Feedback Control and Synthetic Biology: Constraints on Design ⋆

Harrison C.B. Steel * **Antonis Papachristodoulou** **

* *Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK (e-mail: harrison.steel@eng.ox.ac.uk).*
** *Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK (e-mail: antonis@eng.ox.ac.uk).*

Abstract: Synthetic Biology is an emerging field at the interface of biology and engineering, concerned with the design and implementation of synthetic biological parts, devices and systems. With applications ranging from industrial biosynthesis of chemicals to treatment and prevention of disease, Synthetic Biology holds great promise, but faces several challenges due to the uncertainties and noise inherent in biological systems. In this paper we review recent progress in the design and testing of biological control systems that aim to overcome these limitations. We then use classical control theory to derive a number of design constraints for implementation of linear control systems that achieve adaptation and disturbance rejection. Finally, we design a linear system for rejection of ramp-type disturbances, and from this demonstrate how the derived linear system constraints can be embedded in a more realistic non-linear biological context.

Keywords: Bio control, Constraints, Control systems, Feedback control, Synthetic Biology.

1. INTRODUCTION

Synthetic Biology is an emerging field concerned with the rational design and engineering of biological systems. It has risen sharply in prominence following landmark results of the early 21st century, such as the demonstration of a Toggle Switch (Collins et al. (2000)), and the Repressilator (Elowitz and Leibler (2000)). Whilst these advances demonstrated an ability to re-wire natural genetic components to achieve user-defined goals, attaining reliability and consistency of performance remains challenging (Nielsen et al. (2013)). Difficulties arise for reasons including the random fluctuations and noise inherent in cellular processes (Balazsi et al. (2011)), as well as unforeseen dependencies and interactions between synthetic constructs and native cellular machinery (Del Vecchio (2015)). Study of wild-type cell behaviours has demonstrated that many such challenges are overcome in nature by systems that have evolved feedback architectures similar to those commonly used in control engineering (Yi et al. (2000)). In recent years this has inspired researchers to begin designing and implementing analogous synthetic biological control systems (Del Vecchio et al. (2016)).

The bacterial chemotaxis system provides a model example of a natural control system, and as such has attracted extensive study. In the presence of extracellular stimulus it can robustly (that is, with minimal sensitivity to properties of its constituent components, see Alon et al. (1999)) return its state to pre-stimulus levels, a process commonly known as *adaptation*. It has been demonstrated that the underlying structure of this bacterial control system contains integral feedback control (Yi et al. (2000)), which has led to the propositioning of a range of synthetic signalling pathways to achieve similar goals (Iglesias and Levchenko (2001); Ma et al. (2009)). Networks with capabilities such as this are of increasing interest to synthetic biologists as they aim to realise systems in varying and non-ideal conditions (Briat et al. (2016a)) found when attempting to move synthetic constructs out of the laboratory.

In addition to adaptation, biological networks with many favourable properties can be constructed via the inclusion of feedback loops. For example, feedback can be exploited to implement high dynamic-range gene circuits for computation (Daniel et al. (2013)). In the simple form of negative auto-regulation, it can improve the response time of gene networks (Rosenfeld et al. (2002)), and reduce heterogeneity of gene expression between cells (Nevozhay et al. (2009)). Moving beyond biological feedback, *in silico* feedback systems which measure cell behaviour, calculate appropriate control signals, and manipulate cells in response have been demonstrated (Milias-Argeitis et al. (2011); Menolascina et al. (2014)). As desired control architectures become more complex it will be increasingly necessary to develop modular biological constructs that approximate the fundamental components of traditional control systems, such as integration, gain, and summation gates (Oishi and Klavins (2011); Daniel et al. (2013)). This will facilitate implementation of standard control structures such as Lead-Lag (Harris et al. (2015)) and Proportional-Integral-Derivative (PID) variant (Briat et al. (2016b)) controllers, whose favourable properties have led to their ubiquity across engineering disciplines (Nise (2010)).

⋆ H.C.B. Steel would like to acknowledge the General Sir John Monash foundation for partial financial support of this work. A. Papachristodoulou was supported in part by EP/M002454/1.

Though the widespread need for synthetic biological control systems has been recognised (Del Vecchio et al. (2016)), and many architectures have been proposed, their implementation remains difficult (Dolan et al. (2012)). This stems from a lack of systematic approaches to design, as well as methods for optimal (or even feasible) implementation (Harris et al. (2015)). Fundamental challenges arise due to uncertainties in the properties of the constituent components of biological circuits, which are exacerbated by the inherent noise of cellular processes. Early attempts to tackle the challenge of creating such circuits focused on the standardisation of synthetic biological "parts" to allow ease (and ideally, predictability) of assembly. However, as parts libraries have grown it has become clear that the context-dependent non-ideal behaviour of almost all components means that larger constructs require substantial experimental fine-tuning, even after extensive *in silico* testing and development (Nielsen et al. (2013)). An obvious work-around for this technical challenge is to design synthetic systems that build upon (and potentially rewire) native cellular processes (Nandagopal and Elowitz (2011)) thus reducing the required network size to achieve a given functionality. Since this approach requires fewer synthetic components, it provides additional benefit by minimising the metabolic load the synthetic control system places upon its host cell (Briat et al. (2016b)).

Once a system has been designed and modelled, the next challenge facing scientists is the selection and tuning of biological parts for its implementation. This decision space is growing in complexity as new research reveals natural cellular regulation mechanisms, which are re-engineered for use in synthetic constructs. Biological systems can now be created over a range of molecular biological levels (for example, using DNA (Sawlekar et al. (2016)), RNA (Chappell et al. (2015)) or Protein (Nevozhay et al. (2009))), as well as at widely varying time-scales (Prescott and Papachristodoulou (2015); Rivera-Ortiz and Del Vecchio (2015)) and species concentrations (Briat et al. (2016b)). However, working with this diversity of components requires extensive expertise across a range of biological disciplines, and even after components are selected extensive experimental fine-tuning is often required to achieve acceptable performance. Previous work (for example, Nielsen et al. (2016)) has automated the component-selection and tuning process for some classes of synthetic constructs, making such systems more accessible to users with a limited knowledge of the underlying biology. However, these tools only treat systems with a restricted set of capabilities, and may provide a sub-optimal implementation in many situations since the automated design process does not fully account for the up- and down-stream systems with which it has to interface (Del Vecchio (2015)).

To address some of these challenges in implementation, in this paper we outline fundamental structural requirements for networks that solve two common control goals: adaptation and disturbance rejection. We provide general guidelines that can be used to simplify the creation of systems with these desirable capabilities. This is done by first designing a linearised control system via application of reasonable assumptions to aid implementation, and then deriving constraints on the network structure and parameters. Some of these results follow from the Internal Model Principle (Francis and Wonham (1976)), but are particulated for the biological systems in question. Once an appropriate linear system has been designed, we demonstrate methods for its embedding in a non-linear model of biological relevance. The final challenge, that of selecting particular biological components with appropriate behaviour and parameter values, is not addressed herein.

In Section 2 methods for modelling of biological systems are discussed, and the general approach to linearisation of a non-linear system is described. In Section 3 results are derived that govern the equilibrium behaviour of linear systems, which are then used in Sections 4 and 5 to derive minimal requirements for networks capable of adaptation and rejection of disturbance respectively. In Section 6 the results of Section 5 are used to create a network able to reject ramp-type disturbances. This network is used to design a non-linear system with these capabilities, which is simulated using illustrative parameters. Section 7 discusses assumptions made in this work and concludes the paper.

2. MODELLING BIOLOGICAL CONTROL SYSTEMS

Approaches to the modelling of biological systems range from stochastic/probabilistic analyses to deterministic differential equation (DE) models (Ingalls (2014)). Here we focus on a subset of the later, utilising both linear and non-linear first-order differential equations to describe the dynamics of individual state variables, each of which may represent individual species, or larger scale properties, of biological systems. Each first order equation will be a (potentially non-linear) function of the system's state variables and any external inputs/disturbances to the system.

A general first-order non-linear system that describes the dynamics of n state variables $\boldsymbol{x}(t) = [x_1(t), \ldots, x_n(t)] \in \mathbb{R}^n$ can be expressed in 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 of time-dependent inputs to the non-linear system's state variables, and $\mathbf{f}(\boldsymbol{x}(t), \mathbf{u}(t)) \in \mathbb{R}^n$ contains n non-linear functions of the state variables and inputs.

To simplify analysis it can be advantageous to examine non-linear systems in a regime in which they can be approximated by a linear system. The system in (1) can be linearised via Taylor series expansion of $\mathbf{f}(\boldsymbol{x}, \mathbf{u})$ about one of its equilibrium points $\boldsymbol{x}_e \in \mathbb{R}^n$ (so, $\mathbf{f}(\boldsymbol{x}_e, \mathbf{0}) = \mathbf{0}$), to give (disregarding 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 $\boldsymbol{x}_e$.

In this work we consider the above linearisation process in the opposite direction: We first derive the necessary form of $\left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_e, \mathbf{0}}$, and then attempt to find a non-linear 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.

3. GENERAL LINEARISED NETWORK STRUCTURE

We can express a general linear system in the form

$$\begin{aligned} \dot{\boldsymbol{x}} &= \mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{u} \\ \boldsymbol{y} &= \mathbf{C}\boldsymbol{x} + \mathbf{D}\boldsymbol{u} \end{aligned} \tag{3}$$

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 2), $\mathbf{u}(t) \in \mathbb{R}^m$ is the system's input and $\mathbf{y}(t) \in \mathbb{R}^p$ the system output. For SISO (single input single output) systems as considered in Sections 3, 4 and 5 of this paper, $m = p = 1$ (for an example with greater input/output dimension see (10)). $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are constant matrices of appropriate corresponding dimensions that define the system, with individual elements in $\mathbb{R}$. In this paper we will generally take $\mathbf{D} = 0$, meaning the output of the system is not a direct function of the input.

Proposition 3.1. Consider a SISO system in the form of (3) with n state variables, $\mathbf{D} = 0$, and in which $\mathbf{A}$ has all of its eigenvalues in the open left-half plane (and is hence asymptotically stable). 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 independent of $u(t)$.
2. $$\lim_{s\to 0} \det \begin{bmatrix} s\mathbf{I} - \mathbf{A} & \mathbf{B} \\ \mathbf{C}/s^{k-1} & 0 \end{bmatrix} = 0 \tag{4}$$
   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 Presented in Steel and Papachristodoulou (2017), omitted here for brevity. □

In the following sections sub-matrix notation will be as follows:

- $\mathbf{A}_{i,j}$ refers to the element in the ith row and jth column of $\mathbf{A}$
- $\mathbf{A}_{i:j,k:l}$ refers to a sub-matrix 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 $\mathbf{A}$.

4. GENERAL NETWORK FOR ADAPTATION

For a system to achieve adaptation the long-term behaviour 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

$$\mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ \vdots \end{bmatrix}, \mathbf{C} = [\,\cdots\ 0\ 1\,]. \tag{5}$$

Definition: Degree of Connectivity (D_g)
If the output of a network with $\mathbf{B}, \mathbf{C}$ as in (5) is to be influenced in some way by the input, there must be a path of connection (in the graph theory sense) between input/output state variables. The length of the shortest such path will be referred to as the Degree of Connectivity (D_g). This length will be defined by the number of *connections* made between state variables, and thus a path of length D_g will include D_g+1 state variables, two of these being the input and output (1st and nth state variables respectively). If the rows of $\mathbf{A}$ are arbitrarily re-arranged such that the first $1, \ldots, D_g$ state variables sequentially form the shortest path to the output (nth) state variable, then the requirement for connectivity is equivalent to the following for $D_g = 1, \ldots, n-1$

$$\mathbf{A}_{n,D_g} \prod_{j=2}^{D_g} \mathbf{A}_{j,j-1} \neq 0, \quad \mathbf{A}_{n,1:D_g-1} = 0 \tag{6}$$

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

Proposition 4.1. Consider a linear system that satisfies the assumptions of Proposition 3.1, with $\mathbf{B}, \mathbf{C}$ as in (5). Suppose that for this system we are unable to set any element of $\mathbf{A}$ to be a function of other elements of $\mathbf{A}$, and that there exists a solution to (6) for some D_g, so that the input and output of the network are connected. The system with the minimal number of state variables (n) that is able to reject an input of the form $u(t) = t^{k-1}$ 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 $i = 2, \ldots, n-1$
3. $\mathbf{A}_{i,n} \neq 0$ for at least one $i = 2, \ldots, n$, and $\mathbf{A}_{1,j} \neq 0$ for at least one $j = 1, \ldots, n-1$.
4. $D_g = 1$, and thus $\mathbf{A}_{n,1} \neq 0$.

Proof Presented in Steel and Papachristodoulou (2017), omitted here for brevity. □

From a biological standpoint, the first constraint in property 2 of Proposition 4.1 means that the degradation/dilution rate of each non-input/output species must be small compared to the other dynamics of our synthetic system. The second constraint dictates that a species i (except for the input species) can not simultaneously be influenced by the input node while directly influencing the output node. This is equivalent to requiring that there are no paths from input to output that go through precisely one other node.

5. GENERAL NETWORK FOR DISTURBANCE REJECTION

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

$$\mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ \vdots \end{bmatrix}, \mathbf{C} = [\,1\ 0\ \cdots\,]. \tag{7}$$

In a biological sense, such a network would be able to stabilise 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*. 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 systems (Mishra et al. (2014)).

Proposition 5.1. Consider a linear system that satisfies the assumptions of Proposition 3.1, with $\mathbf{B}, \mathbf{C}$ as in (7). Suppose that for this system we are unable to set any element of $\mathbf{A}$ to be a function of other elements of $\mathbf{A}$. The system with the minimal number of state variables (n) that is able to reject an input of the form $u(t) = t^{k-1}$ 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 Presented in Steel and Papachristodoulou (2017), omitted here for brevity. □

6. APPLICATION TO BIOLOGICAL SYSTEM DESIGN

6.1 Implementation of parameter constraints

We now apply the above constraints to design a synthetic biological system capable of rejecting a ramp ($k = 2$) disturbance. To achieve a physically realisable system we desire at least one negative element in each row of $\mathbf{A}$, such that an equilibrium can be achieved without necessitating a negative constant input. For a minimal realisation of this system, following Proposition 5.1 we require a network of size $n = 3$, which in linearised form will be expressed as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \mathbf{A}_{1,3} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \mathbf{A}_{2,3} \\ \mathbf{A}_{3,1} & \mathbf{A}_{3,2} & \mathbf{A}_{3,3} \end{bmatrix} \tag{8}$$

Again from Proposition 5.1 we have $\mathbf{A}_{2,2} = \mathbf{A}_{3,3} = 0$ and $\mathbf{A}_{2,3}\mathbf{A}_{3,2} = 0$, for which we will arbitrarily set $\mathbf{A}_{2,3} = 0$, as this choice is symmetric under a change of species index. To avoid degeneracy of $\mathbf{A}$ we then require $\mathbf{A}_{3,1} \neq 0$, and to enforce stability the determinant is constrained as $\mathbf{A}_{3,1}\mathbf{A}_{1,2}\mathbf{A}_{2,3} < 0$ (since the determinant equals the product of three eigenvalues with negative real parts) which can be enforced by setting $\mathbf{A}_{3,1} < 0$. We are left with $\mathbf{A}_{2,1}$ and $\mathbf{A}_{1,3}$ unconstrained, and so will set $\mathbf{A}_{2,1} < 0$ such that there is a negative rate in the second row, and $\mathbf{A}_{1,3} = 0$ to simplify the system. Thus we are left with

$$\mathbf{A} = \begin{bmatrix} -|\mathbf{A}_{1,1}| & |\mathbf{A}_{1,2}| & 0 \\ -|\mathbf{A}_{2,1}| & 0 & |\mathbf{A}_{2,3}| \\ -|\mathbf{A}_{3,1}| & 0 & 0 \end{bmatrix} \tag{9}$$

which forms part of a single-input multiple-output linear system as in (3) given by

$$\begin{aligned} \dot{\tilde{\boldsymbol{x}}} &= \mathbf{A}\tilde{\boldsymbol{x}} + \begin{bmatrix} u_{x_1}(t) \\ 0 \\ 0 \end{bmatrix} \\ \boldsymbol{y} &= \tilde{\boldsymbol{x}} \end{aligned} \tag{10}$$

where as before $\tilde{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{x}_e$, and $u_{x_1}(t)$ is the time-dependent disturbance to species x_1 which we hope to reject. We now seek a non-linear system which can be approximated near to an equilibrium by the constrained linear system in (10).

6.2 Non-linear system selection and correspondance with linear system

To design a non-linear system (in the form of (1)) that can be approximated by (10) we begin by selecting biological mechanisms to account for each of the elements of (9). There are many ways in which this can be done depending on the biological components selected, and we will describe one such implementation.

The element $-\mathbf{A}_{1,1}$ can be easily implemented via a fast degradation tag on x_1, whilst elements $\mathbf{A}_{1,2}$ and $\mathbf{A}_{2,3}$ can be provided by having species x_2 and x_3 non-cooperatively activate the expression of species x_1 and x_2 respectively, with both in the regime for which their concentration is small compared to the activator binding equilibrium constant. Elements $-\mathbf{A}_{2,1}$ and $-\mathbf{A}_{3,1}$ provide the greatest difficulty in the implementing of this system (as negative non-diagonal terms do in general), but can be satisfied by requiring species x_1 be able to prevent species x_2 and x_3 from being able to function as activators. Biologically, this could be achieved by choosing x_1 to be an enzyme operating in the saturation regime (where enzyme concentration is small compared to that of the substrate) that is able to degrade or otherwise inactivate x_2 and x_3. As an alternative, an annihilation reaction between species x_1 and both x_2 and x_3 could be used which would provide a constant term in the linearisation, provided the system's design goal of keeping x_1 constant was achieved.

Using the standard Hill equation (for non-cooperative activation) and Michaels-Menten equation (for enzyme activity) (Murray and Del Vecchio (2010)) the system's non-linear dynamics can therefore be expressed as

$$\mathbf{f}(\boldsymbol{x}, \mathbf{u}) = \begin{bmatrix} \alpha_{x_1} - \delta_1[x_1] + \dfrac{\beta_{x_2}[x_2]/k_1}{1 + [x_2]/k_1} + u_{x_1}(t) \\ \alpha_{x_2} - \dfrac{k_{cat_1}[x_1][x_2]}{K_{M_1} + [x_2]} + \dfrac{\beta_{x_3}[x_3]/k_2}{1 + [x_3]/k_2} \\ \alpha_{x_3} - \dfrac{k_{cat_2}[x_1][x_3]}{K_{M_2} + [x_3]} \end{bmatrix} \tag{11}$$

where we have placed the state variables in square brackets to indicate that they represent concentrations. The α_{x_i} represent basal expression rates of each species, and β_{x_i} the maximal rate increase due to activation. We want activators x_2 and x_3 to operate in the regime where their concentration is small compared to their binding equilibrium constants ($k_1 \gg [x_2]$ and $k_2 \gg [x_3]$ respectively), whilst the enzyme x_1 should operate in the saturation regime (so $K_{M_1} \ll [x_2]$ and $K_{M_2} \ll [x_3]$). This allows the non-linear dynamics to be simplified to

$$\mathbf{f}(\boldsymbol{x}, \mathbf{u}) \approx \begin{bmatrix} \alpha_{x_1} - \delta_1[x_1] + \beta_{x_2}[x_2]/k_1 + u_{x_1}(t) \\ \alpha_{x_2} - k_{cat_1}[x_1] + \beta_{x_3}[x_3]/k_2 \\ \alpha_{x_3} - k_{cat_2}[x_1] \end{bmatrix} \tag{12}$$

From this the Jacobian can be calculated as

$$\left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_e, \mathbf{0}} = \begin{bmatrix} -\delta_1 & \beta_{x_2}/k_1 & 0 \\ -k_{cat_1} & 0 & \beta_{x_3}/k_2 \\ -k_{cat_2} & 0 & 0 \end{bmatrix} \tag{13}$$

We are now able to use (2) to construct a system in the form of (10) by setting

$$\left.\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\right|_{\boldsymbol{x}_e, \mathbf{0}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_e, \mathbf{0}} = \mathbf{A}, -\left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_e, \mathbf{0}} \boldsymbol{x}_e = \begin{bmatrix} \alpha_{x_1} \\ \alpha_{x_2} \\ \alpha_{x_2} \end{bmatrix} \tag{14}$$

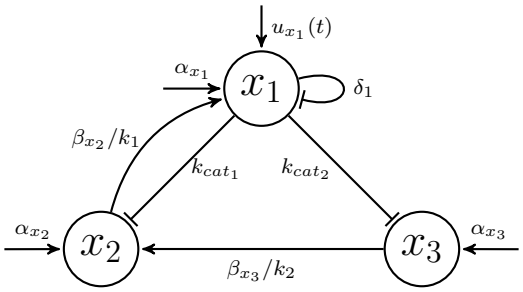

Fig. 1. Network diagram for linear system in (10) with parameter values from the non-linear approximation in (13). Pointed (blunt) arrows indicate a positive (negative) affect of one species' concentration on that of another.

6.3 Parameter selection and implementation

With the equivalences in (14) set, the non-linear system described by (11) is approximated by a linear system as in Figure 1. To model this system, values for the parameters and ratios in (13) (which then define the elements of the linear system (9)) are selected in order to ease implementation, subject to two constraints. First, that the eigenvalues of this matrix all lie in the open left half plane as per Proposition 5.1. Second, that there exists an equilibrium solution to $\mathbf{f}(\boldsymbol{x}_e, \mathbf{0}) = 0$ with $\mathbf{f}$ as in (12), with all elements of $\boldsymbol{x}_e$ and each α_{x_i} term strictly positive.

Table 1. Sample model parameters (unitless)

Parameter	Value	Parameter	Value	Parameter	Value
$[x_{e1}]$	1.5	α_{x_1}	0.5	k_{cat_1}	1.5
$[x_{e2}]$	2.5	α_{x_2}	1.0	k_{cat_2}	1.0
$[x_{e3}]$	2.5	α_{x_3}	1.5	k_1	80
β_{x_2}	80	K_{M_1}	0.01	k_2	80
β_{x_3}	40	K_{M_2}	0.01	δ_1	2.0

Based on the equilibrium concentrations $[\boldsymbol{x}_e]$ the activation binding constants are selected to satisfy $k_1 \gg [x_{e2}]$ and $k_2 \gg [x_{e3}]$, and the Michaelis-Menten constants to satisfy $K_{M_1} \ll [x_{e2}]$ and $K_{M_2} \ll [x_{e3}]$. Values for β_{x_2} and β_{x_3} are then fully defined by the ratios in (13). This process was followed to yield an illustrative set of satisfactory parameters, summarised in Table 1.

Using these values we simulated the linear (10) and non-linear (11) systems using MATLAB, for which results are presented in Figure 2. With all state variable concentrations $[\boldsymbol{x}]$ initially zero, designed equilibrium values (Table 1) are quickly reached. The time-dependent input $u_{x_1}(t)$ applies a negative unit step input at $t = 20$, and from $t = 50$ a linearly decreasing ramp. Though $[x_1]$ departs from equilibrium at each discontinuity of $u_{x_1}(t)$, it quickly returns to the desired value $[x_{e1}]$, achieving the system's design purpose. Agreement between the linear and non-linear systems is good, particularly in terms of the controlled variable x_1, 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 $k_1, k_2, K_{M_1}, K_{M_2}$ can reduce the applicability of assumptions made when linearising our non-linear implementation. This results in departure

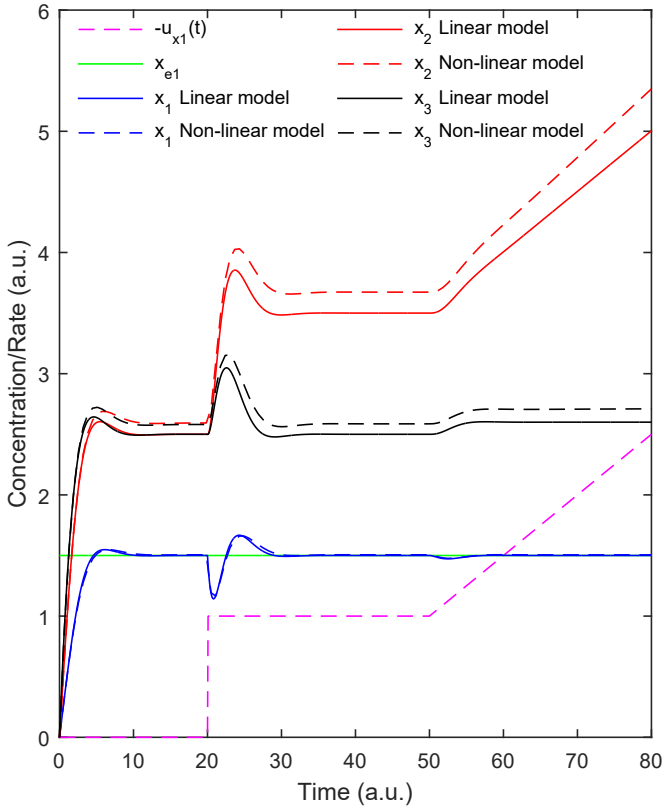

Fig. 2. Comparison of linear (10) and non-linear (11) realisations of a system designed to reject step and ramp disturbances to x_1. Close agreement is found between the value of x_1 in both models and its desired equilibrium value x_{e1}. Model parameters used are as in Table 1. The negative of the 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.).

between the linear and non-linear models, particularly in x_2 and x_3, with the controlled variable x_1 affected to a lesser extent. However, since the assumptions made in the derivation of (12) still loosely hold, the structure of the system is maintained and so rejection of disturbance is still achieved.

7. CONCLUSIONS

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 requirement that elements of $\mathbf{A}$ be independent of one-another can be violated by transformation reactions between state variables such as

$$x_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} x_2 \tag{15}$$

which results in a matrix of the form

$$\mathbf{A} = \begin{bmatrix} -k_1 & k_2 \\ k_1 & -k_2 \end{bmatrix} \tag{16}$$

However, in many cases such reactions can prove challenging to implement within the context of a larger system. For example, it may be difficult to find a biological component with this property that is also able to activate/repress other species in the system. Investigating the properties

of systems in which these assumptions are relaxed will be the topic of future work, and will hopefully lead to the experimental realisation of the synthetic biological circuits described in this work.

Whilst this paper has not addressed the challenge of choosing specific components for synthetic biological control systems, the methods described herein provide a starting framework to simplify this design task. An engineer using these results can avoid simulating and testing a variety of network topologies, instead starting with a valid network structure to which biological components can be fit. As discussed in Section 1 this can be done using an ever-growing tool-kit of biological components, which will aid in the realisation of the complex control structures required for regulation of synthetic biological systems.

REFERENCES

Alon, U., Surette, M.G., Barkai, N., and Leibler, S. (1999). Robustness in bacterial chemotaxis. *Nature*, 397(6715), 168–171.

Balazsi, G., Van Oudenaarden, A., and Collins, J.J. (2011). Cellular decision making and biological noise: From microbes to mammals. *Cell*, 144(6), 910–925.

Briat, C., Gupta, A., and Khammash, M. (2016a). Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell Systems*, 2(1), 15–26.

Briat, C., Zechner, C., and Khammash, M. (2016b). Design of a synthetic integral feedback circuit: dynamic analysis and DNA implementation. *ACS Syn. Biol.*, 5(10), 1108–1116.

Chappell, J., Takahashi, M.K., and Lucks, J.B. (2015). Creating small transcription activating RNAs. *Nat. Chem. Biol.*, 11, 1–9.

Collins, J.J., Gardner, T.S., and Cantor, C.R. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767), 339–342.

Daniel, R., Rubens, J.R., Sarpeshkar, R., and Lu, T.K. (2013). Synthetic analog computation in living cells. *Nature*, 497(7451), 619–623.

Del Vecchio, D. (2015). Modularity, context-dependence, and insulation in engineered biological circuits. *Trends Biotechnol.*, 33(2), 111–119.

Del Vecchio, D., Dy, A.J., and Qian, Y. (2016). Control theory meets synthetic biology. *J. Roy. Soc.*, 13(120), 3–43.

Dolan, J., Anderson, J., and Papachristodoulou, A. (2012). A loop shaping approach for designing biological circuits. *IEEE Decis. Contr. P.*, 3614–3619.

Elowitz, M.B. and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767), 335–338.

Francis, B. and Wonham, W.M. (1976). The internal model principle of control theory the internal model principle of control theory. *Automatica*, 12(5), 457–465.

Harris, A.W.K., Dolan, J.A., Kelly, C.L., Anderson, J., and Papachristodoulou, A. (2015). Designing genetic feedback controllers. *IEEE T. Biomed. Circ. S.*, 9(4), 475–484.

Iglesias, P.A. and Levchenko, A. (2001). A general framework for achieving integral control in chemotactic biological signaling mechanisms. *IEEE Decis. Contr. P.*, 843–848.

Ingalls, B.P. (2014). Mathematical Modelling in Systems Biology : An Introduction. *MIT Press*.

Ma, W., Trusina, A., El-Samad, H., Lim, W.A., and Tang, C. (2009). Defining network topologies that can achieve biochemical adaptation. *Cell*, 138(4), 760–773.

Menolascina, F., Fiore, G., Orabona, E., De Stefano, L., Ferry, M., Hasty, J., di Bernardo, M., and di Bernardo, D. (2014). In-vivo real-time control of protein expression from endogenous and synthetic gene networks. *PLoS Comput. Biol.*, 10(5), e1003625.

Milias-Argeitis, A., Summers, S., Stewart-Ornstein, J., Zuleta, I., Pincus, D., El-Samad, H., Khammash, M., and Lygeros, J. (2011). In silico feedback for in vivo regulation of a gene expression circuit. *Nat. Biotechnol.*, 29(12), 1114–6.

Mishra, D., Rivera, P.M., Lin, A., Del Vecchio, D., and Weiss, R. (2014). A load driver device for engineering modularity in biological networks. *Nat. Biotechnol.*, 32(12), 1268–1275.

Murray, R.M. and Del Vecchio, D. (2010). Biomolecular Feedback Systems. *Bernoulli*.

Nandagopal, N. and Elowitz, M.B. (2011). Synthetic biology: integrated gene circuits. *Science*, 333(6047), 1244–1248.

Nevozhay, D., Adams, R.M., Murphy, K.F., Josic, K., and Balázsi, G. (2009). Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression. *PNAS*, 106(13), 5123–8.

Nielsen, A.A.K., Der, B.S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E.A., Ross, D., Densmore, D., and Voigt, C.A. (2016). Genetic circuit design automation. *Science*, 352(6281), aac7341.

Nielsen, A.A.K., Segall-Shapiro, T.H., and Voigt, C.A. (2013). Advances in genetic circuit design: Novel biochemistries, deep part mining, and precision gene expression. *Curr. Opin. Chem. Biol.*, 17(6), 878–892.

Nise, N.S. (2010). Control Systems Engineering, 6th Ed. *Wiley*.

Oishi, K. and Klavins, E. (2011). Biomolecular implementation of linear I/O systems. *IET Syst. Biol.*, 5(4), 252–260.

Prescott, T.P. and Papachristodoulou, A. (2015). Designing conservation relations in layered synthetic biomolecular networks. *IEEE Trans. Biomed. Circuits Syst.*, 9(4), 572–580.

Rivera-Ortiz, P. and Del Vecchio, D. (2015). Integral action with time scale separation: A mechanism for modularity in biological systems. *IEEE Decis. Contr. P.*, 49–55.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative autoregulation speeds the response times of transcription networks. *J. Mol. Biol.*, 323(5), 785–793.

Sawlekar, R., Montefusco, F., Kulkarni, V.V., and Bates, D.G. (2016). Implementing nonlinear feedback controllers using DNA strand displacement reactions. *IEEE Trans Nanobioscience*.

Steel, H.C.B. and Papachristodoulou, A. (2017). Constraints for biological systems that achieve adaptation and disturbance rejection. *Submitted paper*.

Yi, T.M., Huang, Y., Simon, M.I., and Doyle, J. (2000). Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *PNAS*, 97(9), 4649–53.