Design Constraints for Biological Systems That Achieve Adaptation and Disturbance Rejection

Harrison Steel , *Student Member, IEEE*, and Antonis Papachristodoulou , *Senior Member, IEEE*

Abstract—**Many processes in natural biological systems, such as chemotaxis in bacteria and osmoregulation in yeast, rely on control architectures fundamentally equivalent to commonly used motifs from electrical and control engineering. However, difficulties arise when designing and implementing these architectures in a biological context due to uncertainties inherent in the behavior of biological systems and physical limitations of the available parts. In this paper, we discuss recent developments in the study of biological control systems, which are increasingly necessary for realization of complex synthetic biological constructs, focusing on methods for their design and implementation. We establish a range of desirable properties that ease implementation of biological constructs and apply classical control theory to derive a set of constraints to aid the design of systems that achieve adaptation or disturbance rejection. We demonstrate how these constraints can be used in practice, first deriving the necessary structure for a linear system that achieves adaptation, and then embedding this in a nonlinear model of biological relevance that could be built in the laboratory.**

Index Terms—**Adaptation, biological control, design constraints, disturbance rejection, feedback control, homeostasis, synthetic biology.**

I. INTRODUCTION

SYNTHETIC biology is an emerging field at the interface of biology and engineering, which aims to design and build novel biological systems to solve problems in fields ranging from biomanufacturing to medicine [1]. Over the past 15 years, scientists have successfully used synthetic biology to rewire natural genetic components; however, attaining reliability and consistency of performance remains challenging [2]. Difficulties arise for reasons including unforeseen dependencies and interactions between synthetic constructs and native cellular machinery [3], and the random fluctuations and noise inherent in cellular processes [4]. In natural systems, it has been observed that many such challenges are overcome by feedback architectures similar to those commonly used in control engineering [5], [6]. This has inspired researchers to implement analogous control systems in biological contexts [7], with these systems now playing an increasingly important role in the realization of synthetic biological designs.

The bacterial chemotaxis system is a model example of a natural biological control system. In the presence of extracellular stimuli, it is able to robustly (that is, with minimal sensitivity to the properties of its constituent components [8]) return its state to prestimulus levels [9], a process known as *adaptation*. It has been demonstrated that the chemotaxis system's underlying structure relies on integral feedback control [5], which has encouraged the use of feedback for development of synthetic signaling pathways that perform similar functions [10], [11]. In addition to adaptation, feedback control architectures can provide a range of favorable capabilities for synthetic biological networks [12]. For example, in the simple form of negative autoregulation, feedback can improve the response time of gene networks [13] and reduce heterogeneity of gene expression between cells [14]. Feedback can also be exploited to implement high dynamic-range gene circuits for computation [15] and to improve the efficacy of gene regulatory systems [16]. Networks that utilize feedback control to improve both performance and robustness will be of increasing interest to synthetic biologists as they attempt to move synthetic constructs outside the laboratory [6], where varying and nonideal conditions can impact their performance [17].

Though the widespread need for synthetic biological control systems has been recognized [7] and many architectures have been proposed, their implementation remains challenging [18]. This stems from a lack of systematic approaches to design, as well as methods for optimal (or even feasible) implementation [19]. To this end, it has been recognized that constraint sets must be developed to limit the space of viable designs [20]. Past work has established such constraints for systems with properties such as switch-like responses [21], disturbance rejection (homeostasis) [22], and adaptation [18]. This has been done using a range of techniques including BDC decomposition, which decomposes system Jacobians into a product of three matrices B, D, and C that capture the system structure. Such approaches can be used to determine structural dependencies of variables upon one another [23] (an approach originally popularized in the determination of ecological community matrices [24], [25]) and to reveal adaptive and oscillatory behaviors [26]. State-space approaches have been used in both linear [27] and nonlinear [28] systems to provide necessary conditions for adaptation. Furthermore, "cofactor conditions," which require systems to have attractive steady-state functions for which changes in the input do not affect the output, have been employed to determine design principles for construction of homeostatic systems [29]. However, even though these constraints can substantially narrow

feasible design spaces, they are often mathematically abstract and do not address many of the physical challenges posed by biological components that make the implementation of control structures of any substantial complexity challenging [18].

Therefore, as planned control architectures become more complex, it will be increasingly advantageous to develop biological constructs that approximate the fundamental components of traditional control systems, such as integration, gain, and summation junctions [15], [30]. This will facilitate implementation of standard control structures such as lead-lag [19] and poportional-integral-derivative variant [31] controllers, whose favorable properties have led to their ubiquity across engineering disciplines [6].

Prior attempts to develop modularity for biological components have focused on the standardization of synthetic biological "parts" (for example, via the BioBrick standard [32]) to allow ease and predictability of assembly. However, as parts libraries have grown, it has become clear that the effective design of biological systems requires accounting for the context-dependent influence that up- and down-stream interactions impart upon a given component (though tools have been developed to "insulate" components from these effects [33]), making true modularity challenging. Thus, substantial experimental fine-tuning is often required for constructs of any appreciable size, even after extensive *in silico* testing and development [2]. One approach to minimize the laborious fine-tuning process is to pursue designs that build upon (and potentially rewire) native cellular processes [34], [35], reducing the network size required to achieve a given functionality (and providing the secondary benefit of minimizing the metabolic load that the synthetic system places upon its host cell [31]). However, approaches that build upon native cellular systems are still in their infancy, and the limits of their practicality and applicability are yet unclear.

Though challenges surrounding the modularity and unpredictability of biological systems persist, thanks to fundamental biological research, there is an ever-increasing assortment of biological elements that can be used for synthetic designs. Constructs can be built using systems operating on a range of molecular biological levels (for example, using DNA [36], RNA [37], or protein [14]), as well as at varying time-scales [38]–[40] and species concentrations [31]. However, working with this disparate assortment of components requires extensive interdisciplinary expertise and can substantially increase the necessity for experimental troubleshooting and system fine-tuning. Previous work has developed tools for automating the component-selection and tuning process for some classes of synthetic constructs [41], though as yet they only provide a narrow (primarily logic function-related) range of capabilities.

Working toward addressing some of these challenges in implementation, in this paper, we derive design constraints on the structure of networks that provide two common control capabilities: adaptation and disturbance rejection (homeostasis). We provide guidelines that can be used to simplify the design and implementation of systems with these capabilities, which due to our restrictive constraint set provide a greater level of specificity than previous work. We take a state-space approach to designing the linearized control system, to which we apply assumptions to

aid implementation, and then derive constraints on the network structure and parameters. Some of these results follow from the internal model principle [42], which in the present context dictates that should a feedback system adapt to a class of input signals, then it must contain a subsystem able to generate signals of that class [43]. Our design constraints are narrowed to consider the biological systems in question, and due to the restrictive assumption set proposed yield only the most readily implementable designs. Once an appropriate linear system has been designed, we demonstrate methods for its embedding in a nonlinear model of biological relevance. The final stage of the synthetic biology design process, the selection of particular biological parts for implementation, is not addressed herein.

In Section II, methods for the modeling of biological systems are discussed, and the approach taken to linearization of a general nonlinear system is outlined. Section III describes the paper's aim, as well as examples of functional forms for biological interactions that may be used to implement the systems described in this work. Section III also outlines two general constraints that are used in the rest of the paper to exclude systems with some difficult-to-implement properties. In Section IV, results are derived that govern the equilibrium behavior of linear systems, which are then used in Sections V and VI to derive constraints for networks capable of adaptation and rejection of disturbances, respectively. In Section VII, the results of Section V are employed to derive a linearized network that can adapt to step-type disturbances. This network is used to design a nonlinear system with these capabilities, which is simulated using illustrative parameters. Section VIII discusses the assumptions made in this work and their implications, and concludes the paper.

II. MODELING BIOLOGICAL CONTROL SYSTEMS

Modeling frameworks for biological systems range from stochastic/probabilistic to deterministic differential equation (DE) models [44]. This paper focuses on a subset of the later, utilizing both linear and nonlinear DEs to describe the dynamics of individual state variables, which may represent individual species, or larger scale properties, of biological systems. Each DE will be a (potentially nonlinear) function of the system's state variables, as well as any external inputs or disturbances to the system.

A general first-order nonlinear system describing the dynamics of n state variables $\boldsymbol{x}(t) = [x_1(t), \ldots, x_n(t)] \in \mathbb{R}^n$ as considered in this paper takes the form

$$\dot{\boldsymbol{x}}(t) = \mathbf{f}(\boldsymbol{x}(t), \mathbf{u}(t)) \tag{1}$$

in which $\dot{\boldsymbol{x}}(t) = \frac{d\boldsymbol{x}(t)}{dt} \in \mathbb{R}^n$ is a vector containing the time derivative of each state variable, $\mathbf{u}(t) \in \mathbb{R}^n$ is a vector containing the time-dependent inputs to the nonlinear system's state variables, and $\mathbf{f}(\boldsymbol{x}(t), \mathbf{u}(t)) \in \mathbb{R}^n$ contains n general nonlinear functions of the state variables and inputs.

To simplify analysis, it is often advantageous to examine nonlinear systems in a regime where they can be approximated by a linear system. The system in (1) can be linearized via a Taylor series expansion of $\mathbf{f}(\boldsymbol{x}, \mathbf{u})$ about one of its equilibrium points
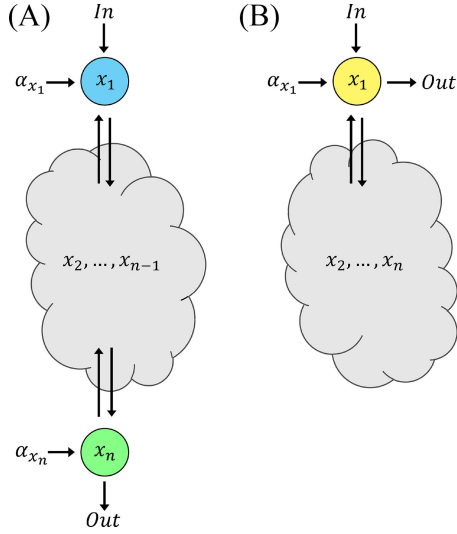
Fig. 1. Schematic representation of systems that achieve (A) adaptation and (B) disturbance rejection. For an adaptive system, the long-term behavior of the output node (x_n, green) is independent of a time-dependent signal applied at the input node (x_1, blue). The output variable of a disturbance-rejecting system (x_1, yellow) is in the long-term independent of a time-dependent input acting upon the same node. The α_{x_i} represent basal expression rates of the system's state variables x_i, and grey clouds represent a yet undefined network of interactions between the system's state variables.

$\boldsymbol{x}_e \in \mathbb{R}^n$ (so $\mathbf{f}(\boldsymbol{x}_e, \mathbf{0}) = \mathbf{0}$)), which gives (ignoring higher order terms)

$$\dot{\tilde{\boldsymbol{x}}} \approx \left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_e, \mathbf{0}} \tilde{\boldsymbol{x}} + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\right|_{\boldsymbol{x}_e, \mathbf{0}} \mathbf{u} \tag{2}$$

where $\left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_e, \mathbf{0}} \in \mathbb{R}^{n \times n}$ is the Jacobian of $\mathbf{f}$ evaluated at $(\boldsymbol{x}_e, \mathbf{0})$, $\left.\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\right|_{\boldsymbol{x}_e, \mathbf{0}} \in \mathbb{R}^{n \times n}$ is a matrix of all partial derivatives of $\mathbf{f}$ with respect to elements of $\mathbf{u}$ evaluated at $(\boldsymbol{x}_e, \mathbf{0})$, and $\tilde{\boldsymbol{x}} = (\boldsymbol{x} - \boldsymbol{x}_e)$ is the deviation in $\boldsymbol{x}$ from $\boldsymbol{x}_e$.

III. AIM AND PROBLEM OUTLINE

In this paper, we aim to provide constraints for the design of biological systems of minimal size that achieve *adaptation* and *disturbance rejection* (defined in Fig. 1) for inputs of the form $u(t) = t^{k-1}, t > 0$ (where $k = 1$ would indicate a step, $k = 2$ a ramp, etc.). Rejection of step inputs is necessary in many biological systems for rebaselining (for example, in chemotaxis [8]), and adaptation to higher order inputs (i.e., $k \geq 2$) may be useful for distinguishing biological processes that create products at a constant rate from those that have higher order dynamics. Furthermore, disturbance rejecting systems can better function as modular components of larger synthetic circuits since they can be designed to have outputs that (for example) maintain a constant concentration while being consumed by downstream processes. Designing nonlinear systems of biological relevance that achieve these goals, and are also feasible for implementation, is challenging. Therefore, we approach this problem by designing a linear system, which is then used to determine the form of an appropriate nonlinear system. As such our design approach considers the linearization process outlined in Section II in the opposite direction: we first derive the necessary structure

of $\left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_e, \mathbf{0}}$ from a set of constraints, and then attempt to find a nonlinear biological system as in (1) possessing these linear dynamics about one of its equilibrium points. For a given linear system, there will be many feasible forms of $\mathbf{f}(\boldsymbol{x}, \mathbf{u})$ depending on the physical components chosen, and so this choice will be narrowed down by selection of implementations that are most biologically tractable.

A. Potential Forms for $\boldsymbol{f}(\boldsymbol{x}, \mathbf{u})$

Though there is a wide (and growing) range of synthetic parts from which biological circuits can be built, some commonly used elements [and their potential mathematical expressions that could appear in $\mathbf{f}(\boldsymbol{x}, \mathbf{u})$ from (1)] are outlined below. In these equations, we have placed the state variables in square brackets to indicate that in this case they refer to concentrations.

1) *Constant expression*: An element of a network for which the concentration increases at a constant rate α:

$$\frac{d[x]}{dt} \propto \alpha. \tag{3}$$

2) *Activated/Repressed expression*: An element for which the creation rate depends on the concentration of a secondary species $[y]$. Activation corresponds to the case in (4) for which large $[y]$ results in increased expression, while repression removes the factor of $[y]$ from the equation's numerator, resulting in a smaller value of $d[x]/dt$ for large $[y]$. The Hill coefficient (n) describes the cooperativity of the activation/repression process ($n = 1$ for a noncooperative process):

$$\frac{d[x]}{dt} \propto \frac{\beta[y]^n}{k^n + [y]^n}. \tag{4}$$

3) *Degradation/Dilution*: Reduction in an element's concentration due to its degradation over time, or dilution (due to cell growth), which is proportional (via a factor δ) to its own concentration:

$$\frac{d[x]}{dt} \propto -\delta[x]. \tag{5}$$

4) *Michaelis–Menten enzyme kinetics*: Reactions in which an enzyme y is used to increase or decrease the concentration of a substrate x can be described using the Michaelis–Menten equation. The rate of this process depends on the enzyme's concentration ($[y]$) and its catalytic rate (k_{cat}):

$$\frac{d[x]}{dt} \propto \frac{k_{\text{cat}}[y][x]}{K_M + [x]}. \tag{6}$$

5) *Mass action kinetics*: Chemical reactions that transform one group of species into another can be described by mass action kinetics. For example, for a reaction of the form

$$\alpha x + \beta y \rightarrow \gamma c \tag{7}$$

we have that

$$\frac{d[x]}{dt}, \frac{d[y]}{dt} \propto -k[x]^\alpha [y]^\beta. \tag{8}$$

Again, the above-mentioned list of potential forms found in $\mathbf{f}(\boldsymbol{x}, \mathbf{u})$ is by no means exhaustive; it aims to provide examples of simple mathematical approximations for the biological systems considered in this work.

B. System Linearization and Assumptions for Implementation

We can express the linearization (2) as a general linear system of the form

$$\dot{\boldsymbol{x}} = \mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{u},$$

$$\boldsymbol{y} = \mathbf{C}\boldsymbol{x} \qquad (9)$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$ represents the system's state variables and $\dot{\boldsymbol{x}}(t)$ their time derivatives (as in Section II). $\mathbf{u}(t) \in \mathbb{R}^m$ is the system's input and $\mathbf{y}(t) \in \mathbb{R}^p$ the system output. For single-input single-output (SISO) systems as considered in Sections IV–VI, $m = p = 1$. $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are matrices of appropriate dimensions that define the system, with a structure that will be constrained in terms of the sign or nullity of their elements. For practical implementation, the elements of these matrices will then take values in $\mathbb{R}$ defined by the system's biological parameters. We focus on SISO systems due to their ease of analysis in the present context, though approaches to expanding this framework to incorporate systems in which inputs affect multiple state variables are discussed in Section VIII.

In the following sections submatrix notation will be as follows.

1) $\mathbf{A}_{i,j}$ refers to the element in the ith row and jth column of $\mathbf{A}$.
2) $\mathbf{A}_{i:j,k:l}$ refers to a submatrix of $\mathbf{A}$ of size $j - i$ by $l - k$, which includes elements from the ith to jth row, and kth to lth column of $\mathbf{A}$.

To provide biologically tractable designs, we make a range of assumptions that limit the properties of designed linear systems. First, we require that $\mathbf{A}$ matrices are Hurwitz (stable in the control theory sense). In order to create systems for which the *structure*, rather than particular parameter values, guarantees performance goals are achieved, we make the assumption of *element independence*, defined in Assumption 3.2. Due to the leaky nature of biological systems, particularly at the promoter level [45], we will constrain the constant term in the DE describing each species' time evolution to be non-negative. This will be referred to as a *basal* level of production (see Constraint 3.3). Furthermore, we wish to avoid systems in which individual components impact other state variables both positively and negatively, and thus require that the partial derivatives with respect to $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ of the DE governing the dynamics of x_i have the same sign for all $\boldsymbol{x}$ (see Constraint 3.4).

Definition III.1: We define (and denote as $\mathbb{S}^{\mathbf{A}}$) a *system of constraints* on a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ to be an $n \times n$ array of functions of elements of $\mathbf{A}$.

For example, we might have that $\mathbb{S}^{\mathbf{A}}_{i,j} = \mathbf{A}_{a,b} + \mathbf{A}_{c,d}\mathbf{A}_{e,f}$ if under a given constraint it was required that the element $\mathbf{A}_{i,j}$ be equal to the combination of other elements, $\mathbf{A}_{a,b} + \mathbf{A}_{c,d}\mathbf{A}_{e,f}$.

Assumption III.2. Element Independence: We say that a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ has the property of *element independence* if, for

a given system of constraints $\mathbb{S}^{\mathbf{A}}$, the following holds:

$$\frac{\partial \mathbb{S}^{\mathbf{A}}_{i,j}}{\partial \mathbf{A}_{l,m}} = 0 \qquad (10)$$

for all $\mathbf{A}_{l,m} \in \mathbf{A}$ subject to $l \neq i$ or $m \neq j$.

Element independence can be thought of as a criterion that eliminates constraint sets that require elements of the system matrix $\mathbf{A}$ to be equal to the sum, product, or any arbitrary nontrivial function, of other elements of $\mathbf{A}$. This criterion must be understood within the context of constraints applied to a matrix's elements, which impose functional relationships between different elements (i.e., $\mathbb{S}^{\mathbf{A}}_{i,j}$) for families of systems that achieve given design goals. We desire element independence in biological contexts in order to aid the realization of designed systems, since tuning parameter relationships for different biological parts of a system is challenging in many cases. Similar requirements to element independence have previously been investigated [46] and are discussed further in Section VIII.

C. Other General Constraints for Implementation

In addition to the mathematical constraints derived above, there are a number of desirable properties that our designed system should possess in order to ease its implementation. These will not be used to prove the results in Sections IV–VI, but are applied during the design of a specific biological implementation in Section VII.

Previously, we assumed that in a biological system, each state variable could have a non-negative basal expression level. If a given state variable x_i is to reach an equilibrium, there must be a balance between positive and negative rate contributions to its nonlinear dynamics. Since we assume each state variable will have a non-negative basal expression rate α_{x_i}, there must be at least one negative element in each row of $\mathbf{A}$. This requirement yields the following constraint.

Constraint III.3: $\mathbf{A}_{i,j} < 0$ for at least one $j = 1, \ldots, n$, for all $i = 1, \ldots, n$.

To aid implementation, we would like to avoid situations in which one element of a network has influences of differing signs on other elements. Either it should have a positive influence (e.g., a transcriptional activator) or a negative influence (e.g., a transcriptional repressor), but not both. Thus, the off diagonal terms in each column of $\mathbf{A}$ should all be strictly non-negative or nonpositive. This requirement yields the following constraint.

Constraint III.4: Either $\mathbf{A}_{i,j} \leq 0 \quad \forall\, i = 1, \ldots, j - 1, j + 1, \ldots, n \quad \forall\, j = 1, \ldots, n$ or $\mathbf{A}_{i,j} \geq 0 \quad \forall\, i = 1, \ldots, j - 1, j + 1, \ldots, n \quad \forall\, j = 1, \ldots, n$.

IV. GENERAL LINEARIZED NETWORK STRUCTURE

Proposition IV.1: Consider an SISO system in the form of (9) with n state variables and in which $\mathbf{A}$ is Hurwitz. The following statements are equivalent.

1) The system rejects inputs of the form $u(t) = t^{k-1}$ with $k \in \mathbb{N}^+$, i.e., $\lim_{t \to \infty} y(t)$ is constant.

2)

$$\lim_{s \to 0} \det \begin{bmatrix} s\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C}/s^{k-1} & 0 \end{bmatrix} = 0 \qquad (11)$$

where s is the frequency domain (Laplace) parameter, $\mathbf{A} \in \mathbb{R}^{n \times n}, \mathbf{B} \in \mathbb{R}^{n \times 1}$, and $\mathbf{C} \in \mathbb{R}^{1 \times n}$, and $\mathbf{I}$ is the $n \times n$ identity matrix.

Proof: Under the stability assumptions of the proposition, from the final value theorem [47], we have that

$$y(\infty) = \lim_{s \to 0} sY(s). \qquad (12)$$

For this system to be unperturbed (in the $t \to \infty$ limit) by the input/disturbance $u(t)$, we, thus, require $\lim_{t \to \infty} y(t) = 0$. For our system, we can express the transfer function from input to output in terms of its Rosenbrock system matrix as [48], [49]

$$G(s) = \frac{1}{\det(s\mathbf{I} - \mathbf{A})} \det \begin{bmatrix} s\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & 0 \end{bmatrix} \qquad (13)$$

which, taking $Y(s) = G(s)U(s)$ and a general input of the form $u(t) = t^{k-1}$, equivalent to $U(s) = (k-1)!/s^k$ in the Laplace domain, gives

$$\lim_{s \to 0} \frac{(k-1)!}{\det(s\mathbf{I} - \mathbf{A}) \times s^{k-1}} \det \begin{bmatrix} s\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & 0 \end{bmatrix} = 0. \qquad (14)$$

Here, $\det\begin{bmatrix} s\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & 0 \end{bmatrix}$ will yield a polynomial in s of maximal order $n - 1$, while the lowest order s term in $\det(s\mathbf{I} - \mathbf{A}) \times s^{k-1}$ will be of order $k - 1$ and equal to $\det(-\mathbf{A}) \times s^{k-1}$. By placing $k - 1$ zeros at the origin, this term causes cancelling of inputs of the form the system aims to reject, as required by the internal model principle. Thus, since we require this limit to approach zero, as $s \to 0$, the denominator's value will be dominated by its lowest order term, and we can replace the condition in (14) with one which is equivalent in the $s \to 0$ limit

$$0 = \lim_{s \to 0} \frac{(k-1)!}{\det(-\mathbf{A}) \times s^{k-1}} \det \begin{bmatrix} s\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & 0 \end{bmatrix}. \qquad (15)$$

Condensing the $1/s^{k-1}$ prefactor into the bottom row of the determinant and eliminating constant terms gives the desired result. ∎

V. NETWORK CONSTRAINTS FOR ADAPTATION

For a system to achieve adaptation [see Fig. 1(A)], the long-term behavior of one of its state variables (the output) must be independent of a time-dependent input applied to a different state variable (the input). We will arbitrarily define the input state variable as the first element in the system, and the output state variable as the last, giving

$$\mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}. \qquad (16)$$

Recall that the *relative degree* (D_g) of a system is defined as the number of times the system's output (y) must be differentiated before the input (u) appears explicitly in its expression.

Equivalently, it is equal to the difference in the order of the polynomials (in s) found in the numerator and denominator of a system's state-space transfer function ($G(s)$). Described in a network context, D_g is the shortest path (in the graph theory sense) between the system's input/output state variables. For a system with relative degree D_g, this path will, thus, include D_g state variables, two of these being the input and output (1st and nth state variables, respectively). If we arbitrarily label the state variables corresponding to the rows of $\mathbf{A}$ such that the first $1, \dots, D_g - 1$ state variables sequentially form the shortest path to the output (nth) state variable, then requiring connectivity between the input and output elements of a network for a given $D_g = 1, \dots, n$ is equivalent to the following:

$$\mathbf{A}_{n,D_g-1} \prod_{j=2}^{D_g-1} \mathbf{A}_{j,j-1} \neq 0,$$

$$\mathbf{A}_{n,1:D_g-2} = 0. \qquad (17)$$

For example, if $D_g = 2$, the input state variable connects directly to the output ($\mathbf{A}_{n,1} \neq 0$), or if $D_g = 3$, the input state variable connects to the second state variable ($\mathbf{A}_{2,1} \neq 0$), which connects to the output ($\mathbf{A}_{n,2} \neq 0$).

Proposition V.1: Consider a linear system (9) that satisfies the assumptions of Proposition 4.1, with $\mathbf{B}, \mathbf{C}$ as in (16). Suppose that for this system, element independence (see Assumption 3.2) of $\mathbf{A}$ holds and that there exists a solution to (17) for some D_g, so that the input and output of the network are connected. If such a system has the minimal number of state variables (n) and is able to reject an input of the form $u(t) = t^{k-1}$, then it has the following properties.

1) $n = k + 2$.
2) $\mathbf{A}_{i,i} = 0$, and the product $\mathbf{A}_{i,1}\mathbf{A}_{n,i} = 0$, for all $i = 2, \dots, n - 1$.
3) $\mathbf{A}_{i,n} \neq 0$ for at least one $i = 2, \dots, n$, and $\mathbf{A}_{1,j} \neq 0$ for at least one $j = 1, \dots, n - 1$.
4) $D_g = 2$, and thus, $\mathbf{A}_{n,1} \neq 0$.

Proof: From (11), this system must satisfy

$$\lim_{s \to 0} \det \begin{bmatrix} s\mathbf{I} - \mathbf{A}_{1:n,1:n} & \begin{matrix} 1 \\ 0 \\ \vdots \end{matrix} \\ \cdots \ 0 \ \frac{1}{s^{k-1}} & 0 \end{bmatrix} = 0. \qquad (18)$$

Using Laplace's formula, this determinant can be expanded along the bottom row and up the right column to give

$$\lim_{s \to 0} \frac{1}{s^{k-1}} \times \det(\mathbf{A}_r) = 0 \qquad (19)$$

where

$$\mathbf{A}_r = \begin{bmatrix} \mathbf{A}_{2:n-1,1} & \mathbf{A}_{2:n-1,2:n-1} - s\mathbf{I} \\ \mathbf{A}_{n,1} & \mathbf{A}_{n,2:n-1} \end{bmatrix}. \qquad (20)$$

Let $\tilde{\lambda}_i(s), i = 1, \dots, n - 1$ denote the eigenvalues of the matrix $\mathbf{A}_r$, then for (19) to be satisfied, we require

$$\lim_{s \to 0} \frac{1}{s^{k-1}} \prod_{i=1}^{n-1} \tilde{\lambda}_i(s) = 0. \qquad (21)$$

We will now examine two cases, the first where one of $\tilde{\lambda}_i(s) = 0$ when s is arbitrary, and the second where this does not hold, thus requiring that a term of order s^k can be factorized out of the product of eigenvalues in (21) [so that (19) is satisfied]. We will rule out the first case due to its inability to satisfy (17) for any value of D_g, and thus conclude the second case must be true.

Case 1: If one of $\tilde{\lambda}_i(s) = 0$ when s is arbitrary, then $\mathbf{A}_{2:n,1:n-1}$ is degenerate. Expanding $\det(\mathbf{A}_r)$ along the bottom row and separating out the first term, we have

$$\mathbf{A}_{n,1} \times \det(\mathbf{A}_{2:n-1,2:n-1} - s\mathbf{I})$$
$$+ \sum_{j=2}^{n-1} (-1)^{j+1} \mathbf{A}_{n,j} \mathbf{M}_{n,j} = 0 \qquad (22)$$

where $\mathbf{M}_{i,j}$ is the standard matrix minor equal to the determinant of the $(n-2) \times (n-2)$ matrix that results from removing the ith row and jth column from $\mathbf{A}_r$. The first term in (22) is the only one proportional to s^{n-2}, and therefore, for this expression to be equal to zero for arbitrary s, we require $\mathbf{A}_{n,1} = 0$, ruling out the $D_g = 2$ case of (17).

Expanding $\det(\mathbf{A}_r)$ down the left column, the first term is given by

$$\mathbf{A}_{2,1} \times \det \begin{bmatrix} \mathbf{A}_{3:n-1,2} & \mathbf{A}_{3:n-1,3:n-1} - s\mathbf{I} \\ \mathbf{A}_{n,2} & \mathbf{A}_{n,3:n-1} \end{bmatrix}. \qquad (23)$$

Due to the assumption of element independence for $\mathbf{A}$, (23) must equal zero for all s. As the $D_g = 2$ case of (17) has already been excluded, $\mathbf{A}_{2,1} \neq 0$, and so expanding this determinant along the bottom row requires $\mathbf{A}_{n,2} = 0$, it being the coefficient of the only s^{n-3} term. Therefore, the $D_g = 3$ case of (17) is likewise excluded. Continuing to expand (23) in a similar manner [for example, in the next stage expanding the determinant in (23) about $\mathbf{A}_{3,2}$] rules out all possible values of D_g, thereby proving case 1 to be impossible.

Case 2: With case 1 excluded, it is necessary that a factor of s^k is shared between the $n-1$ eigenvalues in (21) for the limit to hold. This implies the matrix $\mathbf{A}_{2:n,1:n-1}$ is onefold degenerate (that is, it contains one linearly dependent row/column), which follows from the fact that $\mathbf{A}$ is not degenerate. This also requires that at least one of the values $\mathbf{A}_{2,n}, \ldots, \mathbf{A}_{n,n} \neq 0$, and one of $\mathbf{A}_{1,1}, \ldots, \mathbf{A}_{1,n-1} \neq 0$, to avoid row- and column-wise degeneracy of $\mathbf{A}$, respectively, which gives property 3 of Proposition 5.1.

In order for a factor of s^k to appear in the eigenvalues $\tilde{\lambda}_i$, it is necessary that there are at least k terms including s in $\mathbf{A}_r$. Since $\mathbf{A}_r$ contains $n-2$ terms linear in s, we can constrain the minimal size of the complete network (n) as $n = k + 2$, which gives property 1 of Proposition 5.1. If we express the determinant of $\mathbf{A}_r$ as a polynomial in s

$$\beta_{n-2} s^{n-2} + \beta_{n-3} s^{n-3} + \cdots + \beta_1 s^1 + \beta_0. \qquad (24)$$

Then $\beta_0, \ldots, \beta_{k-1} = 0$ in order for an s^k term to be factored out. By breaking down this determinant to calculate coefficients

of the highest order s terms, we find $\beta_{n-2} = \mathbf{A}_{n,1}$ and

$$\beta_{n-3} = \sum_{i=2}^{n-1} \mathbf{A}_{i,1} \mathbf{A}_{n,i} + \mathbf{A}_{n,1} \sum_{i=2}^{n-1} \mathbf{A}_{i,i}. \qquad (25)$$

Since $n = k + 2$ the only nonzero term in (24) is the one with coefficient $\beta_k = \mathbf{A}_{n,1}$, and thus, the network's input must be connected to its output (i.e., $D_g = 2$, yielding property 4 of Proposition 5.1). The requirement that $\beta_{k-1} = 0$ in (25) dictates that $\mathbf{A}_{i,i} = 0$ and $\mathbf{A}_{i,1} \mathbf{A}_{n,i} = 0$ for all $i = 2, \ldots, n-1$, which gives property 2 of Proposition 5.1. ∎

From a biological standpoint, the first constraint in property 2 of Proposition 5.1 corresponds to a requirement that each non-input/output species has a constant rate of degradation/dilution (i.e., zeroth-order degradation), independent of its concentration. This requirement, which has previously been established for perfectly adapting systems [50], can be realized (for example) by saturation of proteolysis—forcing the protease responsible for degradation of a given species to work at saturation [51]. The second constraint in property 2 of Proposition 5.1 means that a species i (except for the input species) is either directly influenced by the input node or it directly influences the output node, but both conditions are never simultaneously true. This is equivalent to requiring that there are no paths from input to output that go through precisely one other node. For the minimal relative degree ($D_g = 2$), the minimum size of a network that can reject inputs of order $k - 1$ can be equivalently calculated by recalling that rejecting disturbances of the form t^{k-1} requires the system's transfer function to have k zeros at the origin, and therefore, by the definition of relative degree, the system must be of minimal order $k + 2$.

Although the minimally sized network for adaptation to a given input has $D_g = 2$ via Proposition 5.1, in some situations, it may be advantageous to select a larger network in order to relax some constraints on the constituent component dynamics. With each increase in D_g, we must include an additional term in (24), and thus, the minimal size of an adaptive network for a given D_g is $n = D_g + k$.

VI. NETWORK CONSTRAINTS FOR DISTURBANCE REJECTION

To formulate a network that can reject a disturbance in one of its state variables [see Fig. 1(B)], we will consider a system in which the input and output are (arbitrarily) applied to the first state variable, giving

$$\mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}. \qquad (26)$$

In a biological sense, such a network would be able to stabilize the concentration of an output species (whose equilibrium concentration may be a function of inputs elsewhere in the network) as it is being consumed by secondary processes. This system is, therefore, able to function as a *load-driver* [3]. Such systems have been implemented in a synthetic biological context

in recent years and have been shown to provide much-needed modularity for the interfacing of biological constructs [33].

Proposition VI.1: Consider a linear system (9) that satisfies the assumptions of Proposition 4.1, with $\mathbf{B}, \mathbf{C}$ as in (26). Suppose that for this system, element independence (see Assumption 3.2) of $\mathbf{A}$ holds. If such a system has the minimal number of state variables (n) and is able to reject an input of the form $u(t) = t^{k-1}$, then it has the following properties.

1) $n = k + 1$.
2) $\mathbf{A}_{1,1} < 0$, $\mathbf{A}_{i,i} = 0$ for all $i = 2, \ldots, n$, and $\mathbf{A}_{j,i}\mathbf{A}_{i,j} = 0$ for all combinations of $i, j = 2, \ldots, n$.
3) $\mathbf{A}_{i,1} \neq 0$ and $\mathbf{A}_{1,i} \neq 0$ for at least one $i = 2, \ldots, n$.

Proof: Analogously to the process following from (18), it is necessary that

$$\lim_{s \to 0} \frac{1}{s^{k-1}} \times \det(\mathbf{A}_t) = 0 \qquad (27)$$

where

$$\mathbf{A}_t = [s\mathbf{I} - \mathbf{A}_{2:n,2:n}]. \qquad (28)$$

If we express the determinant of $\mathbf{A}_t$ as a polynomial in s, we get

$$s^{n-1} + \beta_{n-2}s^{n-2} + \cdots + \beta_1 s + \beta_0 = 0. \qquad (29)$$

To satisfy (27), an s^k term must be factored out of (29), which sets a minimal network size of $n = k + 1$ (property 1 of Proposition 6.1). Setting $n = k + 1$ in (29), we have

$$s^k + \beta_{k-1}s^{k-1} + \cdots + \beta_1 s + \beta_0 = 0 \qquad (30)$$

which requires $\beta_{k-1}, \ldots, \beta_0 = 0$ in order to leave s^k as the lowest order term. Since $\beta_0 = \det(-\mathbf{A}_{2:n,2:n})$, we have that $\mathbf{A}_{2:n,2:n}$ is degenerate. However, since $\mathbf{A}$ is invertible, $\mathbf{A}_{2:n,2:n}$ has at most one degenerate row/column, and so at least one of $\mathbf{A}_{2,1}, \ldots, \mathbf{A}_{n,1} \neq 0$ and one of $\mathbf{A}_{1,2}, \ldots, \mathbf{A}_{1,n} \neq 0$ to avoid row- and column-wise degeneracy of $\mathbf{A}$, respectively (property 3 of Proposition 6.1). In (30), the term β_{k-1} is given by

$$\beta_{k-1} = (-1)^{k-1}\sum_{i=2}^{k+1}\mathbf{A}_{i,i} \qquad (31)$$

which if $\beta_{k-1} = 0$, given the assumption of element independence for $\mathbf{A}$, gives the second part of property 2 in Proposition 6.1. Biologically, this constraint requires that all state variables bar the input/output must have zeroth-order degradation dynamics [50]. Since all eigenvalues of $\mathbf{A}$ have negative real part, $\mathrm{Trace}(\mathbf{A}) < 0$, and thus, $\mathbf{A}_{1,1} < 0$, giving the first part of property 2 in Proposition 6.1. Since $\mathbf{A}_{i,i} = 0$ for $i = 2, \ldots, n$, we have

$$\beta_{k-2} = (-1)^k\sum_{j=2}^{k+1}\sum_{i=j+1}^{k+1}\mathbf{A}_{j,i}\mathbf{A}_{i,j} \qquad (32)$$

which if $\beta_{k-2} = 0$, given again the assumption of element independence for $\mathbf{A}$, constrains every pair of nonoutput state variables to be joined only by a single link. This gives the final part of property 2 in Proposition 6.1. ∎

Considering a nonminimal network with $n = k + 2$, we still require s^k to be the lowest order term in $\det(\mathbf{A}_t)$, which is now given by

$$s^{k+1} + \beta_k s^k + \cdots + \beta_1 s + \beta_0 = 0. \qquad (33)$$

The term β_k has the same form as in (31), but with the sum index going to $k + 2$. This sum can now be nonzero, and thus, diagonal elements may be nonzero. The next coefficient in (33) is

$$\beta_{k-1} = (-1)^k \left(\sum_{j=2}^{k+2}\sum_{i=j+1}^{k+2}\mathbf{A}_{j,i}\mathbf{A}_{i,j} + \sum_{j=2}^{k+2}\sum_{i=2,i\neq j}^{k+2}\mathbf{A}_{i,i}\mathbf{A}_{j,j} \right) \qquad (34)$$

which if $n = k + 2$ must equal zero. The right-hand sum in (34) indicates that all pairwise products of diagonal terms are zero. Since the nonzero β_k allowed these to be individually nonzero, precisely one of the diagonal terms can be nonzero. The left-hand sum in (34) constrains every pair of nonoutput nodes to be joined only by a single link, as was the case before.

VII. APPLICATION TO BIOLOGICAL SYSTEM DESIGN

A. Implementation of Parameter Constraints

We now apply the derived constraints to design a synthetic biological system capable of adapting to a step ($k = 1$) disturbance. For a minimal realization of this system, following Proposition 5.1, we require a network of size $n = 3$, which in linearized form will be expressed as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \mathbf{A}_{1,3} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \mathbf{A}_{2,3} \\ \mathbf{A}_{3,1} & \mathbf{A}_{3,2} & \mathbf{A}_{3,3} \end{bmatrix}. \qquad (35)$$

From Proposition 5.1, we have $\mathbf{A}_{3,1} \neq 0$ and $\mathbf{A}_{2,2} = 0$. Since all eigenvalues of $\mathbf{A}$ have negative real part, we know $\mathbf{A}_{1,1} + \mathbf{A}_{3,3} < 0$, and so we will set both $\mathbf{A}_{1,1}, \mathbf{A}_{3,3} < 0$ to avoid constraining the relative magnitudes of these parameters (if one were positive, it would need to be lesser in magnitude than the other). In doing this, we have also satisfied property 3 of Proposition 5.1. From property 2 of Proposition 5.1, we have $\mathbf{A}_{2,1}\mathbf{A}_{3,2} = 0$, and will set $\mathbf{A}_{3,2} = 0$ as this choice allows us to set $\mathbf{A}_{2,1} \neq 0$ so that there is a nonzero element in the matrix's second row. Expanding the determinant of $\mathbf{A}$ down its middle column now gives $-\mathbf{A}_{1,2}(-|\mathbf{A}_{3,3}|\mathbf{A}_{2,1} - \mathbf{A}_{2,3}\mathbf{A}_{3,1})$, which must be less than zero for $\mathbf{A}$ to be stable (since this determinant is equal to the product of three negative eigenvalues), and so we set $\mathbf{A}_{2,3} = 0$ and $\mathbf{A}_{1,2}\mathbf{A}_{2,1} < 0$. At this point, we can set $\mathbf{A}_{1,3} = \mathbf{A}_{2,3} = 0$ to simplify the system, since these elements are not required to satisfy other constraints. By Constraint 3.3, we then have that $\mathbf{A}_{2,1} < 0$, and therefore by Constraint 3.4, $\mathbf{A}_{3,1} < 0$. Thus, we are left with

$$\mathbf{A} = \begin{bmatrix} -|\mathbf{A}_{1,1}| & |\mathbf{A}_{1,2}| & 0 \\ -|\mathbf{A}_{2,1}| & 0 & 0 \\ -|\mathbf{A}_{3,1}| & 0 & -|\mathbf{A}_{3,3}| \end{bmatrix} \qquad (36)$$

which forms part of a single-input multiple-output linear system as in (9) given by

$$\dot{\tilde{x}} = \mathbf{A}\tilde{x} + \begin{bmatrix} u_{x_1}(t) \\ 0 \\ 0 \end{bmatrix},$$

$$y = \tilde{x} \tag{37}$$

where as before $\tilde{x} = x - x_e$, and $u_{x_1}(t)$ is the time-dependent disturbance to species x_1 to which we hope x_3 will adapt. We now seek a nonlinear system that can be approximated near to an equilibrium by the constrained linear system in (37).

B. Nonlinear System Selection and Correspondence With a Linear System

To design a nonlinear system [of the form of (1)] that can be approximated by (37), we begin by selecting biological mechanisms to account for each of the elements of (36). There are many ways in which this can be done depending on the biological components selected. We will discuss one such implementation.

The elements $-|\mathbf{A}_{1,1}|$ and $-|\mathbf{A}_{3,3}|$ can be implemented by attaching fast degradation tags to x_1 and x_3, or by introducing a protease external to this system, which acts in saturation upon x_2, giving x_2 zeroth-order degradation [50]. The element $|\mathbf{A}_{1,2}|$ can be provided by having species x_2 noncooperatively activate the expression of species x_1 in the regime in which the concentration of x_2 is small compared to the activator binding equilibrium constant. Elements $-|\mathbf{A}_{2,1}|$ and $-|\mathbf{A}_{3,1}|$ can be implemented by choosing species x_1 to be a protease enzyme operating in the saturation regime (where enzyme concentration is small compared to substrate concentration).

Using the standard Hill equation (for noncooperative activation) and Michaels–Menten equation (for enzyme activity) [52], the system's nonlinear dynamics can be expressed as

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} \alpha_{x_1} - \delta_{x_1}[x_1] + \frac{\beta_{x_2}[x_2]/k_1}{1+[x_2]/k_1} + u_{x_1}(t) \\ \alpha_{x_2} - \frac{k_{cat_1}[x_1][x_2]}{K_{M_1}+[x_2]} \\ \alpha_{x_3} - \delta_{x_3}[x_3] - \frac{k_{cat_2}[x_1][x_3]}{K_{M_2}+[x_3]} \end{bmatrix} \tag{38}$$

where we have placed the state variables in square brackets to indicate that they represent concentrations. α_{x_i} represent the basal expression rates of each species, β_{x_i} represent the maximal rate increase due to activation, and δ_{x_i} represent the first-order degradation rates. K_M are Michaelis constants, and the k_{cat} the corresponding maximal enzyme catalytic rates. The structure of this system's Jacobian is given by

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\bigg|_{\mathbf{x}_e,\mathbf{0}} = \begin{bmatrix} * & * & 0 \\ * & * & 0 \\ * & 0 & * \end{bmatrix}. \tag{39}$$

To set $\mathbf{A}_{2,2} = 0$, we need the activator x_2 to operate in the regime where its concentration is small compared to its binding equilibrium constant (so that $k_1 \gg [x_2]$), making the magnitude of this partial derivative small. Furthermore, forcing the protease enzyme x_1 to operate in its saturation regime gives $K_{M_1} \ll [x_2]$ and $K_{M_2} \ll [x_3]$. With these constraints enforced, the Jacobian
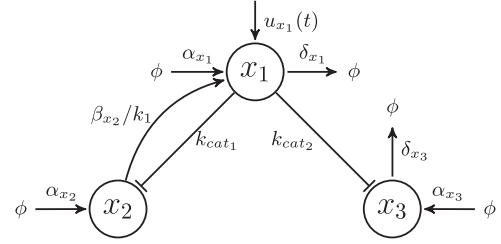

Fig. 2.   Network diagram for linear system in (37) with parameter values from the nonlinear approximation in (40). Pointed (blunt) arrows indicate a positive (negative) effect of one species' concentration on that of another.

TABLE I
SAMPLE MODEL PARAMETERS (UNITLESS)

Parameter	Value	Parameter	Value
$[x_{e_1}]$	2.0	k_{cat_1}	1.5
$[x_{e_2}]$	2.5	k_{cat_2}	1.0
$[x_{e_3}]$	1.5	k_1	80
α_{x_1}	1.5	K_{M_1}	0.01
α_{x_2}	3.0	K_{M_2}	0.01
α_{x_3}	5.0	δ_{x_1}	2.0
β_{x_2}	80	δ_{x_3}	2.0

can be approximated by

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\bigg|_{\mathbf{x}_e,\mathbf{0}} \approx \begin{bmatrix} -\delta_{x_1} & \beta_{x_2}/k_1 & 0 \\ -k_{cat_1} & 0 & 0 \\ -k_{cat_2} & 0 & -\delta_{x_3} \end{bmatrix}. \tag{40}$$

We are now able to use (2) to construct a system in the form of (37) by setting

$$\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\bigg|_{\mathbf{x}_e,\mathbf{0}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\bigg|_{\mathbf{x}_e,\mathbf{0}} = \mathbf{A},$$

$$-\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\bigg|_{\mathbf{x}_e,\mathbf{0}} \mathbf{x}_e = \begin{bmatrix} \alpha_{x_1} \\ \alpha_{x_2} \\ \alpha_{x_2} \end{bmatrix}. \tag{41}$$

C. Parameter Selection and Implementation

With the equivalences in (41) set, the nonlinear system described by (38) is approximated by the linear system, as shown in Fig. 2. To model this system, values for the parameters and ratios in (40) [which then define the elements of the linear system (36)] are selected in order to ease implementation, subject to two constraints. First, the eigenvalues of this matrix all lie in the open left half-plane as per Proposition 5.1. Second, there exists an equilibrium solution to $\mathbf{f}(\mathbf{x}_e, \mathbf{0}) = 0$ with $\mathbf{f}$ as in (38), with all elements of $\mathbf{x}_e$ and each α_{x_i} term strictly positive.

Based on the equilibrium concentrations $[\mathbf{x}_e]$, the activation binding constant is selected to satisfy $k_1 \gg [x_{e_2}]$, and the Michaelis–Menten constants to satisfy $K_{M_1} \ll [x_{e_2}]$ and $K_{M_2} \ll [x_{e_3}]$. Values for β_{x_2} and β_{x_3} are then fully defined by the ratios in (40).

The above-mentioned process was followed to yield an illustrative set of satisfactory parameters, summarized in Table I.
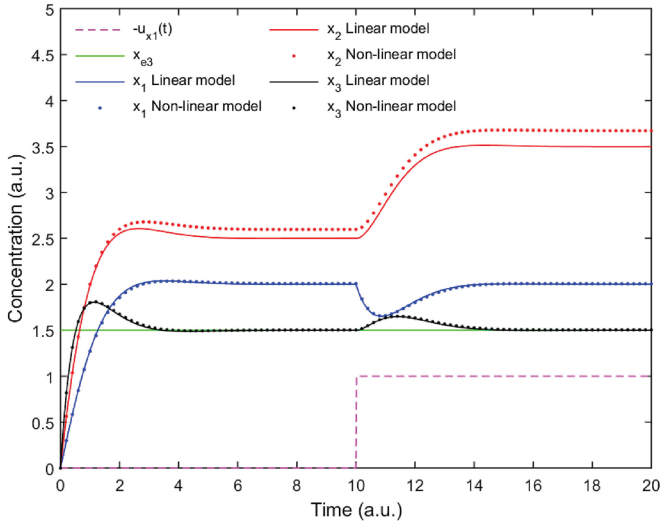
Fig. 3. Comparison of linear (37) and nonlinear (38) realizations of a system designed such that x_3 is able to adapt to step disturbances of x_1. Close agreement is found between the value of x_3 in both models and its desired equilibrium value x_{e3}. Model parameters used are presented in Table I. The negative of the time-dependent input $u_{x_1}(t)$ is plot as a rate, and state variables (x) are plot as concentrations. All values in arbitrary units (a.u.).

Using these values, we simulated the linear (37) and nonlinear (38) systems using MATLAB, for which results are presented in Fig. 3. With all state variable concentrations $[x]$ initially zero, designed equilibrium values (see Table I) are quickly reached. The time-dependent input $u_{x_1}(t)$ applies a negative unit step at $t = 10$. Though $[x_3]$ departs from equilibrium at this discontinuity in $u_{x_1}(t)$, it quickly returns to the desired value $[x_{e_3}]$, achieving the system's design purpose. Agreement between the linear and nonlinear systems is good, particularly in terms of the controlled variable x_3, which was found to be robust to variation of model parameters, though these variations can shift the system's equilibrium position. Substantial (e.g., order of magnitude) changes in k_1, K_{M_1}, K_{M_2} can reduce the applicability of assumptions made when linearizing our nonlinear implementation. This results in departure between the linear and nonlinear models, particularly in x_2, with x_1 and the controlled variable x_3 affected to a lesser extent. However, since the assumptions made in the derivation of (40) still loosely hold, the structure of the system is maintained and so adaptation by x_3 is still achieved.

While the example illustrated in this section demonstrates the application of constraints to design a minimal system for step-disturbance rejection, the result is not unique. Rather than specifying particular parameter values, the constraints outlined in this paper define the structure of $\mathbf{A}$ in terms of the sign (or existence) of its individual elements. When these constraints (and the assumptions that underlie them, such as the stability of $\mathbf{A}$) are satisfied, our results then guarantee that (due to its structural properties) the linear system is able to reject a given class of inputs. After all structural constraints are satisfied, unconstrained elements of $\mathbf{A}$ can then be set based on any other design goals; in our case, we set any uncertain elements to zero in order to minimize the number of required interactions within the system
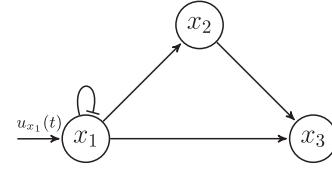

Fig. 4. Simple feedforward network structure, in which an inducer drives the steady-state concentration of x_1. By treating x_1 as a direct proxy for the input to this network, it can be thought of as an input interacting with both x_2 and x_3.

(thereby simplifying implementation). For networks that are designed to reject higher order disturbances $(k > 2)$, the structure imposed by our design constraints will be increasingly sparse, leaving many more elements undetermined (though generally setting these to zero to reduce the complexity of the designed system will be desirable).

VIII. DISCUSSION AND CONCLUSION

The constraints derived in this paper provide a general guideline for designing systems capable of adaptation and disturbance rejection, but do so under a limited set of assumptions which may, in practice, be violated. For example, the element independence requirement (that elements of $\mathbf{A}$ be independent of one-another) was investigated by Drengstig *et al.* [46], who classified adaptive networks as *robust* (those in which perfect adaptation is a property of the network's topology) or *nonrobust* (those in which fine-tuning of rate constants is necessary to achieve adaptation). Our assumptions ruled out all such nonrobust systems; however, their coarseness also eliminated a number of robust strategies. For example, the element independence assumption for $\mathbf{A}$ could be violated in practice by reversible reactions such as

$$x_1 \underset{k_1}{\overset{k_2}{\rightleftharpoons}} x_2 \tag{42}$$

which when linearized gives a matrix of the form

$$\mathbf{A} = \begin{bmatrix} -k_1 & k_2 \\ k_1 & -k_2 \end{bmatrix}. \tag{43}$$

Despite this, we elected to avoid any form of parameter dependence since systems that violate these assumptions (in the nonrobust case) only function within a limited parameter space, or (in the robust case) can be challenging to implement. For example, for reversible reactions as in (42), it may be difficult to tune the relative rate values k_1 and k_2, or to find a biological component with this property that is also able to interact with other species in the system as required. Other derived constraints can also be violated in practice (for example, the λ repressor protein CI can act as both an activator and repressor [53], violating Constraint 3.4), but were included since our goal is to limit the design space to include only the most readily implementable systems.

The present analysis can be extended to consider systems of more than one input by careful selection of node identities and layout. For example, feedforward architectures often found in natural biological systems [54] (illustrated in Fig. 4) can be constructed by having the first node (x_1) act as a proxy for the

system input. By enforcing fast dynamics of this node, it can track the input signal, though it will be subject to consumption/disturbance by downstream processes (as are inputs such as chemical inducers [55]).

The linearized approach to design described in this paper simplifies the process of defining network structures and guiding implementation, but does so at the expense of ruling out a class of nonlinear systems for which the system matrix [as in (40)] would have nonconstant terms. Such systems can be designed to adapt to a wide range of signals while requiring fewer components than those described by our linear design approach. For example, an $n = 2$ dimensional nonlinear system can adapt to all subexponential inputs (including $u(t) = t^k$ for all k) [56], [57]. However, such systems can be challenging to implement using biological components [56], and in fact for many applications, we desire adaptation or disturbance rejection for certain classes of signals but not others [58]. For example, many natural systems (including bacterial chemotaxis [59]) utilize negative feedback architectures that adapt to step inputs, but not to ramps, in order to sense nutrient gradients in their environment. This contrasts with other natural regulatory architectures, such as incoherent feedforward loops, which adapt to a wider range of signals in order to reduce a system's sensitivity on environmental fluctuations [58].

As with all mathematical descriptions of biological systems, it is important to note that our approach involves making a range of approximations, which in reality, depending on circumstance, may be violated. For example, our seemingly simple assumption that it is possible to build a network in which one species does *not* influence another (i.e., there is an off-diagonal zero in $\mathbf{A}$) may not hold: Limited ribosomal resources mean that if one gene is translationally activated, this may have a repressive influence on the translation of others [60]. Deliberate utilization of these secondary effects may provide design approaches that could (for example) violate Constraint 3.4. Furthermore, zeroth-order degradation (zero terms on the diagonal of $\mathbf{A}$) is difficult to achieve in practice, since there will always be some rate of protein dilution and passive degradation. Another approximation arises when linearizing the more biologically relevant nonlinear system, as this approximation is only applicable within a given range of parameter values, thus constraining the regime in which our system can achieve its design specification. Combined, these and other second-order effects (which are prevalent throughout biological systems) mean that biological networks can almost never be *perfectly* adapting or rejecting of disturbances, though they can approximate these capabilities [29].

In this paper, we have described a set of constraints that can be employed to simplify the design and implementation of biological networks that achieve adaptation and disturbance rejection. Using this framework, an engineer can quickly narrow the design space to a set of viable linear network topologies, which can then be embedded in a nonlinear system of biological relevance. Though biologically feasible designs are possible outside of the assumption set we have outlined, our assumptions ensure that our method provides only the most readily implementable architectures. These can then be built in the laboratory from the ever-growing array of components available to synthetic bi-

ologists. In this way, we hope that our work will aid in the realization of the robust control systems that are increasingly necessary for the regulation of complex synthetic biological constructs.

REFERENCES

[1] J. J. Collins and A. S. Khalil, "Synthetic biology: Applications come of age," *Nature Rev. Genetics*, vol. 11, no. 5, pp. 367–379, 2010.

[2] A. A. K. Nielsen, T. H. Segall-Shapiro, and C. A. Voigt, "Advances in genetic circuit design: Novel biochemistries, deep part mining, and precision gene expression," *Curr. Opin. Chem. Biol.*, vol. 17, no. 6, pp. 878–892, 2013.

[3] D. Del Vecchio, "Modularity, context-dependence, and insulation in engineered biological circuits," *Trends Biotechnol.*, vol. 33, no. 2, pp. 111–119, 2015.

[4] G. Balazsi, A. Van Oudenaarden, and J. J. Collins, "Cellular decision making and biological noise: From microbes to mammals," *Cell*, vol. 144, no. 6, pp. 910–925, 2011.

[5] T. M. Yi, Y. Huang, M. I. Simon, and J. Doyle, "Robust perfect adaptation in bacterial chemotaxis through integral feedback control," *Proc. Nat. Acad. Sci. USA*, vol. 97, no. 9, pp. 4649–4653, 2000.

[6] M. Khammash, "An engineering viewpoint on biological robustness," *BMC Biol.*, vol. 14, no. 1, 2016, Art. no. 22.

[7] D. Del Vecchio, A. J. Dy, and Y. Qian, "Control theory meets synthetic biology," *J. Roy. Soc. Interface*, vol. 13, no. 10, 2016, Art. no. 20160380.

[8] U. Alon, M. G. Surette, N. Barkai, and S. Leibler, "Robustness in bacterial chemotaxis," *Nature*, vol. 397, no. 6715, pp. 168–171, Jan. 1999.

[9] H. C. Berg and D. A. Brown, "Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking," *Nature*, vol. 239, no. 5374, pp. 500–504, 1972.

[10] P. A. Iglesias and A. Levchenko, "A general framework for achieving integral control in chemotactic biological signaling mechanisms," in *Proc. 40th IEEE Conf. Decis. Control*, 2001, pp. 843–848.

[11] W. Ma, A. Trusina, H. El-Samad, W. A. Lim, and C. Tang, "Defining network topologies that can achieve biochemical adaptation," *Cell*, vol. 138, no. 4, pp. 760–773, 2009.

[12] S. Chen, P. Harrigan, B. Heineike, J. Stewart-Ornstein, and H. El-Samad, "Building robust functionality in synthetic circuits using engineered feedback regulation," *Curr. Opin. Biotechnol.*, vol. 24, no. 4, pp. 790–796, 2013.

[13] N. Rosenfeld, M. B. Elowitz, and U. Alon, "Negative autoregulation speeds the response times of transcription networks," *J. Mol. Biol.*, vol. 323, no. 5, pp. 785–793, 2002.

[14] D. Nevozhay, R. M. Adams, K. F. Murphy, K. Josic, and G. Balázsi, "Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 13, pp. 5123–5128, 2009.

[15] R. Daniel, J. R. Rubens, R. Sarpeshkar, and T. K. Lu, "Synthetic analog computation in living cells," *Nature*, vol. 497, no. 7451, pp. 619–623, 2013.

[16] J.-C. Lin and D. Thirumalai, "Gene regulation by riboswitches with and without negative feedback loop," *Biophys. J.*, vol. 103, no. 11, pp. 2320–2330, 2012.

[17] C. Briat, A. Gupta, and M. Khammash, "Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks," *Cell Syst.*, vol. 2, no. 1, pp. 15–26, 2016.

[18] J. Ang and D. R. McMillen, "Physical constraints on biological integral control design for homeostasis and sensory adaptation," *Biophys. J.*, vol. 104, no. 2, pp. 505–515, 2013.

[19] A. W. K. Harris, J. A. Dolan, C. L. Kelly, J. Anderson, and A. Papachristodoulou, "Designing genetic feedback controllers," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 4, pp. 475–484, Aug. 2015.

[20] S. Green, "Revisiting generality in biology: Systems biology and the quest for design principles," *Biol. Philos.*, vol. 30, no. 5, pp. 629–652, 2014.

[21] N. A. Shah and C. A. Sarkar, "Robust network topologies for generating switch-like cellular responses," *PLoS Comput. Biol.*, vol. 7, no. 6, 2011, Art. no. e1002085.

[22] T. Drengstig, I. W. Jolma, X. Y. Ni, K. Thorsen, X. M. Xu, and P. Ruoff, "A basic set of homeostatic controller motifs," *Biophys. J.*, vol. 103, no. 9, pp. 2000–2010, 2012.

[23] G. Giordano, C. Cuba Samaniego, E. Franco, and F. Blanchini, "Computing the structural influence matrix for biological systems," *J. Math. Biol.*, vol. 72, no. 7, pp. 1927–1958, 2016.

[24] R. Levins, *Evolution in Changing Environments: Some Theoretical Explorations*. Princeton, NJ, USA: Princeton Univ. Press, 1968.

[25] J. M. Dambacher, H. W. Li, and P. A. Rossignol, "Relevance of community structure in assessing indeterminacy of ecological predictions," *Ecology*, vol. 83, no. 5, pp. 1372–1385, 2002.

[26] F. Blanchini, E. Franco, and G. Giordano, "Determining the structural properties of a class of biological models," in *Proc. 51st IEEE Conf. Decis. Control*, 2012, pp. 5505–5510.

[27] S. U. Wei, "Topologies needed for perfect adaptation," in *Proc. 30th China Control Conf.*, 2011, pp. 5714–5717.

[28] S. U. Wei, "Perfect adaptation of general nonlinear systems," *J. Syst. Sci. Complexity*, vol. 29, no. 1, pp. 61–73, 2016.

[29] Z. F. Tang and D. R. McMillen, "Design principles for the analysis and construction of robustly homeostatic biological networks," *J. Theor. Biol.*, vol. 408, pp. 274–289, 2016.

[30] K. Oishi and E. Klavins, "Biomolecular implementation of linear I/O systems," *IET Syst. Biol.*, vol. 5, no. 4, pp. 252–260, 2011.

[31] C. Briat, C. Zechner, and M. Khammash, "Design of a synthetic integral feedback circuit: Dynamic analysis and DNA implementation," *ACS Synth. Biol.*, vol. 5, no. 10, pp. 1108–1116, 2016.

[32] B. Schulz, "BioBricks foundation." [Online]. Available: http://biobricks.org/.

[33] D. Mishra, P. M. Rivera, A. Lin, D. Del Vecchio, and R. Weiss, "A load driver device for engineering modularity in biological networks," *Nature Biotechnol.*, vol. 32, no. 12, pp. 1268–1275, 2014.

[34] S. Waldherr, S. Streif, and F. Allgöwer, "Design of biomolecular network modifications to achieve adaptation," *IET Syst. Biol.*, vol. 6, no. 6, pp. 223–231, 2012.

[35] N. Nandagopal and M. B. Elowitz, "Synthetic biology: Integrated gene circuits," *Science*, vol. 333, no. 6047, pp. 1244–1248, 2011.

[36] R. Sawlekar, F. Montefusco, V. V. Kulkarni, and D. G. Bates, "Implementing nonlinear feedback controllers using DNA strand displacement reactions," *IEEE Trans. NanoBiosci.*, vol. 15, no. 5, pp. 443–454, Jul. 2016.

[37] J. Chappell, M. K. Takahashi, and J. B. Lucks, "Creating small transcription activating RNAs," *Nature Chem. Biol.*, vol. 11, pp. 214–220, 2015.

[38] B. Mengel, S. Krishna, M. H. Jensen, and A. Trusina, "Nested feedback loops in gene regulation," *Phys. A, Stat. Mech. Appl.*, vol. 391, nos. 1–2, pp. 100–106, 2012.

[39] T. P. Prescott and A. Papachristodoulou, "Designing conservation relations in layered synthetic biomolecular networks," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 4, pp. 572–580, Aug. 2015.

[40] P. Rivera-Ortiz and D. Del Vecchio, "Integral action with time scale separation: A mechanism for modularity in biological systems," in *Proc. IEEE Conf. Decis. Control*, 2015, pp. 49–55.

[41] A. A. K. Nielsen *et al.*, "Genetic circuit design automation," *Science*, vol. 352, no. 6281, 2016, Art. no. aac7341.

[42] B. Francis and W. M. Wonham, "The internal model principle of control theory the internal model principle of control theory," *Automatica*, vol. 12, no. 5, pp. 457–465, 1976.

[43] E. D. Sontag, "Adaptation and regulation with signal detection implies internal model," *Syst. Control Lett.*, vol. 50, no. 2, pp. 119–126, 2003.

[44] B. P. Ingalls, *Mathematical Modelling in Systems Biology: An Introduction*. Cambridge, MA, USA: MIT Press, 2014.

[45] L. Huang *et al.*, "Effects of promoter leakage on dynamics of gene expression," *BMC Syst. Biol.*, vol. 9, 2015, Art. no. 16.

[46] T. Drengstig, H. R. Ueda, and P. Ruoff, "Predicting perfect adaptation motifs in reaction kinetic networks," *J. Phys. Chem. B*, vol. 112, no. 51, pp. 16752–16758, 2008.

[47] N. S. Nise, *Control Systems Engineering*, 6th ed. Hoboken, NJ, USA: Wiley, 2010.

[48] H. H. Rosenbrock, "Transformation of linear constant system equations," *Proc. Inst. Electr. Eng.*, vol. 114, no. 4, pp. 541–544, Apr. 1967.

[49] H. H. Rosenbrock, *State-Space and Multivariable Theory*. London, U.K.: Nelson, 1970.

[50] J. Ang, S. Bagh, B. P. Ingalls, and D. R. McMillen, "Considerations for using integral feedback control to construct a perfectly adapting synthetic gene network," *J. Theor. Biol.*, vol. 266, no. 4, pp. 723–738, 2010.

[51] W. W. Wong, T. Y. Tsai, and J. C. Liao, "Single-cell zeroth-order protein degradation enhances the robustness of synthetic oscillator," *Mol. Syst. Biol.*, vol. 3, 2007, Art. no. 130.

[52] R. M. Murray and D. Del Vecchio, "Biomolecular feedback systems," Princeton, NJ, US: Princeton Univ. Press, 2014.

[53] D. Lewis, P. Le, C. Zurla, L. Finzi, and S. Adhya, "Multilevel autoregulation of lambda repressor protein CI by DNA looping in vitro," *Proc. Nat. Acad. Sci. USA*, vol. 108, no. 36, pp. 14807–14812, 2011.

[54] J. Krishnan, "Effects of saturation and enzyme limitation in feedforward adaptive signal transduction," *IET Syst. Biol.*, vol. 5, no. 3, pp. 208–219, 2011.

[55] A. W. K. Harris, H. Steel, C. L. Kelly, and A. Papachristodoulou, "The autorepressor: A case study of the importance of model selection," in *Proc. 56th IEEE Conf. Decis. Control*, 2017, pp. 1622–1627.

[56] E. D. Sontag, "A dynamic model of immune responses to antigen presentation predicts different regions of tumor or pathogen elimination," *Cell Syst.*, vol. 4, pp. 231–241, 2017.

[57] M. Lang and E. D. Sontag, "Zeros of nonlinear systems with input invariances," *Automatica*, vol. 81, pp. 46–55, 2017.

[58] H. Chang and A. Levchenko, "Adaptive molecular networks controlling chemotactic migration: dynamic inputs and selection of the network architecture," *Philos. Trans. Roy. Soc. B, Biol. Sci.*, vol. 368, no. 1629, 2013, Art. no. 20130117.

[59] C. J. Wang, A. Bergmann, B. Lin, K. Kim, and A. Levchenko, "Diverse sensitivity thresholds in dynamic signaling responses by social amoebae," *Sci. Signaling*, vol. 5, no. 213, 2012, Art. no. ra17.

[60] Y. Qian and D. Del Vecchio, "Mitigation of ribosome competition through distributed sRNA feedback," in *Proc. IEEE Conf. Decis. Control*, 2016, pp. 1–30.

Harrison Steel (S'17) received the B.S. degree in physics and mathematics and the B.Eng. degree in mechanical (space) engineering from the University of Sydney, Sydney, N.S.W., Australia, in 2016. He is currently working toward the Doctor of Philosophy degree at the University of Oxford, Oxford, U.K., supported by the General Sir John Monash Scholarship.

In 2014, he was a Research Assistant with NASA Ames Research Center; in 2015, a Research Assistant with DESY; and in 2016, an Engineer with the Quantum Nanoscience Laboratory, The University of Sydney.

Antonis Papachristodoulou (SM'15) received the M.A./M.Eng. degree in electrical and information sciences from the University of Cambridge, Cambridge, U.K., in 2000 and the Ph.D. degree in control and dynamical systems (with a minor in aeronautics) from the California Institute of Technology, Pasadena, CA, USA, in 2005.

He is currently a Professor of engineering science with the University of Oxford, Oxford, U.K., and a Tutorial Fellow with Worcester College, Oxford. He is an EPSRC Fellow for Growth in Synthetic Biology and the Director of the EPSRC and BBSRC Centre for Doctoral Training in Synthetic Biology. His research interests include large-scale nonlinear systems analysis, sum of squares programming, synthetic and systems biology, networked systems, and flow control.

Dr. Papachristodoulou was the recipient of the 2015 European Control Award for his contributions to robustness analysis and applications to networked control systems and systems biology, and the O. Hugo Schuck Best Paper Award in 2015.