



Research

Cite this article: Enciso GA. 2016 Transient absolute robustness in stochastic biochemical networks. *J. R. Soc. Interface* **13**: 20160475. <http://dx.doi.org/10.1098/rsif.2016.0475>

Received: 15 June 2016

Accepted: 5 August 2016

Subject Category:

Life Sciences—Mathematics interface

Subject Areas:

computational biology, systems biology, biochemistry

Keywords:

chemical reaction networks, absolute robustness, Poisson distribution, deficiency theory

Author for correspondence:

German A. Enciso

e-mail: enciso@uci.edu

Transient absolute robustness in stochastic biochemical networks

German A. Enciso

Department of Mathematics, University of California, Irvine, CA 92697, USA

GAE, 0000-0001-8908-9539

Absolute robustness allows biochemical networks to sustain a consistent steady-state output in the face of protein concentration variability from cell to cell. This property is structural and can be determined from the topology of the network alone regardless of rate parameters. An important question regarding these systems is the effect of discrete biochemical noise in the dynamical behaviour. In this paper, a variable freezing technique is developed to show that under mild hypotheses the corresponding stochastic system has a transiently robust behaviour. Specifically, after finite time the distribution of the output approximates a Poisson distribution, centred around the deterministic mean. The approximation becomes increasingly accurate, and it holds for increasingly long finite times, as the total protein concentrations grow to infinity. In particular, the stochastic system retains a transient, absolutely robust behaviour corresponding to the deterministic case. This result contrasts with the long-term dynamics of the stochastic system, which eventually must undergo an extinction event that eliminates robustness and is completely different from the deterministic dynamics. The transiently robust behaviour may be sufficient to carry out many forms of robust signal transduction and cellular decision-making in cellular organisms.

1. Introduction

Chemical reaction networks lie at the heart of many biological processes at the cell level such as signal transduction, development, gene expression and metabolism. These processes are often under strict regulation to prevent an unintended outcome given large amounts of noise. A typical protein may vary in concentration by tens of per cents from cell to cell even in seemingly identical populations [1–3]. An important scientific problem is to discover the mechanisms used by cells to produce a consistent phenotype in the face of this variability [4,5].

Work by Shinar & Feinberg [6] showed that under certain *structural* hypotheses on the underlying chemical network, the steady-state concentration of a particular species is independent of the total concentration of other proteins in the system. This property is called *absolute robustness* (or absolute concentration robustness), and it can be determined from the reaction network independently of rate constants. Experimental applications include some of the most common signalling pathways in bacteria, such as the Envz–OmpR system in *Escherichia coli* [7], and possibly other systems such as mammalian olfactory receptor development [8].

Protein copy numbers are often low inside each cell, i.e. it is not uncommon for key proteins in the system to have less than 10 copies per cell. The stochastic nature of the chemical events themselves becomes relevant in such systems, and it has been widely studied in a variety of contexts (e.g. [9–12]). A natural question is how the noise affects absolutely robust systems, and determining the stationary distribution of the absolutely robust variable. Anderson *et al.* [13] found that for many of these systems, when chemical noise is taken into account an extinction event takes place in the long term, destroying the absolute robustness property.

In this work, I consider the *transient behaviour* of absolutely robust systems under stochastic chemical noise. Preliminary simulations [13] have already hinted that the distribution of the absolutely robust variable after finite time resembles a Poisson distribution centred around the deterministic mean. Here

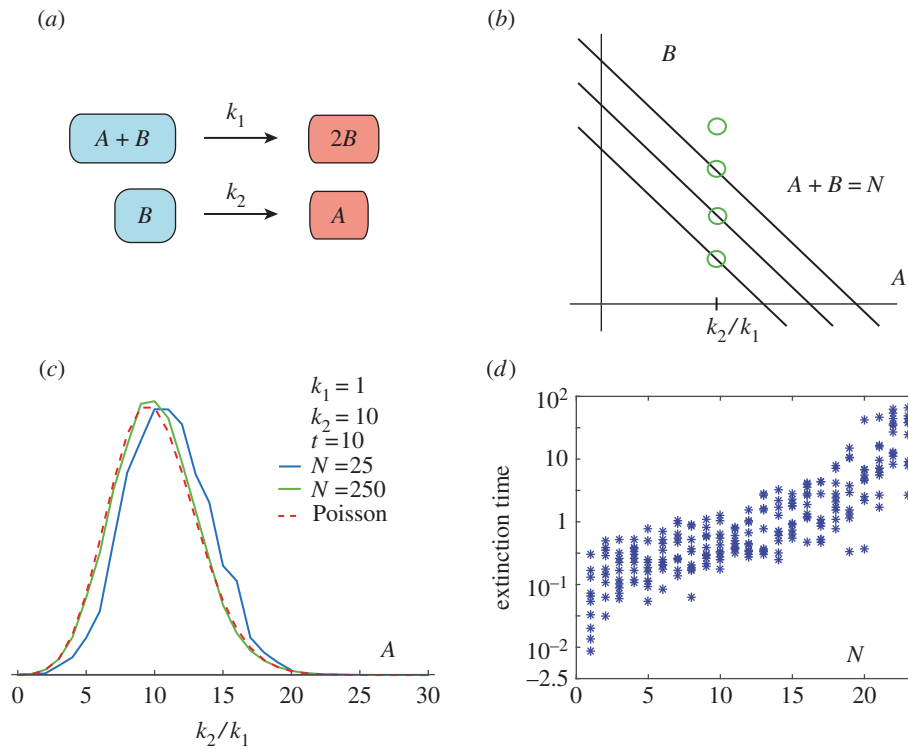
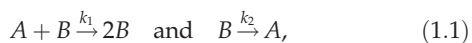


Figure 1. (a) The reactions of the so-called toy model of absolute robustness. (b) Steady states of the toy model for different total protein concentrations N . (c) Poisson distribution (red dotted), and histogram of Gillespie simulations for this model with $k_1 = 1$, $k_2 = 10$ (green) and different values of N . (d) Extinction times for multiple Gillespie simulations using different values of N and initial condition $A = 0$. (Online version in colour.)

this result will be mathematically demonstrated for a large class of systems, effectively establishing a stochastic counterpart to the work by Shinar & Feinberg. Most cellular signalling and decision-making processes take place under a limited timeframe, so that a transient distribution will often form the basis for these decisions when the final distribution takes a long time to be reached.

The concept of absolute robustness, as well as the relevant stochastic effects, can be best explained through a simple example. Consider the chemical reaction



which is described using the system of differential equations

$$A' = -k_1AB + k_2B \quad \text{and} \quad B' = k_1AB - k_2B. \quad (1.2)$$

These reactions satisfy the mass conservation equation $A + B = N$, that is, the total protein amount N is conserved. In this way, A and B can be thought of as different forms of the same protein. At steady state, it holds that $k_2B = k_1AB$, and assuming $B \neq 0$, it follows that $A = k_2/k_1$. Thus at steady state, A is independent of the total concentration N of the protein, a surprising result that does not hold for most reaction networks (figure 1b). This property is called the *absolute robustness* of A in this system. At steady state, the protein $B = N - A$ acts as a buffer and absorbs the excess protein. (We assume here that $N \geq k_2/k_1$, otherwise $B = 0$ and the argument above breaks down.)

Gillespie simulations of this system [14] appear to show that its stochastic behaviour is analogous to that of the deterministic network. In figure 1c, this system was modelled for $t = 10$ and the distribution of A was plotted for two different total protein concentrations, choosing rate parameters so that $k_2/k_1 = 10$. In both cases, the distribution approximates a Poisson distribution with $\lambda = 10$.

1.1. Two-component signalling

The behaviour observed above would be noteworthy but not particularly important if it was only true for that toy model. However, it also holds in larger generality for a family of networks, including a so-called *two-component system* of great importance in bacterial signalling [15–20]. Two-component systems typically consist of a sensory protein X that interacts with a regulatory protein Y . For instance, the EnvZ–OmpR system in *E. coli* regulates osmolarity, which can cause a damaging influx of water across the cell membrane if not properly controlled.

The EnvZ–OmpR osmoregulatory system has been well characterized and can be described as follows (figure 2a). EnvZ has a cytoplasmic osmosensory domain, and it phosphorylates OmpR through a histidine kinase domain. Phosphorylation activates OmpR, which forms a dimer and regulates the transcription of outer membrane porin proteins OmpC and OmpF. These two proteins form pores of two different sizes that regulate the passage of nutrients and other molecules through the cell membrane. Under low osmolarity conditions, the concentration of active OmpR is low, which leads to the expression of OmpF through OmpR binding to high affinity OmpF promoter sites. When osmolarity is high, active OmpR is in high concentration, leading to OmpC expression through OmpR binding to low affinity OmpC promoter sites, and also resulting in OmpF suppression through binding of OmpR to low affinity, OmpF inhibitory sites. In this way, low and high osmolarities lead to OmpF and OmpC expression, respectively.

Importantly, EnvZ not only phosphorylates OmpR, but it can also dephosphorylate it through a separate phosphatase domain [21]. Intuitively, one could think that if total EnvZ concentration were increased, then phosphorylated OmpR concentration might not dramatically change. This is because more EnvZ would more rapidly phosphorylate, but also

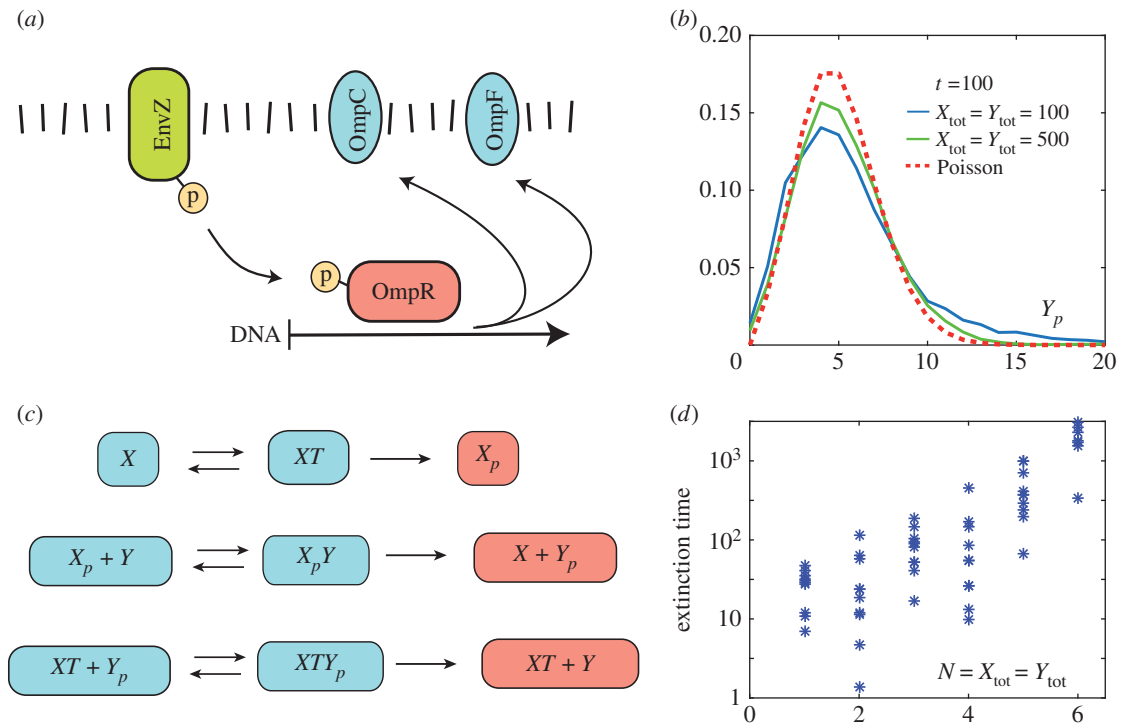


Figure 2. (a) The EnvZ–OmpR two-component pathway regulates osmolarity in *E. coli*. (b) Histograms of Y_p in stochastic simulations of (c) at time $t = 100$ and two different total protein concentrations. Reaction $XT \rightarrow X_p$ has rate $k_3 = 5$, all other reactions have rate one. The deterministic steady state of Y_p is $\lambda_\ell = 5$. (c) Model of EnvZ–OmpR dynamics [6]. (d) Extinction times calculated using different total protein concentrations $X_{\text{tot}} = Y_{\text{tot}} + N$. At the initial condition $XT = Y = N$, all other proteins not present. (Online version in colour.)

more rapidly dephosphorylate, OmpR. This experiment was carried out by Batchelor & Goulian [7], who indeed observed a roughly $\pm 30\%$ change in OmpR-p, over two orders of magnitude change in total EnvZ concentration and under different osmolarity conditions. Surprisingly, OmpR-p concentration was also fairly constant over two orders of magnitude change of *total OmpR itself*. Thus, as total OmpR increases, the active OmpR concentration remains roughly constant while the rest of the protein is in inactive form.

Shinar & Feinberg [6] considered the following model of EnvZ–OmpR dynamics (figure 2c). EnvZ and OmpR are denoted by X and Y , respectively. EnvZ can be dephosphorylated, phosphorylated or simply bound to adenosine triphosphate (ATP) (X , X_p or XT , respectively). When phosphorylated, it has the ability to transfer this phosphorylation to OmpR through a classic Michaelis–Menten reaction. When bound to ATP, it acts as a phosphatase to dephosphorylate OmpR through a similar enzymatic mechanism. Standard mass action rates are used to define the ODE system for this model.

Shinar & Feinberg then demonstrated that OmpR-p is absolutely robust in this system, that is, Y_p has the same concentration at every positive steady state. In that sense, the output of this system is the same regardless of the total amount of the proteins EnvZ and OmpR.

In the current work, it is shown that the OmpR-p concentration has a transiently robust behaviour in the stochastic case. Specifically, the distribution after finite time t for large total protein concentrations $X_{\text{tot}}, Y_{\text{tot}}$ approaches a Poisson distribution

$$\lim_{t \rightarrow \infty} \lim_{X_{\text{tot}}, Y_{\text{tot}} \rightarrow \infty} P(Y_p = i) = e^{-\lambda_\ell} \frac{\lambda_\ell^i}{i!}.$$

Here λ_ℓ is the deterministic steady state of Y_p , that is, the distribution of Y_p in the stochastic system hovers around the

deterministic steady-state λ_ℓ . In particular, $P(Y_p = i) \approx e^{-\lambda_\ell} \lambda_\ell^i / i!$ for reasonably large t and large total protein concentrations.

This is consistent with Gillespie simulations as shown in figure 2b. The distribution of OmpR-p was calculated after a fixed time $t = 100$ and compared with a Poisson distribution centred around the deterministic mean for different total EnvZ and OmpR concentrations. A stationary distribution is found independent of the total protein concentrations.

The long-term picture for any single Gillespie simulation of this system is more complex, since an absorbing boundary event eventually must take place, which destroys the absolute robustness of the system. In this case, an absorbing state is ultimately reached when all EnvZ and OmpR protein is in the phosphorylated state X_p and Y_p , respectively. In that case, none of the reactions in the model can take place, i.e. this is an absorbing steady state that is not robust to total protein concentrations.

However, simulations also indicate that it would likely take a long time to reach this absorbing boundary. Figure 2d shows this through a number of Gillespie simulations with different total protein concentrations. In each case, the system was simulated until it reached extinction, and this time was marked with a blue dot. The fact that the times lie on a rough straight line on a semilog plot is evidence that the average time it takes to reach this absorbing state is a roughly exponential function of total protein concentrations.

In the next section, the main ideas for the proof are outlined in the simple case of the toy model (1.1); the associated Markov process is compared with that of a reduced chemical reaction network that has a simpler dynamics. Section 3 introduces background notation and some existing results for absolute robustness. The main result is stated in detail in §4, and proved in §§5–7 as well as appendix A. Section 8 discusses

several examples in detail, including an analysis of the EnvZ–OmpR model, followed by a discussion.

2. Toy model analysis

The ideas used in the proof of the main result can once again be best introduced using the toy model (1.1). The stochastic system consists of the following continuous-time Markov process [22]:

$$\begin{pmatrix} 0 \\ N \end{pmatrix} \xrightleftharpoons[k_{1 \cdot 1 \cdot (N-1)}]{k_2 \cdot N} \begin{pmatrix} 1 \\ N-1 \end{pmatrix} \xrightleftharpoons[k_{1 \cdot 2 \cdot (N-2)}]{k_2 \cdot (N-1)} \begin{pmatrix} 2 \\ N-2 \end{pmatrix} \cdots \xrightleftharpoons[k_{1 \cdot (N-1) \cdot 1}]{k_2 \cdot 2} \begin{pmatrix} N \\ 0 \end{pmatrix} \\ \times \begin{pmatrix} N-1 \\ 1 \end{pmatrix} \xrightleftharpoons[k_{1 \cdot N \cdot 0}]{k_2 \cdot 1} \begin{pmatrix} N \\ 0 \end{pmatrix}. \quad (2.1)$$

Each node $(i, N-i)$ corresponds to the probability that $A=i$ and $B=N-i$ at a given time. The node $(N, 0)$ is an absorbing state, because the probability to leave it is $k_1 \cdot n \cdot 0 = 0$. This means that at steady distribution, with probability 1 it holds that $A=N, B=0$. Figure 1*d* estimates the mean time before extinction to be roughly exponential as a function of N . However before this state is reached, Gillespie simulations reach a transient distribution similar to a Poisson distribution with mean k_2/k_1 .

To illustrate why this may be the case, consider the rescaled rate parameters $q_1 := k_1 N, q_2 := k_2 N$. In terms of these new parameters, the Markov process has the associated graph

$$\begin{pmatrix} 0 \\ N \end{pmatrix} \xrightleftharpoons[q_{1 \cdot 1 \cdot N-1/N}]{q_2 \cdot N/N} \begin{pmatrix} 1 \\ N-1 \end{pmatrix} \xrightleftharpoons[q_{1 \cdot 2 \cdot N-2/N}]{q_2 \cdot N-1/N} \begin{pmatrix} 2 \\ N-2 \end{pmatrix} \xrightleftharpoons[q_{1 \cdot 3 \cdot N-3/N}]{q_2 \cdot N-2/N} \cdots \quad (2.2)$$

Now, as $N \rightarrow \infty$ this graph formally converges towards a *reduced* system

$$0 \xrightleftharpoons[q_{1 \cdot 1}]{q_2} 1 \xrightleftharpoons[q_{1 \cdot 2}]{q_2} 2 \xrightleftharpoons[q_{1 \cdot 3}]{q_2} 3 \cdots \quad (2.3)$$

The reader might have encountered this graph before—it is the Markov process associated with a simple birth–death model

$$\phi \xrightleftharpoons[q_1]{q_2} A. \quad (2.4)$$

Moreover, a standard calculation shows that this reduced system at steady state has a Poisson distribution

$$P(A=i) = e^{-\lambda} \frac{\lambda^i}{i!},$$

with mean $\lambda := q_2/q_1 = k_2/k_1$. Now, the ODE systems associated with (2.2) and (2.3) cannot be expected to have the same steady-state behaviour, and in fact they behave very differently at steady state. However, as their associated vector fields are similar for large N , one can reasonably expect that after *finite* time t they have similar solutions. Define $x_i(t) := P(A=i, B=N-i)$ at time t for the process (2.2), and similarly $y_i(t) := P(A=i)$ at time t for (2.3). The convergence of (2.2) towards the reduced system (2.3) after fixed finite time t will be rigorously proved in appendix A, in the sense that $\lim_{N \rightarrow \infty} x_i(t) = y_i(t)$ for all i . Since it also holds that $\lim_{t \rightarrow \infty} y_i(t) = e^{-\lambda} \lambda^i / i!$, bringing both equations together

$$\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} x_i(t) = \lim_{t \rightarrow \infty} y_i(t) = e^{-\lambda} \frac{\lambda^i}{i!}.$$

Because the solutions of (2.3) converge globally towards a Poisson distribution, after fixed finite time $t \geq t_0$ the

distribution of the model (2.2) for large N approximates a Poisson distribution centred around λ .

The question remains whether one can systematically associate with other absolutely robust systems a reduced network such as the birth–death process above. In order to do this, recall that the original toy model satisfies the mass conservation equation $A+B=N$. As A approaches k_2/k_1 at steady state, then $B \approx N$ for large values of N . This motivates the step of formally setting $B=N$, i.e. *freezing* the value of B , which leads to its removal from the system since it is now constant. The new chemical reaction system after removal of B and its incorporation into the reaction rates is none other than (2.4).

It will be shown below that this variable freezing procedure can be done in large generality for other networks. In order to freeze a variable, it only needs be part of a mass conservation relation. Multiple variables can be frozen in this way, if there are multiple mass conservation relations available. Note that after freezing a variable, its associated mass conservation relation does not apply anymore.

Regarding the rate parameter scaling and the rate parameters q_i , in this case, one can carry out a rescaling of time so that the values k_i are recovered and the same result holds. But in the more general analysis below, the rate parameters k_i will be defined as functions of N , $k_i := q_i/N$, in such a way that as $N \rightarrow \infty$ the network converges towards a fixed reduced system. A similar treatment of parameter values is sometimes carried out, for example, in stochastic ecological models and even the classic SIR model [23,24].

3. Previous absolute robustness results

To describe the more general results on absolute robustness, including the main result in this paper, we need a few definitions; consider again the EnvZ–OmpR osmolarity system (figure 2*c*). The graph of this network has $C=9$ nodes, known in this literature as *complexes*. The graph also has $\ell=3$ connected components, and the stoichiometry matrix has rank $R=5$. The stoichiometry matrix contains the net number of molecules of each species that are produced or destroyed in each of the reactions [25]. The so-called *deficiency* of the system is defined as $\delta := C - \ell - R = 1$. A complex Q in the network is *terminal* if every random walk along the directed graph edges beginning at Q eventually returns to Q . Thus $X_p + Y$ is *non-terminal*, since such a random walk would eventually reach $X + Y_p$ and not come back.

A steady state of the chemical reaction is *positive* if all its components have positive values. A variable A in the system is *absolutely robust* if all positive steady states have the same value. Assume mass action kinetics for the biochemical networks throughout for simplicity [26]. Also, assume for convenience that the reaction network admits a positive steady state—otherwise there would be nothing to prove about the system.

The main theorem by Shinar & Feinberg [6] for deterministic systems can be stated as follows: given a chemical reaction network such that $\delta=1$ and two different non-terminal complexes differ by exactly one species A , it must follow that A is absolutely robust. In the toy model, it holds $C=4, \ell=2, R=1, \delta=4-2-1=1$, and the non-terminal complexes $A+B$ and B differ by A , therefore A must be absolutely robust. In the EnvZ–OmpR model, $\delta=1$ and

the non-terminal complexes $XT + Y_p$ and XT differ by Y_p , implying that Y_p is absolutely robust.

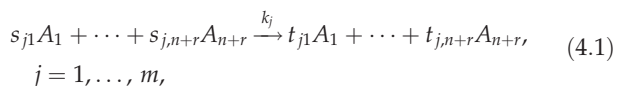
This result is particularly striking in that dynamical information is concluded from structural network information alone, i.e. without knowledge of the rate constants. Of course, the rate constants must be known in order to calculate the absolutely robust value of A concentration, but the fact that it is robust does not require this knowledge.

A chemical reaction network is *conservative* if every species is included in some mass conservation relation. For example, the EnvZ–OmpR system is conservative because the mass conservation equations $X_{\text{tot}} = X_p + X + XT + X_pY + XTY_p$ and $Y_{\text{tot}} = Y_p + Y + X_pY + XTY_p$ include every single species in the system. Given a state of the system with concentrations for each species, a complex in the chemical reaction network is said to be *off* if at least one of the species in the complex has concentration equal to zero.

The main theorem proved by Anderson *et al.* [13] addresses the long-term behaviour of stochastic, absolutely robust networks: given a conservative chemical reaction network such that $\delta = 1$ and two different non-terminal complexes differ by exactly one species A , an extinction event must eventually take place, such that every non-terminal complex turns off and stays off. This can be used to deduce much information about the stationary distribution. For example in the EnvZ–OmpR network, all the complexes labelled in blue must be off at stationary distribution according to the theorem. In particular, the species X , XT , X_pY , XTY_p must have concentration zero at steady state, which implies $X_{\text{tot}} = X_p$ from the mass concentration equation for X_{tot} . Since the complex $X_p + Y$ is also off, it follows that $Y = 0$ because $X_p > 0$. This in turn means that $Y_{\text{tot}} = Y_p$ at steady state, from the mass conservation equation for Y_{tot} . One can conclude that there is a single absorbing state for this system, namely that in which $X_{\text{tot}} = X_p$, $Y_{\text{tot}} = Y_p$ and all other species have concentration zero. The question however remains as to what is the behaviour of such a system before this eventual absorption takes place.

4. Main result statement

Having described the relevant background, the following is the main theorem of this work. Consider a chemical reaction network with $n + r$ species and m reactions,

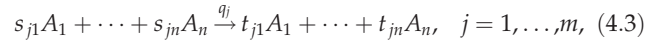


which has r mass conservation relations

$$c_{\gamma 1}A_1 + \dots + c_{\gamma n}A_n + A_{n+\gamma} = d_{\gamma}N, \quad \gamma = 1, \dots, r. \quad (4.2)$$

The variables A_{n+1}, \dots, A_{n+r} are eliminated through a ‘variable freezing’ procedure as illustrated in the Introduction. Before doing this, note that only one variable $A_{n+\gamma}$ appears in each mass conservation equation in (4.2) and the coefficient in front of $A_{n+\gamma}$ is equal to one. This could be done by carrying out linear operations on existing conservation relations as will be shown in the applications. No restrictions are made on the sign of $c_{\gamma i}$; however $d_{\gamma} > 0$, $\gamma = 1 \dots r$. In the case $r = 1$, one can set $d_1 = 1$ simply by rescaling N . We also ask that *at least one of the variables $A_{n+\gamma}$ is turned off at steady state*. This can be easily verified by looking at the reaction network and using the main result in [13].

After formally setting $A_{n+\gamma} = d_{\gamma}N$, the species $A_{n+\gamma}$ becomes constant and disappears as a variable in the system, leading to the reduced network



where

$$q_j = k_j(d_1N)^{s_{j,n+1}} \dots (d_rN)^{s_{j,n+r}} =: k_j\sigma_j(N), \quad (4.4)$$

$$j = 1, \dots, m.$$

To obtain the desired convergence result, fix q_j and define instead $k_j := q_j/\sigma_j(N)$. In practice, one can choose the q_j and N so that the rate parameters k_j have desired values, for instance, to fit experimental data.

Theorem 4.1. *Suppose that a conservative chemical reaction network (4.1) satisfies $\delta = 1$ and has two non-terminal nodes that differ by a single variable A_{ℓ} . Also suppose that after freezing A_{n+1}, \dots, A_{n+r} , the irreducible network (4.3) has deficiency $\delta = 0$. Then the distribution of A_{ℓ} in the stochastic system, for finite t and sufficiently large N , is approximated by a Poisson distribution:*

$$\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} P(A_{\ell} = i) = e^{-\lambda_{\ell}} \frac{\lambda_{\ell}^i}{i!}. \quad (4.5)$$

Moreover, the mean λ_{ℓ} of the distribution is the deterministic steady state of A_{ℓ} in (4.1) in the limit as $N \rightarrow \infty$.

This result is proved by showing that the stochastic system associated with (4.1) has very similar dynamics after fixed finite time to that of (4.3). Since (4.3) has a Poisson stationary distribution for each of its variables, due to a result from chemical reaction network theory [27], then this is also the case for (4.1). The ‘irreducible’ assumption for (4.3) means that the Markov process associated with it has a strongly connected graph. This assumption is necessary to conclude that the solutions of (4.3) converge globally towards the unique stationary distribution.

The overall picture of the dynamical behaviour is described in figure 3. If N is fixed and the stochastic system associated with (4.1) evolves over time, the solution will eventually undergo an extinction event. However for fixed values of t , larger values of N eventually lead to a Poisson-like distribution for A_{ℓ} . The values of t for which this holds need only be larger than or equal to t_0 , the time that it takes for the stochastic reduced network (4.3) to reasonably converge towards a Poisson distribution; that convergence is independent of N .

Regarding the value λ_{ℓ} , it is the steady state of A_{ℓ} in the reduced network (4.3). It is also the limit as $N \rightarrow \infty$ of λ_{ℓ}^N , the steady state of A_{ℓ} for the original system (4.1). As such, it retains the absolute robustness properties of the original network.

Note the similarity with the assumptions of the existing results for deterministic and stochastic absolutely robust networks. The main assumptions made beyond those in the original result by Shinar and Feinberg are that the network is conservative and that after the variable freezing procedure the deficiency is zero, all of which can be verified easily without knowledge of parameter values. The irreducibility of the reduced network can also be readily verified.

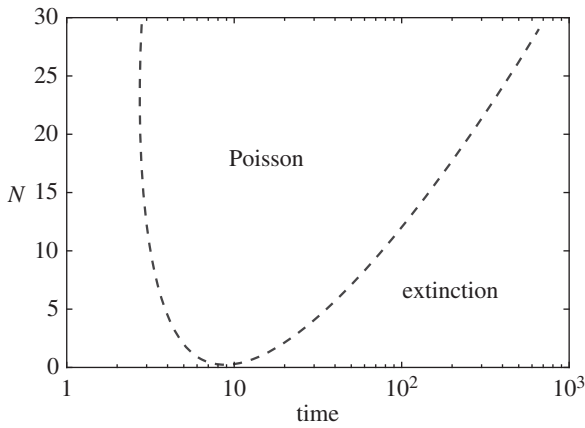


Figure 3. Cartoon depiction of the expected dynamics of the absolutely robust variable in a stochastic simulation, as a function of both time t and the total protein concentration parameter N .

The idea of freezing the value of a variable has been suggested previously in a deterministic context, notably in work by Joshi & Shiu [28].

5. Deterministic network convergence

This section is concerned with proving the convergence of the deterministic dynamics in the network (4.1) towards that of (4.3) after finite time for $N \rightarrow \infty$. It also constitutes a warm up for a similar analysis in the stochastic case. The reduced network (4.3) has the associated system of equations

$$y_i'(t) = \sum_{j=1}^m q_j v_{ji} y_j^{s_j}. \quad (5.1)$$

Here $y_i(t)$ is the concentration of variable A_i at time t , $y = (y_1, \dots, y_n)$, $v_{ji} = t_{ji} - s_{ji}$, and we use the shorthand notation $s_j = (s_{j1}, \dots, s_{jn})$, $y^{s_j} = y_1^{s_{j1}} \dots y_n^{s_{jn}}$.

Also, the original system of reactions (4.1) can be reduced to n species in a standard way by simply writing $A_{n+\gamma}$ in terms of the other variables,

$$A_{n+\gamma} = d_\gamma N - \sum_{\ell=1}^n c_{\gamma\ell} A_\ell, \quad \gamma = 1, \dots, r.$$

The resulting network is described by the system of equations

$$x_i'(t) = \sum_{j=1}^m k_j v_{ji} x_j^{s_j} \prod_{\gamma=1}^r \left(d_\gamma N - \sum_{\ell=1}^n c_{\gamma\ell} x_\ell \right)^{s_{j,n+\gamma}}. \quad (5.2)$$

Now, writing this network using the fixed parameters q_j one obtains

$$x_i'(t) = \sum_{j=1}^m q_j v_{ji} x_j^{s_j} \prod_{\gamma=1}^r \left(\frac{d_\gamma N - \sum_{\ell=1}^n c_{\gamma\ell} x_\ell}{d_\gamma N} \right)^{s_{j,n+\gamma}}. \quad (5.3)$$

This allows one to compare the behaviour of the reduced ODE system (5.1) with that of the original ODE system (5.3) for $N \rightarrow \infty$. As N increases, it is clear that the equation on the right-hand side of (5.3) converges towards that of (5.1). This convergence holds pointwise and is also uniform on compact sets. One can conclude that

$$\lim_{N \rightarrow \infty} x_i(t) = y_i(t),$$

for every $t > 0$ and $i = 1, \dots, n$, provided both systems start with the same initial condition and both systems are defined for that time interval.

6. Stochastic network convergence

A similar argument can be carried out to compare the master equation of both systems in the stochastic case. The state of the network at any time is given by a vector of non-negative integers $I = (i_1, \dots, i_n)$, representing the number of molecules for each species. Denoting $y_I(t) = P(A_1 = i_1, \dots, A_n = i_n)$ at time t for the reduced network (4.3), these probabilities evolve over time according to the ODE system

$$y_I' = \sum_{j=1}^m q_j (I - v_j)^{(s_j)} y_{I-v_j} - \sum_{j=1}^m q_j I^{(s_j)} y_I. \quad (6.1)$$

Here $v_j := (v_{j1}, \dots, v_{jn}) = t_j - s_j$ is the j th reaction vector, $i^{(m)} := i(i-1) \dots (i-m+1)$ for any non-negative integers i , m and $I^{(s)} := i_1^{(s_{11})} \dots i_n^{(s_{n1})}$.

Moving on to the original network (4.1), one can once again eliminate all the variables $A_{n+\gamma}$ by writing them in terms of the other variables, $A_{n+\gamma} = d_\gamma N - \sum_{\ell=1}^n c_{\gamma\ell} A_\ell$. In this case,

$$\begin{aligned} x'_{I,N} &= \sum_{j=1}^m k_j (I - v_j)^{(s_j)} \prod_{\gamma=1}^r \left(d_\gamma N - \sum_{\ell=1}^n c_{\gamma\ell} (i_\ell - v_{j\ell}) \right)^{(s_{j,n+\gamma})} x_{I-v_j,N} \\ &\quad - \sum_{j=1}^m k_j I^{(s_j)} \prod_{\gamma=1}^r \left(d_\gamma N - \sum_{\ell=1}^n c_{\gamma\ell} i_\ell \right)^{(s_{j,n+\gamma})} x_{I,N}. \end{aligned}$$

Using the notation for q_j instead of k_j , this can be written as

$$\begin{aligned} x'_{I,N} &= \sum_{j=1}^m q_j (I - v_j)^{(s_j)} \prod_{\gamma=1}^r \frac{[d_\gamma N - \sum_{\ell=1}^n c_{\gamma\ell} (i_\ell - v_{j\ell})]^{(s_{j,n+\gamma})}}{(d_\gamma N)^{s_{j,n+\gamma}}} x_{I-v_j,N} \\ &\quad - \sum_{j=1}^m q_j I^{(s_j)} \prod_{\gamma=1}^r \frac{[d_\gamma N - \sum_{\ell=1}^n c_{\gamma\ell} i_\ell]^{(s_{j,n+\gamma})}}{(d_\gamma N)^{s_{j,n+\gamma}}} x_{I,N}. \end{aligned} \quad (6.2)$$

One can verify that for any non-negative integers N , i and s , it holds $(N+i)^s / N^s \rightarrow 1$ as $N \rightarrow \infty$. Applying this principle above, as $N \rightarrow \infty$ in the above equation the right-hand side converges pointwise towards that of (6.1).

An argument as described in the Introduction would imply that if the same initial condition is used for both systems (6.2) and (6.1), and if each is integrated over a finite interval $[0, t]$, then for each I it holds that

$$\lim_{N \rightarrow \infty} x_{I,N} = y_I(t). \quad (6.3)$$

The problem is that one of the two systems of equations, that of the reduced network, has infinitely many state variables I , so that this convergence is a non-trivial problem. Equation (6.3) is rigorously proved in appendix A using a technique known as finite-state projection (FSP), originally developed by Munsky [29], Khammash and co-workers [24]. FSP is normally used to carry out numerical approximations of master equation solutions and to establish numerical error bounds. But it can also be helpful as a form of theoretical analysis.

An intuitive explanation of this convergence for finite time t is the following. Suppose that the master equation (6.2) is written in compact form as $x' = Mx$, where $M_{IJ} = \mu_N(I, J)$ is the intensity of the transition from I to J for every two states I, J . Suppose that at time $t=0$ the system is in state J , i.e. $x_J(0) = 1$ and $x_I(0) = 0$ for all $I \neq J$.

One can use the shorthand notation $x(0) = e_j$. Then a k th-order Taylor approximation can be used to accurately approximate $x(t)$ over a finite range of values of t :

$$x(t) = e^{Mt}x(0) \approx \left(I + Mt + \dots + \frac{M^k t^k}{k!} \right) e_j.$$

Recall that the matrix entry $(M^\ell)_{IJ}$ can be written as

$$(M^\ell)_{IJ} = \sum_{I=I_1, I_2, \dots, I_\ell=J} \mu_N(I_1, I_2) \cdots \mu_N(I_{k-1}, I_k),$$

where the sum is taken over all directed paths from I to J of length ℓ . One can (at least formally) write the master equation of the reduced system (6.1) as $y' = Ly$. For a fixed state I , approximate

$$\begin{aligned} x_I(t) &= (e^{Mt}e_j)_I \approx I + M_{IJ}t + \dots + \frac{(M^k)_{IJ}t^k}{k!} \\ &\approx I + L_{IJ}t + \dots + \frac{(L^k)_{IJ}t^k}{k!} \approx y_I(t). \end{aligned} \quad (6.4)$$

As $N \rightarrow \infty$, the paths of length up to k in both systems coincide, and M_{IJ}^ℓ converges to L_{IJ}^ℓ . Although the matrix L is infinite, only finitely many paths are considered in the equation above. As k is increased, the range of values of t for which the approximation applies increases but remains bounded, stressing the fact that this approximation is only valid for finite t . In other words, if N is fixed and t grows indefinitely, equation (6.4) eventually breaks down.

7. Proof of the theorem

One can now finish the proof of the main theorem, including equation (4.5) on the Poisson-like transient distribution of the absolutely robust variable. Recall that the original biochemical network (4.1) has deficiency $\delta = 1$, admits a positive equilibrium, and two non-terminal linkage classes differ by the species A_ℓ . Moreover, after freezing the variables A_{n+1}, \dots, A_{n+r} , it reduces to equation (4.3) which has deficiency $\delta = 0$ and is irreducible as a Markov process.

There is a mature theory for stochastic chemical reaction networks with deficiency zero. But for this theory to apply it is important for the network to have a technical property known as weak reversibility. It will be left to the end of this section to prove that under the given hypotheses this is the case. Now, given $\delta = 0$ and weak reversibility, a result by Anderson *et al.* [27] ensures that the system converges globally towards a product-form Poisson stationary distribution

$$\lim_{t \rightarrow \infty} y_I(t) = \prod_{\ell=1}^n e^{-\lambda_\ell} \frac{\lambda_\ell^{i_\ell}}{i_\ell!}, \quad (7.1)$$

where $\lambda = (\lambda_1, \dots, \lambda_n)$ is the unique steady state of the corresponding deterministic system (4.3). At stationary distribution each of the species A_i in the stochastic reaction network has marginal Poisson distribution and is uncorrelated from the other variables, a remarkable result for a nonlinear system. The product-form Poisson distribution is the global attractor of the master equation because this system is irreducible.

From equations (6.3) and (7.1), it follows that

$$\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} x_{I,N}(t) = \lim_{t \rightarrow \infty} y_I(t) = \prod_{\ell=1}^n e^{-\lambda_\ell} \frac{\lambda_\ell^{i_\ell}}{i_\ell!}.$$

Then the absolutely robust variable A_ℓ has Poisson distribution with mean λ_ℓ in this limit, leading to the main equation (4.5)

$$\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} P(A_\ell = i) = e^{-\lambda_\ell} \frac{\lambda_\ell^i}{i!}$$

for the original system (4.1).

At this point, we know that the transient distribution of A_ℓ approaches a Poisson distribution with mean λ_ℓ , but we still need to relate λ_ℓ with a deterministic steady-state value for (4.1) (recall that λ_ℓ is just the steady state for A_ℓ in the reduced system (4.3)).

For this, we will use some properties of deterministic systems with $\delta = 0$ and weak reversibility. By the deficiency zero theorem [30,31], λ is the unique steady state of (4.3), has only positive entries and is locally stable. This means that small changes to the reduced vector field (5.1) lead to small changes in the steady-state concentrations. Given the convergence of (5.3) towards the reduced system (5.1) as $N \rightarrow \infty$, it follows that there are positive steady states λ^N of the original biochemical network (4.1) satisfying mass conservation equation (4.2), such that

$$\lambda^N \rightarrow \lambda \text{ as } N \rightarrow \infty.$$

That means that $\lambda_\ell^N \rightarrow \lambda_\ell$ as $N \rightarrow \infty$. That is, for large N the steady states of both reaction networks are similar, and the Poisson-like distribution of A_ℓ is approximately centred around the steady state λ_ℓ^N of A_ℓ . This completes the proof of the theorem.

We conclude the section with a proof that the reduced chemical master equation (6.1) is weakly reversible. Since the original deterministic network (4.1) has a positive steady state, it follows [32] that it is consistent, that is, there exist $\gamma_j > 0$, $j = 1, \dots, m$ such that

$$\sum_j \gamma_j v_j = 0,$$

where $v_j := (t_{j1} - s_{j1}, \dots, t_{j,n+r} - s_{j,n+r})$ are the reaction vectors of the system. The reduced network (4.3) has very similar reaction vectors $\bar{v}_j := (t_{j1} - s_{j1}, \dots, t_{j,n} - s_{j,n})$, and obviously it also holds $\sum_j \gamma_j \bar{v}_j = 0$, that is, the reduced network is also consistent. But any consistent network must have a positive equilibrium for some kinetic rates \hat{k}_j (for instance, letting $\hat{k}_j = \gamma_j$ and $x_i = 1$ for all i). Since the reduced network has deficiency zero, it follows from the deficiency zero theorem that this network is weakly reversible [33,34].

8. Examples

In this section, we illustrate the main result in the context of several examples. Consider first a generalization of the toy model (1.1) that has $n + 1$ variables rather than two, given by the network in figure 4a. This network has the sole mass conservation law

$$A_1 + \dots + A_n + 1 = N,$$

which makes it a conservative network because it involves all species. It also has $\ell = 2$ linkage classes, $n + 3$ complexes, rank $R = n$, and $\delta = C - \ell - R = (n + 3) - 2 - n = 1$. Finally, the non-terminal complexes $A_1 + A_{n+1}$ and A_{n+1} differ by A_1 . This means that by the Shinar and Feinberg theorem, A_1 is absolutely robust, and by the Anderson, Enciso and Johnston theorem, eventually an extinction event takes place in which A_2, \dots, A_{n+1} are equal to zero and $A_1 = N$.

By formally freezing the variable $A_{n+1} = N$, we obtain the network from figure 4b, which has deficiency

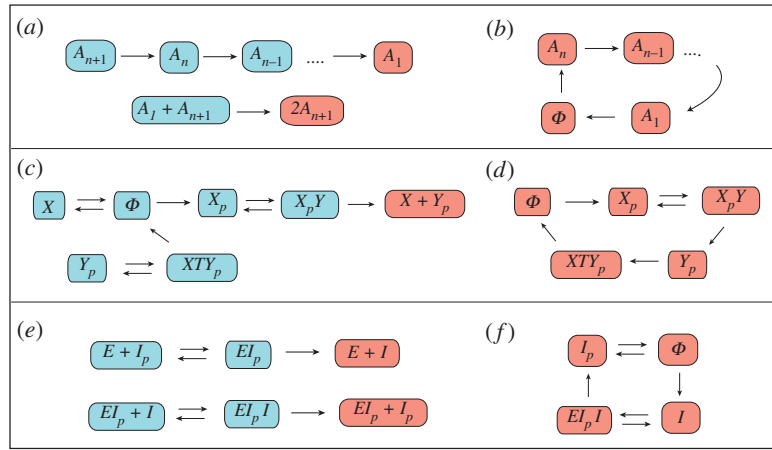


Figure 4. (a,b) A generalization of the toy model using $n+1$ species, and its corresponding reduced network after freezing the variable A_{n+1} . (c) Network resulting from the EnvZ–OmpR model after freezing the variables Y and XT . (d) The same network after removal of X , which does not affect the dynamics of the other variables in (c). (e,f) Core absolutely robust model first used by Shinar & Feinberg [6], and reduced model after freezing E and EI_p . (Online version in colour.)

$\delta = (n+1) - 1 - n = 0$. The parameter scaling is $k_1 = q_1/N$, $k_{n+1} = q_{n+1}/N$ and $k_i = q_i$ for $i = 2 \dots n$.

The main theorem applies to conclude that for the original network in figure 4a

$$\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} P(A_1 = i) = e^{-\lambda_1} \frac{\lambda_1^i}{i!},$$

where $\lambda_1 = q_{n+1}/q_1 = k_{n+1}/k_1$ is the steady state for A_1 in the reduced network. One can also verify by inspection that $\lambda_1^N = k_{n+1}/k_1$ is also the steady state for the original network, for arbitrary N .

The second example is the EnvZ–OmpR model that has been used throughout this work. EnvZ, denoted by X , is a bifunctional sensory protein, and OmpR, denoted by Y , is its associated actuator. Phosphorylated EnvZ activates OmpR by phosphorylation, and EnvZ bound to ATP dephosphorylates OmpR (figure 2c). The mass conservation equations in this case are

$$X_{\text{tot}} = X_p + X + XT + X_p Y + XTY_p \quad \text{and} \\ Y_{\text{tot}} = Y_p + Y + X_p Y + XTY_p.$$

This network is conservative since all species are involved in at least one of the two equations. By calculating the rank of the stoichiometry matrix, $\delta = 9 - 3 - 5 = 1$. Moreover, the complexes XT and $XT + Y_p$ differ by Y_p , hence Y_p is absolutely robust in the deterministic system. As was discussed previously, the long-term behaviour of any stochastic simulation is necessarily $X_p = Y_{\text{tot}}$, $Y_p = Y_{\text{tot}}$.

By trial and error, one can find that freezing the variables Y and XT leads to the reduced network in figure 4c. At first, there seems to be a problem, since this network still has deficiency $\delta = 7 - 1 - 5 = 1$ and is not weakly reversible. However notice that in this system the variable X is decoupled from the rest of the network, i.e. it does not affect any other variable. Without altering the dynamics of the deterministic or stochastic network, one can eliminate this variable and obtain the network in figure 4d, which indeed has deficiency zero and is weakly reversible. Finally, note that the mass conservation equations include exactly one frozen variable each, as required in the format (4.2), and that both of the frozen variables eventually turn off during the extinction event. Setting $XT = X_{\text{tot}} = d_1 N$, $Y = Y_{\text{tot}} = d_2 N$, one can apply the main theorem and conclude that

$$\lim_{t \rightarrow \infty} \lim_{X_{\text{tot}}, Y_{\text{tot}} \rightarrow \infty} P(Y_p = i) = e^{-\lambda_I} \frac{\lambda_I^i}{i!},$$

since $N \rightarrow \infty$ is equivalent to the growth of the total protein concentrations.

The next example is the so-called core absolutely robust module discussed by Shinar & Feinberg [6]. This is a simple model of a two-component signal transduction network, with a bifunctional sensory protein E and a regulatory protein I . Each of the two proteins may be phosphorylated or dephosphorylated, and they can also form dimer and trimer complexes as shown in figure 4e. This system satisfies the two mass conservation equations

$$I_{\text{tot}} = I + I_p + EI_p + 2EI_p I \quad \text{and} \quad E_{\text{tot}} = E + EI_p + EI_p I$$

and is therefore conservative, since every species is in one of these two equations. Also $\delta = 6 - 2 - 3 = 1$ and the non-terminal complexes $EI_p + I$ and EI_p differ by I , so that I is absolutely robust.

By trying out different combinations of species, one can see that freezing the species E and EI_p leads to a network with deficiency $\delta = 4 - 1 - 3 = 0$ (figure 4f). Notice that EI_p forms a non-terminal complex, so it must turn off at stationary distribution for any stochastic simulation. Since the mass conservation equations are not in the correct format (exactly one frozen variable must appear in each equation), one can rearrange as

$$I_{\text{tot}} = I + I_p + 2EI_p I + EI_p = d_1 N \quad \text{and} \\ E_{\text{tot}} - I_{\text{tot}} = -I - I_p - EI_p I + E = d_2 N.$$

By freezing the variables $EI_p = d_1 N$ and $E = d_2 N$, the system is reduced to one with deficiency zero. By the theorem, the transient behaviour of I is approximately Poisson as $N \rightarrow \infty$. Since $I_{\text{tot}} = d_1 N$ and $E_{\text{tot}} = d_2 N + I_{\text{tot}} = (d_2 + d_1)N$, the growth of N is equivalent to that of E_{tot} , I_{tot} :

$$\lim_{t \rightarrow \infty} \lim_{E_{\text{tot}}, I_{\text{tot}} \rightarrow \infty} P(I = i) = e^{-\lambda_I} \frac{\lambda_I^i}{i!}.$$

The mean λ_I approximates the deterministic steady state of I in the core absolutely robust network as the total protein concentrations grow.

9. Fitting to experimental data

Suppose that a modeller is interested in describing the behaviour of a specific biochemical system that appears to be absolutely robust. After writing down the chemical reactions

(4.1), the modeller estimates the value of the rate parameters and total protein concentrations on the right-hand side of (4.2) based on various forms of experimental data. Let us call the baseline rate parameters $k_{j,0}$, and $d_1 N_0, \dots, d_r N_0$ the total protein concentrations.

The idea is to find values of the q_j in such a way that when $N = N_0$, it holds that $k_j = k_{j,0}$ and all parameters reach their experimentally determined values. The parameters k_j have already been defined in (4.4) as $k_j = q_j / \sigma_j(N)$, where

$$\sigma_j(N) = (d_1 N)^{s_{j,1}} \dots (d_r N)^{s_{j,r}}.$$

So let us define

$$q_j := k_{j,0} \sigma_j(N_0) \quad \text{and} \quad k_j := \frac{q_j}{\sigma_j(N)} = k_{j,0} \frac{\sigma_j(N_0)}{\sigma_j(N)}.$$

This way when $N = N_0$, it holds $k_j = k_{j,0}$. Since the main result shows that for fixed $t > 0$ the distribution of A_j is approximately Poisson for large N , it is hoped that N_0 is sufficiently large that this approximation holds.

As it turns out, the absolute robustness of the original network (4.1) provides a form of absolute robustness to the reduced network (4.3). Specifically, the steady-state concentration λ_ℓ of A_ℓ in the reduced network is independent of N_0 . This is surprising because the q_j depend on N_0 , so in principle one would expect that the steady states of that network also depend on this parameter.

To see why this is the case, recall that $\lambda \in \mathbb{R}^n$ is the unique steady state of (4.3), and that there exists a sequence of positive steady states $\lambda^N \in \mathbb{R}^n$ of (4.1) satisfying mass conservation (4.2), such that $\lambda^N \rightarrow \lambda$ as $N \rightarrow \infty$. Also recall that A_ℓ is an absolutely robust species in (4.1).

Denote $\beta^N := \lambda^N$ and $\beta := \lambda$ in the special case $N_0 = 1$. Note that

$$k_j = k_{j,0} \frac{\sigma_j(N_0)}{\sigma_j(N)} = k_{j,0} \frac{\sigma_j(1)}{\sigma_j(N/N_0)}.$$

That is, k_j would have the same value if N_0 was replaced with 1 and N was replaced with N/N_0 . Moreover, by absolute robustness of A_ℓ , for large N one could replace the mass conservation equation (4.2) with

$$c_{\gamma 1} A_1 + \dots + c_{\gamma n} A_n + A_{n+\gamma} = d_\gamma \frac{N}{N_0},$$

without changing the steady-state concentration of A_ℓ . That is, $\lambda_\ell^N = \beta_\ell^{N/N_0}$. As $N \rightarrow \infty$,

$$\lambda_\ell = \lim_{N \rightarrow \infty} \lambda_\ell^N = \lim_{N \rightarrow \infty} \beta_\ell^{N/N_0} = \beta_\ell, \quad (9.1)$$

which does not depend on N_0 .

10. Discussion

This paper has described the stochastic behaviour of systems that are absolutely robust in the deterministic framework, and it has shown that a form of transient robustness is preserved before an eventual population collapse. This dynamics is analogous to the behaviour of stochastic ecological models involving, for example, a predator and prey in an island. While the dynamics can resemble that of the corresponding deterministic system after finite time, in the long run one of the species will die off after an unlikely sequence of events, causing a drastic qualitative change. Because this extinction appears to take place at a very long time scale (e.g. in exponential time as a

function of the population), in many applications the relevant outcomes are determined away from equilibrium.

The main result here is framed in a very similar way as the original result from Shinar and Feinberg, involving networks with deficiency one and mass action kinetics. This was done in part for convenience, because it is now known that absolute robustness holds for many systems of higher deficiency [35]. Simulations appear to show that the conservative assumption is not necessary for the result to hold, which could be proved in future work. Similarly, no attempt was made to generalize the above results to non-mass action systems, e.g. involving different exponents in the kinetic reaction rates.

Importantly, the proof of the main result implies more than was stated, namely, the distribution of other variables in the original system is also approximately Poisson in the limit as $N \rightarrow \infty$. This has a number of consequences for the behaviour of these systems as predicted by methods such as fluctuation dissipation or the linear noise approximation, which will be studied in future work. Overall, it appears that variable freezing can be applied in a wide range of low-deficiency systems to prove that their behaviour transiently approximates the behaviour of zero deficiency systems, both in the deterministic and stochastic cases.

As this paper neared completion the author was made aware of an unpublished manuscript by D. Anderson, D. Cappelletti and T. Kurtz on the topic of stochastic absolute robustness. While there is overlap with that work, there are significant differences in the assumptions and particularly in the methodology of both forms of analysis, so that the two studies are broadly complementary.

Recall that the main result was applied to a specific model of two-component signalling involving the osmolarity regulatory proteins EnvZ and OmpR. To what extent could this result apply to other networks? The key for the application of the result is the existence of a bifunctional enzyme that can carry out a modification (such as phosphorylation) and separately undo it through another reaction domain. This feature is very common in two-component systems, in that the sensory protein is usually a bifunctional histidine kinase/phosphatase. The fact that the same protein is carrying out two separate, similar reactions introduces a redundancy into the system, which decreases the rank of the stoichiometry matrix of the network, and thereby increases the deficiency. For example, in the EnvZ–OmpR model (figure 2c), the three proteins X , X_p , XT can turn into each other and therefore are equivalent in the network for the purposes of rank calculation. One also has the overall reactions



where X_* stands for one of the equivalent proteins X , X_p , XT . There is a linear dependence between these two reactions which reduces the rank of the reaction matrix, and therefore increases the deficiency from zero to one. This relationship between enzyme bifunctionality and deficiency likely also applies to other models of two-component signalling, as well as to other models involving a bifunctional enzyme. Notice that if one replaced the separate phosphorylation and dephosphorylation reactions by a single, reversible enzymatic reaction, this argument would break down and the deficiency of the system would not increase. This could be an important reason for the rather mysterious existence of a separate phosphatase domain in EnvZ and other histidine kinases [21].

It is surprising that although the abstract result does not require protein bifunctionality, most of the biologically motivated examples of this theorem in the literature involve bifunctional enzymes (see also [19,36]). The simplest example of absolute robustness, the toy model described in the Introduction, can also be thought of as involving a bifunctional enzyme B , which catalyses the degradation of A ($A + B \rightarrow 2B$) but also spontaneously promotes it ($B \rightarrow A$). Another interesting example of enzyme 'bifunctionality' is the histone demethylase LSD1 involved in olfactory receptor development [8]. This protein eliminates methylation sites in histones which are associated with olfactory receptor expression. However, some of these methylation sites promote gene expression, while others inhibit it. In this way, LSD1 can carry out both gene activation and gene inhibition. Absolute robustness may play a strong role in this heavily regulated system.

The specific form for the mass conservation equations (4.2) introduces certain restrictions on the total protein concentrations. For example, in the core absolutely robust model example the chosen frozen variables E, E_l lead to a the condition $E_{\text{tot}} = d_2 N + I_{\text{tot}} \geq I_{\text{tot}}$. However, simulations give no indication that such restriction is necessary. One intuitive way to generalize the algorithm is to iteratively freeze one variable after another, and to apply the main result in sequence, first with the very last network, which has only one mass conservation equation left, then with the second to last, and so forth. Each step could then be used to guarantee an approximately Poisson distribution for the network in the previous iteration. Such a procedure could be proved in the future.

Competing interests. The author has no competing interests.

Funding. The author would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme Stochastic Dynamical Systems in Biology, where work on this paper was undertaken. This work was supported by EPSRC grant no. EP/K032208/1. This material is based upon work supported by the National Science Foundation under grant nos. DMS-1122478 and 1616233.

Acknowledgements. I would like to thank Simon Cotter, Michael Cranston, Oleg Igoshin, Badal Joshi, Ilya Kachkovskiy, Jay Newby and Abhyudai Singh for many helpful suggestions. Credit is due to Badal Joshi in particular for the suggested use of consistent networks in this context.

Appendix A

We address in detail equation (6.3), the convergence of the solution $x_i(t)$ of the master equation (6.2) towards the solution $y_i(t)$ of (6.1) as $N \rightarrow \infty$.

The technique known as finite state projection, or FSP for short, was developed originally by Mustafa Khammash and Brian Munsky to truncate the master equation of a chemical reaction and establish error bounds compared with the original solution (e.g. [29,37]). In an FSP analysis, a linear system $v' = Mv$ corresponding to a Markov process with countably many variables is projected to a system $w' = Nw$ which is identical to the original system except that all states outside of a finite region of interest S are clumped together into a single absorbing state G . All the reaction edges and intensities that are internal to S are unchanged, but all fluxes originally leading out of S are now aimed towards G , and all fluxes leading into S from outside S are eliminated from the system altogether. The basic result that we use regarding such projections is that if both systems are initialized at

time $t = 0$ with the same initial condition b with support in S , then (i) $w_i(t) \leq v_i(t)$ for every $t > 0$ and $i \in S$ and (ii) $\sum_{i \in S} |v_i(t) - w_i(t)| \leq G(t)$, for all $t > 0$. See lemma 5.0.1 and theorem 5.0.2 of [29] for full proofs.

Define S_N to be the finite set of non-absorbing states in the master equation (6.2) of the original system. So for instance, in the toy model of the Introduction S_N corresponds to the states $\{(0, N), \dots, (N-1, 1)\}$, excluding the absorbing state $(N, 0)$. Let $z_N = Kz_N$ be the FSP of the reduced system's master equation (6.1) to S_N . Assume that the system $z_N = Kz_N$ and the reduced system (6.1) both have the same initial condition b with finite support.

Proposition A.1. For every fixed $t > 0$ and state $I \in S_N$,

$$\lim_{N \rightarrow \infty} z_{I,N}(t) = y_I(t).$$

Proof. It will be used here that the reduced network associated with the master equation (6.1) has deficiency zero and is weakly reversible. Denote by $G_N \subseteq S_N$ the set of states $I \in S_N$ that have an outedge leading out of S_N . The standard results from FSP theory guarantee that (i) $z_{I,N}(t) \leq y_I(t)$ for every $I \in S_N$ and (ii) $\sum_{I \in S_N} y_I(t) - z_{I,N}(t) \leq G(t)$. Then for every $I \in S_N$,

$$\begin{aligned} |y_I(t) - z_{I,N}(t)| &\leq \sum_{I \in S_N} y_I(t) - z_{I,N}(t) \leq G(t) = \int_0^t G'(s) ds \\ &\leq \int_0^t \sum_{J \in G_N} \sum_{j=1}^m q_j J^{(s)} z_{J,N}(s) ds \leq \int_0^t \sum_{J \in G_N} \sum_{j=1}^m q_j J^{(s)} y_J(s) ds. \end{aligned}$$

We now bound the different terms in this expression: $q_j \leq \max_j q_j$, $J^{(s)} \leq |J|^\alpha$, for $\alpha := n \max_{ij} s_{ji}$ and $|J| := J_1 + \dots + J_n$. Finally, define $\pi(J) = \prod_{i=1}^n e^{-\lambda_i} \lambda_i^{J_i} / (J_i!)$, the product Poisson distribution that is a steady state of the reduced system (6.3) by the deficiency assumptions [27]. Let $\gamma > 0$ be such that $b_j \leq \gamma \pi(J)$ for all states J . Then a monotonicity property for (6.3) implies that $y_J(s) \leq \gamma \pi(J)$ for all $s > 0$ [38]. Putting all together we have

$$|y_I(t) - z_{I,N}(t)| \leq tm \max_j q_j \sum_{J \in G_N} |J|^\alpha \gamma \pi(J), \quad (\text{A } 1)$$

for all $I \in S_N$.

On the other hand, $|I|$ is large for any $J \in G_N$, in the sense that there exists a constant $\beta > 0$ such that $|I| \geq \beta N$ for any such I . This is because there is at least one frozen variable $A_{n+\gamma}$ that is off for every $I \in G_N$, and

$$d_\gamma N = I_{n+\gamma} + \sum_{\ell=1}^n c_{\gamma\ell} I_\ell = 0 + \sum_{\ell=1}^n c_{\gamma\ell} I_\ell \leq (\max_{\gamma\ell} c_{\gamma\ell}) |I|.$$

Therefore, $|I| \geq \beta N$ for $\beta := d_\gamma / \max_{\gamma\ell} c_{\gamma\ell}$.

If I is an absorbing state and $J \in G_N$ is any state that has an edge pointing towards I , then $J = I - s_j + t_j$ for some j , and $|J| = |I - s_j + t_j| \geq |I| - \alpha \geq \beta N - \alpha$. Therefore as $N \rightarrow \infty$, $|J|$ grows linearly for all $J \in G_N$. The right-hand side of (A 1) converges to zero since $\pi(J)$ decreases exponentially as $N \rightarrow \infty$ while $|J|^\alpha$ only grows polynomially. ■

One can now return to equation (6.3), the convergence of the solutions $x_{I,N}$ of the original master equation (6.2) towards the solution $y_I(t)$ of (6.1). Suppose that an identical initial condition $b = (b_i)$ with finite support is used for both systems.

Proposition A.2. For every fixed $t > 0$ and state I ,

$$\lim_{N \rightarrow \infty} x_{I,N} = y_I(t).$$

Proof. Once again it will be important that the reduced network (6.1) has deficiency zero and is weakly reversible. Notice that in the proof of proposition A.1 a stronger result was proved, namely that $\sum_{I \in S_N} |y_I(t) - z_I(t)| \rightarrow 0$ as $N \rightarrow \infty$. Fix $t > 0$ and $\epsilon > 0$, and choose N_1 sufficiently large that

$$\sum_{I \in S_{N_1}} y_I(t) - z_I(t) \leq \frac{\epsilon}{3} \quad \text{and} \quad \sum_{I \in S_{N_1}} y_I(t) \leq \frac{\epsilon}{3}.$$

Now we construct another FSP, that of the original master equation (6.2) with $N = N_2 > N_1$, restricted to the set S_{N_1} , and call the solution of that projection $w_I(t)$. The intuition is that as $N_2 \rightarrow \infty$, it holds $w_I(t) \approx z_I(t) \approx y_I(t)$, and at the same time one can compare $w_I(t)$ with x_{I,N_2} using FSP techniques.

In fact, both the variables $w_I(t)$ and $z_I(t)$ are defined on the finite set S_{N_1} , and as $N_2 \rightarrow \infty$ their infinitesimal generators converge. Suppose that N_2 is sufficiently large that

$$\sum_{I \in S_{N_1}} |w_I(t) - z_I(t)| \leq \frac{\epsilon}{3}.$$

Then

$$\sum_{I \in S_{N_1}} |w_I(t) - y_I(t)| \leq \sum_{I \in S_{N_1}} |w_I(t) - z_I(t)| + \sum_{I \in S_{N_1}} |z_I(t) - y_I(t)| \leq \frac{2\epsilon}{3}.$$

On the other hand,

$$\sum_{I \in S_{N_1}} w_I(t) \geq \sum_{I \in S_{N_1}} y_I(t) - \sum_{I \in S_{N_1}} |w_I(t) - y_I(t)| \geq 1 - \frac{\epsilon}{3} - \frac{2\epsilon}{3} = 1 - \epsilon.$$

Therefore for every $I \in S_{N_1}$ and every sufficiently large N_2 ,

$$\begin{aligned} |x_{I,N_2}(t) - y_I(t)| &\leq |w_I(t) - y_I(t)| + |x_{I,N_2}(t) - w_I(t)| \\ &\leq \frac{2\epsilon}{3} + \sum_{J \in S_{N_1}} x_{J,N_2}(t) - \sum_{J \in S_{N_1}} w_J(t) \leq \frac{2\epsilon}{3} + 1 - (1 - \epsilon) = \frac{5\epsilon}{3}, \end{aligned}$$

where we are using the inequality $x_{J,N_2}(t) \geq w_J(t)$ for all $J \in S_{N_1}$ given by the FSP framework. ■

References

- Alon U. 2007 *An introduction to systems biology: design principles of biological circuits*. Boca Raton, FL: Chapman & Hall/CRC.
- Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002 Stochastic gene expression in a single cell. *Science* **297**, 1183–1186. (doi:10.1126/science.1070919)
- Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. 2002 Regulation of noise in the expression of a single gene. *Nat. Genet.* **31**, 69–73. (doi:10.1038/ng869)
- Altschuler SJ, Wu LF. 2010 Cellular heterogeneity: do differences make a difference? *Cell* **141**, 559–563. (doi:10.1016/j.cell.2010.04.033)
- Bluthgen N, Legewie S. 2013 Robustness of signal transduction pathways. *Cell. Mol. Life Sci.* **70**, 2259–2269. (doi:10.1007/s00018-012-1162-7)
- Shinar G, Feinberg M. 2010 Structural sources of robustness in biochemical reaction networks. *Science* **327**, 1389–1391. (doi:10.1126/science.1183372)
- 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)
- Tian X-J, Zhang H, Xing J. 2015 Three-layer regulation leads to monoallelic olfactory receptor expression. (<http://arxiv.org/abs/1505.05179>)
- Huang HJ, Velazquez JLL. 2013 Bistable stochastic biochemical networks: large chemical networks and systems with many molecules. *J. Math. Chem.* **51**, 2074–2103. (doi:10.1007/s10910-013-0200-5)
- van Kampen NG. 2007 *Stochastic processes in physics and chemistry*, 3rd edn. Amsterdam, The Netherlands: North-Holland Personal Library.
- Lee CH. 2006 Stochastic analysis of biochemical reaction networks. Dissertation, University of Minnesota Mathematics Department.
- Wilkinson DJ. 2012 *Stochastic modelling for systems biology*. Boca Raton, FL: CRC Press.
- Anderson D, Enciso G, Johnston M. 2014 Stochastic analysis of biochemical reaction networks with absolute concentration robustness. *J. R. Soc. Interface* **11**, 20130943. (doi:10.1098/rsif.2013.0943)
- Gillespie DT. 1977 Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361. (doi:10.1021/j100540a008)
- Goulian M. 2010 Two-component signaling circuit structure and properties. *Curr. Opin. Microbiol.* **13**, 184–189. (doi:10.1016/j.mib.2010.01.009)
- Laub MT, Goulian M. 2007 Specificity in two-component signal transduction pathways. *Annu. Rev. Genet.* **41**, 121–145. (doi:10.1146/annurev.genet.41.042007.170548)
- 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)
- Maity AK, Bandyopadhyay A, Chaudhury P, Bank SK. 2014 Role of functionality in two-component signal transduction: a stochastic study. *Phys. Rev. E* **89**, 032713. (doi:10.1103/PhysRevE.89.032713)
- Shinar G, Milo R, Rodriguez Martinez M, 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)
- Straube R. 2014 Reciprocal regulation as a source of ultrasensitivity in two-component systems with bifunctional sensor kinase. *PLoS Comp. Biol.* **10**, e1003614. (doi:10.1371/journal.pcbi.1003614)
- Zhu Y, Qin L, Yoshida T, Inouye M. 2000 Phosphatase activity of histidine kinase EnvZ without kinase catalytic domain. *Proc. Natl Acad. Sci. USA* **97**, 7808–7813. (doi:10.1073/pnas.97.14.7808)
- Liggett TM. 2010 *Continuous time Markov processes*. Providence, RI: American Mathematical Society.
- Heathcote R. 2000 The mathematics of infectious diseases. *SIAM Rev.* **42**, 599–653. (doi:10.1137/S0036144500371907)
- Kuske R, Greenwood P, Gordilla L. 2007 Sustained oscillations via coherence resonance in SIR. *J. Theor. Biol.* **245**, 459–469. (doi:10.1016/j.jtbi.2006.10.029)
- Ingalls B. 2013 *Mathematical modeling in systems biology: an introduction*. Cambridge, MA: MIT Press.
- Horn F, Jackson R. 1972 General mass action kinetics. *Arch. Rat. Mech. Anal.* **47**, 187–194. (doi:10.1007/BF00251225)
- Anderson D, Craciun G, Kurtz TG. 2011 Product-form stationary distributions for deficiency zero chemical reaction networks. *Bull. Math. Biol.* **72**, 1947–1970. (doi:10.1007/s11538-010-9517-4)
- Joshi B, Shiu A. 2013 Atoms of multistationarity in chemical reaction networks. *J. Math. Chem.* **51**, 153–178. (doi:10.1007/s10910-012-0072-0)
- Munsky B. 2008 The finite state projection approach for the solution of the master equation and its applications to stochastic gene regulatory networks. Dissertation, UC Santa Barbara.
- Feinberg M. 1979 *Lectures on chemical reaction networks*. Madison, WI: The Mathematical Research Center, University of Wisconsin. See <http://www.chbmeng.ohio-state.edu/~feinberg/research/>.

31. Gunawardena J. 2003 *Chemical reaction network theory for in silico biologists*. Lecture Notes, Bauer Center for Genomics Research. Cambridge, MA: Harvard University.
32. Feinberg M. 1995 Multiple steady states for chemical reaction networks of deficiency one. *Arch. Rat. Mech. Anal.* **132**, 371–406. (doi:10.1007/BF00375615)
33. Banaji M, Pantea C. 2016 Some results on injectivity and multistationarity in chemical reaction networks. *SIAM J. App. Dyn. Syst.* **15**, 807–869. (doi:10.1137/15M1034441)
34. Joshi B, Shiu A. Which small networks are multistationary? (<http://arxiv.org/abs/1603.06441>)
35. Karp RL, Perez-Millan M, 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)
36. Dexter JP, Gunawardena J. 2013 Dimerization and bifunctionality confer robustness to the isocitrate dehydrogenase regulatory system in *Escherichia coli*. *J. Biol. Chem.* **288**, 5770–5778. (doi:10.1074/jbc.M112.339226)
37. Munsky B, Khammash M. 2008 The finite state projection approach for the analysis of stochastic noise in gene networks. *IEEE Trans. Autom. Contr.* **53**, 201–214. (doi:10.1109/TAC.2007.911361)
38. Smith H. 2014 *Monotone dynamical systems*, Mathematical Surveys and Monographs 41. Providence, RI: American Mathematical Society.