Propagation of fluctuations in biochemical systems, II: nonlinear chains

D.F. Anderson and J.C. Mattingly

Abstract: We consider biochemical reaction chains and investigate how random external fluctuations, as characterised by variance and coefficient of variation, propagate down the chains. We perform such a study under the assumption that the number of molecules is high enough so that the behaviour of the concentrations of the system is well approximated by differential equations. We conclude that the variances and coefficients of variation of the fluxes will decrease as one moves down the chain and, through an example, show that there is no corresponding result for the variances of the concentrations of the chemical species. We also prove that the fluctuations of the fluxes as characterised by their time averages decrease down reaction chains. The results presented give insight into how biochemical reaction systems are buffered against external perturbations solely by their underlying graphical structure and point out the benefits of studying the out-of-equilibrium dynamics of systems.

1 Introduction

In [1, 2] we began a study of biochemical reaction systems subjected to random, external forcing. The question we considered, and continue with here, is the following: if we add random, external forcing to the input of a biochemical reaction system, how do those fluctuations (characterised by their variance and coefficient of variation) propagate through the entire system? The broader aims of this paper are to gain a better understanding of how the network topology of biochemical reaction systems suppresses or otherwise alters the behaviour of fluctuations in reaction systems and to point out the benefits of studying the out-of-equilibrium dynamics of biochemical systems.

In [2], we studied systems under the two simplifying assumptions that the kinetics were all mass action and that each reaction involved turning precisely one species into another species. Therefore each complex consisted of a single species [3, 4]. Thus, we allowed reactions of the form $A \to B$, but not $A + B \to C$. These two assumptions caused the differential equations governing the concentrations of the species to be linear and so we referred to them as linear SSC (single species complex) systems. Considering linear SSC systems decreased some of the technical difficulties of the analysis while still allowing us to probe how different network structures affect the propagation of fluctuations. Under these assumptions we proved that the variances of fluxes decrease down reaction chains and that side reaction systems and feedback loops lower the variance of the flux out of reaction chains. A natural question is whether or not these results from [2] hold when we drop one or both of the simplifying assumptions. The main purpose of this paper is to demonstrate a biologically significant result from [2] that does hold when we drop both the SSC and mass action assumptions: the variances and coefficients of variation of fluxes decrease as one moves down a non-reversible reaction chain.

For an example of a reaction chain, consider the following biochemical system: a species, S, enters the system at a constant rate, $I > 0$. This species then combines with an enzyme, E, to form ES, which is then degraded to some product species P plus the original enzyme. Finally, the product P leaves the system. If the concentration of the enzyme is taken to be so large as to be assumed constant and the reactions are non-reversible then the following graph faithfully models our system

$$\xrightarrow{I} S \xrightarrow{F_1} ES \xrightarrow{F_2} P \xrightarrow{F_3} \tag{1}$$

where F_1, F_2 and F_3 are functions that give the rates of the respective reactions. Let s, es and p be the concentrations of S, ES and P, respectively. Then, if the kinetic functions F_1, F_2 and F_3 are functions of the reactant species only, the differential equations governing the temporal evolution of the concentrations are

$$\begin{aligned} \dot{s}(t) &= I - F_1(s(t)) \\ \dot{es}(t) &= F_1(s(t)) - F_2(es(t)) \\ \dot{p}(t) &= F_2(es(t)) - F_3(p(t)) \end{aligned} \tag{2}$$

If the functions F_i are differentiable, monotone increasing, and satisfy $F_i(0) < I < \lim_{x\to\infty} F_i(x)$, then it is easily seen that, independent of initial conditions, the system (2) will converge to the steady state $(\bar{s}, \bar{es}, \bar{p}) = (F_1^{-1}(I), F_2^{-1}(I), F_3^{-1}(I))$. However, if the input to the system (1) is allowed to fluctuate in time, then each concentration will fluctuate, and, hence, each flux, F_i, will also fluctuate. If the fluctuations are random, we can ask what the variance or coefficient of variation of each flux is with respect to that randomness, and how they relate. It is the goal of this

doi:10.1049/iet-syb:20060063

Paper first received 1st September 2006 and in revised form 12th April 2007

D.F. Anderson is with the Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706, USA

J.C. Mattingly is with the Department of Mathematics, Duke University, Durham, NC 27708, USA

E-mail: anderson@math.wisc.edu

paper to prove

$$\mathrm{Var}(F_1(s(t))) > \mathrm{Var}(F_2(es(t))) > \mathrm{Var}(F_3(p(t))) \quad (3)$$

$$CV(F_1(s(t))) > CV(F_2(es(t))) > CV(F_3(p(t))) \quad (4)$$

where $\mathrm{Var}(\cdot)$ and $CV(\cdot)$ represents variance and coefficient of variation, respectively. (Notice that the inequalities are strict.)

The reaction chain given in (1) is an example of an SSC chain because each node of the network graph consists of a single species. In general, a reaction chain is any biochemical system of the following form

$$\xrightarrow{I} C_1 \xrightarrow{F_1} C_2 \xrightarrow{F_2} \cdots \xrightarrow{F_{n-1}} C_n \xrightarrow{F_n} \quad (5)$$

where $I \in \mathbb{R}_{>0}$ is the constant input to the system, the complexes, C_i, are linear combinations of the species, and F_i: $\mathbb{R}^{m_i}_{\geq 0} \to \mathbb{R}_{\geq 0}$ are the reaction kinetics (where m_i is the number of distinct species composing complex C_i). In [2], we showed that if the constant input I is replaced by the fluctuating in time random process $I + \xi(t, \omega)$, where $\xi(t, \omega)$ is either white noise or a mean zero, finite variance, stationary stochastic process such that $\xi(t, \omega) \geq -I$, and if the system (5) is a linear SSC chain, then for all $i \geq 1$, $\mathrm{Var}(F_i) > \mathrm{Var}(F_{i+1})$, where the variance is computed according to the unique stationary measure to which the distribution of the species converges. In this paper, we prove that this result still holds when we drop the assumption that the kinetics are mass action and the assumption that each complex consists of a single species. The main assumption on the kinetics in this paper will be that they are monotone increasing in each of their dependent variables (so, for example, we may consider Michaelis–Menten kinetics). We will also show that the result still holds when the complexes are composed of multiple species so long as each species appears in precisely one complex. Throughout, we will refer to systems for which complexes can by composed of multiple species yet each species appears in a single complex as MSC (multiple species complexes) systems.

We will show that if I is the average input to a reaction chain, then (once the system has reached its statistical equilibrium) the mean of each flux is also equal to I. Therefore saying that the variances of the fluxes decrease down a reaction chain is equivalent to saying that the coefficients of variation of the fluxes decrease down a reaction chain. That is, (3) is equivalent to (4). Because of this equivalence between the magnitudes of variances and the magnitudes of coefficients of variation, each result in this paper is stated in terms of variance alone and it is understood that each result is still valid if $\mathrm{Var}(\cdot)$ is replaced with $CV(\cdot)$.

Throughout, we allow external perturbations to be white noise processes or mean zero, finite variance, stationary stochastic processes. Considering white noise processes is useful because it allows one to make use of the Itô Calculus with which stronger results (i.e. fewer restrictions on the rate functions, F_i) can be proved than if you solely consider arbitrary perturbations. Also, if the input flux to a reaction system is perturbed by white noise, then all other fluxes are perturbed by mean zero, finite variance, stationary continuous processes. Therefore one may construct continuous, stationary perturbations from white noise processes by allowing a pseudo-species to be perturbed by white noise and considering the output from the pseudo-species as the input to the reaction system of interest. In doing so, one is allowed to use the stronger white noise results as opposed to the stationary noise results. Thus, allowing both types of perturbations is quite natural. However, we point out that we do not feel the choice of external forcing is critical because the broader aim of this paper is to study the out-of-equilibrium dynamics of biochemical systems and both choices of perturbation achieve this aim.

The layout of the paper is as follows. In Section 2, we consider SSC chains. In Section 3, we consider MSC chains. Complementing the main results are two important examples. In Section 2, Example 2.3 is a nonlinear chain perturbed by white noise for which the variance (and CV) of the concentrations of the species increase down the chain. Hence, there is no corresponding 'decreasing fluctuation' result for the concentrations of reaction chains. In Section 3, Example 3.8 demonstrates that the assumption that each species is in precisely one complex is a necessary one. In both examples, we use a Monte Carlo simulation to arrive at our conclusions. The proofs of all the results in this paper are found in Section 7.

This paper is part of a larger research project in which the main biological goal is to understand how network topology affects how biochemical systems react to large-scale, random perturbations to their inputs. There are two distinct approaches we take in trying to achieve this goal. In the first, we apply random fluctuations to in silico representations of specific biological systems. We can then identify reactions, concentrations or whole subsystems that are buffered against the fluctuations, i.e. are homeostatic. We can then take the system apart piece by piece through in silico experimentation to discover the regulatory mechanisms that give rise to the homeostasis. In the second, we prove theorems about how random fluctuations propagate through relatively simple, but biologically relevant, systems. We are interested in how these systems magnify or suppress fluctuations as this may give clues as to why these systems are structured as they are. In this second approach it is the out-of-equilibrium dynamics that is being probed in order to give information on the emergent properties of the system. This paper, like [1] and [2], takes the second approach; for an example of the first, see [5].

Owing to the inherent randomness in the making and breaking of chemical bonds, biochemical reaction systems are, at their most fundamental level, modelled as jump Markov processes [6–10]. However, if one scales up the volume and number of molecules in a system while keeping the initial concentrations constant, then this intrinsic randomness becomes negligible at the scale of concentrations. One is then able to faithfully model the concentrations of the species by a system of differential equations [11]. As in [1, 2], we consider systems in this scaling limit. Thus, the random external forcing in this paper is on the scale of concentrations (and not of individual molecules) and the concentrations of the species are modelled by differential equations. For a more detailed comparison between the randomness in this paper and the inherent randomness of biochemical systems, see [2].

2 SSC chains with random perturbations

2.1 The model

A non-reversible SSC chain with a constant input is a biochemical reaction system with the following graphical structure

$$\xrightarrow{I} X_1 \xrightarrow{F_1} X_2 \xrightarrow{F_2} \cdots \xrightarrow{F_{n-1}} X_n \xrightarrow{F_n} \quad (6)$$

where $I > 0$ is the constant input to the system, X_i are the species (and complexes) of the system and F_i: $\mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$

are the reaction kinetics. For each i, let x_i represent the concentration of species X_i. The temporal evolution of $x(t)$ is governed by the following differential equation

$$\begin{aligned} \dot{x}_1 &= I - F_1(x_1) \\ \dot{x}_2 &= F_1(x_1) - F_2(x_2) \\ &\vdots \\ \dot{x}_n &= F_{n-1}(x_{n-1}) - F_n(x_n) \end{aligned} \tag{7}$$

In the sequel, we make the following standing assumptions on the functions F_i.

Assumption 2.1: Each F_i is a real valued C^1 function of $[0, \infty)$ with the following properties:

(a) $F_i(0) = 0$.
(b) For all $x \in \mathbb{R}_{>0}$, $F_i'(x) > 0$.
(c) $\lim_{x\to\infty} F_i(x) > I$.

Note that condition (c) guarantees that mass will not build up at any point along the chain so long as the input is kept at the constant value I. This assumption is also reasonable for systems for which the input is being perturbed by a mean zero random process and will be used to keep concentrations from escaping to infinity.

We will consider two different classes of random perturbations of the input I. The first will be white in time while the second will be almost surely continuous in time. Since the kinetics, F_i, are defined only on the positive portion of the real line, it is important that any noise we consider as a perturbation to the input will never drive the concentrations of the species into the negative portion of the real line. Hence we will impose restrictions on the perturbations to ensure that the concentrations stay non-negative at all times. Because we consider two different classes of perturbations of the input, we consider two different mechanisms to achieve this goal.

In the case of the white in time random perturbation, we multiply the noise term, $\mathrm{d}B_t$, by a function, $\theta_\delta(\cdot)$, that turns the noise off if x_1 approaches zero. This property of θ_δ combined with the dynamics governing the concentration of X_1 ensures that x_1 remains non-negative for all time which, in turn, keeps all other concentrations non-negative for all time. The concentration of x_1 is now governed by a stochastic differential equation, while the equations for the remaining x_i stay as in (7). That is

$$\begin{aligned} \mathrm{d}x_1 &= (I - F_1(x_1))\,\mathrm{d}t + \sigma\theta_\delta(x_1)\mathrm{d}B(t) \\ \dot{x}_2 &= F_1(x_1) - F_2(x_2) \\ &\vdots \\ \dot{x}_n &= F_{n-1}(x_{n-1}) - F_n(x_n) \end{aligned} \tag{8}$$

where $B(t) = B(t, \omega)$ is standard one-dimensional Brownian motion, $\sigma \in \mathbb{R}_{>0}$, and for some small $\delta > 0$, $\theta_\delta(x) = 1$ for all $x > \delta$, $\theta_\delta(0) = 0$ and $\theta_\delta(x)$ is C^∞ and monotone increasing. Since $F_1(0) = \theta_\delta(0) = 0$ and $I > 0$, $x_1(t) \geq 0$ for all $t > 0$ if $x_1(0) > 0$.

The second class of perturbations to the input considered in this paper are mean zero, finite variance, stationary, random processes, $\xi(t, \omega)$, which are continuous in time at almost every moment of time. To guarantee that the concentrations of the chemical species remain non-negative for all time we only consider perturbations such that $\xi(t, \omega) \geq -I$ for all t and ω (and so we no longer need the function $\theta_\delta(\cdot)$ used in the white in time setting). In order to keep the reaction system away from equilibrium, we make the added restriction that for each choice of ω, $\xi(t, \omega)$ is non-constant on all time intervals larger than some fixed value $a = a(\omega)$. We will typically write $\xi(t)$ instead of $\xi(t, \omega)$. The almost everywhere continuity of $\xi(t)$ allows the possibility of isolated jumps and allows us to use a standard differential equation for x_1 (in contrast to the Itô stochastic differential equation used in (8)). In this case, the equations governing the behaviour of the concentrations are

$$\begin{aligned} \dot{x}_1 &= I - F_1(x_1) + \xi(t) \\ \dot{x}_2 &= F_1(x_1) - F_2(x_2) \\ &\vdots \\ \dot{x}_n &= F_{n-1}(x_{n-1}) - F_n(x_n) \end{aligned} \tag{9}$$

To prove the existence of a stationary state we will further assume that the distribution of the noise's future is completely determined by its past. An example of such a $\xi(t)$ is a Markov processes whose future distribution depends only on its present value.

2.2 Decreasing variance for SSC chains

We are mainly interested in describing the system once it has settled into a statistical equilibrium and any behaviour that is transient in time has passed. Such statistical steady states are characterized by a stationary solution. A solution $x^*(t)$ is stationary if for any collection of times $t_1 < t_2 < \cdots < t_n$ and any s the distribution of the vector $(x^*(t_1 + s), x^*(t_2 + s), \ldots, x^*(t_n + s))$ is independent of s. When the forcing is Brownian as in (8), the solution is a Markov process and the distribution of $x^*(t)$ at any time t is an invariant measure for the associated Markov semigroup. An invariant measure, μ, is a measure on the state space of the system, $\mathbb{R}^n$, such that if the initial condition is chosen according to μ then solutions at any time $t \geq 0$ are also distributed as μ. More precisely, if for all measurable $A \subset \mathbb{R}^n, P(x(0) \in A) = \mu(A)$ implies that $P(x(t) \in A) = \mu(A)$ for all $t \geq 0$, then μ is invariant to the dynamics of the system. Therefore when the forcing is Brownian, a stationary solution exists. When the forcing is a stationary process $\xi(t)$ as in (9) more care must be taken to obtain a stationary solution to the dynamics as the solution need not be a Markov process.

We will concern ourselves with the existence and basic properties of stationary solutions and invariant measures at the end of this section. First we state the principal result of the article and give a few numerical examples to illustrate its use. The following theorem is proved in Section 7.2.

Theorem 2.2 (Decreasing variance down a nonlinear SSC chain): Let $x^*(t)$ be a stationary solution for the dynamics given in either equation (8) or (9). Then for all $1 \leq i \leq n$ and all t

$$\mathrm{Var}(F_i(x_i^*(t))) > \mathrm{Var}(F_{i+1}(x_{i+1}^*(t)))$$

We note that the variances of Theorem 2.2 are computed with respect to the choice of randomness, ω, in $B(t, \omega)$ or $\xi(t, \omega)$. That is, Theorem 2.2 gives the variance as an average over the choice of perturbations. In Section 2.3 we give a similar result except the variance is computed as a time average over a single path. In many natural settings (including those given in the examples below), the two notions are equivalent and therefore give the same intuition about fluctuations down reaction chains.

We now give three examples where the preceding theorem holds. For the moment we will assume that the systems possess a stationary solution to which the statistics of the solutions converge as $t \to \infty$. At the end of the section, we will prove that the preceding assumption holds for any initial condition.

Example 2.3 (Variances of concentrations need not decrease): Consider the following SSC chain with Michaelis–Menten kinetics

$$\xrightarrow{10 + \sigma\theta_\delta(x_1)\mathrm{d}B_t} X_1 \xrightarrow{F_1(x_1)} X_2 \xrightarrow{F_2(x_2)}$$

where $\sigma = 1$, $\delta = 0.001$, $F_1(x_1) = x_1$, $F_2(x_2) = (12x_2)/(1 + x_2)$. We will see in Theorem 2.6 that this system possesses a unique invariant measure to which the statistics of the trajectories converge. Using Matlab to perform a Monte Carlo simulation we computed the means, variances and coefficients of variation of the concentrations and fluxes to be the following.

	x_1	x_2	$F_1(x_1)$	$F_2(x_2)$
Mean	10	5.18	10	10
Variance	0.5	1.19	0.5	0.124
CV	0.07	0.21	0.109	0.035

As guaranteed by Theorem 2.2, the fluctuations of the fluxes decrease down the chain. However, no matter the measure we use (variance or CV), the fluctuations of x_2 are always greater than those of x_1. Therefore there is no counterpart to Theorem 2.2 pertaining to the concentrations of an SSC reaction chain.

To understand why the fluctuations of x_2 are higher than those of x_1, consider the plot of $F_2(x) = 12x/(1 + x)$ in Fig. 1. The horizontal lines at $x_1 = 9$, 10 and 11 represent possible fluxes into species X_2 and the vertical lines show what the equilibrium value of x_2 would be corresponding to that input. While the perturbed system will never settle to an equilibrium, the kinetics will always be driving the concentration of X_2 towards the solution of $12x_2(t)/(1 + x_2(t)) = x_1(t)$. Therefore minor fluctuations in the input to the species X_2 give rise to large fluctuations in x_2.

Example 2.4 (Continuous input and unbounded kinetics): Consider the following SSC chain

$$\xrightarrow{10 + \xi(t)} X_1 \xrightarrow{F_1(x_1)} X_2 \xrightarrow{F_2(x_2)}$$

where $-10 \leq \xi(t)$ is a modified Ornstein–Uhlenbeck process defined in Section 8, $F_1(x_1) = x_1^2$ and $F_2(x_2) = x_2^2/(1 + x_2)$. Because both $F_1(x)$ and $F_2(x)$ are unbounded as $x \to \infty$, we will see in Theorem 2.7 that the system possesses a unique stationary solution to which the statistics of the trajectories converge. Using Matlab to perform a Monte Carlo simulation we computed the means, variances, and coefficient of variation of the fluxes to be the following.

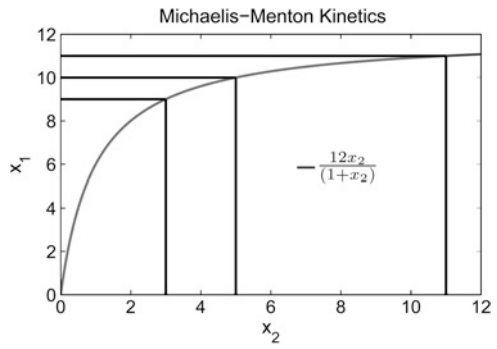

Fig. 1 *Horizontal lines represent possible inputs to species X_2 from species X_1 and the vertical lines represent the value of x_2 that would give an equilibrium to the system for a given input. Therefore we see that minor fluctuations in $F_1(x_1) = x_1$ can correspond with large fluctuations in x_2.*

	$10 + \xi(t)$	$F_1(x_1)$	$F_2(x_2)$
Mean	10	10	10
Variance	8	6.8	3.9
CV	0.28	0.26	0.20

The variances and coefficients of variation of the fluxes decrease down the chain, as guaranteed by Theorem 2.2.

Example 2.5 (Continuous input and Michaelis–Menten kinetics): Consider the following SSC chain

$$\xrightarrow{4 + \xi(t)} X_1 \xrightarrow{F_1(x_1)} X_2 \xrightarrow{F_2(x_2)}$$

where $-4 \leq \xi(t) \leq 4$ is a modified Ornstein–Uhlenbeck process defined in Section 8, $F_1(x_1) = 11x_1/(1 + x_1)$ and $F_2(x_2) = 10x_2/(1 + x_2)$. We will see in Theorem 2.8 that this system possesses a unique stationary solution to which the statistics of the trajectories converge. Using Matlab to perform a Monte Carlo simulation we computed the means, variances and coefficient of variation of the fluxes to be the following:

	$4 + \xi(t)$	$F_1(x_1)$	$F_2(x_2)$
Mean	4	4	4
Variance	4.2	3.3	2.9
CV	0.51	0.46	0.43

The variances and coefficients of variation of the fluxes decrease down the chain, as guaranteed by Theorem 2.2.

We now show the random dynamics given in (8) possesses a unique invariant measure. This invariant measure generates a stationary solution $x^*(t)$ when extended to paths from $-\infty$ to ∞. Similarly, under some additional assumptions on ξ and the F_i's, we show that the dynamics given in (9) possesses a unique stationary solution $x^*(t)$ (the concept of an invariant measure does not directly make sense for (9) since the dynamics are not necessarily Markovian, see Section 7.1). In addition, in both settings we show that the statistics of the trajectories converge to those of the stationary solution $x^*(t)$ as $t \to \infty$. In other words, for any $x(0)$ and measurable $A \subset \mathbb{R}^n$

$$P(x(t) \in A) \to P(x^*(t) \in A) = P(x^*(0) \in A), \text{ as } t \to \infty. \tag{10}$$

This means that the long time statistics of the solutions are independent of the initial condition and the result on the decrease of fluctuations is applicable on long time intervals. Of course in the setting of (8), $x^*(t)$ is distributed as the invariant measure μ so $P(x^*(0) \in A) = \mu(A)$. In the setting of (9), $x^*(t, \xi)$ should be viewed as a function of the entire past of the noise.

The next three results apply, respectively, in the three preceding examples to ensure the existence of an unique stationary solution to whose statistics the statistics of

arbitrary trajectories converge in time. The first result covers the case of white in time forcing while the second two apply to stationary forcing. The proofs of all three are contained in Section 7.3.

Theorem 2.6 (Ergodicity of the SSC chain with white noise): Equation (8) possesses a unique invariant measure, μ, on $\mathbb{R}^n$. Furthermore, the distribution of any solution to (8) converges to μ as $t \to \infty$.

To prove the existence of a stationary solution to (9), we need to assume that the distribution of the future of the noise, $\xi(t)$, is determined by its past (such as for Markov processes). This is made precise in Section 7.1. We also need additional assumptions on the functions F_i. We give two versions of these assumptions.

Theorem 2.7: Let ξ be as in Section 7.1. Under the additional assumption that the rate functions F_i are unbounded as $x \to \infty$, (9) possesses a unique stationary solution, $x^*(t)$. Furthermore, any solution $x(t)$ to (9) converges to $x^*(t)$ as $t \to \infty$.

In the event that any of the F_i are bounded, we need a bound on the size of $\xi(t)$.

Theorem 2.8: Let ξ be as in Section 7.1. Define $K = \min_i \{ \lim_{x\to\infty} F_i(x) - I \}$. Under the additional assumption that $-I \le \xi(t) \le M < K$, for all t and some $M < K$, (9) possesses a unique stationary solution, $x^*(t)$. Furthermore, any solution $x(t)$ to (9), converges to $x^*(t)$ as $t \to \infty$.

In the white in time setting, the system is in fact ergodic and hence by Birkhoff's ergodic theorem we know that for almost every realisation the time average of any statistic converges to the value of the statistic in the invariant measure. Combining this with the strong mixing properties of such a system we have that

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t (F_i(x_i(s)) - I)^2 \, \mathrm{d}s = \mathrm{Var}(F_i(x_i^*(t)))$$

for almost every realisation of the Brownian forcing and every initial condition x_0. For the above equation to hold in the setting of (9), we need to assume in addition that the stationary measure on ξ is ergodic. Even without this assumption, the next section shows that one can say something in general. This underlines the fact that the decrease of variance is really a pathwise phenomenon due to the dynamics.

2.3 Pathwise perturbations

The variance described in the previous subsection is computed with respect to the probability measure of the perturbations. More precisely, if ω is the realisation of the perturbation then $\mathrm{Var}(F_i(x(t, \omega))) = \mathbb{E}_\omega(F(x_i(t, \omega)) - I)^2$, that is, the variance is an average over the realisations of the noise. Another natural way to characterise how perturbations propagate down chains is to consider the time averages of paths. Consider again the dynamics given by (9), except now the only assumptions on $\xi(t)$ are the pathwise assumptions

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \xi(s) \, \mathrm{d}s = 0 \quad \text{and} \quad \limsup_{t\to\infty} \frac{1}{t} \int_0^t \xi(s)^2 \, \mathrm{d}s < \infty. \tag{11}$$

That is, we now assume that the time average for $\xi(t)$ converges to zero and that the time average of the square is bounded above. The following theorem states that the pathwise variances of the fluxes do not increase down reaction chains and is proved in Section 7.4.

Theorem 2.9: Consider (9) where $\xi(t)$ satisfies (11). Then for all $i \ge 1$, the following hold

1. $\lim_{t\to\infty} \frac{1}{t} \int_0^t F_i(x_i(s)) \, \mathrm{d}s = I.$
2. $\liminf_{t\to\infty} \left(\frac{1}{t} \int_0^t \xi(s)^2 \, \mathrm{d}s - \frac{1}{t} \int_0^t (F_i(x_i(s)) - I)^2 \, \mathrm{d}s \right) \ge 0.$
3. $\liminf_{t\to\infty} \left(\frac{1}{t} \int_0^t (F_i(x_i(s)) - I)^2 \, \mathrm{d}s - \frac{1}{t} \int_0^t (F_{i+1}(x_{i+1}(s)) - I)^2 \mathrm{d}s \right) \ge 0.$

3 MSC chains with random perturbations

We now consider MSC chains with random perturbations. We will again allow perturbations that are white in time or that are stationary, mean zero, finite variance and continuous for almost every t and that satisfy the conditions of Section 7.1. Consider a reaction chain (5), where each complex, C_i, consists of m_i unique species and no species is contained in more than one complex. Thus, if $x(t)$ is the vector representing the species concentrations at time t, then $x(t) \in \mathbb{R}^{m_1+\cdots+m_n}$. Let X_i^j represent the jth species in complex i and v_{ij} be the multiplicity of species X_i^j in complex i. For example, if the reaction chain is

$$\longrightarrow X_1^1 + 2X_1^2 \longrightarrow 3X_2^1 \longrightarrow$$

then $v_{11} = 1$, $v_{12} = 2$, and $v_{21} = 3$.

If F_i represents the reaction rate from complex C_i to complex C_{i+1} we have that

$$F_i(x(t)) = F_i(x_i^1, \ldots, x_i^{m_i}) \colon \mathbb{R}^{m_i} \to \mathbb{R}$$

We assume each F_i satisfies Assumption 3.1 which is analogous to Assumption 2.1.

Assumption 3.1: F_i is a real-valued C^1 function of $[0, \infty)^{m_i}$ with the following properties:

(a) If $x_i^j = 0$ for any $1 \le j \le m_i$, then $F_i(x) = 0$.
(b) If $1 \le j \le m_i$ and $x \in \mathbb{R}^{m_i}_{\ge 0}$, then $(\mathrm{d}/\mathrm{d}x_i^j) F_i(x) > 0$.
(c) $\exists M > 0$ such that if $x_i^j > M$ for all species in the ith complex, then $F_i(x) > I$.

As in the SSC case, if we want to add a random perturbation to the input flux of the system, we must only consider perturbations that will never drive concentrations into the negative portion of the real line. We handle this issue in a similar manner as in the SSC case:

If the perturbation is white in time, we multiply the perturbation by a function which will go to zero if the concentration of one of the species in the first complex goes to zero. Therefore let $\theta_\delta(x_1^1, \ldots, x_1^{m_1}) \colon \mathbb{R}^{m_1} \to \mathbb{R}_{\ge 0}$ satisfy the following three properties for some small $\delta > 0$.

1. $\theta_\delta(x) = 1$ when each $x_1^j > \delta$.
2. $\theta_\delta(x) = 0$ if $x_1^j = 0$ for any $1 \le j \le m_1$.
3. θ_δ is C^∞ and is monotone increasing in each of the variables $x_1^1, \ldots, x_1^{m_1}$.

If we add a white noise perturbation multiplied by $\theta_\delta(x)$ to the input of the system, then the dynamics are governed by the stochastic differential equation

$$dx(t) = f(x(t))\,dt + \sigma\theta_\delta(x)\,dB(t)u \tag{12}$$

where $u = [v_{11}, v_{12}, \ldots, v_{1m_1}, 0, \ldots, 0]^T$, $\sigma \in \mathbb{R}_{>0}$, $B(t)$ is standard one dimensional Brownian motion, and $f_1^i(x) = v_{1i}(I - F_1(x))$ for each $1 \le i \le m_1$, $f_2^i = v_{2i}(F_1(x) - F_2(x))$, for each $1 \le i \le m_2, \ldots, f_n^i = v_{ni}(F_{n-1}(x) - F_n(x))$ for each $1 \le i \le m_n$.

If the perturbation is a mean zero, finite variance, stationary process, $\xi(t, \omega)$, that for each ω is continuous for almost all t and that satisfies the conditions of Section 7.1, then in order to keep the concentrations non-negative, we again assume that $\xi(t, \omega) \ge -I$ for all t and ω. In this case the dynamics of the system are governed by the differential equation

$$\dot{x}(t) = f(x(t)) + \xi(t)u \tag{13}$$

where f and u are as above.

The following four theorems are analogous to those in the SSC case and their proofs can be found in Section 7.5.

Theorem 3.2: (Decreasing variance down a nonlinear MSC chain): Let $x^*(t)$ be a stationary solution for the dynamics given by either (12) or (13). Then for all $i \ge 1$ and $t \ge 0$

$$\mathrm{Var}(F_i(x^*(t))) > \mathrm{Var}(F_{i+1}(x^*(t)))$$

Theorem 3.3 (Ergodicity of the MSC chain with white noise): Equation (12) possesses a unique invariant measure, μ, on $\mathbb{R}^n$. Furthermore, the distribution of any solution to (8) converges to μ as $t \to \infty$.

Theorem 3.4: Let ξ be as in Section 7.1. Under the additional assumption that the rate functions F_i are unbounded as $x \to \infty$, (13) possesses a unique stationary solution, $x^*(t)$. Furthermore, any solution $x(t)$ to (13) converges to $x^*(t)$ as $t \to \infty$.

Theorem 3.5: Let ξ be as in Section 7.1. Define $K = \min_i \{\lim_{x\to\infty} F_i(x) - I\}$. Under the additional assumption that $-I \le \xi(t) \le M < K$, for all t and some $M < K$, (13) possesses a unique stationary solution, $x^*(t)$. Furthermore, any solution $x(t)$ to (13) converges to $x^*(t)$ as $t \to \infty$.

Example 3.6 (Sum of two species with mass action kinetics): Consider the following MSC chain with mass action kinetics

$$\xrightarrow{I(t)} Y \xrightarrow{F_1} X_1 + X_2 \xrightarrow{F_2} X_3 + X_4 \xrightarrow{F_3}$$

where $I(t) = 10 + 2\theta_\delta(y)\,dB_t$ (with $\delta = 0.001$), $F_1(y) = y$, $F_2(x_1, x_2) = x_1x_2$, $F_3(x_3, x_4) = x_3x_4$. Using Matlab to perform a Monte Carlo simulation we computed the means, variances and coefficient of variation of the fluxes to be the following

	$F_1(y)$	$F_2(x_1, x_2)$	$F_3(x_3, x_4)$
Mean	10	10	10
Variance	2	1.73	1.62
CV	0.14	0.131	0.127

We note that the variances and coefficients of variation of the fluxes decrease down the chain, as guaranteed by Theorem 3.2.

Example 3.7 (Sum of two species with Michaelis–Menten kinetics): Consider the following MSC chain with Michaelis–Menten kinetics

$$\xrightarrow{I(t)} Y \xrightarrow{F_1} X_1 + X_2 \xrightarrow{F_2} X_3 + X_4 \xrightarrow{F_3}$$

where $I(t)$ and F_1 are as in Example 3.6, and $F_2(x_1, x_2) = 14x_1x_2/[(1 + x_1)(1 + x_2)]$, $F_3(x_3, x_4) = 14x_3x_4/[(1 + x_3)(1 + x_4)]$. Using Matlab to perform a Monte Carlo simulation we computed the means, variances and coefficient of variation of the fluxes to be the following

	$F_1(y)$	$F_2(x_1, x_2)$	$F_3(x_3, x_4)$
Mean	10	10	10
Variance	2	0.72	0.49
CV	0.14	0.085	0.07

As guaranteed by Theorem 3.2 the variances and coefficients of variation of the fluxes decrease down the chain.

Example 3.8 (Species cannot be in more than one complex): Consider the following MSC chain subjected to white noise perturbations for which the species X_1 appears in two complexes (and so this system is not covered by Theorem 3.2)

$$\xrightarrow{I(t)} X_1 + X_2 \xrightarrow{F_1} X_3 \xrightarrow{F_2} X_1 + X_4 \xrightarrow{F_3}$$

where $I(t) = 10 + \theta_\delta(x_1, x_2)\,dB(t)$ (with $\delta = 0.001$), $F_1(x_1, x_2) = 2x_1x_2$, $F_2(x_3) = x_3$ and $F_3(x_1, x_4) = 5x_1x_4$. We performed a Monte Carlo simulation using Matlab to compute

	$F_1(x_1, x_2)$	$F_2(x_3)$	$F_3(x_1, x_4)$
Mean	10	10	10
Variance	4.16	0.45	1.71
CV	0.204	0.067	0.131

Note that $\mathrm{Var}(F_3(x)) > \mathrm{Var}(F_2(x))$ and $CV(F_3(x)) > CV(F_2(x))$. Therefore the assumption in Theorem 3.2 that each species is in precisely one complex is necessary.

4 Discussion

We have proved under a variety of different contexts that if the input to a non-reversible biochemical reaction chain is perturbed by a random process, then the variances and coefficients of variation of the fluxes will decrease as one moves down the chain. The assumptions made on the different choices of perturbations and on the properties of the rate functions were varied, and explicitly spelled out. Further, much care was taken to state precisely what is meant by 'fluctuations decrease down reaction chains.' Owing to this (necessary) mathematical detail, however, it is easy for the over-riding point of the paper to be lost: considering the out-of-equilibrium dynamics of a biochemical system can be an important tool for understanding the dynamical properties of that system.

A comparison of the results of this paper to metabolic control analysis (MCA) [12, 13] sheds light on the importance of considering out-of-equilibrium dynamics. The control coefficient for the flux out of a reaction chain, F,

in terms of the input, I, is

$$C_I^F = \frac{\partial F}{\partial I} \cdot \frac{I}{F}$$

where the values are computed at equilibrium. However, at equilibrium, $F = I$. Therefore independent of the choice of reaction kinetics or the length of the chain, $C_I^F = 1$. This implies that changes in the output of a chain correspond directly with changes to the input. However, by studying the out-of-equilibrium dynamics, we have shown in this paper that the fluctuations in a reaction chain will actually decrease as one moves down a reaction chain and changes to the output of a chain do not correspond directly with changes to the input. The differing results are biologically significant since it is tempting to speculate that this decrease in fluctuations (and, hence, increase in stability) is one reason long reaction chains may be evolutionarily advantageous in cellular systems.

While all of the technical details of the proofs have been relegated to the appendices, we would like to point out that to prove the main results of this paper (with the exception of Theorem 2.9) two things must be shown: (1) the existence of a unique solution whose statistics are stationary for the dynamics and (2) variances of a stationary solution decrease down reaction chains. Typically, a stationary solution can be proved to exist so long as the perturbations to the system do not drive any solutions to infinity. Stability properties of the non-perturbed system can then be used to show uniqueness of the stationary solution. The fact that the variances decrease down reaction chains follows from standard inequalities and the use of Lyapunov type functions. Intuitively, however, the variances decrease down reaction chains because the dynamics are always forcing the output flux from a complex towards the input flux. That is, the dynamics are constantly moving the system towards a shifting equilibrium. There is, however, a natural time delay in its ability to do so. Therefore the output will always be lagging behind the input, which leads to the decrease in variance.

There is still much work to be done in studying biochemical reaction systems subjected to external perturbations. A natural extension of this work and that of [2] would be to attempt to analyse reaction systems with more complicated geometries and more complicated kinetics (like product inhibition). The main technical issues encountered in such a study would be: 1) the extremely weak stability of many such systems [3, 14] would make proving the existence of a stationary solution difficult and 2) it will be difficult to isolate the variances of particular fluxes or concentrations within a complicated system. While both of these problems are formidable in a theoretical study such as in this paper, they become trivial in an in silico study [5].

5 Acknowledgments

This paper is part of a larger research program between the authors, Fred Nijhout, and Michael C. Reed, and we thank them for their input and encouragement. In particular, we thank Michael C. Reed for being the catalyst of this line of research. We also acknowledge useful discussions with Martin Hairer. D.F.A. was supported by the grants NIH R01-CA105437, NSF DMS-0109872, NSF DMS-061670 and NSF DMS-0553687. J.C.M. was partially supported by an NSF CAREER award (DMS0449910) and an Alfred P. Sloan foundation fellowship.

6 References

1 Anderson, D.F.: 'Stochastic perturbations of biochemical reaction systems'. PhD thesis, Duke University, 2005
2 Anderson, D.F., Mattingly, J.C., Nijhout, H.F., and Reed, M.C.: 'Propagation of fluctuations in biochemical systems, I: Linear SSC networks', *Bull. Math. Biol.*, 2007
3 Feinberg, M.: 'Lectures on chemical reaction networks', in 'Delivered at the Mathematics Research Center' (University of Wisconsin-Madison, 1979)
4 Horn, F.J.M., and Jackson, R.: 'General mass action kinetics', *Arch. Rat. Mech. Anal.*, 1972, **47**, pp. 81–116
5 Nijhout, H.F., Reed, M.C., Anderson, D.F., Mattingly, J.C., James, S.J., and Ulrich, C.M.: 'Long-range allosteric interactions between the folate and methionine cycles stabilize DNA methylation reaction rate', *Epigenetics*, 2006, **1**, pp. 81–87
6 Delbrück, M.: 'Statistical fluctuations in autocatalytic reactions', *J. Chem. Phys.*, 1940, **8**, pp. 120–124
7 Gadgil, C., Othmer, H., and Lee, C.H.: 'A stochastic analysis of chemical firstorder reaction networks', *Bull. Math. Biol.*, 2005, **67**, pp. 901–946
8 Gillespie, D.T.: 'A general method for numerically simulating the stochastic time evolution of coupled chemical reactions', *J. Comput. Phys.*, 1976, **22**, pp. 403–434
9 Krieger, I.M., and Gans, P.J.: 'First-order stochastic processes', *J. Chem. Phys.*, 1960, **32**, (1), pp. 247–250
10 Thattai, M., and van Oudenaarden, A.: 'Intrinsic noise in gene regulatory networks', *Proc. Natl. Acad. Sci. USA*, 2001, **98**, (15), pp. 8614–8619
11 Kurtz, T.: 'The relationship between stochastic and deterministic models for chemical reactions', *J. Chem. Phys.*, 1972, **57**, (7), pp. 2976–2978
12 Kacser, H., and Burns, J.A.: 'The control of flux', *Symp. Soc. Exper. Biol.*, 1973, **27**, pp. 65–104
13 Heinrich, R., and Rapoport, T.A.: 'A linear steady-state treatment of enzymatic chains. General properties, control and effector strength', *Eur. J. Biochem.*, 1974, **42**, pp. 89–95
14 Feinberg, M.: 'Chemical reaction network structure and the stability of complex isothermal reactors - I. the deficiency zero and deficiency one theorems: review article 25', *Chem. Eng. Sci.*, 1978, **42**, pp. 2229–2268
15 Hairer, M.: 'Ergodicity of stochastic differential equations driven by fractional Brownian motion', *Ann. Probab.*, 2005, **33**, (2), pp. 703–758
16 Billingsley, P.: 'Convergence of probability measures' (Wiley, New York, 1991)
17 Lasota, A., and Mackey, M.C.: 'Chaos, fractals, and noise: stochastic aspects of dynamics' (Springer-Verlag, New York, 1994, 2nd edn.)
18 Mattingly, J.C., Stuart, A.M., and Highman, D.J.: 'Ergodicity for sdes and approximations: locally Lipschitz vector fields and degenerate noise', *Stochastic Process. Appl.*, 2002, **101**, (2), pp. 185–232
19 Meyn, S.P., and Tweedie, R.L.: 'Markov chains and stochastic stability' (Springer-Verlag, New York, 1993)
20 Rey-Bellet, L., and Thomas, L.E.: 'Exponential convergence to nonequilibrium stationary states in classical statistical mechanics', *Comm. Math. Phys.*, 2002, **225**, (2), pp. 305–329
21 Meyn, S.P., and Tweedie, R.L.: 'Stability of Markovian processes III', *Adv. Appl. Probab.*, 1993, **25**, (3), pp. 518–548
22 Weinan, E., and Mattingly, J.C.: 'Ergodicity for the Navier–Stokes equation with degenerate random forcing: finite-dimensional approximation', *Comm. Pure Appl. Math.*, 2001, **54**, (11), pp. 1386–1402

7 Appendix A. Precise definition of the noise ξ and the proofs

7.1 Assumptions on the noise $\xi(t)$ needed for existence

In addition to the standing assumptions that $\xi(t)$ is stationary with mean zero and finite variance, to prove the existence of a stationary solution for (9), we need to assume that the distribution of $\xi(t)$ is determined entirely by the past of ξ on any interval of time $(-\infty, s]$ with $s \leq t$. Intuitively, we mean that given the value of $\xi(s)$ for $s \in (-\infty, t]$ the distribution on $[t, T]$ is uniquely determined for any $T > t$. This must be done in a way such that if one first adds a segment $[t, T - r]$ and then $[T - r, T]$, the resulting distribution on $[t, T]$ is the same as if one had added the segment $[t, T]$

in one step. For a discussion of some of the issues involved if one does not make such an assumption, see [15] and subsequent works by the author.

Let $C_{ae}((-\infty, 0], \mathbb{R})$ denote the space of almost every where continuous functions, f, endowed with the norm sup $|f(s)|e^{-\alpha|s|}$ for some $\alpha > 0$. Let $\mathcal{P}_t$ be a Markov semigroup on $C_{ae}((-\infty, 0], \mathbb{R})$ which is Feller, has an invariant measure $\mathcal{M}$, and such that for $\mathcal{M}$-almost every $\gamma_0 \in C_{ae}((-\infty, 0], \mathbb{R}), \mathcal{P}_t(\gamma_0, \cdot)$ is concentrated on elements $\gamma_t \in C_{ae}((-\infty, 0], \mathbb{R})$ with $\gamma_t(s) = \gamma_0(s+t)$ for $s \leq -t$. If γ_t, $t \geq 0$, is a realisation of the Markov chain generated with $\mathcal{P}_t$ with γ_0 distributed as $\mathcal{M}$, we define $\xi(s) = \gamma_t(s-t)$ for $s \leq t$. This is well defined since our assumptions on $\mathcal{P}_t$ make $\gamma_s(r-s) = \gamma_t(r-t)$ for $r \leq \min(s, t)$.

The dynamics of such a $\mathcal{P}_t$ can be understood as follows. Given an initial history from $-\infty$ to 0, one adds on a segment of length t, resulting in a trajectory from $-\infty$ to t. After shifting this trajectory back by $-t$, one again obtains a trajectory from $-\infty$ to 0. The distribution of this new trajectory from $-\infty$ to 0 is given by $\mathcal{P}_t$. The conditions above simply insure that the trajectory from $-\infty$ to $-t$ coincide with the initial trajectory from $-\infty$ to zero. If $\xi(t)$ is a Markov process, then it can be constructed as earlier and hence is an example of the type of noise we allow.

7.2 Proof of principal result on variances

Proof of Theorem 2.2: We consider the dynamics given by (8). The proof when the dynamics is given by (9) is identical.

Defining $\xi_1(t) = F_1(x_1^*(t)) - I$, the equations governing x_1^* and x_2^* are

$$\mathrm{d}x_1^* = (I - F_1(x_1^*))\,\mathrm{d}t + \sigma\theta_\delta(x_1^*)\,\mathrm{d}B(t) \tag{14}$$

$$\dot{x}_2^* = F_1(x_1^*) - F_2(x_2^*) \doteq I - F_2(x_2^*) + \xi_1 \tag{15}$$

We claim that for any t, $\mathbb{E}F_1(x_1^*(t)) = \mathbb{E}F_2(x_2^*(t)) = I$. Integrating (14), taking expected values, using that the distribution of $x^*(t)$ is stationary and noting that $\sigma\int_0^t \theta_\delta(x_1^*(s))\mathrm{d}B_s$ is an L^2-martingale gives

$$\mathbb{E}x_1^*(t) = \mathbb{E}x_1^*(0) + \big(I - \mathbb{E}F_1(x_1^*(t))\big)t$$

By the stationarity of the system, $\mathbb{E}x_1^*(t) = \mathbb{E}x_1^*(0)$, so $\mathbb{E}\big(F_1(x_1^*(t)) - I\big) = 0$ as claimed. A similar argument which uses $\mathbb{E}F_1(x_1^*(t)) = I$ (and, hence, $\mathbb{E}\xi_1(t) = 0$) shows that $\mathbb{E}F_2(x_2^*(t)) = I$. Therefore in order to show that $\mathrm{Var}(F_2(x_2^*(t))) < \mathrm{Var}(F_1(x_1^*(t)))$, we need $\mathbb{E}\big(F_2(x_2^*(t)) - I\big)^2 < \mathbb{E}\xi_1(t)^2$.

Let $G_1(x) = 2\int_0^x (F_2(y) - I)\,\mathrm{d}y$. Then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}G_1(x_2^*(t)) &= G_1'(x_2^*(t))\dot{x}_2^*(t)\\ &= 2(F_2(x_2^*(t)) - I)(I - F_2(x_2^*(t)) + \xi_1(t))\\ &= -2(F_2(x_2^*(t)) - I)^2 + 2(F_2(x_2^*(t)) - I)\xi_1(t)\end{aligned} \tag{16}$$

Pick $\bar{t} > 0$ arbitrarily. Integrating (16) up to time $\bar{t}$ and taking expected values gives

$$\begin{aligned}\mathbb{E}G_1(x_2^*(\bar{t})) - \mathbb{E}G_1(x_2^*(0)) &= -2\int_0^{\bar{t}} \mathbb{E}(F_2(x_2^*(s)) - I)^2\,\mathrm{d}s\\ &\quad + 2\int_0^{\bar{t}} \mathbb{E}\big[(F_2(x_2^*(s)) - I)\xi_1(s)\big]\,\mathrm{d}s\end{aligned} \tag{17}$$

Using that $x_2^*(t)$ is stationary, differentiation of (17) together with the inequality $2ab \leq a^2 + b^2$ gives

$$\begin{aligned}0 &= -2\mathbb{E}(F_2(x_2^*(\bar{t})) - I)^2 + 2\mathbb{E}\big[(F_2(x_2^*(\bar{t})) - I)\xi_1(\bar{t})\big]\\ &\leq -\mathbb{E}(F_2(x_2^*(\bar{t})) - I)^2 + \mathbb{E}\xi_1(\bar{t})^2\end{aligned} \tag{18}$$

We claim, however, that the inequality in (18) is strict. To see why, we suppose, in order to find a contradiction, that $2\mathbb{E}\big[(F_2(x_2^*(\bar{t})) - I)\xi_1(\bar{t})\big] = \mathbb{E}(F_2(x_2^*(\bar{t})) - I)^2 + \mathbb{E}\xi_1(\bar{t})^2$. Then $F_2(x_2^*(\bar{t})) - I = \xi_1(\bar{t})$ with probability one. However, this implies $F_2(x_2^*(\bar{t})) = F_1(x_1^*(\bar{t}))$ with probability one. Because $\bar{t}$ was arbitrary, we conclude that with probability one $F_2(x_2^*(t)) = F_1(x_1^*(t))$ for all t in some countably dense subset of $\mathbb{R}$. However, by the continuity of the functions involved, this implies that with probability one $F_2(x_2^*(t)) = F_1(x_1^*(t))$ for all $t \in \mathbb{R}$. Thus, $\dot{x}_2^*(t) = 0$ for all time and $x_2^*(t)$ is a constant. But, $F_2(x_2^*(t)) \equiv F_1(x_1^*(t))$ and so $x_1^*(t)$ is also a constant. However, for any $t > 0$, $P\{x_1^*(s) = \mathrm{const} : s \in [0, t]\} = 0$. Thus, the inequality in (18) is strict, which was the desired result of the Theorem. Therefore the result is shown for the first step in the chain. To complete the proof, one simply repeats the argument down the chain. □

7.3 Existence and uniqueness of stationary solutions and invariant measures

The proofs of Theorems 2.6, 2.7 and 2.8 have the same overall structure. We use the assumptions over the dynamics to obtain a uniform in time bound on some statistic of the concentration vector which can be used to prove that a sequence of time averages is tight. By extracting a convergent sub-sequence we can prove the existence of at least one invariant measure for the white in time setting. For the stationary forcing, we must work on the space to trajectories stretching back to negative infinity and prove the existence of a stationary measure on that space. We then prove that the invariant measure or stationary solution is unique and that the statistics of any solution converge to it under the dynamics of the system.

To prove the needed tightness for Theorem 2.6, we make use of the following Lyapunov function

$$V(x) = \sum_{i=1}^{n} \frac{V_i}{2}\left[\sum_{j=1}^{i}\Big(x_j - \bar{x}_j\Big)\right]^2 \tag{19}$$

where the V_i's are positive numbers yet to be determined and the $\bar{x}_j$ are defined as the solution to $F_j(\bar{x}_j) = I$ (i.e., they are the equilibrium values of the unperturbed problem). As an example, for a chain with $n = 2$ we have

$$V(x) = \frac{V_1}{2}(x_1 - \bar{x}_1)^2 + \frac{V_2}{2}\big[(x_1 - \bar{x}_1) + (x_2 - \bar{x}_2)\big]^2$$

We begin by proving a fact that, while technical, is the crux of the proof of Theorem 2.6.

Lemma 7.1: Let $\mathcal{A}$ be the generator of the SDEs (8). Then there are positive numbers $V_1, V_2, \ldots, V_n$ and positive numbers c, k such that if V(x) is defined by (19) then $\mathcal{A}V(x) \leq c - k|x|$.

Proof: For all $k \leq n$

$$\frac{\partial V}{\partial x_k} = \frac{\partial}{\partial x_k} \sum_{i=1}^{n} \frac{V_i}{2} \left[\sum_{j=1}^{i} \left(x_j - \bar{x}_j \right) \right]^2 = \sum_{i=k}^{n} V_i \left[\sum_{j=1}^{i} \left(x_j - \bar{x}_j \right) \right] \tag{20}$$

Let $F_0 = I$. Using (20), it can be shown that

$$\sum_{k=1}^{n} \frac{\partial V}{\partial x_k} (F_{k-1} - F_k) = \sum_{j=1}^{n} (x_j - \bar{x}_j) \left(\sum_{i=j}^{n} V_i (I - F_i) \right)$$

Therefore

$$\begin{aligned} \mathcal{A}V(x) &= \frac{1}{2} \sigma^2 \theta_\delta(x_1)^2 \frac{\partial^2}{\partial x_1^2} V(x) + \sum_{k=1}^{n} \frac{\partial V}{\partial x_k} (F_{k-1} - F_k) \\ &= \frac{1}{2} \left(\sigma^2 \theta_\delta(x_1)^2 \sum_{i=1}^{n} V_i \right) \\ &\quad + \sum_{j=1}^{n} (x_j - \bar{x}_j) \left(\sum_{i=j}^{n} V_i (I - F_i) \right) \\ &\doteq \frac{1}{2} \left(\sigma^2 \theta_\delta(x_1)^2 \sum_{i=1}^{n} V_i \right) + \sum_{j=1}^{n} s_j(x) \end{aligned}$$

where the last equality is a definition. We now choose the V_j's recursively. Let $V_n = 1$. Because $\lim_{x \to \infty} F_n(x) > I$, s_n is bounded by

$$s_n(x) = (x_n - \bar{x}_n)(I - F_n(x_n)) < c_n - k_n x_n$$

where c_n and k_n are some positive constants. Then s_{n-1} is given by

$$\begin{aligned} s_{n-1}(x) = (x_{n-1} - \bar{x}_{n-1}) &(V_{n-1}(I - F_{n-1}(x_{n-1})) \\ &+ (I - F_n(x_n))) \end{aligned}$$

$F_n(x_n) \geq 0$, so if $x_{n-1} \geq \bar{x}_{n-1}$, then

$$s_{n-1}(x) \leq (x_{n-1} - \bar{x}_{n-1}) \left(V_{n-1}(I - F_{n-1}(x_{n-1})) + I \right)$$

We may therefore choose V_{n-1} to be large enough so that there are positive constants c_{n-1} and k_{n-1} such that

$$s_{n-1}(x) < c_{n-1} - k_{n-1} x_{n-1}$$

Continuing up the chain, we consider s_j for $j < n$. When $x_j > \bar{x}_j$ we have

$$\begin{aligned} s_j(x) &= (x_j - \bar{x}_j) \left(V_j(I - F_j(x_j)) + \sum_{i=j+1}^{n} V_i (I - F_i) \right) \\ &\leq (x_j - \bar{x}_j) \left(V_j(I - F_j(x_j)) + I \sum_{i=j+1}^{n} V_i \right) \end{aligned}$$

Since $V_{j+1}, \ldots, V_n$ have already been defined, we may choose V_j so large that there are positive constants c_j and k_j such that

$$s_j < c_j - k_j x_j$$

Setting

$$c = \frac{1}{2} \sigma^2 \sum_{i=1}^{n} V_i + \sum_{i=1}^{n} c_i$$

we now have that for some k

$$\mathcal{A}V(x) \leq c - \sum_{i=1}^{n} k_i x_i \leq c - k|x|$$

which was the desired result. □

Proof of Theorem 2.6: The proof has two parts. First, we will use Lemma 7.1 and Prohorov's theorem [16, p. 59] to show that there exists a measure which is invariant to the stochastic flow generated by (8). We will then prove that this invariant measure is unique and that all distributions converge to it under the flow of the SDE (8).

Part 1. Let $V(x)$ be defined by (19) where $V_1, V_2, \ldots, V_n$ are given by Lemma 7.1. Then, if k, $c > 0$ are the constants given in the conclusion of Lemma 7.1

$$\mathrm{d}V(x) = \mathcal{A}V(x)\,\mathrm{d}t + \mathrm{d}M(t) \leq (c - k|x|)\,\mathrm{d}t + \mathrm{d}M(t)$$

where $M(t)$ is some L^2-martingale. Integrating gives

$$V(x(t)) \leq V(x(0)) + ct - k \int_0^t |x(s)|\,\mathrm{d}s + M(t) - M(0)$$

where $x(0)$ is some fixed value. Rearranging terms, taking expected values and using the fact that $\mathbb{E}M(t) = \mathbb{E}M(0)$ then yields

$$\frac{1}{t} \int_0^t \mathbb{E}|x(s)|\,\mathrm{d}s \leq \frac{c}{k} + \frac{V(x(0))}{kt}$$

Thus for any $R > 0$ Chebychev's inequality gives

$$\frac{1}{t} \int_0^t P\{|x(s)| > R\}\,\mathrm{d}s \leq \frac{c}{k} \frac{1}{R} + \frac{V(x(0))}{kt} \frac{1}{R} \tag{21}$$

where, again, the initial condition $x(0)$ is fixed. The right side of (21) converges to zero uniformly in $t \geq 1$ as $R \to \infty$. Therefore the sequence of measures on $\mathbb{R}$ defined by

$$\nu_n(A) \doteq \frac{1}{t_n} \int_0^{t_n} P\{x(s) \in A\}\,\mathrm{d}s$$

where $t_n \to \infty$ as $n \to \infty$ and $A \subset \mathbb{R}^n$, is tight [16, p. 59]. By Prohorov's theorem, ν_n is relatively compact and so there exists a subsequence ν_{n_k} and a measure μ, such that $\nu_{n_k} \to \mu$, where the convergence is weak convergence. Thus, for all $A \subset \mathbb{R}$

$$\mu(A) = \lim_{k \to \infty} \frac{1}{t_{n_k}} \int_0^{t_{n_k}} P\{x(s) \in A\}\,\mathrm{d}s$$

For $A \subset \mathbb{R}^n$, let $\phi_T(A) = \{x(T)\colon x(0) \in A\}$. To show that μ is invariant to the flow of (8) we need to demonstrate that for all $T > 0$ and $A \subset \mathbb{R}$, $\mu(\phi_T^{-1})(A) = \mu(A)$, where $\mu(\phi_T^{-1})(A) \doteq \mu(x\colon \phi_T(x) \in A)$. Note that, by definition, $\mu(x\colon \phi_T(x) \in A) = \lim_{k \to \infty} \frac{1}{t_{n_k}} \int_0^{t_{n_k}} P\{x(s+T) \in A\}\,\mathrm{d}s$. Using a change of variable, we then make the following

computation for any $T > 0$ and $A \subset \mathbb{R}$

$$\begin{aligned}\mu(\phi_T^{-1})(A) &= \lim_{k\to\infty} \frac{1}{t_{n_k}} \int_0^{t_{n_k}} P\{x(s+T) \in A\}\, \mathrm{d}s \\ &= \lim_{k\to\infty} \frac{1}{t_{n_k}} \int_0^{t_{n_k}} P\{x(s) \in A\}\, \mathrm{d}s \\ &\quad + \lim_{k\to\infty} \frac{1}{t_{n_k}} \int_{t_{n_k}}^{t_{n_k}+T} P\{x(s) \in A\}\, \mathrm{d}s \\ &\quad - \lim_{k\to\infty} \frac{1}{t_{n_k}} \int_0^{T} P\{x(s) \in A\}\, \mathrm{d}s \\ &= \mu(A) + \lim_{k\to\infty} \frac{1}{t_{n_k}} \left[\int_{t_{n_k}}^{t_{n_k}+T} P\{x(s) \in A\}\, \mathrm{d}s \right. \\ &\quad \left. - \int_0^{T} P\{x(s) \in A\}\, \mathrm{d}s \right]\end{aligned}$$

However

$$\left| \lim_{k\to\infty} \frac{1}{t_{n_k}} \left[\int_{t_{n_k}}^{t_{n_k}+T} P\{x(s) \in A\}\, \mathrm{d}s - \int_0^{T} P\{x(s) \in A\} \mathrm{d}s \right] \right| \le \lim_{k\to\infty} \frac{2T}{t_{n_k}} = 0$$

and so $\mu(\phi_T^{-1})(A) = \mu(A)$. Thus, μ is invariant under the stochastic flow generated by (8).

Part 2. The proof that the invariant measure is unique is not completely straightforward. The noise enters only one species, hence the diffusion is not uniformly elliptic (so arguments such as in [17] do not suffice). The proof we now sketch follows a rather standard line of argument. We refer the reader to [18, 19, 20] for the missing details. The proof has three elements. First, one shows that the generator of the diffusion satisfies Höormander's 'sum of squares' theorem and hence is hypoelliptic. This ensures that the Markov transition density $p_t(x, y)$ is smooth in x and y and hence is a Strong Feller process. This gives the local smoothing needed to ensure that the invariant measure found above is unique. The structure of (8) and the fact that the F_i' do not vanish ensures that the span of the needed Lie brackets is of full dimension. Hence Höormander's theorem holds.

Secondly, we need to provide the global information which ensures open set irreducibility (the fact that processes starting from different initial points have nonzero probability of entering a small neighborhood of each other). The Lyapunov function given by (19) shows that the processes return to a bounded ball $\mathcal{B}$ about the origin eventually. Since there is a globally attracting fix point, if the noise is small for long enough all of the points of $\mathcal{B}$ will enter an arbitrarily small neighbourhood of the fixed point.

Finally, with the above facts in hand, the uniqueness and convergence result follows from standard arguments [18, 20–22]. □

Proof of Theorem 2.7: As in the proof of Theorem 2.6, the proof is split into two parts. In the first we prove the existence of a stationary solution $x^*(t)$ for the dynamics (9). In the second we show that the if $x(t)$ and $y(t)$ are solutions driven by the same noise, then $y(t) \to x(t)$ pathwise. Hence, we conclude there can only be one stationary solution since any two would converge to each other over time.

Part 1. Unlike the previous example, the process $x(t)$ alone is not a Markov process. However, if we include the entire history of ξ then the system does become Markovian. More precisely, from the assumptions in Section 7.1 we know that $\xi(s) = \gamma_t(s-t)$ for $s \le t$ where γ_t is a Feller Markov process on $C_{\mathrm{ae}}((-\infty, 0], \mathbb{R})$ with semi-group $\mathcal{P}_t$ and with γ_0 distributed as the invariant measure $\mathcal{M}$. Then the pair $(x(t), \gamma_t)$ is a Markov process on the expanded state space $\mathbb{R}^n \times C_{\mathrm{ae}}((-\infty, 0], \mathbb{R})$. Let $\hat{\mathcal{P}}_t$ denote the Markov transition semi-group of this system and π_x and π_γ the projection onto the x and γ coordinates, respectively. Since we start γ_t from an invariant measure for its dynamics, we know that the statistics of γ_t are constant in time equal to $\mathcal{M}$ for all t and hence is tight.

Let $x^{(0)}$ be an arbitrary initial condition for $x(t)$. Defining the measure

$$Q_t(\cdot) = \frac{1}{t} \int_0^t \int \hat{\mathcal{P}}_s(x^{(0)}, \gamma, \cdot)\, \mathcal{M}(\mathrm{d}\gamma)\, \mathrm{d}s$$

we need only shows that $Q_t \pi_x^{-1}$ is tight to conclude that Q_t is tight since $Q_t \pi_\gamma^{-1} = \mathcal{M}$ is independent for t. We will do this coordinate by coordinate. Consider the equation governing $x_1(t)$

$$\dot{x}_1(t) = I - F_1(x_1(t)) + \xi(t) \tag{22}$$

Integrating (22) gives

$$\begin{aligned}x_1(t) &= x(0) + It - \int_0^t F_1(x_1(s))\, \mathrm{d}s + \int_0^t \xi(s)\, \mathrm{d}s \\ &\le x(0) + It - \int_0^t F_1(x_1(s)) 1_{\{|x_1(s)|>R\}}\, \mathrm{d}s + \int_0^t \xi(s)\, \mathrm{d}s \\ &\le x(0) + It - F_1(R) \int_0^t 1_{\{|x_1(s)|>R\}}\, \mathrm{d}s + \int_0^t \xi(s)\, \mathrm{d}s\end{aligned} \tag{23}$$

Taking expected values and rearranging terms gives

$$\frac{1}{t} \int_0^t P\{|x_1(s)| > R\}\, \mathrm{d}s \le \frac{I}{F_1(R)} + \frac{\mathbb{E}x_1(0)}{F_1(R)} \tag{24}$$

Note that rearranging (23) and taking expected values gives us the additional bound

$$\frac{1}{t} \int_0^t \mathbb{E}F_1(x_1(s))\, \mathrm{d}s \le I + \frac{\mathbb{E}x_1(0)}{t} \tag{25}$$

Continuing down the chain we consider $x_2(t)$

$$\dot{x}_2(t) = F_1(x_1(t)) - F_2(x_2(t))$$

Integrating gives

$$\begin{aligned}x_2(t) &= x_2(0) + \int_0^t F_1(x_1(s))\, \mathrm{d}s - \int_0^t F_2(x_2(s))\, \mathrm{d}s \\ &\le x_2(0) + \int_0^t F_1(x_1(s))\, \mathrm{d}s - F_2(R) \int_0^t 1_{\{|x_2(s)|>R\}}\, \mathrm{d}s\end{aligned} \tag{26}$$

Rearranging terms as before, taking expected values and using (25) gives

$$\frac{1}{t} \int_0^t P\{|x_2(s)| > R\}\, \mathrm{d}s \le \frac{I}{F_2(R)} + \frac{\mathbb{E}x_1(0)}{tF_2(R)} + \frac{\mathbb{E}x_2(0)}{tF_2(R)}$$

Further, rearranging (26) and using (25) gives

$$\frac{1}{t} \int_0^t \mathbb{E}F_2(x_2(s))\, \mathrm{d}s \le I + \frac{\mathbb{E}x_1(0)}{t} + \frac{\mathbb{E}x_2(0)}{t}$$

We may continue down the chain in a similar manner and conclude that there are positive constants $c_1, c_2, \ldots, c_n$ such that for all $t \geq 1$

$$\begin{aligned} Q_t \pi_x^{-1}(\{y\colon \sup_i |y_i| > R\}) &= \frac{1}{t}\int_0^t P\{\sup_i |x_i(s)| > R\}\,\mathrm{d}s \\ &< \sum_i \frac{1}{t}\int_0^t P\{|x_i(s)| > R\}\,\mathrm{d}s \\ &< \sum_i \frac{c_i}{F_i(R)} \end{aligned} \tag{27}$$

Because each F_i is monotone and unbounded, the right side of inequality (27) converges to zero uniformly in t as $R \to \infty$. Therefore just as in the proof of Theorem 2.6, we may invoke Prohorov's Theorem to guarantee the existence of a measure μ on $\mathbb{R}^n \times C_{\mathrm{ae}}((-\infty, 0], \mathbb{R})$ that is invariant to the dynamics induced by $\hat{\mathcal{P}}_t$. By Kolmogormov's extension theorem we can use this measure to define a measure on pairs of noise ξ and solution trajectories x starting at $-\infty$ and continuing to ∞. The projection of this measure onto the solution coordinate produces a stationary solution for the $x(t)$ dynamics. One should really view this stationary solution x^* along with its noise trajectory ξ which was constructed along with it.

Part 2. Let $x^*(t) = x^*(t, \xi)$ be the stationary solution and matching noise trajectory found above. Let $y(t)$ be the solution starting from an arbitrary initial condition $y(0)$ using the same noise $\xi(t)$.

Consider $x_1^*(t)$ and $y_1(t)$. If $x_1^*(0) = y_1(0)$, then $x_1^*(t) = y_1(t)$ for all time by uniqueness of solutions. Suppose that $x_1^*(0) > y_1(0)$ (if $x_1^*(0) < y_1(0)$ there is a symmetric argument). Then, $x_1^*(t) > y_1(t)$ for all time. Differentiating $x_1^*(t) - y_1(t)$ gives

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(x_1^*(t) - y_1(t)) &= -(F_1(x_1^*(t)) - F_1(y_1(t))) \\ &= -(x_1^*(t) - y_1(t)) \\ &\quad \times \frac{F_1(x_1^*(t)) - F_1(y_1(t))}{x_1^*(t) - y_1(t)} \end{aligned}$$

Defining

$$H_1(t) = \frac{F_1(x_1^*(t)) - F_1(y_1(t))}{x_1^*(t) - y_1(t)} \tag{28}$$

we have that

$$x_1^*(t) - y_1(t) = (x_1^*(0) - y_1(0))\mathrm{e}^{-\int_0^t H_1(s)\,\mathrm{d}s}$$

By the uniform bound given in (27), we know that both $x_1^*(t)$ and $y_1(t)$ spend a positive fraction of time in a compact set on which $H_1(t) > d_1 > 0$ for some $d_1 > 0$ (since H_1 is an approximation to the derivative of F_1). Thus, $x_1^*(t) - y_1(t) \to 0$, as $t \to \infty$.

We next consider $x_2^*(t)$ and $y_2(t)$. Suppose that $x_2^*(0) < y_2(0)$. Let τ_2 be the first time $x_2^*(t) = y_2(t)$. Then, up until time τ_2

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(y_2(t) - x_2^*(t)) &= -(F_1(x_1^*) - F_1(y_1)) \\ &\quad - (F_2(y_2(t)) - F_2(x_2^*(t))) \end{aligned}$$

But, $x_1^* \geq y_1$ so $(F_1(x_1^*) - F_1(y_1)) \geq 0$ and

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(y_2(t) - x_2^*(t)) &\leq -(F_2(y_2(t)) - F_2(x_2^*(t))) \\ &= -(y_2(t) - x_2^*(t)) \\ &\quad \times \frac{F_2(y_2(t)) - F_2(x_2^*(t))}{y_2(t) - x_2^*(t)} \end{aligned}$$

Defining $H_2(t)$ as we did $H_1(t)$ we may conclude from the above that up until time τ_2

$$y_2(t) - x_2^*(t) < (y_2(0) - x_2^*(0))\mathrm{e}^{-\int_0^t H_2(r)\,\mathrm{d}r}$$

Therefore if $\tau_2 = \infty$, then, as in the previous case, we may use the above equation and the bound (27) to conclude that $|x_2^*(t) - y_2(t)| \to 0$, which is the desired result.

If τ_2 is finite, then for all time after τ_2, $x_2^*(t) \geq y_2(t)$. To see this note that if $x_2^*(t) = y_2(t)$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}(x_2^*(t) - y_2(t)) = F_1(x_1^*) - F_1(y_1) \geq 0$$

where the inequality follows since $x_1^* \geq y_1$. Thus, we consider times past τ_2 and redefine our initial condition to be the values $x(\tau_2)$ and $y(\tau_2)$.

We note

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(x_2^*(t) - y_2(t)) &= -\big(F_2(x_2^*) - F_2(y_2)\big) + F_1(x_1^*) - F_1(y_1) \\ &= -(x_2^* - y_2)\frac{F_2(x_2^*) - F_2(y_2)}{x_2^* - y_2} \\ &\quad + F_1(x_1^*) - F_1(y_1) \end{aligned} \tag{29}$$

To gain control over the term $F_1(x_1^*) - F_1(y_1)$ we use the equations governing $x_1{}^*$ and y_1

$$x_1^*(t) - y_1(t) = x_1^*(0) - y_1(0) + \int_0^t (F_1(x_1^*(x)) - F_1(y_1(s)))\,\mathrm{d}s \tag{30}$$

Rearranging and using that $x_1^* \geq y_1$, we have

$$\int_0^t F_1(x_1^*(s)) - F_1(y_1(s))\,\mathrm{d}s \leq x_1^*(0) - y_1(0) \tag{31}$$

Thus, if $\eta_1(t) = \int_0^t F_1(x_1^*(s)) - F_1(y_1(s))\,\mathrm{d}s$, we have that $F_1(x_1^*(t)) - F_1(y_1(t)) = \eta_1'(t)$, and that for all t, $\eta_1(t) < x_1^*(0) - y_1(0)$. Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}(x_2^*(t) - y_2(t)) = -(x_2^* - y_2)H_2(t) + \eta_1'(t)$$

and integrating by parts gives

$$\begin{aligned} x_2^*(t) - y_2(t) &= (x_2^*(0) - y_2(0)\mathrm{e}^{-\int_0^t H_2(r)\,\mathrm{d}r} \\ &\quad + \int_0^t \eta_1(s)\mathrm{e}^{-\int_s^t H_2(r)\mathrm{d}r}\,\mathrm{d}s \\ &= (x_2^*(0) - y_2(0)\mathrm{e}^{-\int_0^t H_2(s)\,\mathrm{d}s} \\ &\quad + \eta_1(t) - \eta_1(0)\mathrm{e}^{-\int_0^t H_2(r)\,\mathrm{d}r} \\ &\quad - \int_0^t \eta_1(s)\mathrm{e}^{-\int_s^t H_2(r)\,\mathrm{d}r} H_2(s)\,\mathrm{d}s \end{aligned}$$

The last two terms are negative and, as before, the exponential terms go to zero as $t \to \infty$. So,

$\lim_{t\to\infty}|x_2^*(t)-y_2(t)|\le\eta_1(t)\le x_1^*(0)-y_1(0)$. However, we can re-scale time (do the above analysis on the interval $[t/2,\ t]$ instead of $[0,\ t]$) to conclude that $\lim_{t\to\infty}|x_2^*(t)-y_2(t)|\le\lim_{t\to\infty}\eta_1(t)\le\lim_{t\to\infty}(x_1^*(t/2)-y_1(t/2))=0$.

Now we continue down the chain in a similar manner and consider x_3 and y_3. Without loss of generality we may assume $x_2^*(t)>y_2(t)$ for all time. If $x_3(t)<y_3(t)$ for all time, we do the same argument as above to conclude that $|x_3(t)-y_3(t)|\to 0$, as $t\to\infty$. Thus, we assume $x_3(t)>y_3(t)$ for all time. The argument is the same as that above, except we should now have control over $F_2(x_2^*(s))-F_2(y_2(s))$. We have

$$\begin{aligned}x_2^*(t)-y_2(t)&=x_2^*(0)-y_2(0)\\&\quad+\int_0^t\big(F_1(x_1^*(s))-F_1(y_1(s))\big)\,\mathrm{d}s\\&\quad-\int_0^t\big(F_2(x_2^*(s))-F_2(y_2(s))\big)\,\mathrm{d}s\end{aligned}$$

Rearranging gives

$$\int_0^t\big(F_2(x_2^*(s))-F_2(y_2(s))\big)\mathrm{d}s=x_2^*(0)-y_2(0)+\eta_1(t)$$

We may define the above integral to be $\eta_2\ (t)$ and perform the same analysis as before. In this way, we continue down the chain and conclude that $\lim_{t\to\infty}|x(t)-y(t)|=0$, which was the desired result. □

Proof of Theorem 2.8: Let $N_1^\varepsilon=F_1^{-1}(I+M+\varepsilon)$, where $\varepsilon<K-M$. If $x_1(t)>N_1^\varepsilon$, then by the monotonicity of F_1 we have

$$\begin{aligned}\dot{x}_1(t)&=I-F(x_1(t))+\xi(t)\\&\le I-F(N_1^\varepsilon)+\xi(t)\\&=-M-\varepsilon+\xi(t)\\&<-\varepsilon\end{aligned}$$

Therefore independent of initial conditions, $x_1(t)\le N_1^\varepsilon$ for all t large enough. However, ε was arbitrary, so $\limsup_{t\to\infty}x_1(t)\le F_1^{-1}(I+M)$. Continuing in this manner down the chain shows $\limsup_{t\to\infty}x_i(t)\le F_i^{-1}(I+M)$, for each i. Thus, for large t, there exists $L>0$ such that $\mathbb{E}|x(t)|<L$. By Chebychev's inequality we then have

$$\frac{1}{t}\int_0^t P\{|x_1(s)|>R\}\mathrm{d}s\le\frac{L}{R}$$

which converges to zero uniformly in t as $R\to\infty$. As in the proof of Theorem 2.7, we need to consider the Markov process on the extended state space $\mathbb{R}^n\times C_{ae}((-\infty,0],\mathbb{R})$. As before, we obtain tightness by using the above estimates on the marginal of this measure in the $x(t)$ variable since the $\xi\ (t)$ variable is stationary and hence already tight. We may again use Prohorov's Theorem to guarantee the existence of an invariant measure. The proof of uniqueness is the same is in the proof of Theorem 2.7. □

7.4 Proof of Theorem 2.9

We begin by showing that $\lim_{t\to\infty}x_1(t)/t=0$. Consider the dynamics governing x_1 where $\xi\ (t)$ satisfies (11)

$$\dot{x}_1=I-F_1(x_1)+\xi(t)\tag{32}$$

Let $\bar{x}_1=F^{-1}(I)$. Then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(x_1(t)-\bar{x}_1)&=I-F_1(x_1(t))+\xi(t)\\&=-\frac{I-F_1(x_1(t))}{\bar{x}_1-x_1(t)}(x_1(t)-\bar{x}_1)+\xi(t)\end{aligned}$$

Setting $H(t)=(I-F_1(x_1(t)))/(\bar{x}_1-x_1(t))>0$ (which is well defined since F_1 is assumed differentiable and is positive by the monotonicity of F_1) and using Duhamel's formula gives us

$$(x_1(t)-\bar{x}_1)=(x_1(0)-\bar{x}_1)\mathrm{e}^{-\int_0^t H(s)\,\mathrm{d}s}+\int_0^t\mathrm{e}^{-\int_s^t H(r)\,\mathrm{d}r}\xi(s)\,\mathrm{d}s$$

Integrating by parts gives

$$\begin{aligned}(x_1(t)-\bar{x}_1)&=(x_1(0)-\bar{x}_1)\mathrm{e}^{-\int_0^t H(s)\,\mathrm{d}s}+\int_0^t\xi(s)\,\mathrm{d}s\\&\quad+\int_0^t\mathrm{e}^{-\int_s^t H(r)\,\mathrm{d}r}H(s)\left(\int_0^s\xi(r)\mathrm{d}r\right)\mathrm{d}s\end{aligned}$$

By the positivity of $H(t)$ and property (11), we then have

$$\lim_{t\to\infty}\frac{x_1(t)}{t}=\lim_{t\to\infty}\frac{1}{t}\int_0^t\mathrm{e}^{-\int_s^t H(r)\mathrm{d}r}H(s)\left(\int_0^s\xi(r)\,\mathrm{d}r\right)\mathrm{d}s$$

Let $\varepsilon>0$. There exists an $S>0$ such that $s>S$ implies $|(1/s)\int_0^s\xi(r)\,\mathrm{d}r|<\varepsilon/2$. There exists a $T=T(s)>0$ such that $t>T$ implies $\sup_{s<S}|(1/t)\int_0^s\xi(r)\,\mathrm{d}r|<\varepsilon/2$. Therefore, if $t>\max\{S,\ T\}$ we have that

$$\begin{aligned}\left|\frac{1}{t}\int_0^t\mathrm{e}^{-\int_s^t H(r)\,\mathrm{d}r}H(s)\left(\int_0^s\xi(r)\,\mathrm{d}r\right)\mathrm{d}s\right|&\le\int_0^t\mathrm{e}^{-\int_s^t H(r)\,\mathrm{d}r}H(s)\\&\quad\times\left|\frac{1}{t}\int_0^s\xi(r)\,\mathrm{d}r\right|(1_{\{s\le S\}}+1_{\{s>S\}})\,\mathrm{d}s\\&\le\varepsilon\int_0^t\mathrm{e}^{-\int_s^t H(r)\,\mathrm{d}r}H(s)\,\mathrm{d}s\le\varepsilon\end{aligned}$$

Thus, $\lim_{t\to\infty}x_1(t)/t=0$.

Integrating (32), dividing by t and taking the limit as $t\to\infty$ now gives us

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t F(x(s))\,\mathrm{d}s=I,$$

which proves Part 1 of Theorem 2.9 for x_1.

Let $G(x)=2\int_0^x(F_1(y)-I)\,\mathrm{d}y$, which, non-coincidentally, is the same function used in the proof of Theorem 2. We have

$$\frac{\mathrm{d}}{\mathrm{d}t}G(x_1(t))=-2(F_1(x_1(t))-I)^2+2(F_1(x_1(t))-I)\xi(t)$$

Integrating and using the inequality $ab\le(1/2)a^2+(1/2)b^2$ gives

$$G(x_1(t))\le G(x_1(0))-\int_0^t(F_1(x_1(s))-I)^2\mathrm{d}s+\int_0^t\xi(s)^2\,\mathrm{d}s$$

Therefore, Part 2 of Theorem 2.9 will be shown for x_1 if $\liminf_{t\to\infty}G(x_1(t))/t\ge 0$. We have

$$\begin{aligned}\liminf_{t\to\infty}\frac{1}{t}G(x_1(t)) &= 2\liminf_{t\to\infty}\frac{1}{t}\int_0^{x_1(t)}(F_1(y)-I)\\ &\quad\times(1_{\{y>\bar{x}_1\}}+1_{\{y\le\bar{x}_1\}})\,\mathrm{d}y\\ &\ge 2\liminf_{t\to\infty}\frac{1}{t}\int_0^{x_1(t)}(F_1(y)-I)1_{\{y\le\bar{x}_1\}}\,\mathrm{d}y\\ &\ge -2I\lim_{t\to\infty}\frac{x_1(t)}{t}=0\end{aligned}$$

so Part 2 is shown for x_1. Note that Parts 1 and 2 of Theorem 2.9 show that $F_1(x_1(t))-I$ satisfy condition (11). Therefore to prove Parts 1, 2 and 3 for all x_i, one simply continues down the chain considering $F_i(x_i(t))$ as the external perturbation of x_{i+1}.

7.5 Proofs of Section 3

The proofs of Theorems 3.2, 3.3, 3.4 and 3.5 can be handled simultaneously.

Proof of Theorems 3.2, 3.3, 3.4 and 3.5: The key to each proof is the recognition that the species in each complex satisfy constant multiples of the same (stochastic) differential equations. Mathematically, this means these species can by grouped and treated as a single species with a redefined kinetics. This reduces us to the case previously studied. More explicitly, there are constants c_{ijk} and d_{ijk} such that $x_i^j(t)=d_{ijk}x_i^k(t)+c_{ijk}$ for all t. Thus, the species of each complex can be solved from knowledge of just one species from that complex. Further, a monotone increase in one translates to a monotone increase in the others. Choosing one species, y_i, from each complex, we may redefine the F_i's (and θ_δ in the white noise case) appropriately so that the vector function $y(t)$ satisfies either (8) or (9) with the F_i's satisfying Assumption 2.1. Therefore applying the theorems of Section 2 completes the proof. □

8 Processes used in examples 2.4 and 2.5

In Example 2.4, $\xi(t)$ is described as a modified Ornstein–Uhlenbeck process such that $-10\le\xi(t)$. More precisely, $\xi(t)$ is governed by the following dynamics

$$d\xi(t)=\begin{cases}-\xi(t)\,\mathrm{d}t+4\mathrm{d}B(t) & \text{if } \xi(t)>-10\\ -\xi(t)\,\mathrm{d}t & \text{if } \xi(t)\le -10\end{cases}$$

This dynamics ensures that if $\xi(0)>-10$ then $\xi(t)\ge-10$ for all t.

In Example 2.5, $\xi(t)$ is built from the Ornstein–Uhlenbeck equation $\mathrm{d}\xi(t)=-\xi(t)\,\mathrm{d}t+3\mathrm{d}B(t)$, with the added condition that if $\xi(t)=-4$ or $\xi(t)=4$, then $\mathrm{d}\xi(t)=-\xi(t)\,\mathrm{d}t$. More precisely, $\xi(t)$ is governed by the following dynamics

$$\mathrm{d}\xi(t)=\begin{cases}-\xi(t)\,\mathrm{d}t & \text{if } \xi(t)\,\mathrm{d}t\ge 4\\ -\xi(t)\,\mathrm{d}t+3\mathrm{d}B(t) & \text{if } -4<\xi(t)<4\\ -\xi(t)\,\mathrm{d}t & \text{if } \xi(t)\,\mathrm{d}t\le -4\end{cases}$$