A hidden integral structure endows absolute concentration robust systems with resilience to dynamical concentration disturbances: A hidden integral structure endows absolute

INTERFACE

royalsocietypublishing.org/journal/rsif

Research

Cite this article: Cappelletti D, Gupta A, Khammash M. 2020 A hidden integral structure endows absolute concentration robust systems with resilience to dynamical concentration disturbances. *J. R. Soc. Interface* **17**: 20200437.
http://dx.doi.org/10.1098/rsif.2020.0437

Received: 4 June 2020
Accepted: 2 October 2020

Subject Category:
Life Sciences–Mathematics interface

Subject Areas:
bioengineering, biomathematics, synthetic biology

Keywords:
integral feedback, reaction networks, absolute concentration robustness, insulators

Author for correspondence:
Mustafa Khammash
e-mail: mustafa.khammash@bsse.ethz.ch

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5172402.

THE ROYAL SOCIETY PUBLISHING

A hidden integral structure endows absolute concentration robust systems with resilience to dynamical concentration disturbances

Daniele Cappelletti, Ankit Gupta and Mustafa Khammash

Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26 4058 Basel, Switzerland

DC, 0000-0003-4259-2772

Biochemical systems that express certain chemical species of interest at the same level at any positive steady state are called 'absolute concentration robust' (ACR). These species behave in a stable, predictable way, in the sense that their expression is robust with respect to sudden changes in the species concentration, provided that the system reaches a (potentially new) positive steady state. Such a property has been proven to be of importance in certain gene regulatory networks and signaling systems. In the present paper, we mathematically prove that a well-known class of ACR systems studied by Shinar and Feinberg in 2010 hides an internal integral structure. This structure confers these systems with a higher degree of robustness than was previously known. In particular, disturbances much more general than sudden changes in the species concentrations can be rejected, and robust perfect adaptation is achieved. Significantly, we show that these properties are maintained when the system is interconnected with other chemical reaction networks. This key feature enables the design of insulator devices that are able to buffer the loading effect from downstream systems—a crucial requirement for modular circuit design in synthetic biology. We further note that while the best performance of the insulators are achieved when these act at a faster timescale than the upstream module (as typically required), it is not necessary for them to act on a faster timescale than the downstream module in our construction.

1. Introduction

The network of chemical interactions of a biochemical system of interest can be complex and involve unknown reaction propensities. One of the main goals of reaction network theory consists of deriving dynamical properties from simpler graphical properties of the model, and independently of the specific value of kinetic parameters [1,2]. The results presented in this paper follow this approach.

A qualitative property of great interest is the capability of a certain chemical species to be expressed with the same concentration at any positive steady state, independently of the initial conditions and of how many steady states are present. Namely, assume that the dynamics of the biochemical system are expressed by the following d-dimensional ordinary differential equation, which may have an infinite number of steady states

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t)).$$

We say that the ith species is *absolute concentration robust* (ACR), if there exists an *ACR value* q independent of the initial condition $x(0)$ such that, whenever $x(t)$ tends to a positive vector $\overline{x}$, we have $\overline{x}_i = q$. In the typical cases of interest, the positive steady state $\overline{x}$ that is reached will depend on the initial condition $x(0)$, while the entry $\overline{x}_i = q$ does not. As noted in [3], the property of absolute concentration

robustness alone does not imply stability of the positive steady states: it only ensures that if a positive steady state $\overline{x}$ exists, then the value of the ACR species at $\overline{x}$ is the ACR value. Under the assumption of stability, ACR provides a reliable, predictive response to environmental changes, since the species of interest reaches the steady-state level corresponding to the new environmental setting, regardless of the previous conditions. The existence and importance of this robustness property for various gene regulatory networks and signal transduction cascades is explored in many papers, including [4–12].

To achieve robustness with respect to some disturbance (in this case changes of initial conditions), the deviations from the ACR value caused by the disturbance needs to be measured first. To this aim, a quantity of interest in the control theory setting is the *integrator*, which is a function ϕ of the system variables whose derivative is exactly the difference between the concentration of the ACR species and its ACR value. At steady state, the derivative of ϕ is zero and so needs to be the distance from the ACR value. Unfortunately, in general such an integrator cannot be found, as shown in §4.1 and as discussed in [13]. However, we will show that a more relaxed concept of integrators can be fruitfully used in this setting.

In the present paper, we systematically study for the first time the connection between ACR systems and integrators. Specifically, our first contribution is related to the existence of a linear combination of chemical species whose derivative is the difference between the ACR species and its ACR value, multiplied by a monomial. Such a linear combination of species is called *constrained integrator* (CI), because it behaves similarly to an integrator given that the monomial does not vanish [13]. We rigorously prove that such a linear CI always exists for a class of models that strictly includes the ACR systems introduced in [14]. This result has some important consequences: first of all, under the assumption of stability, it implies that the expression of ACR species is not only robust to changes in the initial conditions, but also to disturbances that are applied over time.

An important application in synthetic biology concerns the design of *insulators*. A number of biochemical systems are known to express a specific output if given a certain input. The systems can therefore be considered as modules with different functions. In cells, different modules are combined so that more complex responses to external stimuli become possible [15]. In synthetic biology, it is desirable to combine different modules to achieve the same level of complexity [16]. However, when connected, the different modules can affect the dynamics of each other and they can lose the desired dynamical properties they had when considered in isolation [17]. In a simplified framework, an *upstream* module processes an input, and its output is fed to a second, *downstream* module to be further processed. Since the information is passed in the form of molecules, which are then consumed or temporarily sequestrated by the downstream module, the equations governing the upstream module dynamics are perturbed and its functionality can be affected. This effect is commonly called the *loading effect* [18] or *retroactivity* [17,19], and needs to be minimized. In other words, the upstream module needs to be *insulated* from the loading effect caused by the downstream module. We propose two ways in which the robustness of the systems studied in this paper can be used to this aim. The first solution is to simply design an upstream module which is robust to loading effects, modelled as a persistent disturbance over time. The second solution is to design an extra component, called an *insulator*, which transfers the signal from the upstream module to the downstream module while at the same time shielding the dynamics of the upstream module from retroactivity effects.

We will also show how more theoretical results on reaction network models can be obtained as a consequence of our work. In reaction network theory, the study of steady-state invariants constitute an interesting topic of research [1,14,20,21]. In [14], it has been proven that certain graphical properties of the network imply the existence of an ACR species, regardless of the choice of kinetic parameters. Such sufficient conditions are generalized in the present work while they remain simple to check. Moreover, no method to explicitly determine the ACR value was given in [14], and we fill the gap by proposing a fast linear method to calculate it. Furthermore, a substantial effort in the reaction network community is devoted to understanding the conditions under which the dynamical properties of single systems can be extended to larger systems [22–26]. Notably, extensions of ACR systems are addressed in [27]. Our contribution in this sense consists of proving that, under certain conditions, if an ACR system of the class studied in this paper is part of a larger model, the ACR species is still ACR in the larger system and its ACR value is maintained. Finally, it is worth mentioning that in the present work, we consider the possibility of time-dependent rates for the occurrence of chemical transformations. This is more general than what is usually studied in reaction network theory, with the exception of few works explicitly allowing for this scenario [28–33].

2. Examples of absolute concentration robust systems

2.1. An illustrative example

Consider two proteins A and B, whose interaction is described by

$$\begin{aligned} A + B &\xrightarrow{\kappa_1} 2B \\ B &\xrightarrow{\kappa_2} A \end{aligned} \tag{2.1}$$

where the positive constants κ_1 and κ_2 describe the propensity of a reaction to occur. If enough proteins are present and they are homogeneously spread in space, then a good model for the time evolution of the concentrations of proteins A and B is given by mass-action kinetics. Specifically, the concentrations of A and B at time t, denoted by $x_A(t)$ and $x_B(t)$, respectively, are assumed to solve

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_A(t) \\ x_B(t) \end{pmatrix} = \kappa_1 x_A(t) x_B(t) \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \kappa_2 x_B(t) \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{2.2}$$

It is easy to check that the steady states of (2.2) are given by states $(\overline{x}_A, \overline{x}_B)$ such that either $\overline{x}_B = 0$ or $\overline{x}_A = \kappa_2/\kappa_1$. Hence, A is an ACR species because its expression at any positive steady state is the same. It is common during biochemical experiments to be able to control the inflow rate of some species (say B). Some additional chemical species may also be introduced, with the purpose of degrading some of the present components (in this case, species C is introduced to faster degrade species B). After

$$\begin{aligned}
&\text{EnvZ-D} \underset{\kappa_2[\text{ADP}]}{\overset{\kappa_1}{\rightleftharpoons}} \text{EnvZ} \underset{\kappa_4}{\overset{\kappa_3[\text{ATP}]}{\rightleftharpoons}} \text{EnvZ-T} \xrightarrow{\kappa_5} \text{EnvZ-P}\\
&\text{EnvZ-P} + \text{OmpR} \underset{\kappa_7}{\overset{\kappa_6}{\rightleftharpoons}} \text{EnvZ-OmpR-P} \xrightarrow{\kappa_8} \text{EnvZ} + \text{OmpR-P}\\
&\text{EnvZ-D} + \text{OmpR-P} \underset{\kappa_{10}}{\overset{\kappa_9}{\rightleftharpoons}} \text{EnvZ-OmpR-D-P} \xrightarrow{\kappa_{11}} \text{EnvZ-D} + \text{OmpR}
\end{aligned}$$

Figure 1. Proposed model for the EnvZ-OmpR signal transduction system in *Escherichia coli*, which is able to explain the experimentally observed robustness in the expression of phosphorylated OmpR. In the first line of reactions, EnvZ can bind to ADP and ATP, but only when bound to ATP it can gain a phosphoryl group, and the resulting species is denoted by EnvZ-P. In the second line of reactions, EnvZ-P transfers the phosphoryl group to OmpR, through the formation of an intermediate complex. In the last line of reactions, the phosphoryl group is removed from OmpR-P through the action of EnvZ-D. The concentration of ATP and ADP is assumed to be maintained constant in time.

these modifications, (2.1) becomes

$$\begin{aligned}
A + B &\xrightarrow{\kappa_1} 2B\\
0 &\xrightarrow{u_1(t)} B \xrightarrow{\kappa_2} A\\
0 &\underset{\kappa_3}{\overset{u_2(t)}{\rightleftharpoons}} C\\
C + B &\xrightarrow{\kappa_4} 0
\end{aligned} \qquad (2.3)$$

Since we still have

$$\frac{\mathrm{d}}{\mathrm{d}t}x_A(t) = -\kappa_1 x_B(t)\left(x_A(t) - \frac{\kappa_2}{\kappa_1}\right),$$

it is still true that the value of $x_A(t)$ will converge to κ_2/κ_1, as long as the functions u_1, u_2 and the rates κ_3, κ_4 are such that the concentration of the species B is not driven to 0. In this paper, we will prove a general result describing when such robustness to persistent perturbations is present for ACR systems.

2.2. EnvZ-OmpR osmoregulatory system

Consider the mass-action system described in figure 1. The model is proposed and studied in [14,34] as the osmoregulatory system in *Escherichia coli*. It is in accordance with experimental observations discussed in [6,9,12]. According to the model, whose schematics is described in figure 1, the activation rate of the sensor-transmitter protein EnvZ depends on the medium osmolarity. Then, an active form of EnvZ transfers its phosphoryl group to the sensory response protein OmpR, which becomes OmpR-P and promotes the production of the outer membrane porins OmpF and OmpC. Hence, it is important that the concentration of OmpR-P responds in a reliable, predictive way to changes in the osmolarity of the medium (which the rate constants κ_i in the reaction network of figure 1 depend upon), but not on the initial concentration of the different chemical species involved. As a matter of fact, it is shown in [14] that OmpR-P is an ACR species, as a consequence of the theory developed in the paper.

3. Necessary terminology and known results

In order to present the theory we develop, we first need to introduce some terminology. The linear combinations of chemical species appearing on either side of the chemical reactions of interest are called *complexes*, in accordance with the reaction network theory literature. Be aware that the word 'complex' usually has a different meaning in the biology literature. We denote by m the number of complexes present in the network, and by d the number of chemical species. As an example, the complexes of (2.1) are $A+B$, $2B$, B, and A. Here, $d=2$ and $m=4$. In (2.3), the complexes are $A+B$, $2B$, 0, B, A, C and $C+B$, hence $d=3$ and $m=7$. Finally, in the system depicted in figure 1 $d=8$ and $m=10$. Since a complex is a linear combination of species, each complex can be regarded as a vector of length d. For example, for the model (2.1), we can consider $A+B$ as (1, 1), $2B$ as (0, 2), B as (0, 1), and finally A as (1, 0). With this in mind, we can define the *stoichiometric subspace* as

$$\mathcal{S} = \operatorname{span}_{\mathbb{R}}\{y_j - y_i : \text{there is a reaction from } y_i \text{ to } y_j\},$$

where y_n denotes the nth complex, for all $1 \le n \le m$. For example, for (2.1), we have

$$\mathcal{S} = \operatorname{span}_{\mathbb{R}}\left\{\begin{pmatrix}-1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\right\} = \operatorname{span}_{\mathbb{R}}\left\{\begin{pmatrix}-1\\1\end{pmatrix}\right\}.$$

For (2.3), we have $\mathcal{S} = \mathbb{R}^3$.

In the most general formulation of reaction systems, a (time-dependent) rate function λ_{ij} is associated with the reaction from the ith to the jth complex of the network, and the concentration vector of the different chemical species is assumed to solve the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = \sum_{1\le i,j\le m}(y_j - y_i)\lambda_{ij}(x(t), t), \qquad (3.1)$$

where if a reaction from the ith to the jth complex does not exist, then λ_{ij} is the zero function. Note that (3.1) simply sums the contributions to the dynamics given by the different chemical reactions. Since the derivative in (3.1) is an element of $\mathcal{S}$ at all times, every solution to (3.1) is necessarily confined within a translation of $\mathcal{S}$. If for all non-zero propensities λ_{ij} there exists a positive constant κ_{ij} such that

$$\lambda_{ij}(x(t), t) = \kappa_{ij}\prod_{l=1}^{d} x_l(t)^{y_{il}},$$

then the model is a *mass-action system*. In this case, (3.1) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = Y A(\kappa)\Lambda(x(t)), \qquad (3.2)$$

where Y is a $d \times m$ matrix whose ith column is y_i, $A(\kappa)$ is a $m \times m$ matrix given by

$$A(\kappa)_{ij} = \begin{cases}\kappa_{ji} & \text{if } i \ne j\\ -\sum_{l=1}^{m}\kappa_{il} & \text{if } i = j\end{cases}$$

and $\Lambda(x(t))$ is a vector of length m whose ith entry is $\prod_{l=1}^{d} x_l(t)^{y_{il}}$. Examples of mass-action systems are (2.1) and the model in figure 1.

A directed graph can be associated with a reaction network, where the nodes are given by the complexes and the directed edges are given by the reactions. Such a graph is

called a *reaction graph*. As an example, (2.1) is a reaction graph, while (2.3) is not because the complex 0 is repeated. The reaction graph corresponding to (2.3) is

$$\begin{array}{ccccccc} A+B & \xrightarrow{\kappa_1} & 2B & & & & \\ C+B & \xrightarrow{\kappa_4} & 0 & \xrightarrow{u_1(t)} & B & \xrightarrow{\kappa_2} & A \\ & & \kappa_3 \Big\| u_2(t) & & & & \\ & & C & & & & \end{array} \qquad (3.3)$$

We denote by ℓ the number of connected components of the reaction graph associated with the network. For both networks (2.1) and (2.3) $\ell = 2$, as for the EnvZ-OmpR osmoregulatory system of figure 1 we have $\ell = 3$. Then, we define the *deficiency* of a network as

$$\delta = m - \ell - \dim \mathcal{S}.$$

The deficiency of a network has important geometric interpretation, and a collection of classical deficiency theory results can be found in [35]. The deficiency of (2.1) is $\delta = 4 - 2 - 1 = 1$, and the deficiency of (2.3) is $\delta = 7 - 2 - 3 = 2$. Similarly, it can be checked that the deficiency of the EnvZ-OmpR osmoregulatory system in figure 1 is 1.

Finally, we say that a complex y is *terminal* if for all paths in the reaction graph leading from y to another complex y', there is a path leading from y' to y. If a complex is not terminal, then it is called *non-terminal*. As an example, the only terminal complexes for (2.1) and (2.3) are $2B$ and A.

We recall that a species is said to be ACR if its concentration at any positive steady state of (3.1) is the same. We are ready to state the following result, as presented in [14].

Theorem 3.1. *Consider a mass-action system, and assume the following holds:*

- *there are two non-terminal complexes y_i and y_j such that only one entry of $y_j - y_i$ is non-zero;*
- *the deficiency is 1.*
- *a positive steady state exists.*

Then, the species corresponding to the non-zero entry of $y_j - y_i$ is ACR.

Note that a stronger version of theorem 3.1 is proven in [14], which detects more general steady-state invariants than the steady-state concentration of a single species. The stronger version is stated in the electronic supplementary material as theorem B.3, and an extension of it is proven in the present work.

The model (2.1) has deficiency 1, as already observed, has at least one positive steady state and the non-terminal complexes $A + B$ and B differ only for the species A. Hence, theorem 3.1 applies and A is ACR. It is shown in [14] that the EnvZ-OmpR osmoregulatory system in figure 1 also fulfils the hypothesis of theorem 3.1, with the non-terminal complexes EnvZ-D and EnvZ-D+OmpR-P only differing for the species OmpR-P. As a consequence, OmpR-P is ACR. In §4.2, we will develop a method to explicitly calculate the ACR value through symbolic linear algebra. We note that theorem 3.1 cannot be applied to (2.3) for two reasons: the model is not a mass-action system unless $u_1(t)$ and $u_2(t)$ are constant, and its deficiency is 2.

As noted in [3], the positive steady states of a system with an ACR species are not necessarily stable. However, as a consequence of the present work (more precisely, as a consequence of theorem 5.1 with u being the zero function), we know the following: if a mass-action system as in theorem 3.1 has an unstable positive steady state, then either the system oscillates around it, or some chemical species is completely consumed, or some chemical species is indefinitely produced. We give here the formal definition of 'oscillation', as intended in this paper.

Definition 3.2. We say that a function $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ *oscillates* around a value $q \in \mathbb{R}$ if for each $t \in \mathbb{R}_{\geq 0}$ there exist $t_+ > t$ and $t_- > t$ such that

$$g(t_+) > q \quad \text{and} \quad g(t_-) < q.$$

4. A linear constrained integrator

4.1. Control theory background

In control theory, the focus is usually on systems of differential equations of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t), u(t)), \qquad (4.1)$$

where $x\colon \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_x}$ and $u\colon \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_u}$ for some $n_x, n_u \in \mathbb{Z}_{>0}$, and f is a continuous function, such that (4.1) has a unique global solution for a set of initial conditions and a family of functions u of interest. The function u is called the *input of the system*. Furthermore, a quantity of the form $z(t) = a(x(t))$ is of interest, where a is a differentiable function with $a\colon \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}$, for some $n_z \in \mathbb{Z}_{>0}$. The function z is called the *output* of the system. In the usual setting, one needs to find an appropriate function u such that z is close to a desired level $\bar{z} \in \mathbb{R}^{n_z}$, either on average or for $t \to \infty$. To this aim, the existence of a differentiable function $\phi\colon \mathbb{R}^{n_x} \to \mathbb{R}$ such that for all considered solutions x

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(x(t)) = z(t) - \bar{z},$$

is of high importance, and is called an *integrator*. The name derives from $\phi(x(t))$ being the integral of the error that needs to be controlled

$$\phi(x(t)) = \phi(x(0)) + \int_0^t (z(s) - \bar{z})\,\mathrm{d}s.$$

If the function is fed back to the system and is used to tune the input, then an *integral action* or *integral feedback* is in place [36,37]. One of the main features of an integrator is that the derivative of $\phi(x(t))$ is zero if and only if $z(t) = \bar{z}$. If a function $\tilde{\phi}\colon \mathbb{R}^{n_x} \to \mathbb{R}$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\phi}(x(t)) = r(x(t))\Big(z(t) - \bar{z}\Big), \qquad (4.2)$$

for some differentiable function $r\colon \mathbb{R}^{n_x} \to \mathbb{R}$, then $\tilde{\phi}$ is called a *constrained integrator* (CI) [13]. The name derives from the fact that the derivative of $\tilde{\phi}(x(t))$ is zero if and only if $z(t) = \bar{z}$, provided that $r(x(t)) \neq 0$. In biology, it is common to find CIs, and the condition $r(x(t)) \neq 0$ is usually implied by $x(t) \neq 0$ [13]. Note that in [13] an explicit distinction between integrators and integral feedbacks is not made.

In the setting of systems with ACR species, the output z can be considered to be (a power of) the concentration of the ACR species over time, and $\bar{z}$ can be their ACR values. In (2.1), $z(t) = x_A(t)$ and $\bar{z} = \kappa_2/\kappa_1$. A CI (as noted in [13]) is

given by $\tilde{\phi}(x(t)) = x_B(t)$, since

$$\frac{\mathrm{d}}{\mathrm{d}t}x_B(t) = \kappa_1 x_B(t)\left(x_A(t) - \frac{\kappa_2}{\kappa_1}\right).$$

The question of whether an integrator exists can be quickly answered in negative, because any point of the form $(\bar{x}_A, 0)$ is a steady state. If an integrator ϕ existed, then by choosing $x(0) = (\bar{x}_A, 0)$ we would have

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\phi(x(t)) = \bar{x}_A - \frac{\kappa_2}{\kappa_1},$$

which cannot hold except for a specific value of $\bar{x}_A$. An integrator may still exist in a weaker sense, if we restrict its domain. For example, in this case, the function $\hat{\phi}(x(t)) = (1/\kappa_1)\log x_B(t)$ would be an integrator, in the sense that if $x_B(t) > 0$ then

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\phi}(x(t)) = \frac{1}{\kappa_1 x_B(t)}\left(\kappa_1 x_B(t)\left(x_A(t) - \frac{\kappa_2}{\kappa_1}\right)\right) = x_A(t) - \frac{\kappa_2}{\kappa_1}.$$

However, the domain of $\hat{\phi}$ is not the entire $\mathbb{R}^2$. Finally, since linear functions could always be extended continuously to the boundaries of $\mathbb{R}^2_{>0}$, a linear integrator cannot exist for (2.1) even if its domain is restricted.

4.2. Existence and characterization

We state here our result concerning linear CIs. A stronger version is proved in the electronic supplementary material. The result is inspired by the analysis carried on in [14], which is here expanded along lines similar to those of [21].

Our goal is to find a linear function $\tilde{\phi}\colon \mathbb{R}^d \to \mathbb{R}$ of the form $\tilde{\phi}(x) = \langle \hat{\gamma}, x\rangle$ satisfying (4.2) with $z(t)$ being some power of the ACR species, where $\langle \cdot, \cdot \rangle$ is the standard scalar product. For any $n \times l$ real matrix M and real vector v of length n, we denote by $(M \,|\, v)$ the $n \times l+1$ matrix obtained by adding the column v at the right of the matrix M. We further denote by e_k the kth vector in the canonical basis of $\mathbb{R}^m$, whose kth component is 1 and whose other components are 0. If the complexes y_i and y_j only differ in the nth component, and $\gamma \in \mathbb{R}^{d+1}$ satisfies

$$\left(A(\kappa)^\top Y^\top \,|\, e_i\right)\gamma = e_j,$$

then if we denote by $\hat{\gamma}$ the projection onto the first d components of γ and if the system is mass-action, it follows from (3.2) that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{\gamma}, x(t)\rangle &= \hat{\gamma}^\top Y A(\kappa)\Lambda(x(t)) = \Lambda_j(x(t)) - \gamma_{d+1}\Lambda_i(x(t)) \\ &= x(t)^{y_i}\left(x_n(t)^{(y_j - y_i)_n} - \gamma_{d+1}\right).\end{aligned}$$

Hence, at any positive steady state $x_n = \gamma_{d+1}^{1/(y_j - y_i)_n}$, implying that the nth species is ACR. Moreover, the function $\tilde{\phi}(x) = \langle \hat{\gamma}, x\rangle$ is the CI we were looking for.

Inspired by the above analysis, we focus on the study of the set

$$\Gamma_{ij}(\kappa) = \left\{\gamma \in \mathbb{R}^{d+1} : \left(A(\kappa)^\top Y^\top \,|\, e_i\right)\gamma = e_j\right\}, \tag{4.3}$$

for $1 \le i, j \le m$. The projection of $\Gamma_{ij}(\kappa)$ onto the first d coordinates, denoted by $\hat{\Gamma}_{ij}(\kappa)$, will also be of interest as its elements are linear CIs. We state this in the following result, a more general version of which is proven in the electronic supplementary material

Theorem 4.1. *Consider a mass-action system. Assume that there are two complexes y_i and y_j only differing in the nth entry, and that $\Gamma_{ij}(\kappa)$ is non-empty. Let $\gamma \in \Gamma_{ij}(\kappa)$, and define*

$$q = \gamma_{d+1}^{1/(y_j - y_i)_n}.$$

Then, either no positive steady state exists or the nth species is ACR with ACR value q. Moreover,

$$\phi(x) = \sum_{i=1}^{d} \beta_i x_i,$$

is a linear CI with

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(x(t)) = \Lambda_i(x(t))(x_n(t)^{(y_j - y_i)_n} - q^{(y_j - y_i)_n}),$$

for any initial condition $x(0)$ if and only if $\beta \in \hat{\Gamma}_{ij}(\kappa)$.

Note that if $\gamma \in \Gamma_{ij}(\kappa)$, then the vector $v = e_j - \gamma_{d+1}e_i$ is in the rowspan of $Y\,A(\kappa)$, because $A(\kappa)^\top Y^\top \hat{\gamma} = v$. As such, v is an example of 'complex linear invariant on the complexes y_i and y_j', as introduced in [21]. We follow a slightly different approach than in [21], since we do not only care about the steady state invariants but also aim to calculate the CIs given by the set $\hat{\Gamma}_{ij}(\kappa)$, which can be considered as the preimage of v via $Y\,A(\kappa)$. However, methods discussed in [21] can still be fruitfully used to quickly decide whether $\Gamma_{ij}(\kappa)$ is empty, and to efficiently calculate the ACR value encoded in γ_{d+1}. As an example we state the following result, which is an immediate consequence of [21, proposition 1]:

Theorem 4.2. *Let N be the matrix obtained by removing the ith and the jth column from $Y\,A(\kappa)$. Then, $\Gamma_{ij}(\kappa)$ is non-empty if and only if* rank $YA(\kappa) >$ rank N.

The set $\Gamma_{ij}(\kappa)$ can be calculated with symbolic linear algebra. Moreover, if $\xi \in \hat{\Gamma}_{ij}(\kappa)$ then necessarily

$$\{\xi + w : w \in \mathcal{S}^\perp\} \subseteq \hat{\Gamma}_{ij}(\kappa). \tag{4.4}$$

Hence, $\hat{\Gamma}_{ij}(\kappa)$ is connected with $\mathcal{S}^\perp$, which is a set easily described by linear algebra and independent on the rate functions. The reason for (4.4) is perhaps clearer if $\hat{\Gamma}_{ij}(\kappa)$ is regarded to as a preimage via $YA(\kappa)$, as discussed above, since the columns of $YA(\kappa)$ are in $\mathcal{S}$. A formal proof of (4.4) is given in the electronic supplementary material, where we will also give sufficient conditions under which the inclusion is an equality.

As a concrete example, consider the model in figure 1. Using Matlab, we quickly obtain that a vector ξ is in $\Gamma_{18}(\kappa)$, with

$$\xi_9 = \frac{\kappa_1\kappa_3\kappa_5(\kappa_{10} + \kappa_{11})[\mathrm{ATP}]}{\kappa_2(\kappa_4 + \kappa_5)\kappa_9\kappa_{11}[\mathrm{ADP}]}, \tag{4.5}$$

and it is shown in the electronic supplementary material that

$$\hat{\Gamma}_{18}(\kappa) = \left\{\hat{\xi} + \begin{pmatrix} w_1 \\ w_1 \\ w_1 \\ w_1 \\ 0 \\ w_1 \\ 0 \\ w_1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ w_2 \\ w_2 \\ w_2 \\ w_2 \end{pmatrix} : w_1, w_2 \in \mathbb{R}\right\}, \tag{4.6}$$

where $\hat{\xi}$ is the projection of ξ onto its first $d = 8$ coordinates.

The family of models we study in this paper concerns reaction systems with two complexes y_i and y_j differing in just one entry, for which $\hat{\Gamma}_{ij}(\kappa)$ is non-empty. The following result shows that such a family includes the models studied in [14]. The result has been already discussed in [21, section 2.4] with a different formulation, and its proof is basically already present

in the proof of the main result of [14]. However, for the sake of clarity, we propose a proof with the notations of the present paper in the electronic supplementary material.

Theorem 4.3. *Consider a mass-action system, and assume the following holds*:

- *there are two non-terminal complexes y_i and y_j such that only one entry of $y_j - y_i$ is non-zero;*
- *the deficiency is 1.*
- *a positive steady state exists.*

Then, $\Gamma_{ij}(\kappa)$ is non-empty.

As an example of application, we know already from direct calculation that for the EnvZ-OmpR signalling system the set $\Gamma_{18}(\kappa)$ is non-empty. However, we could have also derived this information from theorem 4.3, without explicitly calculating it.

We note here that the converse of theorem 4.3 does not hold. We show this with an example of a multisite phosphorylation signalling system in the electronic supplementary material (figure 6), which does not fall in the setting of [14] but for which we are able to prove absolute concentration robustness regardless of the choice of rate constants, as long as a positive steady state exists. Notably, we are also able to derive information on when this occurs without working directly with the differential equation. As a consequence of this example, the family of models we analyse is proven to be strictly larger than that studied in [14].

A first interesting consequence of theorem 4.1 is that the ACR value of the mass-action systems satisfying the assumption of the theorem can be calculated by finding at least one element of $\Gamma_{ij}(\kappa)$, and this can be done via a simple symbolic linear algebra calculation. As an example, consider the EnvZ-OmpR osmoregulatory system in figure 1. Then, theorem 4.1 implies that the ACR value of OmpR-P is the value given in (4.5). This value is in accordance with the one found in the electronic supplementary material of [14], however, we found it by calculating a single element in the preimage of a matrix, as opposed to working with the rather complicated differential equation associated with the model. An even more involved example is dealt with in the electronic supplementary material.

The existence of a linear CI given by theorem 4.1 is essential to develop the results presented in the next sections. Before unveiling the consequences of theorem 4.1 in terms of disturbance rejection, however, it is important to stress that a CI does not necessarily constitute a feedback, as one may be tempted to think. Consider

$$\begin{aligned} A + B &\xrightarrow{\kappa_1} A \underset{\kappa_3}{\overset{\kappa_2}{\rightleftharpoons}} 0 \\ B &\xrightarrow{\kappa_4} 2B \end{aligned}$$

with $\kappa_1\kappa_3 = \kappa_2\kappa_4$. It can be shown that the system satisfies the conditions of theorems 3.1 and 4.3, with the non-terminal complexes $A + B$ and A differing only in species A. Hence, A is ACR and the assumptions of theorem 4.1 hold. A linear CI as in theorem 4.1 is given by $\phi(x) = -x_B/\kappa_1$, since for this choice

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(x(t)) = x_B(t)\left(x_A(t) - \frac{\kappa_4}{\kappa_1}\right).$$

However, the quantity $\phi(x(t))$ does not regulate the dynamics of A, since

$$\frac{\mathrm{d}}{\mathrm{d}t}x_A(t) = \kappa_2 - \kappa_3 x_A(t)$$

does not depend on $x_B(t)$. Since in this case the CI is not acting on the system, it is not surprising that the existence of positive steady states is lost as soon as $\kappa_1\kappa_3 \neq \kappa_2\kappa_4$.

It is also worth mentioning that not all systems with ACR species have a linear CI: consider the mass-action system

$$\begin{aligned} A + B &\xrightarrow{\kappa_1} B + C \underset{\kappa_3}{\overset{\kappa_2}{\rightleftharpoons}} 2B \\ B &\underset{\kappa_5}{\overset{\kappa_4}{\rightleftharpoons}} 2E \xrightarrow{\kappa_6} 2D \\ C &\xrightarrow{\kappa_7} A \\ D &\xrightarrow{\kappa_8} E \end{aligned}$$

The model is considered in [38], where it is proven that the species A is ACR. We show in the electronic supplementary material that there exists no linear function ϕ whose derivative at time t is of the form $r(x(t))(x_A(t)^\gamma - q)$, for some polynomial r and some real numbers γ, q. Note that in this case theorem 4.3 does not apply because the deficiency of the network is 2.

5. Rejection of persistent disturbances

5.1. The result

We state an important consequence of theorem 4.1, a stronger version of which is proven in the electronic supplementary material:

Theorem 5.1. *Consider a mass-action system with d species, with associated differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t)).$$

Assume that there are two complexes y_i and y_j only differing in the nth entry, and that $\hat{\Gamma}_{ij}(\kappa)$ is non-empty. Let q be the ACR value of the nth species. Consider an arbitrary function u with image in $\mathbb{R}^d$ such that a solution to

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{x}(t) = f(\tilde{x}(t)) + u(\tilde{x}(t), t),$$

exists. Assume that there exists a $\hat{\gamma} \in \hat{\Gamma}_{ij}(\kappa)$ which is orthogonal to the vector $u(x, t)$ for any x, t. Then, for any initial condition $\tilde{x}(0)$, at least one of the following holds:

- *the concentration of some species goes to 0 or infinity, along some sequence of time points;*
- *$\tilde{x}_n(t)$ oscillates around q and $\hat{\gamma}_k \neq 0$ for some $k \neq n$;*
- *the integral*

$$\int_t^\infty |\tilde{x}_n(s) - q|\,\mathrm{d}s,$$

tends to 0, as *t goes to infinity.*

The result implies that if a disturbance orthogonal to a vector $\hat{\gamma} \in \hat{\Gamma}_{ij}(\kappa)$ is applied over time, then, at most, the ACR species can be forced to oscillate around its original ACR value, but it cannot be forced to attain another steady-state level without causing extinction or overexpression of the chemical species present. We analyse the power of theorem 5.1 by showing some examples of applications.

Example 5.2. Consider the mass-action system (2.1), which fulfills the assumptions of theorem 5.1 as already observed. Assume the complexes are ordered as $A+B$, $2B$, B and A, and the species are ordered alphabetically as A, B. Hence, the two non-terminal complexes differing in the ACR species A are the 1st and the 3rd, and it is shown in the electronic supplementary material that

$$\hat{\Gamma}_{13}(\kappa) = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} w \\ w \end{pmatrix} : w \in \mathbb{R} \right\}. \tag{5.1}$$

Hence, by choosing $w = -1$, we have that

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} \in \hat{\Gamma}_{13}(\kappa). \tag{5.2}$$

This vector is clearly orthogonal to any disturbance acting on the production and degradation rates of the species B. Hence, it follows that the stability and the ACR value of the species A is maintained in (2.3), provided that no species is completely removed or indefinitely expressed. Specifically, since the entry of (5.2) corresponding to B is zero, it follows from theorem 5.1 that if all the species concentrations are bounded from below and from above by positive quantities, necessarily the concentration of the species A converges to its ACR value as t goes to infinity, despite the disturbances.

Example 5.3. (EnvZ-OmpR osmoregulatory system). Consider the osmoregulatory system in figure 1, whose features have already been discussed in the paper. In particular, we know the species OmpR-P is ACR with ACR value (4.5). Recall that we ordered the complexes such that the two non-terminal ones differing in OmpR-P are the 1st and the 8th. It follows from (4.6) that for any chemical species, there is a vector in $\hat{\Gamma}_{18}(\kappa)$ with the corresponding entry equal to 0. It follows that even if the production and degradation of any chemical species in the model is tampered with, the stability and the ACR value of the species OmpR-P are maintained, in the sense described by theorem 5.1.

We can push the disturbances further. By appropriately choosing w_1 and w_2 in (4.6), we can see that there is a vector in $\hat{\Gamma}_{18}(\kappa)$ whose entries corresponding to the species OmpR and EnvZ are both 0. Hence, it follows by theorem 5.1 that by tampering with the production and degradation rates of both these species over time, if no extinction and no overexpression occurs, then the concentration of OmpR-P still converges to the value (4.5), or oscillates around it.

As a final remark, we note that (4.4) can be useful in determining whether a vector in $\Gamma_{ij}(\kappa)$ exists, with a specific component equal to 0, say the nth one. In fact, the existence of such a vector can be deduced without calculations, if there is a vector $w \in \mathcal{S}^{\perp}$ whose nth component is different from 0.

5.2. Insulating properties

Here, we discuss how the theory we developed can be used to design ACR modules that serve as effective insulators in mitigating the problem of retroactivity in synthetic biology. As explained in the Introduction, the loading effects caused by a downstream biochemical module can disrupt the functionality of upstream modules (see figure 4 for a concrete example), which prevents the implementation of biochemical circuits by interconnecting biochemical modules with different functions [17]. Our results indicate that certain ACR systems are remarkably tolerant, and the ACR species maintains its ACR value even in the presence of arbitrary time-varying disturbances. We exploit this property to show that insulator ACR modules can be designed to provide inputs to the downstream modules while robustly mirroring the key functional property of upstream modules. In other words, the loading effects generated by downstream modules are *rejected* by the ACR insulator, facilitating modularity. We now explore this idea in greater detail and present an illustrative example.

Assume that a mass-action system has two complexes y_i and y_j, that are only different in the nth component, which corresponds to the species X. Assume further that $\hat{\Gamma}_{ij}(\kappa)$ is non-empty and that a positive steady state exists. Hence, the species X is ACR, with some ACR value q. It further follows from theorem 5.1 that, if there exists $\hat{\gamma} \in \hat{\Gamma}_{ij}(\kappa)$ with $\hat{\gamma}_n = 0$, then the production and degradation rates of the species X can be arbitrarily perturbed over time by an arbitrary function u, without compromising its robustness. Specifically, if the perturbed system is stable and no chemical species is completely consumed, then the concentration of X will still converge to the same ACR value q as in the original mass-action system. We can consider the perturbation u as the loading effect of a downstream module whose input is the concentration of species X. In this case, the loading effect on the original system is rejected and the concentration of X is maintained at a desired level q at steady state. Furthermore, the concentration of X is maintained approximately constant in the transient dynamics as well, if we assume as done in [17] that a separation of dynamics timescale is in place. Specifically, assume

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{x}(t) = \frac{1}{\varepsilon} f(\tilde{x}(t)) + u(\tilde{x}(t), t)e_n,$$

for some small $\varepsilon > 0$, with $f(\tilde{x}(t))$ and $u(\tilde{x}(t), t)$ being of the same order of magnitude. Under the assumption of stability, if ε is very small then the perturbed system will quickly approach the slow manifold defined by

$$0 = f(\tilde{x}(t)),$$

hence the concentration of X is constantly equal to its ACR value. Finally and perhaps more importantly, if part of the disturbance acts on the same timescale as the system, that is if

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{x}(t) = \frac{1}{\varepsilon}(f(\tilde{x}(t)) + u(\tilde{x}(t), t)e_n) + u'(\tilde{x}(t), t)e_n,$$

with $f(\tilde{x}(t))$, $u(\tilde{x}(t), t)$ and $u(\tilde{x}(t), t)$ being of the same order of magnitude, then the slow manifold

$$0 = f(\tilde{x}(t)) + u_t(\tilde{x}(t))e_n,$$

is quickly approached, where u_t is a function from $\mathbb{R}^d$ to $\mathbb{R}^d$ defined by $u_t(x) = u(x, t)$. By theorem 5.1 applied to the disturbance u_t, the species X assumes its ACR value at any positive point of the slow manifold, which is exactly what we wanted.

As an illustrative example consider the EnvZ-OmpR osmoregulatory system in figure 1. It follows from (4.6) that there exists $\hat{\gamma} \in \Gamma_{18}(\kappa)$ with a zero in the entry corresponding to the ACR species OmpR-P. Hence, the production and degradation rates of OmpR-P can be arbitrarily changed over time, without altering its robustness property, in the sense described by theorem 5.1. As observed, the statement still holds true if the perturbation is originated by a downstream module that acts on OmpR-P. Hence, the EnvZ-OmpR osmoregulatory system can be used to maintain the expression of OmpR-P at a desired

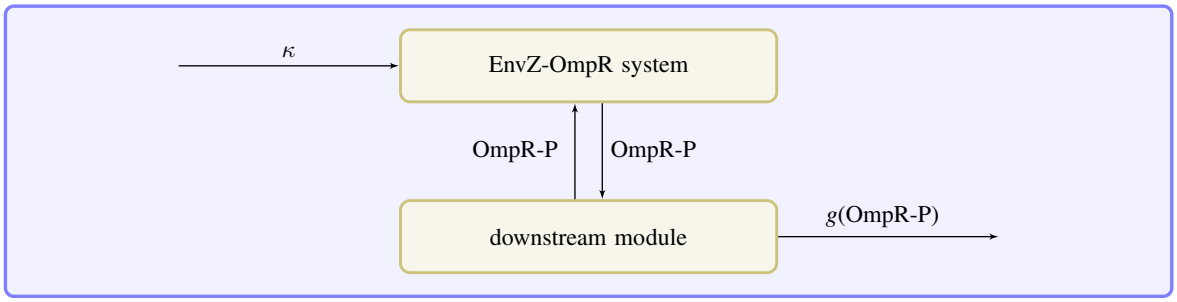

Figure 2. Proposed use of the EnvZ-OmpR signal transduction system of figure 1 as a controller of a downstream module using OmpR-P. The concentration of OmpR-P is regulated by the EnvZ-OmpR signalling system, with steady state given by (4.5). The steady state can be adjusted by modifying the parameters κ_{ij} of the EnvZ-OmpR signalling system (which depend on the medium osmolarity) or the ratio between ADP and ATP present. The output of the downstream module is a function g of the concentration of OmpR-P, which is received as an input.

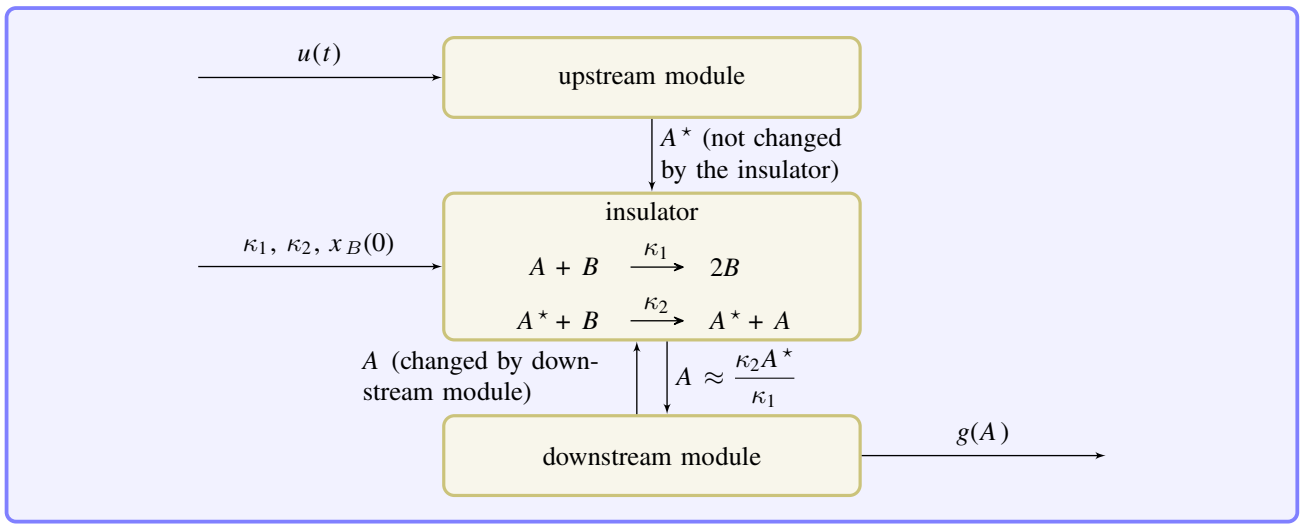

Figure 3. The upstream module expresses the chemical species $A^\star$ as output. The insulator transfers a multiple of the signal from the upstream module to the downstream module, which is modified to accept as an input the concentration of A rather than the concentration of $A^\star$.

level, which depends on the input rate constants, even if the species OmpR-P is used by a downstream module. Moreover, if the osmoregulatory system acts on a fast enough timescale, the concentration of the species OmpR-P is approximately maintained at the target level at any time point. Note that this analysis holds true provided that the disturbed system converges to a positive steady state, which should be checked separately. In figure 2, a diagram describing this situation is proposed.

Consider now the case were an upstream module is affected by loading effects. We show how the theory developed in this paper can be used to design an insulator. Assume that the upstream module accepts $u(t)$ as an input, and modulates the concentration of the species $A^\star$ accordingly. The species $A^\star$ is then used by a downstream module, which returns a function of the concentration of $A^\star$ as output. The action of the downstream module on the species $A^\star$ causes a loading effect on the upstream module. To reduce the loading effects, we propose to modify the downstream module such that it acts on a species A rather than on the species $A^\star$, and to include in the system the following module, where B is a species that is not used by neither the upstream nor the downstream module:

$$A + B \xrightarrow{\kappa_1} 2B$$
$$A^* + B \xrightarrow{\kappa_1} A^* + A.$$

Assume the system is stable and that the species B is not completely consumed. Then, at steady state the concentration level of $A^\star$ is fixed, and the concentration of the ACR species A will converge to its ACR value $\kappa_2 x_{A^\star}/\kappa_1$ regardless of any disturbance applied to the production and degradation rate of A. In fact, a linear CI as in the statement of theorem 4.1 is given by $\phi(x) = x_B/\kappa_1$, and at any time point

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(x(t)) = x_B(t)\left(x_A(t) - \frac{\kappa_2}{\kappa_1}x_{A^\star}(t)\right). \tag{5.4}$$

We further note that if the dynamics of (5.3) occurs on a faster timescale than the upstream module (not necessarily of the downstream module), then a slow manifold is quickly approached where the concentration of the species A is maintained at the level $\kappa_2 x_{A^\star}(t)/\kappa_1$ at any time point. In this case, the module (5.3) approximately outputs a multiple of the concentration of $A^\star$ over the whole time line. The multiplicative constant can be tuned through the parameters κ_1 and κ_2, as well as the timescale that (5.3) operates in. The timescale can be further tuned via the concentration of B, as it also follows from (5.4). In conclusion, the downstream module receives as an input a good approximation of a multiple of the concentration of $A^\star$, and its activity does not affect the upstream module, nor (5.3). Moreover, (5.3) does not affect the upstream module at all, since the species A appears in (5.3) as a catalyst and is not changed in the catalysed reaction. The proposed insulating strategy is illustrated in figure 3, and it is applied to an example discussed in [17] in figure 4.

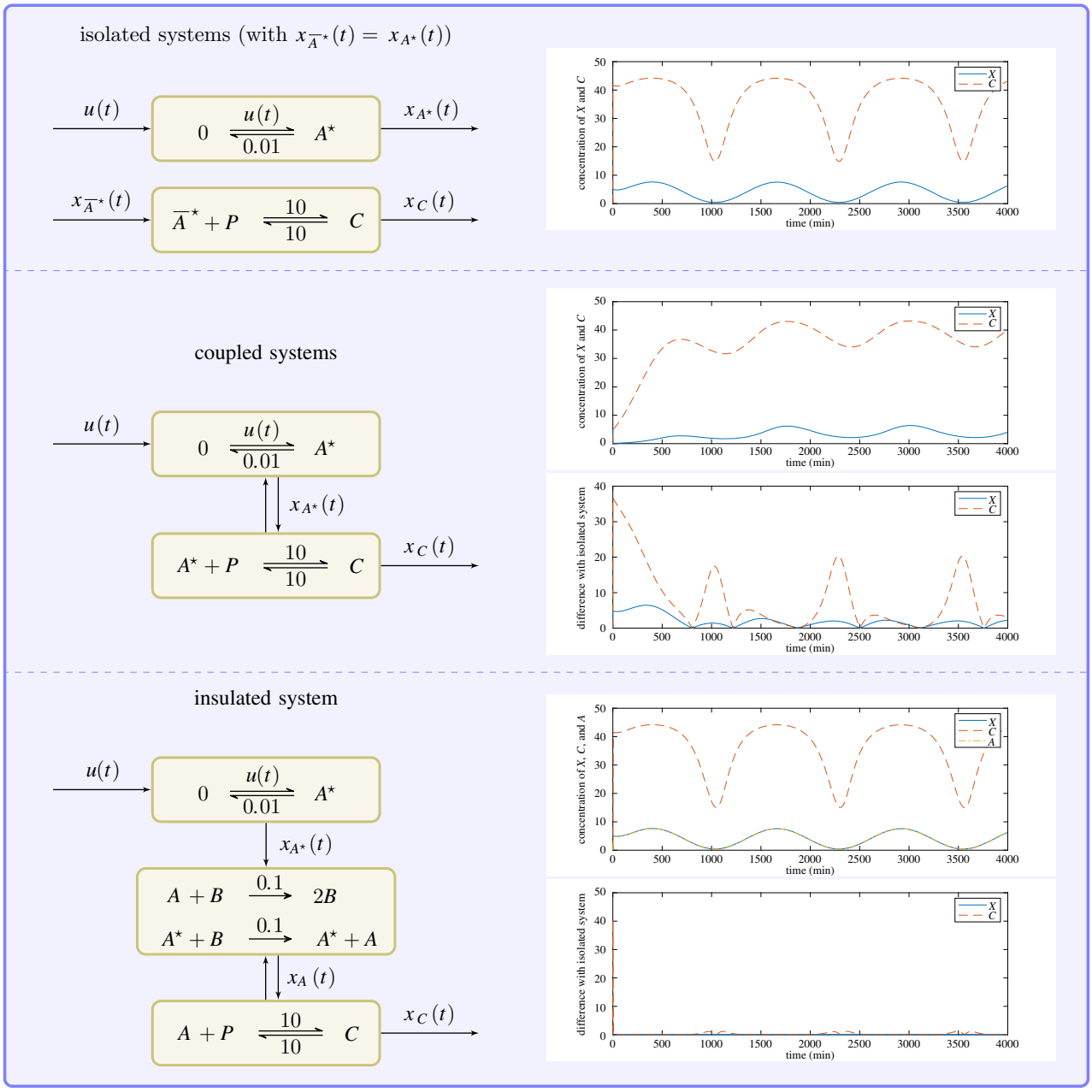

Figure 4. In the example, we let $u(t) = 0.04(1 + \sin(0.005t))$ and $x_{A^\star}(0) = 5$. All the rates are in 1 min. The ODE solution is calculated in Matlab with `ode23s`. In the first panel, the two systems are considered in isolation, with the assumption that the concentration of the species $\overline{A}^\star$ is maintained at the same level as the concentration of species $A^\star$, at any time point. In the second panel, the two systems are linked together, and the species $A^\star$ is directly used by the downstream system. The dynamics are completely disrupted by the loading effects, a plot of the absolute value of the difference of the two solutions over time is proposed. In the third panel, the insulator of figure 3 is used, with $x_A(0) = 0$ and $x_B(0) = 20$. The loading effects are practically removed, despite the choice of low rate constants 0.1. Notably, with this choice the insulator acts on a faster timescale than the upstream module, but on the same timescale as the downstream module. The difference of the concentration of C between the solution of insulated system and the solution of the isolated systems spikes quickly to 40, but it decreases to less than 0.5 within 10 min, after which is maintained low as illustrated in the second plot of the third panel.

5.3. Inclusion in larger systems

In the previous section, we have seen how the absolute concentration robustness and the related stability properties can be transferred to a larger model that is not necessarily mass-action. The larger system may include further chemical transformations and external inputs, whose dynamics may be partly unknown. If seen under a different perspective, this result allows us to further extend the set of sufficient conditions of theorem 3.1 that imply the existence of an ACR species, in the sense of theorem 5.4 below. Before stating the precise result, which is a consequence of theorem 4.1, we need a definition: Given a reaction system $\mathscr{S}$, we say that $\mathscr{S}'$ is a subsystem of $\mathscr{S}$ if it can be obtained from $\mathscr{S}$ by cancelling some reactions, and if the choice of rate functions for the remaining reactions is maintained. Moreover, if $\mathscr{S}$ and $\mathscr{S}'$ have d and d' species, respectively, we let $\pi\colon \mathbb{R}^d \to \mathbb{R}^{d'}$ be the projection onto the species of $\mathscr{S}'$. The following holds:

Corollary 5.4. *Consider a reaction system $\mathscr{S}$, and let $\mathscr{S}'$ be a subsystem. Assume that $\mathscr{S}'$ is a mass-action system with two complexes $\pi(y_i)$ and $\pi(y_j)$ only differing in the entry corresponding to the species X, and for which $\hat{\Gamma}_{ij}(\kappa)$ is non-empty. Moreover, assume there exists $\hat{\gamma} \in \hat{\Gamma}_{ij}(\kappa)$ such that $\pi(y_l - y_k)$ is orthogonal to $\hat{\gamma}$ for all $y_k \to y_l$ that are reactions of $\mathscr{S}$ but not reactions of $\mathscr{S}'$. Then, the species X is ACR for both $\mathscr{S}$ and $\mathscr{S}'$, with the same ACR value.*

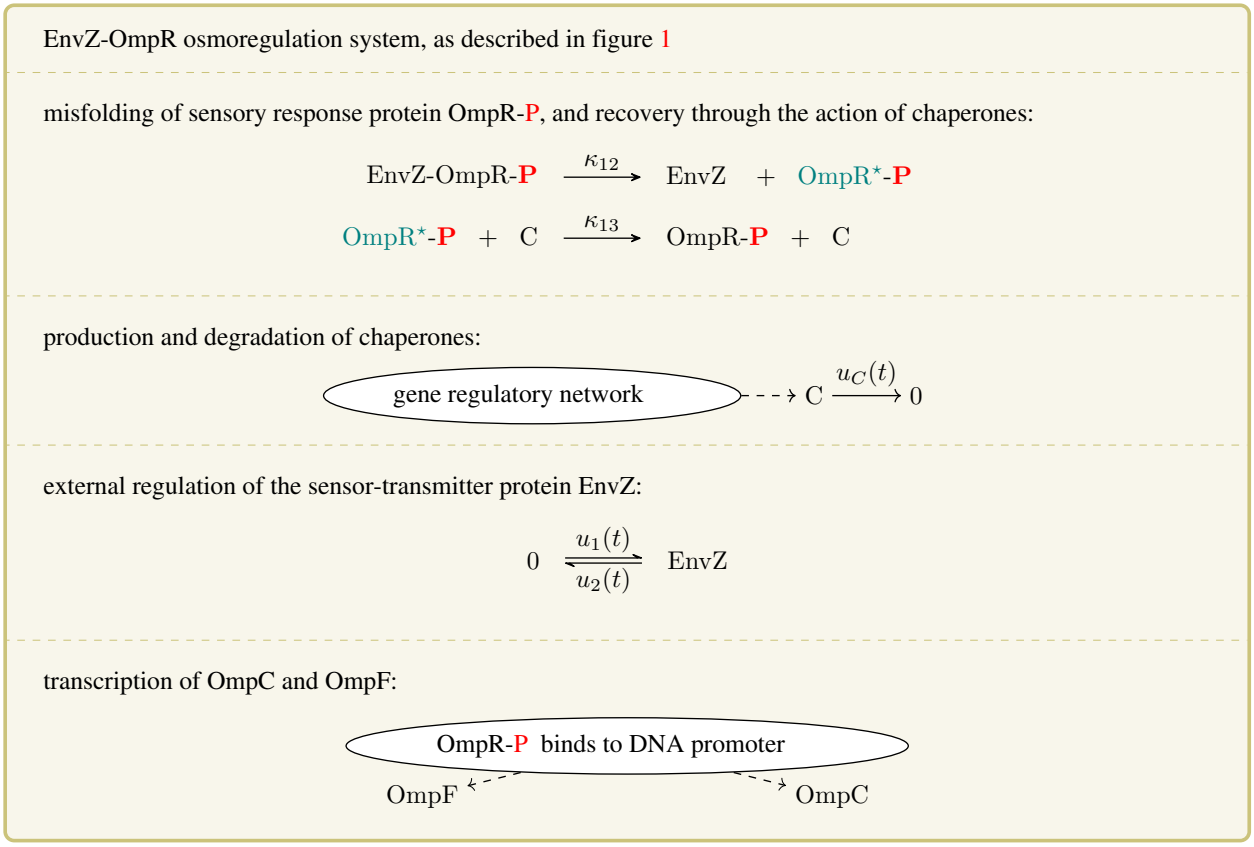

Figure 5. Reaction system including the EnvZ-OmpR osmoregulation system depicted in figure 1 as a subsystem. Parts of the system depicted here are unknown, specifically, no model is given for the production of chaperones or for the transcription of the outer membrane porins OmpF and OmpC.

The proof of a stronger result is in the electronic supplementary material. Here, we illustrate how the corollary can be applied in the case of EnvZ-OmpR signalling system, where we extend the model with reactions whose kinetics is not completely known, and in particular is not required to be of mass-action type. Consider the reaction system $\mathcal{S}$ described in figure 5, which includes the EnvZ-OmpR osmoregulation system described in figure 1 as a subsystem. We assume that a protein can misfold when the phosphoryl group is transferred from EnvZ to OmpR. Such a misfold can be corrected by chaperones, which are proteins assisting the conformational folding of other proteins. We can realistically assume chaperones are independently produced and degraded through an unknown mechanism that does not involve EnvZ or OmpR proteins. We also allow for an arbitrary and persistent external control on the expression level of EnvZ sensor-transmitter protein. Finally, we consider the utilization of OmpR-P as transcription regulatory protein of the outer membrane porins OmpF and OmpC. For our purposes, we assume the details of the transcription mechanism are not known, but that only the protein Ompr-P is involved in the process. As previously done, let the complexes of the EnvZ-OmpR osmoregulation system be ordered from left to right and from top to bottom, such that EnvZ-D and EnvZ-D+OmpR-P are the 1st and the 8th complex, respectively. Also, let the species be ordered according to their appearance from left to right and from top to bottom, as EnvZ-D, EnvZ, EnvZ-T, EnvZ-P, OmpR, EnvZ-OmpR-P, OmpR-P, EnvZ-OmpR-D-P. In particular, EnvZ is the second species, EnvZ-OmpR-P is the sixth species, and OmprR-P is the seventh species. It follows from (4.6) that, by choosing $w_1 = -\xi_2$ and $w_2 = -\xi_7 = 0$, a vector $\hat{\zeta}$ is in $\hat{\Gamma}_{18}(\kappa)$ with:

1. $\hat{\zeta}_2 = \hat{\xi}_2 + w_1 = 0$;
2. $\hat{\zeta}_6 - \hat{\zeta}_2 = \hat{\xi}_6 + w_1 + w_2 - \hat{\xi}_2 - w_1 = \hat{\xi}_6 - \hat{\xi}_2 = 0$;
3. $\hat{\zeta}_7 = \hat{\xi}_7 + w_2 = 0$.

Denote by E_k the vector of $\mathbb{R}^d$ with the kth entry equal to 1 and the other entries equal to zero. The following holds.

Misfolding of OmpR-P. The projection of the difference between EnvZ + OmpR*-P and EnvZ-OmpR-P onto the species of the EnvZ-OmpR signalling system is $E_2 - E_6$, which is orthogonal to $\hat{\zeta}$ by 2. The projection of the difference between OmpR-P + C and OmpR*-P + C is E_7, which is orthogonal to $\hat{\zeta}$ by 3.

Production and degradation of chaperones. By assumption, any chemical reaction $y \to y'$ involved in the production and degradation of chaperones does not consume or produce any chemical species of the EnvZ-OmpR signalling system. Hence, $\pi(y' - y) = 0$, which is orthogonal to $\hat{\zeta}$.

External regulation of EnvZ. The difference between EnvZ and 0 is $\pm E_2$, which is orthogonal to $\hat{\zeta}$ by 1.

Transcription of OmpC and OmpF. We assume that the transcription only involves the species OmpR-P, out of all the species in the EnvZ-OmpR osmoregulation system. Hence, for all the reactions $y \to y'$ involved in the transcription, either $\pi(y' - y) = 0$ or $\pi(y' - y)$ is a multiple of E_7. In either case, $\pi(y' - y)$ is orthogonal to $\hat{\zeta}$.

It follows from corollary 5.4 that the species OmpR-P is ACR in the reaction system of figure 5. Moreover, its ACR value is still given by (4.5), as long as a positive steady

state exists. Note that corollary 5.4 could be applied even if not all chemical reactions are known, and even if the model is not mass-action. It is also worth noting that the deficiency of the model is not known, due to the lack of information on the precise reactions constituting the network, but is certainly greater than 1. Indeed, the deficiency of the subsystem constituted by the EnvZ-OmpR osmoregulation system and by the misfolding of OmpR-P is 2, and the deficiency of a system is necessarily greater than or equal to the deficiency of any subsystem [39, lemma 5].

6. Discussion

Several biological systems of interest exhibit an ACR species, meaning that the steady-state value of the species is invariant to the initial condition, under the assumption of stability. In this paper, we perform control theoretic analysis of biochemical systems that have an ACR species (ACR systems). Our main contribution is to prove the existence of a linear *constrained integrator* (CI) for a family of ACR systems, which strictly includes the models studied in [14] (see theorem 4.1). As we elaborate in the paper, this technical result and its generalization (see electronic supplementary material) have three important biological consequences. Firstly, they provide an easy algebraic way to calculate the steady-state value of ACR species, which is a cumbersome task for complex networks. This method adds to those developed in [21]. Secondly, we show that in the studied systems, ACR species are still ACR even if arbitrary time-varying disturbances are applied (see theorem 5.1). This property can be naturally exploited to design insulators to attenuate loading effects in synthetic biology applications. Lastly, under certain conditions, a large system can inherit an ACR species from an ACR subsystem, and maintain its steady-state value. (see corollary 5.4).

Our results reveal previously unknown facts about ACR systems and they open the door to many interesting problems for future research. Efficient algorithms can be designed in order to check for the existence of portions of the systems that confer absolute concentration robustness to the whole system. To this aim, the connections of $\hat{\Gamma}_{ij}(\kappa)$ with structural properties of the network which we show in the electronic supplementary material can be useful, and theoretical results can be expanded in this direction. As the ACR property depends crucially on stability, further detailed analysis on when stability can be ensured would be welcome. Currently, in the statement of theorem 5.1, we cannot exclude the possibility that some species is completely consumed or indefinitely produced upon tampering with the model, or that oscillations around the ACR value occur. Finding structural conditions able to eliminate this possibility would be a nice and useful contribution.

As a final remark, we think the study of stochastically modelled systems that satisfy the assumptions of theorem 4.1 would be interesting and fruitful. Stochastic models of reaction systems are typically used when few molecules of certain chemical species are available [40,41]. It is proven in [3] that systems satisfying the assumptions of theorem 3.1, when stochastically modelled, undergo an extinction event almost surely. As a consequence, the desirable robustness properties of the ACR systems studied in [14] are completely destroyed in a low molecule copy-number regime. As an example, the model depicted in (2.1) undergoes an almost sure extinction of the chemical species B when stochastically modelled, regardless of the initial conditions. This is caused by the fact that all the molecules of B can be consumed by the reaction $B \to A$, before the occurrence of a reaction $A + B \to 2B$. Robustness at finite time intervals of some stochastically modelled ACR systems is recovered, but only in a multiscale limit sense [42]. Moreover, it is shown in [38] that absolute concentration robustness of the deterministic model does not necessarily imply an extinction event in the corresponding stochastic model, but the connection is still largely unexplored. The results developed in the present paper can help in this direction: consider again (2.1). The extinction of species B cannot occur if production of B is included in the model as in (2.3), or as in

$$\begin{aligned} A + B &\xrightarrow{\kappa_1} 2B \\ B &\xrightarrow{\kappa_2} A \\ 0 &\underset{\kappa_4}{\overset{\kappa_3}{\rightleftharpoons}} B \end{aligned}$$

At the same time, it follows from theorem 5.1 that the stability properties of the species A are maintained both in (2.3) and in (6.1), when deterministically modelled. In particular, the concentration of the species A still converges to the value κ_2/κ_1. It would be interesting to study if in this and in similar cases some form of absolute concentration robustness arise in the long-term dynamics of the stochastic models as well.

Data accessibility. This article has no additional data.

Competing interests. We declare we have no competing interests.

Funding. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme grant agreement no. 743269 (CyberGenetics project).

References

1. Feinberg M. 2019 *Foundations of chemical reaction network theory*. New York, NY: Springer.
2. Tóth J, Nagy AL, Papp D. 2018 *Reaction kinetics: exercises, programs and theorems*. Berlin, Germany: Springer.
3. Anderson DF, Enciso GA, Johnston MD. 2014 Stochastic analysis of biochemical reaction networks with absolute concentration robustness. *J. R. Soc. Interface* **11**, 20130943. (doi:10.1098/rsif.2013.0943)
4. Alon U, Surette MG, Barkai N, Leibler S. 1999 Robustness in bacterial chemotaxis. *Nature* **397**, 168–171. (doi:10.1038/16483)
5. Barkai N, Leibler S. 1997 Robustness in simple biochemical networks. *Nature* **387**, 913–917. (doi:10.1038/43199)
6. Batchelor E, Goulian M. 2003 Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system. *Proc. Natl Acad. Sci. USA* **100**, 691–696. (doi:10.1073/pnas.0234782100)
7. Blanchini F, Franco E. 2011 Structurally robust biological networks. *BMC Syst. Biol.* **5**, 74. (doi:10.1186/1752-0509-5-74)
8. Miyashiro T, Goulian M. 2008 High stimulus unmasks positive feedback in an autoregulated bacterial

signaling circuit. *Proc. Natl Acad. Sci. USA* **105**, 17 457–17 462. (doi:10.1073/pnas.0807278105)

9. Pratt LA, Silhavy TJ. 1995 Porin regulon of *Escherichia coli*. In *Two-component signal transduction* (eds JA Hoch, TJ Silhavy), pp. 105–127. Washington, DC: ASM Press.
10. Shinar G, Alon U, Feinberg M. 2009 Sensitivity and robustness in chemical reaction networks. *SIAM J. Appl. Math.* **69**, 977–998. (doi:10.1137/080719820)
11. Shinar G, Rabinowitz JD, Alon U. 2009 Robustness in glyoxylate bypass regulation. *PLoS Comput. Biol.* **5**, e1000297. (doi:10.1371/journal.pcbi.1000297)
12. Stock AM, Robinson VL, Goudreau PN. 2000 Two-component signal transduction. *Annu. Rev. Biochem.* **69**, 183–215. (doi:10.1146/annurev.biochem.69.1.183)
13. Xiao F, Doyle JC. 2018 Robust perfect adaptation in biomolecular reaction networks. In *2018 IEEE Conf. on Decision and Control* (*CDC*), pp. 4345–4352. New York, NY: IEEE.
14. Shinar G, Feinberg M. 2010 Structural sources of robustness in biochemical reaction networks. *Science* **327**, 1389–1391. (doi:10.1126/science.1183372)
15. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999 From molecular to modular cell biology. *Nature* **402**, C47–C52. (doi:10.1038/35011540)
16. Purnick PEM, Weiss R. 2009 The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell. Biol.* **10**, 410–422. (doi:10.1038/nrm2698)
17. Del Vecchio D, Ninfa AJ, Sontag ED. 2008 Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.* **4**, 161. (doi:10.1038/msb4100204)
18. Mishra D, Rivera PM, Lin A, Del Vecchio D, Weiss R. 2014 A load driver device for engineering modularity in biological networks. *Nat. Biotechnol.* **32**, 1268–1275. (doi:10.1038/nbt.3044)
19. Pantoja-Hernández L, Martínez-García JC. 2015 Retroactivity in the context of modularly structured biomolecular systems. *Front. Bioeng. Biotechnol.* **3**, 85. (doi:10.3389/fbioe.2015.00085)
20. Dexter JP, Dasgupta T, Gunawardena J. 2015 Invariants reveal multiple forms of robustness in bifunctional enzyme systems. *Integr. Biol.* **7**, 883–894. (doi:10.1039/c5ib00009b)
21. Karp RL, Millán MP, Dasgupta T, Dickenstein A, Gunawardena J. 2012 Complex-linear invariants of biochemical networks. *J. Theor. Biol.* **311**, 130–138. (doi:10.1016/j.jtbi.2012.07.004)
22. Banaji M, Pantea C. 2018 The inheritance of nondegenerate multistationarity in chemical reaction networks. *SIAM J. Appl. Math.* **78**, 1105–1130. (doi:10.1137/16M1103506)
23. Feliu E, Cappelletti D, Wiuf C. 2018 Node balanced steady states: unifying and generalizing complex and detailed balanced steady states. *Math. Biosci.* **301**, 68–82. (doi:10.1016/j.mbs.2018.03.002)
24. Feliu E, Wiuf C. 2011 Enzyme-sharing as a cause of multi-stationarity in signalling systems. *J. R. Soc. Interface* **9**, 20110664. (doi:10.1098/rsif.2011.0664)
25. Gross E, Harrington HA, Meshkat N, Shiu A. 2020 Joining and decomposing reaction networks. *J. Math. Biol.* **80**, 1–49.
26. Joshi B, Shiu A. 2013 Atoms of multistationarity in chemical reaction networks. *J. Math. Chem.* **51**, 153–178.
27. Kim J, Enciso G. 2020 Absolutely robust controllers for chemical reaction networks. *J. R. Soc. Interface* **17**, 20200031. (doi:10.1098/rsif.2020.0031)
28. Anderson DF. 2007 A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *J. Chem. Phys.* **127**, 214107. (doi:10.1063/1.2799998)
29. Brunner JD, Craciun G. 2018 Robust persistence and permanence of polynomial and power law dynamical systems. *SIAM J. Appl. Math.* **78**, 801–825. (doi:10.1137/17M1133762)
30. Cappelletti D, Majumder AP, Wiuf C. In preparation. Fixed-time and long-term dynamics of monomolecular reaction networks in stochastic environment.
31. Craciun G, Nazarov F, Pantea C. 2013 Persistence and permanence of mass-action and power-law dynamical systems. *SIAM J. Appl. Math.* **73**, 305–329. (doi:10.1137/100812355)
32. Gopalkrishnan M, Miller E, Shiu A. 2014 A geometric approach to the global attractor conjecture. *SIAM J. Appl. Dyn. Syst.* **13**, 758–797. (doi:10.1137/130928170)
33. Jahnke T, Huisinga W. 2007 Solving the chemical master equation for monomolecular reaction systems analytically. *J. Math. Biol.* **54**, 1–26. (doi:10.1007/s00285-006-0034-x)
34. Shinar G, Milo R, Martínez MR, Alon U. 2007 Input–output robustness in simple bacterial signaling systems. *Proc. Natl Acad. Sci. USA* **104**, 19 931–19 935. (doi:10.1073/pnas.0706792104)
35. Feinberg M. 1987 Chemical reaction network structure and the stability of complex isothermal reactors—I. The deficiency zero and deficiency one theorems. *Chem. Eng. Sci.* **42**, 2229–2268. (doi:10.1016/0009-2509(87)80099-4)
36. Aström KJ, Murray RM. 2010 *Feedback systems: an introduction for scientists and engineers*. Princeton, NJ: Princeton University Press.
37. Doyle JC, Francis BA, Tannenbaum AR. 1992 *Feedback control theory*. New York, NY: Macmillan.
38. Anderson DF, Cappelletti D. 2019 Discrepancies between extinction events and boundary equilibria in reaction networks. *J. Math. Biol.* **79**, 1253–1277. (doi:10.1007/s00285-019-01394-9)
39. Cappelletti D, Wiuf C. 2016 Product-form Poisson-like distributions and complex balanced reaction systems. *SIAM J. Appl. Math.* **76**, 411–432. (doi:10.1137/15M1029916)
40. Anderson DF, Kurtz TG. 2015 *Stochastic analysis of biochemical systems*, vol. 1. New York, NY: Springer.
41. Érdi P, Tóth J. 1989 *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models*. Manchester, UK: Manchester University Press.
42. Anderson DF, Cappelletti D, Kurtz TG. 2017 Finite time distributions of stochastically modeled chemical systems with absolute concentration robustness. *SIAM J. Appl. Dyn. Syst.* **16**, 1309–1339. (doi:10.1137/16M1070773)