buffering species that allows for adaptation to a changing demand of adenosine triphosphate (ATP), thus exploiting the anticipatory action enabled by derivative control (Cloutier and Wellstead, 2010). This observation is a specific example of a broader class of biomolecular processes where the presence of rapid buffering proves to be equivalent to negative derivative feedback (Hancock et al., 2017).

In traditional engineering, differentiators refer to devices capable of applying time differentiation to an input stimulus, for example a mechanical or electrical signal. In the rapidly growing field of synthetic biology, the ability to build reliable biomolecular differentiators would offer considerable advantages (Steel et al., 2017; Del Vecchio et al., 2016; Lu et al., 2009). As an immediate application, such genetic circuits would be able to track the rate of change of the concentration of biomolecules, thus acting as speed biosensors. This is of interest when assessing uptake rates of certain molecules, such as uptake of pollutants into bacteria used for bioremediation (Chen and Wilson, 1997; Pieper and Reineke, 2000). They can also allow for advanced regulation strategies in the cellular environment by enabling the construction of more efficient bio-controllers, e.g., Proportional-Integral-Derivative (PID) control schemes, the workhorses of modern technological process control applications (Åström and Murray, 2021). In general, derivative control can enhance the stability of a feedback system and provide a smoother transient response.

Recent efforts in this rather underexplored research area include the design of a differentiator module consisting of linear input/output functions realized by specific processes of protein production (Halter et al., 2017; Halter et al., 2019). It has further been demonstrated that calculation of time derivatives is possible by using ultrasensitive topologies operating within a negative feedback loop (Samaniego et al., 2019), and a motif capable of computing positive and negative temporal gradients, which includes input delays and

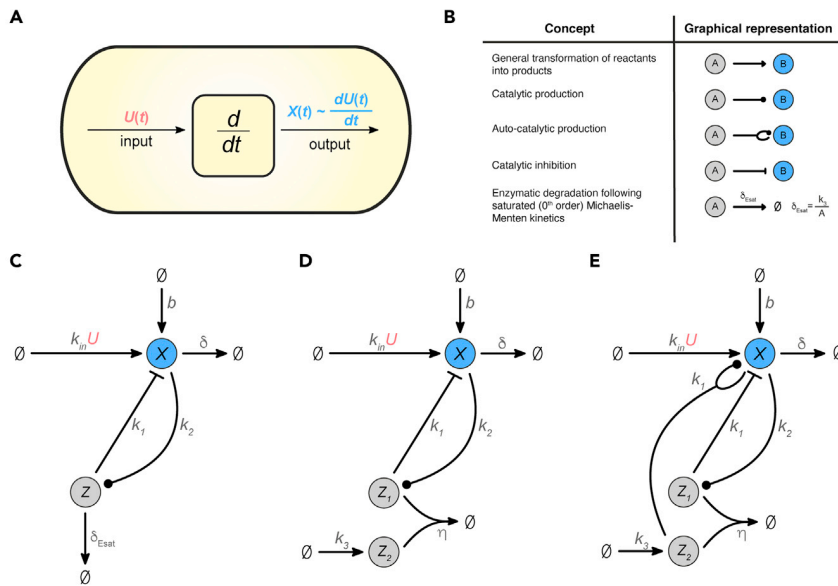[1]Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK

[2]Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, UK

[3]Department of Computer Science, University of Oxford, Oxford OX1 3QD, UK

[4]Lead contact

*Correspondence: antonis@eng.ox.ac.uk

https://doi.org/10.1016/j.isci.2021.103462

 1

Figure 1. Biomolecular structures capable of signal differentiation

(A) Schematic representation of the notion of signal differentiation carried out by a biomolecular device inside the cell.

(B) Graphical representation of the biological concepts found in the signal differentiator motifs. To describe the different kind of biomolecular reactions the following notation is adopted: (→) means that the transformation of reactants into products only happens in the direction of the arrow. (—•) indicates that reactants enable product formation without being consumed. (—|) denotes inhibition of products by a reactant where the reactant is not consumed. In addition, the depicted concept of enzymatic degradation is further analyzed in STAR Methods Equilibria and stability of biomolecular signal differentiators: Biomolecular Signal Differentiator-I.

(C) Topology of Biomolecular Signal Differentiator - I or BioSD-I (Equation (1)).

(D) Topology of Biomolecular Signal Differentiator - II or BioSD-II (Equation (2)).

(E) Topology of Biomolecular Signal Differentiator - III or BioSD-III (Equation (3)).

the idea of an incoherent feedforward loop, has been presented (Samaniego et al., 2020). With the aim of providing derivative action in PID control architectures, networks directly inspired by bacterial chemotaxis (Chevalier et al., 2019) or based on the so-called dual rail encoding have also been proposed (Whitby et al., 2021; Paulino et al., 2019). This approach enables the representation of both positive and negative signals via biomolecular species by decomposing a signal into two non-negative parts (Oishi and Klavins, 2011). Finally, a derivative controller tailored to gene expression is analyzed in (Modi et al., 2019), while in the PID architecture introduced in (Filo and Khammash, 2021), derivative control is carried out with inseparable connection to proportional and integral actions.

In this article, we introduce novel differentiator modules aiming to elucidate unexplored mechanisms that cells potentially exploit to achieve signal differentiation. In parallel, these motifs can pave the way for designing efficient and reliable synthetic signal differentiator devices in a cellular context. Notably, our motifs offer considerable ease of experimental implementation compared to some of the earlier discussed designs which are based on more "artificial" mechanisms such as dual-rail encoding. In addition, the motifs under consideration can function as independent, general-purpose differentiators, which may be a challenging task for other topologies, such as some control-oriented topologies showing derivative action. Moreover, under suitable tuning high accuracy of temporal derivative calculation for a wide range of molecular signals can be guaranteed.

Specifically, we present three biomolecular architectures capable of functioning as signal differentiators around their equilibria. We call them Biomolecular Signal Differentiators (BioSD). Each of these networks can be interpreted as a modular and tunable topology inside the cell that accepts a molecular signal as an input and produces an output signal proportional to the time derivative of the input signal (Figure 1A). The output corresponds to a biochemical species whose concentration can be measured. The proposed architectures provide simple blueprints for the design of synthetic biomolecular differentiators, but can also be interpreted as lenses through which derivative action in natural systems can be identified and studied.

41

We demonstrate the special characteristics and the performance trade-offs of the three BioSD architectures (BioSD-I, II, and III) via theoretical analyses and numerical simulations. We also discuss a major obstacle of both technological and biological differentiators, namely amplification of undesired high-frequency components of the input signal, and propose strategies to overcome this obstacle. Finally, we show the occurrence of one of the BioSD topologies in natural regulatory networks involved in bacterial adaptation to stress conditions and present potential synthetic implementations for all three topologies, highlighting the biological relevance of our designs.

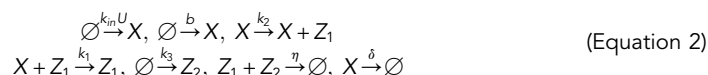## RESULTS

Biological structure

We begin by presenting the molecular interactions in the BioSD circuits as chemical reaction networks (CRNs). These circuits represent three alternative topological entities which, under certain assumptions, realize the same concept of signal differentiation. In the analysis that follows, the input and output signals of the differentiators are generally treated as biomolecular species, namely U and X respectively. Nevertheless, an input signal may also refer to different concepts such as light, temperature or pH.

Figure 1C illustrates the first architecture, BioSD-I, which consists of the following reactions:

$$\varnothing \xrightarrow{k_{in}U} X, \ \varnothing \xrightarrow{b} X, \ X \xrightarrow{k_2} X + Z$$
$$X + Z \xrightarrow{k_1} Z, \ X \xrightarrow{\delta} \varnothing, \ Z \xrightarrow{\delta_{Esat}} \varnothing$$
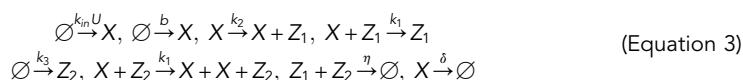
(Equation 1)

Here, the production of output species X depends on two reactions. One of them has a constant rate while the other occurs at a rate proportional to the concentration of input species U. It is convenient to represent such processes via reactions of the form $\varnothing \xrightarrow{r} X$, where r can be a constant or a time-varying quantity, e.g., biomolecular concentration. This allows us to describe general concepts of production without the need to specify their impact on the reactants involved. Furthermore, X also catalyzes the formation of species Z which, in turn, inhibits X. Note that the process of inhibition is interpreted as catalysis of degradation. Finally, the removal rate of X is proportional to its concentration (first-order decay) while, as indicated by the notation δ_{Esat} (defined in Figure 1B), Z adheres to a constant rate of decay (0th-order decay). The latter behavior is attained through enzyme-catalyzed degradation of Z where the enzyme is operating at saturating substrate levels (for more details see STAR Methods Equilibria and stability of biomolecular signal differentiators: biomolecular signal differentiator-I).

In the second architecture, BioSD-II (Figure 1D), the formation process of output species X is the same as in BioSD-I, while Z_1, the production of which is facilitated by X, and Z_2 annihilate each other. Z_1 inhibits X which decays in the same way as in BioSD-I. The reactions that form the corresponding CRN are:

$$\varnothing \xrightarrow{k_{in}U} X, \ \varnothing \xrightarrow{b} X, \ X \xrightarrow{k_2} X + Z_1$$
$$X + Z_1 \xrightarrow{k_1} Z_1, \ \varnothing \xrightarrow{k_3} Z_2, \ Z_1 + Z_2 \xrightarrow{\eta} \varnothing, \ X \xrightarrow{\delta} \varnothing$$

(Equation 2)

Finally, Figure 1E shows the third topology, BioSD-III, which is described by the reactions:

$$\varnothing \xrightarrow{k_{in}U} X, \ \varnothing \xrightarrow{b} X, \ X \xrightarrow{k_2} X + Z_1, \ X + Z_1 \xrightarrow{k_1} Z_1$$
$$\varnothing \xrightarrow{k_3} Z_2, \ X + Z_2 \xrightarrow{k_1} X + X + Z_2, \ Z_1 + Z_2 \xrightarrow{\eta} \varnothing, \ X \xrightarrow{\delta} \varnothing$$

(Equation 3)

This CRN includes an autocatalytic-like reaction: X is able to produce more of itself in the presence of Z_2. The rest of its structure is identical to the CRN of BioSD-II.

Mathematical description

We now derive the dynamics of the proposed BioSD networks using the law of mass action (Del Vecchio and Murray, 2015) unless otherwise stated, adopting the same order of presentation as in the preceding section.

BioSD-I (CRN given by Equation (1)) can be described by the following system of Ordinary Differential Equations (ODEs):

$$\dot{X} = k_{in}U + b - k_1 XZ - \delta X$$

(Equation 4a)

$$\dot{Z} = k_2 X - k_3$$

(Equation 4b)

Note that the enzymatic degradation of Z is assumed to follow saturated (0th-order) Michaelis-Menten kinetics, as previously discussed.

Next, from the CRN given by Equation (2) we obtain the following ODE model for BioSD-II:

$$\dot{X} = k_{in}U + b - k_1XZ_1 - \delta X \qquad \text{(Equation 5a)}$$
$$\dot{Z}_1 = k_2X - \eta Z_1Z_2 \qquad \text{(Equation 5b)}$$
$$\dot{Z}_2 = k_3 - \eta Z_1Z_2 \qquad \text{(Equation 5c)}$$

For the last circuit, BioSD-III, the CRN given by Equation (3) can be modeled using the following ODEs:

$$\dot{X} = k_{in}U + b - k_1XZ_1 + k_1XZ_2 - \delta X \qquad \text{(Equation 6a)}$$
$$\dot{Z}_1 = k_2X - \eta Z_1Z_2 \qquad \text{(Equation 6b)}$$
$$\dot{Z}_2 = k_3 - \eta Z_1Z_2 \qquad \text{(Equation 6c)}$$

By assuming a constant input U^* and setting the derivatives to zero, we can show that each of the BioSD network models has a unique equilibrium. In addition, we can prove through linearization that the equilibrium is locally exponentially stable (a detailed analysis can be found in STAR Methods Equilibria and stability of biomolecular signal differentiators). Near their steady-states, the circuits are able to exhibit derivative action, as shown in the next section. Furthermore, for the purpose of this study we assume that the parameter η in BioSD-II is sufficiently large which can lead to a practically insignificant concentration of species Z_2 (more details can be found in STAR Methods The notion of strong rate of annihilation between Z_1, Z_2 (large η) in biomolecular signal differentiator-II). This constraint does not have to hold for BioSD-III, which includes the same annihilation reaction. Finally, Equations (5b) and (5c) indicate that in case $\dot{Z}_2 \approx 0$, the removal rate of Z_1 is roughly constant and equal to k_3, similar to the 0th-order removal of Z in BioSD-I.

Achieving biological signal differentiation

In order for the proposed biomolecular modules to work as signal differentiators, we desire for their output X to be proportional to the derivative of their input U. This immediately raises the following challenge: Both U and X refer to biomolecular species concentrations and, by extension, represent non-negative signals. However, in the general case, the derivative of a nonnegative signal can take negative values and, as a result, X would need to go below zero. Thus, it could be argued that X is unable to express the rate of change of an arbitrary input signal. An obvious way to overcome this obstacle is to add a bias to the computed derivative. As we demonstrate here, the perfect candidate for realizing this bias is the steady state of X around which derivative action can be achieved.

We are interested in the local behavior of the BioSD networks and, therefore, consider input stimuli that do not force them to operate far away from their equilibrium. Subsequently, we assume that every input signal can be described as:

$$U = U^* + U^{TV} \qquad \text{(Equation 7)}$$

where U^* is constant while U^{TV} is time-varying. Here, we focus on Fourier transformable signals which is typically the case for physical signals in practical applications (for more details see STAR Methods Signals under consideration).

By linearizing and applying appropriate transformations, we can show that the dynamics of the output of any of the three BioSD topologies presented in the previous section can be approximated by the following non-dimensional second - order differential equation (see STAR Methods Behavior analysis of biomolecular signal differentiators):

$$\varepsilon \ddot{x}_n + \varepsilon \dot{x}_n + x_n = \dot{u}_n \qquad \text{(Equation 8)}$$

where x_n and u_n refer to the output and input, respectively and:

$$\varepsilon = \frac{k_2^2}{k_1 k_3^3}(k_{in}U^* + b)^2 \qquad \text{(Equation 9)}$$

Equation (8) represents a signal differentiator accompanied with some filtering action. Indeed, the input/output relation in the Laplace domain can be described by the following transfer function (Oppenheim et al., 1996):

$$\tilde{\Delta}_{BSD}(s) = \frac{\tilde{X}_n(s)}{\tilde{U}_n(s)} = \frac{s}{\varepsilon(s^2 + s) + 1} \qquad \text{(Equation 10)}$$

where $\tilde{X}_n(s)$ and $\tilde{U}_n(s)$ are the Laplace transform of the output x_n and input u_n, respectively and s is the Laplace variable (complex frequency). As can be seen from Equation (10), a BioSD network is the series combination of an ideal differentiator and a second-order low pass filter (Samoilov et al., 2002). Therefore, for a given positive ε, the accuracy of signal differentiation depends on the frequency spectrum of the input signal or, in other words, the range of frequencies contained by it (see STAR Methods Signals under consideration). Accompanying a differentiator with a low-pass filter is a widely used strategy in traditional engineering in order to deal with high-frequency input noise (this topic is analyzed in Response to input signals corrupted by high-frequency noise and A structural addition for enhanced performance).

To gain a deeper insight, we calculate the Fourier transform (Oppenheim et al., 1996) of the output:

$$\tilde{X}_n(j\omega) = \tilde{\Delta}_{BSD}(j\omega)\tilde{U}_n(j\omega) \qquad \text{(Equation 11)}$$

where ω represents the frequency, j is the imaginary unit number ($j = \sqrt{-1}$) and $\tilde{X}_n(j\omega)$, $\tilde{U}_n(j\omega)$ are the Fourier transform of the output x_n and input u_n, respectively. Furthermore, $\tilde{\Delta}_{BSD}(j\omega)$ is the Fourier transform of the system's impulse response, also known as the frequency response of the system. (ibid.). Since we have a linear, asymptotically stable, system we can compute the latter Fourier transform from Equation (10) by setting $s = j\omega$. Thus, we have:

$$\tilde{X}_n(j\omega) = \frac{j\omega}{\varepsilon(-\omega^2 + j\omega) + 1}\tilde{U}_n(j\omega) \qquad \text{(Equation 12)}$$

The operation of (ideal) differentiation in the frequency domain is defined as:

$$\tilde{X}_{nd}(j\omega) = j\omega\tilde{U}_n(j\omega) \qquad \text{(Equation 13)}$$

To compare the output of an ideal differentiator to the one of a BioSD device, we introduce the following performance metric:

$$\tilde{\Lambda}(j\omega) = \frac{\tilde{X}_n(j\omega)}{\tilde{X}_{nd}(j\omega)} = \frac{1}{\varepsilon(-\omega^2 + j\omega) + 1} \qquad \text{(Equation 14)}$$

Using the magnitude-phase representation of Equation (14) we get:

$$|\tilde{\Lambda}(j\omega)| = \frac{1}{\sqrt{\varepsilon^2\omega^2 + (1 - \varepsilon\omega^2)^2}} \qquad \text{(Equation 15)}$$

and

$$\angle\,\tilde{\Lambda}(j\omega) = \arctan\left(\frac{-\varepsilon\omega}{1 - \varepsilon\omega^2}\right)$$

Signal differentiation of high accuracy is carried out when $\tilde{\Lambda}(j\omega)$ is close to $1\angle0°$. As shown in Figure 2, there is a "low-frequency" range where this is true, but as ε decreases the aforementioned range expands toward "higher frequencies". In the time domain this entails that for a given positive ε, a BioSD device can work as an accurate signal differentiator for sufficiently slow input signals and, in that case, the BioSD output can be approximated by (see STAR Methods Behavior analysis of biomolecular signal differentiators):

$$X = \frac{k_{in}}{k_1 k_3}\dot{U} + \frac{k_3}{k_2} \qquad \text{(Equation 16)}$$

There is a family of input signals for which the BioSD topologies are able to provide accurate differentiation regardless of the exact value of ε (see STAR Methods Behavior analysis of biomolecular signal differentiators). More specifically, this holds for input signals for which the term U^{TV} in Equation (7) is of the form:

$$U^{TV} = \xi_1 e^{-\xi_3 t} + \xi_2 t, \qquad \text{(Equation 17)}$$

44
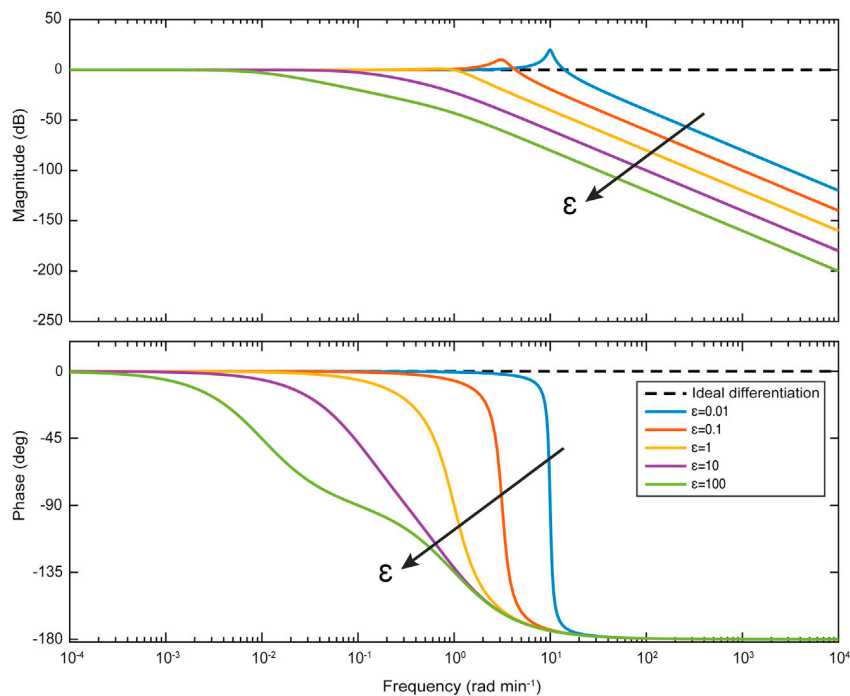
Figure 2. A performance metric for Biomolecular Signal Differentiators in the frequency domain

Bode plot of the metric given by Equation (14). Different colors represent the magnitude and the phase of the corresponding transfer function for different values of ε. The case of ideal differentiation corresponds to $\varepsilon = 0$ and the direction in which the latter increases indicated by an arrow.

where ξ_1, ξ_2 are arbitrary constants and $\xi_3 = \frac{k_2}{k_3}(k_{in}U^* + b)$. If ξ_2 is not zero which implies linear growth over time, we assume that the above holds as long as the system stays near its equilibrium. This means that the term $\xi_2 t$ is sufficiently small. Indeed, several biological processes can generate (bounded) signals some part of which can be viewed as linear growth (Del Vecchio and Murray, 2015). We study such a scenario in Sensing the response speed of biomolecular networks.

As Equation (8) states, the response of a BioSD network, is given as the solution of a second-order non-ho-mogeneous differential equation with constant coefficients where the forcing function is $\dot{u}_n$. The response can therefore be seen as the sum of two terms: a "transient" term which highly depends on the initial con-ditions and dies out with time; and a "steady-state" term which, under the conditions discussed above, can approximate the derivative of the input signal (Zill, 2012). Therefore, for input signals applied for a suffi-ciently long time, the BioSD output practically coincides with the latter since the effect of the former is negligible. However, this may not be always the case for short duration input signals where any undesired initial transient phenomena can greatly compromise the accuracy of the differentiator output.

From Equation (16), we can see that the BioSD modules use the biomolecular concentration $\frac{k_3}{k_2}$ as a bias. Around this point they can operate as signal differentiators, producing an output signal component which is proportional to the derivative of the input. The bias therefore depends only on two parameters which, ideally, can be adjusted as desired. This provides us with the freedom of choosing any (fixed) concentration of X as a bias, which will remain unchanged regardless of the rest of the model parameters, the input stimulus, or potential constant disturbances on the output. To appreciate this further, we recall the production reaction for X with constant rate b, which is included in each of the proposed CRNs. Besides its role as a structural requirement, this production reaction can also represent an external constant disturbance applied on X; this, however, does not affect the zero-level we choose for our measurements. Once the concentration of X reaches this level, it will stay there until an input excitation appears and it will come back once the excitation stops. Hence, the previously mentioned fixed con-centration can also be seen as a "rest position" for the differentiators.

The feature just described is of key importance and stems mainly from the following two sources: the stability that characterizes BioSDs and the fact that the steady-state of the output coincides with the aforementioned zero-level concentration. The latter is achieved due to integration carried out by the 'memory' function which is realized via species Z within BioSD-I and the quantity $Z_1 - Z_2$ within BioSD-II, III.

Tunability and accuracy

It is convenient for the circuit designer who aims to implement the BioSD topologies to be able to choose the parameter values and ensure that the resulting differentiators meet the expected performance requirements. Nonetheless, there may be cases where the number of system parameters that can be suitably tuned is limited, for instance due to constraints related to the cellular processes involved in the circuits under investigation. Even in this case, the architecture of our circuits allows for some tunability as long as the designer can choose some crucial parameters.
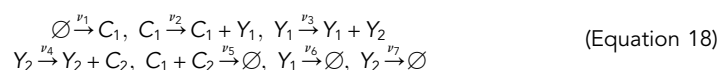
Consider for example the extreme scenario where only one of the model parameters can be regulated. If this parameter is k_3, then, according to Equation (16), its appropriate tuning may result in an acceptable gain by which the output signal is multiplied (output gain) and bias based on which this signal is measured. At the same time, Equation (9) reveals that (contrary to other parameters) a small change in k_3 can affect ε significantly since the latter is inversely proportional to the cube of k_3.

It immediately emerges from the above that the way we tune the BioSD networks defines the level of accuracy regarding their derivative action. Indeed, ε is subject to almost all parameter rates in these networks and, as pointed out in the previous section, the value of ε defines the range of frequencies over which BioSDs can accurately compute the rate of change of a biological signal.

Sensing the response speed of biomolecular networks

We now demonstrate through an example the ability of BioSD modules to compute the temporal derivative of biological signals. At the same time, we highlight one of their potential applications discussed above, namely as rate-of-change detectors or speed biosensors.

We consider the antithetic motif (Figure 3) (Briat, Gupta, and Khammash, 2016, 2018; Chevalier et al., 2019; Olsman et al., 2019a, 2019b; Olsman and Forni, 2020; Baetica et al., 2020):

$$\varnothing \xrightarrow{\nu_1} C_1, \ C_1 \xrightarrow{\nu_2} C_1 + Y_1, \ Y_1 \xrightarrow{\nu_3} Y_1 + Y_2$$
$$Y_2 \xrightarrow{\nu_4} Y_2 + C_2, \ C_1 + C_2 \xrightarrow{\nu_5} \varnothing, \ Y_1 \xrightarrow{\nu_6} \varnothing, \ Y_2 \xrightarrow{\nu_7} \varnothing$$

(Equation 18)

Species Y_1, Y_2 represent an arbitrary biological process whose output, Y_2, can be robustly steered toward a desired value $\left(\dfrac{\nu_1}{\nu_4}\right)$. This is feasible through the feedback integral control which is implemented via species C_1, C_2, thus achieving robust perfect adaptation. Depending on the parameter rates, the dynamics of the above architecture can be either stable or unstable. Nonetheless, even in a stable system, the species of interest, Y_2, sometimes displays a long-lasting transient response with damped oscillations before it settles to a steady-state. This provides an opportunity to assess the ability of the BioSD networks to calculate the speed at which these oscillations evolve.

In order for a BioSD device to function as a biosensor for the CRN given by (Equation 18), a suitable interconnection between these circuits is required while preserving the modularity of the two networks and avoiding any loading problems, i.e., effects of retroactivity (Del Vecchio and Murray, 2015; Del Vecchio et al., 2008, 2016). One way to accomplish this is through the reaction:

$$Y_2 \xrightarrow{k_{in}} Y_2 + X$$

(Equation 19)

where Y_2 plays the role of the input species U without being consumed. Alternatively, in case the nature of Y_2 prevents it from directly producing X, we can use a separate sensory species S which is capable of participating in the formation of X. In particular, we assume that S is co-expressed with and decays at the same rate as Y_2, i.e.:

$$Y_1 \xrightarrow{\nu_3} Y_1 + Y_2 + S, \ S \xrightarrow{k_{in}} S + X, \ S \xrightarrow{\nu_7} \varnothing$$

(Equation 20)
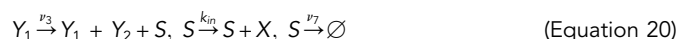
Adopting the second interconnection as the most general one, we demonstrate in Figure 3 that the rate of change of the concentration of Y_2 can be accurately represented by the output of the BioSD networks. We also
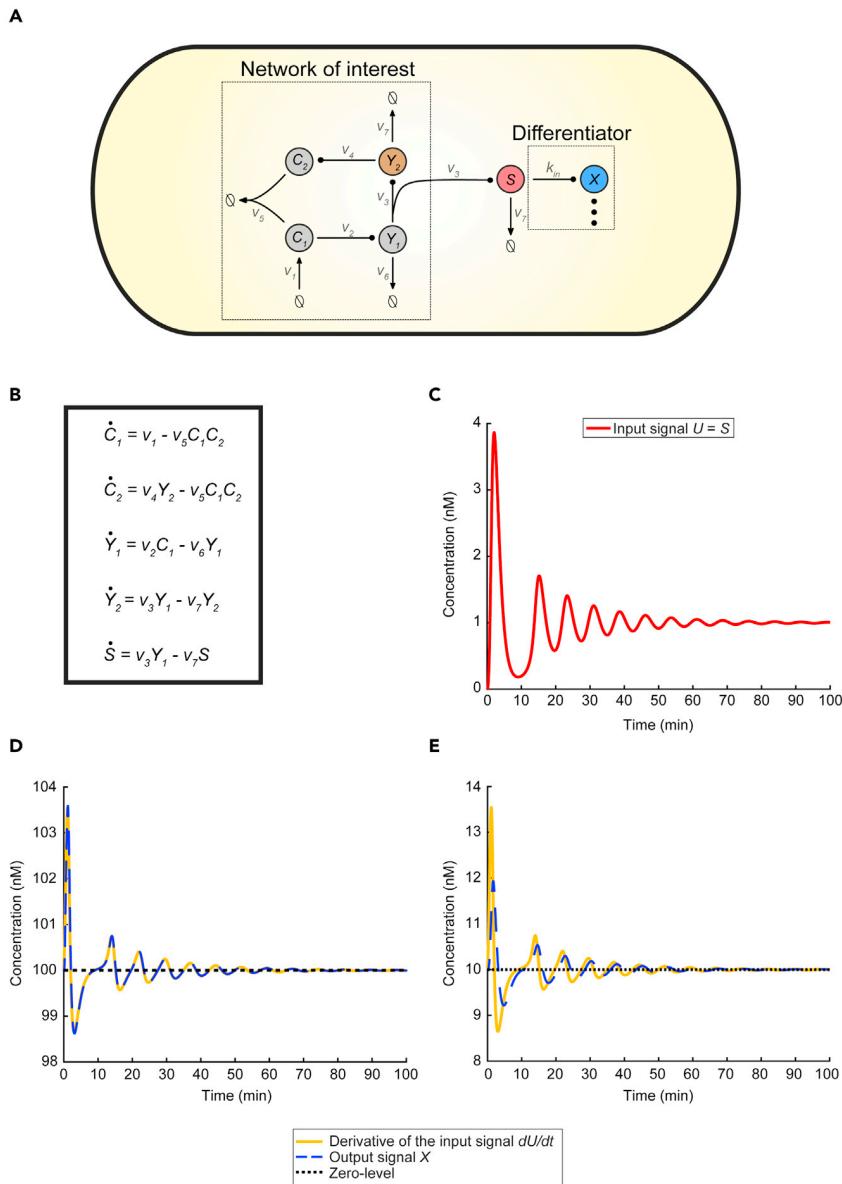
Figure 3. Sensing the rate-of-change of the output of a synthetic regulatory biomolecular network through a Biomolecular Signal Differentiator
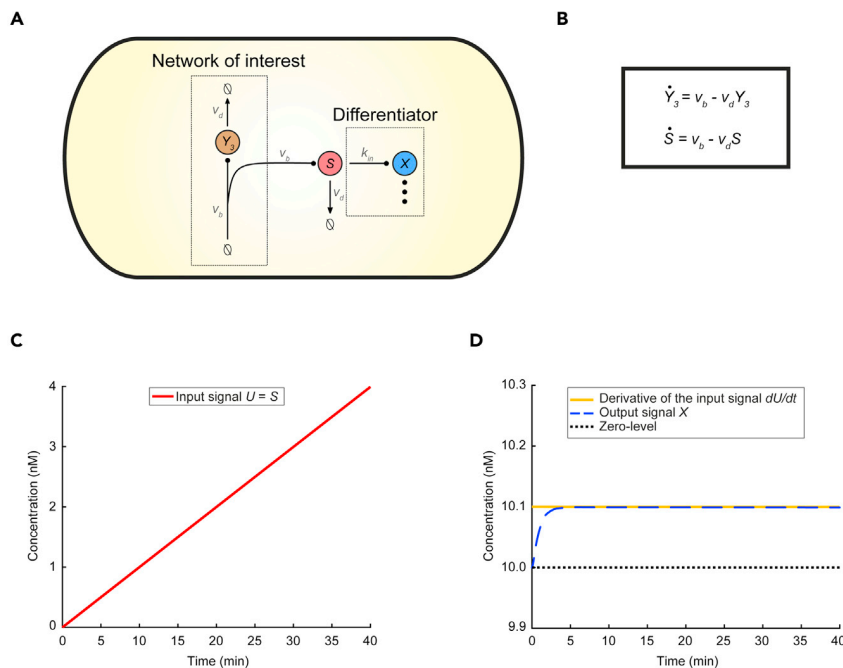
(A) Schematic of CRN (18) (network of interest) accompanied by a BioSD device (differentiator) which measures the speed of the output, Y_2 of the network via the sensing mechanism in Equation (20). We adopt the same arrow notation as in Figure 1 while the symbol (:) represents any of the three BioSD devices.

(B) ODE model capturing the dynamics of the topology given by Equations (18) and (20). As anticipated, the behavior of species Y_2 and S is described by the same equation.

(C) Input U of the differentiator coincides with species S and results from the simulation of the ODE model depicted in (B) with the following parameters: $\nu_1 = 2$ nM min^{-1}, $\nu_2 = \nu_4 = 2$ min^{-1}, $\nu_3 = 4$ min^{-1}, $\nu_5 = 12$ nM^{-1} min^{-1}, $\nu_6 = \nu_7 = 1$ min^{-1}.

(D) Simulation of BioSD-I (Equations (4a) and (4b)) response to the input shown in (C) using the following parameters: $k_{in} = 100$ min^{-1}, $k_3 = b = 100$ nM min^{-1}, $k_1 = 1$ nM^{-1} min^{-1}, $k_2 = 1$ min^{-1}, $\delta = 0.5$ min^{-1}. Equation (9) therefore yields $\varepsilon = 0.01$. As can be seen, the output, X, of the differentiator is an accurate replica of the derivative of input U.

(E) The simulation in (D) is repeated after replacing the value of both k_{in} and k_3 with 10. Equation (9) therefore yields $\varepsilon = 10$. Although the output, X, of the differentiator remains close to the derivative of input U, there is some loss of accuracy compared to (D). The respective simulations regarding BioSD-II and BioSD-III are presented in Figure S1. As expected, their responses are identical to those of BioSD-I.

Figure 4. Sensing the rate-of-change of a production - removal biomolecular process through a Biomolecular Signal Differentiator

(A) Schematic of CRN (21) (network of interest) accompanied by a BioSD device (differentiator), which measures the speed of the output of the network (Y_3) via the sensing mechanism in Equation (20). We adopt the same arrow notation as in Figure 1 while the symbol (:) represents any of the three BioSD devices.

(B) ODE model capturing the dynamics of the topology given by Equations (20) and (21). As anticipated, the behavior of species Y_3 and S is described by the same equation.

(C) Input U of the differentiator coincides with species S and results from the simulation of the ODE model depicted in B with the following parameter values: $v_b = 0.1$ nM min^{-1}, $v_d = 0.001$ min^{-1}.

(D) Simulation of the BioSD-I (Equation (4a),(4b)) response to the input presented in (C) using the following parameters: $k_{in} = 10$ min^{-1}, $k_3 = 10$ nM min^{-1}, $b = 100$ nM min^{-1}, $k_1 = 1$ nM^{-1} min^{-1}, $k_2 = 1$ min^{-1}, $\delta = 0.5$ min^{-1} (same as in Figure 3E, $\varepsilon = 10$). The output, X, of the differentiator is now an accurate replica of the derivative of input U. The latter (shown in C) belongs to the class of signals defined by Equations (7) and (17). The respective simulations regarding BioSD-II and BioSD-III are presented in Figure S2. As expected, their responses are identical to those of BioSD-I.

demonstrate that, for a given input signal, there exist sufficiently large values of ε for which the BioSD performance may not be satisfactory due to some loss of accuracy (discussed in Achieving biological signal differentiation).

We now replace the circuit described by (18) with the general production-removal process:

$$\varnothing \xrightarrow{v_b} Y_3, \ Y_3 \xrightarrow{v_d} \varnothing \qquad \text{(Equation 21)}$$

maintaining the same type of interconnection, as illustrated in Figure 4. Although the response of this process eventually converges to an equilibrium, for some period of time it practically increases linearly with time. Here, we focus on this linear regime of the response which is clearly aligned with Equation (17). Thus, as can be seen from Figure 4, BioSD networks are now able to provide accurate signal differentiation regardless of the high value of ε which, in the case of Figure 3, lead to a noticeable loss of accuracy.

Response to input signals corrupted by high-frequency noise

Potentially the most important problem of differentiator devices is their sensitivity to high-frequency noise components which the applied input signal may contain (Åström and Murray, 2021). To this end, we consider an input signal with a time-varying component

$$U^{TV} = \underbrace{A_u \sin(\omega_u t + \varphi_u)}_{\text{useful information}} + \underbrace{A_d \sin(\omega_d t + \varphi_d)}_{\text{noise}} \qquad \text{(Equation 22)}$$
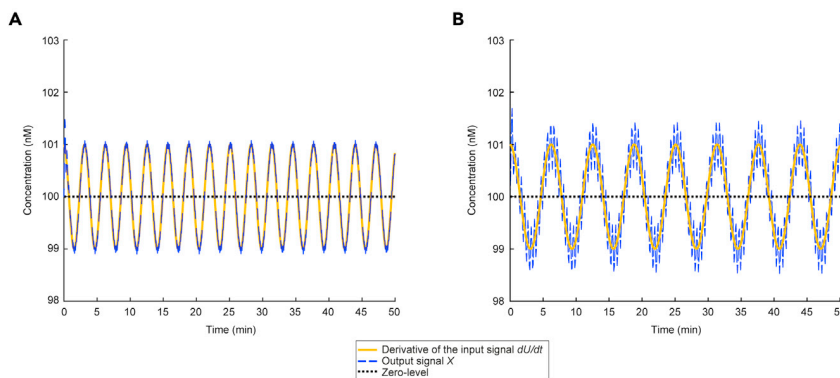
Figure 5. Response of Biomolecular Signal Differentiators to input signals with undesired high frequency components

(A) Without loss of generality we select BioSD-I (Equations (4a) and (4b)) to plot: A simulated response to an input of the form given by Equations (7) and (22) using the following parameters: $U^* = 1.2$ nM, $A_u = 1$ nM $\omega_u = 1$ rad min^{-1}, $A_d = 0.2$ nM, $\omega_d = 400$ rad min^{-1}, $\varphi_u = \varphi_d = 0$ rad, $k_{in} = 100$ min^{-1}, $k_3 = b = 100$ nM min^{-1}, $k_1 = 1$ nM^{-1} min^{-1}, $k_2 = 1$ min^{-1}, $\delta = 0.5$ min^{-1}. Equation (9) therefore yields $\varepsilon = 0.0484$. Consequently, with respect to the input signal, the frequency of the undesired component (noise) is 400 times higher than that of the component of interest (useful information). It is evident that significant noise attenuation takes place and the accuracy of signal differentiation therefore remains very high.

(B) The simulation in (A) is repeated after changing the value of ω_d to 50 which makes the noise 50 times faster compared to the useful information. As can be seen, there is a decrease in the accuracy level of signal differentiation since the input noise of this frequency cannot be filtered adequately. For demonstration purposes, in both (A) and (B) we have chosen a baseline (around of which derivative action is carried out) much larger than the amplitudes of the (ideal) derivatives regarding all the input stimuli. The useful information is represented by a signal component whose (ideal) derivative has an amplitude much smaller than the one of the (ideal) derivative of the noise. Consequently, the former can be drowned out by the latter if no noise attenuation is performed.

where the actual signal we want to differentiate–useful information–is accompanied by undesired fluctuations (noise) arising, for instance, from unintended cross-talk interactions (Del Vecchio and Murray, 2015). Note that although we model both the useful information and the noise as sinusoids, this is without loss of generality as they can be thought of as Fourier components of more general signals (see STAR Methods Signals under consideration). Assuming perfect differentiation, we get:

$$\dot{U}^{TV} = \underbrace{\omega_u A_u \sin\left(\omega_u t + \varphi_u + \frac{\pi}{2}\right)}_{\text{derivative of useful information}} + \underbrace{\omega_d A_d \sin\left(\omega_d t + \varphi_d + \frac{\pi}{2}\right)}_{\text{derivative of noise}}$$

(Equation 23)

Hence, even if the level of input corruption is low (e.g., A_d is much smaller than A_u - Equation (22)), the damage in the output of a perfect differentiator may be detrimental in case of a rapidly fluctuating noise signal (ω_d high). That is, $\omega_d A_d$ can be made arbitrarily large compared to $\omega_u A_u$ (Equation (23)) and, therefore, it is possible for the derivative of the useful signal to be completely drowned out by the derivative of some high frequency input noise. It is also apparent that the behavior of such an ideal differentiator module in the cellular environment is undesirable since it can lead to generation of greatly amplified output signals, which can be catastrophic.

Interestingly, the BioSD topologies allow us to deal with this noise amplification by suitably adjusting ε. As already discussed, BioSDs possess a low-pass filtering property defined by ε (see Equation (10)). Although this may be viewed as an "imperfection" in terms of their signal differentiation ability, it turns out to be a saving feature of great significance. Recalling the performance metric given by Equation (10) which coincides with the frequency response of the embedded filter and the Bode plot of Figure 2, we can see that there is a range of high frequencies over which signal attenuation can be effectively performed (see also STAR Methods Behavior analysis of biomolecular signal differentiators and Figure S3). This implies that Equation (15) approaches zero. Moreover, as ε increases, this range expands toward lower frequencies. Nevertheless, between the aforementioned range and the low-frequency one corresponding to signal differentiation, we can detect the existence of a relatively narrow frequency band where BioSD circuits may not be able to differentiate or attenuate input signals with satisfactory accuracy. The characteristics described above are demonstrated in Figure 5.

49
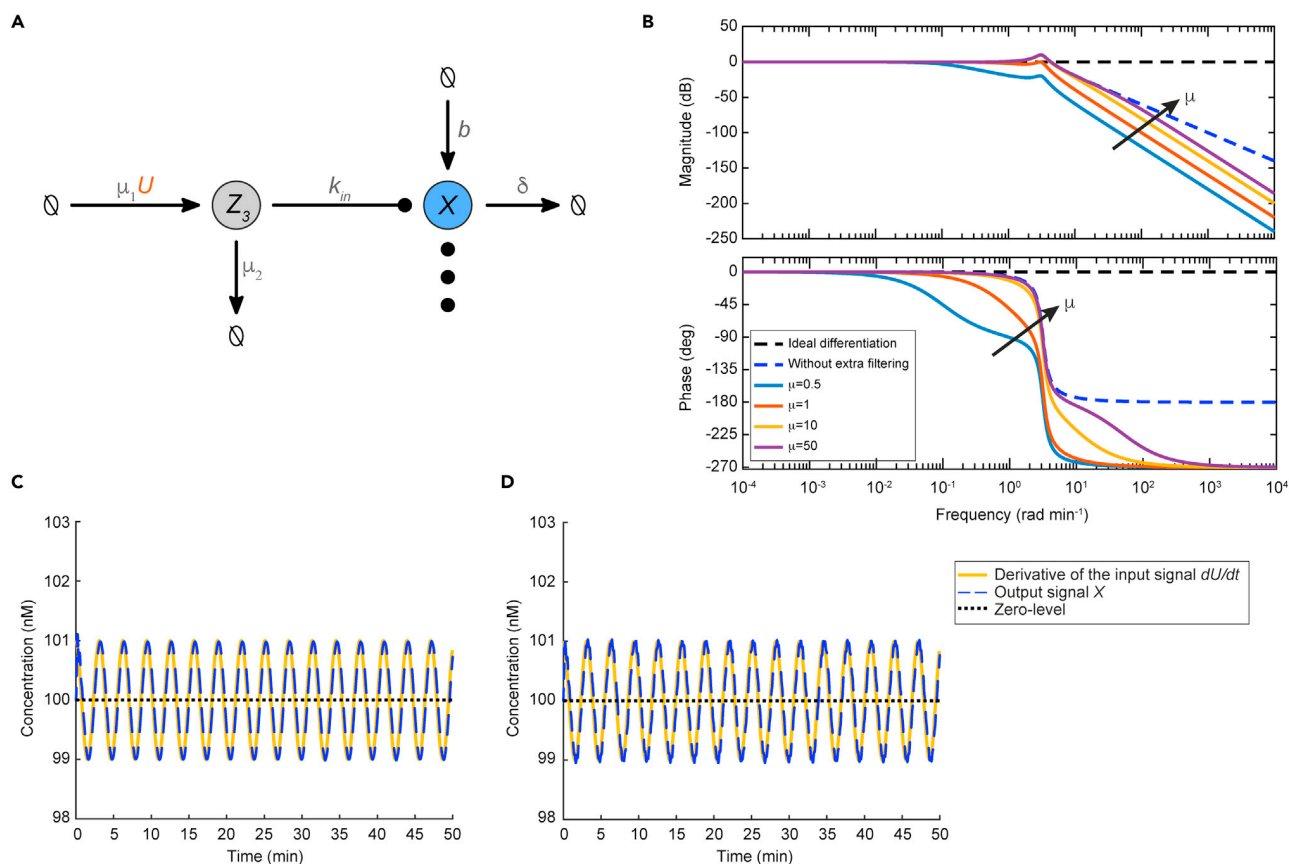
Figure 6. An alternative version of Biomolecular Signal Differentiators with an enhanced capability of input noise filtering

(A) Schematic structure of BioSD^F. We adopt the same arrow notation as in Figure 1 while the symbol (∶) represents the remaining reactions composing any of the three BioSD devices. (B) Bode plot of the performance metric given by Equation (25) with $\varepsilon = 0.1$. We consider different values of μ, where $\mu = \mu_1 = \mu_2$, that correspond to solid lines of different colors while the increasing direction of μ indicated by an arrow. We also depict the bode plot (magnitude and phase) of Equation (14) for the same value of ε and the case of ideal differentiation which are represented by blue and black dashed lines, respectively. In addition, for comparison purposes, we focus on a BioSD^F device based on BioSD-I to re-plot the simulation of **c** Figures 5A and **d** Figure 5B for the same values of the mutual parameters and $\mu_1 = \mu_2 = 5 \text{ min}^{-1}$. It is apparent that in both (C) and (D) very strong input noise attenuation takes place and the differentiation of the useful signal is thus conducted with significantly high accuracy.

A structural addition for enhanced performance

In case there are increased requirements for noise reduction that cannot be easily met via parameter tuning, we present an alternative version of the BioSD networks with higher noise insensitivity, which we call BioSD^F (Figure 6A). These topologies are described by the same CRNs presented in the section Biological structure, but amended appropriately.

More analytically, recalling the CRNs given by Equations (1), (2), and (3), we see that input signals are applied to BioSD modules through the reaction:

$$\varnothing \xrightarrow{k_{in}U} X$$

In BioSD^F topologies, the above is replaced by the following set of reactions:

$$\varnothing \xrightarrow{\mu_1 U} Z_3, \ Z_3 \xrightarrow{k_{in}} Z_3 + X, \ Z_3 \xrightarrow{\mu_2} \varnothing$$
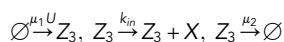
The additional species Z_3 is produced by the input species and degrades in the traditional manner while it catalyzes the formation of the output species. This structural addition is inspired by the work in (Samoilov et al., 2002; Laurenti et al., 2018), where biomolecular concepts from the area of signal processing were studied. In the following, we briefly present the main features of BioSD^F modules – a comprehensive

analysis of their behavior can be found in STAR Methods An alternative version of biomolecular signal differentiators (Figures S4 and S5).

The input/output relation of BioSD^F networks in the Laplace domain can be described by the transfer function:

$$\tilde{\Delta}_{BSD^F}(s) = \frac{\mu_1}{s + \mu_2} \cdot \frac{s}{\varepsilon(s^2 + s) + 1}$$

(Equation 24)

Similarly to BioSDs, we introduce the (normalized) performance metric:

$$\tilde{\Lambda}_F(j\omega) = \frac{\mu_2}{\mu_1} \frac{\tilde{X}_{nF}(j\omega)}{\tilde{X}_{nd}(j\omega)} = \frac{\mu_2}{j\omega + \mu_2} \frac{1}{\varepsilon(-\omega^2 + j\omega) + 1}$$

(Equation 25)

where $\tilde{X}_{nF}(j\omega)$ refers to the output of a BioSD^F network.

Using the magnitude-phase representation of Equation (25) we get:

$$|\tilde{\Lambda}_F(j\omega)| = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\mu_2}\right)^2}} \frac{1}{\sqrt{\varepsilon^2\omega^2 + (1 - \varepsilon\omega^2)^2}}$$

(Equation 26)

and

$$\angle\,\tilde{\Lambda}_F(j\omega) = \arctan\left(\frac{-\varepsilon\omega}{1 - \varepsilon\omega^2}\right) + \arctan\left(\frac{-\omega}{\mu_2}\right)$$

When $\tilde{\Lambda}_F(j\omega)$ is close to $1\angle 0°$ signal differentiation of high accuracy is achieved (Figure 6B) and the BioSD^F output can be approximated by:

$$X = \frac{\mu_1 k_{in}}{\mu_2 k_1 k_3}\dot{U} + \frac{k_3}{k_2}$$

(Equation 27)

Compared to the original BioSD topologies (Equation (16)), we now have two additional tuning parameters (μ_1, μ_2) with respect to the output differentiation gain when it comes to the low-frequency regime. However, the major advantage of this version of differentiators is an enhanced capability of noise filtering. In fact, we can have a greatly extended frequency range across which very strong attenuation of high frequency input noise can be achieved (Figures 6C and 6D). In that case, Equation (26) approaches zero. At the same time, the width of this frequency band depends on μ_2 and can be adjusted appropriately. As Equation (24) immediately reveals, the latter advantage stems from the fact that compared to BioSD circuits, BioSDs^F are equipped with an additional low-pass filter.
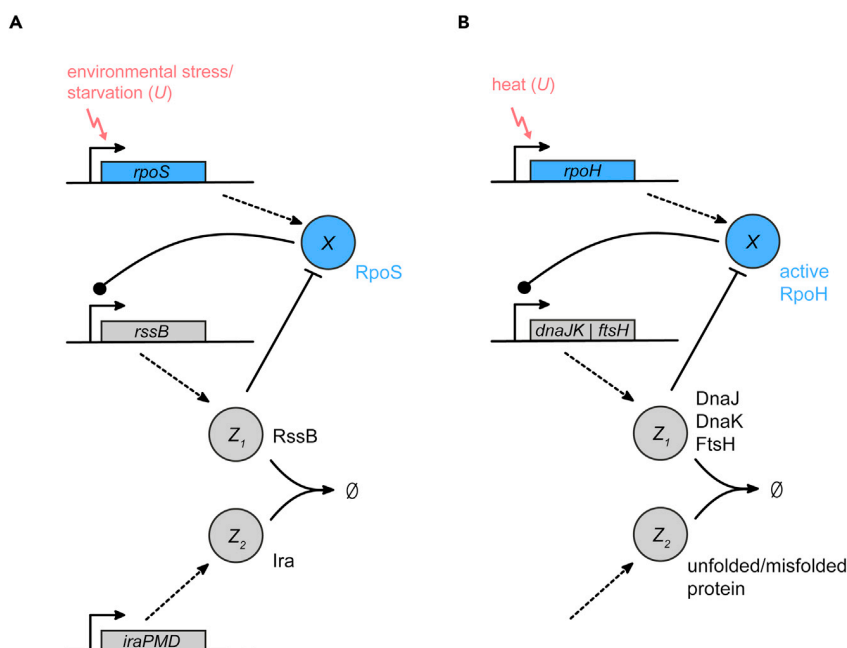
Biomolecular signal differentiators in natural regulatory networks

As outlined in the introduction, derivative action appears to be an important mechanism in various biological systems. To explore the biological relevance of the proposed BioSDs for cellular adaptations to environmental changes, we identified two naturally occurring and well-investigated regulatory network motifs that resemble the BioSD-II network. Note that these natural topologies are operating in the larger context of complex regulatory networks involving a plethora of signaling factors, some of which remain to be identified. We therefore describe the relevant motifs but do not comprehensively detail all interactions occurring in the biological system.

Stationary phase and starvation response - RpoS regulatory network

As shown in Figure 7A, we found the BioSD-II motif in the context of adaptation to nutrient starvation and entry into stationary phase, which is mediated by the sigma factor RpoS in *E. coli* and related bacteria (reviewed in (Battesti et al., 2011; Hengge-Aronis, 2002)). Stress conditions, such as nutrient depletion or high pH, serve as the input U. While RpoS is present at low levels (b) in exponentially growing cells, its expression is substantially increased through both transcriptional and post-transcriptional regulation in response to environmental stresses or starvation (Battesti et al., 2011). One of the genes whose expression is dependent on RpoS is *rssB*, which encodes a response regulator. RssB binds to RpoS and mediates its degradation by the ClpXP protease (Pruteanu and Hengge-Aronis, 2002), thus functioning as Z_1. Nutrient starvation also induces the expression of several anti-adaptor proteins (Ira; inhibitor of RssB activity). These proteins

Figure 7. Examples of the Biomolecular Signal Differentiator-II motif in natural systems
Simplified schematics of BioSD-II topologies occurring as part of (A) the RpoS-mediated stress response and (B) the RpoH-mediated heat shock response in *Escherichia coli*. Corresponding components of BioSD-II are indicated.

bind to RssB and prevent RpoS degradation (Battesti et al., 2013), which corresponds to the action of Z_2 in BioSD-II.

Heat shock response - RpoH regulatory network

A second example for the BioSD-II motif was identified in the regulatory network of the sigma factor RpoH, which coordinates the heat shock response in *E. coli* (Figure 7B) (Straus et al., 1987; Roncarati and Scarlato, 2017). Upon heat shock, cellular RpoH levels rise above their low baseline concentration (b), inducing the expression of several chaperones (e.g. DnaKJ and GroELS) and proteases (e.g. FtsH and Lon). DnaK and DnaJ can bind to RpoH and facilitate its degradation by FtsH (Straus, Walter, and Gross, 1989a; Gamer et al., 1992), thereby acting as Z_1. Unfolded or misfolded proteins will sequester chaperones and proteases (Gamer et al., 1992), thus increasing the stability of RpoH and fulfilling the function of Z_2. In this network, the amount of *active* RpoH (as opposed to the total amount of RpoH) should be considered as X, since it has been found that the activity rather than the concentration of RpoH inside the cell drops during temperature downshifts (Straus, Walter, and Gross, 1989b).

Guidelines for experimental implementation of biomolecular signal differentiators

In addition to the natural regulatory networks described in the preceding section, here we outline possible synthetic implementations for all BioSD circuits inside a living cell and, in particular, in *E. coli* (Figure 8). Inducible expression of species X can be achieved from any well-characterized promoter, such as the IPTG-inducible P_{lac}. Leakiness of the lac promoter will ensure nonzero expression levels (b) even in the absence of inducer. Alternatively, if higher baseline expression levels are required, X could additionally be expressed from a weak constitutive promoter. To minimize undesirable interference with other cellular processes, X should be an orthogonal sigma factor, such as σ^F from *Bacillus subtilis* (Bervoets et al., 2018). A translational fusion of X to GFP will allow for easy tracking of the system output. σ^F will then induce expression of a Lon protease (Z in BioSD-I, Z_1 in BioSD-II and III) from its cognate promoter P_{F1}. In this case, a Lon$^-$ strain of *E. coli* would be used to avoid interference of naturally present Lon protease. Addition of a degradation tag to σ^F will target it for degradation by the Lon protease. To approximate 0th-order degradation of Z in BioSD-I, an *ssrA* tag will be fused to the Lon protease as described in (Wong et al., 2007; Ang et al., 2010).
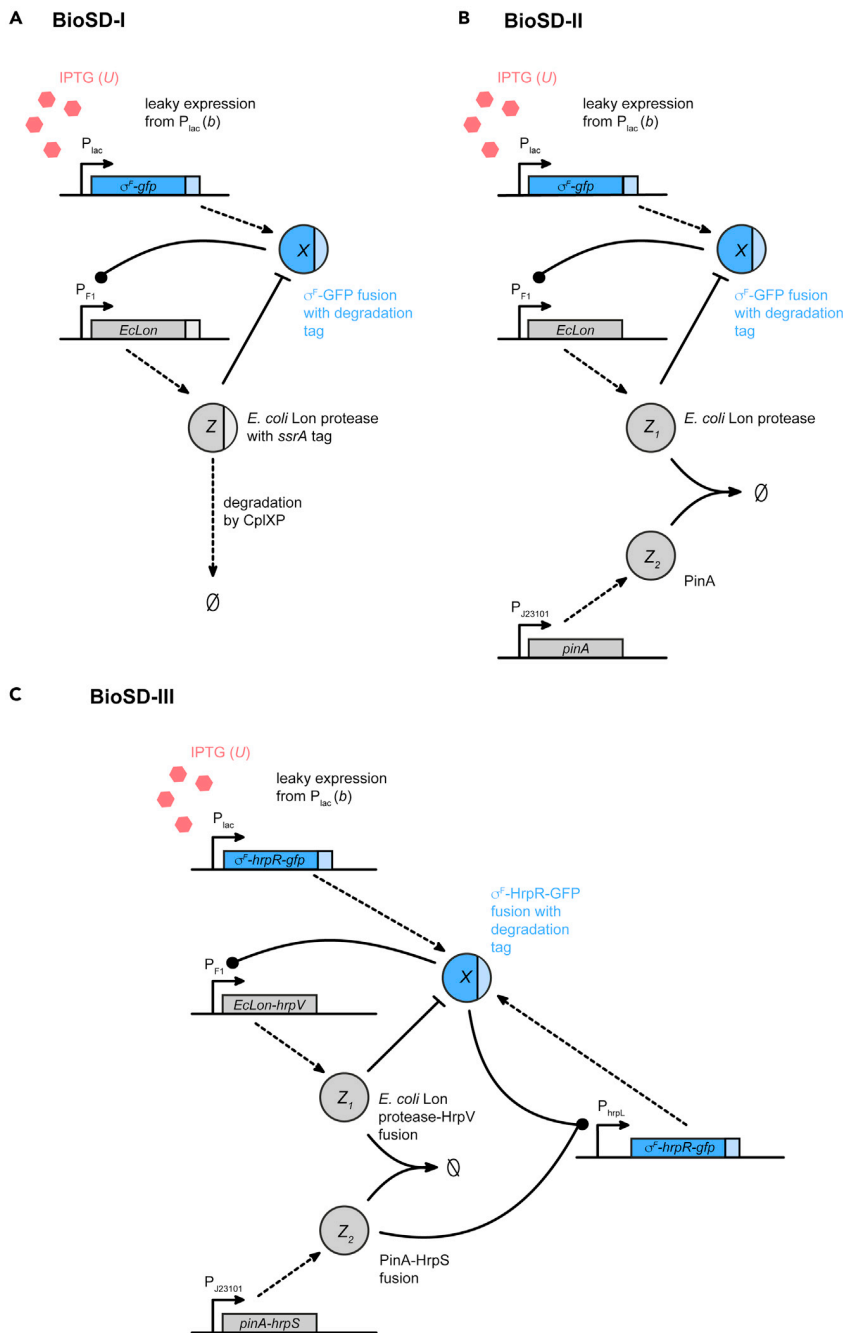
Figure 8. Possible experimental implementations of Biomolecular Signal Differentiators

Schematic representation of synthetic designs for ABioSD-I, BBioSD-II and CBioSD-III.

For BioSD-II, we additionally introduce constitutive expression of the protease inhibitor PinA from phage T4 (Z_2), which has been shown to specifically inhibit the Lon protease in *E. coli* with high affinity (Hilliard et al., 1998). A synthetic promoter from the BioBrick collection (Kelly et al., 2009) may be used to achieve the desired expression level of Z_2. Ideally, an orthogonal Lon protease should be used (e.g. Lon protease from *Mesoplasma florum* (Aoki et al., 2019)) to prevent cross-talk with other cellular proteins. However, since the interaction of PinA with proteases has been characterized only in *E. coli* so far, we have suggested use of the *E. coli* Lon protease.

Due to the number of required interactions in BioSD-III, it will likely be necessary to introduce auxiliary species for X, Z_1 and Z_2, which we refer to as X_{aux}, $Z_{1,aux}$ and $Z_{2,aux}$, respectively. These auxiliary species would ideally have identical behavior to the main species X, Z_1 and Z_2, even though simulations indicate that completely identical behavior is not required (see STAR Methods Analysis of the experimental topology of Biomolecular Signal Differentiator-III and Figure S6). One option is to augment the design for BioSD-II with the Hrp system from *Pseudomonas syringae*, which has previously been implemented in synthetic biology studies (Wang et al., 2014). HrpR (X_{aux}) is expressed from P_{lac} together with σ^F, and HrpS ($Z_{2,aux}$) is expressed as a protein fusion with PinA. HrpR and HrpS are both required to induce additional production of σ^F and HrpR from P_{hrpL}. At the same time, HrpV ($Z_{1,aux}$) binds HrpS rendering it inactive. The structural addition required for BioSD^F can be implemented by, for example, expressing X from a T7 promoter and expressing T7 RNA polymerase (Z_3) from a separate inducible promoter.

DISCUSSION

In this study, we propose three biomolecular topologies that are able to act as highly accurate signal differentiators inside the cell. These designs provide guidance for building cellular devices capable of computing time derivatives of molecular signals. At the same time, they reveal concepts that are found in natural biological networks implementing differentiation and derivative feedback.

More specifically, we introduce three general biomolecular architectures BioSD-I, II, and III. Their generality lies in the fact that they are represented by CRNs without being restricted by the biological identity of reactants and products and, by extension, the corresponding biological pathway. Important structural components of the BioSDs are a negative feedback loop created by a special process of excitation and inhibition between two species (Iglesias and Shi, 2014), an enzymatic degradation of zero-order kinetics (BioSD-I), an autocatalytic-like reaction (BioSD-III) and an antithetic-like motif based on annihilation (Oishi and Klavins, 2011; Briat et al., 2016) (BioSD-II, BioSD-III). We theoretically analyze their features and show the conditions under which high performance can be guaranteed. Among others, important concepts such as stability, tunability, and accuracy are discussed in detail.

Special emphasis is placed on the expected sensitivity of differentiators to input signals corrupted by high-frequency noise. We demonstrate that this issue can be resolved to a certain extent through suitable parameter tuning. Nevertheless, for cases in which stronger noise attenuation is needed, we present a structural modification that gives rise to three slightly different architectures, namely BioSD^F-I, II and III, with enhanced capabilities. However, the price for this improvement is the addition of an extra biomolecular species, which implies an increase in structural complexity. Moreover, we introduce performance metrics both for BioSDs and BioSDs^F based on which the circuit designer can assess the quality of signal differentiation and attenuation. These metrics take into account both the frequency content of the input signal and the reaction rates involved in the circuits, thus facilitating tuning according to the expected performance standards.

The ability to perform time differentiation is of central importance in various biological systems, contributing to stability and fast adaptation to changing conditions (Barkai and Leibler, 1997; Bazellières et al., 2015; Cloutier and Wellstead, 2010). Owing to the generality of the presented topologies, we anticipate that the present study will facilitate the investigation of naturally occurring systems capable of derivative action. In this study, we discuss the regulatory networks of two bacterial sigma factors, RpoS and RpoH, which play a central role in the response and adaptation to stress conditions and heat shock, respectively. Interestingly, these networks share structural characteristics with one of the proposed topologies, BioSD-II.

In addition, the motifs presented here provide building blocks that can be both implemented in stand-alone applications, such as speed biosensors, and also combined with existing biochemical control structures in a modular fashion, e.g., for building biomolecular PID controllers (Chevalier et al., 2019). We describe potential designs for synthetic experimental implementation of all three BioSDs, which can be readily adapted depending on the nature of the system and available biological parts. To realize the antithetic motif in BioSD-II and III, we propose the use of a protease/protease inhibitor pair as an alternative to the previously described systems using sigma and anti-sigma factors (Aoki et al., 2019) or sRNAs and mRNAs (Huang et al., 2018; Kelly et al., 2018). To allow for greater flexibility in choosing the biomolecular species, we introduce a concept of auxiliary species whose usefulness is demonstrated through BioSD-III. Furthermore, to enhance the biological significance of our work in STAR Methods Modeling a more realistic

case of Biomolecular Signal Differentiator-II (Figures S7 and S8, Table S1), we investigate the behavior of one of the differentiator modules, namely BioSD-II, under more realistic conditions stemming from our experimental designs.

Stochasticity is an essential characteristic of biomolecular systems which operate in a noisy environment (Del Vecchio and Murray, 2015; Laurenti et al., 2018; Raj and Van Oudenaarden, 2008; Eldar and Elowitz, 2010; Cardelli et al., 2016; Warne et al., 2019). The biomolecular motifs introduced in the current study were analyzed through ODE models (deterministic analysis) which generally approximate well the dynamics of CRNs whose species are present in high copy-numbers. It therefore remains an interesting endeavor to identify the probabilistic effects of the molecular reactions involved that may have a significant impact on the behavior of these motifs when the biomolecular counts are low.

The speed or higher derivatives of the output of a system offers important information about its properties. For an electromechanical system this is not difficult, but it has been a challenging question for biological systems. In this article, we provide an approach to gain access to this information, which will be invaluable for assessing and improving the performance of biological systems. We believe that our BioSD topologies will expand the tools available for understanding and engineering biological systems for robustness and reliability.

Limitations of the study

As emphasized in the Discussion, the behavior of the topologies presented here is studied via deterministic mathematical analysis and simulations; the effect of inherent stochasticity of living systems stemming from the random nature of molecular reactions on these topologies is left for future work.

SUPPORTING CITATIONS

The following references appear in the supplemental information: Buchler and Louis, 2008; Fekkes et al., 1995; Gur et al., 2012; Schlosshauer and Baker, 2004.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Signals under consideration
  - Equilibria and stability of biomolecular signal differentiators
  - The notion of strong rate of annihilation between Z_1, Z_2 (large η) in Biomolecular Signal Differentiator-II
  - Behavior analysis of biomolecular signal differentiators
  - An alternative version of biomolecular signal differentiators
  - Analysis of the experimental topology of Biomolecular Signal Differentiator-III
  - Modeling a more realistic case of Biomolecular Signal Differentiator-II

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.103462.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

Conceptualization and methodology, E.A., C.C.M.S., A.P., L.C.; Formal analysis and Software: E.A., Writing, E.A., C.C.M.S., A.P., L.C.; Supervision: A.P., L.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Alon, U. (2019). An Introduction to Systems Biology: Design Principles of Biological Circuits (CRC press).

Ang, J., Bagh, S., Ingalls, B.P., and McMillen, D.R. (2010). Considerations for using integral feedback control to construct a perfectly adapting synthetic gene network. J. Theor. Biol. 266, 723–738.

Aoki, S.K., Lillacci, G., Gupta, A., Baumschlager, A., Schweingruber, D., and Khammash, M. (2019). A universal biomolecular integral feedback controller for robust perfect adaptation. Nature 570, 533–537.

Åström, K.J., and Murray, R.M. (2021). Feedback Systems: An Introduction for Scientists and Engineers (Princeton University Press).

Baetica, A.A., Leong, Y.P., and Murray, R.M. (2020). Guidelines for designing the antithetic feedback motif. Phys. Biol. 17, 055002.

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

Battesti, A., Hoskins, J.R., Tong, S., Milanesio, P., Mann, J.M., Kravats, A., Tsegaye, Y.M., Alexandre, B., Wickner, S., and Gottesman, S. (2013). Anti-adaptors provide multiple modes for regulation of the RssB adaptor protein. Genes Develop. 27, 2722–2735.

Battesti, Aurelia, NadimMajdalani, and Gottesman, S. (2011). The RpoS-mediated general stress response in Escherichia coli. Annu. Rev. Microbiol. 65, 189–213.

Bazellie'res, E., Conte, V., Elosegui-Artola, A., Serra-Picamal, X., Bintanel-Morcillo, M., Roca-Cusachs, P., Muñoz, J.J., Sales-Pardo, M.a., Guimerà, R., and Trepat, X. (2015). Control of cell–cell forces and collective cell dynamics by the intercellular adhesome. Nat. Cell Biol. 17, 409–420.

Bervoets, I., Van Brempt, M., Van Nerom, K., Bob Van Hove, Maertens, J., De Mey, M., and Charlier, D.A. (2018). A sigma factor toolbox for orthogonal gene expression in Escherichia coli. Nucleic Acids Res. 46, 2133–2144.

Block, S.M., Segall, J.E., and Berg, H.C. (1983). Adaptation kinetics in bacterial chemotaxis. J. Bacteriol. 154, 312–323.

Briat, C., Gupta, A., and Khammash, M. (2016). Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. Cell Syst. 2, 15–26.

Briat, C., Gupta, A., and Khammash, M. (2018). Antithetic proportional-integral feedback for reduced variance and improved control performance of stochastic reaction networks. J. R. Soc. Interf. 15, 20180079.

Buchler, N.E., and Louis, M. (2008). Molecular titration and ultrasensitivity in regulatory networks. J. Mol. Biol. 384, 1106–1119.

Cardelli, L., Kwiatkowska, M., and Laurenti, L. (2016). Stochastic analysis of chemical reaction networks using linear noise approximation. Biosystems 149, 26–33.

Chen, S., and Wilson, D.B. (1997). Construction and characterization of Escherichia coli genetically engineered for bioremediation of Hg^{2+}-contaminated environments. Appl. Environ. Microbiol. 63, 2442–2445.

Chevalier, M., Gómez-Schiavon, M., Ng, A.H., and El-Samad, H. (2019). Design and analysis of a proportional-integral-derivative controller with biological molecules. Cell Syst. 9, 338–353.

Cloutier, M., and Wellstead, P. (2010). The control systems structures of energy metabolism. J. R. Soc. Interf. 7, 651–665.

Del Vecchio, D., Dy, A.J., and Qian, Y. (2016). Control theory meets synthetic biology. J. R. Soc. Interf. 13, 20160380.

Del Vecchio, D., and Murray, R.M. (2015). Biomolecular Feedback Systems (Princeton University Press).

Del Vecchio, D., Ninfa, A.J., and Eduardo, D.S. (2008). Modular cell biology: retroactivity and insulation. Mol. Syst. Biol. 4, 161.

Eldar, A., and Elowitz, M.B. (2010). Functional roles for noise in genetic circuits. Nature 467, 167–173.

Fekkes, P.r., den Blaauwen, T., and Driessen, A.J.M. (1995). Diffusion-limited interaction between unfolded polypeptides and the Escherichia coli chaperone SecB. Biochemistry 34, 10078–10085.

Filo, M., and Khammash, M. (2021). A class of simple biomolecular antithetic

proportionalintegral-derivative controllers. bioRxiv. https://doi.org/10.1101/2021.03.21.436342.

Gamer, J., Hermann, B., and Bukau, B. (1992). Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor σ^{32}. Cell 69, 833–842.

Gur, E., Vishkautzan, M., and Sauer, R.T. (2012). Protein unfolding and degradation by the AAA+ Lon protease. Protein Sci. 21, 268–278.

Halter, W., Murray, R.M., and Frank, A. (2019). Analysis of primitive genetic interactions for the design of a genetic signal differentiator. Synth.Biol. 4, ysz015.

Halter, W., Tuza, Z.A., and Frank, A. (2017). Signal differentiation with genetic networks. IFAC-PapersOnLine 50, 10938–10943.

Hancock, E.J., Jordan, A., Papachristodoulou, A., and Stan, G.B. (2017). The interplay between feedback and buffering in cellular homeostasis. Cell Syst. 5, 498–508.

Hengge-Aronis, R. (2002). Signal transduction and regulatory mechanisms involved in control of the σS (RpoS) subunit of RNA polymerase. Microbiol. Mol. Biol. Rev. 66, 373–395.

Hilliard, J.J., Maurizi, M.R., and Lee, D.S. (1998). Isolation and characterization of the phage T4 PinA protein, an inhibitor of the ATP-dependent Lon protease of Escherichia coli. J. Biol. Chem. 273, 518–523.

Huang, H.H., Qian, Y., and Del Vecchio, D. (2018). A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. Nat. Commun. 9, 1–12.

Iglesias, P.A., and Devreotes, P.N. (2008). Navigating through models of chemotaxis. Curr.Opin.Cell Biol. 20, 35–40.

Iglesias, P.A., and Shi, C. (2014). Comparison of adaptation motifs: temporal, stochastic and spatial responses. IET Syst. Biol. 8, 268–281.

Kelly, C.L., Harris, A.W.K., Harrison, S., Hancock, E.J., Heap, J.T., and Papachristodoulou, A. (2018). Synthetic negative feedback circuits using engineered small RNAs. Nucleic Acids Res. 46, 9875–9889.

Kelly, J.R., Rubin, A.J., Davis, J.H., Ajo-Franklin, C.M., Cumbers, J., Czar, M.J., de Mora, K.,

56

Glieberman, A.L., Monie, D.D., and Drew, E. (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. J. Biol. Eng. *3*, 1611–1754.

Khalil, H.K. (2002). Nonlinear Systems, 3rd Ed. (Prentice-Hall, Inc).

Lathi, B.P. (1998). Signal Processing and Linear Systems (Oxford University Press).

Laurenti, L., Csikasz-Nagy, A., Kwiatkowska, M., and Cardelli, L. (2018). Molecular filters for noise reduction. Biophys. J. *114*, 3000–3011.

Lu, T.K., Khalil, A.S., and Collins, J.J. (2009). Next-generation synthetic gene networks. Nat. Biotechnol. *27*, 1139.

Macnab, R.M., and Koshland, D.E. (1972). The gradient-sensing mechanism in bacterial chemotaxis. Proc. Natl. Acad. Sci. *69*, 2509–2512.

Modi, S., Dey, S., and Singh, A. (2019). Proportional and derivative controllers for buffering noisy gene expression. In 2019 IEEE 58th Conference on Decision and Control (CDC) (IEEE), pp. 2832–2837.

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

Olsman, N.h., Baetica, A.A., Xiao, F., Leong, Y.P., Murray, R.M., and Doyle, J.C. (2019a). Hard limits and performance tradeoffs in a class of antithetic integral feedback networks. Cell Syst. *9*, 49–63.

Olsman, N., Xiao, F., and Doyle, J.C. (2019b). Architectural principles for characterizing the performance of antithetic integral feedback networks. Iscience *14*, 277–291.

Olsman, N., and Forni, F. (2020). Antithetic integral feedback for the robust control of monostable and oscillatory biomolecular circuits. IFAC-PapersOnLine *53*, 16826–16833.

Oppenheim, A.V., Willsky, A.S., and Hamid Nawab, S. (1996). Signals Systems, 2nd Ed. (Prentice-Hall, Inc).

Paulino, N.M.G., Foo, M., Kim, J., and Bates, D.G. (2019). PID and state feedback controllers using

DNA strand displacement reactions. IEEE Control Syst. Lett. *3*, 805–810.

Pieper, D.H., and Reineke, W. (2000). Engineering bacteria for bioremediation. Curr.Opin.Biotechnol. *11*, 262–270.

Pruteanu, M., and Hengge-Aronis, R. (2002). The cellular level of the recognition factor RssB is rate-limiting for σS proteolysis: implications for RssB regulation and signal transduction in σS turnover in Escherichia coli. Mol. Microbiol. *45*, 1701–1713.

Qian, Y., and Del Vecchio, D. (2018). Realizing 'integral control' in living cells: how to overcome leaky integration due to dilution? J. R. Soc. Interf. *15*, 20170902.

Raj, A., and Van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. Cell *135*, 216–226.

Roncarati, D., and Scarlato, V. (2017). Regulation of heat-shock genes in bacteria: from signal sensing to gene expression output. FEMS Microbiol. Rev. *41*, 549–574.

Samaniego, C.C., Giordano, G., and Franco, E. (2019). Practical differentiation using ultrasensitive molecular circuits. In 2019 18th European Control Conference (ECC) (IEEE), pp. 692–697.

Samaniego, C.C., Kim, J., and Franco, E. (2020). Sequestration and delays enable the synthesis of a molecular derivative operator. In 2020 59th IEEE Conference on Decision and Control (CDC) (IEEE), pp. 5106–5112.

Samoilov, M., Adam, A., and Ross, J. (2002). Signal processing by simple chemical systems. J. Phys. Chem. A *106*, 10205–10221.

Schlosshauer, M., and Baker, D. (2004). Realistic protein–protein association rates from a simple diffusional model neglecting long-range interactions, free energy barriers, and landscape ruggedness. Protein Sci. *13*, 1660–1669.

Shimizu, T.S., Tu, Y., and Berg, H.C. (2010). A modular gradient-sensing network for chemotaxis in Escherichia coli revealed by responses to time-varying stimuli. Mol. Syst. Biol. *6*, 382.

Steel, H., Lillacci, G., Khammash, M., and Papachristodoulou, A. (2017). Challenges at the interface of control engineering and synthetic biology. In 2017 IEEE 56th Annual Conference on Decision and Control (CDC) (IEEE), pp. 1014–1023.

Straus, D.B., Walter, W.A., and Gross, C.A. (1987). The heat shock response of E. coli is regulated by changes in the concentration of σ32. Nature *329*, 348–351.

Straus, D.B., Walter, W.A., and Gross, C.A. (1989a). DnaK, DnaJ, and GrpE heat shock proteins negatively regulate heat shock gene expression by controlling the synthesis and stability of σ32. Genes Develop. *4*, 2202–2209.

Straus, D.B., AWalter, W., and Gross, C.A. (1989b). The activity of σ32 is reduced under conditions of excess heat shock protein production in Escherichia coli. Genes Development *3*, 2003–2010.

Wang, B., Barahona, M., and Buck, M. (2014). Engineering modular and tunable genetic amplifiers for scaling transcriptional signals in cascaded gene networks. Nucleic Acids Res. *42*, 9484–9492.

Warne, D.J., Baker, R.E., and Simpson, M.J. (2019). Simulation and inference algorithms for stochastic biochemical reaction networks: from basic concepts to state-of-the-art. J. R. Soc. Interf. *16*, 20180943.

Whitby, M., Cardelli, L., Kwiatkowska, M., Laurenti, L., Tribastone, M., and Tschaikowski, M. (2021). PID control of biochemical reaction networks. IEEE Trans. Automat.Control *1*, 1.

Wilson, H.R. (1999). Spikes, Decisions, and Actions: The Dynamical Foundations of Neurosciences (Oxford University Press).

Wong, W.W., Tsai, T.Y., and Liao, J.C. (2007). Single-cell zeroth-order protein degradation enhances the robustness of synthetic oscillator. Mol. Syst. Biol. *3*, 1–8.

Zill, D.G. (2012). A First Course in Differential Equations with Modeling Applications (Cengage Learning).

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
MATLAB	Mathworks	www.mathworks.com
Matlab code used for simulations	This study	https://github.com/emgalox/BioS-Differentiators

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Antonis Papachristodoulou (antonis@eng.ox.ac.uk).

Materials availability

This study did not generate new unique materials.

Data and code availability

All numerical simulations were performed in MATLAB R2020 using the ODE solver ode23s except for those in Figure 5 where the ODE solver ode113 was used. Simulation parameter values can be found in the figure captions. Initial conditions for the biomolecular species involved are considered zero except for BioSDs and BioSDs^F where the corresponding equilibria ("rest-positions") are used (see STAR Methods Equilibria and stability of Biomolecular Signal Differentiators and An alternative version of Biomolecular Signal Differentiators). The corresponding programming code is available at: https://github.com/emgalox/BioS-Differentiators.

METHOD DETAILS

Signals under consideration

In this study we consider Fourier-transformable signals (unless otherwise stated) (Lathi, 1998; Oppenheim et al., 1996). The Fourier transform exists for any signal, $s(t)$, satisfying the following conditions, also known as Dirichlet conditions:

- $s(t)$ is absolutely integrable, i.e.:

$$\int_{-\infty}^{+\infty} |s(t)| dt < \infty$$

- $s(t)$ has a finite number of maxima and minima within any finite interval.

- $s(t)$ has a finite number of discontinuities within any finite interval. In addition, each of these discontinuities must be finite.

The Dirichlet conditions are sufficient but not necessary for the existence of Fourier transform of a signal. Moreover, it should be noted that the Fourier transform of periodic signals can be computed from their Fourier series representation (assuming it exists) with the help of impulse functions.

The main idea behind Fourier analysis is the decomposition of a signal into a sum of sinusoids, the relative amplitudes and phases of which are determined by the Fourier spectrum of that signal. In the case of a linear, time invariant system, transmission of a signal can be therefore treated as transmission of its constituent sinusoids. Moreover, the frequency-domain description of such a system using its frequency response is an alternative to the time-domain description based on convolution.

Finally, in the current study we focus on physical signals that can be generated in a cellular environment. Such naturally-occurring signals typically satisfy the Dirichlet conditions and, thus, have a Fourier representation - signals that do not satisfy these conditions do not normally arise in practical applications. Further details on the above can be found in (Lathi, 1998; Oppenheim et al., 1996).

Equilibria and stability of biomolecular signal differentiators

We assume that all biomolecular circuits in this study are represented by chemical reaction networks (CRNs) whose dynamics are described by the law of mass action unless otherwise stated. For the purposes of deterministic modeling, we consider inputs $U(t)$ that are bounded, non-negative, continuous-time signals of finite duration, the time derivatives of which exist and are also bounded and continuous. This is clearly aligned with the biological nature of $U(t)$ which can correspond, for example, to the concentration of a biomolecular species.

Biomolecular signal differentiator-I. Biomolecular Signal Differentiator-I (BioSD-I) is described by the CRN:

$$\varnothing \xrightarrow{k_{in}U} X, \ \varnothing \xrightarrow{b} X, \ X \xrightarrow{k_2} X + Z$$
$$X + Z \xrightarrow{k_1} Z, \ X \xrightarrow{\delta} \varnothing, \ Z \xrightarrow{k_3/Z} \varnothing \qquad \text{(Equation S1)}$$

where $k_{in}, b, k_2, k_1, k_3, \delta \in \mathbb{R}_+$. Note that the removal rate of Z is constant and equal to k_3. To achieve this we assume that Z participates in an enzyme-catalyzed degradation process which is traditionally described by Michaelis-Menten kinetics. More precisely, the removal rate of Z is equal to

$$k_3 \frac{Z}{Z + K_m} \qquad \text{(Equation S2)}$$

where $K_m \in \mathbb{R}_+$ is the Michaelis-Menten constant. When the enzyme that catalyzes the degradation process is saturated by its substrate, we have:

$$Z \gg K_m \qquad \text{(Equation S3)}$$

which entails, in effect, zero-order kinetics since Equation (S2) becomes practically equal to k_3.

The dynamics of the above CRN (Equation (S1)) are given by the following system of Ordinary Differential Equations (ODEs):

$$\dot{X} = k_{in}U + b - k_1 XZ - \delta X \qquad \text{(Equation S4)}$$
$$\dot{Z} = k_2 X - k_3 \qquad \text{(Equation S5)}$$

For any constant input U^*, a steady state (X^*, Z^*) of the system given by Equations (S4) and (S5) exists and is finite. By setting the time derivatives of this system to zero, we can obtain the following unique steady-state:

$$X^* = \frac{k_3}{k_2} \qquad \text{(Equation S6)}$$

$$Z^* = \frac{k_2(k_{in}U^* + b)}{k_1 k_3} - \frac{\delta}{k_1} \qquad \text{(Equation S7)}$$

Clearly X^* is positive while, due to Equation (S3), the same is true for Z^* (in fact: $Z^* \gg 0$).

To study the local stability of the above equilibrium, we linearize Equations (S4) and (S5) around (X^*, Z^*) for a constant input U^* to get:

$$\begin{bmatrix} \dot{X} \\ \dot{Z} \end{bmatrix} = \underbrace{\begin{bmatrix} -\dfrac{k_2(k_{in}U^* + b)}{k_3} & -\dfrac{k_1 k_3}{k_2} \\ k_2 & 0 \end{bmatrix}}_{G_1} \begin{bmatrix} X \\ Z \end{bmatrix} \qquad \text{(Equation S8)}$$
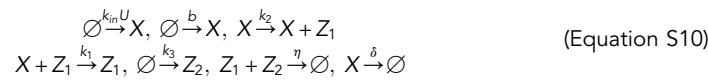
As far as the linear system described by Equation (S8) is concerned, the steady state (X^*, Z^*) is exponentially stable since matrix G_1 is Hurwitz. To prove this, we find the characteristic polynomial of G_1 as:

$$P_1(s) = \det(sI - G_1) = s^2 + \frac{k_2}{k_3}(k_{in}U^* + b)s + k_1k_3 \qquad \text{(Equation S9)}$$

According to Routh-Hurwitz criterion, the second-order polynomial given by Equation (S9) has both roots in the open left half plane if, and only if, both $\frac{k_2(k_{in}U^* + b)}{k_3}$ and k_1k_3 are positive, which is always true. Consequently, (X^*, Z^*) is a positive locally exponentially stable steady state for the nonlinear system given by Equations (S4) and (S5).

Following the same procedure, we next analyze BioSD-II and BioSD-III.

Biomolecular signal differentiator-II. The CRN that corresponds to Biomolecular Signal Differentiator-II (BioSD-II) is:

$$\emptyset \xrightarrow{k_{in}U} X, \ \emptyset \xrightarrow{b} X, \ X \xrightarrow{k_2} X + Z_1$$
$$X + Z_1 \xrightarrow{k_1} Z_1, \ \emptyset \xrightarrow{k_3} Z_2, \ Z_1 + Z_2 \xrightarrow{\eta} \emptyset, \ X \xrightarrow{\delta} \emptyset \qquad \text{(Equation S10)}$$

where $k_{in}, b, k_2, k_1, \delta, \eta \in \mathbb{R}_+$.

The dynamics of the above CRN (Equation (S10)) are described by the set of ODEs:

$$\dot{X} = k_{in}U + b - k_1XZ_1 - \delta X \qquad \text{(Equation S11)}$$
$$\dot{Z}_1 = k_2X - \eta Z_1Z_2 \qquad \text{(Equation S12)}$$
$$\dot{Z}_2 = k_3 - \eta Z_1Z_2 \qquad \text{(Equation S13)}$$

For any constant input U^*, provided that:

$$k_2(k_{in}U^* + b) > \delta k_3, \qquad \text{(Equation S14)}$$

we have a unique positive (finite) steady state:

$$X^* = \frac{k_3}{k_2} \qquad \text{(Equation S15)}$$
$$Z_1^* = \frac{k_2(k_{in}U^* + b)}{k_1k_3} - \frac{\delta}{k_1} \qquad \text{(Equation S16)}$$
$$Z_2^* = \frac{k_3}{\eta\left(\frac{k_2(k_{in}U^* + b)}{k_1k_3} - \frac{\delta}{k_1}\right)} \qquad \text{(Equation S17)}$$

We now linearize Equations (S11), (S12), and (S13) around the fixed point defined by Equations (S15), (S16), and (S17) to obtain:

$$\begin{bmatrix} \dot{X} \\ \dot{Z}_1 \\ \dot{Z}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} -\frac{k_2(k_{in}U^* + b)}{k_3} & -\frac{k_1k_3}{k_2} & 0 \\ k_2 & -\eta Z_2^* & -\eta Z_1^* \\ 0 & -\eta Z_2^* & -\eta Z_1^* \end{bmatrix}}_{G_2} \begin{bmatrix} X \\ Z_1 \\ Z_2 \end{bmatrix}$$

The characteristic polynomial of G_2 is:

$$P_2(s) = \det(sI - G_2) = s^3 + \alpha_2 s^2 + \alpha_1 s + \alpha_0 \qquad \text{(Equation S18)}$$

where

$$\alpha_2 = \sigma + \eta(Z_1^* + Z_2^*) \qquad \text{(Equation S19)}$$
$$\alpha_1 = k_1k_3 + \sigma\eta(Z_1^* + Z_2^*) \qquad \text{(Equation S20)}$$
$$\alpha_0 = k_1k_3\eta Z_1^* \qquad \text{(Equation S21)}$$

and

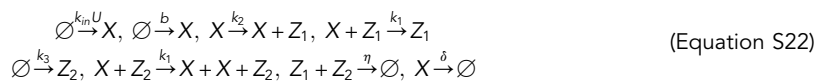$$\sigma = \frac{k_2(k_{in}U^* + b)}{k_3}$$

The polynomial given by Equation (S18) has all roots in the open-half plane if and only if α_2, α_0 are positive and $\alpha_2\alpha_1 > \alpha_0$ (Routh-Hurwitz criterion). Indeed:

$$\left(\sigma + \eta\left(Z_1^* + Z_2^*\right)\right)\left(k_1 k_3 + \sigma\eta\left(Z_1^* + Z_2^*\right)\right) > \eta k_1 k_3 z_1^*$$

or

$$\sigma k_1 k_3 + \sigma^2\eta\left(Z_1^* + Z_2^*\right) + k_1 k_3 \eta\left(Z_1^* + Z_2^*\right) + \sigma\eta^2\left(Z_1^* + Z_2^*\right)^2 > \eta k_1 k_3 z_1^*$$

or

$$\sigma k_1 k_3 + \sigma^2\eta\left(Z_1^* + Z_2^*\right) + \sigma\eta^2\left(Z_1^* + Z_2^*\right)^2 + \eta k_1 k_3 Z_2^* > 0$$

which is always true since all the quantities involved are positive. Therefore, (X^*, Z_1^*, Z_2^*) is a positive locally exponentially stable steady state (G_2 is Hurwitz) for the nonlinear system described by Equations (S11), (S12), and (S13).

Note that outside the parameter regime defined by Equation (S14) BioSD-II is unable to reach equilibrium. In particular, assuming non-negative initial conditions for Equations (S11), (S12), and (S13) (which is always the case because the variables involved represent biomolecular concentrations) the states of the latter remain always non-negative (as expected from mass action kinetics). Indeed, when $X = 0$, Equation (S11) implies $\dot{X} = k_{in}U + b > 0$. Furthermore, when $Z_1 = 0$, Equation (S12) results in $\dot{Z}_1 = k_2 X \geq 0$ and, finally, when $Z_2 = 0$, Equation (S13) imposes $\dot{Z}_2 = k_3 > 0$. However, outside the parameter regime in question, one of the following must hold: $k_2(k_{in}U^* + b) < \delta k_3$ or $k_2(k_{in}U^* + b) = \delta k_3$. In the first scenario, it is apparent from Equations (S16) and (S17) that the steady state of Z_1, Z_2 becomes negative while in the second case Equation (S17) indicates that Z_2 tends to infinity - thus, BioSD-II cannot approach a finite steady state.

Biomolecular signal differentiator-III. Biomolecular Signal Differentiator-III (BioSD-III) is represented by the CRN:

$$\varnothing \xrightarrow{k_{in}U} X, \ \varnothing \xrightarrow{b} X, \ X \xrightarrow{k_2} X + Z_1, \ X + Z_1 \xrightarrow{k_1} Z_1$$
$$\varnothing \xrightarrow{k_3} Z_2, \ X + Z_2 \xrightarrow{k_1} X + X + Z_2, \ Z_1 + Z_2 \xrightarrow{\eta} \varnothing, \ X \xrightarrow{\delta} \varnothing$$

(Equation S22)

where k_{in}, b, k_2, k_1, δ, $\eta \in \mathbb{R}_+$.

The corresponding ODE model describing the dynamics is

$$\dot{X} = k_{in}U + b - k_1 X Z_1 + k_1 X Z_2 - \delta X \qquad \text{(Equation S23)}$$
$$\dot{Z}_1 = k_2 X - \eta Z_1 Z_2 \qquad \text{(Equation S24)}$$
$$\dot{Z}_2 = k_3 - \eta Z_1 Z_2 \qquad \text{(Equation S25)}$$

For any constant input U^*, we have a unique positive steady state (providing that it exists and is finite):

$$X^* = \frac{k_3}{k_2} \qquad \text{(Equation S26)}$$

$$Z_1^* = \frac{1}{2}\left[\frac{k_2(k_{in}U^* + b)}{k_1 k_3} - \frac{\delta}{k_1}\right] + \frac{1}{2}\sqrt{\left[\frac{k_2(k_{in}U^* + b)}{k_1 k_3} - \frac{\delta}{k_1}\right]^2 + 4\frac{k_3}{n}} \qquad \text{(Equation S27)}$$

$$Z_2^* = -\frac{1}{2}\left[\frac{k_2(k_{in}U^* + b)}{k_1 k_3} - \frac{\delta}{k_1}\right] + \frac{1}{2}\sqrt{\left[\frac{k_2(k_{in}U^* + b)}{k_1 k_3} - \frac{\delta}{k_1}\right]^2 + 4\frac{k_3}{n}} \qquad \text{(Equation S28)}$$

Linearizing the system given by Equations (S23), (S24), and (S25) around its steady state (Equations (S26), (S27), and (S28)) yields:

$$\begin{bmatrix} \dot{X} \\ \dot{Z}_1 \\ \dot{Z}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} -\dfrac{k_2(k_{in}U^* + b)}{k_3} & -\dfrac{k_1 k_3}{k_2} & \dfrac{k_1 k_3}{k_2} \\ k_2 & -\eta Z_2^* & -\eta Z_1^* \\ 0 & -\eta Z_2^* & -\eta Z_1^* \end{bmatrix}}_{G_3} \begin{bmatrix} X \\ Z_1 \\ Z_2 \end{bmatrix}$$

The characteristic polynomial of G_3 is:

$$P_3(s) = \det(sI - G_3) = s^3 + \alpha\prime_2 s^2 + \alpha\prime_1 s + \alpha\prime_0 \qquad \text{(Equation S29)}$$

where $\alpha\prime_2$, $\alpha\prime_1$ are identical to α_2 (Equation S19), α_1 (Equation S20), respectively and:

$$\alpha\prime_0 = k_1 k_3 \eta \left(Z_1^* + Z_2^* \right)$$

In order to show that G_3 is Hurwitz we need to verify that $\alpha\prime_2 \alpha\prime_1 > \alpha\prime_0$ (from Routh-Hurwitz criterion).

This inequality is satisfied because:

$$\left(\sigma + \eta \left(Z_1^* + Z_2^* \right) \right) \left(k_1 k_3 + \sigma\eta \left(Z_1^* + Z_2^* \right) \right) > \eta k_1 k_3 \left(Z_1^* + Z_2^* \right)$$
or
$$\sigma k_1 k_3 + \sigma^2 \eta \left(Z_1^* + Z_2^* \right) + k_1 k_3 \eta \left(Z_1^* + Z_2^* \right) + \sigma\eta^2 \left(Z_1^* + Z_2^* \right)^2 > k_1 k_3 \eta \left(Z_1^* + Z_2^* \right)$$
or
$$\sigma k_1 k_3 + \sigma^2 \eta \left(Z_1^* + Z_2^* \right) + \sigma\eta^2 \left(Z_1^* + Z_2^* \right)^2 > 0$$

which is always true as a sum of positive quantities. Hence, (X^*, Z_1^*, Z_2^*) is a positive locally exponentially stable steady state for the nonlinear system described by Equations (S23), (S24), and (S25).

The notion of strong rate of annihilation between Z_1, Z_2 (large η) in Biomolecular Signal Differentiator-II

This reaction describes a process where species Z_1, Z_2 bind to each other irreversibly to form a product which can be considered as biologically inactive. In other words, this product does not participate in any of the reactions in BioSD-II. Here we demonstrate that the steady state of Z_2 as well as its deviation from it is practically negligible if the formation rate, η, of the product in question is sufficiently high. At the same time, the effect of Z_2 on the dynamics of BioSD-II can be considered insignificant, too.

By adopting the coordinate transformations: $u = U - U^*$, $x = X - X^*$, $z_1 = Z_1 - Z_1^*$, $z_2 = Z_2 - Z_2^*$ which denote small perturbations around (U^*, X^*, Z_1^*, Z_2^*), we obtain through linearization of Equations (S11), (S12), and (S13):

$$\begin{bmatrix} \dot{x} \\ \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -\dfrac{k_2 (k_{in} U^* + b)}{k_3} & -\dfrac{k_1 k_3}{k_2} & 0 \\ k_2 & -\eta Z_2^* & -\eta Z_1^* \\ 0 & -\eta Z_2^* & -\eta Z_1^* \end{bmatrix} \begin{bmatrix} x \\ z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} k_{in} \\ 0 \\ 0 \end{bmatrix} u \qquad \text{(Equation S30)}$$

We now introduce the non-dimensional variables:

$$t_n = \beta_1 t \qquad \text{(Equation S31)}$$

$$x_n = \frac{1}{\beta_2} \overline{x} \qquad \text{(Equation S32)}$$

$$z_{1n} = \frac{\beta_1}{\beta_2 k_2} z_1 \qquad \text{(Equation S33)}$$

$$z_{2n} = \frac{\beta_1}{\beta_2 k_2} z_2 \qquad \text{(Equation S34)}$$

$$u_n = \frac{k_{in}}{\beta_1 \beta_2} u \qquad \text{(Equation S35)}$$

where

$$\beta_1 = \frac{k_3}{\left(\dfrac{k_2 (k_{in} U^* + b)}{k_1 k_3} - \dfrac{\delta}{k_1} \right)} \qquad \text{(Equation S36)}$$

and β_2 is an arbitrary scaling parameter that carries the same units as x_n. In addition, we introduce the non-dimensional parameters:

$$\lambda_1 = \frac{\beta_1^2}{\eta k_3} \qquad \text{(Equation S37)}$$

$$\lambda_2 = \frac{k_2 (k_{in} U^* + b)}{\beta_1 k_3} \qquad \text{(Equation S38)}$$

$$\lambda_3 = \frac{k_1 k_3}{\beta_1 k_2} \qquad \text{(Equation S39)}$$

By substituting Equations (S31), (S32), (S33), (S34), (S35), (S36), (S37), (S38), and (S39) into the model given by Equation (S30), we obtain:

$$\dot{x}_n = u_n - \lambda_2 x_n - \lambda_3 z_{1n}$$
$$\dot{z}_{1n} = k_2 x_n - z_{1n} - \frac{1}{\lambda_1} z_{2n}$$
$$\dot{z}_{2n} = -z_{1n} - \frac{1}{\lambda_1} z_{2n}$$

or

$$\dot{x}_n = u_n - \lambda_2 x_n - \lambda_3 z_{1n}$$
$$\lambda_1 \dot{z}_{1n} = \lambda_1 x_n - \lambda_1 z_{1n} - z_{2n}$$
$$\lambda_1 \dot{z}_{2n} = -\lambda_1 z_{1n} - z_{2n}$$

We now introduce the linear transformation $g_n = z_{1n} - z_{2n}$ resulting in the following mathematically equivalent system:

$$\dot{x}_n = u_n - \lambda_2 x_n - \lambda_3 g_n - \lambda_3 z_{2n} \qquad \text{(Equation S40)}$$
$$\dot{g}_n = x_n \qquad \text{(Equation S41)}$$
$$\lambda_1 \dot{z}_{2n} = -\lambda_1 g_n - (1 + \lambda_1) z_{2n} \qquad \text{(Equation S42)}$$

According to Equation (S37), $\lambda_1 \to 0$ as $\eta \to \infty$. This means that we can make λ_1 negligible by choosing a large value for η:

$$\eta \gg \frac{\beta_1^2}{k_3} \qquad \text{(Equation S43)}$$

We now regard λ_1 as a singular perturbation parameter and use Theorem 11.1 in (H. K. Khalil, 2002). From Equations (S40), (S41), and (S42) we obtain the following reduced model for $\lambda_1 = 0$:

$$\dot{x}_n = u_n - \lambda_2 x_n - \lambda_3 g_n \qquad \text{(Equation S44)}$$
$$\dot{g}_n = x_n \qquad \text{(Equation S45)}$$

since $z_{2n} = 0$.

For a finite time interval $[0, t_f]$ of interest, Equations (S44) and (S45) produce a unique solution $\overline{x}_n(t), \overline{g}_n(t)$ taking into account the initial conditions of the system. In addition, the origin is an exponentially stable equilibrium point of the boundary layer model:

$$\frac{dz_{2n}}{d\tau} = -z_{2n}$$

where $\tau = t_n / \lambda_1$.

Thus, according to Tikhonov's theorem (Theorem 11.1 in (Khalil, 2002)), there exist a positive constant λ_1^* such that for $0 < \lambda_1 < \lambda_1^*$ the singular perturbation problem of Equations (S40), (S41), and (S42) has a unique solution $x_n(t, \lambda_1)$, $g_n(t, \lambda_1)$, $z_{2n}(t, \lambda_1)$ on $[0, t_f]$ and

$$x_n(t, \lambda_1) - \overline{x}_n(t) = \mathcal{O}(\lambda_1)$$
$$g_n(t, \lambda_1) - \overline{g}_n(t) = \mathcal{O}(\lambda_1)$$

Moreover, given any $t_b > 0$, there is λ_1^{**} such that

$$z_{2n}(t, \lambda_1) = \mathcal{O}(\lambda_1)$$

whenever $\lambda_1 < \lambda_1^{**}$.

Finally, combining Equations (S16), (S17), (S36), and (S37) results in:

$$\lambda_1 = \frac{Z_2^*}{Z_1^*}$$

Assuming that Z_1^* corresponds to some finite (nonzero) concentration, $Z_2^* \to 0$ as $\lambda_1 \to 0$.

Behavior analysis of biomolecular signal differentiators

Here we prove that, near their equilibria, BioSD networks are capable of signal differentiation.

We begin with BioSD-I whose dynamics close to its steady state are derived via linearization of Equations (S4) and (S5) as:

$$\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} -\dfrac{k_2(k_{in}U^* + b)}{k_3} & -\dfrac{k_1 k_3}{k_2} \\ k_2 & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} k_{in} \\ 0 \end{bmatrix} u \qquad \text{(Equation S46)}$$

assuming the coordinate transformations: $u = U - U^*$, $x = X - X^*$, $z = Z - Z^*$ which represent small perturbations around (U^*, X^*, Z^*). Note that u represents U^{TV} of the main text. We next consider the non-dimensional variables:

$$t_n = c_1 t \qquad \text{(Equation S47)}$$

$$x_n = \frac{1}{c_3} x \qquad \text{(Equation S48)}$$

$$z_n = \frac{c_1}{k_2 c_3} z \qquad \text{(Equation S49)}$$

$$u_n = \frac{c_1 k_{in}}{k_2 c_2 c_3} u \qquad \text{(Equation S50)}$$

where

$$c_1 = \frac{k_2(k_{in}U^* + b)}{k_3} \qquad \text{(Equation S51)}$$

$$c_2 = \frac{k_1 k_3}{k_2} \qquad \text{(Equation S52)}$$

and c_3 is an arbitrary scaling parameter that carries the same units as x_n. We also introduce the non-dimensional parameter:

$$\varepsilon = \frac{c_1^2}{k_2 c_2} \qquad \text{(Equation S53)}$$

Substituting Equations (S47), (S48), (S49), (S50), (S51), (S52), and (S53) into the system (S46) results in:

$$\dot{x}_n = -x_n - \frac{1}{\varepsilon}z_n + \frac{1}{\varepsilon}u_n$$
$$\dot{z}_n = x_n$$

or

$$\varepsilon\dot{x}_n = -\varepsilon x_n - z_n + u_n \qquad \text{(Equation S54)}$$
$$\dot{z}_n = x_n \qquad \text{(Equation S55)}$$

The system described by Equations (S54), (S55) is mathematically equivalent to the following second-order differential equation:

$$\varepsilon\ddot{x}_n + \varepsilon\dot{x}_n + x_n = \dot{u}_n \qquad \text{(Equation S56)}$$

We see immediately that if $\varepsilon(\ddot{x}_n + \dot{x}_n) = 0$ then $x_n = \dot{u}_n$ which gives through Equations (S47), (S48), (S50), and (S52):

$$x = \frac{k_{in}}{k_1 k_3}\dot{u} \qquad \text{(Equation S57)}$$

By recalling Equation (S6) and our initial coordinate transformations, this relationship can be rewritten as:

$$X = \frac{k_{in}}{k_1 k_3}\dot{U} + \frac{k_3}{k_2} \qquad \text{(Equation S58)}$$

Having this in mind and taking into account that ε is positive as a combination of positive parameters (Equation (S53)) we calculate the general solution of $\ddot{x}_n + \dot{x}_n = 0$ as:

$$x_n = \theta_1 e^{-t_n} + \theta_2 \qquad \text{(Equation S59)}$$

where θ_1, θ_2 are arbitrary constants. Subsequently, from Equations (S47), (S48), (S57), and (S59) we get:

$$u = f_1 e^{-c_1 t} + f_2 t + f_3 \qquad \text{(Equation S60)}$$

where f_1, f_2, f_3 are arbitrary constants.

To study the behavior of BioSD-I in the more general case where the input signal does not satisfy Equation (S60) we consider the following transfer function describing the system defined by Equations (S54) and (S54) in the Laplace domain:

$$\tilde{\Delta}_{BSD}(s) = \frac{\tilde{X}_n(s)}{\tilde{U}_n(s)} = \frac{s}{\varepsilon(s^2 + s) + 1} \qquad \text{(Equation S61)}$$

where $\tilde{X}_n(s)$ and $\tilde{U}_n(s)$ are the Laplace transform of the output x_n and input u_n, respectively and s is the complex frequency. As can be seen, BioSD-I can compute the derivative of the input signal filtered by a second - order low - pass filter.

As pointed out in Signals under consideration, Fourier transform is a powerful tool that allows the decomposition of a signal into its constituent sinusoids. Thus, focusing on the frequency response of the system, we set $s = j\omega$ (where $j = \sqrt{-1}$) in Equation (S61) to get:

$$\tilde{\Delta}_{BSD}(j\omega) = \frac{j\omega}{\varepsilon(-\omega^2 + j\omega) + 1} \qquad \text{(Equation S62)}$$

which can be equivalently represented by:

$$|\tilde{\Delta}_{BSD}(j\omega)| = \frac{\omega}{\sqrt{\varepsilon^2 \omega^2 + (1 - \varepsilon\omega^2)^2}} \qquad \text{(Equation S63)}$$

and

$$\angle \tilde{\Delta}_{BSD}(j\omega) = \arctan\left(\frac{1}{\varepsilon\omega} - \omega\right) \qquad \text{(Equation S64)}$$

As shown in Figure S3, for a given ε, there is a low-frequency range over which BioSD-I functions as a pure signal differentiator and, by extension Equation (S58) holds (the filtering action is practically zero), and a high-frequency one over which it works as a signal attenuator instead. At the same time, there is a narrow frequency band in between where the aforementioned operations may not be carried out with the expected accuracy. The behavior of BioSD-I therefore depends on the value of ε as well as on "how fast" an input signal varies over time.

Following the same procedure, we study the local dynamics of BioSD-III by linearizing Equations (S23), (S24), and (S25):

$$\begin{bmatrix} \dot{x} \\ \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -\dfrac{k_2(k_{in}U^* + b)}{k_3} & -\dfrac{k_1 k_3}{k_2} & \dfrac{k_1 k_3}{k_2} \\ k_2 & -\eta Z_2^* & -\eta Z_1^* \\ 0 & -\eta Z_2^* & -\eta Z_1^* \end{bmatrix} \begin{bmatrix} x \\ z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} k_{in} \\ 0 \\ 0 \end{bmatrix} u$$

where the variables $u = U - U^*$, $x = X - X^*$, $z_1 = Z_1 - Z_1^*$, $z_2 = Z_2 - Z_2^*$ refer to small perturbations around the equilibrium (U^*, X^*, Z_1^*, Z_2^*). Introducing the linear transformation $g = z_1 - z_2$ results in the following mathematically equivalent system:

$$\begin{bmatrix} \dot{x} \\ \dot{g} \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -\dfrac{k_2(k_{in}U^* + b)}{k_3} & -\dfrac{k_1 k_3}{k_2} & 0 \\ k_2 & 0 & 0 \\ 0 & -\eta Z_2^* & -\eta(Z_1^* + Z_2^*) \end{bmatrix} \begin{bmatrix} x \\ g \\ z_2 \end{bmatrix} + \begin{bmatrix} k_{in} \\ 0 \\ 0 \end{bmatrix} u \qquad \text{(Equation S65)}$$

We notice that the dynamics of x and g of the system given by Equation (S65) are identical to that of x and z of the system given by Equation (S46), respectively. Hence, the output, x, of BioSD-III behaves in the exact same way as the one of previously analyzed BioSD-I.

Subsequently, we recall Equation (S30) describing the dynamics of BioSD-II near its equilibrium. It is evident that using the linear transformation $g = z_1 - z_2$ again and assuming a sufficiently large η (Equation (S43)
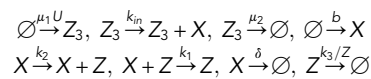
holds), the dynamics of x and g in BioSD-II are described by Equation (S46), namely the dynamics of BioSD-I (see The notion of strong rate of annihilation between Z_1, Z_2 (large η) in Biomolecular Signal Differentiator-II). By extension, the output behavior of these two circuits is identical.

An alternative version of biomolecular signal differentiators

Here we analyze a slightly different version of the previously studied BioSD networks which we call Biomolecular Signal Differentiators^F (BioSDs^F) that include an additional biomolecular species, Z_3. In particular, we describe the following three biomolecular topologies:
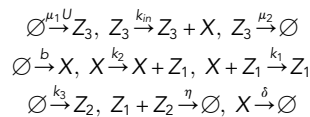
- BioSD^F-I

We have the CRN:

$$\varnothing \xrightarrow{\mu_1 U} Z_3, \ Z_3 \xrightarrow{k_{in}} Z_3 + X, \ Z_3 \xrightarrow{\mu_2} \varnothing, \ \varnothing \xrightarrow{b} X$$
$$X \xrightarrow{k_2} X + Z, \ X + Z \xrightarrow{k_1} Z, \ X \xrightarrow{\delta} \varnothing, \ Z \xrightarrow{k_3/Z} \varnothing$$

where μ_1, μ_2, k_{in}, b, k_2, k_1, δ, $k_3 \in \mathbb{R}_+$. The 0th-order removal of Z is the result of enzymatic degradation following saturated Michaelis - Menten kinetics (see Equilibria and stability of biomolecular signal differentiators: biomolecular signal differentiator-I).

The corresponding ODE model is:

$$\dot{Z}_3 = \mu_1 U - \mu_2 Z_3$$
$$\dot{X} = k_{in} Z_3 + b - k_1 XZ - \delta X$$
$$\dot{Z} = k_2 X - k_3$$

- BioSD^F-II

We have the CRN:

$$\varnothing \xrightarrow{\mu_1 U} Z_3, \ Z_3 \xrightarrow{k_{in}} Z_3 + X, \ Z_3 \xrightarrow{\mu_2} \varnothing$$
$$\varnothing \xrightarrow{b} X, \ X \xrightarrow{k_2} X + Z_1, \ X + Z_1 \xrightarrow{k_1} Z_1$$
$$\varnothing \xrightarrow{k_3} Z_2, \ Z_1 + Z_2 \xrightarrow{\eta} \varnothing, \ X \xrightarrow{\delta} \varnothing$$
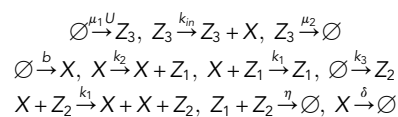
where μ_1, μ_2, k_{in}, b, k_2, k_1, δ, $\eta \in \mathbb{R}_+$. We assume that the parameter rate η is sufficiently large (see The notion of strong rate of annihilation between Z_1, Z_2 (large η) in Biomolecular Signal Differentiator-II).

The corresponding ODE model is:

$$\dot{Z}_3 = \mu_1 U - \mu_2 Z_3$$
$$\dot{X} = k_{in} Z_3 + b - k_1 XZ_1 - \delta X$$
$$\dot{Z}_1 = k_2 X - \eta Z_1 Z_2$$
$$\dot{Z}_2 = k_3 - \eta Z_1 Z_2$$

- BioSD^F-III

We have the CRN:

$$\varnothing \xrightarrow{\mu_1 U} Z_3, \ Z_3 \xrightarrow{k_{in}} Z_3 + X, \ Z_3 \xrightarrow{\mu_2} \varnothing$$
$$\varnothing \xrightarrow{b} X, \ X \xrightarrow{k_2} X + Z_1, \ X + Z_1 \xrightarrow{k_1} Z_1, \ \varnothing \xrightarrow{k_3} Z_2$$
$$X + Z_2 \xrightarrow{k_1} X + X + Z_2, \ Z_1 + Z_2 \xrightarrow{\eta} \varnothing, \ X \xrightarrow{\delta} \varnothing$$

where μ_1, μ_2, k_{in}, b, k_2, k_1, δ, $\eta \in \mathbb{R}_+$.

The corresponding ODE model is:

$$\dot{Z}_3 = \mu_1 U - \mu_2 Z_3$$

$$\dot{X} = k_{in}Z_3 + b - k_1XZ_1 + k_1XZ_2 - \delta X$$
$$\dot{Z}_1 = k_2X - \eta Z_1Z_2$$
$$\dot{Z}_2 = k_3 - \eta Z_1Z_2$$

Each of the above circuits can be seen as the interconnection of two subsystems. More specifically, we have the linear, asymptotically stable, subsystem (the first equation in each of above ODE models):

$$\dot{Z}_3 = \mu_1 U - \mu_2 Z_3 \qquad \text{(Equation S66)}$$

which receives the signal U we want to differentiate as input and produces an output Z_3. This is, in turn, applied as input to a second subsystem whose output is X. While the first subsystem is the same in all BioSD^F topologies, the second one differs. In fact, the latter is identical to BioSD-I, BioSD-II, BioSD-III (see previous sections) for BioSD^F-I, BioSD^F-II, BioSD^F-III, respectively, with the only difference lying in the input, which is now Z_3 (instead of U as before).

For a constant input U^* the first subsystem defined by Equation (S66) has a unique positive steady state (assuming it exists and is finite):

$$Z_3 = \frac{\mu_1 U^*}{\mu_2} \qquad \text{(Equation S67)}$$

Since Equation (S66) is linear and (μ_2) is always positive then Equation (S67) is a globally exponentially stable equilibrium point.

We now concentrate on the local behavior of BioSD^F modules and, consequently, we consider the coordinate transformations: $u = U - U^*$, $x = X - X^*$, $z = Z - Z^*$, $z_1 = Z_1 - Z_1^*$, $z_2 = Z_2 - Z_2^*$, $z_3 = Z_3 - Z_3^*$ denoting small perturbations around the corresponding equilibria of BioSD^F networks - (U^*, X^*, Z^*, Z_3^*) for BioSD^F-I and ($U^*, X^*, Z_1^*, Z_2^*, Z_3^*$) for BioSD^F-II, BioSD^F-III (the steady states of the last two networks do not necessarily coincide).

First, we study Equation (S66) separately. In the Laplace domain, we have:

$$\tilde{\Delta}_{LPF}(s) = \frac{\tilde{Z}_3(s)}{\tilde{U}(s)} = \frac{\mu_1}{s + \mu_2} \qquad \text{(Equation S68)}$$

where $\tilde{Z}_3(s)$, $\tilde{U}(s)$ are the Laplace transform of z_3, u, respectively. Focusing on the frequency response, we get:

$$\tilde{\Delta}_{LPF}(j\omega) = \frac{\mu_1}{\mu_2} \frac{1}{j\dfrac{\omega}{\mu_2} + 1} \qquad \text{(Equation S69)}$$

This is a transfer function of a first-order low-pass filter which is capable of preserving low-frequency signals and rejecting high-frequency signals. Indeed, the magnitude and the phase of the system in question are given by:

$$\left| \tilde{\Delta}_{LPF}(j\omega) \right| = \frac{\mu_1}{\mu_2} \frac{1}{\sqrt{1 + \left(\frac{\omega}{\mu_2} \right)^2}}$$

and

$$\angle \tilde{\Delta}_{LPF}(j\omega) = -\arctan \frac{\omega}{\mu_2},$$

respectively.

We can easily see that in practice, when $\omega \ll \mu_2$, there is a constant input/output gain $\left(\frac{\mu_1}{\mu_2} \right)$ and no phase lag. On the other hand, for $\omega^2 \gg \mu_2^2$ strong attenuation takes place. The general behavior of the filter can be easily understood through the Bode diagram in Figure S4.

67

CellPress
OPEN ACCESS

We now consider a BioSD^F design which can be described by the transfer function of the series connection of the previously studied filter and a BioSD design (as already outlined in Behavior analysis of biomolecular signal differentiators, all three BioSD circuits are described by the same transfer function), i.e.:

$$\tilde{\Delta}_{BSD^F}(s) = \frac{\tilde{X}_n(s)}{\tilde{U}_n(s)} = \tilde{\Delta}_{LPF}(s)\tilde{\Delta}_{BSD}(s)$$

or

$$\tilde{\Delta}_{BSD^F}(s) = \frac{\mu_1}{s+\mu_2} \cdot \frac{s}{\varepsilon(s^2+s)+1} \qquad \text{(Equation S70)}$$

where $\tilde{\Delta}_{LPF}(s) = \frac{\tilde{Z}_3(s)}{\tilde{U}_n(s)}$, $\tilde{\Delta}_{BSD}(s) = \frac{\tilde{X}_n(s)}{\tilde{Z}_{3n}(s)}$ with $\tilde{Z}_{3n}(s) = p\tilde{Z}_3(s)$ and $p = \frac{u_n}{u}$ (see Behavior analysis of biomolecular signal differentiators).

Shifting our focus on the frequency response we have:

$$\tilde{\Delta}_{BSD^F}(j\omega) = \tilde{\Delta}_{LPF}(j\omega)\tilde{\Delta}_{BSD}(j\omega) \qquad \text{(Equation S71)}$$

for which:

$$\left|\tilde{\Delta}_{BSD^F}(j\omega)\right| = |\tilde{\Delta}_{LPF}(j\omega)||\tilde{\Delta}_{BSD}(j\omega)|$$

and

$$\angle\tilde{\Delta}_{BSD^F}(\omega) = \angle\tilde{\Delta}_{LPF}(j\omega) + \angle\tilde{\Delta}_{BSD}(j\omega)$$

Consequently, for a given ε, BioSD^F circuits are characterized by an enhanced capability of high-frequency signal attenuation compared to BioSD ones. In fact, as demonstrated in Figure S5, we can extend the frequency band where strong signal attenuation is carried out by appropriately tuning the filter module. In other words, we can adjust the bandwidth of the extra filter as desired through the parameter rate μ_2. The price we pay for this significant improvement is the increase in structural complexity due to the addition of the species Z_3 via which the additional filtering is accomplished. Finally, in the low-frequency regime, where only signal differentiation takes place (the filtering action is practically zero), the BioSD^F output can be approximated in the time domain as (recall Behavior Analysis of Biomolecular Signal Differentiators):

$$X = \frac{\mu_1 k_{in}}{\mu_2 k_1 k_3}\dot{U} + \frac{k_3}{k_2}$$

Analysis of the experimental topology of Biomolecular Signal Differentiator-III

Here we further analyze the proposed synthetic design of BioSD-III, the behavior of which may be more complicated due to the use of three auxiliary species (see Guidelines for experimental implementation of biomolecular signal differentiators).

The biomolecular topology shown in Figure 8C can be described by the following set of ODEs:

$$\dot{X} = k_{in}U + b - k_1 X Z_1 + k_{1a}X_{aux}Z_{2,aux} - \delta X \qquad \text{(Equation S72)}$$
$$\dot{X}_{aux} = k_{in}U + b - k_{1b}X_{aux}Z_1 + k_{1a}X_{aux}Z_{2,aux} - \delta_a X_{aux} \qquad \text{(Equation S73)}$$
$$\dot{Z}_1 = k_2 X - \eta Z_1 Z_2 \qquad \text{(Equation S74)}$$
$$\dot{Z}_{1,aux} = k_2 X - \eta_a Z_{1,aux}Z_{2,aux} \qquad \text{(Equation S75)}$$
$$\dot{Z}_2 = k_3 - \eta Z_1 Z_2 \qquad \text{(Equation S76)}$$
$$\dot{Z}_{2,aux} = k_3 - \eta_a Z_{1,aux}Z_{2,aux} \qquad \text{(Equation S77)}$$

where $k_{in}, b, k_2, k_1, k_{1a}, k_{1b}, \delta, \delta_a, \eta, \eta_a \in \mathbb{R}_+$.

In order for the behavior of X (measured output species) in the system described by Equations (S72)-(S77) to perfectly match the one of X in the model given by Equations (S23)-(S25), we need: $k_1 = k_{1a} = k_{1b}$, $\delta = \delta_a$ and $\eta = \eta_a$. Nevertheless, non-satisfaction of the aforementioned conditions does not necessarily entail considerable loss of accuracy regarding signal differentiation (Figure S6).

Modeling a more realistic case of Biomolecular Signal Differentiator-II

Here we study the behavior of Biomolecular Signal Differentiator-II under more realistic conditions resulting from the corresponding experimental design discussed in Guidelines for experimental implementation of biomolecular signal differentiators.

First, we consider the ODE model:

$$\dot{X} = k_{in}U + b - k_1XZ_1 - \delta X \qquad \text{(Equation S78)}$$

$$\dot{Z}_1 = V_{max}\frac{X}{X + K_m} - \eta Z_1Z_2 \qquad \text{(Equation S79)}$$

$$\dot{Z}_2 = k_3 - \eta Z_1Z_2 \qquad \text{(Equation S80)}$$

For a constant input U^*, provided that:

$$(k_{in}U^* + b) > \delta X^*,$$

we have a unique positive (finite) steady state:

$$X^* = \frac{k_3 K_m}{V_{max} - k_3} \qquad \text{(Equation S81)}$$

$$Z_1^* = \frac{(k_{in}U^* + b)}{k_1 X^*} - \frac{\delta}{k_1} \qquad \text{(Equation S82)}$$

$$Z_2^* = \frac{k_3}{\eta Z_1^*} \qquad \text{(Equation S83)}$$

Compared to the original model of BioSD-II (Equations (S11), (S12), and (S13)), we now use a Michaelis-Menten function to describe the activation of species Z_1 by species X (Equation S79) through gene expression (Aoki et al., 2019). It is evident that, assuming small perturbations around (U^*, X^*, Z_1^*, Z_2^*), linearization of Equations (S78), (S79), and (S80) yields a system of the same form as Equation (S30). Consequently, we can follow a similar analysis to study its local behavior as the one used for the original model (see The notion of strong rate of annihilation between Z_1, Z_2 (large η) in biomolecular signal differentiator-II and Behavior analysis of biomolecular signal differentiators). Nevertheless, it should be emphasized that when no saturation occurs and the slope of the Michaelis-Menten function is approximately linear, the corresponding production rate can be effectively considered proportional to the concentration of the regulator species (ibid.). In that case, the results of our original analysis can be used directly.

Implementation of BioSD-II in living cells implies the existence of an additional degradation mechanism due to cell growth affecting all the biomolecules involved, known as dilution (Aoki et al., 2019; Qian and Del Vecchio, 2018). This can lead to a "leaky" integration process realized by species Z_1, Z_2 and, by extension, it can affect the output response (see Achieving biological signal differentiation). To this end, we consider the following, more complex, ODE model:

$$\dot{X} = k_{in}U + b - k_1XZ_1 - (\delta + \gamma)X \qquad \text{(Equation S84)}$$

$$\dot{Z}_1 = V_{max}\frac{X}{X + K_m} - \eta Z_1Z_2 - \gamma Z_1 \qquad \text{(Equation S85)}$$

$$\dot{Z}_2 = k_3 - \eta Z_1Z_2 - \gamma Z_2 \qquad \text{(Equation S86)}$$

where γ represents a dilution rate constant.

In general, linearization of Equations (S84), (S85), and (S86) around their steady-state (which is obviously different than before) results in a system which does not have the same form as Equation (S30) and, thus, the procedures of our original analysis are not valid here. Nevertheless, if the dilution effect is not strong, it can be seen from simulations that the behavior of this model approaches the one of Equations (S78), (S79), and (S80).

Note that the above structural "perturbations" appear also in the natural systems discussed in Biomolecular signal differentiators in natural regulatory networks. In parallel, activation of species X by Z and Z_1 in BioSD-I and BioSD-III, respectively is also done through gene expression (see Guidelines for experimental implementation of biomolecular signal differentiators). In addition, dilution is present when realizing the latter topologies in living cells. Consequently, we can draw similar conclusions about them as with BioSD-II.

We now numerically investigate the behavior of BioSD-II. Figure S7A shows the response of the system given by Equations (S78), (S79), and (S80) to the input presented in Figure 3C using the parameter rates in Table S1, except for the dilution rate γ which is considered zero. As can be seen, BioSD-II can accurately calculate the rate of change of the input applied.

Note also the following:

- From Equation (S85) and Table S1 we calculate the steady-state concentration of species X which is equal to 20 nM. Based on the values of V_{max}, K_m and taking into account that X moves around the aforementioned point, the production rate of species Z_1 can be approximated well by the term $k_2 X$, where $k_2 \approx 1$ (no saturation occurs).

- To facilitate the comparison of the BioSD output with the derivative of the input we choose a value for k_{in} equal to the value of the quantity $k_1 k_3$ (see Equation (S58)). At the same time, here input U represents an actuator species whose concentration is related linearly with the corresponding production rate of output species X (which may result from the linear regime of a Hill function as discussed above). Nevertheless, in the general case the term $k_{in}U$ can represent any (nonlinear) function describing the activating mechanism of the output species.

- From Equation (S53) we get $\varepsilon \approx 0.125$. Moreover, η can be considered sufficiently large since $\eta = 425$ nM^{-1} min$^{-1} \gg \frac{\beta_1^2}{k_3} \approx 14.18$ nM^{-1} min^{-1} (see Equation (S36)). Consequently, Equation (S43) holds.

- Protein production rates regarding gene expression can be easily adjusted, for example, by changing gene copy number and, thus, a wide range of values can be achieved - a typical parameter range for *E. coli* is $0.5 - 10^4$ nM nM (Aoki et al., 2019). This implies extensive tunability which is important for meeting different performance standards (see Tunability and accuracy) since a considerable number of parameter rates in BioSD-II is associated with gene expression, i.e. b, k_2 (which is related to V_{max}, K_m), k_3 and k_{in}.

Figure S7B shows the response of the system given by Equations (S84), (S85), and (S86) to the same input stimulus. We also use the same parameters rates as before except for the dilution rate which is now nonzero and equal to a typical value for *E. coli* (see Table S1). It is evident that the output remains an accurate replica of the derivative of the input.

Subsequently, in Figures S7C and S7D we further investigate the impact of dilution on the output of BioSD-II by repeating the simulation of Figure S7B with a 5 and 10 times larger dilution rate, respectively. We notice that as this rate gets stronger the actual response moves away from the zero-level "bias" which coincides with the corresponding output steady-state. Moreover, although the accuracy drops to some extent, the form of the output remains close to the one of the ideal derivative.

As already pointed out, the annihilation rate η is chosen to be sufficiently large so that the condition given by Equation (S43) is satisfied (only BioSD-II entails such a requirement). More specifically, η is approximately 30 times larger than the quantity $\frac{\beta_1^2}{k_3}$. Nevertheless, it remains unclear to us if such suitable values of η can be always guaranteed *in vitro* by the interaction between the pair of protease/protease inhibitor proposed in Guidelines for experimental implementation of biomolecular signal differentiators. It is therefore important to investigate the behavior of the differentiator module in the case where η is not as large as our theoretical analysis demands. As shown in Figure S8, non-satisfaction of the condition given by Equation (S43) does not necessarily entail significant loss of accuracy regarding signal differentiation. Note also that the quantity $\frac{\beta_1^2}{k_3}$ can be easily adjusted to a suitable value by appropriately tuning the protein production rates involved in BioSD-II (discussed earlier).

Finally, to make the above analysis even more realistic (Del Vecchio and Murray, 2015), one could model gene expression as a multi-stage process, thus capturing the dynamics of transcription and translation. At the same time, the dynamics of complexes participating in intermediate stages of inhibition and annihilation reactions could also be considered. Nonetheless, it is important to emphasize that such an approach would increase the complexity of the resulting mathematical models.

# Supplemental information

Biomolecular mechanisms for signal differentiation

**Emmanouil Alexis, Carolin C.M. Schulte, Luca Cardelli, and Antonis Papachristodoulou**

	Description	Value	Unit	Comments	Source
γ	Dilution rate	0.028	min^{-1}	Value for *E. coli*, assuming 25 min doubling time	Aoki et al., 2019
δ	Degradation rate	0.1	min^{-1}	Unspecified mechanism (disturbance) contributing to degradation	Aoki et al., 2019
η	Annihilation rate	425	nM^{-1} min^{-1}	Value based on binding rates for protein-protein interactions that are diffusion-limited	Schlosshauer and Baker, 2004; Fekkes, Blaauwen, and Driessen, 1995
k_1	Catalytic inhibition rate	1.6	nM^{-1} min^{-1}	Value based on the action of Lon protease	Gur, Vishkautzan, and Sauer, 2012
k_3	Constitutive production rate	20	nM min^{-1}		Buchler and Louis, 2008; Aoki et al., 2019
b	Constitutive production rate	40	nM min^{-1}		Buchler and Louis, 2008; Aoki et al., 2019
V_{max}	Maximal production rate	900	nM min^{-1}		Buchler and Louis, 2008; Aoki et al., 2019
K_m	Michaelis-Menten constant	880	nM		Buchler and Louis, 2008; Aoki et al., 2019

Table S1: Simulation parameters for STAR Methods **Modelling a more realistic case of Biomolecular Signal Differentiator-II**
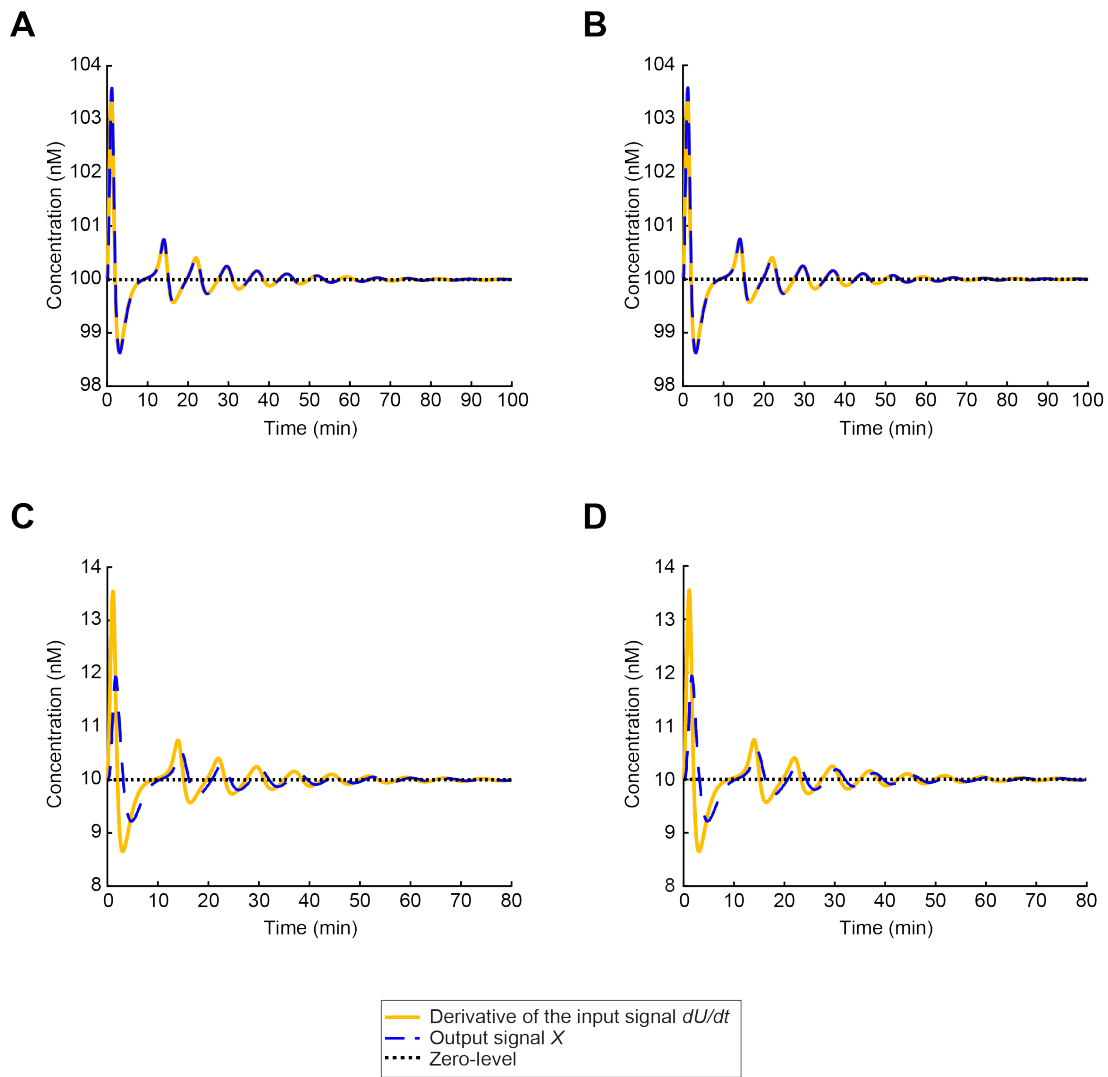
Figure S1: Sensing the rate-of-change of a synthetic regulatory biomolecular network through a Biomolecular Signal Differentiator. Related to Figure 3.

a Simulation of the BioSD-II (Equations (S11)-(S13)) response to the input presented in Figure 3c with $\eta = 3000$ nM^{-1} min^{-1}, $b = 150$ nM min^{-1} and the remaining parameters same as those used in Figure 3d. η can be characterized as sufficiently large since condition (S43) is satisfied. **b** Simulation of the BioSD-III (Equations (S23)-(S25)) response to the input presented in Figure 3c with $\eta = 30$ nM^{-1} min^{-1} and the remaining parameters same as those used in Figure 3d. **c** The simulation in **a** is repeated with the values of k_{in}, k_3, b set to 10, 10 and 100, respectively. **d** The simulation in **b** is repeated with the values of both k_{in} and k_3 set to 10. As can be seen, the behaviour of both BioSD-II and BioSD-III is identical to that of BioSD-I depicted in the main text. As a result, the conclusions drawn with respect to the latter circuit are valid for the other designs as well.

A

B

Legend:
- Derivative of the input signal dU/dt
- Output signal X
- Zero-level
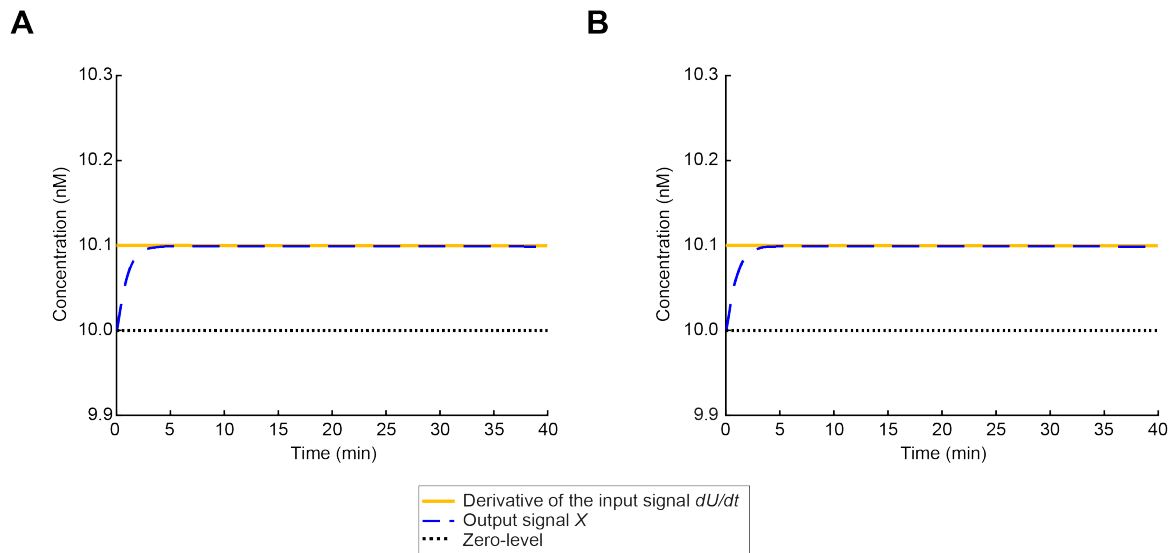
Figure S2: Sensing the rate-of-change of a birth-death biomolecular process through a Biomolecular Signal Differentiator. Related to Figure 4.

a Simulation of the BioSD-II (Equations (S11)-(S13)) response to the input presented in Figure 4c with $\eta = 3000$ nM^{-1} min^{-1} and the remaining parameter same as those used in Figure 4d. η can be described as sufficiently large since condition (S43) is satisfied. **b** Simulation of the BioSD-III (Equations (S23)-(S25)) response to the input presented in Figure 4c with $\eta = 30$ nM^{-1} min^{-1} and the remaining parameter same as those used in Figure 4d. As can be seen, the behaviour of both BioSD-II and BioSD-III is identical to that of BioSD-I depicted in the main text. As a result, the conclusions drawn with respect to the latter circuit are valid for the other designs as well.
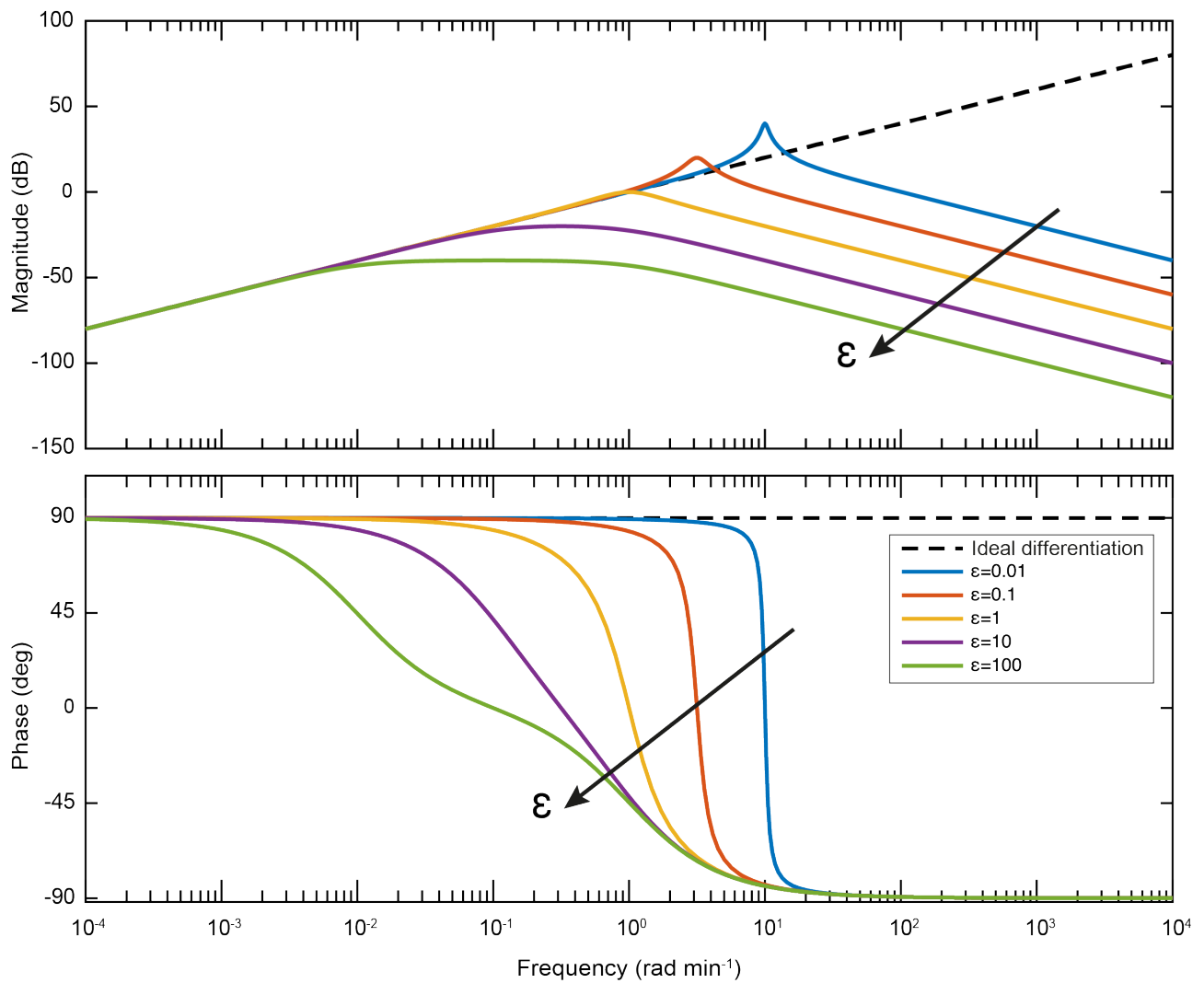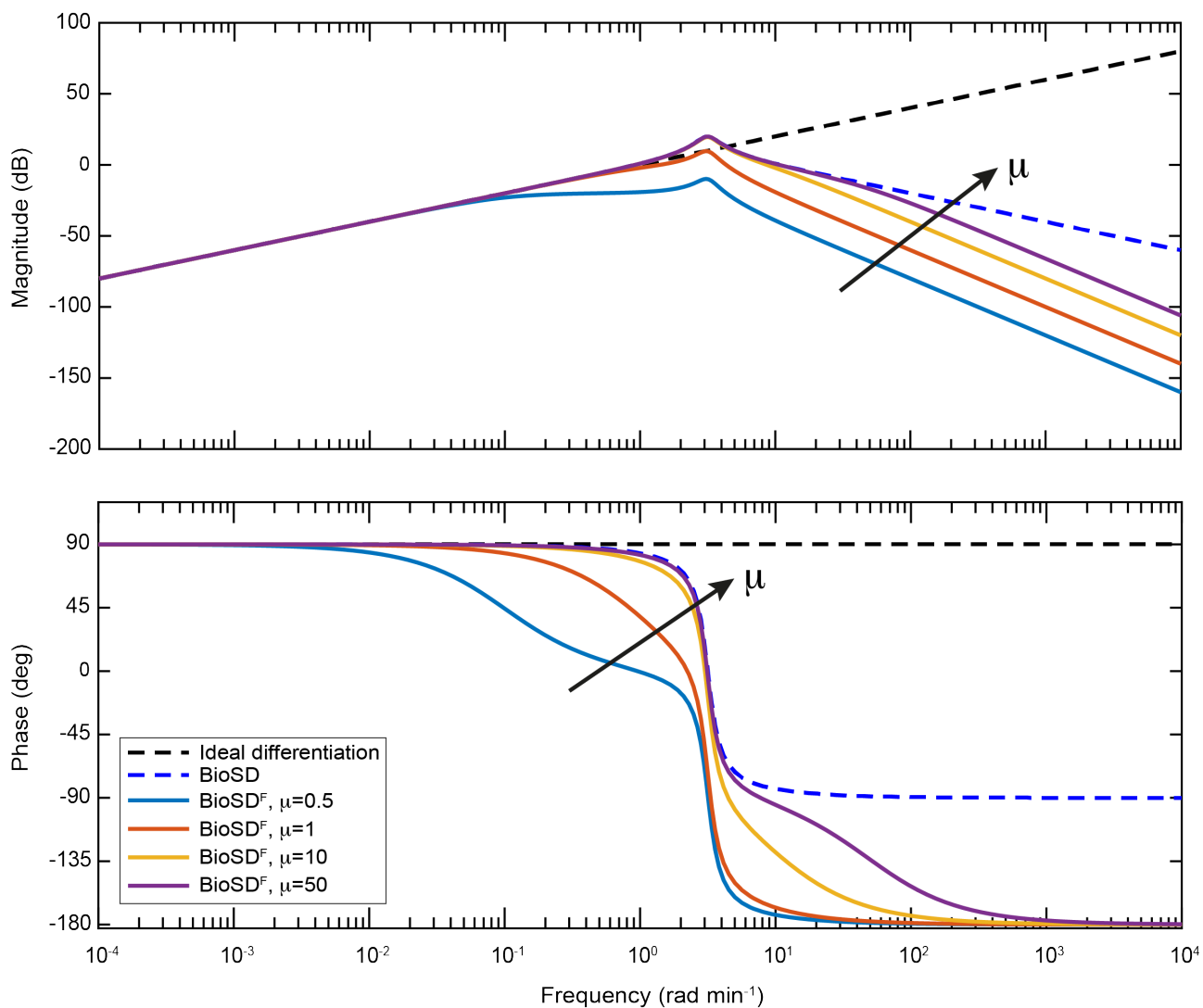
Figure S3: Frequency response analysis of Biomolecular Signal Differentiators. Related to STAR Methods.
Bode plot of a BioSD differentiator (Equation (S62)). The magnitude and the phase of its transfer function are depicted for different values of ε via distinct colours. The case of ideal differentiation corresponds to $\varepsilon = 0$ and the direction in which the latter increases indicated by an arrow..
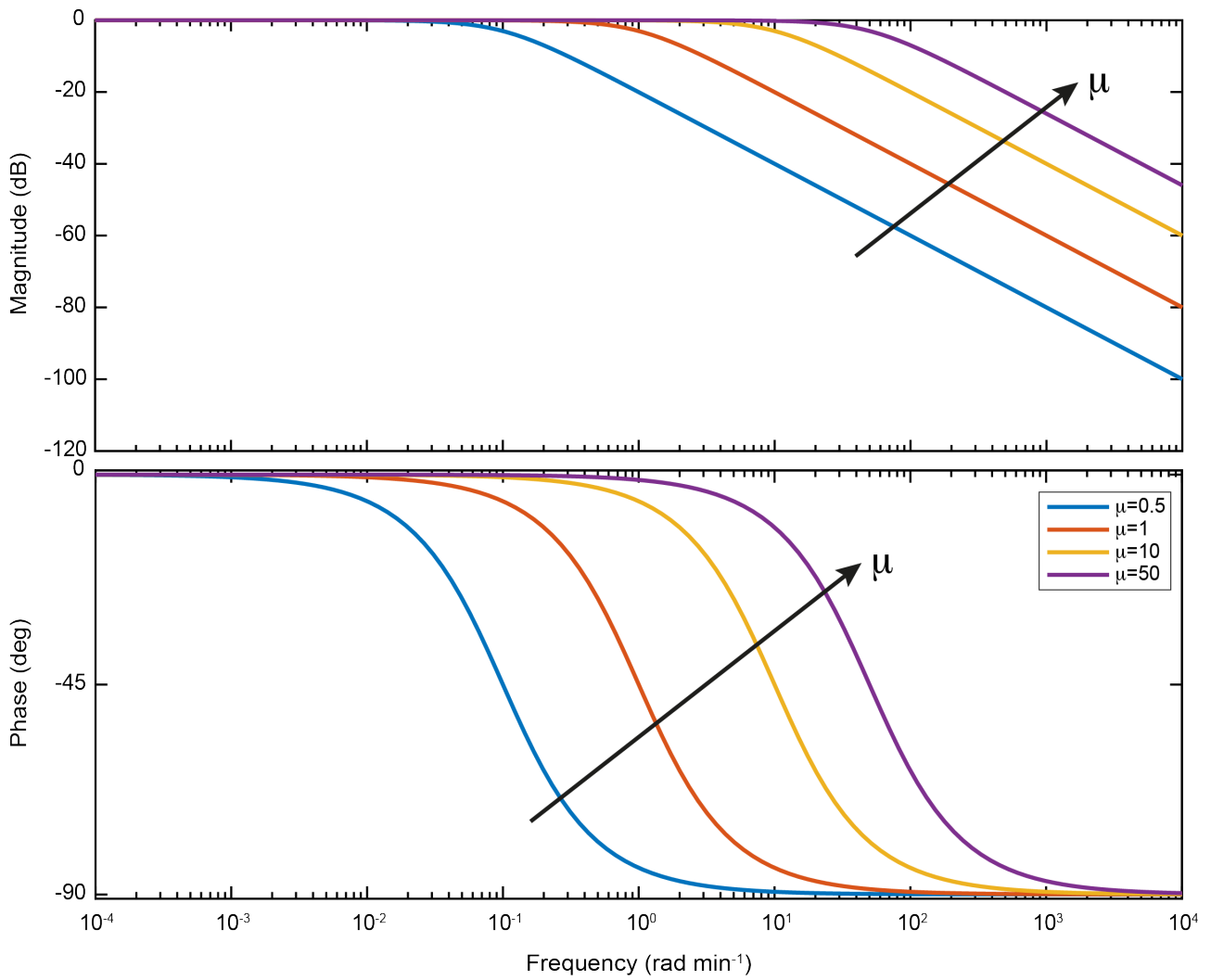
Figure S4: Frequency response analysis of the subsystem that receives the input signal U. Related to STAR Methods.

Bode diagram of the filter module described by Equation (S69). The magnitude and and the phase lag of its frequency response for different values of μ are shown in different colours where $\mu = \mu_1 = \mu_2$. The increasing direction of μ indicated by an arrow.

Figure S5: Frequency response analysis of Biomolecular Signal Differentiators^F. Related to STAR Methods.
Bode diagram depicting the magnitude and phase shift regarding the frequency response of a BioSD^F differentiator (Equation (S71)) with $\varepsilon = 0.1$. We consider different values of μ, where $\mu = \mu_1 = \mu_2$, that correspond to solid lines of different colours while the increasing direction of μ indicated by an arrow. For comparison purposes, we also depict the Bode plot (magnitude and phase) of a BioSD differentiator (Equation (S62)) with $\varepsilon = 0.1$ and the one of an ideal differentiator which are represented by blue and black dashed lines, respectively.

A
B

Derivative of the input signal *dU/dt*
Output signal *X*
Zero-level
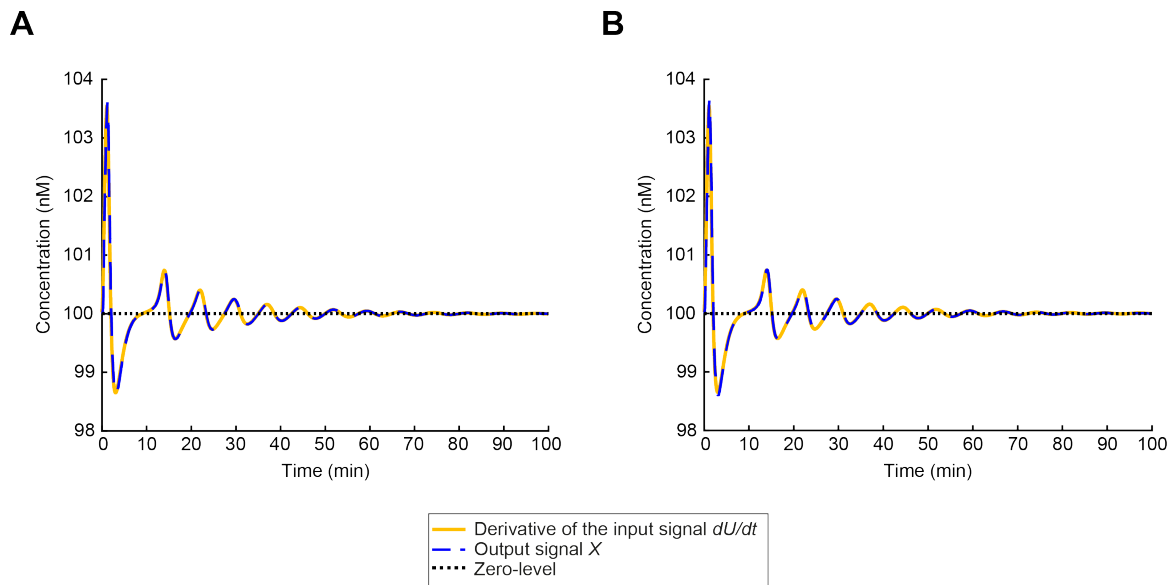
Figure S6: Sensing the rate-of-change of a synthetic regulatory biomolecular network through the proposed (experimental) circuit of Biomolecular Signal Differentiator-III. Related to STAR Methods.
**a** Simulation of the circuit given by Equations (S72)-(S77) using the input presented in Figure 3c and the following parameters: $k_{in} = 100$ min^{-1}, $k_3 = b = 100$ nM min^{-1}, $k_1 = k_{1a} = k_{1b} = 1$ nM^{-1} min^{-1}, $k_2 = 1$ min^{-1}, $\eta = \eta_a = 30$ nM^{-1} min^{-1}, $\delta = \delta_a = 0.5$ min^{-1} (this scenario corresponds to the simulation depicted in Figure S1b). **b** We repeat the simulation in **a** with the values of k_{1a}, k_{1b}, η_a, δ_a set to 1.5 (increase by 50%), 1.25 (increase by 25%), 45 (increase by 50%), 0.75 (increase by 50%), respectively.
It is evident that in both **a** (ideal case) and **b** the output, *X*, of the differentiator is an accurate replica of the derivative of input *U* - the loss of accuracy in **b** is negligible.
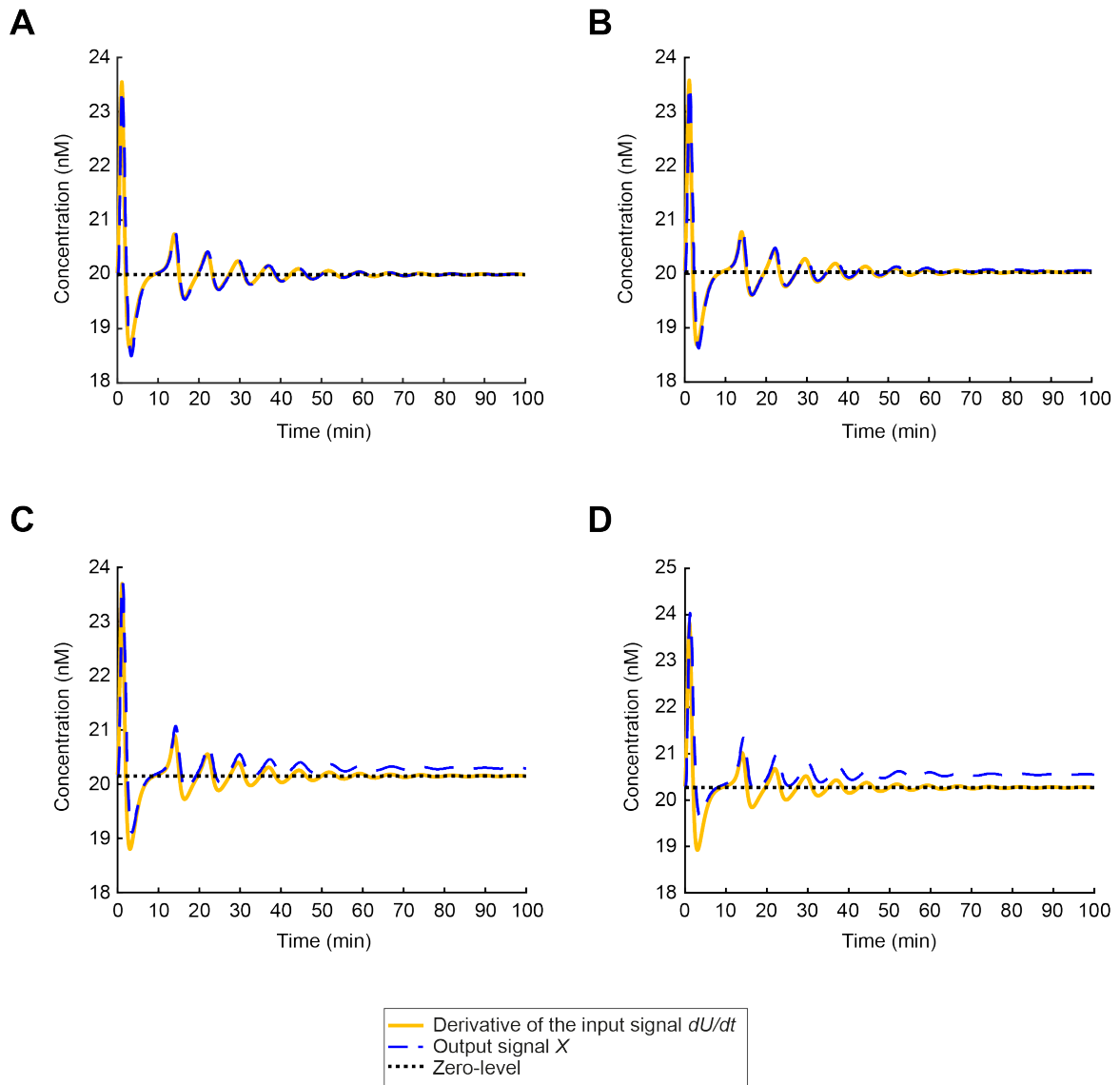
Figure S7: Sensing the rate-of-change of a synthetic regulatory biomolecular network through a more realistic model of Biomolecular Signal Differentiator-II. Related to STAR Methods.

a Simulation of the system given by Equations (S78)-(S80) using the input presented in Figure 3c and the parameters of Table S1 (no dilution). **b** Simulation of the system given by Equations (S84)-(S86) using the input presented in Figure 3c and the parameters of Table S1. **c** The simulation in **b** is repeated with a five times larger dilution rate, i.e. $\gamma = 0.14$ min^{-1}. **d** The simulation in **b** is repeated with a ten times larger dilution rate, i.e. $\gamma = 0.28$ min^{-1}.

In all the simulations we assume that the value of k_{in} is equal to the value of the quantity $k_1 k_3$.
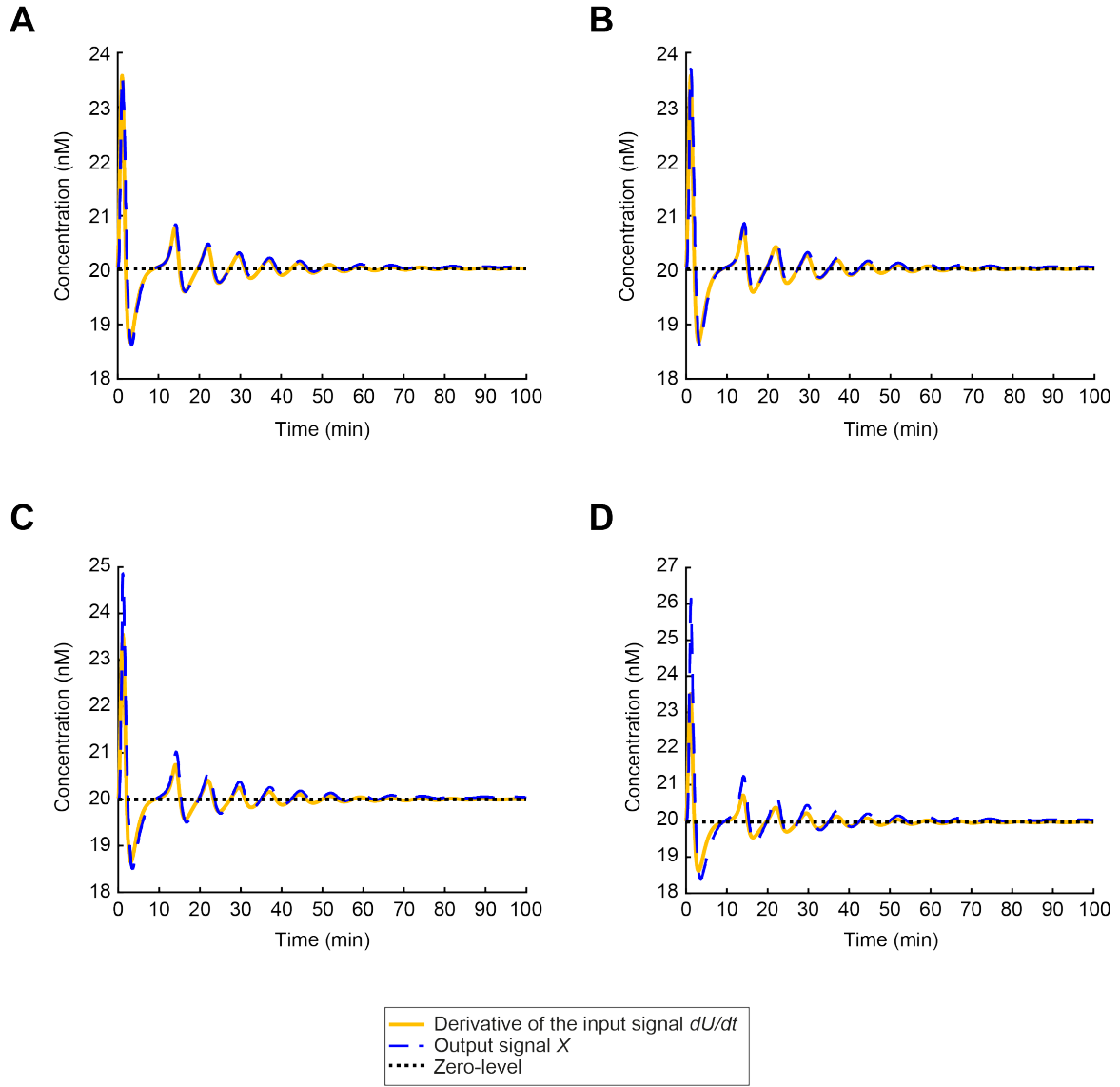
Figure S8: Sensing the rate-of-change of a synthetic regulatory biomolecular network through a more realistic model of Biomolecular Signal Differentiator-II with a lower annihilation rate, η, than the one of Table S1. Related to STAR Methods.

Simulation of the system given by Equations (S84)-(S86) using the input presented in Figure 3c and **a** $\eta = 10\frac{\beta_1^2}{k_3}$, **b** $\eta = 5\frac{\beta_1^2}{k_3}$, **c** $\eta = \frac{\beta_1^2}{k_3}$, **d** $\eta = 0.5\frac{\beta_1^2}{k_3}$ which correspond to 141.8 nM^{-1} min^{-1}, 70.9 nM^{-1} min^{-1}, 14.18 nM^{-1} min^{-1} and 7.09 nM^{-1} min^{-1}, respectively (see STAR Methods **Modelling a more realistic case of Biomolecular Signal Differentiator-II**). In addition, the value of k_{in} is assumed to be equal to the value of the quantity $k_1 k_3$ while rest of the parameters are in accordance with Table S1.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Biomolecular mechanisms for signal differentiation
Publication Status	Published
Publication Details	Alexis, E., Schulte, C. C., Cardelli, L., & Papachristodoulou, A. (2021). Biomolecular mechanisms for signal differentiation. *Iscience*, *24*(12), 103462.

Student Confirmation

Student Name:	Emmanouil Alexis		
Contribution to the Paper	Conceptualization and methodology, Formal analysis and Software, Writing		
Signature		Date	14th February 2023

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Professor Antonis Papachristodoulou		
Supervisor comments. The candidate made a substantial contribution to the publication. The description above is accurate.		
Signature	Date	16 March 2023

This completed form should be included in the thesis, at the end of the relevant chapter.

Chapter 4

On the design of a PID bio-controller with set point weighting and filtered derivative action

On the Design of a PID Bio-controller with Set Point Weighting and Filtered Derivative Action

Emmanouil Alexis, Luca Cardelli, Antonis Papachristodoulou, *Fellow, IEEE*

Abstract— **Effective and robust regulation of biomolecular processes is crucial for designing reliable synthetic bio-devices functioning in uncertain and constantly changing biological environments. Proportional-Integral-Derivative (PID) controllers are undeniably the most common way of implementing feedback control in modern technological applications. Here, we introduce a highly tunable PID bio-controller with *set point weighting* and filtered derivative action presented as a chemical reaction network with mass action kinetics. To demonstrate its effectiveness, we apply our PID scheme on a simple biological process of two mutually activated species, one of which is assumed to be the output of interest. To highlight its performance advantages we compare it to PI regulation using numerical simulations in both the deterministic and stochastic setting.**

Index Terms— **PID control, biomolecular systems, synthetic biology**

I. INTRODUCTION

SYNTHETIC Biology aims to engineer biomolecular systems with novel and useful functionalities in order to tackle a long list of pressing, real-world problems [1]–[4]. One of the main challenges of building synthetic bio-devices operating in the uncertain cellular environment is achieving a reliable and predictable behaviour. Feedback control theory provides a large variety of tools that have proven to be of fundamental importance in regulating such devices, optimizing their function and rendering them robust to disturbances [5]– [10].

Proportional - Integral - Derivative (PID) feedback controllers are regarded as the workhorses of control engineering [11], [12]. They are often called "three - term" controllers due to their triple control action accounting for the past, present and future. More specifically, integral control (I-term) accounts for the history of the error between the set point (desired target value) and the output of interest by accumulating it over time. An important characteristic of the I-term is its ability to eliminate the steady-state error, provided that the feedback system is stable. The present is represented by the P-term which produces a control signal proportional to the current value of the error. Lastly, derivative control (D-term) provides

anticipatory action by estimating future values of the error via linear extrapolation.

Because of the pervasiveness of PID control in technological applications, the biomolecular implementation of PID controllers has seen great interest in Synthetic Biology and several successful research efforts. Notably, the authors in [13] present a hierarchical library of nonlinear PID controllers consisting of up to four biomolecular species with a first-order low-pass filter accompanying some or all the three control terms (P-, I- and D-term). The PID architecture proposed in [14] exploits different variations of Michaelis-Menten functions. Furthermore, the PID designs studied in [15], [16] use the so-called dual rail encoding [17], by which a signal is decomposed into two non-negative components and, thus, both positive and negative signals can be represented via biomolecular species. Lastly, [18] analyzes the noise suppression properties of individual proportional, integral and derivative controllers tailored to gene expression.

In this paper, we introduce an alternative biomolecular network functioning as a PID controller around the nominal operation of the resulting closed-loop system. This local approach is also adopted in [13]. The biomolecular interactions involved are defined by general chemical reaction networks (CRNs) based purely on mass action kinetics [19] and without using dual rail encoding. At the same time, our bio-controller acts solely on the target species (output of interest) without considering other species or reactions of the network to be controlled (open-loop system) as happens, for instance, in [13]. To achieve enhanced dynamic performance we adopt a special form of *set point weighting* commonly used in technological applications and we accompany derivative control with the strong filtering action of a second-order low-pass filter. Moreover, our PID configuration includes six controller species that allow us to build each of the P-, I-, D- terms almost independently providing significant tuning flexibility regarding controller gains, *set point weights* and filtering. Finally, the proposed PID configuration can be used for controlling any open-loop biological process assuming the existence of a biologically meaningful equilibrium and asymptotic stability for the resulting closed-loop system. Here we only consider scenarios where this condition holds.

Section II presents some background concepts on PID control, biomolecular interactions, modelling tools and essential biomolecular motifs. Section III analyzes the main characteristics of the proposed PID bio-controller. Subsequently, an application example including a comparison between PI and PID control in the deterministic and stochastic setting

is provided in Sections IV and V, respectively. Section VI concludes our work and discusses future research directions.

II. BACKGROUND

Here we first outline key properties of PID control, review principles of biomolecular modelling and then present two important biomolecular motifs that implement integral and derivative action.

A. Key points on PID control

Here we briefly present some important features of PID control action [11], [12] based on which our PID bio-controller has been developed.

First, recall that the "traditional", ideal PID algorithm (Fig. 1) is described as follows:

$$u(t) = k_p e(t) \ + \ k_i \int_0^t e(\tau)d\tau \ + \ k_d \frac{de(t)}{dt} \qquad (1)$$

where $u(t)$, $e(t)$ represent the control input signal and the control error, respectively. The latter is defined as $e(t) = y_{sp} - y(t)$, where y_{sp} is the set point and y is the process output.

A major problem of the (ideal) derivative action in (1) is its sensitivity to high-frequency signal components. This can lead to excessively high gains and, by extension, large variations in terms of the control signal. A common strategy to overcome this obstacle is to accompany the derivative term with a low-pass filter.

Another challenge is *derivative kick*: When the set point is constant, the derivative of the error in (1) becomes $\frac{de(t)}{dt} = -\frac{dy(t)}{dt}$ since $\frac{dy_{sp}}{dt} = 0$. Abrupt changes of the set point (when the set point is adjusted) make the aforementioned derivative very large causing undesirable transients in the control signal (derivative kick). To avoid this, we replace $e(t)$ with $-y(t)$ in the derivative term of (1).

The behaviour of the controller can be further improved by modifying appropriately the error quantity on which the proportional action acts. To this end, we consider an alternative PID control law with *set point weighting*:

$$u(t) = k_p(\lambda y_{sp} - y) \ + \ k_i \int_0^t e(\tau)d\tau \ - \ k_d \frac{dy(t)}{dt} \qquad (2)$$

A PID controller based on (2) with $\lambda = 1$ and $\lambda = 0$ is often referred to as a PI-D and I-PD controller, respectively. Finally, the error quantity in the integral term needs to remain unchanged in order for the error to go to zero at steady-state.

B. Biomolecular interactions and modelling

In Fig. 2(a) we present all different types of biomolecular interactions as well as their graphical notation used in this paper. These interactions can be divided into two main categories: non-catalytic reactions where the reactants are consumed in order for products to be formed and catalytic ones where species facilitate production/inhibition processes without being consumed. For deterministic analysis of the biomolecular
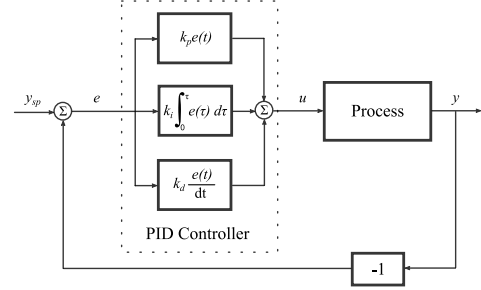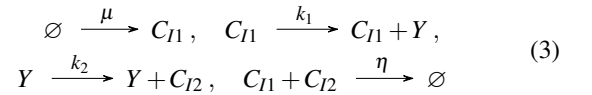

Fig. 1: Ideal PID control of a process based on error feedback [11], [12].

networks in this paper (Sections II-C, III, IV) we use Ordinary Differential Equations (ODEs) models based on the law of mass action [19]. For stochastic analysis (Section V) we use the Linear Noise Approximation (LNA) of the Chemical Master Equation (CME) [21], [22]. LNA stochastic simulations are performed using [21] which provides analytical results that can be exploited in further work.

C. Two important biomolecular motifs

We now review two basic biomolecular motifs from the literature, which are constituent elements of our PID architecture under appropriate modifications.

Fig. 2(b) shows the *antithetic motif* introduced in [23], which is realized by controller species C_{I1}, C_{I2}, regulating a target (output) species, Y, which is part of an arbitrary biological process - "cloud" network. This mechanism can achieve *robust perfect adaptation* (RPA) through integral feedback control. To see this, focusing on the controller species, we have the CRN:

$$\varnothing \xrightarrow{\mu} C_{I1}, \quad C_{I1} \xrightarrow{k_1} C_{I1} + Y, $$
$$Y \xrightarrow{k_2} Y + C_{I2}, \quad C_{I1} + C_{I2} \xrightarrow{\eta} \varnothing \qquad (3)$$

which can be modelled by the following set of ODEs:

$$\dot{C}_{I1} = \mu - \eta C_{I1} C_{I2} \qquad (4a)$$
$$\dot{C}_{I2} = k_2 Y - \eta C_{I1} C_{I2} \qquad (4b)$$

where μ, k_1, k_2, $\eta \in \mathbb{R}_+$.

Integration is carried out by a hidden "memory" variable. Subtract (4a) - (4b) and integrate to obtain:

$$(C_{I1} - C_{I2})(t) = k_2 \int_0^t \left(\frac{\mu}{k_2} - Y(\tau) \right) d\tau$$

Assuming closed-loop stability, at the steady state:

$$Y^* = \frac{\mu}{k_2}$$

where the $*$ notation denotes the steady state of a variable.

Fig. 2(c) shows a topology known as *BioSD-III* which we introduced in [20]. This topology can function as a signal differentiator module around its nominal operation. In particular, it receives an input signal, U, and calculates its filtered
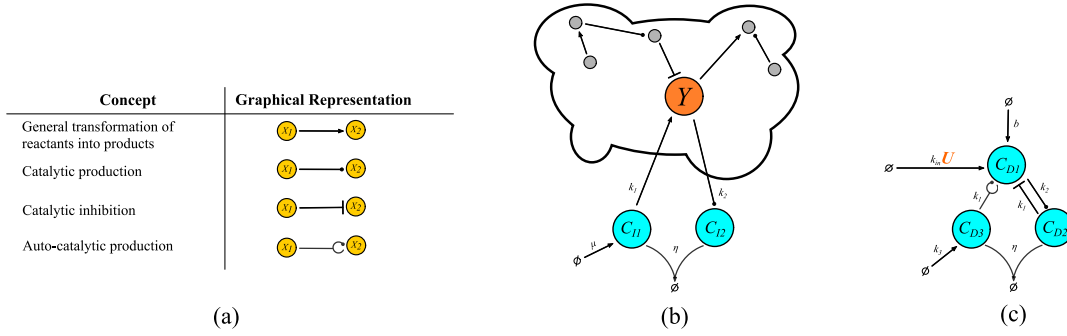
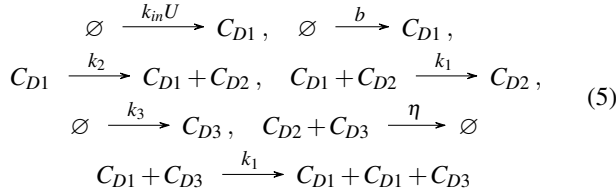Fig. 2: (a) Table with the different types of biomolecular interactions adopted from our previous work [20]. (b) *Antithetic integral controller regulating a target species which is part of an arbitrary biological process - "cloud" network (CRN (3)). (c) BioSD-III differentiator module (CRN (5)).*

derivative in the output species C_{D1}. Contrary to the *antithetic motif*, this is not a regulatory topology but, as shown in the following section, it can be used for developing derivative control with respect to a target species.

The CRN for *BioSD-III* consists of the reactions:

$$\varnothing \xrightarrow{k_{in}U} C_{D1}, \quad \varnothing \xrightarrow{b} C_{D1},$$
$$C_{D1} \xrightarrow{k_2} C_{D1}+C_{D2}, \quad C_{D1}+C_{D2} \xrightarrow{k_1} C_{D2}, \quad (5)$$
$$\varnothing \xrightarrow{k_3} C_{D3}, \quad C_{D2}+C_{D3} \xrightarrow{\eta} \varnothing$$
$$C_{D1}+C_{D3} \xrightarrow{k_1} C_{D1}+C_{D1}+C_{D3}$$

where k_{in}, b, k_2, k_1, $\eta \in \mathbb{R}_+$. The degradation rate of C_{D1} considered in [20] is assumed to be zero.

The dynamics of CRN (5) can be modelled as:

$$\dot{C}_{D1} = k_{in}U + b - k_1 C_{D1}C_{D2} + k_1 C_{D1}C_{D3} \quad (6a)$$
$$\dot{C}_{D2} = k_2 C_{D1} - \eta C_{D2}C_{D3} \quad (6b)$$
$$\dot{C}_{D3} = k_3 - \eta C_{D2}C_{D3} \quad (6c)$$

As shown in [20], for any non-negative constant input U^*, we obtain a positive locally exponentially stable steady state (X^*, Z_1^*, Z_2^*).

Through (Jacobian) linearization of (6), we have the local dynamics of *BioSD-III*:

$$\begin{bmatrix} \dot{c}_{D1} \\ \dot{c}_{D2} \\ \dot{c}_{D3} \end{bmatrix} = \begin{bmatrix} -\dfrac{k_2(k_{in}U^*+b)}{k_3} & -\dfrac{k_1 k_3}{k_2} & \dfrac{k_1 k_3}{k_2} \\ k_2 & -\eta C_{D3}^* & -\eta C_{D2}^* \\ 0 & -\eta C_{D3}^* & -\eta C_{D2}^* \end{bmatrix} \begin{bmatrix} c_{D1} \\ c_{D2} \\ c_{D3} \end{bmatrix} + \begin{bmatrix} k_{in} \\ 0 \\ 0 \end{bmatrix} u$$

where variables $u = U - U^*$, $c_{D1} = C_{D1} - C_{D1}^*$, $c_{D2} = C_{D2} - C_{D2}^*$, $c_{D3} = C_{D3} - C_{D3}^*$ represent small perturbations around (X^*, Z_1^*, Z_2^*). The corresponding input/output relation in the Laplace domain is:

$$\tilde{T}(s) = \frac{\tilde{c}_{D1}(s)}{\tilde{u}(s)} = \frac{k_{in}}{k_1 k_3} \frac{s}{\varepsilon(s^2+s)+1} \quad (7)$$

where:

$$\varepsilon = \frac{k_2^2}{k_1 k_3^3}(k_{in}U^*+b)^2 \quad (8)$$

and s is the Laplace variable (complex frequency).

Equation (7) is an ideal signal differentiator multiplied by

a constant gain in series with a second-order low pass filter. The filtering action can be adjusted to meet our performance requirements by appropriately tuning the dimensionless parameter (8). Thus, moving to the time domain, for a given value of (8), there are sufficiently slow input signals yielding:

$$c_{D1} = \frac{k_{in}}{k_1 k_3}\dot{u} \quad (9)$$

The structural complexity of the differentiator module can be reduced by removing the reaction $C_{D1}+C_{D3} \xrightarrow{k_1} C_{D1}+C_{D1}+C_{D3}$ in CRN (5) while the input/output behaviour remains the same [20]. This results in ODE model (6) without the term $+k_1 C_{D1}C_{D3}$ in (6a). However, this simplification comes with the cost of imposing the following constraint:

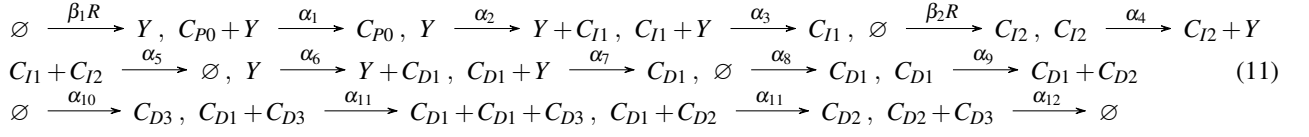$$\eta \gg \frac{k_1^2 k_3^3}{k_2^2(k_{in}U^*+b)^2} \quad (10)$$

Finally, computing the time derivatives of molecular signals as species concentrations constitutes a fundamental difference compared to the differentiators used in [13] where the derivatives in question correspond to reaction rates.

III. STRUCTURE AND BEHAVIOUR OF THE PID ARCHITECTURE

Fig. 3 shows our PID controller regulating a target (output) species, Y, of an abstract "cloud" network. This "cloud" network represents a general biomolecular network with arbitrary number of species/interactions accounting also for potential time delays [19]. The reactions that form the corresponding CRN are given by (11) where $\alpha_i \in \mathbb{R}_+$ with $i \in \mathbb{N}$ and $1 \leq i \leq 12$. R is a non-negative reference signal that can vary over time while β_1, β_2 are non-negative scaling parameters: R, β_1, β_2 can be controlled externally. C_{P0} can be considered as an auxiliary species with constant concentration that catalyzes the degradation of the target species Y. The modified version of the *antithetic motif* with an additional inhibitory reaction as formed by species C_{I1}, C_{I2} has been studied in [24].

A. Achieving PID control

To gain a deeper understanding of the proposed topology (Fig. 3), we study the corresponding dynamics which can be

$$\varnothing \xrightarrow{\beta_1 R} Y, \; C_{P0}+Y \xrightarrow{\alpha_1} C_{P0}, \; Y \xrightarrow{\alpha_2} Y+C_{I1}, \; C_{I1}+Y \xrightarrow{\alpha_3} C_{I1}, \; \varnothing \xrightarrow{\beta_2 R} C_{I2}, \; C_{I2} \xrightarrow{\alpha_4} C_{I2}+Y$$
$$C_{I1}+C_{I2} \xrightarrow{\alpha_5} \varnothing, \; Y \xrightarrow{\alpha_6} Y+C_{D1}, \; C_{D1}+Y \xrightarrow{\alpha_7} C_{D1}, \; \varnothing \xrightarrow{\alpha_8} C_{D1}, \; C_{D1} \xrightarrow{\alpha_9} C_{D1}+C_{D2} \tag{11}$$
$$\varnothing \xrightarrow{\alpha_{10}} C_{D3}, \; C_{D1}+C_{D3} \xrightarrow{\alpha_{11}} C_{D1}+C_{D1}+C_{D3}, \; C_{D1}+C_{D2} \xrightarrow{\alpha_{11}} C_{D2}, \; C_{D2}+C_{D3} \xrightarrow{\alpha_{12}} \varnothing$$

described by the following set of ODEs:

$$\dot{Y} = F + \underbrace{\beta_1 R - \alpha_1 C_{P0} Y + \alpha_4 C_{I1} - \alpha_3 C_{I2} Y - \alpha_7 C_{D1} Y}_{\text{control input signal}} \tag{12a}$$

$$\dot{C}_{I1} = \beta_2 R - \alpha_5 C_{I1} C_{I2} \tag{12b}$$

$$\dot{C}_{I2} = \alpha_2 Y - \alpha_5 C_{I1} C_{I2} \tag{12c}$$

$$\dot{C}_{D1} = \alpha_6 Y + \alpha_8 - \alpha_{11} C_{D1} C_{D2} + \alpha_{11} C_{D1} C_{D3} \tag{12d}$$

$$\dot{C}_{D2} = \alpha_9 C_{D1} - \alpha_{12} C_{D2} C_{D3} \tag{12e}$$

$$\dot{C}_{D3} = \alpha_{10} - \alpha_{12} C_{D2} C_{D3} \tag{12f}$$

where F represents potential interactions associated with the output species Y in the cloud network.

We assume the existence of a (locally) asymptotically stable and biologically meaningful equilibrium for the overall closed loop system for some constant value, R^*, of the reference signal of interest, R. We consider the local behaviour of our bio-controller by adopting coordinate transformations of the form $x = X - X^*$ which denote small perturbations around the equilibrium - X and X^* represent any variable involved in the system under consideration and its corresponding steady state, respectively. Thus, we obtain via (Jacobian) linearization of (12a)-(12f):

$$\dot{y} = f + u_{PID} \tag{13a}$$

$$\dot{c}_{I1} = \beta_2 r - \alpha_5 C_{I2}^* c_{I1} - \alpha_5 C_{I1}^* c_{I2} \tag{13b}$$

$$\dot{c}_{I2} = \alpha_2 y - \alpha_5 C_{I2}^* c_{I1} - \alpha_5 C_{I1}^* c_{I2} \tag{13c}$$

$$\dot{c}_{D1} = \alpha_6 y - \alpha_{11}(C_{D2}^* - C_{D3}^*)c_{D1} - \alpha_{11} C_{D1}^* c_{D2} + \alpha_{11} C_{D1}^* c_{D3} \tag{13d}$$

$$\dot{c}_{D2} = \alpha_9 c_{D1} - \alpha_{12} C_{D3}^* c_{D2} - \alpha_{12} C_{D2}^* c_{D3} \tag{13e}$$

$$\dot{c}_{D3} = -\alpha_{12} C_{D3}^* c_{D2} - \alpha_{12} C_{D2}^* c_{D3} \tag{13f}$$

where f is the "linearized version" of F around the equilibrium and the control input signal is given by:

$$u_{PID} = \beta_1 r - (\alpha_1 C_{P0} + \alpha_3 C_{I2}^* + \alpha_7 C_{D1}^*)Y + \alpha_4 c_{I1} - \alpha_3 Y^* c_{I2} - \alpha_7 Y^* c_{D1} \tag{14}$$

From Equations (12b)-(12c) we get at the steady state :

$$Y^* = \frac{\beta_2 R^*}{\alpha_2}$$

Moreover, species C_{D1}, C_{D2}, C_{D3} form a *BioSD-III* module with $u = y$ (see (12d)-(12f)). Thus, taking into account (7), (8), (10), we have for the input/output relation in the Laplace domain:

$$\tilde{T}(s) = \frac{\tilde{c}_{D1}(s)}{\tilde{y}(s)} = \frac{\alpha_6}{\alpha_{10}\alpha_{11}} \frac{s}{\varepsilon(s^2+s)+1}$$

where:

$$\varepsilon = \frac{\alpha_9^2}{\alpha_{10}^3 \alpha_{11}}(\alpha_6 Y^* + \alpha_8)^2$$

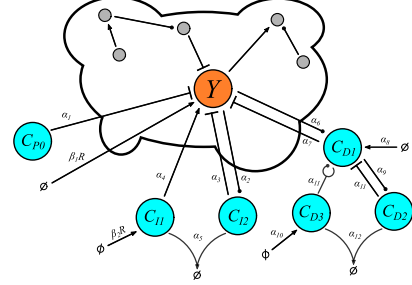Fig. 3: The proposed PID bio-controller regulating a target species which is part of an arbitrary biological process - "cloud" network (CRN (11)).

while the parameter constraint for the simplified *BioSD-II* module becomes:

$$\alpha_{12} \gg \frac{\alpha_{10}^3 \alpha_{11}^2}{\alpha_9^2(\alpha_6 Y^* + \alpha_8)^2} \tag{15}$$

Setting now

$$\delta = \frac{\beta_2(\alpha_1 C_{P0} + \alpha_3 C_{I2}^* + \alpha_7 C_{D1}^*)}{\alpha_2}$$

and assuming $\beta_2 R^* = \dfrac{\alpha_2 \alpha_4}{\alpha_3}$ (16), the control input signal (14) can be rewritten as:

$$u_{PID} = k_p(\lambda y_{sp} - y) + k_i \int_0^t (y_{sp} - y)\, d\tau - k_d c_{D1} \tag{17}$$

with:

$$y_{sp} = \frac{\beta_2 r}{\alpha_2}, \; \lambda = \frac{\beta_1}{\delta}, \; k_p = \frac{\alpha_2 \delta}{\beta_2}, \; k_i = \frac{\alpha_4}{\alpha_2}, \; k_d = \frac{\alpha_4 \alpha_7}{\alpha_3}$$

In addition, taking into account (9), we have:

$$c_{D1} = \frac{\alpha_6}{\alpha_{10}\alpha_{11}}\dot{y}$$

assuming y is sufficiently slow.

Equation (17) describes a PID control law with *set point weighting* and filtered derivative action. Our architecture offers considerable tunability since the controller gains, the reference signal (including the ratio in (16)), the set point, the *set point weight* regarding proportional control as well as the filtering action regarding derivative control can be tuned separately as desired. In addition, setting $\beta_1 = 1$ or $\beta_1 = 0$ leads to a PI-D or I-PD control law, respectively.
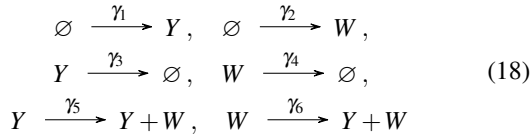
Local closed-loop stability can be assessed by studying the Jacobian matrix resulting from (13a)-(12f). If this matrix is Hurwitz, i.e. the real parts of its eigenvalues are strictly negative, then the equilibrium in question is a locally asymptotically stable equilibrium for the nonlinear system (12a)-(12f).

Necessary and sufficient conditions can be found using the Routh-Hurwitz criterion. We can find suitable parameters by taking into account any parameter constraints stemming from our performance standards or the experimental implementation of interest as well as using the rich toolkit of PID tuning techniques [11], [12].

IV. REGULATING A SPECIFIC BIOLOGICAL PROCESS

In this section we investigate the properties of our PID controller on a specific biological process. In particular, we replace the abstract cloud network of Fig. 3 with a biological process of two mutually activated species, Y and W, with the first species being the target species on which we apply PID control (Fig. 4(a)). This process is based on a positive feedback loop which is a very common concept in biological systems [25], [26].

The open-loop process under consideration consists of the following reactions:

$$\varnothing \xrightarrow{\gamma_1} Y, \quad \varnothing \xrightarrow{\gamma_2} W,$$
$$Y \xrightarrow{\gamma_3} \varnothing, \quad W \xrightarrow{\gamma_4} \varnothing, \quad (18)$$
$$Y \xrightarrow{\gamma_5} Y + W, \quad W \xrightarrow{\gamma_6} Y + W$$

Taking into account CRNs (11) and (18), the dynamics of the resulting closed-loop system can be modelled as:

$$\dot{Y} = \gamma_1 - \gamma_3 Y_1 + \gamma_6 Y_2 + \beta_1 R - \alpha_1 C_{P0} Y$$
$$+ \alpha_4 C_{I1} - \alpha_3 C_{I2} Y - \alpha_7 C_{D1} Y \quad (19a)$$

$$\dot{W} = \gamma_2 - \gamma_4 Y_2 + \gamma_5 Y_1 \quad (19b)$$
$$\dot{C}_{I1} = \beta_2 R - \alpha_5 C_{I1} C_{I2} \quad (19c)$$
$$\dot{C}_{I2} = \alpha_2 Y - \alpha_5 C_{I1} C_{I2} \quad (19d)$$
$$\dot{C}_{D1} = \alpha_6 Y + \alpha_8 - \alpha_{11} C_{D1} C_{D2} + \alpha_{11} C_{D1} C_{D3} \quad (19e)$$
$$\dot{C}_{D2} = \alpha_9 C_{D1} - \alpha_{12} C_{D2} C_{D3} \quad (19f)$$
$$\dot{C}_{D3} = \alpha_{10} - \alpha_{12} C_{D2} C_{D3} \quad (19g)$$

In Fig. 4(b) we present the response of output species, Y, using PI and PID control, respectively. In both scenaria identical integral action takes place and, thus, Y converges to the same value. Nevertheless, the transient response in the first case shows a significant overshoot and oscillations which are eliminated due to the anticipatory action of derivative control in the second case. Moreover, as can be seen, the output response remains the same regardless of the signal differentiator module used in the PID bio-controller.

V. STOCHASTIC SIMULATIONS

The random nature of biomolecular reactions makes biological systems inherently stochastic [19], [28], [29]. The deterministic approach we have followed so far can offer a satisfactory insight into the average biological behaviour when biomolecular populations are sufficiently large. However, this may not be always the case and, as a consequence, analysis of the probabilistic effects may be needed. We focus here on the stochastic evolution of the closed-loop system shown in Fig. 4 over time using the Linear Noise Approximation (LNA). Fig. 5 shows the time evolution of the standard deviation,
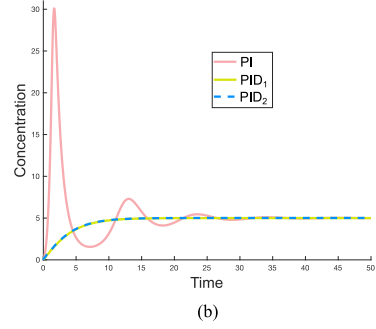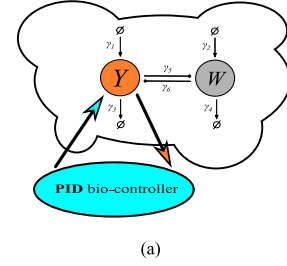

(a)

(b)

Fig. 4: (a) The proposed PID bio-controller regulates a target species (Y) of a network consisting of two mutually activated species (CRNs (11), (18)). (b) Simulated response of the target species (Y) regarding the topology in (a) described by ODE model (19). For PI case, we use only (19a)-(19d) with the following parameter values: $\gamma_1 = 0.5$, $\gamma_2 = 1$, $\gamma_3 = 1$, $\gamma_4 = 2$, $\gamma_5 = 4$, $\gamma_6 = 4$, $\beta_1 R = 1$, $\beta_2 R = 5$, $\alpha_1 C_{P0} = 0.2$, $\alpha_2 = 1$, $\alpha_3 = 0.4$, $\alpha_4 = 2$, $\alpha_5 = 10$. Additionally, the term $-\alpha_7 C_{D1} Y$ in (19a) is removed since there is no derivative action. For PID$_1$, derivative control takes place through *BioSD-III*. Here we use (19a)-(19g) with the following parameter values: $\alpha_6 = 100$, $\alpha_7 = 0.15$, $\alpha_8 = 100$, $\alpha_9 = 1$, $\alpha_{10} = 100$, $\alpha_{11} = 1$, $\alpha_{12} = 10$ with the rest of the parameter values the same as in PI case. For PID$_2$ we replace *BioSD-III* with *BioSD-II* which results in ODE model (19) without the term $+\alpha_{11} C_{D1} C_{D3}$ in (19e). We also use the same parameter values as in PID$_1$ except for $\alpha_{12} = 500$ so that condition (15) is satisfied. The simulations depicted in this figure were performed in MATLAB (Mathworks).

denoted here as σ, with respect to the output species for both PI and PID control. As can be seen, PID control leads to a considerably smaller σ compared to PI control at the steady state, demonstrating the noise reduction capability of derivative control through BioSD modules. Attenuation of stochastic fluctuations through derivative action has been also demonstrated in [13], [18].

VI. CONCLUSION

This paper proposed a highly tunable CRN architecture capable of applying PID feedback control locally using *set point weights* and derivative control filtering. Notable characteristics of our design are the "antithetic integration" and "BioSD signal differentiation". For the latter, we consider two differentiator modules of different structure but identical input/output behaviour. Proportional control is realized through a special birth-death process to which the integral and derivative parts also contribute. To demonstrate the performance benefits
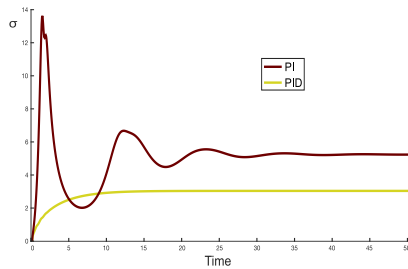
Fig. 5: Time evolution of the standard deviation, σ of the target species Y of the closed-loop system shown in Fig. 4. The models and the parameter sets for PI and PID cases correspond to the cases of PI and PID_1 of Fig. 4. Also, the case of PID_2 results in identical behaviour to PID_1. The simulations depicted in this figure were performed in Kaemika [27] using LNA.

of our PID control strategy, we apply it to an (open-loop) process of two mutually activated species and compare it to PI regulation. We show through deterministic simulations that the concentration of the output species of interest exhibits a significantly improved transient response with PID compared to PI control. At the same time, using LNA we show that the addition of *BioSD* derivative action can reduce the standard deviation at the steady state.

In the future it would be interesting to study the stochastic behaviour of the proposed controller using other, more accurate methods [30], and compare the results with LNA. Moreover, as our PID bio-controller is experimentally realizable, it would be interesting to implement it *in vitro* via molecular programming. In particular, our topology relies purely on mass action kinetics and, thus, can be translated into a DNA strand displacement system [31]–[33]. Finally, the analytical results presented in this work have been obtained by applying linear perturbation analysis since our PID controller is supposed to work around the nominal operation and their accuracy is supported by simulations of the actual nonlinear systems under consideration. A possible extension to this would be the analysis regarding the non-local behaviour of our controller (large signal analysis) and the comparison with the results herein.

REFERENCES

[1] G. M. Church, M. B. Elowitz, C. D. Smolke, C. A. Voigt, and R. Weiss, "Realizing the potential of synthetic biology," *Nature Reviews Molecular Cell Biology*, vol. 15, no. 4, pp. 289–294, 2014.

[2] M. P. McNerney, K. E. Doiron, T. L. Ng, T. Z. Chang, and P. A. Silver, "Theranostic cells: emerging clinical applications of synthetic biology," *Nature Reviews Genetics*, vol. 22, no. 11, pp. 730–746, 2021.

[3] F. Meng and T. Ellis, "The second decade of synthetic biology: 2010–2020," *Nature Communications*, vol. 11, no. 1, pp. 1–4, 2020.

[4] C. A. Voigt, "Synthetic biology 2020–2030: six commercially-available products that are changing our world," *Nature Communications*, vol. 11, no. 1, pp. 1–6, 2020.

[5] H. Steel, G. Lillacci, M. Khammash, and A. Papachristodoulou, "Challenges at the interface of control engineering and synthetic biology," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 1014–1023.

[6] D. Del Vecchio, Y. Qian, R. M. Murray, and E. D. Sontag, "Future systems and control research in synthetic biology," *Annual Reviews in Control*, vol. 45, pp. 5–17, 2018.

[7] T. Frei and M. Khammash, "Adaptive circuits in synthetic biology," *Current Opinion in Systems Biology*, vol. 28, p. 100399, 2021.

[8] C. D. McBride, T. W. Grunberg, and D. Del Vecchio, "Design of genetic circuits that are robust to resource competition," *Current Opinion in Systems Biology*, vol. 28, p. 100357, 2021.

[9] A. Sootla, N. Delalez, E. Alexis, A. Norman, H. Steel, G. Wadhams, and A. Papachristodoulou, "Dichotomous feedback: A signal sequestration-based feedback mechanism for biocontroller design," *Journal of the Royal Society Interface*, vol. 19, no. 189, p. 20210737, 2022.

[10] C. Kelly, A. Harris, H. Steel, E. Hancock, J. Heap, and A. Papachristodoulou, "Synthetic negative feedback circuits using engineered small RNAs," *Nucleic Acids Research*, vol. 46, no. 18, pp. 9875–9889, 2018.

[11] K. J. Åström and R. M. Murray, *Feedback systems: an introduction for scientists and engineers*. Princeton University Press, 2021.

[12] Åström, K. Johan, and T. Hägglund, *Advanced PID control*. ISA-The Instrumentation, Systems, and Automation Society Research Triangle Park, 2006, vol. 461.

[13] M. Filo, S. Kumar, and M. Khammash, "A hierarchy of biomolecular proportional-integral-derivative feedback controllers for robust perfect adaptation and dynamic performance," *Nature Communications*, vol. 13, no. 1, pp. 1–19, 2022.

[14] M. Chevalier, M. Gómez-Schiavon, A. H. Ng, and H. El-Samad, "Design and analysis of a proportional-integral-derivative controller with biological molecules," *Cell Systems*, vol. 9, no. 4, pp. 338–353, 2019.

[15] M. Whitby, L. Cardelli, M. Kwiatkowska, L. Laurenti, M. Tribastone, and M. Tschaikowski, "PID control of biochemical reaction networks," *IEEE Transactions on Automatic Control*, 2021.

[16] N. M. Paulino, M. Foo, J. Kim, and D. G. Bates, "PID and state feedback controllers using dna strand displacement reactions," *IEEE Control Systems Letters*, vol. 3, no. 4, pp. 805–810, 2019.

[17] K. Oishi and E. Klavins, "Biomolecular implementation of linear I/O systems," *IET Systems Biology*, vol. 5, no. 4, pp. 252–260, 2011.

[18] S. Modi, S. Dey, and A. Singh, "Noise suppression in stochastic genetic circuits using PID controllers," *PLoS Computational Biology*, vol. 17, no. 7, p. e1009249, 2021.

[19] D. Del Vecchio and R. M. Murray, *Biomolecular feedback systems*. Princeton University Press Princeton, NJ, 2015.

[20] E. Alexis, C. C. M. Schulte, L. Cardelli, and A. Papachristodoulou, "Biomolecular mechanisms for signal differentiation," *iScience*, vol. 24, no. 12, 2021.

[21] L. Cardelli, M. Kwiatkowska, and L. Laurenti, "Stochastic analysis of chemical reaction networks using linear noise approximation," *Biosystems*, vol. 149, pp. 26–33, 2016.

[22] N. G. Van Kampen, *Stochastic processes in physics and chemistry*. Elsevier, 1992, vol. 1.

[23] C. Briat, A. Gupta, and M. Khammash, "Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks," *Cell Systems*, vol. 2, no. 1, pp. 15–26, 2016.

[24] A. Gupta and M. Khammash, "An antithetic integral rein controller for bio-molecular networks," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 2808–2813.

[25] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*. CRC Press, 2019.

[26] E. Alexis, C. C. M. Schulte, L. Cardelli, and A. Papachristodoulou, "Regulation strategies for two-output biomolecular networks," *bioRxiv*, 2022.

[27] L. Cardelli, "Kaemika app: Integrating protocols and chemical simulation," in *International Conference on Computational Methods in Systems Biology*. Springer, 2020, pp. 373–379.

[28] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins, "Stochasticity in gene expression: from theories to phenotypes," *Nature Reviews Genetics*, vol. 6, no. 6, pp. 451–464, 2005.

[29] A. Eldar and M. B. Elowitz, "Functional roles for noise in genetic circuits," *Nature*, vol. 467, no. 7312, pp. 167–173, 2010.

[30] D. F. Anderson and T. G. Kurtz, *Stochastic analysis of biochemical systems*. Springer, 2015, vol. 674.

[31] L. Cardelli, "Two-domain DNA strand displacement," *Mathematical Structures in Computer Science*, vol. 23, no. 2, pp. 247–271, 2013.

[32] D. Soloveichik, G. Seelig, and E. Winfree, "DNA as a universal substrate for chemical kinetics," *Proceedings of the National Academy of Sciences*, vol. 107, no. 12, pp. 5393–5398, 2010.

[33] F. C. Simmel, B. Yurke, and H. R. Singh, "Principles and applications of nucleic acid strand displacement reactions," *Chemical reviews*, vol. 119, no. 10, pp. 6326–6369, 2019.

Chapter 5

Regulation strategies for two-output biomolecular networks

Regulation strategies for two-output biomolecular networks

Emmanouil Alexis[1], Carolin CM Schulte[1,2,#], Luca Cardelli[3], and Antonis Papachristodoulou[1,*]

[1]Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK

[2]Department of Biology, University of Oxford, Oxford OX1 3RB, UK

[3]Department of Computer Science, University of Oxford, Oxford OX1 3QD, UK

[#]Current affiliation: Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA

[*]Correspondence: antonis@eng.ox.ac.uk

Abstract

Feedback control theory facilitates the development of self-regulating systems with desired performance which are predictable and insensitive to disturbances. Feedback regulatory topologies are found in many natural systems and have been of key importance in the design of reliable synthetic bio-devices operating in complex biological environments. Here, we study control schemes for biomolecular processes with two outputs of interest, expanding previously described concepts based on single-output systems. Regulation of such processes may unlock new design possibilities but can be challenging due to coupling interactions; also potential disturbances applied on one of the outputs may affect both. We therefore propose architectures for robustly manipulating the ratio/product and linear combinations of the outputs as well as each of the outputs independently. To demonstrate their characteristics, we apply these architectures to a simple process of two mutually activated biomolecular species. We also highlight the potential for experimental implementation by exploring synthetic realizations both *in vivo* and *in vitro*. This work presents an important step forward in building bio-devices capable of sophisticated functions.

1 Introduction

For more than two decades we have witnessed significant advances in the highly interdisciplinary field of synthetic biology whose goal it is to harness engineering approaches in order to realize genetic networks that produce user-defined cell behaviour. These advances have the potential to transform several aspects of our life by providing efficient solutions to many global challenges related to food security, healthcare, energy and the environment [1–6]. A fundamental characteristic of living systems is the presence of multi-scale feedback mechanisms facilitating their functioning and survival [7, 8]. Feedback control enables a self-regulating system to adjust its current and future actions by sensing the state of its outputs, thus maintaining an acceptable response even in the face of unintended and unknown changes. This can be the answer to a number of major challenges [9–11] that prevent the successful implementation of synthetic genetic circuits and keep innovative endeavours in the field trapped at the laboratory stage. Control theory offers a rich toolkit of powerful techniques to design and manipulate biological systems and enable the reliable function of next-generation synthetic biology applications [12–16].

94

Engineering synthetic gene circuits aims at constructing modular biomolecular devices which are able to operate in a controllable and predictable way in constantly changing environments with a high level of metabolic burden and interactions (cross-talk) with endogenous signaling systems. It is therefore a requirement for them to be resilient to context-dependent effects and adapt to external environmental perturbations. Several control approaches inspired by both natural and technological systems have recently been proposed allowing for effective and robust regulation of biological networks *in vivo* and/or *in vitro* [17–23]. Despite some conceptual differences, all of these studies focus on biomolecular systems with one output of interest, such as the expression of a single protein.

Building advanced bio-devices capable of performing more sophisticated computations and tasks requires the design of genetic circuits where multiple inputs are applied and multiple outputs are measured. In control engineering these types of systems are also known as multi-input multi-output or MIMO systems [24]. This may be the key for achieving control of the whole cell, which can be regarded as a very complex MIMO bio-device itself. Regulation of processes comprising multiple interacting variables of interest can be challenging since there may be interactions between inputs and outputs. Thus, a change in any input may affect all outputs. At the same time, when attempting to apply feedback control by "closing the loop", a quandary arises as to which input should be connected with which output (input-output pairing problem). Addressing such problems therefore requires alternative, suitably adjusted regulation schemes which take into account the presence of mutual internal interactions in the network to be controlled (open-loop system).

The research area of MIMO control bio-systems has up until now remained relatively unexplored. There have been only a few studies towards this direction, associated with cybergenetic approaches where a computer is a necessary part of the control feedback loop [25, 26]. In contrast, substantial progress has been made in a closely related area, namely MIMO logic bio-circuits which are able to realize Boolean functions [27, 28] while "multi-layer/level" control concepts for one-output processes [29, 30] and resource allocation in gene expression [31] have also been proposed.

In this paper, we investigate regulation strategies for biomolecular networks with two outputs of interest which can correspond, for example, to the concentration of two different proteins inside the cell, assuming the presence of mutual interactions. Both the open-loop and the closed-loop system (open-loop system within a feedback control configuration) are represented by chemical reaction networks (CRNs) obeying the law of mass action [8]. Consequently, the entire regulation process takes place

95

in the biological context of interest without the use of computer-aided methods. Our designs take advantage of the antithetic integral motif which was first introduced in [32] and whose properties and performance trade-offs have been extensively studied in various single-output biomolecular systems [33–51] . The antithetic integral motif is able to achieve robust steady-state tracking, which is equivalent to the biological principle of robust perfect adaptation (RPA) [17, 52, 53], via integral feedback control. A core element of this motif is an (ideally) irreversible sequestration reaction between two species representing a comparison operation at the molecular level. The memory function, necessary for any integral controller, is performed by a memory variable accumulating, through (mathematical) integration, the error between an output and a set-point of interest over time. In the general case, this memory variable is "hidden" in the sense that it corresponds to a non-physical quantity defined as a (mathematical) combination of the (physical) controller species. The efficacy of this biomolecular mechanism has also been demonstrated experimentally in living cells, at both the cell population and the single-cell level, and in cell-free environments using either external (*in silico*) or embedded single-output control schemes [33, 35, 49, 50, 54–60]. Furthermore, in recent years, considerable attention has been given to topologies combining the antithetic integral controller with proportional and derivative control action or biomolecular buffering [35, 49, 50, 55, 61–65]. Such efforts seek to resolve commonly encountered issues associated with the standalone antithetic integral controller, such as instability, poor transient dynamics including overshoots, and long-lasting oscillations or increased variance.

One of the main objectives of this work is to show how this molecular sequestration mechanism can be utilized to regulate biomolecular processes with more than one output, expanding existing theoretical single-output approaches. Thus, we introduce novel strategies of biomolecular interconnections which are able to efficiently control multi-output biological systems in several ways and discuss important challenges and phenomena arising in such contexts. Focusing primarily on two-output biological systems, we present regulatory designs exploiting "multi-loop" concepts based on two independent feedback loops as well as concepts where the control action is carried out jointly considering both outputs simultaneously. Our designs are scalable and, with appropriate modifications, can handle biological systems with an arbitrary number of outputs.

Specifically, we present regulatory architectures, which we refer to as regulators, capable of achieving one of the following control objectives: robustly driving a) the ratio/product of the outputs; b) a

96

linear combination of the outputs; and c) each of the outputs to a desired value (set-point). At steady state, the architectures of a) and b) result in two coupled outputs which can still affect each other, albeit in a specific way dictated by the respective control approach. On the other hand, the architectures for c) achieve steady-state decoupling, thus making the two outputs independent of each other. Our control schemes can be used for regulation of any arbitrary open-loop process provided that the resulting closed-loop system has a finite, positive steady state and the closed-loop system converges to that steady state as time goes to infinity (closed-loop (asymptotic) stability). Thus, the present analysis focuses exclusively on such scenarios. Furthermore, we mathematically and computationally demonstrate their special characteristics by applying these schemes to a simple, monomolecular, biological process of two mutually activating species. Finally, to highlight their biological relevance and motivate further experimental investigation, we explore potential implementations of our designs.
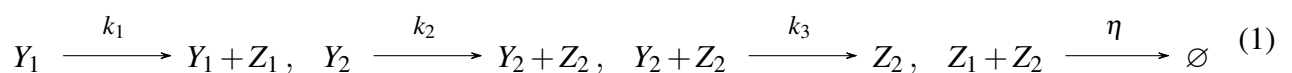
Results

2 Control schemes with steady-state coupling

In Figure 1A we show a general biomolecular process with two outputs of interest for which we first present two bio-controllers aiming to regulate the ratio and an arbitrary linear combination of the outputs, respectively. The different types of biomolecular reactions as well as their graphical representations used in this work are presented in Figure 1B.

2.1 Regulating the ratio of outputs

Figure 1C illustrates a motif which we call Ratio-Regulator (R-Regulator) and consists of the following reactions:

$$Y_1 \xrightarrow{k_1} Y_1 + Z_1, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2, \quad Y_2 + Z_2 \xrightarrow{k_3} Z_2, \quad Z_1 + Z_2 \xrightarrow{\eta} \varnothing \quad (1)$$

This controller consists of two species, Z_1 and Z_2, which annihilate each other. The production of Z_1, Z_2 is catalyzed by the target species Y_1, Y_2, respectively while Y_2 is also inhibited by Z_2.

The dynamics of the R-Regulator are described by the following system of Ordinary Differential

Equations (ODEs):

$$\dot{Z}_1 = k_1 Y_1 - \eta Z_1 Z_2 \tag{2a}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta Z_1 Z_2 \tag{2b}$$

Equations (2a)-(2b) give rise to a non-physical memory variable which enables integration, i.e.:

$$\dot{Z}_1 - \dot{Z}_2 = k_1 Y_1 - k_2 Y_2$$

or

$$(Z_1 - Z_2)(t) = k_1 \int_0^t \left(Y_1(\tau) - \frac{k_2}{k_1} Y_2(\tau) \right) d\tau \tag{3}$$

As a result, assuming closed-loop stability ($\dot{Z}_1, \dot{Z}_2, \dot{Y}_1, \dot{Y}_2 \to 0$ as t $\to \infty$), we get:

$$\frac{Y_1^*}{Y_2^*} = \frac{k_2}{k_1} \tag{4}$$

where the * notation indicates the steady state concentration of a species. As can be seen, the integrand in Equation (3) corresponds to an error quantity which converges to zero over time, thus guaranteeing that the output ratio $\left(\frac{Y_1^*}{Y_2^*} \right)$ will converge to the set-point $\left(\frac{k_2}{k_1} \right)$. It is important to note that the aforementioned stability depends on the structure of the open-loop process, which is unknown here, as well as the set of the reaction rates/parameter values we select for the closed-loop system.
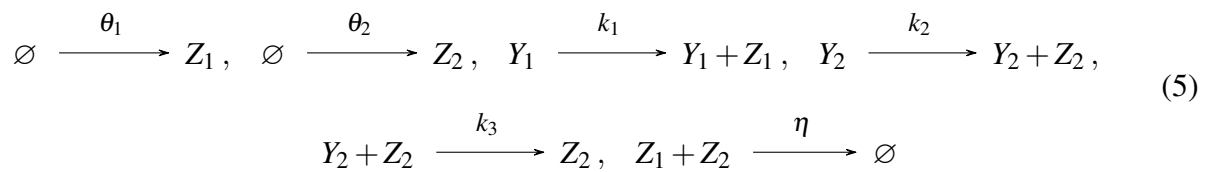
As revealed by Equation (4), the R-Regulator is characterized by a dynamic set-point tracking property regarding species Y_1 and Y_2. This property becomes more apparent if we examine the resulting closed-loop architecture from a different viewpoint. Imagine, for instance, that Y_1 represents an input species through which a "reference signal" is applied while Y_2 represents an output (target) species. Then, Y_2^* is able to track the changes of the set-point $\left(\frac{k_1 Y_1^*}{k_2} \right)$ (and *vice versa*).

A modified version of the above control scheme can be obtained by replacing $Y_1 \xrightarrow{k_1} Y_1 + Z_1$, $Y_2 \xrightarrow{k_2} Y_2 + Z_2$ with $\varnothing \xrightarrow{k_1} Z_1$, $Y_1 + Y_2 \xrightarrow{k_2} Z_2$ in CRN (1). As a result, the memory variable becomes $\dot{Z}_1 - \dot{Z}_2 = k_1 - k_2 Y_1 Y_2$ leading to $Y_1^* Y_2^* = \frac{k_1}{k_2}$. This modified R-Regulator is able to regulate the product of two outputs, assuming both outputs represent species concentrations. Equivalently, this can be seen as regulation of the ratio of two outputs where one of them represents

a species concentration and the other one the reciprocal of a species concentration (see also Section

S12 of the supplementary material for further demonstration).

2.2 Regulating a linear combination of the outputs

In Figure 1D a second motif, which we call Linear Combination - Regulator (LC-Regulator), is depicted. The only difference to the R-Regulator is that species Z_1, Z_2 are also produced through two independent processes with constant rates θ_1, θ_2, respectively. More specifically, the corresponding reaction network is:

$$\varnothing \xrightarrow{\theta_1} Z_1, \quad \varnothing \xrightarrow{\theta_2} Z_2, \quad Y_1 \xrightarrow{k_1} Y_1 + Z_1, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2,$$
$$Y_2 + Z_2 \xrightarrow{k_3} Z_2, \quad Z_1 + Z_2 \xrightarrow{\eta} \varnothing \tag{5}$$

The dynamics of LC-Regulator is given by the set of ODEs:

$$\dot{Z}_1 = \theta_1 + k_1 Y_1 - \eta Z_1 Z_2 \tag{6a}$$

$$\dot{Z}_2 = \theta_2 + k_2 Y_2 - \eta Z_1 Z_2 \tag{6b}$$

Similar to before, in order to see the memory function involved, we subtract Equations (6a) - (6b) and integrate to get:

$$(Z_1 - Z_2)(t) = \int_0^t \left(\left(k_1 Y_1(\tau) - k_2 Y_2(\tau) \right) - \left(\theta_2 - \theta_1 \right) \right) d\tau$$

Under the assumption of closed-loop stability ($\dot{Z}_1, \dot{Z}_2, \dot{Y}_1, \dot{Y}_2 \to 0$ as t $\to \infty$), we have at steady state:

$$k_1 Y_1^* - k_2 Y_2^* = \theta_2 - \theta_1 \tag{7}$$

An interesting feature of LC-Regulator is that Equation (7) can be adjusted as desired by modifying the production reactions regarding Z_1, Z_2. A more general formulation of this control scheme providing a full characterization of the possible (steady-state) output combinations is discussed in Section S2 of the of the supplementary material.

Finally, in Figure 1E we show an alternative version of the controllers presented above. Specifically,

the inhibitory reaction $Y_1 + Z_1 \xrightarrow{k_4} Z_1$ has been added to the R- or LC-Regulator. Note that this additional reaction does not change the dynamics of the controllers - Equations (2a)-(2b) and (6a) - (6b) still hold for R-Regulator and LC-Regulator, respectively. Despite the increase in complexity, the additional reaction strengthens the regulatory ability of the controllers in the sense that control action is now applied on both target species. This could, for example, be useful to make closed-loop stability more robust. These slightly modified motifs are further discussed from a stability viewpoint in Section **Closed-loop stability** and Section S9 of the supplementary material.

3 Control schemes with steady-state decoupling

We now present three alternative bio-controllers, which we call Decoupling - Regulator (D-Regulator) I, II and III, capable of achieving independent control of each output in the arbitrary biomolecular process (Figure 1A). In particular, D-Regulators are able to drive each output species to a desired steady-state concentration unaffected by the behaviour of the other species.

D-Regulators I, II follow a decentralized approach exploiting a "multi-loop" control strategy. More analytically, each of them uses two single-input single-output (SISO) integral controllers which can be constructed separately. This might be advantageous in certain applications in the sense that already-existing, successful SISO implementation techniques can be utilized. However, in the general case, the two SISO controllers cannot be analyzed or tuned independently due to the existence of coupling interactions in the network to be controlled. D-Regulators I, II and, by extension, their resulting closed-loop architectures are MIMO systems and should be studied as such in order for a desirable overall behaviour to be achieved – for instance, in terms of closed-loop stability or dynamic performance of both output responses. Furthermore, in a later section, we investigate a "pairing problem" between actuator and sensor species using a simple example based on one of the above regulators. Problems of such nature are very common in multi-loop contexts and can be difficult to address, especially for complex, strongly coupled networks. Moreover, in D-Regulator I, II the individual SISO controllers constitute alternative realizations of the antithetic integral motif [32]. We choose to focus on these specific versions because of their essential structural differences which can play a crucial role for a circuit designer implementation-wise. At the same time, other well-studied realizations of the antithetic integral motif in the literature appear to be more complex and use the aforementioned
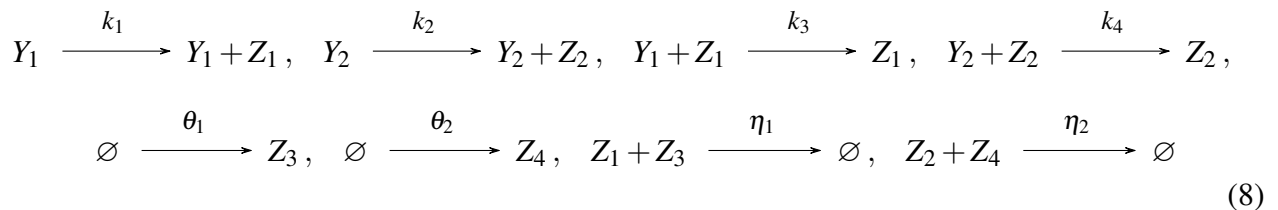
100

versions as a structural basis. A characteristic example is the (SISO) rein controller presented in [46], which is implemented as part of a D-Regulator discussed in Section S3 of the supplementary material.

On the other hand, D-Regulator III follows a centralized approach where some parts of the architecture jointly contribute to the realization of integral control on both output species. This control strategy can result in a structurally simpler topology with fewer controller species. Nevertheless, building such a topology might require more sophisticated biomolecular components.

3.1 D-Regulator I

The set of reactions describing D-Regulator I (Figure 2A) is:

$$Y_1 \xrightarrow{k_1} Y_1 + Z_1 , \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2 , \quad Y_1 + Z_1 \xrightarrow{k_3} Z_1 , \quad Y_2 + Z_2 \xrightarrow{k_4} Z_2 ,$$

$$\varnothing \xrightarrow{\theta_1} Z_3 , \quad \varnothing \xrightarrow{\theta_2} Z_4 , \quad Z_1 + Z_3 \xrightarrow{\eta_1} \varnothing , \quad Z_2 + Z_4 \xrightarrow{\eta_2} \varnothing \tag{8}$$

This design comprises four controller species. The target species Y_1, Y_2 catalyze the formation of two of them, Z_1, Z_2, which, in turn, inhibit the former. In addition, Z_3, Z_4, which are produced independently at a constant rate, participate in annihilation reactions with Z_1 and Z_2, respectively.

The dynamics of D-Regulator I can be modelled using the following set of ODEs:

$$\dot{Z}_1 = k_1 Y_1 - \eta_1 Z_1 Z_3 \tag{9a}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta_2 Z_2 Z_4 \tag{9b}$$

$$\dot{Z}_3 = \theta_1 - \eta_1 Z_1 Z_3 \tag{9c}$$

$$\dot{Z}_4 = \theta_2 - \eta_2 Z_2 Z_4 \tag{9d}$$

In contrast to the regulation strategies presented in the preceding section, D-Regulator I includes two memory variables which carry out integral action independently. Indeed, combining Equations (9a), (9c) results in:

$$(Z_3 - Z_1)(t) = k_1 \int_0^t \left(\frac{\theta_1}{k_1} - Y_1 \right) d\tau \tag{10}$$
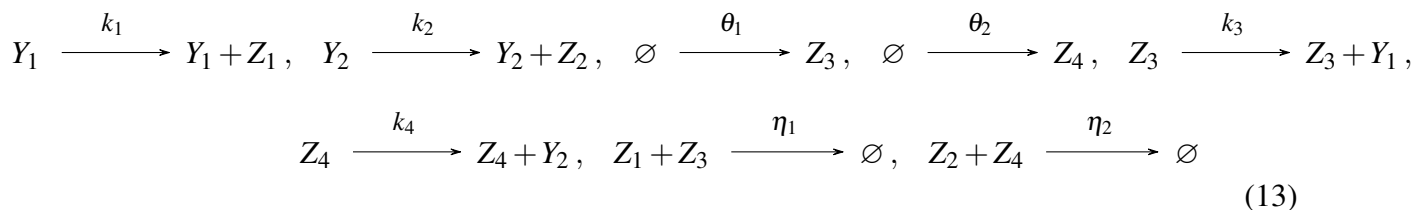
101

while combining Equations (9b), (9d) gives:

$$(Z_4 - Z_2)(t) = k_2 \int_0^t \left(\frac{\theta_2}{k_2} - Y_2 \right) d\tau \tag{11}$$

Consequently, the steady-state output concentrations under the assumption of closed-loop stability $(\dot{Z}_1, \dot{Z}_2, \dot{Z}_3, \dot{Z}_4, \dot{Y}_1, \dot{Y}_2 \to 0$ as t $\to \infty$) are:

$$Y_1^* = \frac{\theta_1}{k_1} , \quad Y_2^* = \frac{\theta_2}{k_2} \tag{12}$$

3.2 D-Regulator II

By using four controller species as before, we construct D-Regulator II (Figure 2B) consisting of the following reactions:
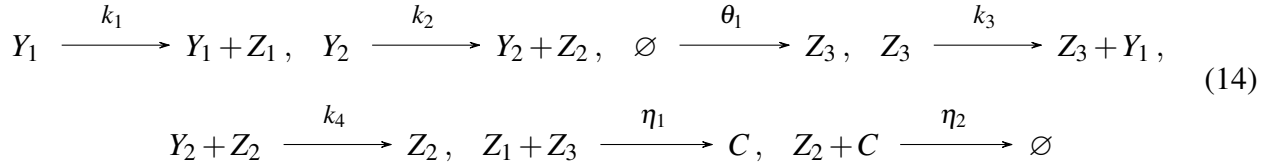
$$Y_1 \xrightarrow{k_1} Y_1 + Z_1 , \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2 , \quad \varnothing \xrightarrow{\theta_1} Z_3 , \quad \varnothing \xrightarrow{\theta_2} Z_4 , \quad Z_3 \xrightarrow{k_3} Z_3 + Y_1 ,$$

$$Z_4 \xrightarrow{k_4} Z_4 + Y_2 , \quad Z_1 + Z_3 \xrightarrow{\eta_1} \varnothing , \quad Z_2 + Z_4 \xrightarrow{\eta_2} \varnothing \tag{13}$$

In this case, species Z_3, Z_4 catalyze the formation of the target species Y_1, Y_2, respectively, and Z_3, Z_4 are produced at a constant rate. Furthermore, species Z_1, Z_2 are catalytically produced by Y_1, Y_2, respectively, while the pairs Z_1-Z_3 and Z_2-Z_4 participate in an annihilation reaction.

Note that the species of D-Regulator II are described by the same ODE model as D-Regulator I (Equations (9a)-(9d)). Thus, the memory variables involved (Equations (10), (11)) as well as the steady-state output behaviour (Equation (12)) are identical in these two motifs (provided that closed-loop stability is guaranteed). Nonetheless, in general, regulating the same open-loop process via the aforementioned controllers results in different output behaviour until an equilibrium is reached or, in other words, the transient responses differ. This is because of the different topological characteristics of the two motifs which cannot be captured by focusing only on the controller dynamics: considering closed-loop dynamics is required, which is addressed in a later section.

3.3 D-Regulator III

215 The last bio-controller presented in this study is D-Regulator III (Figure 2C) whose structure is com-

216 posed of the following reactions:

$$Y_1 \xrightarrow{k_1} Y_1 + Z_1, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2, \quad \varnothing \xrightarrow{\theta_1} Z_3, \quad Z_3 \xrightarrow{k_3} Z_3 + Y_1,$$

$$Y_2 + Z_2 \xrightarrow{k_4} Z_2, \quad Z_1 + Z_3 \xrightarrow{\eta_1} C, \quad Z_2 + C \xrightarrow{\eta_2} \varnothing \tag{14}$$

217   Here there are three controller species. Z_1, Z_3 interact with the target species Y_1 as well as with

218 each other in the same way as in D-Regulator II. The complex C, which is formed by the binding of

219 Z_1, Z_3, and the third controller species, Z_2, can annihilate each other. Finally, the target species Y_2

220 catalyzes the production of Z_2 which, in turn, inhibits Y_2 analogous to D-Regulator I.

221   The dynamics of D-Regulator III can be described by the following set of ODEs:

$$\dot{Z}_1 = k_1 Y_1 - \eta_1 Z_1 Z_3 \tag{15a}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta_2 Z_2 C \tag{15b}$$

$$\dot{Z}_3 = \theta_1 - \eta_1 Z_1 Z_3 \tag{15c}$$

$$\dot{C} = \eta_1 Z_1 Z_3 - \eta_2 Z_2 C \tag{15d}$$

    Similar to the other D-Regulators, the memory function responsible for the regulation of the output Y_1 is carried out by the (non-physical) quantity $Z_3 - Z_1$ (Equation (10)). However, the memory variable related to the output Y_2 is realized in a different way than before. More specifically, combining Equations (15b)-(15d) yields:

$$\dot{Z}_3 + \dot{C} - \dot{Z}_2 = \theta_1 - k_2 Y_2$$

or

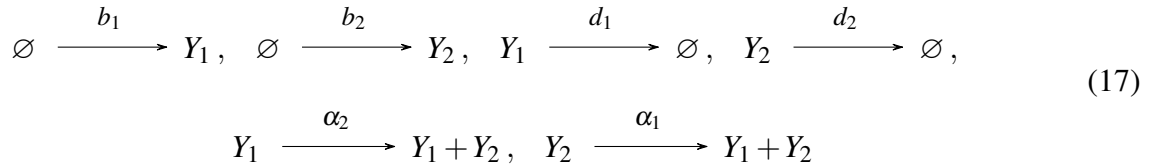$$(Z_3 + C - Z_2)(t) = k_2 \int_0^t \left(\frac{\theta_1}{k_2} - Y_2 \right) d\tau$$

222   Therefore, assuming closed-loop stability, i.e. $\dot{Z}_1, \dot{Z}_2, \dot{Z}_3, \dot{C}, \dot{Y}_1, \dot{Y}_2 \to 0$ as $t \to \infty$, the steady-state

223 output behaviour is:

$$Y_1^* = \frac{\theta_1}{k_1}, \quad Y_2^* = \frac{\theta_1}{k_2} \tag{16}$$

4 Specifying the biological network to be controlled

We now turn our focus to a specific two-output open-loop network which will henceforward take the place of the abstract "cloud" process in the preceding sections. This will allow us to implement *in silico* the proposed control motifs and demonstrate the properties discussed above (see **Implementing the proposed regulation strategies**). In addition, we will explore potential experimental realizations of the resulting closed-loop networks (see **Experimental realization**).

Figure 3A illustrates a simple biological network comprised of two general birth-death processes involving two target species, Y_1, Y_2. These species are coupled in the sense that each of them is able to catalyze the formation of the other. Such motifs of positive feedback action are ubiquitous in biological systems [66–68]. In particular, we have the reactions:

$$\varnothing \xrightarrow{b_1} Y_1 \,, \quad \varnothing \xrightarrow{b_2} Y_2 \,, \quad Y_1 \xrightarrow{d_1} \varnothing \,, \quad Y_2 \xrightarrow{d_2} \varnothing \,, \tag{17}$$

$$Y_1 \xrightarrow{\alpha_2} Y_1 + Y_2 \,, \quad Y_2 \xrightarrow{\alpha_1} Y_1 + Y_2$$

which can be modelled as:

$$\dot{Y}_1 \;=\; b_1 - d_1 Y_1 + \alpha_1 Y_2 \tag{18a}$$

$$\dot{Y}_2 \;=\; b_2 - d_2 Y_2 + \alpha_2 Y_1 \tag{18b}$$

For any $d_1 d_2 > \alpha_1 \alpha_2$, ODE system (18a)-(18b) has the following unique positive steady state:

$$Y_1^* = \frac{\alpha_1 b_2 + b_1 d_2}{d_1 d_2 - \alpha_1 \alpha_2} \,, \quad Y_2^* = \frac{\alpha_2 b_1 + b_2 d_1}{d_1 d_2 - \alpha_1 \alpha_2} \tag{19}$$

which is (globally) exponentially stable (see Section S4 of the supplementary material).

Note that for this system, a change in any of the reaction rates of network (17) due to, for instance, undesired disturbances, will affect the behaviour of both species Y_1 and Y_2 (Figure 3B).

5 Implementing the proposed regulation strategies

We now demonstrate the efficiency of the bio-controllers introduced in **Control schemes with steady-state coupling** and **Control schemes with steady-state decoupling** by regulating the open-loop net-

work (17) presented in **Specifying the biological network to be controlled** (see also **Discussion** for regulation of a more complex network). A detailed analysis of the steady-state behaviour of the resulting closed-loop processes can be found in Section S5 of the supplementary material.

We show in Figure 4 that R-Regulator and LC-Regulator are capable of driving the ratio and a desired linear combination of the output species to the set-point of our choice in the presence of constant disturbances, respectively. Similarly, we illustrate in Figure 5 the ability of D-Regulators to robustly steer each of the output species towards a desired value independently, thus cancelling the steady-state coupling.

In the topology shown in Figure 5B there are two actuation reactions realized though Z_3 and Z_4. Due to the existence of coupling interactions in the network that we aim to control, it is evident that these actuator species act on both Y_1 and Y_2 simultaneously. Consequently, one could argue that an alternative way of closing the loop would be through a different species pairing (Figure 6). In particular, an annihilation (comparison) reaction between Z_1, Z_4 and Z_2, Z_3 could be used instead (Z_1, Z_2 can be considered as sensor species measuring the outputs Y_1, Y_2, respectively). However, it can be demonstrated (see Section S6 of the supplementary material) that this control strategy is not feasible since there is no realistic parameter set that can ensure closed-loop stability.

Finally, in the supplementary material, using three closed-loop architectures (one for each regulator type - R, LC and D), we demonstrate through simulations the robust steady-state tracking property of the systems by perturbing several model parameters (see Section S7 of the supplementary material). At the same time, we computationally investigate the effect of controller species degradation on their performance and how the latter can be mitigated via appropriate parameter tuning (see Section S8 of the supplementary material).

6 Closed-loop stability

As already emphasized, assuming the existence of a finite, positive equilibrium, the proposed regulation strategies require asymptotic closed-loop stability, at least around that equilibrium (locally). A commonly used approach to assess local stability of a nonlinear system is through (Jacobian) linearization. Specifically, we can study the resulting Jacobian matrix [8]. If its eigenvalues have strictly negative real parts, i.e. the matrix is Hurwitz, then the aforementioned equilibrium is locally asymptotically stable. Necessary and sufficient conditions for that can be determined via the Routh-Hurwitz

criterion (see the sections of the supplementary material associated with **Specifying the biological network to be controlled** and **Implementing the proposed regulation strategies**).

Instead of analyzing the system as a whole, we can alternatively examine it as an interconnection of two (or more) subsystems [69, 70]. It is often possible to assess the overall stability by studying those subsystems separately. This could be beneficial when only an input-output property of the system to be controlled is known. To demonstrate this, we consider the R-Regulator and LC-Regulator with two inhibitory reactions controlling a general "cloud" network in a negative feedback configuration, as shown in Figure 1E. Focusing on the behaviour around an equilibrium of interest, we can show in both cases that if $k_2 k_4 Z_1^* = k_1 k_3 Z_2^*$, then the "controller block" corresponds to a positive real (PR) system. It is also known [69] that the negative feedback interconnection of a PR block and a weakly strictly PR (WSPR) one yields an overall asymptotically stable system. Consequently, for every WSPR "cloud block", asymptotic closed-loop stability can be guaranteed. Further details including definitions of PR and WSPR concepts as well as proofs can be found in Section S9 of the supplementary material.

7 Experimental realization

To highlight the feasibility of experimentally realizing the proposed control schemes, this section describes potential *in vivo* and *in vitro* implementations of the open-loop and closed-loop circuits introduced earlier. We first focus on implementations using biological parts that have been characterized in *Escherichia coli* and further discuss a molecular programming approach.

Following the description in **Specifying the biological network to be controlled**, the biological network to be controlled can be realized as shown in Figure 7. In this implementation, Y_1 and Y_2 are heterologous sigma factors [71], which are fused to fluorescent proteins (GFP and mCherry) to facilitate tracking of the output. While genes encoding fusion proteins are shown for simplicity, bicistronic constructs could also be used and may be preferred in practice to avoid impairment of sigma factor activity by fusion to a fluorescent protein. Through a suitable choice of promoters, Y_1 mediates the expression of Y_2 and *vice versa*. Low levels of Y_1 and Y_2 are continuously produced from constitutive promoters, such as promoters from the BioBrick collection [72]. In all following figures, the biological parts underlying these interactions are not explicitly shown.

106

7.1 R-Regulator and LC-Regulator

For the proposed implementation of the R-Regulator (Figure 8), Y_2 mediates expression of the hepatitis C virus protease NS3 fused to maltose-binding protein (MBP) (Z_2). Y_1 facilitates expression of a MBP-single-chain antibody (scFv) fusion (Z_1) that specifically binds to and thus inhibits NS3 protease. Inhibition of NS3 protease activity through coexpression with single-chain antibodies in the cytoplasm of *E. coli* has been demonstrated previously [73]. Adding a recognition sequence to Y_2 will further allow for its degradation by NS3. Importantly, this will require identification of sites in the Y_2 protein that allow for integration of the NS3 recognition sequence without compromising the catalytic activity of Y_2. An additional requirement for the LC-Regulator would be constitutive expression of *malE-scFv* and *malE-scNS3* as indicated in the dashed boxes in Figure 8. It is important to note that binding between the biomolecular species realising the annihilation reaction should ideally be irreversible, which would likely require targeted engineering of a suitable antibody [74] or exploration of alternative protease-protease inhibitor pairs with exceptionally strong binding.

7.2 D-Regulators

Similar to R- and LC-Regulator, the implementation for D-Regulator I makes use of the interaction between NS3 protease and a suitable single-chain antibody (Figure 9A). However, the antibody is solely expressed from a constitutive promoter in this case. As a second protease-protease inhibitor pair, we suggest the *E. coli* Lon protease and the phage T4 protease inhibitor PinA as discussed in our previous work [75]. For this purpose, a suitable degradation tag should be added to Y_1 and to avoid leaky integration due to endogenous Lon protease, a Lon-deficient *E. coli* strain, such as BL21(DE3) [76] should be used. Note that the latter protease-protease inhibitor pair can also be used for realizing the R-Regulator and LC-Regulator.

To realize the two annihilation reactions in D-Regulator II (Figure 9B), we propose the use of sigma factors and anti-sigma factors as described previously [33, 77]. Specifically, Z_3 could be the sigma factor SigW, which is constitutively expressed and mediates expression of SigF (Y_1). SigF mediates expression on the anti-sigma factor RsiW (Z_1), which binds to SigW. Analogous reactions are realized using SigM (Y_2), SigB (Z_4) and RsbW (Z_2).

The design for D-Regulator III may be more difficult to implement experimentally due to the requirement of a two-stage complex formation by three biomolecules (Z_1, Z_2 and Z_3) in addition to

the requirement of Z_3 catalysing the production of Y_1 and Z_2 inhibiting Y_2. While it may be possible to achieve the desired behaviour of biomolecules using protein fusions and/or protein engineering, an alternative method to implement this design (as well as all the others) would be via molecular programming as discussed in the following section. In section S10 of the supplementary material, we further discuss some challenges and limitations of such *in vivo* implementations accompanied by simulations based on more realistic (non-ideal) conditions.

7.3 Molecular programming implementation

In molecular programming, an abstract reaction network is realized by designing a concrete chemical reaction network using engineered molecules, so that the latter network emulates the kinetics of the former. At the edges of the abstract network, appropriate chemical transducers must be introduced to interface the abstract network with the environment. While such transducers are specific to each application, the core network is generic, and DNA (natural or synthetic) is commonly used to construct it. These systems are typically tested *in vitro* in controlled environments, with the eventual aim of embedding them in living cells, in synthetic cells [78], or in other deployable physical media. We refer to [21], Section IV, for details of concrete synthetic DNA schemes in the context of biochemical regulation, and for literature overview. Suffices to say that all the reactions used in this paper can be systematically compiled into networks of synthetic molecules that well approximate the required mass action kinetics [79]. In particular, Section S11 of the supplementary material details the DNA strand-displacement realization of a bimolecular reaction A + B $\rightleftharpoons$ C + D. A collection of such reactions (and their unimolecular special cases) can then realize the chemical reaction networks used in this paper. Tools are available to simulate strand displacement systems, e.g., to evaluate their fidelity to the corresponding chemical reaction networks [21, 22, 79].

8 Discussion

In this paper, we address the challenge of regulating biomolecular processes with two outputs of interest which are, in the general case, co-dependent due to coupling interactions. This co-dependence means that disturbances applied to one of the outputs will also affect the other - each of the output species may be part of a separate, independent network and, by extension, be subject to different per-

turbations. Thus, we propose control schemes for efficient and robust manipulation of such processes adopting concepts based on both output steady-state coupling and decoupling. The proposed regulators describe biomolecular configurations with appropriate feedback interconnections which, under some assumptions, result in closed-loop systems where different types of output regulation can be achieved.

In particular, we present a variety of bio-controllers for regulating the ratio (R-Regulators) and linear combinations of the outputs (LC-Regulators) as well as each of the outputs individually (D-Regulators). At the core of their functioning lies a "hidden" integral feedback action realized in suitable ways in order to meet the control objectives for each case. Integral control is one of the most widely used strategies in traditional control engineering since it guarantees zero control error and constant disturbance rejection at the steady state. This is based on the fact that with this type of control, the existence of a positive/negative error, regardless of its magnitude, always generates an increasing/decreasing control signal. Essential structural components of these designs are production-inhibition loops [75] and/or annihilation reactions [32]. Moreover, to get a more practical insight, we consider a two-output biomolecular network with positive feedback coupling interactions. Treating the network as an open-loop system, we use our control designs to successfully manipulate its outputs under constant parameter perturbations and non-ideal conditions. At the same time, we discuss an alternative way of closing the loop in D-Regulator-II via a different controller species "pairing". Although it may seem reasonable, we show that this feedback configuration leads to an unstable closed-loop system.

Assuming a biologically meaningful equilibrium, the proposed designs can be used to regulate arbitrary biological processes provided that the closed-loop topologies are asymptotically stable. We therefore anticipate that they will be useful for building complex pathways that robustly respond to environmental perturbations in synthetic biology applications. To this end, we extensively discuss ways of achieving local closed-loop asymptotic stability while, for R- and LC- Regulator, we also present specific sufficient conditions based on the concept of positive realness. Furthermore, we describe possible experimental implementations of all regulators using either biomolecular species in *E. coli* or molecular programming.

The regulation strategies presented in this work can be easily adapted to more complex networks to be controlled, than the one introduced in **Specifying the biological network to be controlled**. In

109

Section S12 of the supplementary material, we demonstrate the scalability of our control schemes to networks to be controlled with both monomolecular and bimolecular reactions, a high number of (strongly coupled) species and more than two outputs of interest. We also show that, in general, our control schemes do not require for actuator species to act directly on output species, as happens with the architectures discussed in the main text. In networks to be controlled with high number of accessible species, this can offer significant design flexibility, as different variations of our control schemes might be feasible. Our regulation strategies can be implemented through different biomolecular interconnections provided that the latter result in a stable closed-loop system with suitable memory variables and, by extension, in a desired steady-state output behaviour.

Biological networks are inherently stochastic due to the probabilistic nature of biomolecular interactions [8, 80–83]. In the present study, we use deterministic mathematical analysis and simulations which offer a good approximation of the CRN dynamics when the biomolecular counts are sufficiently high. Thus, an interesting future endeavour would be to investigate the behaviour of our topologies within a stochastic mathematical framework examining, for instance, both the stationary mean and variance [63, 83–86]. Another interesting extension of our work would be to study the non-local behaviour of our topologies. For example, the region of attraction for an equilibrium point of interest can be estimated via Lyapunov functions [87]. Additionally, treating those topologies as interconnections of suitably selected subsystems, dissipativity theory approaches based on storage functions can be used to assess the corresponding (local or global) stability [69, 70].

Data availability

The programming codes supporting this work can be found at: `https://github.com/emgalox/` `MIMO-bio-controllers`.
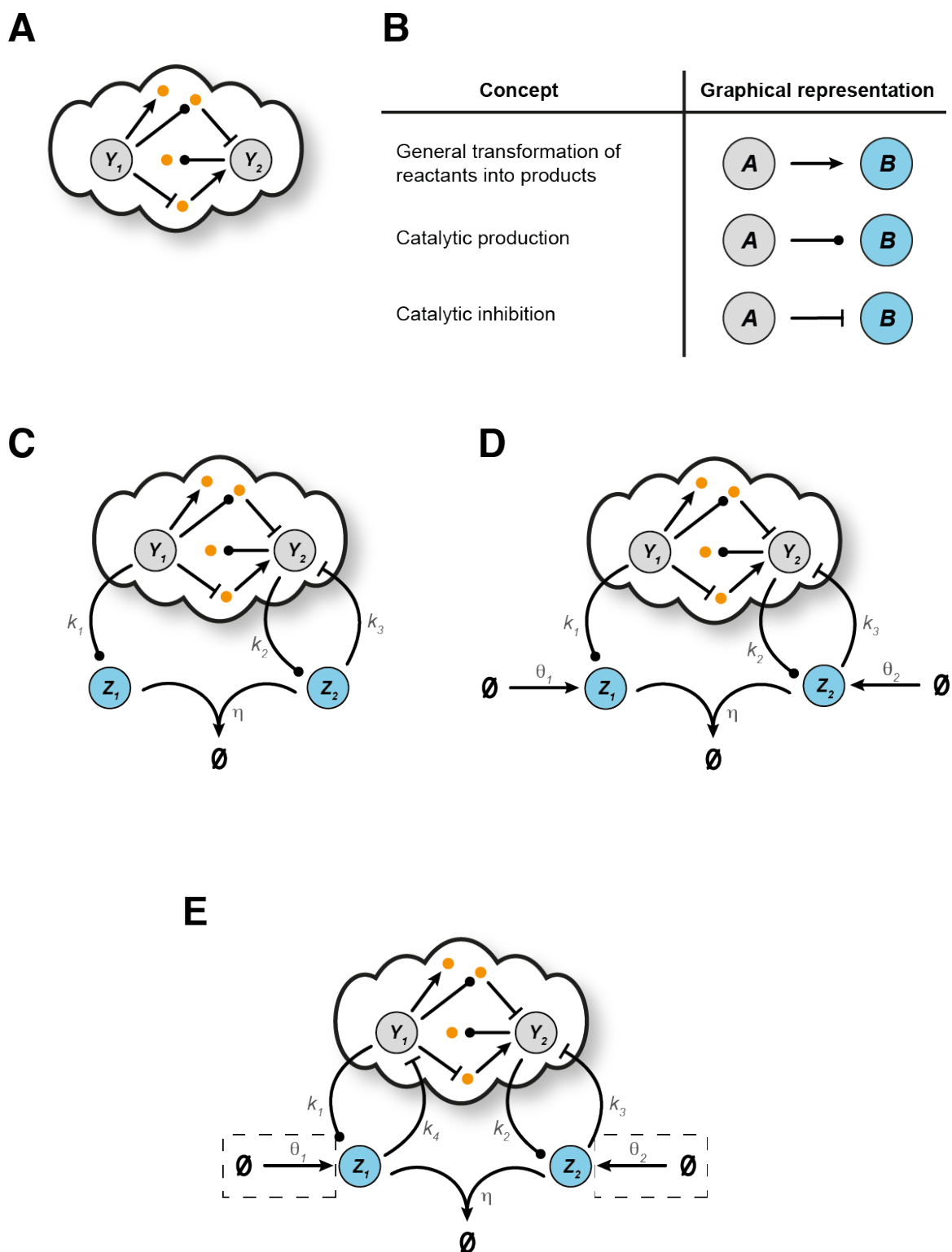
Author contributions

Conceptualization and methodology, E.A., C.C.M.S., A.P., L.C.; Formal analysis and Software: E.A., Writing, E.A., C.C.M.S., A.P., L.C.; Supervision: A.P., L.C.

# Competing interests

The authors declare no competing interests.

Funding

Figure 1: Open-loop biomolecular network and control architectures with steady-state coupling.
A Schematic representation of a general biomolecular network with two output species of interest, Y_1, Y_2, and an arbitrary number of other species and/or biomolecular interactions. **B** Graphical representation of the different types of biochemical reactions adopted from our previous work [75]: general transformation of reactants into products ($A \longrightarrow B$), catalytic production ($A \longrightarrow A + B$), catalytic inhibition ($A + B \longrightarrow A$). Schematic representation of a general closed-loop architecture using **C** R-Regulator (CRN (1)), **D** LC-Regulator (CRN (5)), **E** R- and LC- Regulator with an additional inhibitory reaction (the biological parts enclosed in dashed boxes are only required for LC-Regulator).
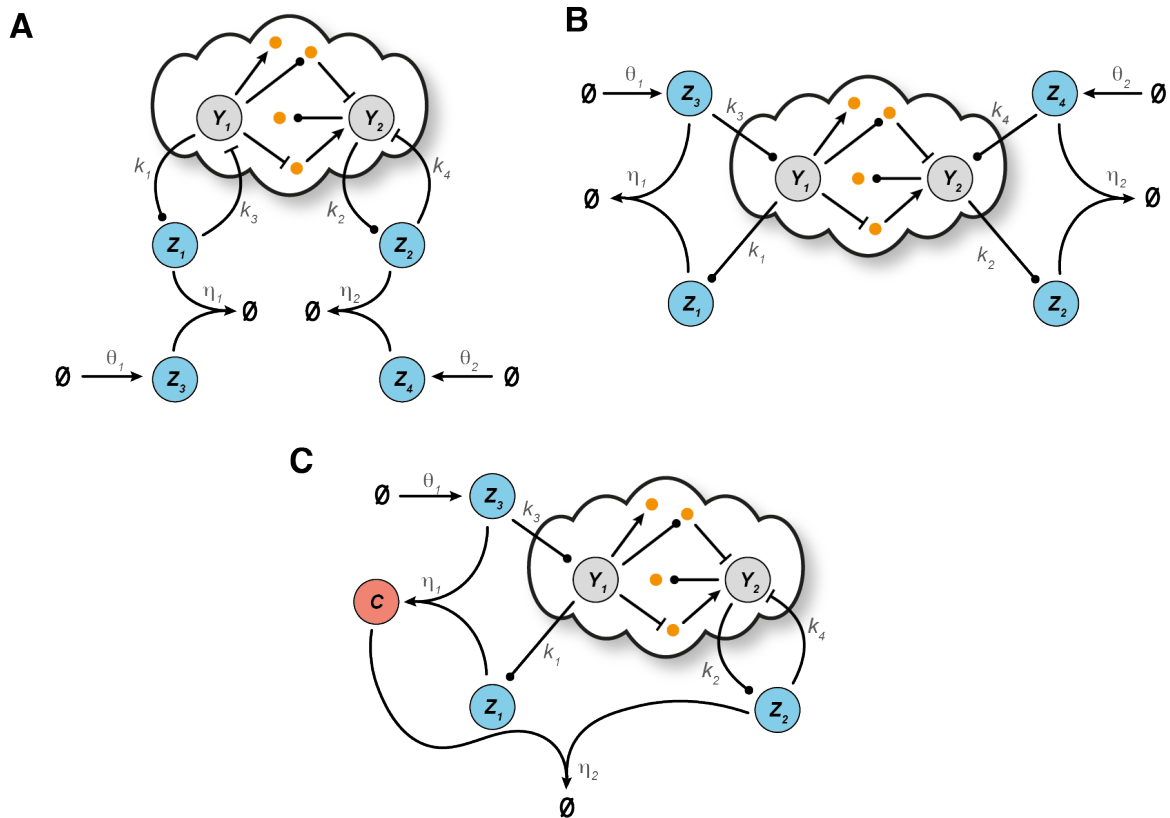
Figure 2: Control architectures with steady-state decoupling.
Schematic representation of a general closed-loop architecture using **A** D-Regulator I (CRN (8)), **B** D-Regulator II (CRN (13)) and **C** D-Regulator III (CRN (14)).
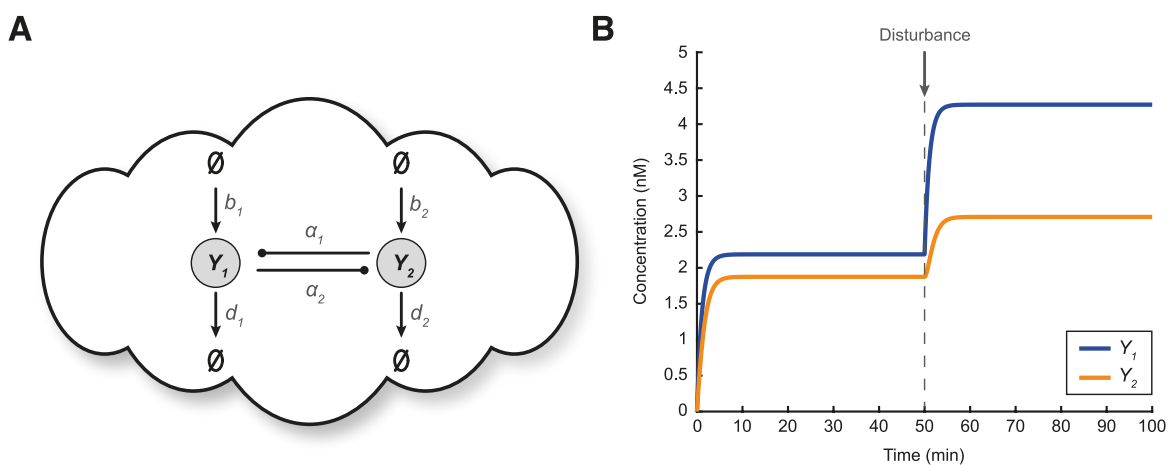
Figure 3: Specifying the open-loop biomolecular network.
A A simple biological process with two mutually activating output species Y_1, Y_2, described by CRN (17). **B** Simulated response of the topology in **A** using the ODE model (18) with the following parameters: $b_1 = 2$ nM min^{-1}, $b_2 = 1$ nM min^{-1}, $d_1 = d_2 = 1$ min^{-1}, $\alpha_1 = 0.1$ min^{-1}, $\alpha_2 = 0.4$ min^{-1}. At time $t = 50$ min, a disturbance on Y_1 is introduced which affects both output species. More specifically, the value of parameter b_1 changes from 2 to 4.
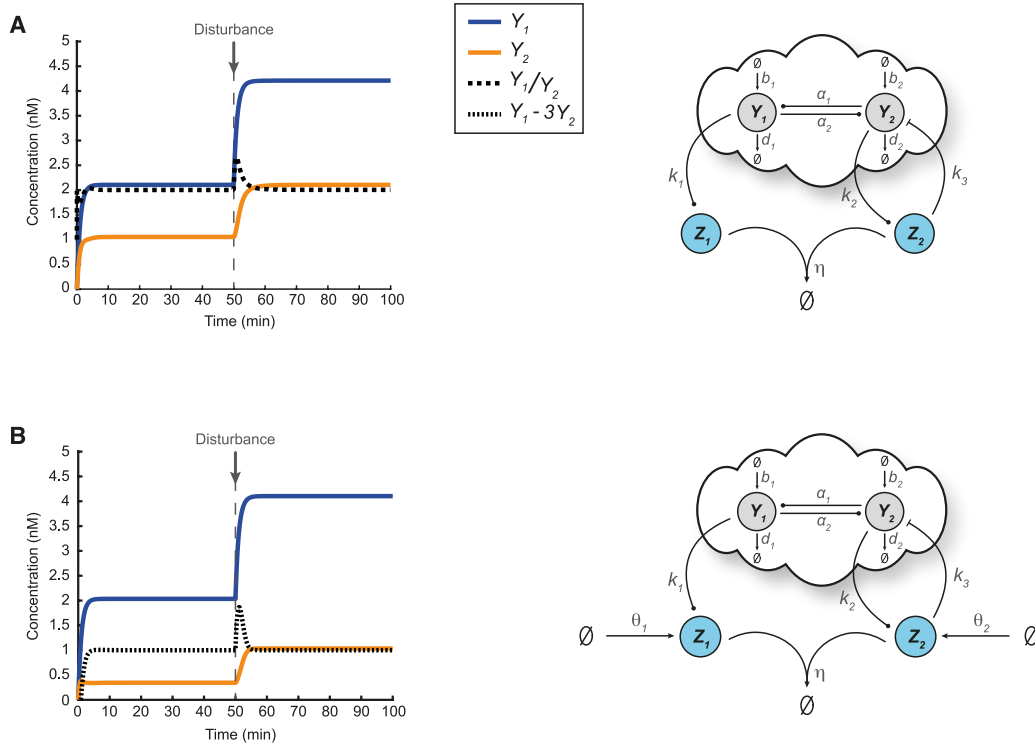
Figure 4: Regulating the ratio and an arbitrary linear combination of the outputs.
A A closed-loop architecture based on the open-loop network shown in Figure 3A and R-Regulator. For the simulated response presented here the following parameters are used: $k_1 = 0.5$ min^{-1}, $k_2 = 1$ min^{-1}, $k_3 = 2$ nM^{-1} min^{-1}, $\eta = 10$ nM^{-1} min^{-1} while the rest of the parameters (associated with the open-loop network) are the same as the ones used in Figure 3B. At time $t = 50$ min, a disturbance is applied (same as in Figure 3B) which alters the output steady states. Nevertheless, $\frac{Y_1^*}{Y_2^*} = \frac{k_2}{k_1} = 2$ always holds (Equation (4)). **B** A closed-loop architecture based on the open-loop network shown in Figure 3A and LC-Regulator. For the simulated response presented here the following parameters are used: $k_1 = 1$ min^{-1}, $k_2 = 3$ min^{-1}, $k_3 = 2$ nM^{-1} min^{-1}, $\eta = 10$ nM^{-1} min^{-1}, $\theta_1 = 4$ nM min^{-1}, $\theta_2 = 5$ nM min^{-1}. The rest of the parameters (associated with the open-loop network) as well as the type of the disturbance (including the time of entry) remain the same as in **A**. Although the output steady states change due to the presence of the disturbance, $k_1 Y_1^* - k_2 Y_2^* = \theta_2 - \theta_1$ or $Y_1^* - 3Y_2^* = 1$ always holds (Equation (7)).
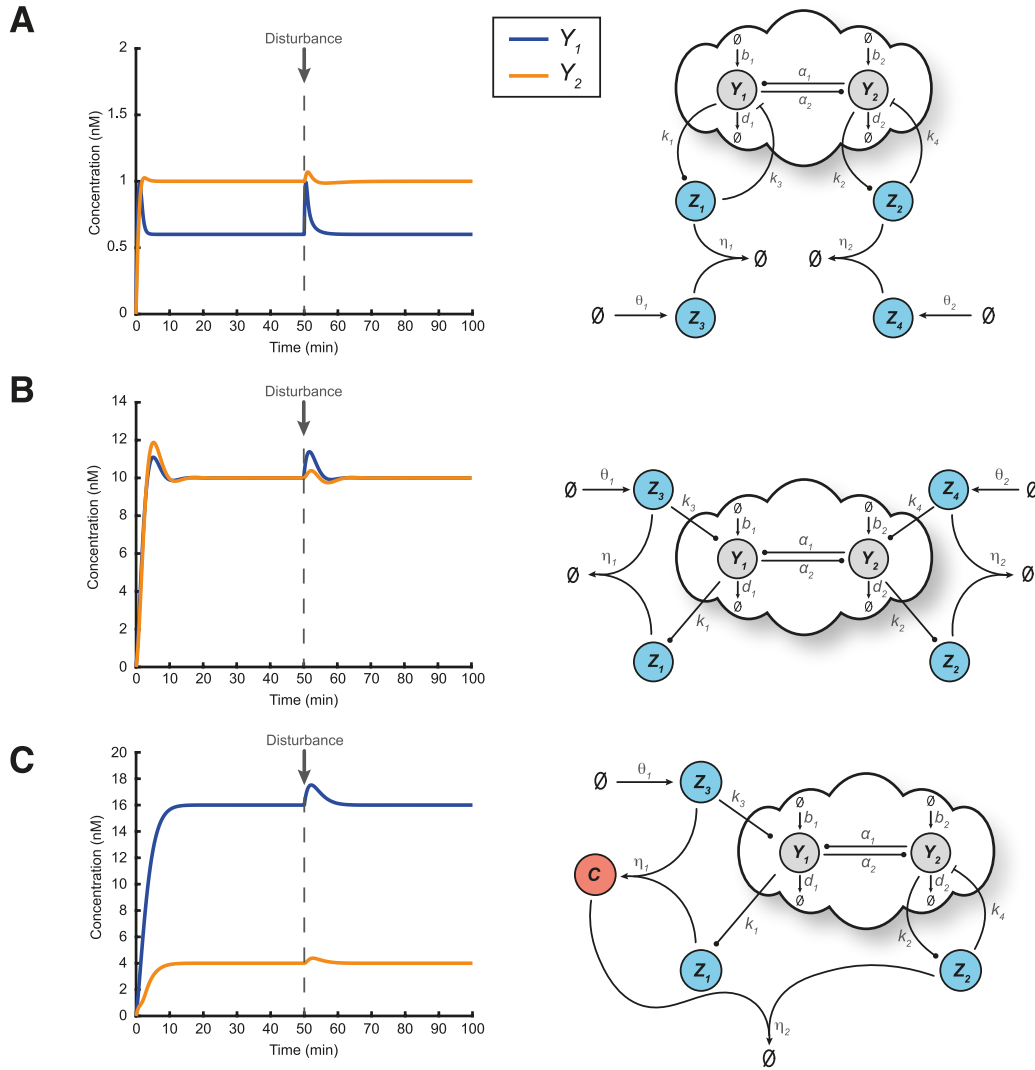
Figure 5: Regulating each output independently.

A A closed-loop architecture based on the open-loop network shown in Figure 3A and D-Regulator I. For the simulated response presented here the following parameters are used: $k_1 = 2.5$ min^{-1}, $k_2 = 0.5$ min^{-1}, $k_3 = 2$ nM^{-1} min^{-1}, $k_4 = 2$ nM^{-1} min^{-1}, $\eta_1 = \eta_2 = 10$ nM^{-1} min^{-1}, $\theta_1 = 1.5$ nM min^{-1}, $\theta_2 = 0.5$ nM min^{-1} while the rest of the parameters (associated with the open-loop network) are the same as the ones used in Figure 3B. Despite the presence of a disturbance, $Y_1^* = \frac{\theta_1}{k_1} = 0.6$ nM, $Y_2^* = \frac{\theta_2}{k_2} = 1$ nM always hold (Equation (12)). **B** A closed-loop architecture based on the open-loop network shown in Figure 3A and D-Regulator II. For the simulated response presented here the following parameters are used: $k_1 = 1$ min^{-1}, $k_2 = 0.8$ min^{-1}, $k_3 = k_4 = 0.5$ min^{-1}, $\eta_1 = \eta_2 = 0.5$ nM^{-1} min^{-1}, $\theta_1 = 10$ nM min^{-1}, $\theta_2 = 8$ nM min^{-1} while the rest of the parameters (associated with the open-loop network) are the same as the ones used in Figure 3B. Despite the presence of a disturbance, $Y_1^* = \frac{\theta_1}{k_1} = 10$ nM, $Y_2^* = \frac{\theta_2}{k_2} = 10$ nM always hold (Equation (12)).
C A closed-loop architecture based on the open-loop network shown in Figure 3A and D-Regulator III. For the simulated response presented here the following parameters are used: $k_1 = 0.5$ min^{-1}, $k_2 = 2$ min^{-1}, $k_3 = 0.5$ min^{-1}, $k_4 = 2$ nM^{-1} min^{-1}, $\eta_1 = 0.5$ nM^{-1} min^{-1}, $\eta_2 = 10$ nM^{-1} min^{-1}, $\theta_1 = 8$ nM min^{-1} while the rest of the parameters (associated with the open-loop network) are the same as the ones used in Figure 3B. Despite the presence of a disturbance, $Y_1^* = \frac{\theta_1}{k_1} = 16$ nM, $Y_2^* = \frac{\theta_1}{k_2} = 4$ nM always hold (Equation (16)). The choice of the set-points in **A**, **B** and **C** is arbitrary while the type of the disturbance (including the time of entry) is the same as in Figure 3B.
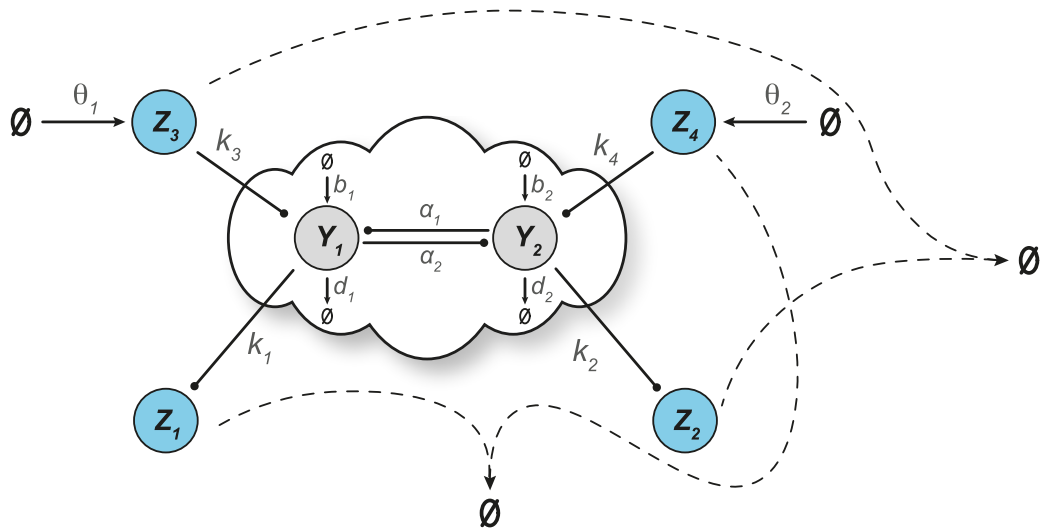
Figure 6: A different feedback configuration regarding the topology shown in Figure 5B. It is based on D-Regulator II with a different actuator-sensor species pairing which leads to instability.
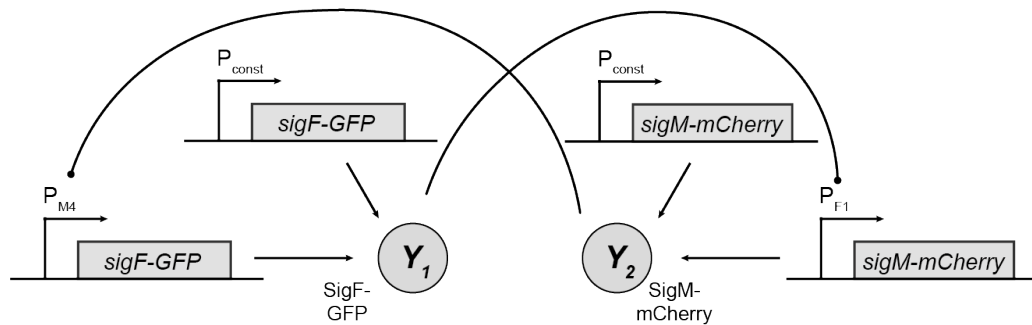
Figure 7: Experimental realization of the network to be controlled described by CRN (17). It constitutes a gene expression system of two mutually activated output species based on a positive feedback loop.
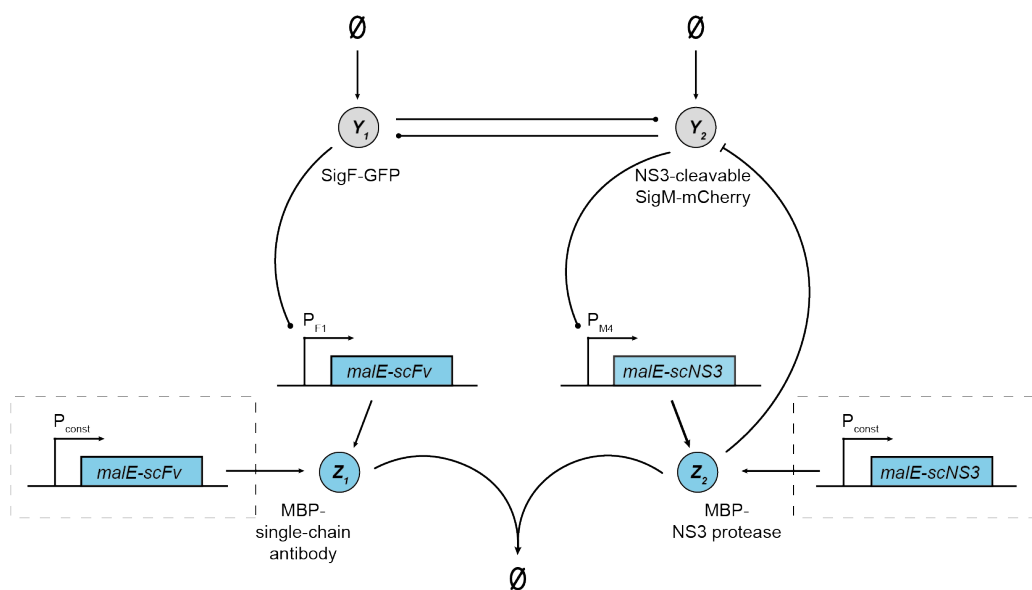
Figure 8: Experimental realization of the closed-loop architecture based on the open-loop network shown in Figure 7 and R-Regulator or LC-Regulator. The biological parts enclosed in dashed boxes are only required for LC-Regulator.
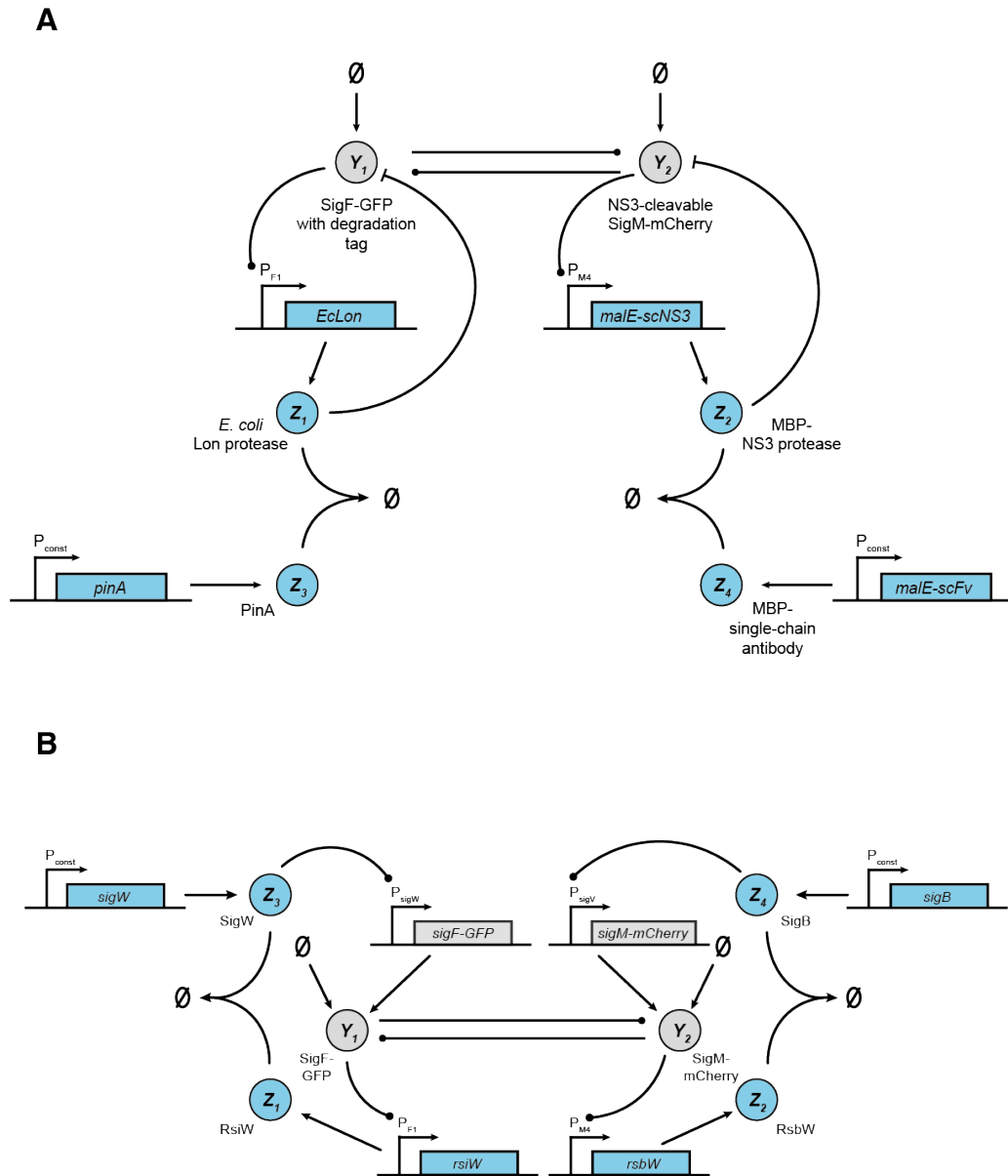
Figure 9: Experimental realization of the closed-loop architecture based on the open-loop network shown in Figure 7 and A D-Regulator I, B D-Regulator II.

References

[1]   Drew Endy. "Foundations for engineering biology". In: *Nature* 438.7067 (2005), pp. 449–453.

[2]   Ahmad S Khalil and James J Collins. "Synthetic biology: applications come of age". In: *Nature Reviews Genetics* 11.5 (2010), pp. 367–379.

[3]   Warren C Ruder, Ting Lu, and James J Collins. "Synthetic biology moving into the clinic". In: *Science* 333.6047 (2011), pp. 1248–1252.

[4]   George M Church et al. "Realizing the potential of synthetic biology". In: *Nature Reviews Molecular Cell Biology* 15.4 (2014), pp. 289–294.

[5]   Fankang Meng and Tom Ellis. "The second decade of synthetic biology: 2010–2020". In: *Nature Communications* 11.1 (2020), pp. 1–4.

[6]   Christopher A Voigt. "Synthetic biology 2020–2030: six commercially-available products that are changing our world". In: *Nature Communications* 11.1 (2020), pp. 1–6.

[7]   Hana El-Samad. "Biological feedback control—Respect the loops". In: *Cell Systems* 12.6 (2021), pp. 477–487.

[8]   Domitilla Del Vecchio and Richard M Murray. *Biomolecular feedback systems*. Princeton University Press Princeton, NJ, 2015.

[9]   Domitilla Del Vecchio, Alexander J Ninfa, and Eduardo D Sontag. "Modular cell biology: retroactivity and insulation". In: *Molecular systems biology* 4.1 (2008), p. 161.

[10] Stefano Cardinale and Adam Paul Arkin. "Contextualizing context for synthetic biology–identifying causes of failure of synthetic biological systems". In: *Biotechnology journal* 7.7 (2012), pp. 856–866.

[11] Yili Qian et al. "Resource competition shapes the response of genetic circuits". In: *ACS synthetic biology* 6.7 (2017), pp. 1263–1272.

[12] Domitilla Del Vecchio, Aaron J Dy, and Yili Qian. "Control theory meets synthetic biology". In: *Journal of The Royal Society Interface* 13.120 (2016), p. 20160380.

[13] Harrison Steel et al. "Challenges at the interface of control engineering and synthetic biology". In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE. 2017, pp. 1014–1023.

118

[14] Victoria Hsiao, Anandh Swaminathan, and Richard M Murray. "Control theory for synthetic biology: recent advances in system characterization, control design, and controller implementation for synthetic biology". In: *IEEE Control Systems Magazine* 38.3 (2018), pp. 32–62.

[15] Domitilla Del Vecchio et al. "Future systems and control research in synthetic biology". In: *Annual Reviews in Control* 45 (2018), pp. 5–17.

[16] Iacopo Ruolo et al. "Control engineering meets synthetic biology: Foundations and applications". In: *Current Opinion in Systems Biology* 28 (2021), p. 100397.

[17] Mustafa H Khammash. "Perfect adaptation in biology". In: *Cell Systems* 12.6 (2021), pp. 509–521.

[18] Nika Shakiba et al. "Context-aware synthetic biology by controller design: Engineering the mammalian cell". In: *Cell Systems* 12.6 (2021), pp. 561–592.

[19] Jinsu Kim and German Enciso. "Absolutely robust controllers for chemical reaction networks". In: *Journal of the Royal Society Interface* 17.166 (2020), p. 20200031.

[20] Aivar Sootla et al. "Dichotomous Feedback: A Signal Sequestration-based Feedback Mechanism for Biocontroller Design". In: *Journal of the Royal Society Interface* 19.189 (2022), p. 20210737.

[21] Max Whitby et al. "PID control of biochemical reaction networks". In: *IEEE Transactions on Automatic Control* (2021).

[22] Nuno MG Paulino et al. "PID and state feedback controllers using DNA strand displacement reactions". In: *IEEE Control Systems Letters* 3.4 (2019), pp. 805–810.

[23] Yitong Ma et al. "Synthetic mammalian signaling circuits for robust cell population control". In: *Cell* 185.6 (2022), pp. 967–979.

[24] Sigurd Skogestad and Ian Postlethwaite. *Multivariable feedback control: analysis and design.* Vol. 2. Citeseer, 2007.

[25] Agostino Guarino, Davide Fiore, and Mario Di Bernardo. "In-silico feedback control of a MIMO synthetic toggle switch via pulse-width modulation". In: *2019 18th European Control Conference (ECC)*. IEEE. 2019, pp. 680–685.

[26]   Jean-Baptiste Lugagne et al. "Balancing a genetic toggle switch by real-time feedback control and periodic forcing". In: *Nature communications* 8.1 (2017), pp. 1–8.

[27]   Benjamin H Weinberg et al. "Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells". In: *Nature biotechnology* 35.5 (2017), pp. 453–462.

[28]   Deepro Bonnerjee, Sayak Mukhopadhyay, and Sangram Bagh. "Design, fabrication, and device chemistry of a 3-input-3-output synthetic genetic combinatorial logic circuit with a 3-input AND gate in a single bacterial cell". In: *Bioconjugate Chemistry* 30.12 (2019), pp. 3013–3020.

[29]   Chelsea Y Hu and Richard M Murray. "Layered Feedback Control Overcomes Performance Trade-off in Synthetic Biomolecular Networks". In: *bioRxiv* (2021).

[30]   F Veronica Greco et al. "Harnessing the central dogma for stringent multi-level control of gene expression". In: *Nature communications* 12.1 (2021), pp. 1–11.

[31]   Alexander PS Darlington and Declan G Bates. "Architectures for Combined Transcriptional and Translational Resource Allocation Controllers". In: *Cell Systems* 11.4 (2020), pp. 382–392.

[32]   Corentin Briat, Ankit Gupta, and Mustafa Khammash. "Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks". In: *Cell Systems* 2.1 (2016), pp. 15–26.

[33]   Stephanie K. Aoki et al. "A universal biomolecular integral feedback controller for robust perfect adaptation". In: *Nature* 570.7762 (2019), pp. 533–537.

[34]   Lorenzo Duso, Tommaso Bianucci, and Christoph Zechner. "Shared antithetic integral control for dynamic cell populations". In: *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE. 2021, pp. 2053–2058.

[35]   Stanislav Anastassov et al. "A cybergenetic framework for engineering intein-mediated integral feedback control systems". In: *Nature Communications* 14.1 (2023), p. 1337.

[36]   Corentin Briat and Mustafa Khammash. "Perfect adaptation and optimal equilibrium productivity in a simple microbial biofuel metabolic pathway using dynamic integral control". In: *ACS synthetic biology* 7.2 (2018), pp. 419–431.

[37]   Deepak K Agrawal et al. "Mathematical modeling of RNA-based architectures for closed loop control of gene expression". In: *ACS synthetic biology* 7.5 (2018), pp. 1219–1228.

[38]   Ania-Ariadna Baetica, Yoke Peng Leong, and Richard M Murray. "Guidelines for designing the antithetic feedback motif". In: *Physical Biology* 17.5 (2020), p. 055002.

[39]   Yadira Boada et al. "Extended metabolic biosensor design for dynamic pathway regulation of cell factories". In: *Iscience* 23.7 (2020), p. 101305.

[40]   Mathias Foo, Ozgur E Akman, and Declan G Bates. "Restoring circadian gene profiles in clock networks using synthetic feedback control". In: *NPJ systems biology and applications* 8.1 (2022), p. 7.

[41]   Noah Olsman and Fulvio Forni. "Antithetic integral feedback for the robust control of monostable and oscillatory biomolecular circuits". In: *IFAC-PapersOnLine* 53.2 (2020), pp. 16826–16833.

[42]   F Nóbel et al. "Reference Conditioning Anti-windup for the Biomolecular Antithetic Controller". In: *IFAC-PapersOnLine* 52.26 (2019), pp. 156–162.

[43]   Yili Qian and Domitilla Del Vecchio. "A singular singular perturbation problem arising from a class of biomolecular feedback controllers". In: *IEEE Control Systems Letters* 3.2 (2018), pp. 236–241.

[44]   Noah Olsman et al. "Hard limits and performance tradeoffs in a class of antithetic integral feedback networks". In: *Cell Systems* 9.1 (2019), pp. 49–63.

[45]   Noah Olsman, Fangzhou Xiao, and John C Doyle. "Architectural principles for characterizing the performance of antithetic integral feedback networks". In: *Iscience* 14 (2019), pp. 277–291.

[46]   Ankit Gupta and Mustafa Khammash. "An antithetic integral rein controller for bio-molecular networks". In: *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE. 2019, pp. 2808–2813.

[47]   Yili Qian and Domitilla Del Vecchio. "Realizing 'integral control' in living cells: how to overcome leaky integration due to dilution?" In: *Journal of The Royal Society Interface* 15.139 (2018), p. 20170902.

[48]   Christian Cuba Samaniego and Elisa Franco. "Ultrasensitive molecular controllers for quasi-integral feedback". In: *Cell Systems* 12.3 (2021), pp. 272–288.

[49] Maurice Filo, Sant Kumar, and Mustafa Khammash. "A hierarchy of biomolecular proportional-integral-derivative feedback controllers for robust perfect adaptation and dynamic performance". In: *Nature communications* 13.1 (2022), pp. 1–19.

[50] Maurice Filo et al. "Exploiting the nonlinear structure of the antithetic integral controller to enhance dynamic performance". In: *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE. 2022, pp. 1294–1299.

[51] Armin M Zand, Mohammad Saleh Tavazoei, and Nikolay V Kuznetsov. "Chaos and its degradation-promoting-based control in an antithetic integral feedback circuit". In: *IEEE Control Systems Letters* 6 (2021), pp. 1622–1627.

[52] Robyn P Araujo and Lance A Liotta. "The topological requirements for robust perfect adaptation in networks of any size". In: *Nature communications* 9.1 (2018), pp. 1–12.

[53] Fangzhou Xiao and John C Doyle. "Robust perfect adaptation in biomolecular reaction networks". In: *2018 IEEE conference on decision and control (CDC)*. IEEE. 2018, pp. 4345–4352.

[54] Victoria Hsiao et al. "Design and implementation of a biomolecular concentration tracker". In: *ACS synthetic biology* 4.2 (2015), pp. 150–161.

[55] Timothy Frei et al. "A genetic mammalian proportional–integral feedback control circuit for robust and precise gene regulation". In: *Proceedings of the National Academy of Sciences* 119.00 (2022), e2122132119.

[56] Sant Kumar, Marc Rullan, and Mustafa Khammash. "Rapid prototyping and design of cyber-genetic single-cell controllers". In: *Nature communications* 12.1 (2021), p. 5651.

[57] Ciarán L Kelly et al. "Synthetic negative feedback circuits using engineered small RNAs". In: *Nucleic acids research* 46.18 (2018), pp. 9875–9889.

[58] Barbara Shannon et al. "In vivo feedback control of an antithetic molecular-titration motif in escherichia coli using microfluidics". In: *ACS Synthetic Biology* 9.10 (2020), pp. 2617–2624.

[59] Deepak K Agrawal et al. "In vitro implementation of robust gene regulation in a synthetic biomolecular integral controller". In: *Nature communications* 10.1 (2019), pp. 1–12.

[60] Hsin-Ho Huang, Yili Qian, and Domitilla Del Vecchio. "A quasi-integral controller for adaptation of genetic modules to variable ribosome demand". In: *Nature communications* 9.1 (2018), pp. 1–12.

[61] Emmanouil Alexis, Luca Cardelli, and Antonis Papachristodoulou. "On the Design of a PID Bio-controller with Set Point Weighting and Filtered Derivative Action". In: *IEEE Control Systems Letters* (2022), pp. 1–1.

[62] Vittoria Martinelli et al. "Multicellular PI Control for Gene Regulation in Microbial Consortia". In: *IEEE Control Systems Letters* 6 (2022), pp. 3373–3378.

[63] Corentin Briat, Ankit Gupta, and Mustafa Khammash. "Antithetic proportional-integral feedback for reduced variance and improved control performance of stochastic reaction networks". In: *Journal of The Royal Society Interface* 15.143 (2018), p. 20180079.

[64] Michael Chevalier et al. "Design and analysis of a proportional-integral-derivative controller with biological molecules". In: *Cell systems* 9.4 (2019), pp. 338–353.

[65] Edward J Hancock and Diego A Oyarzún. "Stabilization of antithetic control via molecular buffering". In: *Journal of the Royal Society Interface* 19.188 (2022), p. 20210762.

[66] Eduardo D Sontag. "Monotone and near-monotone biochemical networks". In: *Systems and synthetic biology* 1.2 (2007), pp. 59–87.

[67] Alexander Y Mitrophanov and Eduardo A Groisman. "Positive feedback in cellular control systems". In: *Bioessays* 30.6 (2008), pp. 542–555.

[68] Uri Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. CRC Press, 2019.

[69] Bernard Brogliato et al. *Dissipative Systems Analysis and Control: Theory and Applications*. Springer, 2020.

[70] Horacio J Marquez. *Nonlinear control systems: analysis and design*. Vol. 161. John Wiley Hoboken^ eN. JNJ, 2003.

[71] Indra Bervoets et al. "A sigma factor toolbox for orthogonal gene expression in Escherichia coli". In: *Nucleic Acids Research* 46.4 (2018), pp. 2133–2144.

[72] Jason R. Kelly et al. "Measuring the activity of BioBrick promoters using an in vivo reference standard". In: *Journal of Biological Engineering* 3.1 (Mar. 2009), p. 4.

[73] Limor Nahary, Alla Trahtenherts, and Itai Benhar. "Isolation of scFvs that Inhibit the NS3 Protease of Hepatitis C Virus by a Combination of Phage Display and a Bacterial Genetic Screen". In: *Antibody Phage Display: Methods and Protocols*. Ed. by Robert Aitken. Totowa, NJ: Humana Press, 2009, pp. 115–132.

[74] Christilyn P. Graff et al. "Directed evolution of an anti-carcinoembryonic antigen scFv with a 4-day monovalent dissociation half-time at 37°C". In: *Protein Engineering, Design and Selection* 17.4 (Apr. 2004), pp. 293–304. ISSN: 1741-0126. DOI: 10.1093/protein/gzh038. eprint: https://academic.oup.com/peds/article-pdf/17/4/293/4694368/gzh038.pdf. URL: https://doi.org/10.1093/protein/gzh038.

[75] Emmanouil Alexis et al. "Biomolecular mechanisms for signal differentiation". In: *iScience* 24.12 (2021).

[76] Julien Ratelade et al. "Production of Recombinant Proteins in the *lon*-Deficient BL21(DE3) Strain of *Escherichia coli* in the Absence of the DnaK Chaperone". In: *Applied and Environmental Microbiology* 75.11 (2009), pp. 3803–3807. DOI: 10.1128/AEM.00255-09. eprint: https://journals.asm.org/doi/pdf/10.1128/AEM.00255-09. URL: https://journals.asm.org/doi/abs/10.1128/AEM.00255-09.

[77] David Chen and Adam P Arkin. "Sequestration-based bistability enables tuning of the switching boundaries and design of a latch". In: *Molecular Systems Biology* 8 (2012), p. 620.

[78] Catherine Fan et al. "Chromosome-free bacterial cells are safe and programmable platforms for synthetic biology". In: *Proceedings of the National Academy of Sciences* 117.12 (2020), pp. 6752–6761.

[79] David Soloveichik, Georg Seelig, and Erik Winfree. "DNA as a universal substrate for chemical kinetics". In: *Proceedings of the National Academy of Sciences* 107.12 (2010), pp. 5393–5398.

[80] Avigdor Eldar and Michael B Elowitz. "Functional roles for noise in genetic circuits". In: *Nature* 467.7312 (2010), pp. 167–173.

[81] Arjun Raj and Alexander Van Oudenaarden. "Nature, nurture, or chance: stochastic gene expression and its consequences". In: *Cell* 135.2 (2008), pp. 216–226.

[82] Mads Kaern et al. "Stochasticity in gene expression: from theories to phenotypes". In: *Nature Reviews Genetics* 6.6 (2005), pp. 451–464.

[83] Corentin Briat and Mustafa Khammash. "Noise in Biomolecular Systems: Modeling, Analysis, and Control Implications". In: *Annual Review of Control, Robotics, and Autonomous Systems* 6 (2022).

[84] Luca Laurenti et al. "Molecular filters for noise reduction". In: *Biophysical Journal* 114.12 (2018), pp. 3000–3011.

[85] Luca Cardelli, Marta Kwiatkowska, and Luca Laurenti. "Stochastic analysis of chemical reaction networks using linear noise approximation". In: *Biosystems* 149 (2016), pp. 26–33.

[86] David J Warne, Ruth E Baker, and Matthew J Simpson. "Simulation and inference algorithms for stochastic biochemical reaction networks: from basic concepts to state-of-the-art". In: *Journal of the Royal Society Interface* 16.151 (2019), p. 20180943.

[87] Hassan K. Khalil. *Nonlinear systems (3nd Ed.)* Prentice-Hall, Inc., 2002.

Regulation strategies for two-output biomolecular networks

Supplementary Material

Emmanouil Alexis[1], Carolin CM Schulte[1,2,#], Luca Cardelli[3], and Antonis Papachristodoulou[1,*]

[1]Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK

[2]Department of Biology, University of Oxford, Oxford OX1 3RB, UK

[3]Department of Computer Science, University of Oxford, Oxford OX1 3QD, UK

[#]Current affiliation: Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA

[*]Correspondence: antonis@eng.ox.ac.uk

S1 Modelling Assumptions

The molecular interactions of the topologies presented in this work are described by chemical reaction networks (CRNs) under mass-action kinetics [1], unless otherwise stated.

S2 LC-Regulator

A general formulation of LC-Regulator (see Section 2.2 **Regulating a linear combination of the outputs** of the main text) providing a full characterization of the possible adjustments is given by the following system of Ordinary Differential Equations (ODEs):

$$\dot{Z}_1 = f_1(Y_1, Y_2) - \eta Z_1 Z_2$$

$$\dot{Z}_2 = f_2(Y_1, Y_2) - \eta Z_1 Z_2$$

where $f_i = \theta_i + k_{i1}Y_1 + k_{i2}Y_2$ and $i = 1, 2$.

Provided that closed-loop stability ($\dot{Z}_1, \dot{Z}_2, \dot{Y}_1, \dot{Y}_2 \to 0$ as t $\to \infty$) can be achieved, we have at steady state:

$$f_1(Y_1, Y_2) - f_2(Y_1, Y_2) = 0$$

or

$$(k_{11} - k_{21})Y_1^* + (k_{12} - k_{22})Y_2^* = \theta_2 - \theta_1 \tag{S1}$$

All possible versions of LC-Regulator can be obtained by selectively setting any of the rates k_{i1}, k_{i2} ($i = 1, 2$) in Equation (S1) to zero. Note that, for a given network to be controlled, not all versions might be capable of ensuring closed-loop stability.

S3 *Rein* D-Regulator

Exploiting the concept introduced in [2], we present *rein* D-Regulator. Assuming the general biomolecular process depicted in Figure S1 as the network to be controlled, *rein* D-Regulator is composed of

the following set of reactions:

$$Y_1 \xrightarrow{k_1} Y_1 + Z_1, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2, \quad \varnothing \xrightarrow{\theta_1} Z_3, \quad \varnothing \xrightarrow{\theta_2} Z_4, \quad Z_3 \xrightarrow{k_3} Z_3 + Y_1,$$

$$Z_4 \xrightarrow{k_4} Z_4 + Y_2, \quad Z_1 + Z_3 \xrightarrow{\eta_1} \varnothing, \quad Z_2 + Z_4 \xrightarrow{\eta_2} \varnothing,$$

$$Y_1 + Z_1 \xrightarrow{k_5} Z_1, \quad Y_2 + Z_2 \xrightarrow{k_6} Z_2$$
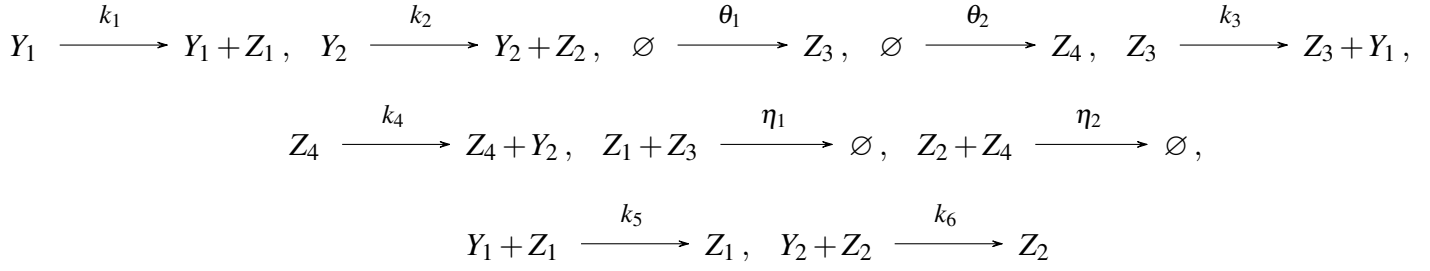
As can be seen, *rein* D-Regulator is essentially a combination of D-Regulator I and II presented in Section 3 **Control schemes with steady-state decoupling** of the main text (see CRNs (8) and (13)) and its behaviour can be described by the same equations as the latter regulators (see Equations (9a)-(9d), (10), (11), (12)).

S4 Open-loop biological network

The open-loop biological network introduced in Section 4 **Specifying the biological network to be controlled** of the main text is represented by the CRN (see Figure 3A) :

$$\varnothing \xrightarrow{b_1} Y_1, \quad \varnothing \xrightarrow{b_2} Y_2, \quad Y_1 \xrightarrow{d_1} \varnothing, \quad Y_2 \xrightarrow{d_2} \varnothing, \tag{S2}$$

$$Y_1 \xrightarrow{\alpha_2} Y_1 + Y_2, \quad Y_2 \xrightarrow{\alpha_1} Y_1 + Y_2$$
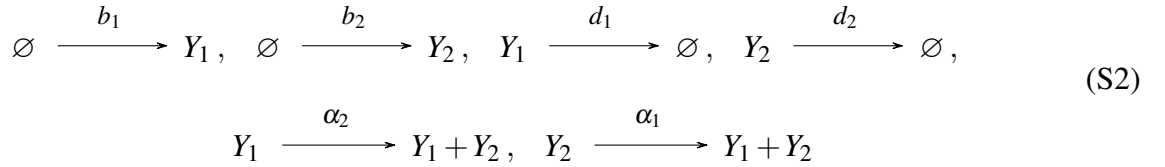
where $b_1, b_2, d_1, d_2, \alpha_1, \alpha_2 \in \mathbb{R}_+$.

The dynamics of CRN (S2) are described by the system of ODEs:

$$\begin{bmatrix} \dot{Y}_1 \\ \dot{Y}_2 \end{bmatrix} = \begin{bmatrix} -d_1 & \alpha_1 \\ \alpha_2 & -d_2 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{S3}$$

Using the linear transformations $y_1 = Y_1 - Y_1^*$, $y_2 = Y_2 - Y_2^*$, we get the following mathematically equivalent system:

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} -d_1 & \alpha_1 \\ \alpha_2 & -d_2 \end{bmatrix}}_{G} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{S4}$$

where $\left(Y_1^* = \dfrac{\alpha_1 b_2 + b_1 d_2}{d_1 d_2 - \alpha_1 \alpha_2}, Y_2^* = \dfrac{\alpha_2 b_1 + b_2 d_1}{d_1 d_2 - \alpha_1 \alpha_2} \right)$ is the unique positive steady state of (Y_1, Y_2) for any $d_1 d_2 > \alpha_1 \alpha_2$.

The characteristic polynomial of system matrix G_1 is:

$$P_o(s) = \det(G_1 - sI) = s^2 + (d_1 + d_2)s + d_1 d_2 - \alpha_1 \alpha_2$$

Both $d_1 + d_2$ and $d_1 d_2 - \alpha_1 \alpha_2$ are positive and, thus, matrix G_1 is Hurwitz (Routh-Hurwitz criterion). Consequently, since system (S4) is linear, the origin is a globally exponentially stable steady state of (S4). By extension, (Y_1^*, Y_2^*) is a globally exponentially stable steady state of system (S3).

S5 Closed-loop biological networks

Here we analyze the behaviour of the closed-loop systems presented in Section 5 **Implementing the proposed regulation strategies** of the main text.

R-Regulator

We have the CRN (see Figure 4A):

$$\varnothing \xrightarrow{b_1} Y_1, \quad \varnothing \xrightarrow{b_2} Y_2, \quad Y_1 \xrightarrow{d_1} \varnothing, \quad Y_2 \xrightarrow{d_2} \varnothing,$$

$$Y_1 \xrightarrow{\alpha_2} Y_1 + Y_2, \quad Y_2 \xrightarrow{\alpha_1} Y_1 + Y_2, \quad Y_1 \xrightarrow{k_1} Y_1 + Z_1,$$

$$Y_2 \xrightarrow{k_2} Y_2 + Z_2, \quad Y_2 + Z_2 \xrightarrow{k_3} Z_2, \quad Z_1 + Z_2 \xrightarrow{\eta} \varnothing$$
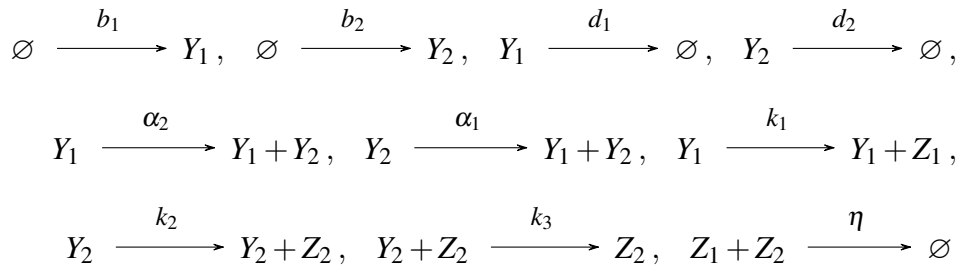
where $b_1, b_2, d_1, d_2, \alpha_1, \alpha_2, k_1, k_2, k_3, \eta \in \mathbb{R}_+$.

The corresponding ODE model is :

$$\dot{Y}_1 = b_1 - d_1 Y_1 + \alpha_1 Y_2 \tag{S5a}$$

$$\dot{Y}_2 = b_2 - d_2 Y_2 + \alpha_2 Y_1 - k_3 Y_2 Z_2 \tag{S5b}$$

$$\dot{Z}_1 = k_1 Y_1 - \eta Z_1 Z_2 \tag{S5c}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta Z_1 Z_2 \tag{S5d}$$

System (S5) has the following unique positive steady state if and only if $\lambda_1 = d_1 k_2 - \alpha_1 k_1 > 0$ and $\lambda_2 = b_1(\alpha_2 k_2 - d_2 k_1) + b_2(d_1 k_2 - \alpha_1 k_1) > 0$:

$$Y_1^* = \frac{b_1 k_2}{\lambda_1} \tag{S6a}$$

$$Y_2^* = \frac{b_1 k_1}{\lambda_1} \tag{S6b}$$

$$Z_1^* = \frac{b_1^2 k_1^2 k_2 k_3}{\eta \lambda_1 \lambda_2} \tag{S6c}$$

$$Z_2^* = \frac{\lambda_2}{b_1 k_1 k_3} \tag{S6d}$$

Combining Equations (S6a), (S6b) yields:

$$\frac{Y_1^*}{Y_2^*} = \frac{k_2}{k_1}$$

By linearizing system (S5) around its steady state (S6) we get:

$$
\begin{bmatrix} \dot{Y}_1 \\ \dot{Y}_2 \\ \dot{Z}_1 \\ \dot{Z}_2 \end{bmatrix}
=
\underbrace{\begin{bmatrix}
-d_1 & \alpha_1 & 0 & 0 \\
\alpha_2 & -(d_2 + k_3 Z_2^*) & 0 & -k_3 Y_2^* \\
k_1 & 0 & -\eta Z_2^* & -\eta Z_1^* \\
0 & k_2 & -\eta Z_2^* & -\eta Z_1^*
\end{bmatrix}}_{G_R}
\begin{bmatrix} Y_1 \\ Y_2 \\ Z_1 \\ Z_2 \end{bmatrix}
\tag{S7}
$$

If all the eigenvalues of system (S7) have a negative real part - matrix G_R is Hurwitz - , then (S6) is a locally exponentially stable steady state for system (S5). This stability criterion can be easily checked for a given set of parameters and was taken into account in all the simulations depicted in Section 5 **Implementing the proposed regulation strategies** of the main text. Of course, as shown in this section, different closed-loop networks may result in different stability matrices. Finally, parameter regimes that guarantee local stability in each case can be found by applying the Routh-Hurwitz criterion.

LC-Regulator

We have the CRN (see Figure 4B):

$$\varnothing \xrightarrow{b_1} Y_1\,, \quad \varnothing \xrightarrow{b_2} Y_2\,, \quad Y_1 \xrightarrow{d_1} \varnothing\,, \quad Y_2 \xrightarrow{d_2} \varnothing \quad Y_1 \xrightarrow{\alpha_2} Y_1+Y_2\,,$$

$$Y_2 \xrightarrow{\alpha_1} Y_1+Y_2\,, \quad \varnothing \xrightarrow{\theta_1} Z_1\,, \quad \varnothing \xrightarrow{\theta_2} Z_2 \quad Y_1 \xrightarrow{k_1} Y_1+Z_1\,,$$

$$Y_2 \xrightarrow{k_2} Y_2+Z_2\,, \quad Y_2+Z_2 \xrightarrow{k_3} Z_2\,, \quad Z_1+Z_2 \xrightarrow{\eta} \varnothing$$
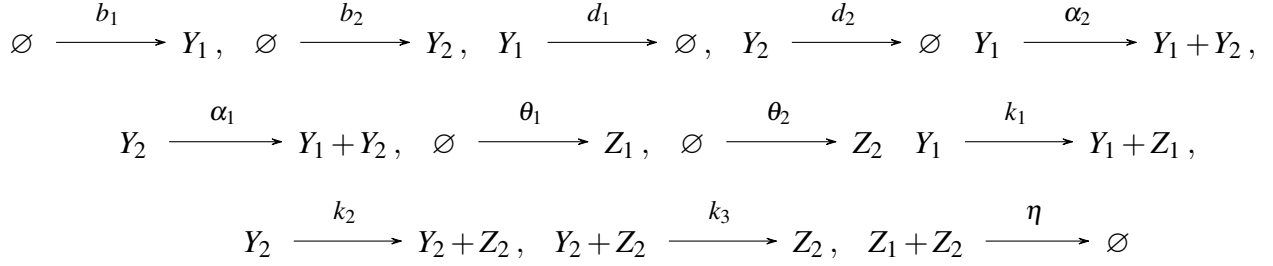
where b_1, b_2, d_1, d_2, α_1, α_2, θ_1, θ_2, k_1, k_2, k_3, $\eta \in \mathbb{R}_+$.

The corresponding ODE model is :

$$\dot{Y}_1 \;=\; b_1 - d_1 Y_1 + \alpha_1 Y_2 \tag{S8a}$$

$$\dot{Y}_2 \;=\; b_2 - d_2 Y_2 + \alpha_2 Y_1 - k_3 Y_2 Z_2 \tag{S8b}$$

$$\dot{Z}_1 \;=\; \theta_1 + k_1 Y_1 - \eta Z_1 Z_2 \tag{S8c}$$

$$\dot{Z}_2 \;=\; \theta_2 + k_2 Y_2 - \eta Z_1 Z_2 \tag{S8d}$$

with the following unique steady state:

$$Y_1^* \;=\; \frac{\lambda_3}{\lambda_1} \tag{S9a}$$

$$Y_2^* \;=\; \frac{\lambda_4}{\lambda_1} \tag{S9b}$$

$$Z_1^* \;=\; \frac{k_3 \lambda_4 \lambda_5}{\eta \lambda_1 \lambda_6} \tag{S9c}$$

$$Z_2^* \;=\; \frac{\lambda_6}{k_3 \lambda_4} \tag{S9d}$$

where $\lambda_3 = b_1 k_2 - \alpha_1(\theta_2 - \theta_1)$, $\lambda_4 = b_1 k_1 - d_1(\theta_2 - \theta_1)$, $\lambda_5 = b_1 k_1 k_2 - \alpha_1 k_1 \theta_2 + d_1 k_2 \theta_1$, $\lambda_6 = \lambda_2 + (d_1 d_2 - \alpha_1 \alpha_2)(\theta_2 - \theta_1)$. Here, we are interested in parameter regimes for which the steady state (S9) is positive.

Using Equations (S9a), (S9b) we calculate:

$$k_1 Y_1^* - k_2 Y_2^* = \frac{1}{\lambda_1}(k_1 \lambda_3 - k_2 \lambda_4) \tag{S10}$$

Taking into account the definitions of λ_1, λ_3 and λ_4 above, relationship (S10) can be rewritten as:

$$k_1 Y_1^* - k_2 Y_2^* = \frac{k_1 k_2 b_1 - k_1 k_2 b_1 + (\theta_2 - \theta_1)(d_1 k_2 - \alpha_1 k_1)}{d_1 k_2 - \alpha_1 k_1}$$

or

$$k_1 Y_1^* - k_2 Y_2^* = \theta_2 - \theta_1$$

Moreover, linearizing system (S8) around its steady state (S9) results in system (S7).

D-Regulator-I

We have the CRN (see Figure 5A):

$$\varnothing \xrightarrow{b_1} Y_1, \quad \varnothing \xrightarrow{b_2} Y_2, \quad Y_1 \xrightarrow{d_1} \varnothing, \quad Y_2 \xrightarrow{d_2} \varnothing \quad Y_1 \xrightarrow{\alpha_2} Y_1 + Y_2,$$

$$Y_2 \xrightarrow{\alpha_1} Y_1 + Y_2, \quad Y_1 \xrightarrow{k_1} Y_1 + Z_1, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2, \quad Y_1 + Z_1 \xrightarrow{k_3} Z_1,$$

$$Y_2 + Z_2 \xrightarrow{k_4} Z_2 \quad \varnothing \xrightarrow{\theta_1} Z_3, \quad \varnothing \xrightarrow{\theta_2} Z_4, \quad Z_1 + Z_3 \xrightarrow{\eta_1} \varnothing,$$

$$Z_2 + Z_4 \xrightarrow{\eta_2} \varnothing$$

where b_1, b_2, d_1, d_2, α_1, α_2, θ_1, θ_2, k_1, k_2, k_3, k_4, η_1, $\eta_2 \in \mathbb{R}_+$.

The corresponding ODE model is :

$$\dot{Y}_1 \;=\; b_1 - d_1 Y_1 + \alpha_1 Y_2 - k_3 Y_1 Z_1 \tag{S11a}$$

$$\dot{Y}_2 \;=\; b_2 - d_2 Y_2 + \alpha_2 Y_1 - k_4 Y_2 Z_2 \tag{S11b}$$

$$\dot{Z}_1 \;=\; k_1 Y_1 - \eta_1 Z_1 Z_3 \tag{S11c}$$

$$\dot{Z}_2 \;=\; k_2 Y_2 - \eta_2 Z_2 Z_4 \tag{S11d}$$

$$\dot{Z}_3 \;=\; \theta_1 - \eta_1 Z_1 Z_3 \tag{S11e}$$

$$\dot{Z}_4 \;=\; \theta_2 - \eta_2 Z_2 Z_4 \tag{S11f}$$

System (S11) has the following unique positive steady state if and only if $\lambda_7 = b_1 k_1 k_2 + \alpha_1 k_1 \theta_2 -$

132

74    $d_1 k_2 \theta_1 > 0$ and $\lambda_8 = b_2 k_1 k_2 + \alpha_2 k_2 \theta_1 - d_2 k_1 \theta_2 > 0$:

$$Y_1^* \ = \ \frac{\theta_1}{k_1} \tag{S12a}$$

$$Y_2^* \ = \ \frac{\theta_2}{k_2} \tag{S12b}$$

$$Z_1^* \ = \ \frac{\lambda_7}{k_2 k_3 \theta_1} \tag{S12c}$$

$$Z_2^* \ = \ \frac{\lambda_8}{k_1 k_4 \theta_2} \tag{S12d}$$

$$Z_3^* \ = \ \frac{k_2 k_3 \theta_1^2}{\eta_1 \lambda_7} \tag{S12e}$$

$$Z_4^* \ = \ \frac{k_1 k_4 \theta_2^2}{\eta_2 \lambda_8} \tag{S12f}$$

Linearization of system (S11) around its steady state (S12) gives:

$$\begin{bmatrix} \dot{Y}_1 \\ \dot{Y}_2 \\ \dot{Z}_1 \\ \dot{Z}_2 \\ \dot{Z}_3 \\ \dot{Z}_4 \end{bmatrix} = \underbrace{\begin{bmatrix} -(d_1 + k_3 Z_1^*) & \alpha_1 & -k_3 Y_1^* & 0 & 0 & 0 \\ \alpha_2 & -(d_2 + k_4 Z_2^*) & 0 & -k_4 Y_2^* & 0 & 0 \\ k_1 & 0 & -\eta_1 Z_3^* & 0 & -\eta_1 Z_1^* & 0 \\ 0 & k_2 & 0 & -\eta_2 Z_4^* & 0 & -\eta_2 Z_2^* \\ 0 & 0 & -\eta_1 Z_3^* & 0 & -\eta_1 Z_1^* & 0 \\ 0 & 0 & 0 & -\eta_2 Z_4^* & 0 & -\eta_2 Z_2^* \end{bmatrix}}_{G_{DI}} \begin{bmatrix} Y_1 \\ Y_2 \\ Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{bmatrix}$$

75   **D-Regulator-II**

We have the CRN (see Figure 5B):

$$\varnothing \xrightarrow{b_1} Y_1 , \quad \varnothing \xrightarrow{b_2} Y_2 , \quad Y_1 \xrightarrow{d_1} \varnothing , \quad Y_2 \xrightarrow{d_2} \varnothing \quad Y_1 \xrightarrow{\alpha_2} Y_1 + Y_2 ,$$

$$Y_2 \xrightarrow{\alpha_1} Y_1 + Y_2 , \quad Y_1 \xrightarrow{k_1} Y_1 + Z_1 , \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2 , \quad \varnothing \xrightarrow{\theta_1} Z_3 ,$$

$$\varnothing \xrightarrow{\theta_2} Z_4 , \quad Z_3 \xrightarrow{k_3} Z_3 + Y_1 \quad Z_4 \xrightarrow{k_4} Z_4 + Y_2 \quad Z_1 + Z_3 \xrightarrow{\eta_1} \varnothing ,$$

$$Z_2 + Z_4 \xrightarrow{\eta_2} \varnothing$$
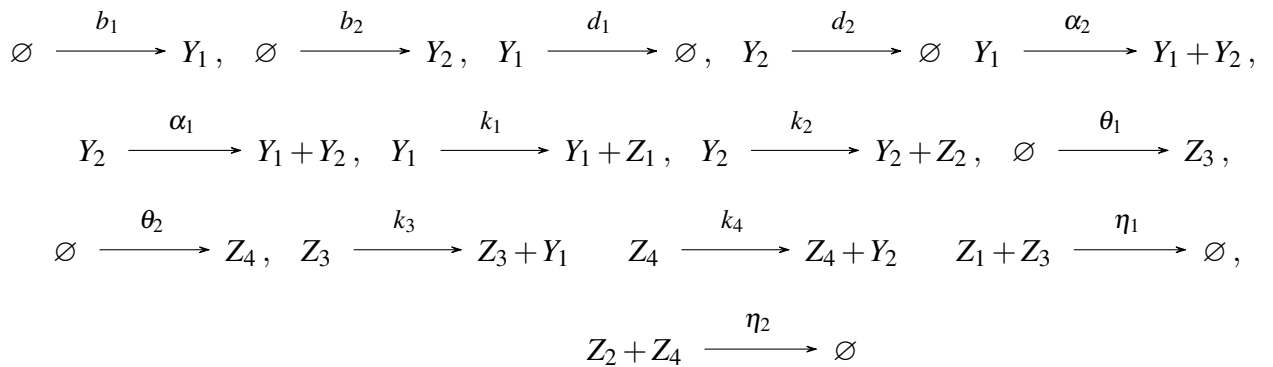
76   where $b_1, b_2, d_1, d_2, \alpha_1, \alpha_2, \theta_1, \theta_2, k_1, k_2, k_3, k_4, \eta_1, \eta_2 \in \mathbb{R}_+$.

The corresponding ODE model is :

$$\dot{Y}_1 = b_1 - d_1 Y_1 + \alpha_1 Y_2 + k_3 Z_3 \tag{S13a}$$

$$\dot{Y}_2 = b_2 - d_2 Y_2 + \alpha_2 Y_1 + k_4 Z_4 \tag{S13b}$$

$$\dot{Z}_1 = k_1 Y_1 - \eta_1 Z_1 Z_3 \tag{S13c}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta_2 Z_2 Z_4 \tag{S13d}$$

$$\dot{Z}_3 = \theta_1 - \eta_1 Z_1 Z_3 \tag{S13e}$$

$$\dot{Z}_4 = \theta_2 - \eta_2 Z_2 Z_4 \tag{S13f}$$

System (S13) has the following unique positive steady state if and only if $\lambda_7 = b_1 k_1 k_2 + \alpha_1 k_1 \theta_2 - d_1 k_2 \theta_1 < 0$ and $\lambda_8 = b_2 k_1 k_2 + \alpha_2 k_2 \theta_1 - d_2 k_1 \theta_2 < 0$:

$$Y_1^* = \frac{\theta_1}{k_1} \tag{S14a}$$

$$Y_2^* = \frac{\theta_2}{k_2} \tag{S14b}$$

$$Z_1^* = -\frac{k_1 k_2 k_3 \theta_1}{\eta_1 \lambda_7} \tag{S14c}$$

$$Z_2^* = -\frac{k_1 k_2 k_4 \theta_2}{\eta_2 \lambda_8} \tag{S14d}$$

$$Z_3^* = -\frac{\lambda_7}{k_1 k_2 k_3} \tag{S14e}$$

$$Z_4^* = -\frac{\lambda_8}{k_1 k_2 k_4} \tag{S14f}$$

We linearize system (S13) around its steady state (S14) to obtain:

$$\begin{bmatrix} \dot{Y}_1 \\ \dot{Y}_2 \\ \dot{Z}_1 \\ \dot{Z}_2 \\ \dot{Z}_3 \\ \dot{Z}_4 \end{bmatrix} = \underbrace{\begin{bmatrix} -d_1 & \alpha_1 & 0 & 0 & k_3 & 0 \\ \alpha_2 & -d_2 & 0 & 0 & 0 & k_4 \\ k_1 & 0 & -\eta_1 Z_3^* & 0 & -\eta_1 Z_1^* & 0 \\ 0 & k_2 & 0 & -\eta_2 Z_4^* & 0 & -\eta_2 Z_2^* \\ 0 & 0 & -\eta_1 Z_3^* & 0 & -\eta_1 Z_1^* & 0 \\ 0 & 0 & 0 & -\eta_2 Z_4^* & 0 & -\eta_2 Z_2^* \end{bmatrix}}_{G_{DII}} \begin{bmatrix} Y_1 \\ Y_2 \\ Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{bmatrix}$$

134

D-Regulator-III

We have the CRN (see Figure 5C):

$$\varnothing \xrightarrow{b_1} Y_1\,, \quad \varnothing \xrightarrow{b_2} Y_2\,, \quad Y_1 \xrightarrow{d_1} \varnothing\,, \quad Y_2 \xrightarrow{d_2} \varnothing \quad Y_1 \xrightarrow{\alpha_2} Y_1 + Y_2\,,$$

$$Y_2 \xrightarrow{\alpha_1} Y_1 + Y_2\,, \quad Y_1 \xrightarrow{k_1} Y_1 + Z_1\,, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2\,, \quad \varnothing \xrightarrow{\theta_1} Z_3\,,$$

$$Z_3 \xrightarrow{k_3} Z_3 + Y_1\,, \quad Y_2 + Z_2 \xrightarrow{k_4} Z_2\,, \quad Z_1 + Z_3 \xrightarrow{\eta_1} C\,, \quad Z_2 + C \xrightarrow{\eta_2} \varnothing$$
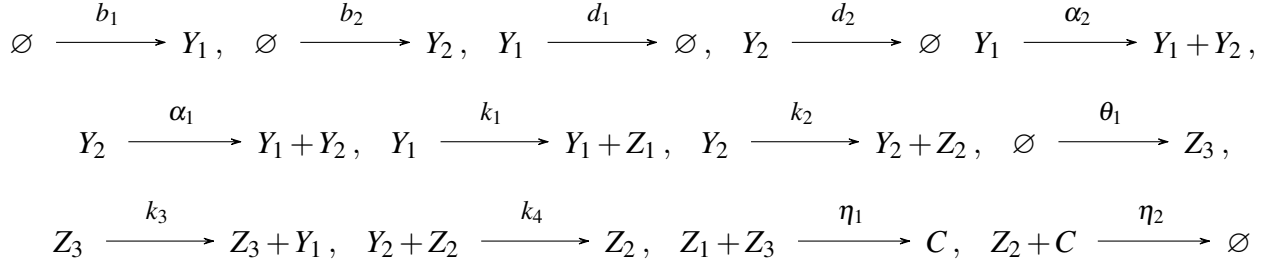
where $b_1, b_2, d_1, d_2, \alpha_1, \alpha_2, \theta_1, k_1, k_2, k_3, k_4, \eta_1, \eta_2 \in \mathbb{R}_+$.

The corresponding ODE model is :

$$\dot{Y}_1 = b_1 - d_1 Y_1 + \alpha_1 Y_2 + k_3 Z_3 \tag{S15a}$$

$$\dot{Y}_2 = b_2 - d_2 Y_2 + \alpha_2 Y_1 - k_4 Y_2 Z_2 \tag{S15b}$$

$$\dot{Z}_1 = k_1 Y_1 - \eta_1 Z_1 Z_3 \tag{S15c}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta_2 Z_2 C \tag{S15d}$$

$$\dot{Z}_3 = \theta_1 - \eta_1 Z_1 Z_3 \tag{S15e}$$

$$\dot{C} = \eta_1 Z_1 Z_3 - \eta_2 Z_2 C \tag{S15f}$$

System (S15) has the following unique positive steady state if and only if $\lambda_9 = -b_1 k_1 k_2 + \theta_1 \lambda_1 > 0$

and $\lambda_{10} = b_2 k_1 k_2 + \theta_1(\alpha_2 k_2 - d_2 k_1) > 0$:

$$Y_1^* = \frac{\theta_1}{k_1} \tag{S16a}$$

$$Y_2^* = \frac{\theta_1}{k_2} \tag{S16b}$$

$$Z_1^* = \frac{k_1 k_2 k_3 \theta_1}{\eta_1 \lambda_9} \tag{S16c}$$

$$Z_2^* = \frac{\lambda_{10}}{k_1 k_4 \theta_1} \tag{S16d}$$

$$Z_3^* = \frac{\lambda_9}{k_1 k_2 k_3} \tag{S16e}$$

$$C^* = \frac{k_1 k_4 \theta_1^2}{\eta_2 \lambda_{10}} \tag{S16f}$$

Linearization of system (S15) around its steady state (S16) yields:

$$
\begin{bmatrix} \dot{Y}_1 \\ \dot{Y}_2 \\ \dot{Z}_1 \\ \dot{Z}_2 \\ \dot{Z}_3 \\ \dot{C} \end{bmatrix} = \underbrace{\begin{bmatrix} -d_1 & \alpha_1 & 0 & 0 & k_3 & 0 \\ \alpha_2 & -(d_2 + k_4 Z_2^*) & 0 & -k_4 Y_2^* & 0 & 0 \\ k_1 & 0 & -\eta_1 Z_3^* & 0 & -\eta_1 Z_1^* & 0 \\ 0 & k_2 & 0 & -\eta_2 C^* & 0 & -\eta_2 Z_2^* \\ 0 & 0 & -\eta_1 Z_3^* & 0 & -\eta_1 Z_1^* & 0 \\ 0 & 0 & \eta_1 Z_3^* & -\eta_2 C^* & \eta_1 Z_1^* & -\eta_2 Z_2^* \end{bmatrix}}_{G_{DIII}} \begin{bmatrix} Y_1 \\ Y_2 \\ Z_1 \\ Z_2 \\ Z_3 \\ C \end{bmatrix}
$$

S6 D-Regulator-II: A different feedback configuration

In this section we explore a different way of "closing the loop" in D-Regulator-II (see Section S5 **D-Regulator-II**). More specifically, we "pair" species Z_1, Z_4 and Z_2, Z_3 by assuming they can annihilate each other.

The resulting CRN is (see Figure 6):

$$\varnothing \xrightarrow{b_1} Y_1, \quad \varnothing \xrightarrow{b_2} Y_2, \quad Y_1 \xrightarrow{d_1} \varnothing, \quad Y_2 \xrightarrow{d_2} \varnothing \quad Y_1 \xrightarrow{\alpha_2} Y_1 + Y_2,$$

$$Y_2 \xrightarrow{\alpha_1} Y_1 + Y_2, \quad Y_1 \xrightarrow{k_1} Y_1 + Z_1, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2, \quad \varnothing \xrightarrow{\theta_1} Z_3,$$

$$\varnothing \xrightarrow{\theta_2} Z_4, \quad Z_3 \xrightarrow{k_3} Z_3 + Y_1 \quad Z_4 \xrightarrow{k_4} Z_4 + Y_2 \quad Z_1 + Z_4 \xrightarrow{\eta_1} \varnothing,$$

$$Z_2 + Z_3 \xrightarrow{\eta_2} \varnothing$$

where b_1, b_2, d_1, d_2, α_1, α_2, θ_1, θ_2, k_1, k_2, k_3, k_4, η_1, $\eta_2 \in \mathbb{R}_+$.
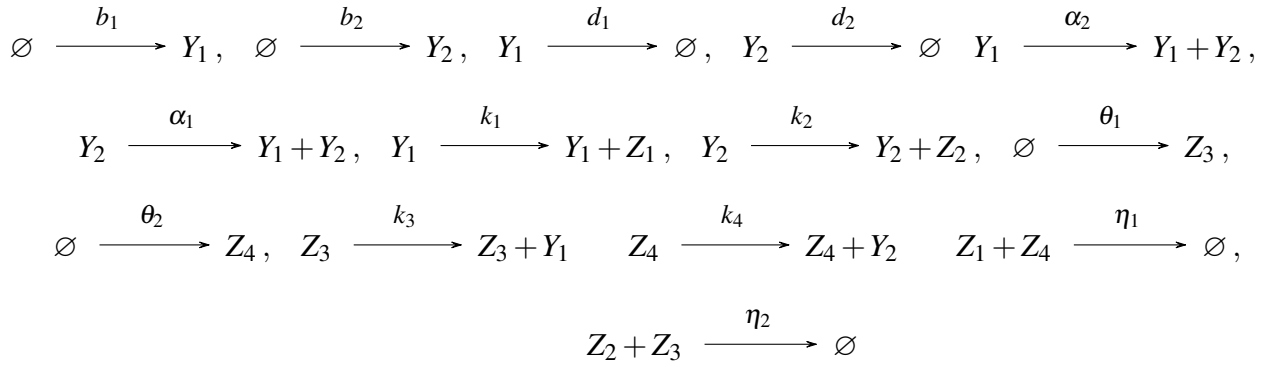
The corresponding ODE model is :

$$\dot{Y}_1 = b_1 - d_1 Y_1 + \alpha_1 Y_2 + k_3 Z_3 \tag{S17a}$$

$$\dot{Y}_2 = b_2 - d_2 Y_2 + \alpha_2 Y_1 + k_4 Z_4 \tag{S17b}$$

$$\dot{Z}_1 = k_1 Y_1 - \eta_1 Z_1 Z_4 \tag{S17c}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta_2 Z_2 Z_3 \tag{S17d}$$

$$\dot{Z}_3 = \theta_1 - \eta_2 Z_2 Z_3 \tag{S17e}$$

$$\dot{Z}_4 = \theta_2 - \eta_1 Z_1 Z_4 \tag{S17f}$$

For any $\lambda_{11} = d_2 k_1 \theta_1 - \alpha_2 k_2 \theta_2 - b_2 k_1 k_2 > 0$ and $\lambda_{12} = d_1 k_2 \theta_2 - \alpha_1 k_1 \theta_1 - b_1 k_1 k_2 > 0$, system (S17) has a unique positive steady state:

$$Y_1^* = \frac{\theta_2}{k_1} \tag{S18a}$$

$$Y_2^* = \frac{\theta_1}{k_2} \tag{S18b}$$

$$Z_1^* = \frac{k_1 k_2 k_4 \theta_2}{\eta_1 \lambda_{11}} \tag{S18c}$$

$$Z_2^* = \frac{k_1 k_2 k_3 \theta_1}{\eta_2 \lambda_{12}} \tag{S18d}$$

$$Z_3^* = \frac{\lambda_{12}}{k_1 k_2 k_3} \tag{S18e}$$

$$Z_4^* = \frac{\lambda_{11}}{k_1 k_2 k_4} \tag{S18f}$$

By linearizing system (S17) around its steady state (S18) we get:

$$
\begin{bmatrix} \dot{Y}_1 \\ \dot{Y}_2 \\ \dot{Z}_1 \\ \dot{Z}_2 \\ \dot{Z}_3 \\ \dot{Z}_4 \end{bmatrix}
=
\underbrace{\begin{bmatrix}
-d_1 & \alpha_1 & 0 & 0 & k_3 & 0 \\
\alpha_2 & -d_2 & 0 & 0 & 0 & k_4 \\
k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\
0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\
0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\
0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11}
\end{bmatrix}}_{G_{DFII}}
\begin{bmatrix} Y_1 \\ Y_2 \\ Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{bmatrix}
\tag{S19}
$$

where $\mu_{14} = \eta_1 Z_4^*$, $\mu_{11} = \eta_1 Z_1^*$, $\mu_{23} = \eta_2 Z_3^*$ and $\mu_{22} = \eta_2 Z_2^*$.

137

The determinant of matrix G_{DFII} can be calculated as follows:

$$\det G_{DFII} = \begin{vmatrix} -d_1 & \alpha_1 & 0 & 0 & k_3 & 0 \\ \alpha_2 & -d_2 & 0 & 0 & 0 & k_4 \\ k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \end{vmatrix} = (-1) \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ \alpha_2 & -d_2 & 0 & 0 & 0 & k_4 \\ -d_1 & \alpha_1 & 0 & 0 & k_3 & 0 \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \end{vmatrix}$$

$$= (-1) \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & -d_2 & \frac{\alpha_2\mu_{14}}{k_1} & 0 & 0 & \frac{k_1 k_4 + \alpha_2\mu_{11}}{k_1} \\ -d_1 & \alpha_1 & 0 & 0 & k_3 & 0 \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \end{vmatrix} = (-1) \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & -d_2 & \frac{\alpha_2\mu_{14}}{k_1} & 0 & 0 & \frac{k_1 k_4 + \alpha_2\mu_{11}}{k_1} \\ 0 & \alpha_1 & -\frac{d_1\mu_{14}}{k_1} & 0 & k_3 & -\frac{d_1\mu_{11}}{k_1} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \end{vmatrix}$$

$$= (-1)^2 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & \alpha_1 & -\frac{d_1\mu_{14}}{k_1} & 0 & k_3 & -\frac{d_1\mu_{11}}{k_1} \\ 0 & -d_2 & \frac{\alpha_2\mu_{14}}{k_1} & 0 & 0 & \frac{k_1 k_4 + \alpha_2\mu_{11}}{k_1} \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \end{vmatrix}$$

$$= (-1)^2 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\frac{d_1\mu_{14}}{k_1} & \frac{\alpha_1\mu_{23}}{k_2} & \frac{k_2 k_3 + \alpha_1\mu_{22}}{k_2} & -\frac{d_1\mu_{11}}{k_1} \\ 0 & -d_2 & \frac{\alpha_2\mu_{14}}{k_1} & 0 & 0 & \frac{k_1 k_4 + \alpha_2\mu_{11}}{k_1} \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \end{vmatrix}$$

138

$$
= (-1)^2 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\frac{d_1\mu_{14}}{k_1} & \frac{\alpha_1\mu_{23}}{k_2} & \frac{k_2k_3+\alpha_1\mu_{22}}{k_2} & -\frac{d_1\mu_{11}}{k_1} \\ 0 & 0 & \frac{\alpha_2\mu_{14}}{k_1} & -\frac{d_2\mu_{23}}{k_2} & -\frac{d_2\mu_{22}}{k_2} & \frac{k_1k_4+\alpha_2\mu_{11}}{k_1} \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \end{vmatrix}
$$

$$
= (-1)^3 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & 0 & \frac{\alpha_2\mu_{14}}{k_1} & -\frac{d_2\mu_{23}}{k_2} & -\frac{d_2\mu_{22}}{k_2} & \frac{k_1k_4+\alpha_2\mu_{11}}{k_1} \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\frac{d_1\mu_{14}}{k_1} & \frac{\alpha_1\mu_{23}}{k_2} & \frac{k_2k_3+\alpha_1\mu_{22}}{k_2} & -\frac{d_1\mu_{11}}{k_1} \end{vmatrix}
$$

$$
= (-1)^3 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & 0 & 0 & -\frac{d_2\mu_{23}}{k_2} & -\frac{d_2\mu_{22}}{k_2} & k_4 \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\frac{d_1\mu_{14}}{k_1} & \frac{\alpha_1\mu_{23}}{k_2} & \frac{k_2k_3+\alpha_1\mu_{22}}{k_2} & -\frac{d_1\mu_{11}}{k_1} \end{vmatrix}
= (-1)^3 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & 0 & 0 & -\frac{d_2\mu_{23}}{k_2} & -\frac{d_2\mu_{22}}{k_2} & k_4 \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & 0 & \frac{\alpha_1\mu_{23}}{k_2} & \frac{k_2k_3+\alpha_1\mu_{22}}{k_2} & 0 \end{vmatrix}
$$

$$
= (-1)^4 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & 0 & 0 & \frac{\alpha_1\mu_{23}}{k_2} & \frac{k_2k_3+\alpha_1\mu_{22}}{k_2} & 0 \\ 0 & 0 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & 0 & -\frac{d_2\mu_{23}}{k_2} & -\frac{d_2\mu_{22}}{k_2} & k_4 \end{vmatrix}
= (-1)^4 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & 0 & 0 & \frac{\alpha_1\mu_{23}}{k_2} & \frac{k_2k_3+\alpha_1\mu_{22}}{k_2} & 0 \\ 0 & 0 & 0 & 0 & \frac{k_2k_3}{\alpha_1} & 0 \\ 0 & 0 & 0 & -\frac{d_2\mu_{23}}{k_2} & -\frac{d_2\mu_{22}}{k_2} & k_4 \end{vmatrix}
$$

139

$$= (-1)^4 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & 0 & 0 & \frac{\alpha_1\mu_{23}}{k_2} & \frac{k_2k_3+\alpha_1\mu_{22}}{k_2} & 0 \\ 0 & 0 & 0 & 0 & \frac{k_2k_3}{\alpha_1} & 0 \\ 0 & 0 & 0 & 0 & \frac{d_2k_3}{\alpha_1} & k_4 \end{vmatrix} = (-1)^5 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & 0 & 0 & \frac{\alpha_1\mu_{23}}{k_2} & \frac{k_2k_3+\alpha_1\mu_{22}}{k_2} & 0 \\ 0 & 0 & 0 & 0 & \frac{d_2k_3}{\alpha_1} & k_4 \\ 0 & 0 & 0 & 0 & \frac{k_2k_3}{\alpha_1} & 0 \end{vmatrix}$$

$$= (-1)^5 \begin{vmatrix} k_1 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & k_2 & 0 & -\mu_{23} & -\mu_{22} & 0 \\ 0 & 0 & -\mu_{14} & 0 & 0 & -\mu_{11} \\ 0 & 0 & 0 & \frac{\alpha_1\mu_{23}}{k_2} & \frac{k_2k_3+\alpha_1\mu_{22}}{k_2} & 0 \\ 0 & 0 & 0 & 0 & \frac{d_2k_3}{\alpha_1} & k_4 \\ 0 & 0 & 0 & 0 & 0 & -\frac{k_2k_4}{d_2} \end{vmatrix}$$

or

$$\det G_{DFII} = -k_1 k_2 k_3 k_4 \mu_{14} \mu_{23} \tag{S20}$$

since the determinant of a triangular matrix is equal to the product of the entries on the main diagonal.

Note also the following:

- The degree of the characteristic polynomial of matrix G_{DFII}, $P_{DFII}(s) = \det(G_{DFII} - sI)$, and, by extension, the number of the eigenvalues of system (S19) is 6 (counting multiplicities).

- The product of the eigenvalues of system (S19) is equal to $det G_{DFII}$.

- All the entries of matrix G_{DFII} are real. Consequently, its complex eigenvalues (if they exist) occur in conjugate pairs.

- The product of a complex number and its conjugate is a real, non-negative number.

- As already discussed, the parameters and the steady state of the system under consideration are positive. Thus, Equation (S20) indicates that $\det G_{DFII} < 0$

We therefore conclude that at least one of the eigenvalues of system (S19) is real and positive. This implies that steady state (S18) cannot be stable and, thus, the feedback configuration in question does

not constitute an efficient regulation strategy.

S7 Perturbation of multiple kinetic parameters

We consider the following closed-loop ODE models (see section S5):

R-Regulator

$$
\begin{aligned}
\dot{Y}_1 &= b_1 - d_1 Y_1 + \alpha_1 Y_2 - k_4 Y_1 Z_1 \\
\dot{Y}_2 &= b_2 - d_2 Y_2 + \alpha_2 Y_1 - k_3 Y_2 Z_2 \\
\dot{Z}_1 &= k_1 Y_1 - \eta Z_1 Z_2 \\
\dot{Z}_2 &= k_2 Y_2 - \eta Z_1 Z_2
\end{aligned}
$$

LC-Regulator

$$
\begin{aligned}
\dot{Y}_1 &= b_1 - d_1 Y_1 + \alpha_1 Y_2 - k_4 Y_1 Z_1 \\
\dot{Y}_2 &= b_2 - d_2 Y_2 + \alpha_2 Y_1 - k_3 Y_2 Z_2 \\
\dot{Z}_1 &= \theta_1 + k_1 Y_1 - \eta Z_1 Z_2 \\
\dot{Z}_2 &= \theta_2 + k_2 Y_2 - \eta Z_1 Z_2
\end{aligned}
$$

D-Regulator-III

$$
\begin{aligned}
\dot{Y}_1 &= b_1 - d_1 Y_1 + \alpha_1 Y_2 + k_3 Z_3 \\
\dot{Y}_2 &= b_2 - d_2 Y_2 + \alpha_2 Y_1 - k_4 Y_2 Z_2 \\
\dot{Z}_1 &= k_1 Y_1 - \eta_1 Z_1 Z_3 \\
\dot{Z}_2 &= k_2 Y_2 - \eta_2 Z_2 C \\
\dot{Z}_3 &= \theta_1 - \eta_1 Z_1 Z_3 \\
\dot{C} &= \eta_1 Z_1 Z_3 - \eta_2 Z_2 C
\end{aligned}
$$

141

Note that for R- and LC-Regulator, we also take into account an additional inhibitory reaction $Y_1 + Z_1 \xrightarrow{k_4} Z_1$ (see Section 2 **Control schemes with steady-state coupling**).

In Figure S2 we illustrate the robust steady-state tracking property of the above regulators by perturbing multiple parameters (reaction rates). The regulators are able to track set-point changes and reject disturbances applied on the system parameters (not involved in the set-point).

Depending on the experimental setting (see Section 7 **Experimental realization** of the main text), different biomolecular processes might be less/more prone to disturbances. Here, we exemplify the potential biological relevance of disturbances (stemming from a natural or synthetic source) that can directly affect the output species (Y_1, Y_2) in common *in vivo* experimental settings:

- Production/activation reactions involving the output species. These reactions might represent gene expression processes where disturbances can be applied as an increase/decrease of the corresponding regulator (activator/repressor) species. The relevant mathematical terms in the above ODE models are: b_1, b_2, $\alpha_1 Y_2$, $\alpha_2 Y_1$.

- Degradation reactions involving the output species. Assuming that the output species represent some proteins of interest, disturbances can be applied via the action of proteases. The relevant mathematical terms in the above ODE models are : $-d_1 Y_1$, $-d_2 Y_2$.

S8 Controller species degradation and adaptation

Our analysis has focused so far on biomolecular architectures where the controller species are not affected by degradation mechanisms - they are only lost due to annihilation/antithetic reactions. This is a requirement for constructing ideal integral controllers and might be able to approximate well experimental realizations where such degradation mechanisms, if present, can be considered practically negligible. Nevertheless, this is not often the case when it comes to implementations in living systems (see Section S10). Degradation of controller species generally leads to a phenomenon known as "leaky" integration, affecting the adaptation property of the system and, thus, inducing steady-state errors [3–5].

We now consider the closed-loop architectures discussed in Section S7 and we appropriately modify the corresponding ODE models to incorporate the action of controller species degradation. We therefore have:

R-Regulator

$$\dot{Y}_1 = b_1 - d_1Y_1 + \alpha_1Y_2 - k_4Y_1Z_1$$

$$\dot{Y}_2 = b_2 - d_2Y_2 + \alpha_2Y_1 - k_3Y_2Z_2$$

$$\dot{Z}_1 = k_1Y_1 - \eta Z_1Z_2 - \gamma Z_1$$

$$\dot{Z}_2 = k_2Y_2 - \eta Z_1Z_2 - \gamma Z_2$$

LC-Regulator

$$\dot{Y}_1 = b_1 - d_1Y_1 + \alpha_1Y_2 - k_4Y_1Z_1$$

$$\dot{Y}_2 = b_2 - d_2Y_2 + \alpha_2Y_1 - k_3Y_2Z_2$$

$$\dot{Z}_1 = \theta_1 + k_1Y_1 - \eta Z_1Z_2 - \gamma Z_1$$

$$\dot{Z}_2 = \theta_2 + k_2Y_2 - \eta Z_1Z_2 - \gamma Z_2$$

D-Regulator-III

$$\dot{Y}_1 = b_1 - d_1Y_1 + \alpha_1Y_2 + k_3Z_3$$

$$\dot{Y}_2 = b_2 - d_2Y_2 + \alpha_2Y_1 - k_4Y_2Z_2$$

$$\dot{Z}_1 = k_1Y_1 - \eta_1Z_1Z_3 - \gamma Z_1$$

$$\dot{Z}_2 = k_2Y_2 - \eta_2Z_2C - \gamma Z_2$$

$$\dot{Z}_3 = \theta_1 - \eta_1Z_1Z_3 - \gamma Z_3$$

$$\dot{C} = \eta_1Z_1Z_3 - \eta_2Z_2C - \gamma C$$

where $\gamma \in \mathbb{R}_+$ represents a degradation rate constant. In D-Regulator-III, we assume that the complex C is degraded, too.

To observe the "leakiness" regarding the integral action taking place within the above regulators, one can calculate the resulting memory variables, which are different from the corresponding (ideal) ones presented in Section 2 **Control schemes with steady-state coupling** and Section 3 **Control schemes with steady-state decoupling** of the main text:

R-Regulator

$$(Z_1 - Z_2)(t) = k_1 \int_0^t \left(Y_1(\tau) - \frac{k_2}{k_1} Y_2(\tau) - \frac{\gamma}{k_1} \Big(Z_1(\tau) - Z_2(\tau) \Big) \right) d\tau$$

LC-Regulator

$$(Z_1 - Z_2)(t) = \int_0^t \left(\Big(k_1 Y_1(\tau) - k_2 Y_2(\tau) \Big) - \Big(\theta_2 - \theta_1 \Big) - \gamma \Big(Z_1(\tau) - Z_2(\tau) \Big) \right) d\tau$$

D-Regulator-III

$$(Z_3 - Z_1)(t) = k_1 \int_0^t \left(\frac{\theta_1}{k_1} - Y_1 - \frac{\gamma}{k_1} \Big(Z_3(\tau) - Z_1(\tau) \Big) \right) d\tau$$

and

$$(Z_3 + C - Z_2)(t) = k_2 \int_0^t \left(\frac{\theta_1}{k_2} - Y_2 - \frac{\gamma}{k_2} \Big(Z_3(\tau) + C(\tau) - Z_2(\tau) \Big) \right) d\tau$$

In Figures S3, S4 and S5-S6 we computationally investigate the negative effect of controller species degradation on the behaviour of R-Regulator, LC-Regulator and D-Regulator III, respectively. More specifically, we show that as γ increases, the behaviour of these systems deviates from the ideal one and their capacity to reject disturbances diminishes. At the same time, it is demonstrated that the aforementioned effect can be mitigated via appropriate parameter tuning. Here, we focus on the following parameters: inhibitory rate k_3 and annihilation rate η regarding R-Regulator and LC-Regulator, production rate k_3 and annihilation rates η_1, η_2 regarding D-Regulator-III.

S9 Feedback interconnection and closed-loop stability

Useful mathematical concepts

- Here we deal with linear, time-invariant systems whose input-output relationship in the Laplace domain can be described by a proper, rational and square transfer function matrix $H(s)$, where $s \in \mathbb{C}$ is the Laplace variable.

- A state-space realization (A, B, C, D) of $H(s)$ is said to be a minimal realization of $H(s)$ if A

has the smallest possible dimension (i.e. the fewest number of states). The smallest dimension is called the McMillan degree of $H(s)$. A mode is hidden if it is not state controllable or observable and thus does not appear in the minimal realization (Definition 4.3 in [6]). Moreover, a state-space realization is minimal if and only if (A,B) is state controllable and (A,C) is state observable [6]. Here, we consider only such state-space realizations.

- The transfer matrix $H(s) \in \mathbb{C}^{m \times m}$ is positive real (PR) if i) $H(s)$ has no pole in $\mathbf{Re}[s] > 0$, ii) $H(s)$ is real for all positive real s, iii) $H(s) + H^H(s) \succcurlyeq 0$ for all $\mathbf{Re}[s] > 0$. (Definition 2.34 in [7]).

- A rational transfer matrix $H(s) \in \mathbb{C}^{m \times m}$ is weakly strictly positive real (WSPR) if i) $H(s)$ is analytic in $\mathbf{Re}[s] \geq 0$, ii) $H(j\omega) + H^T(-j\omega) \succ 0$ for all $\omega \in \mathbb{R}$. (Definition 2.77 in [7]).

The notations ^{H} and ^{T} indicate the conjugate transpose and the transpose of a matrix, respectively while $\succ$ ($\succcurlyeq$) indicates a positive definite (positive-semidefinite) matrix.

R- and LC- Regulator and closed-loop behaviour

We consider a general ("cloud") biomolecular process (see Figure 1A) consisting of q species, $Y_1, Y_2, \ldots, Y_q$ which participate in an arbitrary number of chemical reactions following mass action kinetics. The dynamics of the process can be represented as:

$$\dot{Y} = f(Y) \tag{S27}$$

where $Y = [Y_1 \ Y_2 \ \ldots \ Y_q]^T$. Species Y_1, Y_2 are treated as the species of interest.

We now consider the feedback configuration depicted in Figure 1E where R-Regulator and LC-Regulator are used to control the target (output) species, Y_1, Y_2, of the aforementioned "cloud" process. Given the analysis of Section 2 **Control schemes with steady-state coupling** of the main text and Equation (S27), we have for the closed-loop dynamics:

• R-Regulator case

$$\dot{Y} = f(Y) - \xi_1 k_4 Y_1 Z_1 - \xi_2 k_3 Y_2 Z_2 \tag{S28a}$$

$$\dot{Z}_1 = k_1 Y_1 - \eta_1 Z_1 Z_3 \tag{S28b}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta_2 Z_2 Z_4 \tag{S28c}$$

• LC-Regulator case

$$\dot{Y} = f(Y) - \xi_1 k_4 Y_1 Z_1 - \xi_2 k_3 Y_2 Z_2 \tag{S29a}$$

$$\dot{Z}_1 = \theta_1 + k_1 Y_1 - \eta_1 Z_1 Z_3 \tag{S29b}$$

$$\dot{Z}_2 = \theta_2 + k_2 Y_2 - \eta_2 Z_2 Z_4 \tag{S29c}$$

where $\xi_1 = [1 \ 0 \ \dots \ 0]^T$, $\xi_2 = [0 \ 1 \ \dots \ 0]^T \in \mathbb{Z}^q$ and $b_1, b_2, d_1, d_2, \alpha_1, \alpha_2, \theta_1, \theta_2, k_1, k_2, k_3, k_4, \eta$ $\in \mathbb{R}_+$.

We assume a finite, positive steady state (equilibrium) of interest $E = (Y_1^*, Y_2^*, \dots, Y_q^*, Z_1^*, Z_2^*)$ and we focus on the behaviour of the above closed-loop systems around it. We therefore adopt the coordinate transformations $y_1 = Y_1 - Y_1^*, y_2 = Y_2 - Y_2^*, \dots, y_q = Y_q - Y_q^*, z_1 = Z_1 - Z_1^*, z_2 = Z_2 - Z_2^*$ which denote small perturbations around the aforementioned steady state. The resulting linearized dynamics of both systems (S28) and (S29) are described as:

$$\begin{bmatrix} \dot{y} \\ \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} A_p & -\xi_1 k_4 Y_1^* & -\xi_2 k_3 Y_2^* \\ \xi_1^T k_1 & -\eta Z_2^* & -\eta Z_1^* \\ \xi_2^T k_2 & -\eta Z_2^* & -\eta Z_1^* \end{bmatrix} \begin{bmatrix} y \\ z_1 \\ z_2 \end{bmatrix} \tag{S30}$$

where $y = [y_1 \ y_2 \ \dots \ y_q]^T$ and $A_p = \left. \dfrac{\partial f}{\partial y} \right|_E - \xi_1 k_4 Z_1^* - \xi_2 k_3 Z_2^*$.

System (S30) can be seen as the negative feedback interconnection of two subsystems representing the (linearized) "cloud" process and the controller, respectively. More specifically, we have:

$$\dot{y} = A_p y + B_p u_p \tag{S31a}$$

$$w_p = C_p y + D_p u_p \tag{S31b}$$

146

and

$$\dot{z} = A_c z + B_c u_c \tag{S32a}$$

$$w_c = C_c z + D_c u_c \tag{S32b}$$

where $z = [z_1 \ z_2]^T$, $u_p = [u_{1p} \ u_{2p}]^T$, $u_c = [u_{1c} \ u_{2c}]^T$, $A_c = \begin{bmatrix} -\eta Z_2^* & -\eta Z_1^* \\ -\eta Z_2^* & -\eta Z_1^* \end{bmatrix}$ $B_p = [\xi_1 \ \xi_2]$, $B_c = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$, $C_p = [\xi_1 \ \xi_2]^T$, $C_c = \begin{bmatrix} k_4 & 0 \\ 0 & k_3 \end{bmatrix}$, $D_p = D_c = 0$. In addition, $u_p = -w_c$ and $u_c = w_p$.

We now calculate the transfer function matrix corresponding to state-space model (S32) as $H_c(s) = C_c(sI - A_c)^{-1} + D_c$ to obtain:

$$H_c(s) = \begin{bmatrix} \dfrac{k_1 k_4 (s + \eta Z_1^*)}{s(s + \eta(Z_1^* + Z_2^*))} & \dfrac{-k_2 k_4 \eta Z_1^*}{s(s + \eta(Z_1^* + Z_2^*))} \\ \dfrac{-k_1 k_3 \eta Z_2^*}{s(s + \eta(Z_1^* + Z_2^*))} & \dfrac{k_2 k_3 (s + \eta Z_2^*)}{s(s + \eta(Z_1^* + Z_2^*))} \end{bmatrix} \tag{S33}$$

Here $W_c(s) = H_c(s)U_c(s)$, where $W_c(s)$ and $U_c(s)$ are the Laplace transform of w_c and u_c, respectively.

Theorem If $k_2 k_4 Z_1^* = k_1 k_3 Z_2^*$, then the transfer function matrix $H_c(s)$ (Equation (S33)) is positive real (PR).

Proof. For $k_2 k_4 Z_1^* = k_1 k_3 Z_2^*$, Equation (S33) can be written as:

$$H_c(s) = \begin{bmatrix} \dfrac{k_1 k_4 (s + \eta Z_1^*)}{s(s + \eta(Z_1^* + Z_2^*))} & \dfrac{-k_2 k_4 \eta Z_1^*}{s(s + \eta(Z_1^* + Z_2^*))} \\ \dfrac{-k_2 k_4 \eta Z_1^*}{s(s + \eta(Z_1^* + Z_2^*))} & \dfrac{k_2 k_3 (s + \eta \frac{k_2 k_4}{k_1 k_3} Z_1^*)}{s(s + \eta(Z_1^* + Z_2^*))} \end{bmatrix} \tag{S34}$$

Transfer function matrix (S34) has no poles in $\mathbf{Re}[s] > 0$.

We also calculate:

$$H_c(j\omega) + H_c^H(j\omega) = \begin{bmatrix} \dfrac{1}{k_3} \dfrac{2k_2 k_4^2 \eta Z_1^*}{\omega^2 + \eta^2(Z_1^* + Z_2^*)^2} & \dfrac{2k_2 k_4 \eta Z_1^*}{\omega^2 + \eta^2(Z_1^* + Z_2^*)^2} \\ \dfrac{2k_2 k_4 \eta Z_1^*}{\omega^2 + \eta^2(Z_1^* + Z_2^*)^2} & \dfrac{2k_2 k_3 \eta Z_1^*}{\omega^2 + \eta^2(Z_1^* + Z_2^*)^2} \end{bmatrix}$$

$H_c(j\omega) + H_c^H(j\omega) \succcurlyeq 0$ since $tr(H_c(j\omega) + H_c^H(j\omega)) > 0$ and $det(H_c(j\omega) + H_c^H(j\omega)) = 0$ for all ω.

In addition, $j\omega_0$ is a simple pole of transfer function matrix (S34) with $\omega_0 = 0$ while the corre-

sponding residual is:

$$K_0 = \lim_{s \to 0} sH_c(s) = \begin{bmatrix} \frac{k_1 k_4 \eta Z_1^*}{\eta(Z_1^* + Z_2^*)} & \frac{-k_2 k_4 \eta Z_1^*}{\eta(Z_1^* + Z_2^*)} \\ \frac{-k_2 k_4 \eta Z_1^*}{\eta(Z_1^* + Z_2^*)} & \frac{k_2^2 k_4 \eta Z_1^*}{k_1 \eta(Z_1^* + Z_2^*)} \end{bmatrix}$$
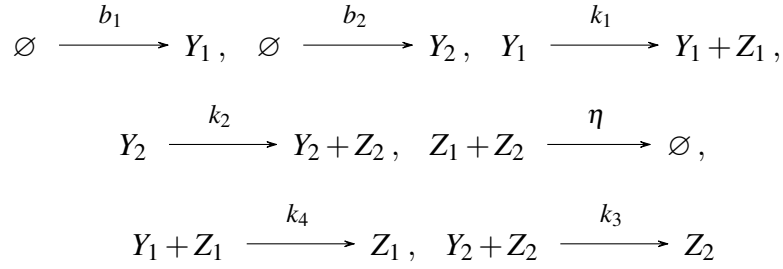
218   $K_0 \succcurlyeq 0$ since $tr(K_0) > 0$ and $det(K_0) = 0$.

219   Thus, according to Theorem 2.48 in [7] transfer function matrix (S34) is PR. $\square$

220   Let now $H_p(s)$ be the transfer function matrix corresponding to state-space model (S31). According

221   to Lemma 3.67 in [7], if $H_p(s)$ is WSPR, then the closed-loop system (S30) is asymptotically stable.

Toy example

We consider a closed-loop system based on R-Regulator described by the following CRN:

$$\varnothing \xrightarrow{b_1} Y_1, \quad \varnothing \xrightarrow{b_2} Y_2, \quad Y_1 \xrightarrow{k_1} Y_1 + Z_1,$$

$$Y_2 \xrightarrow{k_2} Y_2 + Z_2, \quad Z_1 + Z_2 \xrightarrow{\eta} \varnothing,$$

$$Y_1 + Z_1 \xrightarrow{k_4} Z_1, \quad Y_2 + Z_2 \xrightarrow{k_3} Z_2$$

223   For simplicity, we assume unitary kinetic parameter values and obtain the following ODE model for

224   the dynamics:

$$\dot{Y}_1 = 1 - Y_1 Z_1 \tag{S35a}$$

$$\dot{Y}_2 = 1 - Y_2 Z_2 \tag{S35b}$$

$$\dot{Z}_1 = Y_1 - Z_1 Z_2 \tag{S35c}$$

$$\dot{Z}_2 = Y_2 - Z_1 Z_2 \tag{S35d}$$

225   The point $E = (1,1,1,1)$ is a steady state for system (S35). The linearized dynamics about E is given

226   by:

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ z_1 \\ z_2 \end{bmatrix} \tag{S36}$$

148

227   As can be seen, $k_2 k_4 Z_1^* = k_1 k_3 Z_2^* = 1$.

228   Moreover:

$$H_p(s) = C_p(sI - A_p)^{-1} + D_p = \begin{bmatrix} \frac{1}{s+1} & 0 \\ 0 & \frac{1}{s+1} \end{bmatrix} \tag{S37}$$

229   which is analytic in $\mathbf{Re}[s] \geq 0$.

We also calculate:

$$H_p(j\omega) + H_p^H(j\omega) = \begin{bmatrix} \frac{2}{\omega^2+1} & 0 \\ 0 & \frac{2}{\omega^2+1} \end{bmatrix}$$

230   $H_p(j\omega) + H_p^H(j\omega) \succcurlyeq 0$ since $tr\big(H_c(j\omega) + H_c^H(j\omega)\big)$, $det\big(H_c(j\omega) + H_c^H(j\omega)\big) > 0$ for all ω.

231   Consequently, transfer function matrix (S37) is WSPR (see the respective definition in **Useful**

232   **mathematical concepts**). We therefore conclude that closed-loop system (S36) is asymptotically

233   stable. To confirm this, we compute the eigenvalues of its dynamics matrix: $-1.5 \pm j0.87$ and

234   $-0.5 \pm j0.87$ (they all have negative real parts).

235   Finally, note that in case we had R-Regulator with only one inhibitory reaction (either $Y_1 + Z_1 \xrightarrow{k_4} Z_1$

236   or $Y_2 + Z_2 \xrightarrow{k_3} Z_2$) we can immediately see from ODE model (S35) that one of the target

237   species - Y_2 or Y_1 respectively - would go to infinity since the corresponding derivative would al-

238   ways be positive.

239

S10 *In vivo* implementations

Practical considerations

242   Here we discuss some key challenges/limitations with respect to the experimental implementations in

243   *Escherichia coli* presented in Section 7 **Experimental realization** of the main text. It is worth noting

244   that the points raised below are also relevant to potential implementations in other types of organisms,

245   such as yeast or mammalian cells.

246   • Biochemical reactions of the form $A \xrightarrow{r} A + B$ are realized via gene expression processes.

247   Although this is a common approach, it is important to emphasize that it is valid only in a

248   specific regime due to the limited capacity of promoters [5, 8]. In particular, in the reaction

under consideration the formation rate of product B is proportional to the concentration of the regulator (activator) species A. In other words, the formation rate is a linear function of A, i.e. $\dfrac{dB}{dt} = rA$ ($r \in \mathbb{R}_+$). Taking into account the process of gene expression, this rate can be modelled via a Michaelis-Menten function of the form:

$$g(A) = V_{max} \frac{A}{A + K_m}$$

where V_{max}, $K_m \in \mathbb{R}_+$ is the maximal production rate and the Michaelis-Menten constant, respectively. Consequently, the system needs to operate in its first-order regime (linear range) which can be achieved if $K_m \gg A$. On the other hand, if saturation occurs, i.e. $K_m \ll A$, then $g(A) \approx V_{max}$ and the formation rate of B becomes effectively independent of A.

- Biochemical reactions of the form $A + B \xrightarrow{\text{r}} \varnothing$ (annihilation/antithetic reactions) correspond to processes where two species are able to bind to each other, forming an inert complex, i.e. $A + B \xrightarrow{\text{r}_1} A : B$. However, in reality the reverse reaction, $A : B \xrightarrow{\text{r}_2} A + B$ also takes place which can compromise the performance of the overall circuit if its rate is not sufficiently small.

  The actual *in vivo* binding/unbinding rates regarding the "antithetic pairs" used in the proposed experimental implementations are generally not well-defined in the literature. In fact, only the sigma/anti-sigma factors SigW/RsiW (in D-Regulator II) have been successfully tested in living cells for realizing antithetic integral feedback [5]. It therefore remains unclear if the rest of the "antithetic pairs" are suitable for this purpose in practice.

- The species of the annihilation/antithetic reactions ($A + B \xrightarrow{\text{r}} \varnothing$) are supposed to be lost solely due to these reactions. In case they participate in additional decay processes, then the performance of the overall circuit might be affected (see Section S8). Nevertheless, the presence of such decay mechanisms in living cells is, to some extent, unavoidable. A characteristic example is the phenomenon of dilution caused by cell growth [3, 5].

For a genetic circuit operating under non-ideal conditions, such as the above, appropriate parameter tuning is often required to achieve an acceptable performance. The most important feature that the latter entails is achieving sufficiently small output steady-state errors in the presence of disturbances.

An often convenient way to identify operating regimes that include this feature (assuming such operating regimes exist) is the following [5]: computing the differences between the output steady-states quantities of interest in the absence of a disturbance and the ones in the presence of a disturbance and, given the available parameter ranges, minimizing the former.

Realistic simulations

Taking into account the above considerations, in Figure S7 we successfully simulate the response of the genetic circuits depicted in Figure 7 (open-loop system) and Figure 9B (closed-loop system) of the main text under more realistic conditions - our results are aligned with the corresponding (ideal) ones of the main text. The simulations are based on the following ODE models and the parameters in Table S1.

Open-loop system

$$
\begin{aligned}
\dot{Y}_1 &= b_1 - (d + \gamma)Y_1 + V_{max}\frac{Y_2}{Y_2 + K_1} \\
\dot{Y}_2 &= b_2 - (d + \gamma)Y_2 + V_{max}\frac{Y_1}{Y_1 + K_2}
\end{aligned}
$$

Closed-loop system

$$
\begin{aligned}
\dot{Y}_1 &= b_1 - (d + \gamma)Y_1 + V_{max}\frac{Y_2}{Y_2 + K_1} + V_{max}\frac{Z_3}{Z_3 + K_3} \\
\dot{Y}_2 &= b_2 - (d + \gamma)Y_2 + V_{max}\frac{Y_1}{Y_1 + K_2} + V_{max}\frac{Z_4}{Z_4 + K_4} \\
\dot{Z}_1 &= V_{max}\frac{Y_1}{Y_1 + K_5} - \eta Z_1 Z_3 - \gamma Z_1 + k_u Z_5 \\
\dot{Z}_2 &= V_{max}\frac{Y_2}{Y_2 + K_6} - \eta Z_2 Z_4 - \gamma Z_2 + k_u Z_6 \\
\dot{Z}_3 &= \theta_1 - \eta_1 Z_1 Z_3 - \gamma Z_3 + k_u Z_5 \\
\dot{Z}_4 &= \theta_2 - \eta_2 Z_2 Z_4 - \gamma Z_4 + k_u Z_6 \\
\dot{Z}_5 &= \eta Z_1 Z_3 - k_u Z_5 - \gamma Z_5 \\
\dot{Z}_6 &= \eta Z_2 Z_4 - k_u Z_6 - \gamma Z_6
\end{aligned}
$$

where Z_5, Z_6 represent the complex $Z_1 : Z_3$, $Z_2 : Z_4$, respectively.

151

S11 Molecular programming realization of an abstract reaction

We describe how to implement a bimolecular reaction A + B $\rightleftharpoons$ C + D by DNA strand-displacement. Note that any finite chemical reaction network can be reduced to a collection of such reactions and their special cases. Here A, B, C, D are meant as abstract species and not as specific chemicals. These abstract species are then mapped to specially designed DNA molecules that implement the desired reaction kinetics by their interactions. That is, we are interested in representing the kinetics of a desired chemical reaction network by choosing (designing) the species involved, and not (directly) in manipulating existing chemicals.

Each of the A, B, C, D abstract species is represented (Figure S8) by a *3-domain*: a single-stranded DNA sequence logically subdivided into three *domains* (nucleotide subsequences), of which the middle one is short (≈ 6 bases) and the others are long (≈ 20 bases). Short domains are such that they bind reversibly to their Watson-Crick complements (indicated by *), while long domains bind irreversibly. A 3-domain is composed of a long *history* domain (left), which participated in past interactions but does not affect future interactions. Next is a short *toehold* domain, which is used to initiate interactions between 3-domains and *gates* that implement the reactions. Next is a long *identity* domain that is the one that identifies the chemical species (right). The same short sequence t can be used for all toehold occurrences, as successful bindings are determined by matching identity domains. However, different toehold can be chosen, for example, to fine tune reaction rates.

A gate is a double-stranded DNA structure that includes backbone breaks on the top strand; when two breaks or strand-ends are in close proximity, they form an open (i.e., single-stranded) toehold within the double-strand. A gate accepts 3-domains (the inputs to the reaction) that bind to its open toeholds, and through *strand displacement* releases other 3-domains (the outputs of the reaction). Strand displacement is a reversible random walk that starts at an open toehold and gradually replaces a domain with another identical domain within a double strand. At the end of the random walk, a whole single strand can detach from the double strand.

In summary, the species in a reaction networks can be uniquely assigned to domains (i.e., to specific sequences of nucleotides) and then a gate can be constructed for each desired reaction. The 3-domain structure is uniformly accepted and produced by the gates, so reactions can be composed.

S12 Regulating complex networks

We consider an open-loop biological network (network to be controlled) represented by the CRN (Figure S9A):

$$\varnothing \xrightarrow{b_1} Y_4 \,, \quad \varnothing \xrightarrow{b_2} Y_5 \,, \quad \varnothing \xrightarrow{b_3} Y_6 \,, \quad Y_4 \xrightarrow{\alpha_1} Y_4 + Y_1 \quad Y_5 \xrightarrow{\alpha_2} Y_5 + Y_2 \,,$$

$$Y_6 \xrightarrow{\alpha_3} Y_6 + Y_3 \,, \quad Y_3 + Y_4 \xrightarrow{\alpha_4} Y_3 \,, \quad Y_3 + Y_5 \xrightarrow{\alpha_5} Y_3 \quad Y_1 + Y_6 \xrightarrow{\alpha_6} Y_1 \,,$$

$$Y_2 + Y_6 \xrightarrow{\alpha_7} Y_2 \,, \quad Y_1 \xrightarrow{d_1} \varnothing \,, \quad Y_2 \xrightarrow{d_2} \varnothing \,, \quad Y_3 \xrightarrow{d_3} \varnothing \,, \quad Y_4 \xrightarrow{d_4} \varnothing \,,$$

$$Y_5 \xrightarrow{d_5} \varnothing \,, \quad Y_6 \xrightarrow{d_6} \varnothing$$

$$\tag{S40}$$
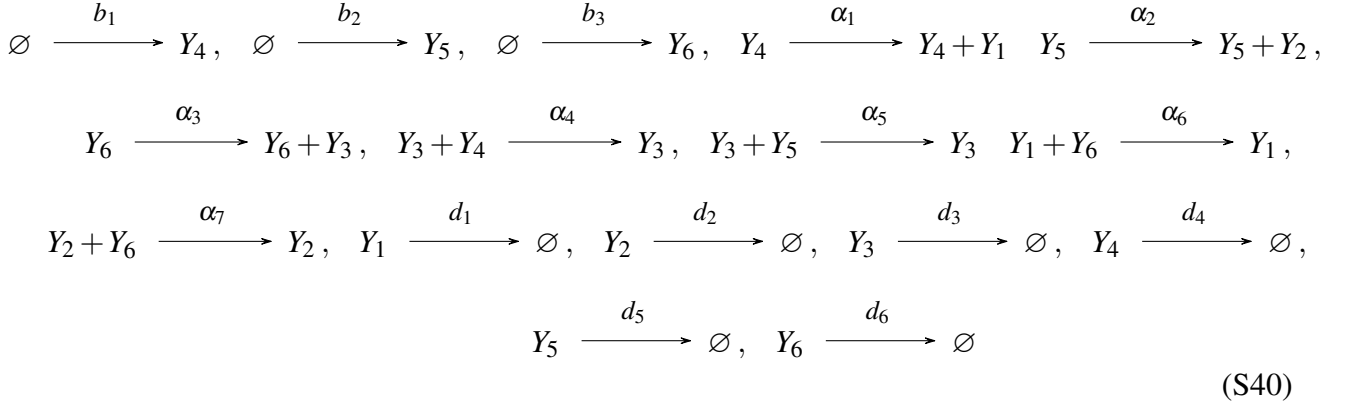
where b_1, b_2, b_3, α_1, α_2, α_3, α_4, α_5, α_6, α_7, d_1, d_2, d_3, d_4, d_5, $d_6 \in \mathbb{R}_+$. We treat Y_1, Y_2, Y_3 as the target species we aim to regulate.

The corresponding ODE model is :

$$\dot{Y}_1 = \alpha_1 Y_4 - d_1 Y_1 \tag{S41a}$$

$$\dot{Y}_2 = \alpha_2 Y_5 - d_2 Y_2 \tag{S41b}$$

$$\dot{Y}_3 = \alpha_3 Y_6 - d_3 Y_3 \tag{S41c}$$

$$\dot{Y}_4 = b_1 - d_4 Y_4 - \alpha_4 Y_3 Y_4 \tag{S41d}$$

$$\dot{Y}_5 = b_2 - d_5 Y_5 - \alpha_5 Y_3 Y_5 \tag{S41e}$$

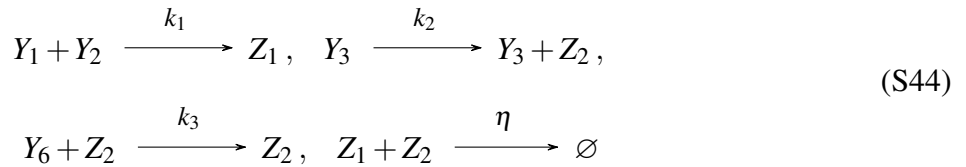$$\dot{Y}_6 = b_3 - d_6 Y_6 - \alpha_6 Y_1 Y_6 - \alpha_7 Y_2 Y_6 \tag{S41f}$$

Figure S9B shows the response of Y_1, Y_2, Y_3 and and how they are affected by an, arbitrarily chosen, disturbance applied on Y_5 (corresponding to an increase of its birth reaction rate).

We now discuss some examples of the controllers which can be build exploiting the regulation strategies introduced in this work. We also plot the output responses of the resulting closed-loop systems (based on CRN(S40)) considering the same disturbance as before.

R-Regulator

We have the controller CRN (Figure S10A):

$$Y_1 \xrightarrow{k_1} Y_1 + Z_1, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_1, \quad Y_3 \xrightarrow{k_3} Y_3 + Z_2,$$

$$Y_6 + Z_2 \xrightarrow{k_4} Z_2, \quad Z_1 + Z_2 \xrightarrow{\eta} \varnothing \tag{S42}$$
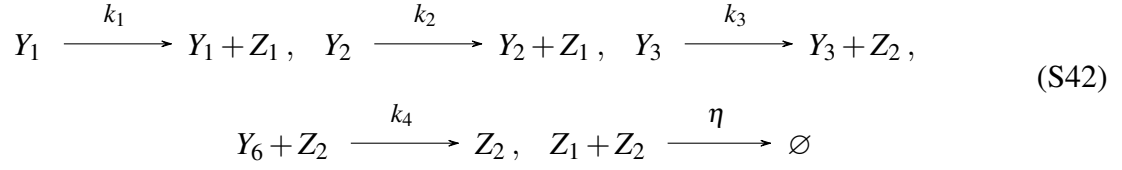
where $k_1, k_2, k_3, k_4, \eta \in \mathbb{R}_+$.

The closed-loop dynamics can be described by:

$$\dot{Y}_1 = \alpha_1 Y_4 - d_1 Y_1 \tag{S43a}$$

$$\dot{Y}_2 = \alpha_2 Y_5 - d_2 Y_2 \tag{S43b}$$

$$\dot{Y}_3 = \alpha_3 Y_6 - d_3 Y_3 \tag{S43c}$$

$$\dot{Y}_4 = b_1 - d_4 Y_4 - \alpha_4 Y_3 Y_4 \tag{S43d}$$

$$\dot{Y}_5 = b_2 - d_5 Y_5 - \alpha_5 Y_3 Y_5 \tag{S43e}$$

$$\dot{Y}_6 = b_3 - d_6 Y_6 - \alpha_6 Y_1 Y_6 - \alpha_7 Y_2 Y_6 - k_4 Y_6 Z_2 \tag{S43f}$$

$$\dot{Z}_1 = k_1 Y_1 + k_2 Y_2 - \eta Z_1 Z_2 \tag{S43g}$$

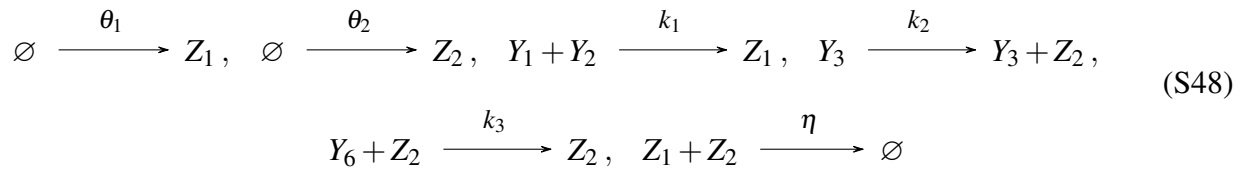$$\dot{Z}_2 = k_3 Y_3 - \eta Z_1 Z_2 \tag{S43h}$$

Steady-state behaviour (Figure S10B): $\dot{Z}_1 - \dot{Z}_2 = 0$ or $\dfrac{Y_1^* + \dfrac{k_2}{k_1} Y_2^*}{Y_3^*} = \dfrac{k_3}{k_1}$

A different version of this controller is given by the following CRN (Figure S11A):

$$Y_1 + Y_2 \xrightarrow{k_1} Z_1, \quad Y_3 \xrightarrow{k_2} Y_3 + Z_2,$$

$$Y_6 + Z_2 \xrightarrow{k_3} Z_2, \quad Z_1 + Z_2 \xrightarrow{\eta} \varnothing \tag{S44}$$

where $k_1, k_2, k_3, \eta \in \mathbb{R}_+$.

The closed-loop dynamics can be described by:

$$\dot{Y}_1 = \alpha_1 Y_4 - d_1 Y_1 - k_1 Y_1 Y_2 \qquad \text{(S45a)}$$

$$\dot{Y}_2 = \alpha_2 Y_5 - d_2 Y_2 - k_1 Y_1 Y_2 \qquad \text{(S45b)}$$

$$\dot{Y}_3 = \alpha_3 Y_6 - d_3 Y_3 \qquad \text{(S45c)}$$

$$\dot{Y}_4 = b_1 - d_4 Y_4 - \alpha_4 Y_3 Y_4 \qquad \text{(S45d)}$$

$$\dot{Y}_5 = b_2 - d_5 Y_5 - \alpha_5 Y_3 Y_5 \qquad \text{(S45e)}$$

$$\dot{Y}_6 = b_3 - d_6 Y_6 - \alpha_6 Y_1 Y_6 - \alpha_7 Y_2 Y_6 - k_3 Y_6 Z_2 \qquad \text{(S45f)}$$

$$\dot{Z}_1 = k_1 Y_1 Y_2 - \eta Z_1 Z_2 \qquad \text{(S45g)}$$

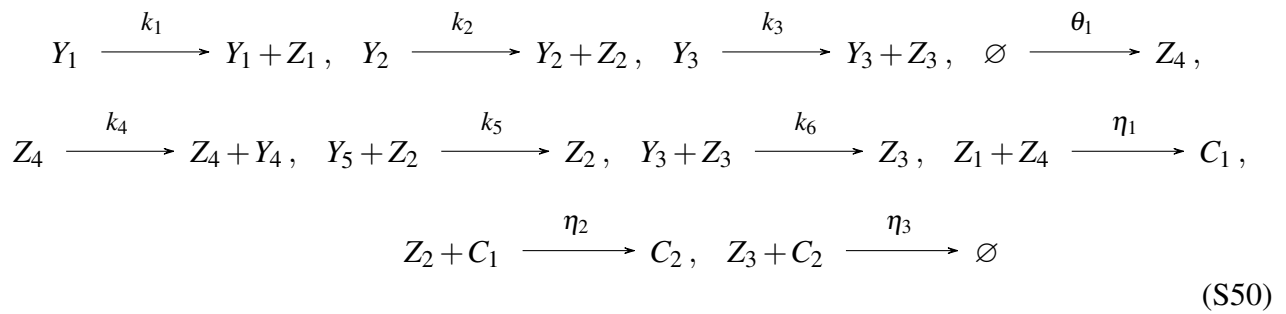$$\dot{Z}_2 = k_2 Y_3 - \eta Z_1 Z_2 \qquad \text{(S45h)}$$

$$\text{(S45i)}$$

Steady-state behaviour (Figure S11B): $\dot{Z}_1 - \dot{Z}_2 = 0$ or $\dfrac{Y_1^* Y_2^*}{Y_3^*} = \dfrac{k_2}{k_1}$

Note that the steady-state behaviour remains the same if we replace $Y_1 + Y_2 \xrightarrow{k_1} Z_1$ with

$Y_1 + Y_2 \xrightarrow{k_1} Y_1 + Y_2 + Z_1$ (catalytic production) in CRN (S44).

LC-Regulator

We have the controller CRN (Figure S12A):

$$\varnothing \xrightarrow{\theta_1} Z_1 \,, \quad \varnothing \xrightarrow{\theta_2} Z_2 \,, \quad Y_1 \xrightarrow{k_1} Y_1 + Z_1 \,, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_1 \,,$$
$$Y_3 \xrightarrow{k_3} Y_3 + Z_2 \,, \quad Y_6 + Z_2 \xrightarrow{k_4} Z_2 \,, \quad Z_1 + Z_2 \xrightarrow{\eta} \varnothing \qquad \text{(S46)}$$
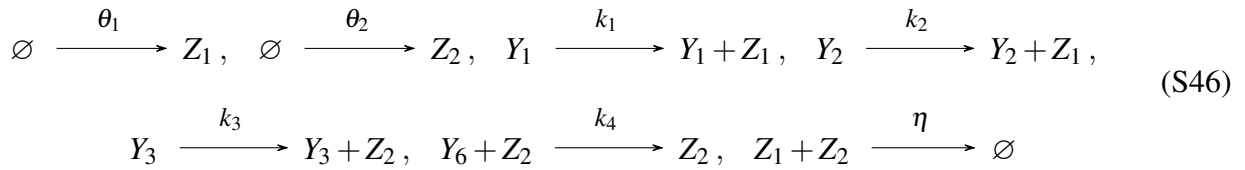
where $\theta_1, \theta_2, k_1, k_2, k_3, k_4, \eta \in \mathbb{R}_+$.

155

The closed-loop dynamics can be described by:

$$\dot{Y}_1 = \alpha_1 Y_4 - d_1 Y_1 \tag{S47a}$$

$$\dot{Y}_2 = \alpha_2 Y_5 - d_2 Y_2 \tag{S47b}$$

$$\dot{Y}_3 = \alpha_3 Y_6 - d_3 Y_3 \tag{S47c}$$

$$\dot{Y}_4 = b_1 - d_4 Y_4 - \alpha_4 Y_3 Y_4 \tag{S47d}$$

$$\dot{Y}_5 = b_2 - d_5 Y_5 - \alpha_5 Y_3 Y_5 \tag{S47e}$$

$$\dot{Y}_6 = b_3 - d_6 Y_6 - \alpha_6 Y_1 Y_6 - \alpha_7 Y_2 Y_6 - k_4 Y_6 Z_2 \tag{S47f}$$

$$\dot{Z}_1 = \theta_1 + k_1 Y_1 + k_2 Y_2 - \eta Z_1 Z_2 \tag{S47g}$$

$$\dot{Z}_2 = \theta_2 + k_3 Y_3 - \eta Z_1 Z_2 \tag{S47h}$$

Steady-state behaviour (Figure S12B): $\dot{Z}_1 - \dot{Z}_2 = 0$ or $k_1 Y_1 + k_2 Y_2 - k_3 Y_3 = \theta_2 - \theta_1$

A combination of R- and LC-Regulator

We have the controller CRN (Figure S13A):

$$\varnothing \xrightarrow{\theta_1} Z_1, \quad \varnothing \xrightarrow{\theta_2} Z_2, \quad Y_1 + Y_2 \xrightarrow{k_1} Z_1, \quad Y_3 \xrightarrow{k_2} Y_3 + Z_2,$$

$$Y_6 + Z_2 \xrightarrow{k_3} Z_2, \quad Z_1 + Z_2 \xrightarrow{\eta} \varnothing \tag{S48}$$

where θ_1, θ_2, k_1, k_2, k_3, $\eta \in \mathbb{R}_+$.

156

The closed-loop dynamics can be described by:

$$\dot{Y}_1 \;=\; \alpha_1 Y_4 - d_1 Y_1 - k_1 Y_1 Y_2 \tag{S49a}$$

$$\dot{Y}_2 \;=\; \alpha_2 Y_5 - d_2 Y_2 - k_1 Y_1 Y_2 \tag{S49b}$$

$$\dot{Y}_3 \;=\; \alpha_3 Y_6 - d_3 Y_3 \tag{S49c}$$

$$\dot{Y}_4 \;=\; b_1 - d_4 Y_4 - \alpha_4 Y_3 Y_4 \tag{S49d}$$

$$\dot{Y}_5 \;=\; b_2 - d_5 Y_5 - \alpha_5 Y_3 Y_5 \tag{S49e}$$

$$\dot{Y}_6 \;=\; b_3 - d_6 Y_6 - \alpha_6 Y_1 Y_6 - \alpha_7 Y_2 Y_6 - k_3 Y_6 Z_2 \tag{S49f}$$

$$\dot{Z}_1 \;=\; \theta_1 + k_1 Y_1 Y_2 - \eta Z_1 Z_2 \tag{S49g}$$

$$\dot{Z}_2 \;=\; \theta_2 + k_2 Y_3 - \eta Z_1 Z_2 \tag{S49h}$$

$$\tag{S49i}$$

Steady-state behaviour (Figure S13B): $\dot{Z}_1 - \dot{Z}_2 = 0$ or $k_1 Y_1 Y_2 - k_2 Y_3 = \theta_2 - \theta_1$.

Note that the steady-state behaviour remains the same if we replace $Y_1 + Y_2 \xrightarrow{k_1} Z_1$ with
$Y_1 + Y_2 \xrightarrow{k_1} Y_1 + Y_2 + Z_1$ (catalytic production) in CRN (S48).

D-Regulator III

We have the controller CRN (Figure S14A):

$$Y_1 \xrightarrow{k_1} Y_1 + Z_1 \,, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2 \,, \quad Y_3 \xrightarrow{k_3} Y_3 + Z_3 \,, \quad \varnothing \xrightarrow{\theta_1} Z_4 \,,$$

$$Z_4 \xrightarrow{k_4} Z_4 + Y_4 \,, \quad Y_5 + Z_2 \xrightarrow{k_5} Z_2 \,, \quad Y_3 + Z_3 \xrightarrow{k_6} Z_3 \,, \quad Z_1 + Z_4 \xrightarrow{\eta_1} C_1 \,,$$

$$Z_2 + C_1 \xrightarrow{\eta_2} C_2 \,, \quad Z_3 + C_2 \xrightarrow{\eta_3} \varnothing \tag{S50}$$

where $\theta_1, k_1, k_2, k_3, k_4, k_5, \eta_1, \eta_2, \eta_3 \in \mathbb{R}_+$.

The closed-loop dynamics can be described by:

$$\dot{Y}_1 = \alpha_1 Y_4 - d_1 Y_1 \tag{S51a}$$

$$\dot{Y}_2 = \alpha_2 Y_5 - d_2 Y_2 \tag{S51b}$$

$$\dot{Y}_3 = \alpha_3 Y_6 - d_3 Y_3 - k_6 Y_3 Z_3 \tag{S51c}$$

$$\dot{Y}_4 = b_1 - d_4 Y_4 - \alpha_4 Y_3 Y_4 + k_4 Z_4 \tag{S51d}$$

$$\dot{Y}_5 = b_2 - d_5 Y_5 - \alpha_5 Y_3 Y_5 - k_5 Y_5 Z_2 \tag{S51e}$$

$$\dot{Y}_6 = b_3 - d_6 Y_6 - \alpha_6 Y_1 Y_6 - \alpha_7 Y_2 Y_6 \tag{S51f}$$

$$\dot{Z}_1 = k_1 Y_1 - \eta_1 Z_1 Z_4 \tag{S51g}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta_2 Z_2 C_1 \tag{S51h}$$

$$\dot{Z}_3 = k_3 Y_3 - \eta_3 Z_3 C_2 \tag{S51i}$$

$$\dot{Z}_4 = \theta_1 - \eta_1 Z_1 Z_4 \tag{S51j}$$

$$\dot{C}_1 = \eta_1 Z_1 Z_4 - \eta_2 Z_2 C_1 \tag{S51k}$$

$$\dot{C}_2 = \eta_2 Z_2 C_1 - \eta_3 Z_3 C_2 \tag{S51l}$$

Steady-state behaviour (Figure S14B): $\dot{Z}_1 - \dot{Z}_4 = 0$ or $Y_1^* = \dfrac{\theta_1}{k_1}$, $\dot{Z}_4 + \dot{C}_1 - \dot{Z}_2 = 0$ or $Y_2^* = \dfrac{\theta_1}{k_2}$, $\dot{Z}_4 + \dot{C}_1 + \dot{C}_2 - \dot{Z}_3 = 0$ or $Y_3^* = \dfrac{\theta_1}{k_3}$.

Constructing control schemes based on D-Regulator I, II (see Section 3 **Control schemes with steady-state decoupling** of the main text) and *Rein* D-Regulator (see Section S3) requires three SISO control loops and, thus, it is quite straightforward. Note though that for this approach we would need, at least, 6 controller species - two more compared to D-Regulator III.
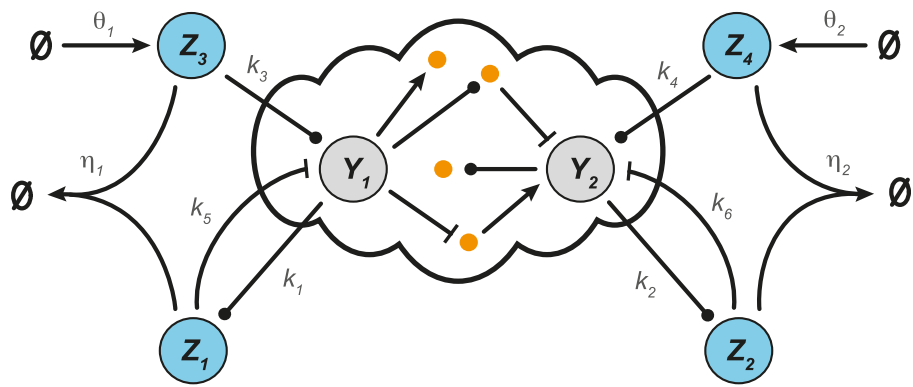
Figure S1: *Rein* **D-Regulator**
Schematic representation of a general closed-loop architecture based on the D-Regulator described in Section S3.

159
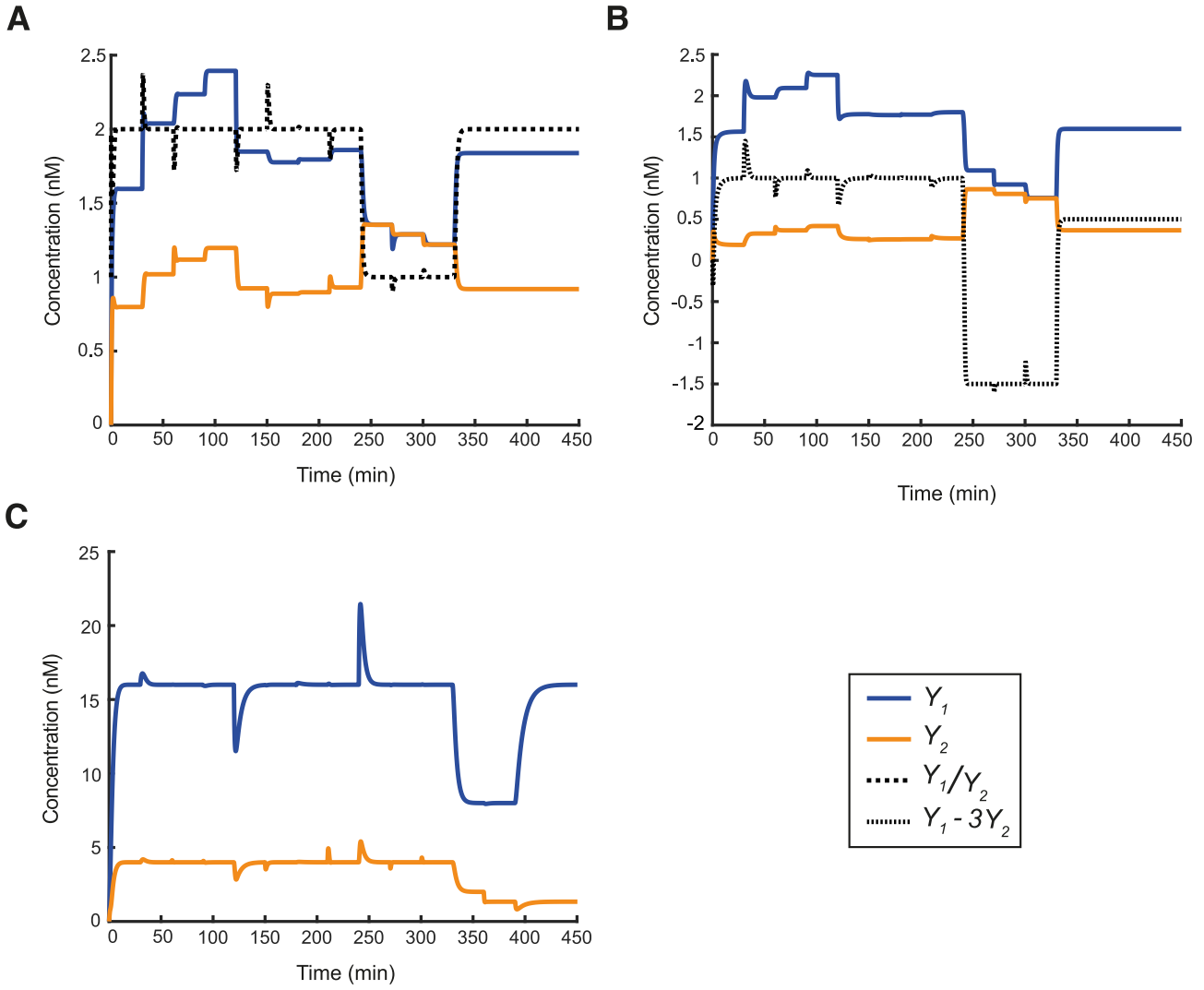
Figure S2: Behaviour in the presence of multiple parameter perturbations.
A Simulated response of R-Regulator presented in Section S7 using the following parameters: $b_1 = 2$ nM min^{-1}, $b_2 = 1$ nM min^{-1}, $\eta = 10$ nM^{-1} min^{-1}, $d_1 = 1$ min^{-1}, $d_2 = 1$ min^{-1}, $\alpha_1 = 0.1$ min^{-1}, $\alpha_2 = 0.4$ min^{-1}, $k_2 = 1$ min^{-1}, $k_4 = 2$ nM^{-1} min^{-1}, $k_3 = 2$ nM^{-1} min^{-1}, $k_1 = 0.5$ min^{-1}. Every 30 min, one of these parameters is pertrubed by 50% (in the order they appear above). $\dfrac{Y_1^*}{Y_2^*} = \dfrac{k_2}{k_1}$ always holds. **B** Simulated response of LC-Regulator presented in Section S7 using the following parameters: $b_1 = 2$ nM min^{-1}, $b_2 = 1$ nM min^{-1}, $\eta = 10$ nM^{-1} min^{-1}, $d_1 = 1$ min^{-1}, $d_2 = 1$ min^{-1}, $\alpha_1 = 0.1$ min^{-1}, $\alpha_2 = 0.4$ min^{-1}, $\theta_2 = 5$ nM min^{-1}, $k_4 = 2$ nM^{-1} min^{-1}, $k_3 = 2$ nM^{-1} min^{-1}, $\theta_1 = 4$ nM min^{-1}, $k_1 = 1$ min^{-1}, $k_2 = 3$ min^{-1}. Every 30 min, one of these parameters (apart from k_1, k_2) is pertrubed by 50% (in the order they appear above). $k_1 Y_1^* - k_2 Y_2^* = \theta_2 - \theta_1$ always holds. **C** Simulated response of D-Regulator-III presented in Section S7 using the following parameters: $b_1 = 2$ nM min^{-1}, $b_2 = 1$ nM min^{-1}, $\eta_1 = 0.5$ nM^{-1} min^{-1}, $d_1 = 1$ min^{-1}, $d_2 = 1$ min^{-1}, $\alpha_1 = 0.1$ min^{-1}, $\alpha_2 = 0.4$ min^{-1}, $k_3 = 0.5$ min^{-1}, $k_4 = 2$ nM^{-1} min^{-1}, $\eta_2 = 10$ nM^{-1} min^{-1}, $\theta_1 = 8$ nM min^{-1}, $k_2 = 2$ min^{-1}, $k_1 = 0.5$ min^{-1}. Every 30 min, one of these parameters is pertrubed by 50% (in the order they appear above). $Y_1^* = \dfrac{\theta_1}{k_1}$ nM, $Y_2^* = \dfrac{\theta_1}{k_1}$ nM always hold.
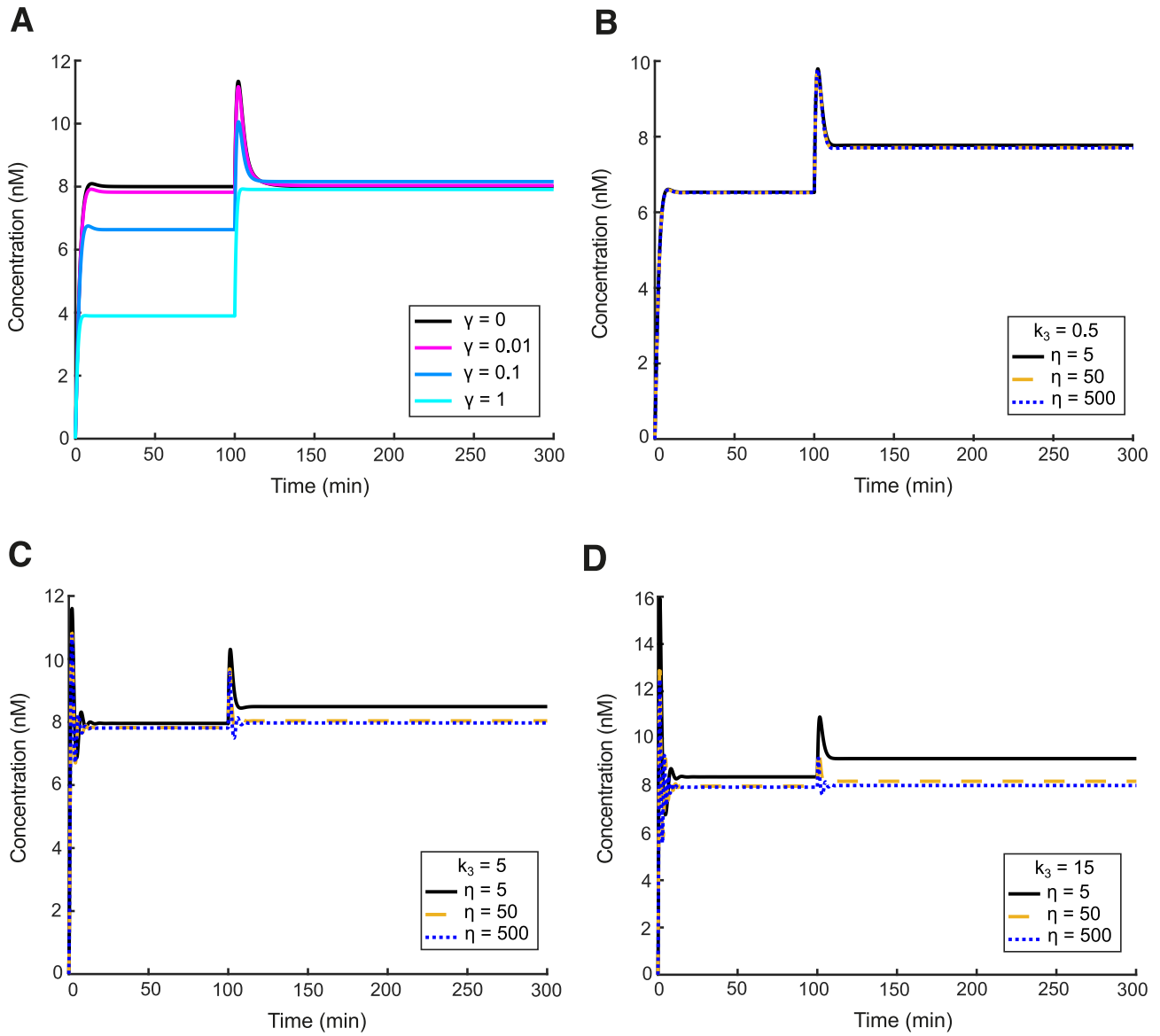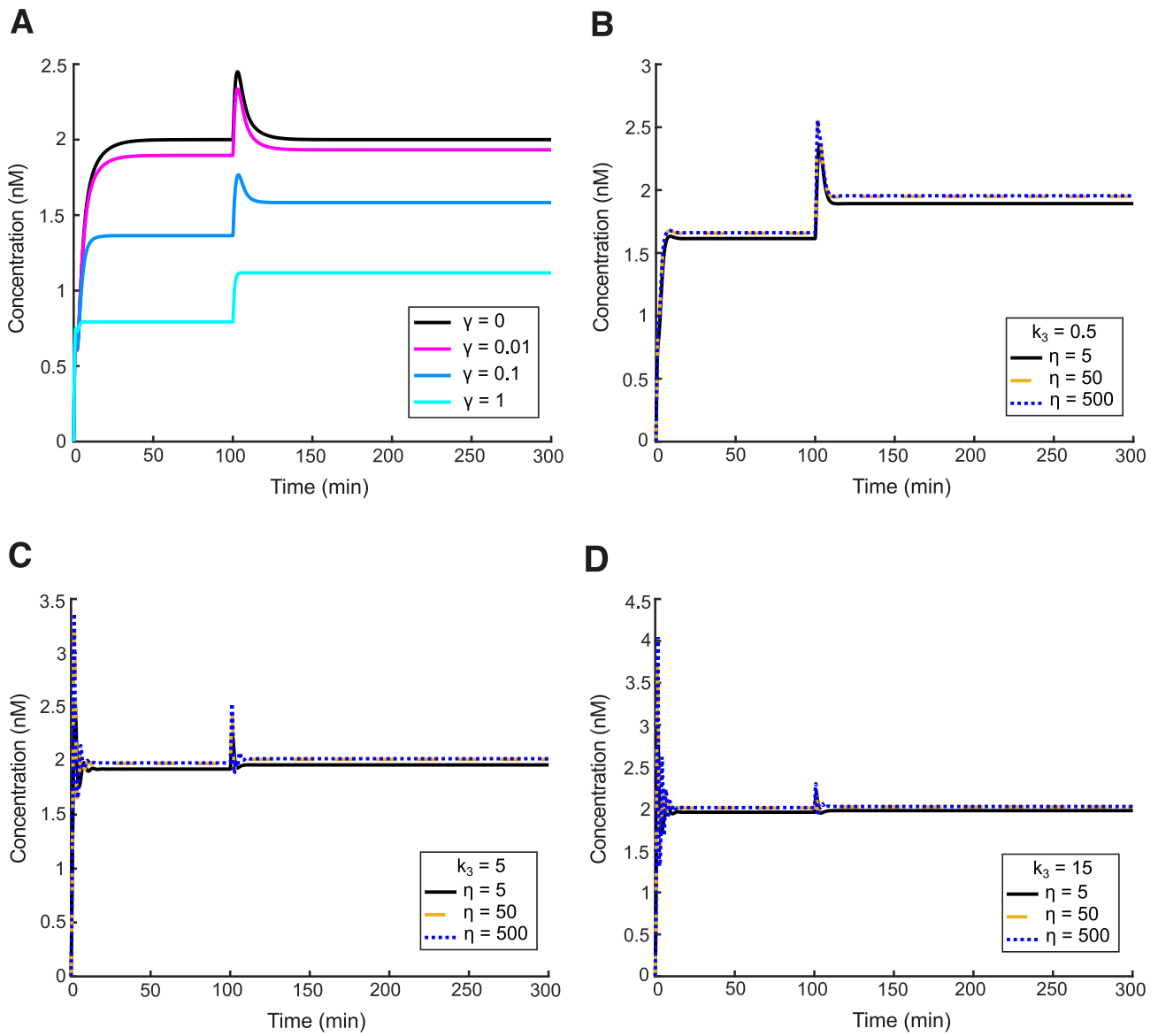
160
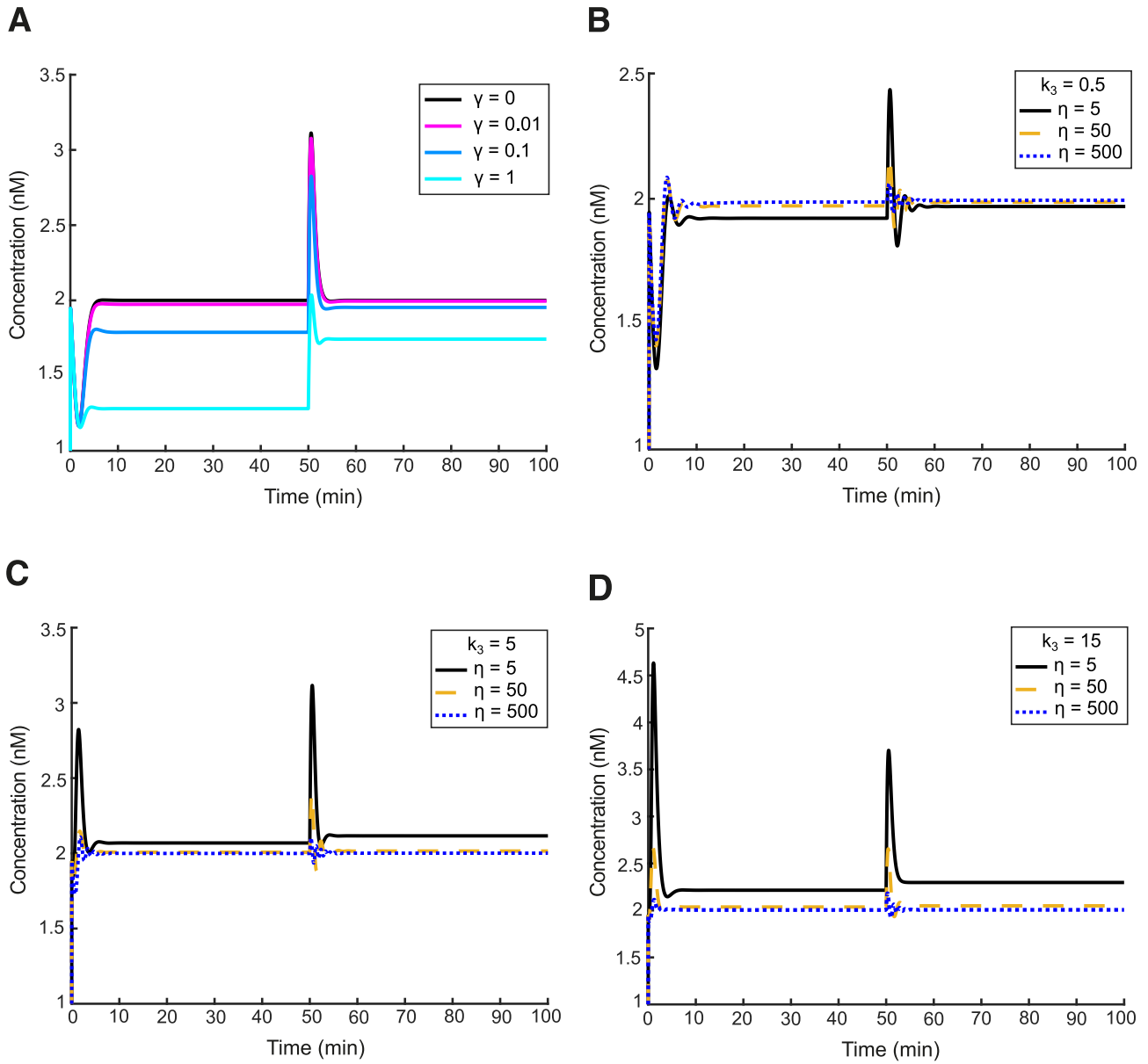
Figure S3: Controller species degradation and adaptation: R-Regulator.
Simulated response of $\frac{Y_1}{Y_2}$ with respect to R-Regulator presented in Section S8 with **A** $b_1 = 2$ nM min^{-1}, $b_2 = 1$ nM min^{-1}, $d_1 = 0.1$ min^{-1}, $d_2 = 0.1$ min^{-1}, $\alpha_1 = 0.1$ min^{-1}, $\alpha_2 = 0.4$ min^{-1}, $k_1 = 1$ min^{-1}, $k_2 = 2$ min^{-1}, $k_3 = 0.5$ nM^{-1} min^{-1}, $k_4 = 1$ nM^{-1} min^{-1}, $\eta = 0.5$ nM^{-1} min^{-1} while γ varying as shown. **B**, **C**, **D** $\gamma = 0.1$ min^{-1}, k_3 and η varying as shown while the rest of the parameters remaining the same as in **A**. In all the above simulations, a disturbance is introduced at time $t = 50$ min in the form of an increase regarding parameter b_1, i.e. its value changes from 2 to 6.

161
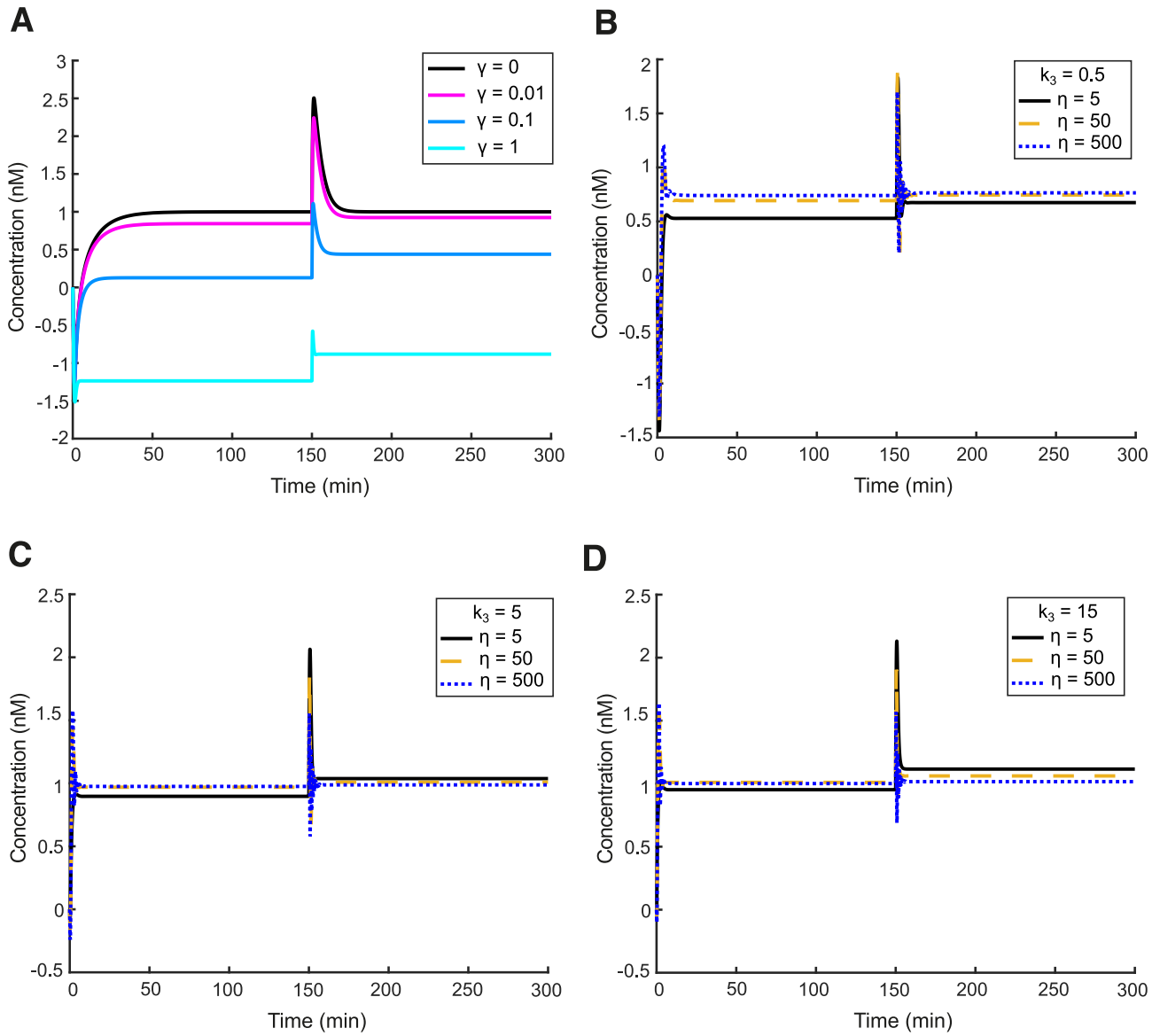
Figure S4: Controller species degradation and adaptation: LC-Regulator.
Simulated response of $Y_1 - 3Y_2$ with respect to LC-Regulator presented in Section S8 with **A** $b_1 = 2$ nM min^{-1}, $b_2 = 1$ nM min^{-1}, $d_1 = 0.1$ min^{-1}, $d_2 = 0.1$ min^{-1}, $\alpha_1 = 0.1$ min^{-1}, $\alpha_2 = 0.4$ min^{-1}, $k_1 = 1$ min^{-1}, $k_2 = 3$ min^{-1}, $k_3 = 0.5$ nM^{-1} min^{-1}, $k_4 = 2$ nM^{-1} min^{-1}, $\eta = 0.5$ nM^{-1} min^{-1}, $\theta_1 = 4$ nM min^{-1}, $\theta_2 = 5$ nM min^{-1} while γ varying as shown. **B**, **C**, **D** $\gamma = 0.1$ min^{-1}, k_3 and η varying as shown while the rest of the parameters remaining the same as in **A**. In all the above simulations, the same disturbance as in Figure S3 is introduced at time $t = 150$ min.

Figure S5: Controller species degradation and adaptation: output species Y_1 of D-Regulator-III.
Simulated response of output species Y_1 with respect to D-Regulator-III presented in Section S8 with **A** $b_1 = 2$ nM min^{-1}, $b_2 = 1$ nM min^{-1}, $d_1 = 0.9$ min^{-1}, $d_2 = 0.9$ min^{-1}, $\alpha_1 = 0.1$ min^{-1}, $\alpha_2 = 0.4$ min^{-1}, $k_1 = 0.5$ min^{-1}, $k_2 = 2$ min^{-1}, $k_3 = 0.5$ min^{-1}, $k_4 = 2$ nM^{-1} min^{-1}, $\eta = \eta_1 = \eta_2 = 0.5$ nM^{-1} min^{-1}, $\theta_1 = 4$ nM min^{-1} while γ varying as shown. **B**, **C**, **D** $\gamma = 0.1$ min^{-1}, k_3 and η varying as shown while the rest of the parameters remaining the same as in **A**. In all the above simulations, the same disturbance as in Figure S3 is introduced at time $t = 100$ min.

Figure S6: Controller species degradation and adaptation: output species Y_2 of D-Regulator-III.
Simulated response of output species Y_2 with respect to D-Regulator-III presented in Section S8 following the exact same concept as in Figure S5.
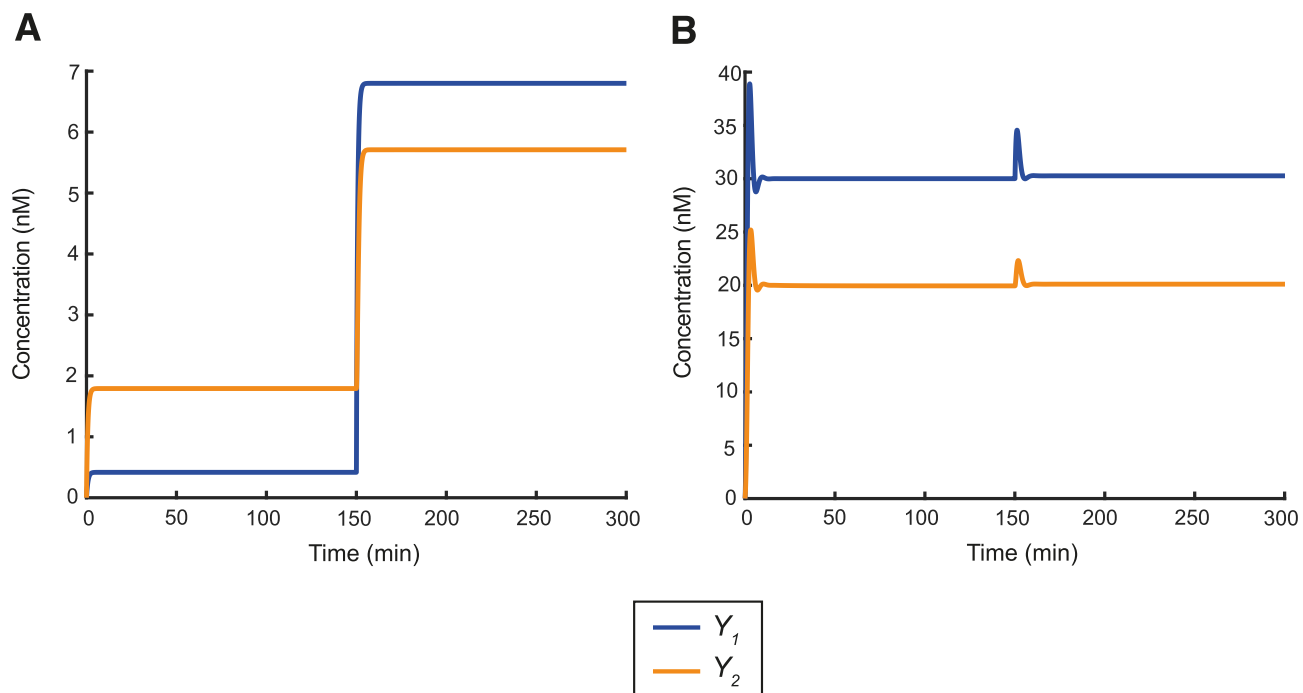
Figure S7: Realistic simulations.
Simulated response of the **A** open-loop system **B** closed-loop system considered in Section S10 using the parameters of Table S1. In both simulations, a disturbance is introduced at time $t = 150$ min in the form of an increase regarding parameter b_1, i.e. its value changes from 0.5 to 10.5 .
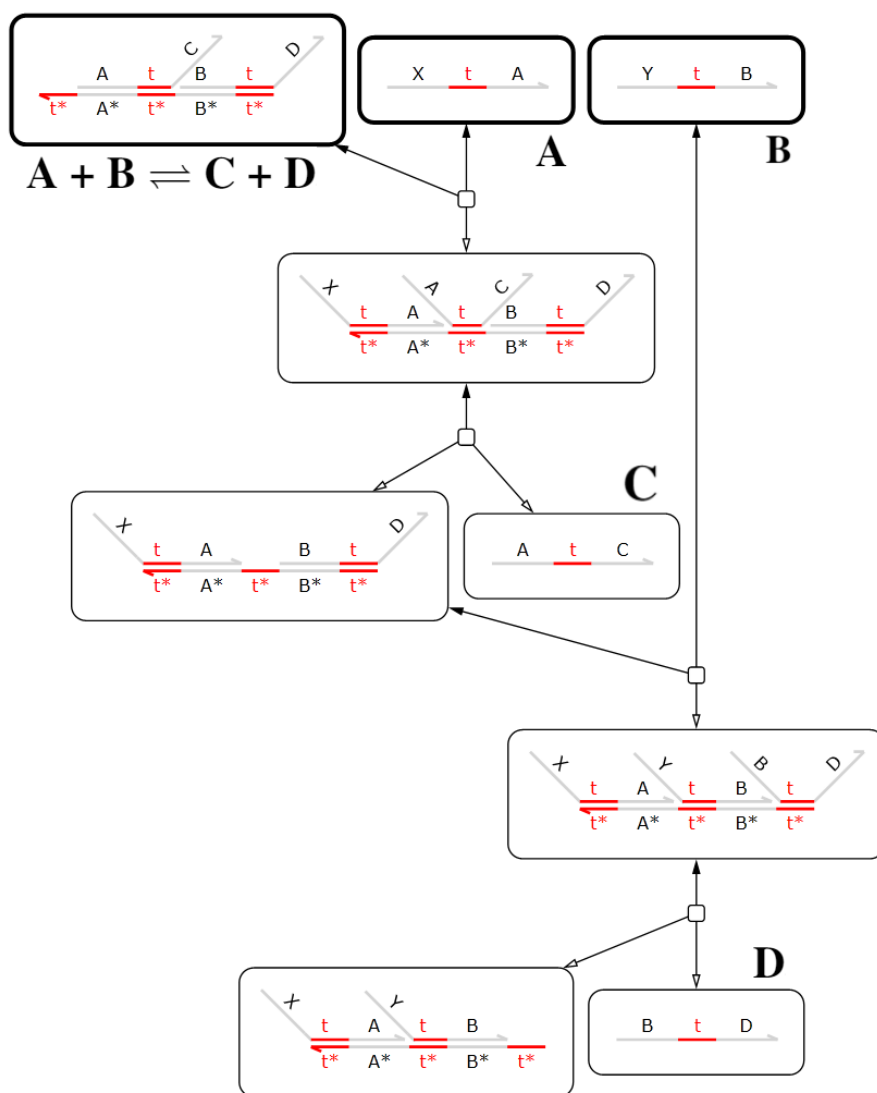
Figure S8: DNA strand-displacement representation of the reaction A + B ⇌ C + D. A, B, C, D and t are unique DNA sequences chosen not to bind to each other. The initial DNA structures are indicated by a boldface border: these are the single-stranded inputs marked A and B and the roughly double-stranded "gate" structure marked A + B ⇌ C + D; the intended outputs are the structures marked C and D. The graph details the chemical interactions that happen between these DNA structures. Reactions between DNA structures (small squares) have hollow heads for direct reactions and filled heads for reverse reactions. A, B, C, D need not be distinct in this scheme, i.e., a reaction like A + B ⇌ 2A would work as expected. The A + B ⇌ C + D reaction described above is reversible: the outputs can bind back through the open toeholds on the right. However, it is easy to convert this to an irreversible A + B → C + D reaction by attaching a double stranded domain to the right of the gate (not shown), with an auxiliary single strand that irreversibly binds to the right toehold once it is exposed and to the new domain, preventing the outputs from binding back to the gate since no open toeholds are left.
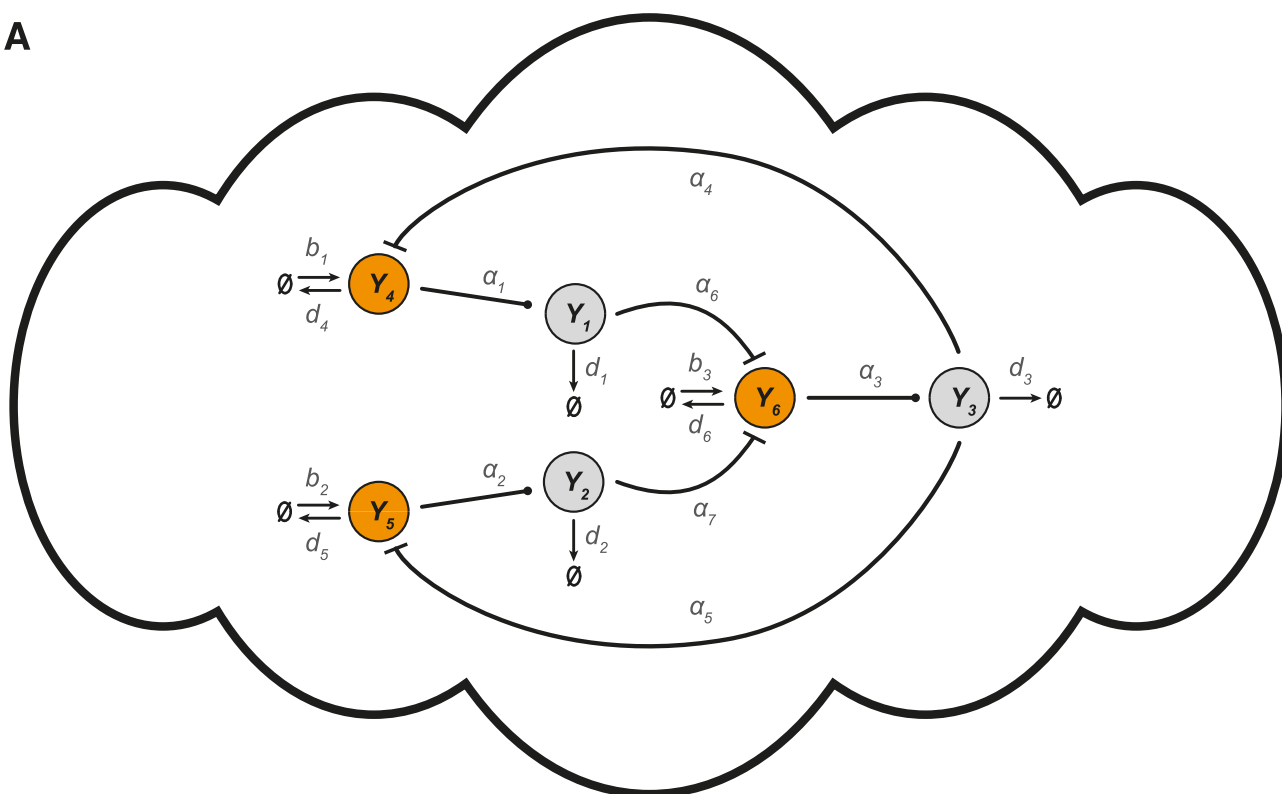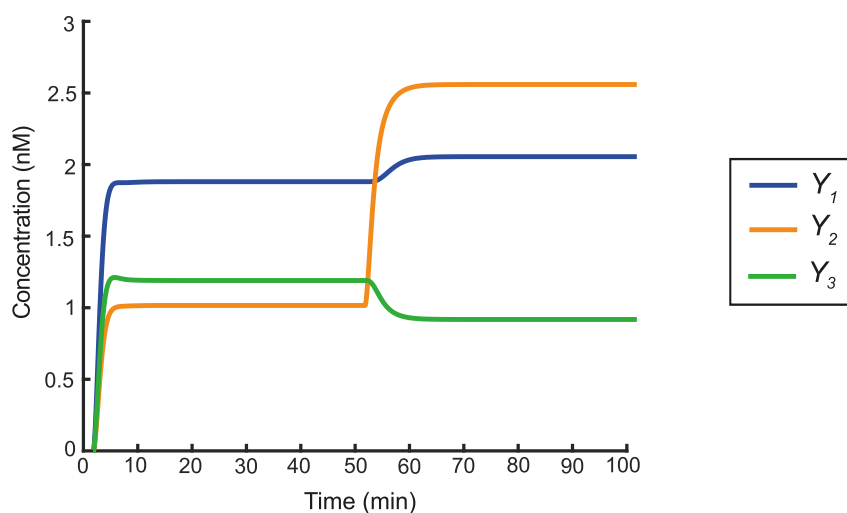
Figure S9: A three-output open-loop biomolecular network.
A Schematic representation of the network to be controlled described by CRN (S40). Y_1, Y_2, Y_3 are considered the output species of interest. **B** Simulated response of the topology in **A** using the ODE model (S41) with the following parameters: $b_1 = 2$ nM min^{-1}, $b_2 = 1.5$ nM min^{-1}, $b_3 = 1$ nM min^{-1}, $d_1 = d_2 = d_3 = d_4 = d_5 = d_6 = 1$ min^{-1}, $\alpha_1 = 1.5$ min^{-1}, $\alpha_2 = 1$ min^{-1}, $\alpha_3 = 2$ min^{-1}, $\alpha_4 = 0.5$ nM^{-1} min^{-1}, $\alpha_5 = 0.4$ nM^{-1} min^{-1}, $\alpha_6 = 0.2$ nM^{-1} min^{-1}, $\alpha_7 = 0.3$ nM^{-1} min^{-1}. At time $t = 50$ min, a disturbance is introduced in the form of an increase regarding parameter b_2, i.e. its value changes from 1.5 to 3.5 .
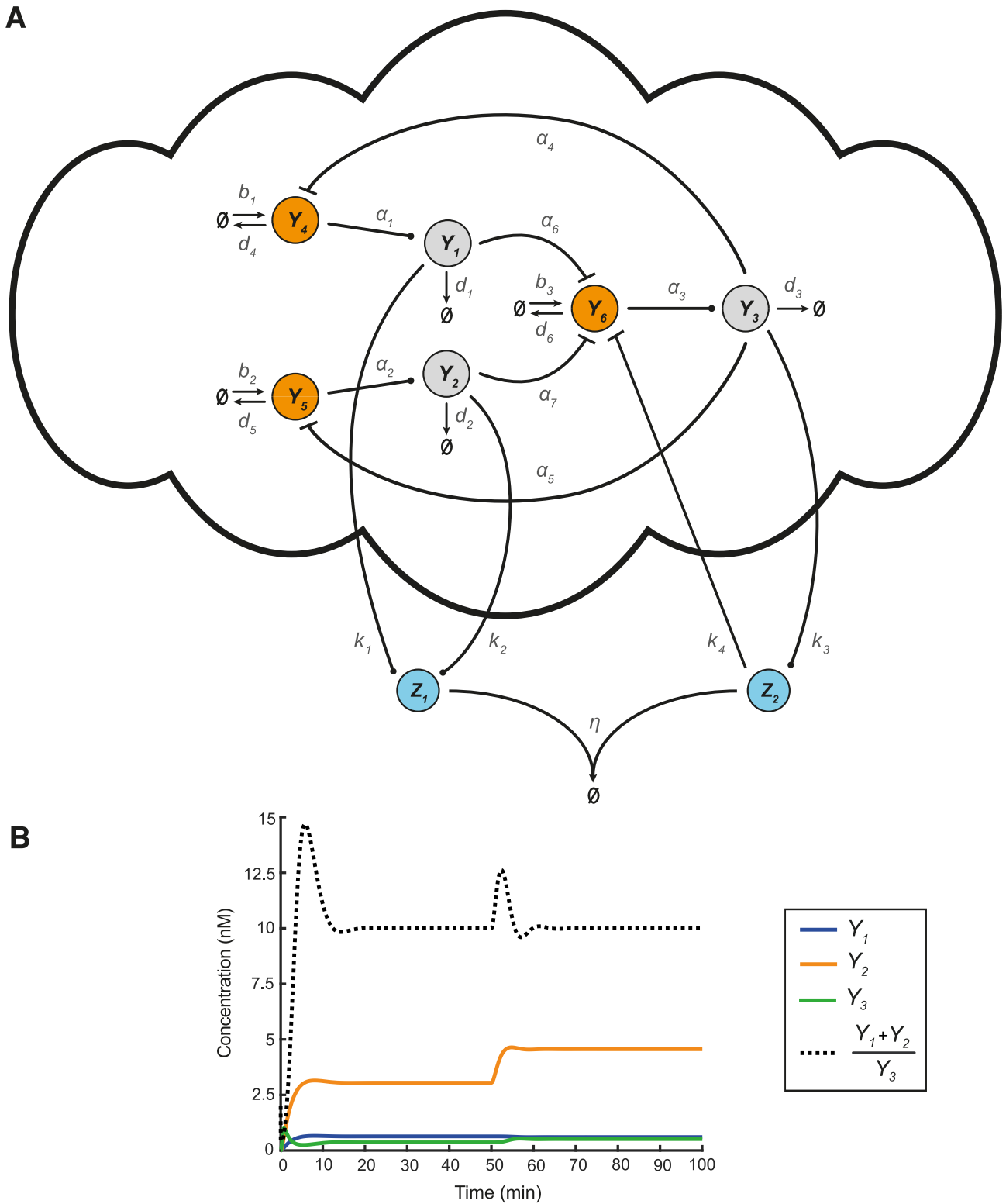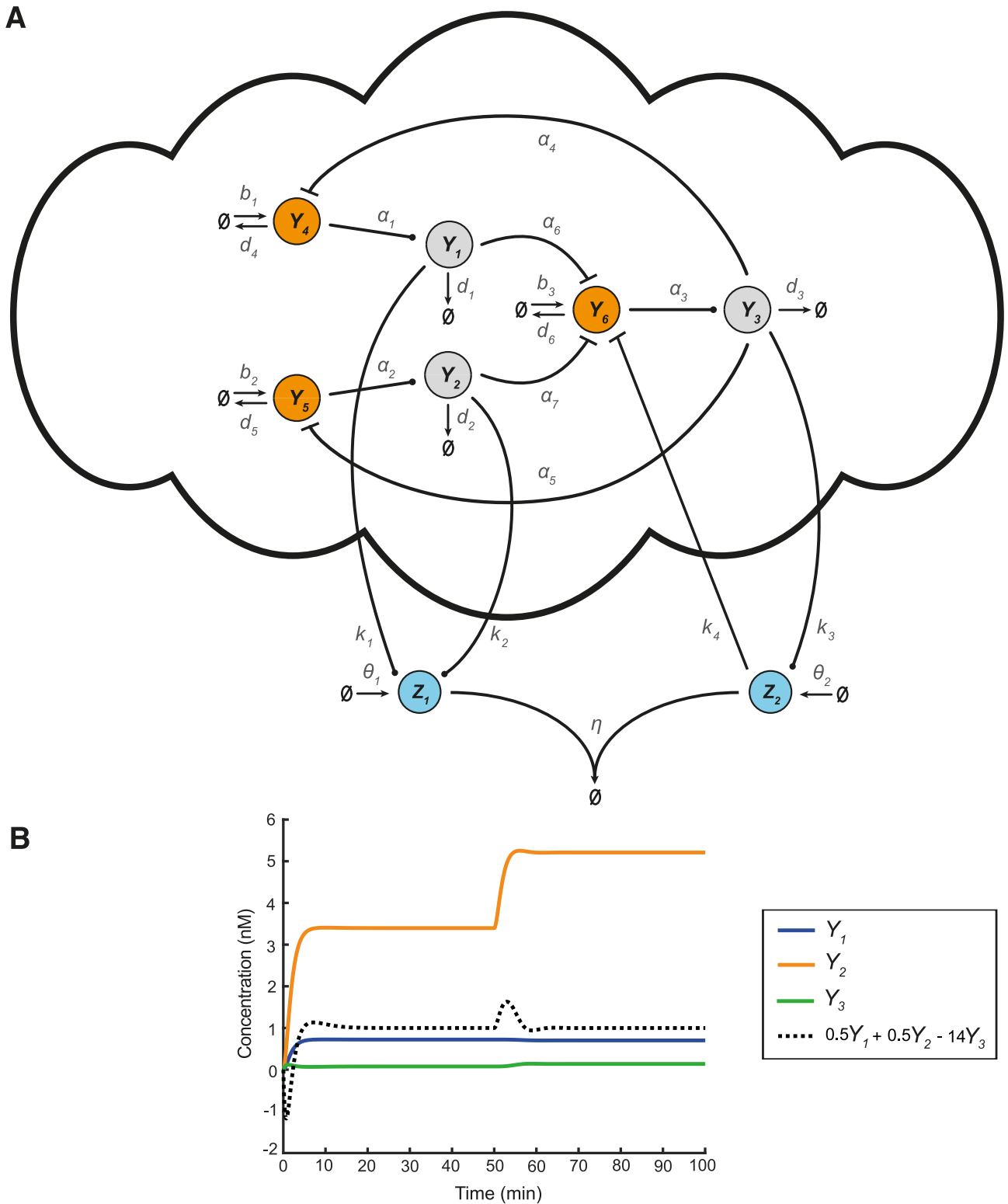
Figure S10: R-Regulator.
A Schematic representation of a closed-loop architecture based on the network to be controlled shown in Figure S9 and R-Regulator described by CRN (S42). **B** Simulated response of the topology in **A** using the ODE model (S43) with the following parameters: $k_1 = 0.5$ min^{-1}, $k_2 = 0.5$ min^{-1}, $k_3 = 4$ min^{-1}, $k_4 = 6$ nM^{-1} min^{-1}, $\eta = 10$ nM^{-1} min^{-1} while the rest of the parameters as well as the type of the disturbance introduced (including the time of entry) remain the same as in Figure S9.
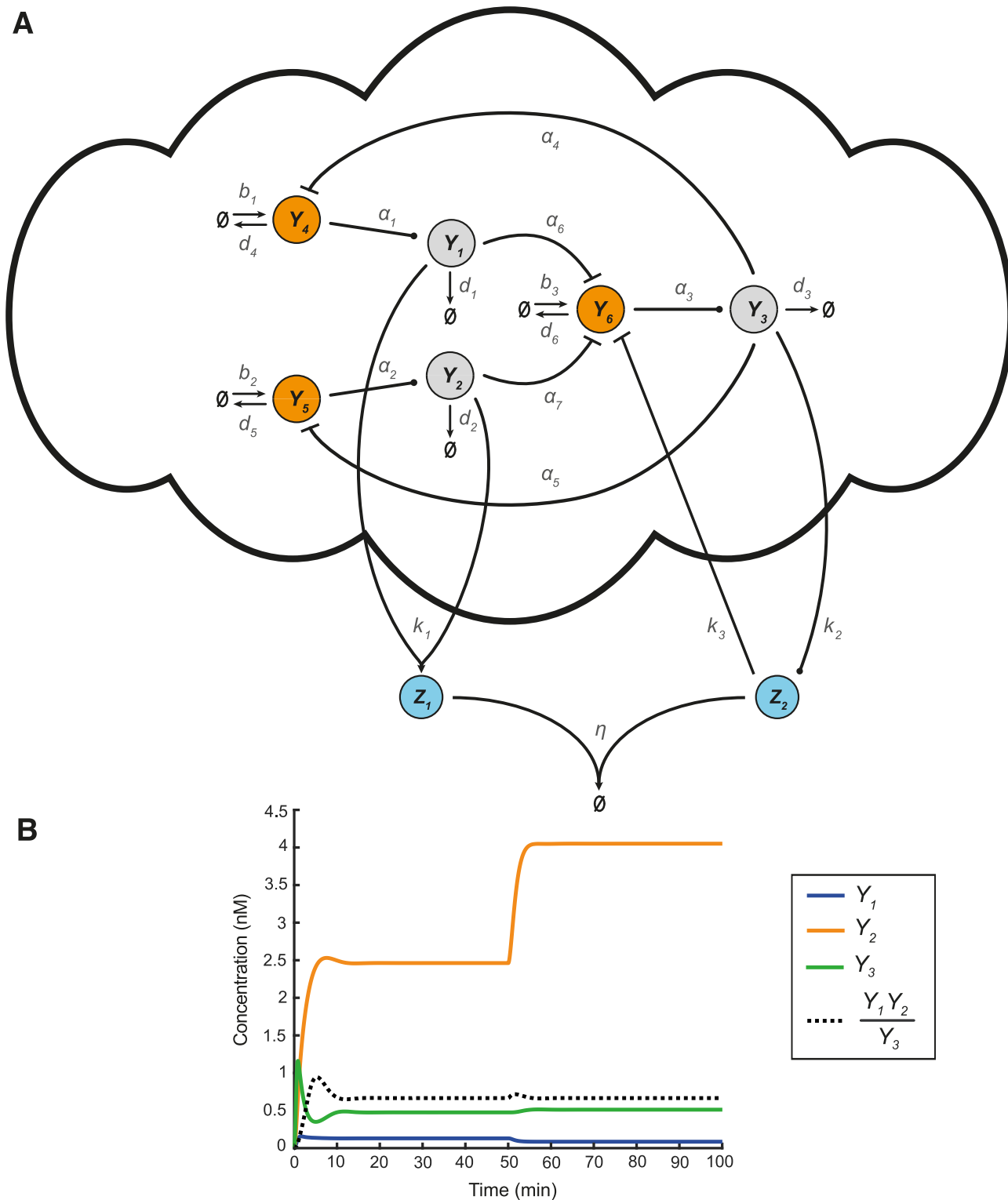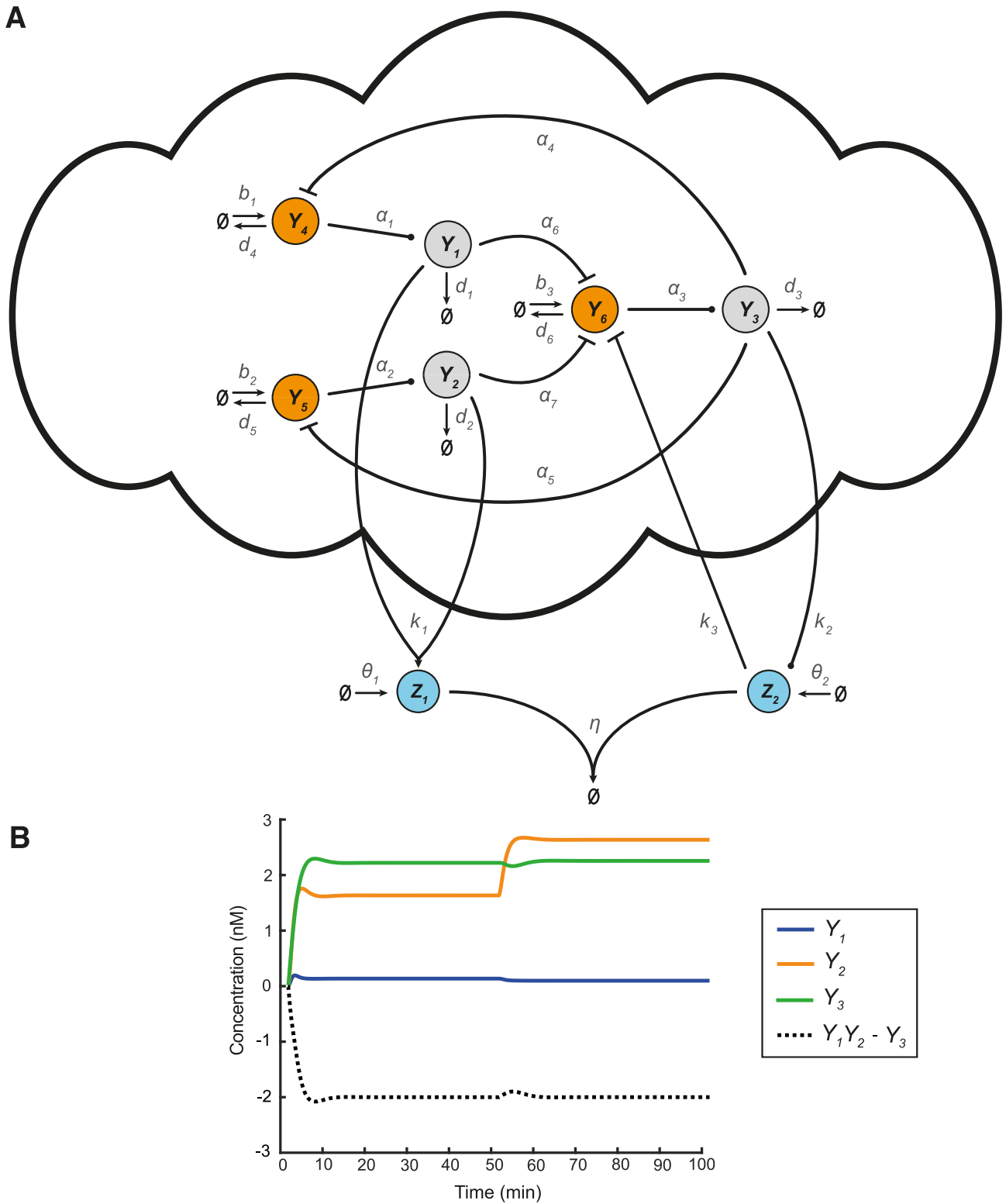
Figure S11: An alternative version of R-Regulator.
A Schematic representation of a closed-loop architecture based on the network to be controlled shown in Figure S9 and R-Regulator described by CRN (S44). **B** Simulated response of the topology in **A** using the ODE model (S45) with the following parameters: $k_1 = 1.5$ nM^{-1} min^{-1}, $k_2 = 1$ min^{-1}, $k_3 = 20$ nM^{-1} min^{-1}, $\eta = 10$ nM^{-1} min^{-1} while the rest of the parameters as well as the type of the disturbance introduced (including the time of entry) remain the same as in Figure S9.

Figure S12: LC-Regulator.
A Schematic representation of a closed-loop architecture based on the network to be controlled shown in Figure S9 and R-Regulator described by CRN (S46). **B** Simulated response of the topology in **A** using the ODE model (S47) with the following parameters: $k_1 = 0.5$ nM^{-1} min^{-1}, $k_2 = 0.5$ min^{-1}, $k_3 = 14$ nM^{-1} min^{-1}, $k_4 = 30$ nM^{-1} min^{-1}, $\eta = 10$ nM^{-1} min^{-1}, $\theta_1 = 15$ nM min^{-1}, $\theta_2 = 16$ nM min^{-1} while the rest of the parameters as well as the type of the disturbance introduced (including the time of entry) remain the same as in Figure S9.
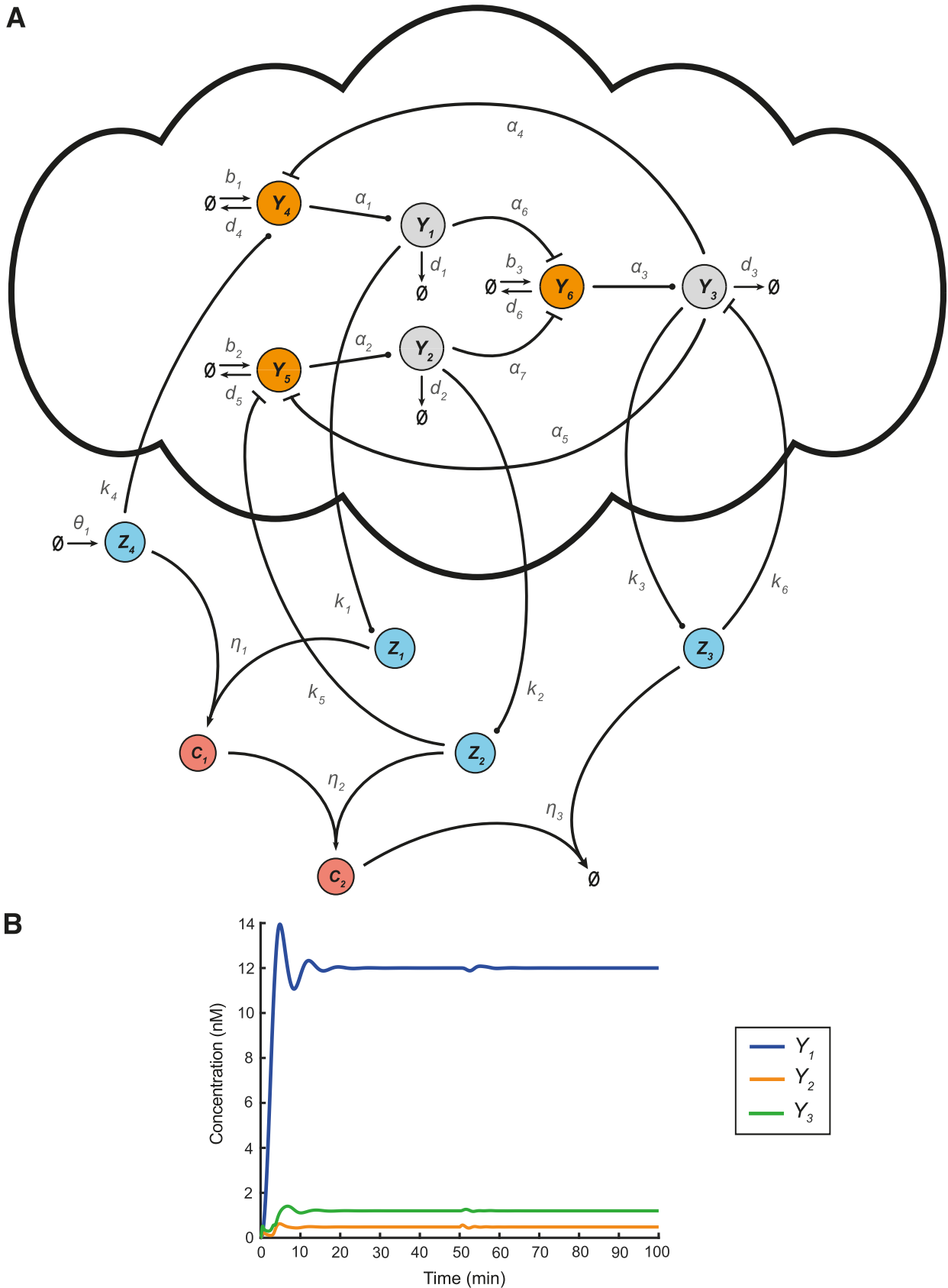
Figure S13: A combination of R- and LC-Regulator.
A Schematic representation of a closed-loop architecture based on the network to be controlled shown in Figure S9 and R-Regulator described by CRN (S48). **B** Simulated response of the topology in **A** using the ODE model (S49) with the following parameters: $k_1 = 1$ nM^{-1} min^{-1}, $k_2 = 1$ min^{-1}, $k_3 = 10$ nM^{-1} min^{-1}, $\eta = 10$ nM^{-1} min^{-1}, $\theta_1 = 10$ nM min^{-1}, $\theta_2 = 8$ nM min^{-1} while the rest of the parameters as well as the type of the disturbance introduced (including the time of entry) remain the same as in Figure S9.

Figure S14: D-Regulator-III.
A Schematic representation of a closed-loop architecture based on the network to be controlled shown in Figure S9 and R-Regulator described by CRN (S50). **B** Simulated response of the topology in **A** using the ODE model (S51) with the following parameters: $k_1 = 1$ min^{-1}, $k_2 = 25$ min^{-1}, $k_3 = 10$ min^{-1}, $k_4 = 0.5$ min^{-1}, $k_5 = 10$ nM^{-1} min^{-1}, $k_6 = 5$ nM^{-1} min^{-1}, $\eta_1 = 10$ nM^{-1} min^{-1}, $\eta_2 = 10$ nM^{-1} min^{-1}, $\eta_3 = 10$ nM^{-1} min^{-1}, $\theta_1 = 12$ nM min^{-1} while the rest of the parameters as well as the type of the disturbance introduced (including the time of entry) remain the same as in Figure S9.

Description	Parameter value	Unit	Comments	Source
Maximal production rate	$V_{max} = 10^4$	nM min^{-1}		[5]
Constitutive production rate	$b_1 = 0.5$, $b_2 = 2.5$, $\theta_1 = 60$, $\theta_2 = 35$	nM min^{-1}		[5]
Michaelis-Menten constant	$K_1 = 10^5$, $K_2 = 10^4$, $K_3 = 10^4$, $K_4 = 10^5$, $K_5 = 5 \cdot 10^3$, $K_6 = 2.5 \cdot 10^3$	nM		[5]
Dilution rate	$\gamma = 0.028$	min^{-1}	Assuming 25min doubling time in bacterial growth	[5]
Degradation rate	$d = 1.6$	min^{-1}	Unspecified degradation mechanism (disturbance); value based on the action of a protease	[9]
Binding rate	$\eta = 0.05$	nM^{-1} min^{-1}		[5]
Unbinding rate	$k_u = 0.0096$	min^{-1}		[10]

Table S1: Simulation parameters for the ODE models in Section S10.

References

[1]   Domitilla Del Vecchio and Richard M Murray. *Biomolecular feedback systems*. Princeton University Press Princeton, NJ, 2015.

[2]   Ankit Gupta and Mustafa Khammash. "An antithetic integral rein controller for bio-molecular networks". In: *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE. 2019, pp. 2808–2813.

[3]   Yili Qian and Domitilla Del Vecchio. "Realizing 'integral control' in living cells: how to overcome leaky integration due to dilution?" In: *Journal of The Royal Society Interface* 15.139 (2018), p. 20170902.

[4]   Noah Olsman et al. "Hard limits and performance tradeoffs in a class of antithetic integral feedback networks". In: *Cell Systems* 9.1 (2019), pp. 49–63.

[5]   Stephanie K. Aoki et al. "A universal biomolecular integral feedback controller for robust perfect adaptation". In: *Nature* 570.7762 (2019), pp. 533–537.

[6]   Sigurd Skogestad and Ian Postlethwaite. *Multivariable feedback control: analysis and design*. Vol. 2. Citeseer, 2007.

[7]   Bernard Brogliato et al. *Dissipative Systems Analysis and Control: Theory and Applications*. Springer, 2020.

[8]   Emmanouil Alexis et al. "Biomolecular mechanisms for signal differentiation". In: *iScience* 24.12 (2021).

[9]   Eyal Gur, Marina Vishkautzan, and Robert T Sauer. "Protein unfolding and degradation by the AAA+ Lon protease". In: *Protein Science* 21 (2012), pp. 268–278.

[10]   Nicolas E Buchler and Matthieu Louis. "Molecular titration and ultrasensitivity in regulatory networks". In: *Journal of Molecular Biology* 384 (2008), pp. 1106–1119.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Regulation strategies for two-output biomolecular networks
Publication Status	Submitted for Publication
Publication Details	Alexis, E., Schulte, C. C., Cardelli, L., & Papachristodoulou, A. (2022). Regulation strategies for two-output biomolecular networks. bioRxiv, 2022-02

Student Confirmation

Student Name:	Emmanouil Alexis		
Contribution to the Paper	Conceptualization and methodology, Formal analysis and Software, Writing		
Signature		Date	14th February 2023

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Professor Antonis Papachristodoulou		
Supervisor comments The candidate made a substantial contribution to the publication. The description above is accurate.		
Signature	Date	16 March 2023

This completed form should be included in the thesis, at the end of the relevant chapter.

Chapter 6

Conclusion and Outlook

The overarching aim of the present thesis is to address key challenges in the design and control of synthetic biological systems. Our contributions are founded on mathematical and computational approaches and encompass the proposal of theory-guided experimental implementations.

In Chapter 3 we propose biological mechanisms for realizing the mathematical operation of differentiation with respect to molecular signals. Estimating time derivatives of signals involved in a system can generally offer useful insights into its function. Nonetheless, unlike electromechanical systems, such a task might be difficult in biological settings. To this end, we introduce three biomolecular topologies which can accept a molecular input signal, such as the concentration of a biomolecule of interest, and produce an output signal which is proportional to the derivative of the former. More precisely, they can successfully differentiate (Fourier transformable) signals of sufficiently long duration around a desired steady state of the output species. The latter cab be interpreted as an elevated "x-axis" (zero-level concentration) and is robust to constant disturbances. Their characteristics and performance trade-offs are mathematically and computationally analyzed. Special emphasis is given to their performance in the presence of input signals with high-frequency components (high-frequency input noise), as this can lead to undesired output signal am-

plification, which is probably the most important problem of both technological and biological differentiators. It is shown that, under appropriate tuning, the proposed topologies are able to "distinguish" input signals based on their frequency content and differentiate only the ones in the frequency range of interest with high accuracy. We further present a structurally modified version of these topologies with enhanced noise-filtering capabilities. In addition, using *E.coli* as a model organism, we investigate natural and synthetic networks which can potentially work efficiently as signal differentiators. Notably, one of our differentiator modules bears resemblance to networks involved in the RpoS-mediated stress response and the RpoH-mediated heat shock response in *E.coli*. Our results aim not only to expand our understanding of cell behaviour but also to pave the way for designing reliable synthetic differentiator modules inside the cell for a variety of applications. Examples include development of speed biosensors as well as realization of biological regulation schemes based on derivative feedback control.

In Chapter 4 we turn our focus to Proportional-Integral-Derivative (PID) control which is the predominant type of feedback control in modern industrial control applications. More specifically, we design an advanced PID controller via biochemical reactions equipped with set point weights and filtered derivative action. It involves an antithetic integrator, thus, achieving Robust Perfect Adaptation (RPA) by eliminating the steady-state error. In parallel, its derivative action, realized by the motifs developed in Chapter 3, considers the rate of change in the control error and takes an anticipatory action to amend the manipulated variable. It can therefore diminish overshoot and expedite the convergence to steady state. Moreover, the aforementioned parts in conjunction with a birth-death process generate proportional action with respect to the control error, providing a simple correction to the manipulated variable which can accelerate the system's response. Our PID controller yields enhanced dynamic performance, reduces stochasticity, and is able to overcome common obstacles such as adverse fluctuations of the control signal owning to sudden changes of the set point or the presence of high-frequency noise. It is also characterized by significant

tunability, as all of its features can be independently adjusted as desired. Our control feedback scheme constitutes a convenient solution for tight regulation of biological processes where both transient and steady-state behaviour is of interest. Finally, our architecture comprises unimolecular and bimolecular reactions governed by mass action kinetics, making it well-suited for *in vitro* experimental implementations – for instance, it could be effectively compiled into a strand displacement DNA-based device. However, embedding this circuit into a cell might present considerable challenges – for example, its intricate structure would impose a significant burden on the host organism.

In Chapter 5 we delve into the problem of regulating biomolecular systems with multiple outputs. In the general case, the outputs affect each other due to coupling (internal) interactions while all can be subject to disturbances from the external environment. Such biomolecular systems provide several capabilities as well as challenges that are not present in the classical single-input single output approaches on which the recent research efforts in this area have concentrated. In particular, we introduce regulatory architectures for processes with two outputs of interest that are able to robustly manipulate the ratio/product of the latter, a linear combination of them, and each of them independently. Our architectures utilize integral feedback action within either centralized or decentralized control schemes, expanding upon the previously described SISO antithetic controller. Their behaviour is thoroughly analyzed via mathematical analysis and simulations with particular emphasis on structural stability. We also highlight their experimental feasibility both *in vivo*, considering *E.coli* as a model organism, and *in vitro* via molecular programming. Note that the challenges and limitations discussed with regard to the former are also relevant to other types of organisms, such as yeast or mammalian cells. The regulation strategies introduced in that chapter signify the inaugural research attempt to manipulate multi-output biological processes with coupling interactions where both the network to be controlled and the controller are embedded in the same biological context. Although our results focus on two-output processes, we demonstrate that our regulation strategies are scalable and can

be easily adapted to more complex processes with a higher number of outputs. The establishment of such multi-output control concepts holds great importance for building the next generation of bio-devices capable of performing sophisticated tasks. Furthermore, our architectures can potentially provide useful insights into the functioning of natural feedback topologies, considering that most of the cellular networks can be viewed as MIMO systems.

There are numerous compelling future research paths, some of which are currently underway, that can build upon the work presented in this thesis. The most important of these directions are outlined below:

- **Multi-output PID control** : A natural extension of our work is the design of PID controllers for multi-output biological processes. To accomplish this, we can integrate the control schemes developed in Chapter 5 with ideas in Chapters 4, 6 as well as other relevant concepts in the literature [23, 87, 17]. One of the main hurdles here is to design efficient regulatory architectures capable of handling potential coupling interactions between inputs/outputs while keeping the corresponding structural complexity at a realistic level.

- **Non-local analysis** : The mathematical analysis presented herein centers on the local behaviour of the systems under consideration which, in most cases, are nonlinear. Thus, we often study the properties of the systems, such as stability, in the vicinity of their nominal operation point through the use of linear perturbation analysis. This approach is favoured due to the plethora of design and analysis methods available for linear systems. Nevertheless, the behaviour of a nonlinear system when operating away from its nominal operation point may deviate significantly from the results obtained via the above approach. It is therefore of particular interest to expand our analysis in order to include large signal analysis as well [43, 12, 57, 80]. In instances where the behaviour of systems away from its equilibria falls short of our performance standards, it is imperative to investigate performance-enhancing techniques,

such as structural modifications or tuning strategies.

- **Biological stochasticity** : The dynamics of the biological systems in this thesis are analyzed through deterministic modelling based on Ordinary Differential Equations (ODEs). Chapter 4 constitutes an exception where, additionally, the standard deviation with respect to the output of the closed-loop system under consideration is estimated by using Van Kampen's Linear Noise Approximation (LNA) of the Chemical Master Equation (CME) [84]. In general, stochastic fluctuations are inevitable in biomolecular environments and can play a crucial role when these environments are characterized by small volumes and molecular counts [39, 22]. Thus, there is a need to systematically study the behaviour of the systems introduced herein using stochastic methods. The latter include Monte Carlo simulations such as Gillespie's stochastic simulation algorithm (SSA) [29], stochastic differential equation approximations such as LNA, and others.

- **Experimental validation** : One of the most impactful extensions of our work is the *in vivo* and/or *in vitro* experimental testing of the proposed architectures. We believe that this will pave the way for our architectures to fulfill their potential by enabling their practical utilization in biotechnology, biomedicine, and other related areas. It is noteworthy that, in recent years, the development of hybrid *in vivo-in silico* platforms has significantly accelerated the experimental evaluation of biomolecular topologies in cell populations or single cells. It is possible for such platforms to implement part of the topology under investigation in a computer while the rest corresponds to a process within cells [48, 47]. Interestingly, such hybrid approaches can yield valuable insights and guidance for fully *in vivo* realizations. Finally, using our theoretical models to obtain realistic predictions and optimally design experiments with respect to different biological environments might require introducing appropriate structural modifications and adopting suitable methods of parameter estimation and sensitivity

analysis [39, 51].

- **Multicellular control** : An alternative approach to implementing our biomolecular control systems is to distribute their constituent parts responsible for sensing, actuation, and computation across different interacting cellular populations [24, 58, 74]. This is particularly convenient for biomolecular designs with high structural complexity whose *in vivo* implementation in a single-cell embedded fashion can result in excessive metabolic burden. For example, this approach might be advantageous in the implementation of our single- or multi-output PID strategies. Multicellular feedback control can generally offer considerable modularity and tuning flexibility in regulating cellular behaviour within microbial consortia.

- **Different control objectives** : Our biomolecular control schemes along with other similar research efforts in the literature focus on constant-in-time disturbance rejection (concerning the output species) which is achieved via output feedback control. However, time-varying disturbances presenting themselves, for instance, as oscillatory signals, are ubiquitous in biological environments. Adapting systems capable of mitigating such disturbances would therefore be of particular interest [33]. Another open problem worth exploring is the development of more advanced regulatory topologies realizing non-linear control concepts based, for example, on state feedback control, which are commonly employed in technological applications. This remains an immense challenge due to several distinct peculiarities in terms of the structure and function of biomolecular networks which are absent in the aforementioned applications [18].

Bibliography

[1] Deepak K Agrawal, Ryan Marshall, Vincent Noireaux, and Eduardo D Sontag. *In vitro* implementation of robust gene regulation in a synthetic biomolecular integral controller. *Nature communications*, 10(1):1–12, 2019.

[2] Pedro Albertos and Sala Antonio. *Multivariable control systems: an engineering approach*. Springer Science & Business Media, 2006.

[3] Bruce Alberts. Molecular biology of the cell. 2017.

[4] Emmanouil Alexis, Luca Cardelli, and Antonis Papachristodoulou. On the design of a PID bio-controller with set point weighting and filtered derivative action. *IEEE Control Systems Letters*, pages 1–1, 2022.

[5] Emmanouil Alexis, Carolin C. M. Schulte, Luca Cardelli, and Antonis Papachristodoulou. Biomolecular mechanisms for signal differentiation. *iScience*, 24(12), 2021.

[6] Emmanouil Alexis, Carolin C. M. Schulte, Luca Cardelli, and Antonis Papachristodoulou. Regulation strategies for two-output biomolecular networks. *bioRxiv*, 2022.

[7] Stephanie K Aoki, Gabriele Lillacci, Ankit Gupta, Armin Baumschlager, David Schweingruber, and Mustafa Khammash. A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature*, 570(7762):533–537, 2019.

[8] Karl Johan Åström and Richard M Murray. *Feedback systems: an introduction for scientists and engineers*. Princeton University Press, 2021.

[9] Roslen Bondi, Francesca Longo, Marco Messina, Francesca D'Angelo, Paolo Visca, Livia Leoni, and Giordano Rampioni. The multi-output incoherent feedforward loop constituted by the transcriptional regulators lasr and rsal confers robustness to a subset of quorum sensing genes in pseudomonas aeruginosa. *Molecular BioSystems*, 13(6):1080–1089, 2017.

[10] Deepro Bonnerjee, Sayak Mukhopadhyay, and Sangram Bagh. Design, fabrication, and device chemistry of a 3-input-3-output synthetic genetic combinatorial logic circuit with a 3-input and gate in a single bacterial cell. *Bioconjugate Chemistry*, 30(12):3013–3020, 2019.

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

[12] Bernard Brogliato, R. Lozano, Bernhard Maschke, and Olav Egeland. *Dissipative Systems Analysis and Control: Theory and Applications*. Springer, 2020.

[13] Stefano Cardinale and Adam Paul Arkin. Contextualizing context for synthetic biology–identifying causes of failure of synthetic biological systems. *Biotechnology journal*, 7(7):856–866, 2012.

[14] Francesca Ceroni, Alice Boo, Simone Furini, Thomas E Gorochowski, Olivier Borkowski, Yaseen N Ladak, Ali R Awan, Charlie Gilbert, Guy-Bart Stan, and Tom Ellis. Burden-driven feedback control of gene expression. *Nature methods*, 15(5):387–393, 2018.

185

[15] Mark R Charbonneau, Vincent M Isabella, Ning Li, and Caroline B Kurtz. Developing a new class of engineered live bacterial therapeutics to treat human diseases. *Nature Communications*, 11(1):1–11, 2020.

[16] Vijaysekhar Chellaboina, Sanjay P Bhat, Wassim M Haddad, and Dennis S Bernstein. Modeling and analysis of mass-action kinetics. *IEEE Control Systems Magazine*, 29(4):60–78, 2009.

[17] Michael Chevalier, Mariana Gómez-Schiavon, Andrew H Ng, and Hana El-Samad. Design and analysis of a proportional-integral-derivative controller with biological molecules. *Cell systems*, 9(4):338–353, 2019.

[18] Domitilla Del Vecchio and Richard M Murray. *Biomolecular feedback systems*. Princeton University Press Princeton, NJ, 2015.

[19] Domitilla Del Vecchio, Alexander J Ninfa, and Eduardo D Sontag. Modular cell biology: retroactivity and insulation. *Molecular systems biology*, 4(1):161, 2008.

[20] H El-Samad, JP Goff, and M Khammash. Calcium homeostasis and parturient hypocalcemia: an integral feedback perspective. *Journal of theoretical biology*, 214(1):17–29, 2002.

[21] Michael B Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, 2000.

[22] Radek Erban and S Jonathan Chapman. *Stochastic modelling of reaction–diffusion processes*, volume 60. Cambridge University Press, 2020.

[23] Maurice Filo, Sant Kumar, and Mustafa Khammash. A hierarchy of biomolecular proportional-integral-derivative feedback controllers for robust perfect adaptation and dynamic performance. *Nature communications*, 13(1):1–19, 2022.

[24] Gianfranco Fiore, Antoni Matyjaszkiewicz, Fabio Annunziata, Claire Grierson, Nigel J Savery, Lucia Marucci, and Mario Di Bernardo. In-silico analysis and implementation of a multicellular feedback control strategy in a synthetic bacterial consortium. *ACS Synthetic Biology*, 6(3):507–517, 2017.

[25] Timothy Frei, Federica Cella, Fabiana Tedeschi, Joaquín Gutiérrez, Guy-Bart Stan, Mustafa Khammash, and Velia Siciliano. Characterization and mitigation of gene expression burden in mammalian cells. *Nature communications*, 11(1):1–14, 2020.

[26] Timothy Frei, Ching-Hsiang Chang, Maurice Filo, Asterios Arampatzis, and Mustafa Khammash. A genetic mammalian proportional–integral feedback control circuit for robust and precise gene regulation. *Proceedings of the National Academy of Sciences*, 119(00):e2122132119, 2022.

[27] Timothy S Gardner, Charles R Cantor, and James J Collins. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403(6767):339–342, 2000.

[28] Charlie Gilbert and Tom Ellis. Biological engineered living materials: growing functional materials with genetically programmable properties. *ACS synthetic biology*, 8(1):1–15, 2018.

[29] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.

[30] Reyna Berenice González-González, Elda A Flores-Contreras, Everardo González-González, Nora E Torres Castillo, Roberto Parra-Saldívar, and Hafiz MN Iqbal. Biosensor constructs for the monitoring of persistent emerging pollutants in environmental matrices. *Industrial & Engineering Chemistry Research*, 2022.

[31] F Veronica Greco, Amir Pandi, Tobias J Erb, Claire S Grierson, and Thomas E Goro-chowski. Harnessing the central dogma for stringent multi-level control of gene expression. *Nature communications*, 12(1):1–11, 2021.

[32] Agostino Guarino, Davide Fiore, and Mario Di Bernardo. In-silico feedback control of a mimo synthetic toggle switch via pulse-width modulation. In *2019 18th European Control Conference (ECC)*, pages 680–685. IEEE, 2019.

[33] Ankit Gupta and Mustafa Khammash. The internal model principle for biomolecular control theory. *IEEE Open Journal of Control Systems*, 2:63–69, 2023.

[34] Wolfgang Halter, Richard M Murray, and Frank Allgöwer. Analysis of primitive genetic interactions for the design of a genetic signal differentiator. *Synthetic Biology*, 4(1):ysz015, 2019.

[35] Wolfgang Halter, Zoltan A Tuza, and Frank Allgöwer. Signal differentiation with genetic networks. *IFAC-PapersOnLine*, 50(1):10938–10943, 2017.

[36] Patrick Harrigan, Hiten D Madhani, and Hana El-Samad. Real-time genetic compensation defines the dynamic demands of feedback control. *Cell*, 175(3):877–886, 2018.

[37] Chelsea Y Hu and Richard M Murray. Layered feedback control overcomes performance trade-off in synthetic biomolecular networks. *Nature Communications*, 13(1):1–13, 2022.

[38] Hsin-Ho Huang, Yili Qian, and Domitilla Del Vecchio. A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *Nature communications*, 9(1):1–12, 2018.

[39] Pablo A Iglesias and Brian P Ingalls. *Control theory and systems biology*. MIT press, 2010.

[40] Brian P Ingalls. *Mathematical modeling in systems biology: an introduction.* MIT press, 2013.

[41] Ross D Jones, Yili Qian, Velia Siciliano, Breanna DiAndreth, Jin Huh, Ron Weiss, and Domitilla Del Vecchio. An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. *Nature communications*, 11(1):1–16, 2020.

[42] Ciarán L Kelly, Andreas W K Harris, Harrison Steel, Edward J Hancock, John T Heap, and Antonis Papachristodoulou. Synthetic negative feedback circuits using engineered small rnas. *Nucleic acids research*, 46(18):9875–9889, 2018.

[43] Hassan K. Khalil. *Nonlinear systems (3nd Ed.).* Prentice-Hall, Inc., 2002.

[44] Petar Kokotović, Hassan K Khalil, and John O'reilly. *Singular perturbation methods in control: analysis and design.* SIAM, 1999.

[45] Nicholas Kottenstette, Michael J McCourt, Meng Xia, Vijay Gupta, and Panos J Antsaklis. On relationships among passivity, positive realness, and dissipativity in linear systems. *Automatica*, 50(4):1003–1016, 2014.

[46] Jocelyn E Krebs, Elliott S Goldstein, and Stephen T Kilpatrick. *Lewin's genes XII.* Jones & Bartlett Learning, 2017.

[47] Sant Kumar and Mustafa Khammash. Platforms for optogenetic stimulation and feedback control. *Frontiers in Bioengineering and Biotechnology*, 10, 2022.

[48] Sant Kumar, Marc Rullan, and Mustafa Khammash. Rapid prototyping and design of cybergenetic single-cell controllers. *Nature communications*, 12(1):1–13, 2021.

[49] B. P. Lathi. *Signal processing and linear systems.* Oxford University Press, 1998.

[50] James C Liao, Luo Mi, Sammy Pontrelli, and Shanshan Luo. Fuelling the future: microbial engineering for the production of sustainable biofuels. *Nature Reviews Microbiology*, 14(5):288–304, 2016.

[51] Gabriele Lillacci and Mustafa Khammash. Parameter estimation and model selection in computational biology. *PLoS computational biology*, 6(3):e1000696, 2010.

[52] Zihe Liu, Kai Wang, Yun Chen, Tianwei Tan, and Jens Nielsen. Third-generation biorefineries as the means to produce fuels and chemicals from CO_2. *NatureCatalysis*, $3(3) : 274 - -288, 2020$.

[53] Harvey Lodish, Arnold Berk, Chris A Kaiser, Chris Kaiser, Monty Krieger, Matthew P Scott, Anthony Bretscher, Hidde Ploegh, Paul Matsudaira, et al. *Molecular cell biology*. Macmillan, 2008.

[54] Jean-Baptiste Lugagne, Sebastián Sosa Carrillo, Melanie Kirch, Agnes Köhler, Gregory Batt, and Pascal Hersen. Balancing a genetic toggle switch by real-time feedback control and periodic forcing. *Nature communications*, 8(1):1–8, 2017.

[55] Yitong Ma, Mark W Budde, Michaëlle N Mayalu, Junqin Zhu, Andrew C Lu, Richard M Murray, and Michael B Elowitz. Synthetic mammalian signaling circuits for robust cell population control. *Cell*, 185(6):967–979, 2022.

[56] Dan B Marghitu. *Mechanical engineer's handbook*. Elsevier, 2001.

[57] Horacio J Marquez. *Nonlinear control systems: analysis and design*, volume 161. John Wiley Hoboken^ eN. JNJ, 2003.

[58] Vittoria Martinelli, Davide Salzano, Davide Fiore, and Mario di Bernardo. Multi-cellular pi control for gene regulation in microbial consortia. *IEEE Control Systems Letters*, 6:3373–3378, 2022.

[59] Monica P McNerney, Kailyn E Doiron, Tai L Ng, Timothy Z Chang, and Pamela A Silver. Theranostic cells: emerging clinical applications of synthetic biology. *Nature Reviews Genetics*, 22(11):730–746, 2021.

[60] Saurabh Modi, Supravat Dey, and Abhyudai Singh. Noise suppression in stochastic genetic circuits using PID controllers. *PLoS Computational Biology*, 17(7):e1009249, 2021.

[61] Kevin J Morey, Mauricio S Antunes, Matt J Barrow, Fernando A Solorzano, Keira L Havens, J Jeff Smith, and June Medford. Crosstalk between endogenous and synthetic components–synthetic signaling meets endogenous components. *Biotechnology Journal*, 7(7):846–855, 2012.

[62] 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.

[63] Kevin Oishi and Eric Klavins. Biomolecular implementation of linear I/O systems. *IET Systems Biology*, 5(4):252–260, 2011.

[64] Alan V. Oppenheim, Alan S. Willsky, and S. Hamid Nawab. *Signals Systems (2nd Ed.)*. Prentice-Hall, Inc., 1996.

[65] Abhilash Patel, Richard M Murray, and Shaunak Sen. Assessment of robustness to temperature in a negative feedback loop and a feedforward loop. *ACS Synthetic Biology*, 9(7):1581–1590, 2020.

[66] Nuno MG Paulino, Mathias Foo, Jongmin Kim, and Declan G Bates. Pid and state feedback controllers using dna strand displacement reactions. *IEEE Control Systems Letters*, 3(4):805–810, 2019.

[67] Johan Paulsson. Summing up the noise in gene networks. *Nature*, 427(6973):415–418, 2004.

[68] Elisa Pedone, Irene De Cesare, Criseida G Zamora-Chimal, David Haener, Lorena Postiglione, Antonella La Regina, Barbara Shannon, Nigel J Savery, Claire S Grierson, Mario Di Bernardo, et al. Cheetah: a computational toolkit for cybergenetic control. *ACS Synthetic Biology*, 10(5):979–989, 2021.

[69] Elisa Pedone, Lorena Postiglione, Francesco Aulicino, Dan L Rocca, Sandra Montes-Olivas, Mahmoud Khazim, Diego di Bernardo, Maria Pia Cosma, and Lucia Marucci. A tunable dual-input system for on-demand dynamic gene expression regulation. *Nature communications*, 10(1):1–13, 2019.

[70] Giansimone Perrino, Sara Napolitano, Francesca Galdi, Antonella La Regina, Davide Fiore, Teresa Giuliano, Mario di Bernardo, and Diego di Bernardo. Automatic synchronisation of the cell cycle in budding yeast through closed-loop feedback control. *Nature communications*, 12(1):1–12, 2021.

[71] Pascal A Pieters, Bryan L Nathalia, Ardjan J van der Linden, Peng Yin, Jongmin Kim, Wilhelm TS Huck, and Tom FA de Greef. Cell-free characterization of coherent feed-forward loop-based synthetic genetic circuits. *ACS synthetic biology*, 10(6):1406–1416, 2021.

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

[73] Jakub Leszek Radzikowski, Silke Vedelaar, David Siegel, Álvaro Dario Ortega, Alexander Schmidt, and Matthias Heinemann. Bacterial persistence is an active σS stress response to metabolic flux limitation. *Molecular systems biology*, 12(9):882, 2016.

[74] Xinying Ren, Ania-Ariadna Baetica, Anandh Swaminathan, and Richard M Murray. Population regulation in microbial consortia using dual feedback control. In *2017 IEEE 56th annual conference on decision and control (CDC)*, pages 5341–5347. IEEE, 2017.

[75] Christian Cuba Samaniego, Giulia Giordano, and Elisa Franco. Practical differentiation using ultrasensitive molecular circuits. In *2019 18th European Control Conference (ECC)*, pages 692–697. IEEE, 2019.

[76] Christian Cuba Samaniego, Jongmin Kim, and Elisa Franco. Sequestration and delays enable the synthesis of a molecular derivative operator. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 5106–5112. IEEE, 2020.

[77] Elena Shchepakina, Vladimir Sobolev, and Michael P Mortell. Singular perturbations. *Introduction to system order reduction methods with applications*, 2014.

[78] Shuobo Shi, Zhihui Wang, Lirong Shen, and Han Xiao. Synthetic biology: a new frontier in food production. *Trends in Biotechnology*, 2022.

[79] Sigurd Skogestad and Ian Postlethwaite. *Multivariable feedback control: analysis and design*, volume 2. Citeseer, 2007.

[80] Jean-Jacques E Slotine, Weiping Li, et al. *Applied nonlinear control*, volume 199. Prentice hall Englewood Cliffs, NJ, 1991.

[81] Hye-In Son, Andrea Weiss, and Lingchong You. Design patterns for engineering genetic stability. *Current opinion in biomedical engineering*, 19:100297, 2021.

[82] Harrison Steel, Robert Habgood, Ciarán L Kelly, and Antonis Papachristodoulou. In situ characterisation and manipulation of biological systems with Chi. Bio. *PLoS biology*, 18(7):e3000794, 2020.

[83] Peter S Swain, Michael B Elowitz, and Eric D Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20):12795–12800, 2002.

[84] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

[85] Ophelia S Venturelli, Hana El-Samad, and Richard M Murray. Synergistic dual positive feedback loops established by molecular sequestration generate robust bimodal response. *Proceedings of the National Academy of Sciences*, 109(48):E3324–E3333, 2012.

[86] Benjamin H Weinberg, NT Hang Pham, Leidy D Caraballo, Thomas Lozanoski, Adrien Engel, Swapnil Bhatia, and Wilson W Wong. Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nature biotechnology*, 35(5):453–462, 2017.

[87] Max Whitby, Luca Cardelli, Marta Kwiatkowska, Luca Laurenti, Mirco Tribastone, and Max Tschaikowski. PID control of biochemical reaction networks. *IEEE Transactions on Automatic Control*, 2021.

[88] Tau-Mu Yi, Yun Huang, Melvin I Simon, and John Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences*, 97(9):4649–4653, 2000.