UNIVERSITY OF
BATH

University of Bath

Stochastic pattern formation and spontaneous polarisation: the linear noise approximation and beyond

Alan J. McKane[1], Tommaso Biancalani[1] and Tim Rogers[1,2]

[1]Theoretical Physics Division, School of Physics and Astronomy, University of Manchester, Manchester M13 9PL, United Kingdom

[2]Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom

Abstract

We review the mathematical formalism underlying the modelling of stochasticity in biological systems. Beginning with a description of the system in terms of its basic constituents, we derive the mesoscopic equations governing the dynamics which generalise the more familiar macroscopic equations. We apply this formalism to the analysis of two specific noise-induced phenomena observed in biologically-inspired models. In the first example, we show how the stochastic amplification of a Turing instability gives rise to spatial and temporal patterns which may be understood within the linear noise approximation. The second example concerns the spontaneous emergence of cell polarity, where we make analytic progress by exploiting a separation of time-scales.

1. Introduction

The evidence from the talks delivered at this meeting should leave little doubt that stochastic models are required to understand a wide range of biological phenomena. The models do need to be carefully formulated, with the nature of the interactions between the constituents clearly specified, but when this is carried out in an unambiguous fashion, their analysis can begin. In the great majority of cases this will be numerical, and many of the contributions at this meeting were concerned with the development and implementation of efficient algorithms for carrying this analysis out.

In this paper we will take a different path. We will similarly take care in formulating the precise form of the model, stressing the need to begin from a microscopic individual-based model (IBM). However we will be primarily be interested in using mathematical analysis to understand the model. To facilitate this, the interactions between the constituents of the model will be represented by chemical-like reactions occurring at specified rates (Black and McKane, 2012). If the underlying stochastic process is Markov, which will usually be the case, a master equation can be written down which will describe the time evolution of the probability of the system being in a given state (van Kampen, 2007).

One analytic procedure that can always be carried out on the master equation is to calculate the equations describing the time evolution of the averages of the state variables. These deterministic equations give a macroscopic description of the dynamics of the system. They are the other major methodology for the theoretical study of biological systems, and their use is exemplified by the book by Murray (2008). This older tradition involved both the study of simple, analytically tractable, models and dynamical systems theory. The former was concerned with the mathematical investigation of specific differential equations of few variables and the latter with general results on stability of attractors, topological notions, bifurcation theory, and so on (Wiggins, 2003).

A parallel methodology for the mathematical analysis of the full stochastic model also exists, however it is much less widely appreciated than that for the corresponding deterministic analysis. Stochastic differential equations (SDEs) can be derived when the number of constituents is large (but not infinite); the stochasticity originating from the discreteness of the underlying IBM. Techniques from the theory of stochastic processes, as well as general results from this theory, can be used to understand these equations analytically, just as in the deterministic case.

Here we will describe this methodology with reference to two specific models in order to illustrate the basic ideas and techniques. We will begin in section 2 with the definition of the state variables of the model, and the specification of the interactions between them in terms of chemical reactions. This allows us to write down the master equation for the process. In section 3 we derive the macroscopic and mesoscopic equations governing the dynamics of the process from this master equation. We then go on to illustrate the use of this formalism, within the linear noise approximation (LNA), to the particular example of the Brusselator model in section 4. However, the LNA is not always able to capture the details of stochastic ordering phenomena, and in section 5 we show that, in some situations, we are able to go beyond the LNA. This is illustrated on the spontaneous emergence of cell polarity; this example being motivated by the talk given by Linda Petzold at the meeting. Finally we conclude with an overview of the techniques and applications that we have discussed in section 6.

2. Definition of the models and the master equation

The stochastic effects that will interest us here occur when the number of constituents, N, is large but finite. In this case a *mesoscopic* description is required: the state variables are assumed continuous — unlike in the microscopic IBM description — but the effect of the discreteness is not lost: it is manifested in the form of Gaussian distributed 'noise'. The mesoscopic form of the model is not obvious; we have to begin from a microscopic IBM and derive it as an approximation which holds when N is large. It cannot be found from a macroscopic description, since this consists of time evolution equations for average quantities, whereas the micro- or mesoscopic formulations describe the time evolution of the entire probability density function (pdf). But many stochastic processes have the same average, so there is in principle no unique way of determining the micro- or mesoscopic model from the macroscopic one. This will be especially clear later, when we derive the mesoscopic form of the equations and compare them to their macroscopic counterparts.

To set up the IBM we first need to decide what are the variables which describe the state of the system. For the simplest case of a single type of constituent with no spatial, class, or other structure, it would be a single integer $n = 1, 2, \ldots, N$ representing the number of constituents in the system. In ecological models this could be the number of individuals in the population. There would then be a transition rate from state n to state n' caused by births, deaths, competition, predation, etc. This rate will be denoted by $T(n'|n)$, with the initial state on the left (the other convention with the initial state on the right is also sometimes used).

The probability of finding the system in state n at time t, $P_n(t)$, changes according to the master equation:

$$\frac{\mathrm{d}P_n(t)}{\mathrm{d}t} = \sum_{n' \neq n} T(n|n') \, P_{n'}(t) - \sum_{n' \neq n} T(n'|n) \, P_n(t). \tag{1}$$

The first term on the right-hand side describes the rate at which $P_n(t)$ increases due to transitions into the state n from all other states n' while the second term on the right-hand side describes the rate at which $P_n(t)$ decreases due to transitions out of the state n to all other states n'. The net change then gives $dP_n(t)/dt$. Although it is intuitively clear, it can also be derived from the Chapman-Kolmogorov equation for Markov processes (Gardiner, 2009). It gives the 'microscopic' description of the system and will be the starting point for deriving the meso- and macroscopic descriptions.

All this generalises to several types of constituents with numbers $\ell, m, \ldots$ at a given time. Spatial structure can also be introduced with constituents located in a particular small volume $j = 1, 2 \ldots$ at a given time. For notational simplicity we will combine the index i labelling these small volumes with an index s which labels the types or classes of constituent, into one label $J = \{j, s\}$. Later, when carrying out the analysis of the differential equations we will separate the indices, and may also take the continuum limit in which volume labels become continuous. An additional comment is worth making at this stage. There is no agreed nomenclature for the small volumes: describing them as 'cells' is potentially confusing in a biochemical context and the term 'patches' is usually only used in an ecological context when the constituents are individuals. We will use the neutral term 'domain' and talk about their 'volume' V, even when they are one- or two-dimensional, as will be the case with the models we discuss in this paper.

If we now write n_I for the number of constituents of a particular type in a particular domain, we can specify the state of the system through the vector of integers $\boldsymbol{n} = (n_1, n_2, \ldots)$. The master equation is then simply Eq. (1) with n replaced by $\boldsymbol{n}$:

$$\frac{\mathrm{d}P_{\boldsymbol{n}}(t)}{\mathrm{d}t} = \sum_{\boldsymbol{n'} \neq \boldsymbol{n}} T(\boldsymbol{n}|\boldsymbol{n'}) P_{\boldsymbol{n'}}(t) - \sum_{\boldsymbol{n'} \neq \boldsymbol{n}} T(\boldsymbol{n'}|\boldsymbol{n}) P_{\boldsymbol{n}}(t). \tag{2}$$

Having decided what the fundamental constituents of the system are, and so how the states of the system are defined, the next step is to give the transition rates $T(\boldsymbol{n}|\boldsymbol{n'})$. This will define the model, and is best done through examples. We will naturally choose as examples the models that will be analysed later in the paper. The first is the Brusselator

(Glansdorff and Prigogine, 1971) and the second is a model of cell polarity introduced by Altschuler et al. (2008).

The notation used to describe the chemical types and the rates in the Brusselator model follows that of Cross and Greenside (2009). In every domain i molecules of two species X and Y interact through the reactions of the Brusselator model (Glansdorff and Prigogine, 1971):

$$
\begin{aligned}
\emptyset &\xrightarrow{a} X_i, \\
X_i &\xrightarrow{b} Y_i, \\
2X_i + Y_i &\xrightarrow{c} 3X_i, \\
X_i &\xrightarrow{d} \emptyset.
\end{aligned}
\tag{3}
$$

In order, these reactions describe: (i) the creation of a new X molecule, (ii) an X molecule spontaneously transforming into a Y molecule, (iii) two X molecules reacting with a Y, changing it to an X, and (iv) X molecules being removed from the system. The rates at which the reactions occur are denoted by a, b, c and d, and X_i and Y_i are the molecules that are in domain i at the time that the reaction occurs. Each of these reactions are assumed to occur independently, and without memory of the previous states of the system. In addition to the reactions given in Eq. (3), migration reactions, which describe molecular diffusion from one domain to another, have to be specified. For every pair of neighbouring domains i and j, molecules of the two species X and Y may diffuse from one domain to the other according to

$$
X_i \xrightarrow{\alpha} X_j, \quad Y_i \xrightarrow{\beta} Y_j \ .
\tag{4}
$$

The second example consists of a two-dimensional 'cell' enclosed within a perfectly circular membrane (Altschuler et al., 2008). The centre of the cell is known as the 'cytoplasmic pool' and contains signalling molecules which we denote by C. In our implementation, the one-dimensional circular membrane is divided up into domains labelled by an index i; molecules lying within domains i are denoted by M_i. The following reactions take place:

$$
\begin{aligned}
C &\xrightarrow{k_{\text{on}}} M_i, \\
M_i &\xrightarrow{k_{\text{off}}} C, \\
M_i + C &\xrightarrow{k_{fb}} 2M_i.
\end{aligned}
\tag{5}
$$

The first two reactions describe a molecule in the cytoplasmic pool attaching itself to a random domain on the membrane and a molecule detaching and returning to the cytoplasmic pool, respectively. The third reaction represents a molecule on the membrane attempting to recruit another molecule from the cytoplasmic pool to the same domain. As for the Brusselator, there is also diffusion, in this case along the membrane:

$$
M_i \xrightarrow{\alpha} M_j,
\tag{6}
$$

for any pair of neighbouring membrane domains i and j. The reaction scheme is illustrated in Fig. 1 (after Altschuler *et al*).
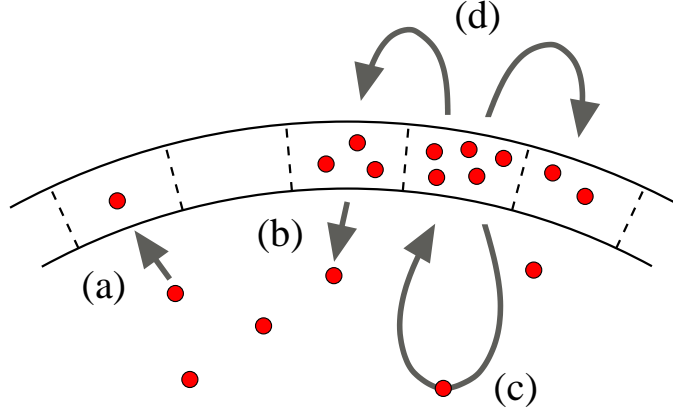
Figure 1: Illustration of the reactions of the cell polarisation model. The three reactions are (a) attachment of signalling molecules to the membrane, (b) the return of a membrane molecule to the cytoplasmic pool, (c) recruitment of new signalling molecules by those already on the membrane, and (d) diffusion along the membrane, which we implement as migration between neighbouring domains.

Once the reaction scheme has been picked and laid out as in Eqs. (3)-(6), the transition rates $T(\boldsymbol{n}|\boldsymbol{n}')$ can be chosen. This effectively specifies the model. When writing the transition rates, we will only list the variables in the domains that are involved in the reaction.

For the Brusselator, the number of X_i and Y_i will be denoted by ℓ_i and m_i, respectively. We then invoke mass action, that is, the rate of the transition for a given reaction is proportional to the product of the densities of the reactants, to write the transition rates as

$$
\begin{aligned}
T_1(\ell_i + 1, m_i | \ell_i, m_i) &= a, \\
T_2(\ell_i - 1, m_i + 1 | \ell_i, m_i) &= b\,\frac{\ell_i}{V}, \\
T_3(\ell_i + 1, m_i - 1 | \ell_i, m_i) &= c\,\frac{\ell_i{}^2 m_i}{V^3} \\
T_4(\ell_i - 1, m_i | \ell_i, m_i) &= d\,\frac{\ell_i}{V},
\end{aligned}
\tag{7}
$$

where the subscripts on the rates refer to the four reactions in Eq. (3) and where V is the volume of the domain. These transition rates are as in the usual Brusselator model (Boland et al., 2009; Biancalani et al., 2011), although it is worth noting that we have not imposed a fixed limit on the number of molecules permitted in a domain (in contrast with some previous work (Lugo and McKane, 2008; Biancalani et al., 2010); this is reflected in the fact we use the inverse volume of a domain, V^{-1}, rather than the inverse total number of molecules in a domain, as the expansion parameter).

The reaction rates which describe molecular diffusion reactions in Eq. (4) are given by

$$
\begin{aligned}
T_5(\ell_i - 1, \ell_j + 1 | \ell_i, \ell_j) &= \alpha\,\frac{\ell_i}{zV}, \\
T_6(m_i - 1, m_j + 1 | m_i, m_j) &= \beta\,\frac{m_i}{zV}.
\end{aligned}
\tag{8}
$$

Here z is the number of nearest neighbours of a domain and the index j denotes a nearest neighbour of domain i.

To model spontaneous cell polarisation we denote the number of C and M_i molecules by ℓ and m_i, respectively. The transition rates are then taken to be

$$
\begin{aligned}
T_1(\ell - 1, m_i + 1 | \ell, m_i) &= k_{\text{on}} \frac{\ell}{V} \\
T_2(\ell + 1, m_i - 1 | \ell, m_i) &= k_{\text{off}} \frac{m_i}{V}, \\
T_3(\ell - 1, m_i + 1 | \ell, m_i) &= k_{\text{fb}} \frac{\ell m_i}{V^2},
\end{aligned}
\tag{9}
$$

for the reactions (5) and

$$
T_4(m_i - 1, m_j + 1 | m_i, m_j) = \alpha \frac{m_i}{2V}, \tag{10}
$$

for the reactions (6).

The transition rates for these two models can be substituted into Eq. (2) which can then, together with suitable initial and boundary conditions, be used to solve for $P_{\boldsymbol{n}}(t)$. They can also be used as the basis for setting up a simulation using the Gillespie algorithm (Gillespie, 1976, 1977). Our approach in this paper will be to devise approximation schemes for the master equation and to check the results obtained from such schemes with simulations based on the Gillespie algorithm.

We end this section by generalising the above formulation so that it applies to a general biochemical network and by extension to any network of interacting agents. To do this, suppose that there are L different constituents in the system. These could be labelled by their type, the domain they occupy, etc. We will denote them by $Z_I, I = 1, \ldots, L$ and at a given time there will be n_I of them, so that the state of the system can be specified by $\boldsymbol{n} = (n_1, \ldots, n_L)$, as described earlier in this section. We suppose that there are M reactions which interconvert species:

$$
\sum_{I=1}^{L} r_{I\mu} Z_I \longrightarrow \sum_{I=1}^{L} p_{I\mu} Z_I, \quad \mu = 1, 2, \ldots M. \tag{11}
$$
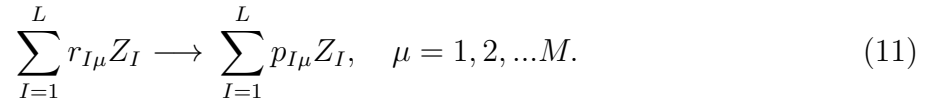
Here the numbers $r_{I\mu}$ and $p_{I\mu}$ ($I = 1, \ldots, L; \mu = 1, \ldots, M$) describe respectively the population of the reactants and the products involved in the reaction. The reactions Eqs. (3)-(6) are simple examples of this general set of reactions.

A quantity which is central to the structure of both the mesoscopic and macroscopic equations is the stoichiometry matrix, $\nu_{I\mu} \equiv r_{I\mu} - p_{I\mu}$, which describes how many molecules of species Z_I are transformed due to the reaction μ. In the notation introduced above for the master equation, $\boldsymbol{n}' = \boldsymbol{n} - \boldsymbol{\nu}$, where $\boldsymbol{\nu}_\mu = (\nu_{1\mu}, \ldots, \nu_{L\mu})$ is the stoichiometric vector corresponding to reaction μ. Therefore the master equation (2) may be equivalently written as

$$
\frac{\mathrm{d}P_{\boldsymbol{n}}(t)}{\mathrm{d}t} = \sum_{\mu=1}^{M} \left[T_\mu(\boldsymbol{n} | \boldsymbol{n} - \boldsymbol{\nu}_\mu) P_{\boldsymbol{n} - \boldsymbol{\nu}_\mu}(t) - T_\mu(\boldsymbol{n} + \boldsymbol{\nu}_\mu | \boldsymbol{n}) P_{\boldsymbol{n}}(t) \right]. \tag{12}
$$

Many models of interest involve only a handful of different reactions; in this situation, it is often convenient to rewrite the master equation as a sum over reactions as in Eq. (12), rather than over pairs of states $\boldsymbol{n}$ and $\boldsymbol{n'}$ as in Eq. (2). In the next section we will describe how the mesoscopic description of the models can be obtained from the master equation (12). These can then be used as the basis for the calculational schemes which we wish to implement.

3. Derivation of the macro- and mesoscopic equations

3.1. Macroscopic equation

Before we derive the mesoscopic equations from the master equation, we will carry out the simpler task of deriving the macroscopic equations. This will be done for the case of the general biochemical network described in section 2.

This is achieved by multiplying Eq. (12) by $\boldsymbol{n}$, and summing over all possible values of $\boldsymbol{n}$. After making the change of variable $\boldsymbol{n} \to \boldsymbol{n} + \boldsymbol{\nu}$ in the first summation, one finds that

$$\frac{\mathrm{d}\langle \boldsymbol{n}(t) \rangle}{\mathrm{d}t} = \sum_{\mu=1}^{M} \boldsymbol{\nu}_\mu \langle T_\mu(\boldsymbol{n} + \boldsymbol{\nu}_\mu | \boldsymbol{n}) \rangle, \tag{13}$$

where the angle brackets define the expectation value:

$$\langle \cdots \rangle = \sum_{\boldsymbol{n}} (\cdots) P_{\boldsymbol{n}}(t). \tag{14}$$

In the limit where both the particle numbers and the volume become large, we will take the state variables to be the concentration of the constituents n_I/V, rather than their number n_I. These will be assumed to have a finite limit as $V \to \infty$. Specifically, the state of the system will be determined by the new variables

$$y_I = \lim_{V \to \infty} \frac{\langle n_I \rangle}{V}, \quad \text{where} \quad I = 1, \ldots, L.$$

From Eq. (13) we have that

$$\frac{\mathrm{d}y_I}{\mathrm{d}\tau} = \sum_{\mu=1}^{M} \nu_{I\mu} f_\mu(\boldsymbol{y}), \quad I = 1, \ldots, L, \tag{15}$$

where $\tau = t/V$ and where

$$\begin{aligned}
f_\mu(\boldsymbol{y}) &= \lim_{V \to \infty} \langle T_\mu(\boldsymbol{n} + \boldsymbol{\nu}_\mu | \boldsymbol{n}) \rangle \\
&= \lim_{V \to \infty} T_\mu(\langle \boldsymbol{n} \rangle + \boldsymbol{\nu}_\mu | \langle \boldsymbol{n} \rangle) \\
&= \lim_{V \to \infty} T_\mu(V\boldsymbol{y} + \boldsymbol{\nu}_\mu | V\boldsymbol{y}).
\end{aligned} \tag{16}$$

In the above we have used the fact that in the macroscopic limit the probability distribution functions are Dirac delta functions and so, for instance, $\langle n^m \rangle = \langle n \rangle^m$, for any integer m.

The equation

$$\frac{\mathrm{d}y_I}{\mathrm{d}\tau} = A_I(\boldsymbol{y}), \tag{17}$$

where

$$A_I(\boldsymbol{y}) \equiv \sum_{\mu=1}^{M} \nu_{I\mu} f_\mu(\boldsymbol{y}), \quad I = 1, \ldots, L, \tag{18}$$

is the macroscopic equation corresponding to the microscopic master equation (12). It can be calculated from a knowledge of the stoichiometric matrix $\nu_{I\mu}$ and the transition rates $T_\mu(\boldsymbol{n} + \boldsymbol{\nu}_\mu | \boldsymbol{n})$. We scaled time by a factor of V simply because the choice we made for the transition rates (7)-(10) were finite as $V \to \infty$, but we could have easily incorporated an extra factor of V in these rates through a time rescaling. We also chose particularly simple forms for these transition rates in that they were all functions of the species concentration n_I/V. More generally, they might separately be functions of n_I and V, which become functions of the species concentration n_I/V only when both the particle numbers and the volume become large, so that in the limit $V \to \infty$ they become functions of the macroscopic state variable $\boldsymbol{y}$.

3.2. Mesoscopic equation

It is perhaps useful at this stage to recall precisely what is meant by the terms 'microscopic', 'mesoscopic' and 'macroscopic'. The microscopic description is the one based on the fundamental constituents whose reactions are described by relations such as those in Eqs. (3)-(6). The dynamics of the processes are described by the master equation or by the Gillespie algorithm. The macroscopic description has been derived from this microscopic description above: it only involves average quantities and their time evolution. In between these two levels of description is the 'meso'-level description, where the continuous variable of the macro-description is used, but where the stochastic effects due to the discrete nature of the individuals is retained. Some other authors include master equations at the meso-level leaving the micro-level for the world of atoms and molecules, but this does not seem such a useful assignment in the biological context in which we are working. The derivation of the mesoscopic equation follows similar lines to the calculation above, with the important difference that we do not take an average, or equivalently, do not take the limit $V \to \infty$.

We begin by substituting $y_I = n_I/V$ directly into the master equation. Since, as discussed above, our transition rates are all functions of n_I/V we simply replace $T_\mu(\boldsymbol{n} + \boldsymbol{\nu}_\mu | \boldsymbol{n})$ by $f_\mu(\boldsymbol{y})$ in the notation of Eq. (16). In addition we will denote the pdf $P_{\boldsymbol{n}}(t)$ where $\boldsymbol{n}$ has been replaced by $V\boldsymbol{y}$ as $P(\boldsymbol{y}, t)$. With these changes we may write the master equation (12) as

$$\frac{\partial P(\boldsymbol{y}, t)}{\partial t} = \sum_{\mu=1}^{M} \left[f_\mu \left(\boldsymbol{y} - \frac{\boldsymbol{\nu}_\mu}{V} \right) P \left(\boldsymbol{y} - \frac{\boldsymbol{\nu}_\mu}{V}, t \right) - f_\mu(\boldsymbol{y}) P(\boldsymbol{y}, t) \right]$$

For V large, the steps $\boldsymbol{\nu}_\mu/V$ are likely to be very small, suggesting that we may expand the

functions P and f as Taylor series around $\boldsymbol{y}$. Truncating at order $\mathcal{O}(V^{-2})$, we arrive at

$$\frac{\partial P(\boldsymbol{y}, \tau)}{\partial \tau} = -\sum_{\mu=1}^{M} \sum_{I} \nu_{I\mu} \frac{\partial}{\partial y_I} \left[f_\mu(\boldsymbol{y}) P(\boldsymbol{y}, \tau) \right] \tag{19}$$

$$+ \sum_{\mu=1}^{M} \frac{1}{2V} \sum_{I,J} \nu_{I\mu} \nu_{J\mu} \frac{\partial^2}{\partial y_I \partial y_J} \left[f_\mu(\boldsymbol{y}) P(\boldsymbol{y}, \tau) \right],$$

where as before we have absorbed a factor of V into the rescaled time variable $\tau = t/V$. This is a Fokker-Planck equation which can be cast into the standard form (Risken, 1989; Gardiner, 2009)

$$\frac{\partial P(\boldsymbol{y}, \tau)}{\partial \tau} = -\sum_{I} \frac{\partial}{\partial y_I} \left[A_I(\boldsymbol{y}) P(\boldsymbol{y}, \tau) \right] + \frac{1}{2V} \sum_{I,J} \frac{\partial^2}{\partial y_I \partial y_J} \left[B_{IJ}(\boldsymbol{y}) P(\boldsymbol{y}, \tau) \right], \tag{20}$$

where $A_I(\mathbf{y})$ is defined by Eq. (18) and where

$$B_{IJ}(\boldsymbol{y}) = \sum_{\mu=1}^{M} \nu_{I\mu} \nu_{J\mu} f_\mu(\boldsymbol{y}), \quad I, J = 1, \ldots, L. \tag{21}$$

In the Fokker-Planck equation (20), the continuous nature of the state variables indicates that the individual nature of the constituents has been lost. However, the stochasticity due to this discreteness has not: it now manifests itself through the function $B_{IJ}(\boldsymbol{y})$. We can see this is the case through the presence of the factor $1/V$.

One might ask if this approach is consistent with the previous macroscopic derivation. As $V \to \infty$, the Fokker-Planck equation reduces to the Liouville equation

$$\frac{\partial P(\boldsymbol{y}, \tau)}{\partial \tau} = -\sum_{I} \frac{\partial}{\partial y_I} \left[A_I(\boldsymbol{y}) P(\boldsymbol{y}, \tau) \right]. \tag{22}$$

One can check by direct substitution that the solution to this equation is $P(\boldsymbol{y}, \tau) = \delta(\boldsymbol{y}(\tau) - \boldsymbol{y})$ where δ is the Dirac delta function and where $\boldsymbol{y}(\tau)$ is the solution of the macroscopic system (17); see (Gardiner, 2009) for details.

It is also natural to ask if it is useful to include higher order terms in V^{-1}. There are sound mathematical reasons for not going to higher order, for instance the pdf may become negative (Risken, 1989). As we will see, for the problems that we are interested in here (and many others) very good agreement with simulations can be found by working with the Fokker-Planck equation (20).

The Fokker-Planck equation (20) provides a mesoscopic description of the system but, like the master equation (2) from which it originated, it is an equation for a pdf. It is therefore quite distinct from the macroscopic equation (17), which is an equation for the state variables themselves. There does, however, exist an equation for the state variables which is completely equivalent to the Fokker-Planck equation (20) (Gardiner, 2009). This equation takes the form

$$\frac{\mathrm{d}y_I}{\mathrm{d}\tau} = A_I(\boldsymbol{y}) + \frac{1}{\sqrt{V}} \sum_{J} g_{IJ}(\boldsymbol{y}) \eta_J(\tau), \tag{23}$$

where the $\eta_J(\tau)$ are Gaussian white noises with zero mean and correlator

$$\langle \eta_I(\tau)\eta_J(\tau') \rangle = \delta_{IJ}\delta(\tau - \tau'), \tag{24}$$

and where $g_{IJ}(\boldsymbol{y})$ is related to $B_{IJ}(\boldsymbol{y})$ by

$$B_{IJ}(\boldsymbol{y}) = \sum_K g_{IK}(\boldsymbol{y})g_{JK}(\boldsymbol{y}). \tag{25}$$

The mesoscopic equation (23) generalises the macroscopic ordinary differential equation (17) with the addition of noise terms $\boldsymbol{\eta}(\tau)$ and so is a stochastic differential equation (SDE). As we will discuss below we need to specify that it is to be interpreted in the sense of Itō (Gardiner, 2009). Notice the direct relationship between this SDE and the macroscopic ODE: sending $V \to \infty$ in Eq. (23) immediately yields equation (17).

It is important to point out that the matrices $g_{IJ}(\boldsymbol{y})$ which define the behaviour of the noise cannot be found from the macroscopic equations, and a knowledge of the microscopic stochastic dynamics is essential if one is to understand the effects of noise. It is not permissible in this context to simply 'add noise terms' to the macroscopic equations to obtain a mesoscopic description, as some authors have done in the past. The only situation in which it is permissible to do this is if the noise is external to the system, that is, it does not originate from the internal dynamics of the system.

We end this section with two general comments on the mesoscopic equation (23). The first is that while there are no strong restrictions on the form of $A_I(\boldsymbol{y})$, there are on $B_{IJ}(\boldsymbol{y})$. From Eq. (21) we see that the matrix B is symmetric, but also that for any non-zero vector $\boldsymbol{w}$,

$$\sum_{I,J} w_I B_{IJ} w_J = \sum_{\mu=1}^{M} (\boldsymbol{w} \cdot \boldsymbol{\nu})^2 f_\mu(\boldsymbol{y}) \geq 0, \tag{26}$$

since $f_\mu(\boldsymbol{y}) \geq 0$. Thus B is positive semi-definite (Mehta, 1989). It follows that $B = g\,g^{\mathrm{T}}$ for some non-singular matrix g, where T denotes transpose (Mehta, 1989). One way of constructing such a matrix is to note that since B is symmetric, it can be diagonalised by an orthogonal transformation defined through a matrix O_{IJ}. Then since B is positive semi-definite, its eigenvalues are non-negative, and so

$$B = O\Lambda O^{\mathrm{T}} = g\,g^{\mathrm{T}}, \quad \text{where } g = O\Lambda^{1/2}, \tag{27}$$

and where Λ and $\Lambda^{1/2}$ are the diagonal matrices with respectively the eigenvalues and square root of the eigenvalues of B as entries. We can take the positive roots of the eigenvalues without loss if generality, since the sign can always be absorbed in the η_J factor in Eq. (23) (its distribution is Gaussian and so invariant under sign changes). It should also be pointed out that we can go further and make an orthogonal transformation on the noise, $\zeta_J = \sum_I S_{IJ}\eta_J$, and leave Eq. (24), and so its distribution, unchanged. The transformation matrix S can then be used to define a new matrix $G_{IJ} = \sum_K g_{IK}S_{JK}$, so that the form of the mesoscopic equation (23) is unchanged. So while the procedure outlined above gives us a way of constructing $g_{IJ}(\boldsymbol{y})$ from $B_{IJ}(\boldsymbol{y})$, it is not unique.

The second comment relates to the statement made earlier, that Eq. (23) is to be interpreted in the Itō sense. The singular nature of white noise means that in some cases SDEs are not uniquely defined by simply writing down the equation, but have to be supplemented with the additional information on how the singular nature of the process is to be interpreted (van Kampen, 2007; Gardiner, 2009). This happens when g_{IJ} depends on the state of the system $\boldsymbol{y}$; the noise is then said to be multiplicative. As we will see in the next section, this subtlety is not relevant within the LNA, since there the g_{IJ} is evaluated at a fixed point of the dynamics, and so ceases to depend on the state of the system. However, when going beyond the LNA, it is an important consideration. If one wishes to manipulate a multiplicative noise SDE like Eq. (23), then one must employ modified rules of calculus which take into account the contribution from the noise. We refer to Gardiner (2009) for details, and a complete discussion on the relationship between the Itō formulation and the alternatives.

This completes our general discussion of the derivation and form of the mesoscopic equation. In the next two sections we will apply it to the two models which we introduced in Section 2. In the first case we will consider use of the LNA is sufficient for our requirements, but in the second case we have to go beyond the LNA.

4. Stochastic patterns in population models

One of the reasons why deterministic reaction-diffusion systems are interesting is the fact that they may give rise to ordered structures, either in space or time. Many models displaying several different kinds of patterns have been extensively discussed in the literature (Murray, 2008; Cross and Greenside, 2009). However, the mathematical mechanisms which are responsible for the pattern formation are few and universal, and they can be conveniently analysed using the simplest models. Perhaps the most famous of these mechanisms was put forward by Turing in his pioneering study of morphogenesis (Turing, 1952), and it is now referred to as the 'Turing instability' (Murray, 2008; Cross and Greenside, 2009). It is invoked as a central paradigm in various areas of science to explain the emergence of steady spatial structures which typically look like 'stripes', 'hexagons' or 'spots' (Murray, 2008; Cross and Greenside, 2009).

In this section, we will be interested in reaction-diffusion systems exhibiting a Turing instability which are composed of discrete entities as described in section 2. The intrinsic noise in the system will render it stochastic. As we shall see, by means of the LNA, one is able to make analytical progress and so clarify the role of demographic noise in the pattern formation process. We shall show that systems of this kind display 'stochastic patterns' in addition to the conventional 'Turing patterns'. It has been suggested that stochastic patterns are responsible for the robustness of the patterning observed in population systems (Butler and Goldenfeld, 2009, 2011; Biancalani et al., 2010) and they have been applied in several ecological models (Butler and Goldenfeld, 2009; Bonachela et al., 2012), in the dynamics of hallucination (Butler et al., 2012) and in a biological model with stochastic growth (Woolley et al., 2011). In a similar way, the emergence of stochastic travelling waves has been studied (Biancalani et al., 2011), which has found application in a marine predator-prey system (Datta et al., 2010). There is also an existing literature on stochastic patterning in arid ecosystems (Ridolfi et al., 2011b) where the origin of the noise is extrinsic rather than

intrinsic.

The following analysis employs the Brusselator model to exemplify the general theory. This is a reaction scheme introduced by Lefever and Prigogine in the 1960s (Glansdorff and Prigogine, 1971) as a model of biochemical reactions which showed oscillatory behaviour. For our purposes, its interest lies in the fact that its spatial version is one of the simplest models which exhibits a Turing instability.

Before we begin the analysis of this model we need to discuss some aspects of the notation we will use. As explained in section 2 the labels that we have been using so far combine a spatial index with an index for the type of constituent, for example $J = \{j, s\}$. This was done in order to reduce the clutter of indices. However in the analysis below we will need to separate them, since we will assume a regular lattice structure for the domains, which will allow us to use Fourier analysis to effectively diagonalise the spatial part of the system. The Fourier components of a given function will be labelled through the argument of that function, leaving only the index (e.g. s) labelling the type. Specifically we will choose the spatial structure to be a regular D-dimensional hypercubic lattice with periodic boundary conditions and domain length l. Following the conventions of (Chaikin and Lubensky, 2000), the discrete spatial Fourier transform is then defined as:

$$\tilde{f}_k = l^D \sum_{j=1}^{\Omega} e^{-ilk\cdot j} f_j, \ \text{with} \ \ f_j = l^{-D}\Omega^{-1} \sum_{k=1}^{\Omega} e^{ilk\cdot j} \tilde{f}_k, \tag{28}$$

where Ω is the number of lattice points, j is a D-dimensional spatial vector and k is its Fourier conjugate. Note that i here is imaginary unit and not a spatial index, and although both j and k are D-dimensional vectors, we do not explicitly show this, in line with the notation adopted in section 2.

We begin the analysis by deriving the well-known deterministic Brusselator equations from this microscopic description using Eqs. (17) and (18):

$$\begin{aligned} \frac{du_i}{d\tau} &= a - (b+d)\, u_i + cu_i^2 v_i + \alpha \Delta u_i, \\ \frac{dv_i}{d\tau} &= bu_i - cu_i^2 v_i + \beta \Delta v_i, \end{aligned} \tag{29}$$

where $u_i = \ell_i / V$ and $v_i = m_i / V$, where u is the density of chemical species X and v is the density of chemical species Y. The symbol Δ represents the discrete Laplacian operator $\Delta f_j = (2/z) \sum_{j' \in \partial j} (f_j - f_{j'})$ where $j' \in \partial j$ indicates that the domain j' is a nearest neighbour of the domain j and z is the co-ordination number of the lattice. The spatial Fourier transform of the discrete Laplacian operator reads (Lugo and McKane, 2008):

$$\tilde{\Delta}_k = \frac{2}{D} \sum_{s=1}^{D} \left[\cos\left(k_s\, l\right) - 1 \right]. \tag{30}$$

It is possible to obtain a continuous spatial description, in the deterministic limit, by taking the limit of small domain length scale, $l \to 0$ (Lugo and McKane, 2008). By doing so,

one can recover the traditional partial differential equations for reaction-diffusion systems:

$$\begin{aligned}
\frac{\partial u}{\partial \tau} &= a - (b + d)\, u + cu^2 v + D_1 \nabla^2 u, \\
\frac{\partial v}{\partial \tau} &= bu - cu^2 v + D_2 \nabla^2 v,
\end{aligned} \tag{31}$$

where D_1 and D_2 are obtained by scaling the diffusivities α and β according to:

$$\frac{1}{2D} l^2 \alpha \mapsto D_1, \quad \frac{1}{2D} l^2 \beta \mapsto D_2. \tag{32}$$

However, we shall keep the space discrete in the following analysis because the theory is simpler to describe and it is most convenient for carrying out stochastic simulations. We shall also set $l = 1$, since this simply amounts to a choice of length scale, and this is the simplest choice.

The macroscopic equations (29) are Eq. (17) for the particular case of the Brusselator model. To find the corresponding mesoscopic equations we need to find the particular form of Eq. (23) for the Brusselator. This we can do by calculating $B_{IJ}(\boldsymbol{y})$, defined in Eq. (21), but we will find that we do not need to utilise the non-linear equation (23) to take the fluctuations into account; it is sufficient to use only a linearised form. This is the LNA, and is implemented by writing

$$y_I(t) = \langle y_I(t) \rangle + \frac{\xi_I(t)}{\sqrt{V}}, \tag{33}$$

where $\langle y_I(t) \rangle$ satisfies the macroscopic equation (17). Substituting Eq. (33) into Eq. (23), we expand in powers of $1/\sqrt{V}$. The terms which are proportional to $1/\sqrt{V}$ give an equation for ξ_I:

$$\frac{\mathrm{d}\xi_I}{\mathrm{d}\tau} = \sum_J \mathcal{J}_{IJ}(\langle \boldsymbol{y} \rangle)\xi_J + \sum_J g_{IJ}(\langle \boldsymbol{y} \rangle)\eta_J(\tau), \tag{34}$$

where $\mathcal{J}$ is the Jacobian of the system.

In many situations, including the one we are describing here, we are only interested in the fixed points of the macroscopic equation, in which case $\langle \boldsymbol{y} \rangle = \boldsymbol{y}^*$ and the matrices $\mathcal{J}$ and g can be replaced by their values at the fixed point $\boldsymbol{y}^*$. The SDE (34) now involves only constant matrices:

$$\frac{\mathrm{d}\xi_I}{\mathrm{d}\tau} = \sum_J \mathcal{J}_{IJ}^* \xi_J + \sum_J g_{IJ}^* \eta_J(\tau). \tag{35}$$

For the specific case of the Brusselator the index I includes the spatial index i and an index $s = 1, 2$ which distinguishes between the variables u and v. If we take the spatial Fourier transform of Eq. (35), translational invariance implies that the matrices $\mathcal{J}^*$ and g^* are diagonalised in the spatial variables, and so this equation becomes

$$\frac{\partial \tilde{\xi}_\gamma(k, \tau)}{\partial \tau} = \sum_{\delta=1}^{2} \mathcal{J}_{\gamma\delta}^*(k)\, \tilde{\xi}_\delta(k, \tau) + \sum_{\delta=1}^{2} g_{\gamma\delta}^*(k)\, \tilde{\eta}_\delta(k, \tau). \tag{36}$$

We are now in a position to discuss both the classical Turing patterns found in deterministic equations such as Eqs. (31) and the stochastic Turing patterns found in the corresponding mesoscopic equations. The homogeneous fixed point of Eqs. (31) is given by

$$u^* = \frac{a}{d}, \quad v^* = \frac{bd}{ac},$$ (37)

although from now on we will set $c = d = 1$, as is common in the literature (Glansdorff and Prigogine, 1971). In the deterministic case, the linear stability analysis about this fixed point is a special case of that carried out above in the stochastic picture, and corresponds to ignoring the noise term in Eq. (36). Therefore the small, deterministic, spatially inhomogeneous, perturbations $\tilde{\xi}_\gamma(k, \tau)$ satisfy the equation

$$\frac{\partial \tilde{\xi}_\gamma(k, \tau)}{\partial \tau} = \sum_{\delta=1}^{2} \mathcal{J}_{\gamma\delta}^*(k) \, \tilde{\xi}_\gamma(k, \tau),$$ (38)

where the Jacobian is found to be

$$\mathcal{J}^*(k) = \begin{pmatrix} b - 1 + \alpha\tilde{\Delta}_k & a^2 \\ -b & -a^2 + \beta\tilde{\Delta}_k \end{pmatrix}.$$ (39)

The eigenvalues of the Jacobian, $\lambda_\gamma(k)$ ($\gamma = 1, 2$), give information about the stability of the homogeneous state. In particular, the perturbations $\tilde{\xi}_\gamma(k, \tau)$ grow like linear combinations of $e^{\lambda_\gamma(k)t}$, therefore if $\mathrm{Re}[\lambda_\gamma(k)]$ is positive for some k and some γ, then the perturbation will grow with time and the homogeneous state will be unstable for this value of k. Turing's insight was that the pattern eventually formed as a result of the perturbation is characterised by this value of k. The overall scenario is complicated by the nature of the boundary conditions, the presence of other attractors and the effect of the non-linearities (Cross and Greenside, 2009), but in the following we shall ignore these, and consider only the simplest case in order to understand the main concepts.

In this most straightforward situation, the small perturbation which excites the unstable k-th Fourier mode will cause the concentrations u and v to develop a sinusoidal profile about their fixed point values characterised by the wave-number k. The pattern is steady or pulsating depending on whether or not the imaginary part $\mathrm{Im}[\lambda_\gamma(k)]$ is zero. In both cases, the amplitude of sinusoidal profile increases exponentially with a time-scale $1/\mathrm{Re}[\lambda_\gamma(k)]$, and so clearly the eigenvalue with the largest real part will dominate. By moving away from the homogeneous state the linear approximation will eventually lose its validity and the effect of the non-linear terms will become relevant. If the system admits no solutions which diverge, the growth will be damped by the non-linearities to some non-zero value, which defines the final amplitude of the spatial pattern.

Typically, the interesting case occurs when a control parameter triggers the pattern formation by making the real part of one of the eigenvalues positive. For the Brusselator this is illustrated in Fig. 2, where the relevant eigenvalue of $\mathcal{J}$, (i.e. the one which becomes positive) is shown for different values of the parameter b. Here b is the control parameter with the other free parameter a fixed and equal to 1.5. For $b < b_c \approx 2.34$, the real part of both eigenvalues is negative and thus the homogeneous state is stable. This corresponds
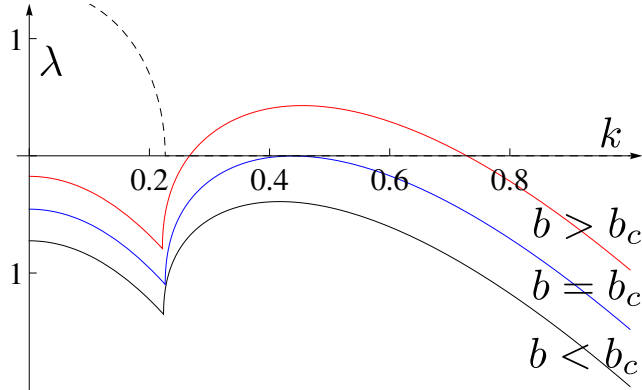
Figure 2: Real part (solid lines) and imaginary part (dashed line) of the most unstable eigenvalue of matrix $\mathcal{J}$. Parameter values are $a = 1.5$, $\alpha = 2.8$ and $\beta = 22.4$, (Cross and Greenside, 2009). Solid lines correspond to $b = 1.8$ (black), $b \equiv b_c = 2.34$ (blue) and $b = 2.9$ (red). The imaginary part is shown for $b = 2.34$ only, although it looks qualitatively the same for the range of b values displayed.

to the situation where there are no patterns. The critical value of b, b_c, occurs when the real part of one of the eigenvalues is tangent to the k-axis. For values of b larger than b_c, a window of unstable modes sets in, delimited by the intersections of $\text{Re}[\lambda]$ with the k-axis. Each mode contributes to the pattern, although the wavelength which maximises $\text{Re}[\lambda]$ is the one with bigger amplitude as it grows faster than the other modes.

As we have mentioned already no Turing patterns emerge from the deterministic equations if both eigenvalues of $\mathcal{J}^*$ have a negative real part for all k, since then the perturbations decay away. However, when the system is described as an IBM, intrinsic noise is present which acts as a continuous perturbation on the homogeneous state, exciting each k-mode. Given that the homogeneous state is stable for all k, every excitation decays, although each with a different time-scale given by $1/\text{Re}[\lambda_\gamma(k)]$. Thus, the closer $\text{Re}[\lambda_\gamma(k)]$ is to zero, the slower the relaxation of the corresponding k-mode. The situation is illustrated in Fig. 2, where the value of k at which this occurs for $b = 1.8 < b_c$ is seen to be the maximum of the curve of $\text{Re}[\lambda_\gamma(k)]$ versus k. The net effect is that only a window of modes around the maximum of $\text{Re}[\lambda_\gamma(k)]$ is visible in the dynamics, the others having died away. This patterns have been called stochastic Turing patterns (Biancalani et al., 2010). Figure 3 shows the result of numerically simulating a two-dimensional Brusselator model using the Gillespie algorithm in a parameter regime where the homogeneous state is stable to perturbations for all k.

From the argument given above we would expect that in order to study stochastic Turing patterns it would be sufficient to linearise about the homogeneous state as before, but now to include noise. In addition, we would expect that the detailed nature of the noise would be unimportant, only its strength. This justifies the use of the LNA in the analysis of stochastic Turing patterns and implies that the arguments which we have used can be made quantitative through the use of Eq. (36). To do this we first take the temporal Fourier
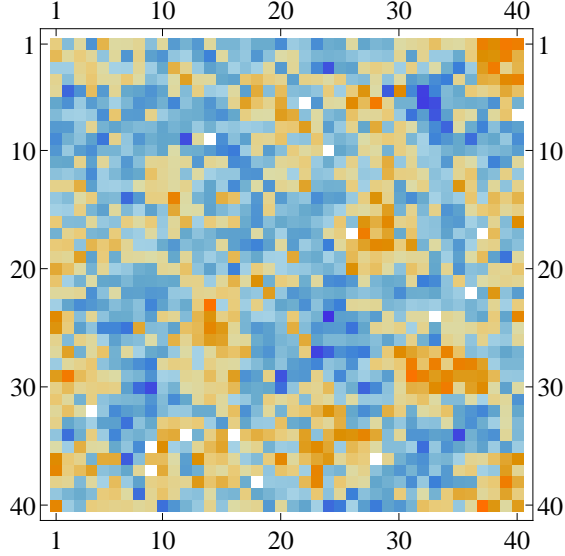
Figure 3: (Colour online) Snapshot of two-dimensional stochastic Turing patterns for species X. The system consists of 40×40 domains with periodic boundary conditions. The parameters are the same as in Fig. 2, except that $b = 2$ and $V = 500$. Simulations started close to (u^*, v^*) and ran for $t/V = 15$. We have used warm colours for values of $u > u^*$ and cold colours for $u < u^*$. White pixels indicate the fixed point value, u^*.

transform of this equation to obtain

$$\tilde{\xi}_\gamma(k, \omega) = \sum_{\delta, \sigma} \Phi_{\gamma\delta}^{-1}(k, \omega)\, g_{\delta\sigma}^*(k)\, \tilde{\eta}_\sigma(k, \omega), \tag{40}$$

where $\Phi_{\gamma\delta}(k, \omega) = -i\omega\delta_{\gamma\delta} - \mathcal{J}_{\gamma\delta}^*$. From Eq. (40) we can find an expression for the power spectrum of the fluctuations which is the quantity we use to analyse the patterns:

$$P_\gamma(k, \omega) = \langle |\tilde{\xi}_\gamma(k, \omega)|^2 \rangle = \sum_{\delta, \sigma} \Phi_{\gamma\delta}^{-1}(k, \omega)\, \tilde{B}_{\delta\sigma}^*(k)\, (\Phi^\dagger)_{\sigma\gamma}^{-1}(k, \omega), \tag{41}$$

where $\tilde{B}^*(k)$ is obtained by Fourier transforming in space the matrix $B(\boldsymbol{y})$ given by Eq. (25), evaluated at the fixed point. The details of this calculation can be found in Lugo and McKane (2008); here we simply state the final formulae — which holds for any two-species system which has one spatial dimension:

$$\begin{aligned} \tilde{B}_{11}^*(k) &= B_{11}^* - 2u^*\alpha\tilde{\Delta}_k, \\ \tilde{B}_{12}^*(k) &= B_{12}^*, \quad \tilde{B}_{12}^*(k) = B_{12}^*, \\ \tilde{B}_{22}^*(k) &= B_{22}^* - 2v^*\beta\tilde{\Delta}_k. \end{aligned} \tag{42}$$

Here, the matrix B^* indicates the correlation matrix of Eq. (25) calculated at the fixed point for the corresponding non-spatial system. For instance, in the case of the Brusselator this is obtained by considering only the reactions (3) without those of Eqs. (4), which yields: $B_{11}^* = 2a(1 + b)$, $B_{12}^* = B_{21}^* = -2ab$ and $B_{22}^* = 2ab$.
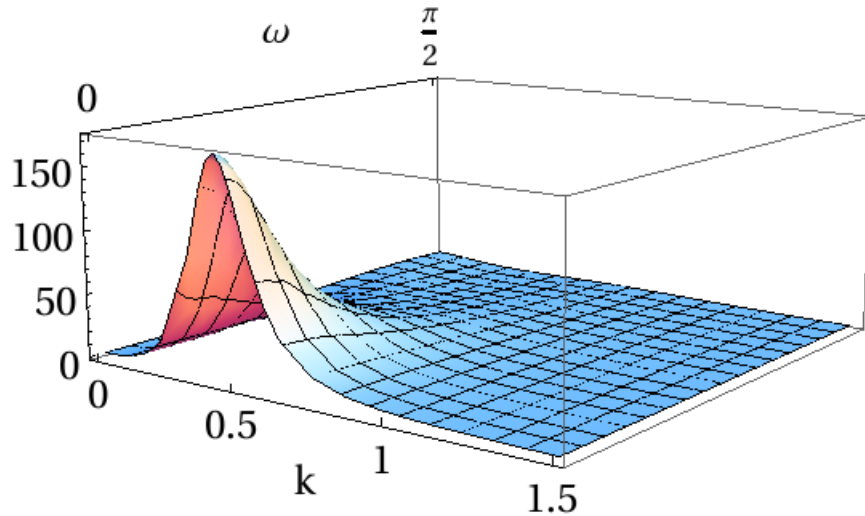
Figure 4: (Colour online) Analytical power spectrum of species X, obtained with the same parameter values as in Fig. 3, and from the analytical expression (41).

The expression (41) for the power is plotted in Fig. 4. We have also measured the numerical power spectrum via the Gillespie algorithm, and found good agreement with the analytical expression, confirming that the dynamics is captured within the approximation scheme we have used. The power spectrum shows a peak at $k \neq 0$ and $\omega = 0$ which indicates the presence of stochastic Turing patterns of length scale characterised by k. As shown in previous studies (Butler and Goldenfeld, 2009; Biancalani et al., 2010), one can compute the region of parameters for which stochastic Turing patterns arise by looking at when $P_\gamma(k, 0)$ has a maximum for some non-zero k. It has been found that those regions are greatly enlarged, making the pattern formation a much more robust mechanism. Specifically, stochastic patterns may appear even for equal diffusivities, a condition for which deterministic patterns cannot occur (Butler and Goldenfeld, 2009; Biancalani et al., 2010; Butler and Goldenfeld, 2011).

Notice that unlike their deterministic analogue, stochastic patterns are not steady but they continuously decay whilst they are re-created by the effect of the noise (Scott et al., 2011; Biancalani et al., 2011). In Fig. 5 this effect is shown by means of the dynamics of a one-dimensional Brusselator. The noisy nature of patterns makes them hard to detect by the naked eye, and the emergence of a length scale only becomes clear by means of a Fourier analysis.

The theory we have presented here is rather general and also applies to other types of
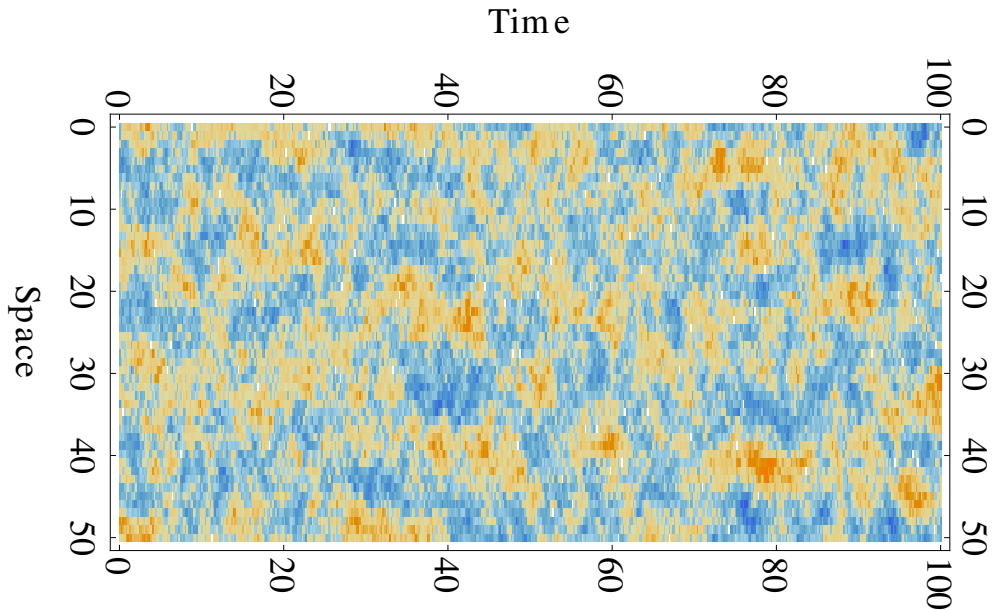
Figure 5: (Colour online) Dynamics of a one-dimensional system of 50 domains run for $2*10^3$ time ($\tau = t/V$) units. Parameter values are the same as in Fig. 3.

pattern instabilities. For instance, if the power spectrum showed a peak at $k \neq 0$ and $\omega \neq 0$ the overall pattern would consist of stochastic travelling waves (Biancalani et al., 2011). Finally, it should be mentioned that the amplitude of stochastic patterns scales as $1/\sqrt{V}$ and it therefore vanishes in the limit $V \to \infty$, in which the deterministic picture is recovered. Stochastic patterns arise because of the noise related to the discreteness of the populations and they are therefore less relevant for populations in which the number of individuals is macroscopic.

5. Spontaneous Emergence of Cell Polarity

Many important biological functions require cells to break their rotational symmetry and form a distinguished 'nose' and 'tail'. The emergence of this symmetry-breaking is known as polarisation, and the precise mechanisms responsible are not yet fully understood. A few years ago Altschuler et al. (2008) proposed a very simple model of polarisation in which signalling molecules self-recruit to the cell membrane, before diffusing away. They showed through simulations that, depending on the choice of model parameters, the membrane molecules may spontaneously aggregate.

Further investigations have followed several different lines, including more detailed simulations (Petzold et al, 2012) and mathematically rigorous studies (Gupta, 2012). In this section, we will show how the mesoscopic framework developed in previous sections may be used to analytically describe the spontaneous emergence of cell polarity in this model. This effect is stronger than those described by the LNA, and it will require a different theoretical approach.

Before beginning the analysis, it is worth noticing that in this model the total number of molecules does not change; there is thus no need to distinguish between this and the volume, so we write $N = V$. Moreover, the variables ℓ (giving the number of molecules in the cytoplasmic pool) and m_i (the number of molecules in membrane domain i) are related by: $\ell + \sum_i m_i = V$.

As usual, we first explore the behaviour of the macroscopic equations. Substituting the transition rates (5) and (6) into equations (17) and (16), we arrive at

$$\frac{du}{d\tau} = (k_{\text{off}} - u\, k_{\text{fb}}) \sum_i v_i - u\, k_{\text{on}}$$

$$\frac{dv_i}{d\tau} = u\, k_{\text{on}} + (u\, k_{\text{fb}} - k_{\text{off}})\, v_i + \alpha \Delta v_i \,.$$

Conservation of the total number of molecules implies that $u + \sum_j v_j = 1$, so we may eliminate u from the above system to obtain

$$\frac{dv_i}{d\tau} = \left(1 - \sum_j v_j\right)\left(k_{\text{on}} + v_i\, k_{\text{fb}}\right) - v_i k_{\text{off}} + \alpha \Delta v_i \,. \tag{43}$$

In the appropriate continuum limit, this equation agrees with that of Altschuler et al. (2008).

As with the Brusselator, there is a homogeneous fixed point. Putting $v_i \equiv v$ and setting $dv_i/d\tau = 0$, we find the quadratic equation

$$(1 - \Omega\, v)(k_{\text{on}} + v\, k_{\text{fb}}) - v\, k_{\text{off}} = 0 \,. \tag{44}$$

For simplicity, we consider the case in which $k_{\text{on}} \approx 0$ (that is, almost all membrane molecules exist as result of the feedback mechanism). In this limit, the homogeneous fixed point is given by $v = v^*/\Omega$, where v^* is the mean fraction of molecules on the membrane:

$$v^* = \begin{cases} \left(1 - \dfrac{k_{\text{off}}}{k_{\text{fb}}}\right) & \text{if} \quad k_{\text{fb}} > k_{\text{off}} \\ 0 & \text{otherwise.} \end{cases} \tag{45}$$

To gain further insight, we pass to Fourier space, where Eq. (43) with $k_{\text{on}} = 0$ becomes

$$\frac{d\tilde{v}_k}{dt} = \tilde{v}_k \left[k_{\text{fb}}(1 - l^{-1}\tilde{v}_0) - k_{\text{off}} + \alpha\big(\cos(lk) - 1\big) \right]. \tag{46}$$

The Jacobian for this system is diagonal, and it is straightforward to read off the eigenvalues at the fixed point $\tilde{v}_k = \delta_{w,0}\, l\, v^*$ as

$$\lambda_k = \begin{cases} k_{\text{off}} - k_{\text{fb}} \,, & \text{if} \quad k = 0 \\ \alpha\big(\cos(lk) - 1\big) & \text{if} \quad k \neq 0 \,. \end{cases}$$

We can conclude from this analysis that provided $k_{\text{fb}} > k_{\text{off}}$ the homogeneous fixed point is non-zero and stable. This is a puzzle: the homogeneous state corresponds to the signalling

molecules being spread uniformly around the membrane, if this state is stable, then how can polarisation occur?

We postulate that the answer lies in the following observation. Notice that if the diffusion coefficient α is small then the modes with wave number $k \neq 0$ are only marginally stable; their associated eigenvalues are close to zero. In this regime, a small random perturbation (resulting from intrinsic noise, for example) may be enough to push the system very far from its equilibrium state. Moreover, the stochastic dynamics in this regime cannot be understood within the framework of the LNA, for the simple reason that when the system has been pushed far from its steady state, linearisation around that state is no longer representative of the true dynamics. To make analytical progress, we will need to deal with the non-linearity of the model some other way.

We begin by writing down the mesoscopic equations. For our purposes, the Fokker-Planck equation (20) is the most useful, with A and B given by

$$A_i(\boldsymbol{v}) = \left(1 - \sum_j v_j\right)\left(k_{\mathrm{on}} + v_i\, k_{\mathrm{fb}}\right) - v_i k_{\mathrm{off}} + \alpha \Delta v_i \,,$$

$$B_{ij}(\boldsymbol{v}) = \delta_{ij}\left[\left(1 - \sum_k v_k\right)\left(k_{\mathrm{on}} + v_i\, k_{\mathrm{fb}}\right) + v_i k_{\mathrm{off}}\right].$$

Note that we have neglected terms of order α/V from the noise, as we are interested in behaviour when α is small and V large, so α/V is negligible. As with the macroscopic equations, this system is easier to analyse in Fourier space. We introduce the distribution $P(\tilde{\boldsymbol{v}}, \tau)$ of Fourier variables $\tilde{v}_k$, which satisfies the Fokker-Planck equation

$$\frac{\partial P(\tilde{\boldsymbol{v}}, \tau)}{\partial \tau} = -\sum_k \frac{\partial}{\partial \tilde{v}_k}\left[\tilde{A}_k(\tilde{\boldsymbol{v}})P(\tilde{\boldsymbol{v}}, \tau)\right] \tag{47}$$

$$+ \frac{1}{2V}\sum_{k,k'} \frac{\partial}{\partial \tilde{v}_k}\frac{\partial}{\partial \overline{\tilde{v}_{k'}}}\left[\tilde{B}_{k,k'}(\tilde{\boldsymbol{v}})P(\tilde{\boldsymbol{v}}, \tau)\right] \,,$$

where

$$\tilde{A}_k(\tilde{\boldsymbol{v}}) = \tilde{v}_k\left[k_{\mathrm{fb}}(1 - l^{-1}\tilde{v}_0) - k_{\mathrm{off}} + \alpha\big(\cos(lk) - 1\big)\right]$$

$$\tilde{B}_{k,k'}(\tilde{\boldsymbol{v}}) = \tilde{v}_{k+k'}\, l\left[k_{\mathrm{fb}}(1 - l^{-1}\tilde{v}_0) + k_{\mathrm{off}}\right]. \tag{48}$$

Note that the Fourier modes may take complex values, and we use $\partial/\partial\overline{\tilde{v}_{k'}}$ to denote differentiating with respect to the complex conjugate. For later convenience, we assume Ω is odd and number the modes by $k \in \{-(\Omega-1)/2, \ldots, (\Omega-1)/2\}$, so that $\tilde{v}_{-k} = \overline{\tilde{v}_k}$.

Our remaining analysis is informed by two observations. First, we note that the non-linearity in equation (48) arises only from terms involving $\tilde{v}_0$. Second, in the interesting regime $k_{\mathrm{fb}} - k_{\mathrm{off}} \gg \alpha$, we have that the eigenvalues of the macroscopic system satisfy $\lambda_0 \ll \lambda_k < 0$ and thus $\tilde{v}_0$ is (comparatively) very stable near v^*. This implies a separation of time-scales in the problem: we expect $\tilde{v}_0$ to relax very quickly to a value near its equilibrium, whilst the other modes fluctuate stochastically on much slower time-scales. Combining these

facts suggests the following strategy: we restrict our attention to only those trajectories in which $\tilde{v}_0$ is held constant at lv^*.

Conditioning on the value of $\tilde{v}_0$ alters the structure of the noise correlation matrix for the other modes; details of the general formulation are given in Appendix B of Rogers et al. (2012b). The result is a Fokker-Planck equation for the distribution $P^{(c)}(\tilde{\boldsymbol{v}}, \tau)$ of the remaining Fourier modes $\tilde{v}_k$ with $k \neq 0$, conditioned on $\tilde{v}_0$ taking the value lv^*:

$$\frac{\partial P^{(c)}(\tilde{\boldsymbol{v}}, \tau)}{\partial \tau} = -\sum_{ki\neq 0} \frac{\partial}{\partial \tilde{v}_i} \left[\tilde{A}_k^{(c)}(\tilde{\boldsymbol{v}}) P(\tilde{\boldsymbol{v}}, \tau) \right] \tag{49}$$
$$+ \frac{1}{2V} \sum_{k,k'\neq 0} \frac{\partial}{\partial \tilde{v}_k} \frac{\partial}{\partial \tilde{v}_{k'}} \left[\tilde{B}_{k,k'}^{(c)}(\tilde{\boldsymbol{v}}) P(\tilde{\boldsymbol{v}}, \tau) \right] ,$$

where

$$\tilde{A}_k^{(c)}(\tilde{\boldsymbol{v}}) = \alpha \tilde{v}_k \left(\cos(lk) - 1 \right)$$
$$\tilde{B}_{k,k'}^{(c)}(\tilde{\boldsymbol{v}}) = 2k_{\text{off}} \left[l \tilde{v}_{k+k'} - \frac{\tilde{v}_k \tilde{v}_{k'}}{v^*} \right] . \tag{50}$$

Multiplying (50) by $\tilde{v}_k$ and integrating over all $\tilde{\boldsymbol{v}}$ we obtain a differential equation for the mode averages:

$$\frac{d}{d\tau} \langle \tilde{v}_k \rangle = \alpha \langle \tilde{v}_k \rangle \left(\cos(lk) - 1 \right) . \tag{51}$$

Thus, for all $\alpha > 0$ we have

$$\langle \tilde{v}_k \rangle \to \delta_{k,0} \, lv^* . \tag{52}$$

For the second-order moments the behaviour is not so trivial. Multiplying (50) this time by $\tilde{v}_k \tilde{v}_{k'}$ and integrating yields

$$\frac{d}{d\tau} \langle \tilde{v}_k \tilde{v}_{k'} \rangle = \frac{lk_{\text{off}}}{V} \langle \tilde{v}_{k+k'} \rangle + \langle \tilde{v}_k \tilde{v}_{k'} \rangle \left[\alpha \left(\cos(lk) + \cos(lk') - 2 \right) - \frac{1}{V} \frac{k_{\text{off}}}{v^*} \right] . \tag{53}$$

The equilibrium values are thus $\langle \tilde{v}_k \tilde{v}_{k'} \rangle \to 0$ if $k + k' \neq 0 \mod \Omega$, and

$$\langle |\tilde{v}_k|^2 \rangle \to \frac{(lv^*)^2}{1 + \alpha V (1 - \cos(lk)) v^* / k_{\text{off}}} . \tag{54}$$

This is our main result, although further analysis is required to interpret the implications for the behaviour of the model.

In Altschuler et al. (2008) it was observed that simulations of a continuum version of this model exhibited the curious phenomenon of the membrane molecules grouping together, despite the macroscopic equations suggesting they should be spread uniformly around the membrane. We introduce a summary statistic to measure this effect. Suppose the membrane molecules are distributed according to an angular density field $v(x)$, and let Λ denote the mean angular separation of two molecules:

$$\Lambda = \left(\frac{1}{v^*} \right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} d(x, y) \langle v(x) v(y) \rangle \, dx \, dy , \tag{55}$$

where

$$d(x,y) = \begin{cases} |x-y| & \text{if} \quad |x-y| < \pi \\ 2\pi - |x-y| & \text{otherwise.} \end{cases} \tag{56}$$

To compute Λ from our result (54) we pass to the continuum limit, taking $\Omega \to \infty$ and using the modes $\tilde{v}_k$ as the coefficients of the Fourier series of the membrane angular density field $v(x)$. Taking an angular prescription for the membrane so that $l = 2\pi/\Omega$, we renormalise by a factor of $1/2\pi l$ and reverse the Fourier transform to obtain

$$v(x) = \lim_{\Omega \to \infty} \frac{1}{2\pi l} \sum_{k=1}^{\Omega} e^{ikx} \tilde{v}_k, \quad \text{for } x \in [-\pi, \pi). \tag{57}$$

The calculation proceeds thus:

$$\begin{aligned} \Lambda &= 2\pi \left(\frac{1}{v^*}\right)^2 \int_{-\pi}^{\pi} |x| \langle v(x)v(0) \rangle \, dx \\ &= \frac{1}{2\pi} \left(\frac{1}{lv^*}\right)^2 \sum_{k,k'} \langle \tilde{v}_k \tilde{v}_{k'} \rangle \int_{-\pi}^{\pi} |x| e^{ikx} \, dx \\ &= \frac{\pi}{2} + \sum_{k \neq 0} \frac{1}{1 + \varphi^2 k^2} \int_{-\pi}^{\pi} |x| e^{ikx} \, dx, \end{aligned} \tag{58}$$

where $\varphi = (\alpha V l^2 v^* / k_{\text{off}})^{1/2}$, which we assume to have a finite value in the limit $l \to 0$. The first equality above comes from the rotational invariance of the model meaning that we may fix $y = 0$, after which we employ Eq. (57) and Eq. (54) in turn. Now, for $k \neq 0$

$$\int_{-\pi}^{\pi} |x| e^{ikx} \, dx = \begin{cases} -4k^{-2} & \text{for odd } k \\ 0 & \text{for even } k. \end{cases}$$

Also, the following infinite series (Bromwich, 1926) will be useful:

$$\sum_{k \text{ odd}} \frac{1}{x^2 + k^2} = \frac{\pi}{2x} \tanh\left(\frac{\pi x}{2}\right).$$

Altogether, we have

$$\begin{aligned} \Lambda(\varphi) &= \frac{\pi}{2} - \frac{2}{\pi} \sum_{k \text{ odd}} \left(\frac{1}{1 + \varphi^2 k^2}\right) \left(\frac{1}{k^2}\right) \\ &= \frac{\pi}{2} - \frac{2}{\pi} \left(\sum_{k \text{ odd}} \frac{1}{k^2} - \sum_{k \text{ odd}} \frac{1}{\varphi^{-2} + k^2}\right) \\ &= \varphi \tanh\left(\frac{\pi}{2\varphi}\right). \end{aligned} \tag{59}$$

The limits are $\Lambda(\infty) = \pi/2$, which corresponds to molecules spread uniformly around the membrane, and $\Lambda(0) = 0$, which corresponds to complete localisation. In figure 6 we compare
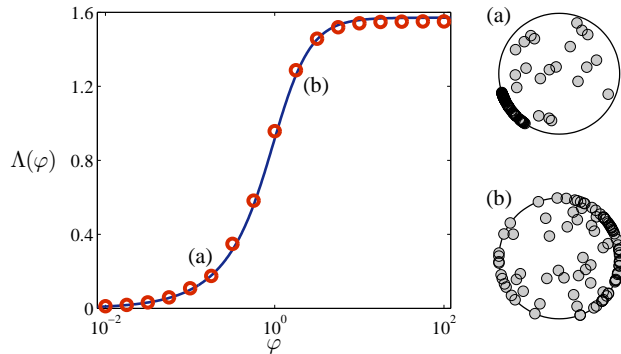
Figure 6: (Colour online) Mean angular distance between membrane molecules, as φ is varied over several orders of magnitude. Red circles are the results of simulations, the blue line shows the result from Eq. (59). On the right are two snapshots from simulations, with φ corresponding to the points (a) and (b) in the main figure. Clearly (a) is polarised and (b) is not, as predicted by the theory.

this theoretical prediction with the result of simulations for φ varying over several orders of magnitude. For small values of φ the membrane molecules cluster into a tight group, meaning that the cell has become polarised. As φ is increased (caused by an over-abundance of signalling molecules in relation to their rate of diffusion around the membrane) this effect is weakened, and the cell loses polarity.

Finally, we discuss the physical meaning of the parameter φ. The average number of molecules on the membrane at equilibrium is Vv^*, the typical lifetime of a membrane molecule is $1/k_{\mathrm{off}}$, and the continuum diffusion coefficient is αl^2. The quantity φ can thus be interpreted as the mean total distance travelled by all the membrane molecules during their lifetime: if this quantity is small, they must remain localised.

6. Discussion and Conclusion

The main aim of this paper has been to show that the analysis of stochastic models in biology need not only be numerical: a range of analytical techniques are available, just as for deterministic models. In fact the treatment of stochastic models may be simpler, since in many cases the noise can be considered to be a linear perturbation (the LNA) to the deterministic form of the dynamical equations. Linear equations such as these are easier to solve, especially when the fluctuations are around a stationary state.

The deterministic, or macroscopic, description of the system is valid when the individual nature of the constituents is unimportant, for example when they are so numerous as to effectively be infinite in number. The ensemble average of the stochastic variables will also obey the same deterministic equation. The general form of this equation is Eq. (17), that is, $\dot{y}_I = A_I(\boldsymbol{y})$, where the dot denotes a time derivative and the index I includes both a position label and another identifying the constituent type. The function A_I can be calculated from the elementary reactions of the process using Eq. (18). The mesoscopic, or stochastic, description of the system which uses the same variables as the macroscopic description, is in principle no different. The general form of this equation is (23), that

is, $\dot{y}_I = A_I(\boldsymbol{y}) + V^{-1/2}\sum_J g_{IJ}(\boldsymbol{y})\eta_J$, where η_J is a Gaussian white noise with zero mean and unit strength. The only additional function which appears over and above that in the deterministic equation is g_{IJ}, which can, like A_I, be calculated from the elementary reactions of the process using Eqs. (21) and (25). Although Eq. (23) is a straightforward generalisation of Eq. (17), it is much less well-known.

There are several reasons for the perceived difficulty of using Eq. (23), probably the most important being the unfamiliarity of many biologists with the general theory of stochastic processes. We have tried to show in this paper that the stochastic theory which is required need not be more complicated than that of dynamical systems theory, which is applicable to equations such as (17). This is especially true if the LNA is a valid approximation for the particular system under study. If the multiplicative nature of the noise cannot be neglected, as in section 5, then care is required because of the singular nature of white noise. However, even in this case, a systematic theory has been developed that may be applicable in situations in which there is a separation of time-scales (Rogers et al., 2012a; Biancalani et al., 2012; Rogers et al., 2012b).

We applied Eq. (23) to two sets of processes, one for which the LNA was applicable and one for which it was not. The former situation was discussed in section 4 where we revisited the problem of the emergence of spatial structures for systems of populations, in the paradigmatic example of the Brusselator model. Intrinsic fluctuations, which are intuitively thought of as a disturbing source, appear instead to be critical for the emergence of spatial order. More specifically, we showed how Turing patterns can arise for parameter values for which the macroscopic equations predict relaxation to a homogeneous state. We called these patterns 'stochastic patterns', as they are generated by the continuous action of noise present in the system. However, it can be argued that the amplitude of stochastic patterns might be so small that they can hardly be observed in a real population, given that the amplitude of the noise is small as well. Whilst this might be true for some systems, a recent study (Ridolfi et al., 2011a) has suggested that the response to a small perturbation in a pattern-forming system can be unexpectedly large, if the system displays a sufficient degree of 'non-normality'. The connection between non-normality and stochastic patterns is so far largely unexplored, and constitutes a possible further investigation in this line of research.

In section 5 we discussed an example of a stochastic phenomenon which goes beyond what can be understood within the LNA. The stochastic patterns appearing in the Brusselator are noise-driven perturbations around the homogeneous state, having characteristic magnitude $1/\sqrt{V}$ (and thus disappearing in the limit $V \to \infty$). By contrast, the spontaneous emergence of cell polarity in the model of Altschuler *et al* requires the noise to have a more complex structure, which can lead the system to a state very far removed from the homogeneous fixed point of the deterministic equations. To characterise this process, it was necessary to study the full effect of the non-linear terms in the mesoscopic equations. To achieve this, we exploited the natural separation of time-scales occurring between the dynamics of the zeroth Fourier mode (which relaxes quickly to its equilibrium value) and the remaining degrees of freedom. This is a non-standard technique, however, it can be made relatively systematic and general, as will be outlined in a forthcoming paper. The LNA has played an important role in boosting the recognition of the importance of stochastic effects in the literature; we

hope that methods employing the separation of time-scales may provide the next theoretical advance.

References

Altschuler, S.J., Angenent, S.B., Wang, Y., Wu, L.F., 2008. On the spontaneous emergence of cell polarity. Nature 454, 886–889.

Biancalani, T., Fanelli, D., Di Patti, F., 2010. Stochastic Turing patterns in the Brusselator model. Phys. Rev. E 81, 046215. doi:`10.1103/PhysRevE.81.046215`.

Biancalani, T., Galla, T., McKane, A.J., 2011. Stochastic waves in a Brusselator model with nonlocal interaction. Phys. Rev. E 84, 026201. doi:`10.1103/PhysRevE.84.026201`.

Biancalani, T., Rogers, T., McKane, A.J., 2012. Noise-induced metastability in biochemical networks. Phys. Rev. E 86, 010106(R). doi:`10.1103/PhysRevE.86.010106`.

Black, A.J., McKane, A.J., 2012. Stochastic formulation of ecological models and their applications. Trends Ecol. Evol. 27, 337–345. doi:`10.1016/j.tree.2012.01.014`.

Boland, R.P., Galla, T., McKane, A.J., 2009. Limit cycles, complex Floquet multipliers and intrinsic noise. Phys. Rev. E. 79, 051131.

Bonachela, J.A., Munoz, M.A., Levin, S.A., 2012. Patchiness and demographic noise in three ecological examples. J. Stat. Phys. 148, 723–739.

Bromwich, T., 1926. An Introduction to the Theory of Infinite Series. Chelsea Publishing Company, London.

Butler, T.C., Benayounc, M., Wallace, E., van Drongelenc, W., Goldenfeld, N., Cowane, J., 2012. Evolutionary constraints on visual cortex architecture from the dynamics of hallucinations. PNAS 109, 606–609. doi:`10.1073/pnas.1118672109`.

Butler, T.C., Goldenfeld, N., 2009. Robust ecological pattern formation induced by demographic noise. Phys. Rev. E 80, 030902(R). doi:`10.1103/PhysRevE.80.030902`.

Butler, T.C., Goldenfeld, N., 2011. Fluctuation-driven Turing patterns. Phys. Rev. E 84, 011112. doi:`10.1103/PhysRevE.84.011112`.

Chaikin, P.M., Lubensky, T.C., 2000. Principles of Condensed Matter Physics. Third ed., Cambridge University Press, Cambridge.

Cross, M.C., Greenside, H.S., 2009. Pattern Formation and Dynamics in Non-Equilibrium Systems. Cambridge University Press, New York.

Datta, S., Delius, G.W., Law, R., 2010. A jump-growth model for predator-prey dynamics: derivation and application to marine ecosystems. Bull. Math. Biol. 72, 1361–1382. doi:`10.1007/s11538-009-9496-5`.

Gardiner, C.W., 2009. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. Fourth ed., Springer, New York.

Gillespie, D.T., 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. 22, 403–434.

Gillespie, D.T., 1977. Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. 81, 2340–2361.

Glansdorff, P., Prigogine, I., 1971. Thermodynamic Theory of Structure, Stability and Fluctuations. Wiley-Interscience, Chichester.

Gupta, A., 2012. Stochastic model for cell polarity. Annals of Applied Probability 22, 827–859.

van Kampen, N.G., 2007. Stochastic Processes in Physics and Chemistry. Third ed., Elsevier Science, Amsterdam.

Lugo, C.A., McKane, A.J., 2008. Quasi-cycles in a spatial predator-prey model. Phys. Rev. E. 78, 051911.

Mehta, M.L., 1989. Matrix Theory. Hindustan Publishing Corporation, India.

Murray, J.D., 2008. Mathematical Biology vol. II. Third ed., Springer-Verlag, Berlin.

Petzold et al, 2012. Submitted .

Ridolfi, L., Camporeale, C., D'Odorico, P., Laio, F., 2011a. Transient growth induces unexpected deterministic spatial patterns in the Turing process. Europhys. Lett. 95, 18003. doi:`10.1209/0295-5075/95/18003`.

Ridolfi, L., D'Odorico, P., Laio, F., 2011b. Noise-Induced Phenomena in the Environmental Sciences. Cambridge University Press, Cambridge.

Risken, H., 1989. The Fokker-Planck Equation - Methods of Solution and Applications. Second ed., Springer, Berlin.

Rogers, T., McKane, A.J., Rossberg, A.G., 2012a. Demographic noise can lead to the spontaneous formation of species. Europhys. Lett. 97, 40008. doi:`10.1209/0295-5075/97/40008`.

Rogers, T., McKane, A.J., Rossberg, A.G., 2012b. Spontaneous genetic clustering in populations of competing organisms. To appear in Physical Biology .

Scott, M., Poulin, F.J., Tang, H., 2011. Approximating intrinsic noise in continuous multi-species models. Proc. R. Soc. A (London) 467, 718–737. doi:`10.1098/rspa.2010.0275`.

Turing, A.M., 1952. The Chemical Basis of Morphogenesis. Phil. Trans. R. Soc. B (London) 237, 37–72. doi:10.1098/rstb.1952.0012.

Wiggins, S., 2003. Introduction to Applied Nonlinear Dynamical Systems and Chaos. Second ed., Springer, Berlin.

Woolley, T.E., Baker, R.E., Gaffney, E.A., Maini, P.K., 2011. Stochastic reaction and diffusion on growing domains: Understanding the breakdown of robust pattern formation. Phys. Rev. E 84, 046216. doi:10.1103/PhysRevE.84.046216.