Spatial-Stochastic Modelling of Synthetic Gene Regulatory Networks

Cicely K Macnamara[a], Elaine I Mitchell[b], Mark AJ Chaplain[a,*]

[a]*School of Mathematics and Statistics, Mathematical Institute, University of St Andrews, United Kingdom, KY16 9SS*
[b]*Division of Mathematics, University of Dundee, Dundee DD1 4HN, Scotland*

Abstract

Transcription factors are important molecules which control the levels of mRNA and proteins within cells by modulating the process of transcription (the mechanism by which mRNA is produced within cells) and hence translation (the mechanism by which proteins are produced within cells). Transcription factors are part of a wider family of molecular interaction networks known as gene regulatory networks (GRNs) which play an important role in key cellular processes such as cell division and apoptosis (e.g. the p53-Mdm2, NFκB pathways). Transcription factors exert control over molecular levels through feedback mechanisms, with proteins binding to gene sites in the nucleus and either up-regulating or down-regulating production of mRNA. In many GRNs, there is a negative feedback in the network and the transcription rate is reduced. Typically, this leads to the mRNA and protein

*Corresponding author.
Phone: +44 (0)1334 463723

Email addresses: ckm@st-andrews.ac.uk (Cicely K Macnamara), emitchell@maths.dundee.ac.uk (Elaine I Mitchell), majc@st-andrews.ac.uk (Mark AJ Chaplain)

Author accepted manuscript

levels oscillating over time and also spatially between the nucleus and cytoplasm. When experimental data for such systems is analysed, it is observed to be noisy and in many cases the actual numbers of molecules involved are quite low. In order to model such systems accurately and connect with the data in a quantitative way, it is therefore necessary to adopt a stochastic approach as well as take into account the spatial aspect of the problem. In this paper, we extend previous work in the area by formulating and analysing stochastic spatio-temporal models of synthetic GRNs e.g. repressilators and activator-repressor systems.

Keywords: synthetic gene regulatory networks, repressilators, activator-repressor systems, spatial-stochastic modelling

---

1. Introduction

Cellular processes (e.g. cell division, apoptosis and adhesion) are governed by a cell's DNA through interactions of RNA and protein. Any subset of these products and their interactions can be thought of as a network and are customarily called gene regulatory networks (GRNs). A specific group of proteins, called transcription factors, are a common feature of GRNs: in response to signals or stimuli they alter the transcription rate of genes in order to affect protein levels. Such GRNs typically employ feedback mechanisms; for example, when a protein represses the transcription rate of its own mRNA there is said to be negative feedback. Negative feedback loops typically lead to fluctuating levels of protein and are implemented in many different biological processes (e.g. inflammation, meiosis, apoptosis and the heat shock response, Lahav et al., 2004). Within the emerging field of

synthetic biology, GRNs are of particular interest and following the ground-breaking work of Becskei and Serrano (2000) and Elowitz and Leibler (2000), they have been investigated both from a practical, experimental (Balagadde et al., 2008; Chen et al., 2012; Yordanov et al., 2014) and a theoretical modelling viewpoint (Purcell et al., 2010; O'Brien et al., 2012).

Interest in the mathematical modelling of GRNs first began in the 1960s (Goodwin, 1965; Griffith, 1968). In these papers ODE models of a simple mRNA-protein feedback system were given and analysed for oscillatory behaviour. In such "closed-loop" negative feedback systems the protein inhibits the production of its own mRNA; intuitively one would expect this to lead to fluctuating levels of both molecules, however, periodic behaviour could not be derived. More recently authors have introduced delay mechanisms into differential equations in order to achieve periodic fluctuations in the mRNA and protein concentrations. These models considered either generic or "synthetic" GRN systems (Mackey and Glass, 1977; Smolen et al., 1999, 2001, 2002; Purcell et al., 2010; O'Brien et al., 2012) or models of actual biological pathways e.g. the Hes1 system, the p53-Mdm2 system and the NF-κB system (Tiana et al., 2002; Jensen et al., 2003; Lewis, 2003; Monk, 2003; Bernard et al., 2006).

The first spatial models of generic intracellular systems were developed in the 1970s and 1980s (Glass and Kauffman, 1970; Shymko and Glass, 1974; Busenberg and Mahaffy, 1985; Mahaffy, 1988; Mahaffy and Pao, 1984). One-dimensional reaction-diffusion PDEs were designed and examined through steady states and stability analysis. The geometry of the system was shown to be important and the term "spatial switching" was introduced to refer

to the fact that the system geometry can lead to different dynamical behaviour. Such an approach has more recently been adopted and extended by Naqib et al. (2012). Certain models have incorporated spatial aspects by introducing compartments to account for the fact that different cellular processes occur in different locations within a cell, for example mRNA is produced in the nucleus and then translates into protein in the cytoplasm (e.g. Momiji and Monk, 2008; Cangiani and Natalini, 2010; Sturrock et al., 2011, 2012). More recent explicit spatial models include those of Dimitrio et al. (2013); Eliaš and Clairambault (2014); Eliaš et al. (2014a,b); Szymańska et al. (2014). Spatial modelling has provided insight into the importance of spatial aspects in deriving periodically fluctuating mRNA and protein behaviour, and recently Chaplain et al. (2015) proved rigorously that molecular diffusion causes oscillations in the Hes1 system.

In order that such continuum approaches be valid, however, it is assumed that the molecular numbers of each species are high enough such that they could be reduced to concentrations. In reality regulator numbers of both mRNA and transcription factors are low, and as such a deterministic differential equation approach (ODE or PDE) is not the most appropriate in order to capture the effects of stochasticity in a single cell. In this paper, then, we consider the more biologically relevant discrete, spatial-stochastic approach derived from the spatial-stochastic model of the Hes1 GRN put forward by Sturrock et al. (2013). PDE models for repressilators and activator-inhibitors showed that oscillations may be achieved provided the relationship between the spatial location of the gene site and diffusion coefficient is optimised (Macnamara and Chaplain, 2016). We will investigate similar themes here,

discussing how spatio-temporal dynamics change as we vary the location of the gene site(s) and the diffusion coefficient of the mRNA and protein species. Note the term repressilator (introduced by Elowitz and Leibler, 2000) has historically been reserved for a system of three genes which couple to form a cycle of negative feedback, however, for ease of reference we choose to use this terminology, for any n-gene system for which the protein of any given gene inhibits the production of the mRNA for the subsequent gene. According to our terminology activator-repressor systems couple positive and negative feedback.

The paper is structured as follows. In Section 2 we layout the specific model(s) to be investigated and give details of how simulations are carried out. In Section 3 we provide results for repressilator systems; first revisiting the Hes1 system, or one-gene repressilator, (as detailed by Sturrock et al., 2013) to discuss how changes to spatial aspects affect the molecular dynamics and then extending the approach to a two-gene repressilator system. In Section 4 we present simulation results for a two-gene activator-repressor system which contains both positive and negative feedback. Discussions, conclusions and directions for future work in this area are given in Section 5.

2. Model

Throughout this paper we investigate synthetic gene regulatory network systems using stochastic reaction-diffusion models. These models are based on one given for the Hes1 system by Sturrock et al. (2013). Details of this type of modelling have been given both by Sturrock et al. (2013) and in Szymanska et al. (2018), for example, but we formulate the general model

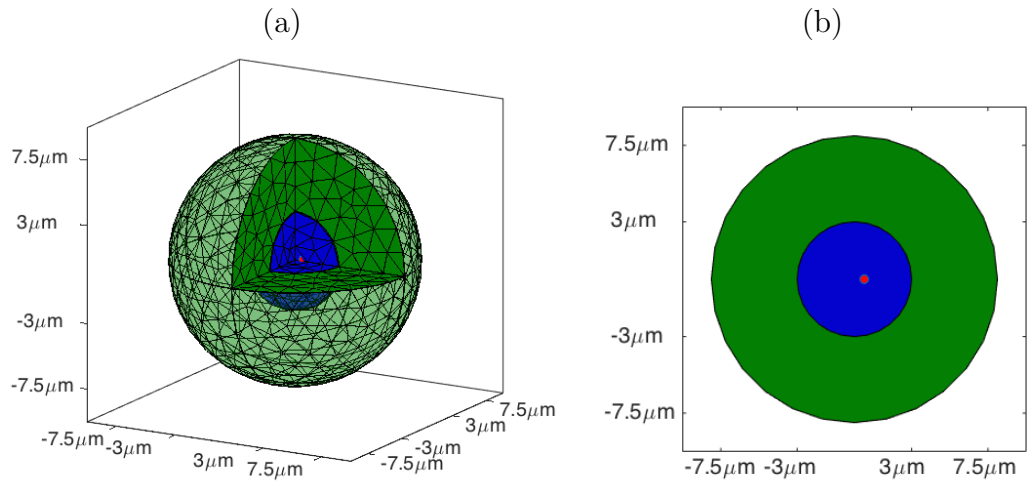

Figure 1: (a) computational 3D cellular domain with imposed tetrahedral mesh, (b) a cross-section of the geometry. The cell has radius 7.5μm comprised of a central nucleus (blue) of radius 3μm and cytoplasm (green). This domain will be used in the spatial-stochastic simulations with a gene site(s) defined somewhere within the nucleus, for example, at the point indicated in red.

fully here for ease of reference for the reader. The biochemical reaction and diffusion processess occur within a 3D spherical domain, as depicted in Figure 1(a). We approximate the cell as two concentric spheres centred on the origin corresponding to the nucleus (blue in Figure 1) and cytoplasm (green in Figure 1) with radii 7.5μm and 3μm, respectively. The computational domain and imposed tetrahedral mesh is generated using COMSOL; note we specify that the maximum mesh element size is 0.8μm. Chemical reactions, at the voxel level, or the movement (diffusion jumps) of a molecule, between neighbouring voxels, govern the changes to the state of the system between

time steps.

2.1. Biochemical Reactions

We consider three distinct types of molecular species: mRNAs, proteins and promoters which interact via reactions within each voxel of our domain; different reactions occur in different voxels depending on where in the domain a voxel is located. The reactions for our models are shown in Table 1, for i cyclically coupled genes where $i = \{1, 2, 3, ...n\}$ mod n. If a molecule of mRNA of a gene i, m_i, enters any voxel in the cytoplasm it may be translated producing protein of that same gene, p_i, at a rate α_p. A promoter for each gene i sits within a single voxel located within the nucleus (for a defined gene site point - illustrated, for example, by the red point in Figure 1, we select the whole voxel which contains it). Transcription of mRNA, m_i, occurs within this ith gene site voxel; the rate of transcription is affected by feedback from the protein of the proceeding gene in the cycle, i.e. p_{i-1}. If a molecule of protein p_{i-1} enters the ith gene-site voxel it may bind to the (free) promoter, fp_i, at a rate k_1, occupying the promoter, becoming op_i; the opposite reaction takes place with the protein uncoupling from the promoter at a rate k_2. When the promoter is free, mRNA is produced at the baseline rate α_m. When a promoter is occupied, the rate of mRNA production is affected by the factor $1/\gamma_i$; for cases of repression, $\gamma_i > 1$, in order that the production of mRNA is reduced from its baseline value; for activation, $0 < \gamma_i < 1$ increasing mRNA production from this baseline value. To complete the system of reactions we consider that mRNA and protein molecules are removed from any voxel within the domain (i.e. degrade) at rates μ_m and μ_p, respectively. The initial rate constants used in simulations for re-

pressilators are given in Table 1; these are taken as values consistent with the robust parameter regime determined by Sturrock et al. (2013). Note that when $n = 1$, this system of equations is exactly that of the Hes1 system, as given in Sturrock et al. (2013).

2.2. Molecular Diffusion

Molecular diffusion is prescribed by the movement of mRNA and protein molecules between voxels, i.e. from a voxel, ψ_j, to a randomly selected adjacent voxel, ψ_k. We model it as a first-order event and treat the diffusive process in much the same way as the above reactions. Specifically, we consider

$$\mathrm{S_{ij}} \xrightarrow{d_{ijk}} \mathrm{S_{ik}} \tag{1}$$

where S_{ij} denotes a species (either mRNA or protein) of gene i located in voxel ψ_j. Hence diffusion is governed by a "jump" rate constants, d_{ijk}, which depends on the macroscopic diffusion coefficient, D, and the shape and size of voxels ψ_j and ψ_k. Note $d_{ijk} = 0$ for unconnected mesh elements since molecules can only "jump" between neighbouring voxels. Gene site species i.e. the free and occupied promoters, fp_i and op_i are confined to their gene site voxels, and are thus given a diffusion coefficient of zero. For the purposes of our numerical investigations, we will assume that the diffusion rates of all mRNA and protein species are the same, and will be denoted by D. However, this value will be varied in certain computational simulations. The biochemical reactions and diffusion jumps are governed by a reaction-diffusion master equation, for the full mathematical formalism of this please see Appendix A.

Cytoplasmic Reaction	Description	Parameter value
$m_i \xrightarrow{\alpha_p} m_i + p_i$	translation of protein	$\alpha_p = 3\text{min}^{-1}$
ith gene site Reactions	Description	Parameter value
$fp_i + p_{i-1} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} op_i$	binding/unbinding of protein with the ith promoter	$k_1 = 1 \times 10^8\text{M}^{-1}\text{min}^{-1}$ $k_2 = 0.05\text{min}^{-1}$
$fp_i \xrightarrow{\alpha_m} fp_i + m_i$	basal transcription of mRNA	$\alpha_m = 3\text{min}^{-1}$
$op_i \xrightarrow{\alpha_m/\gamma_i} op_i + m_i$	altered transcription of mRNA	$\alpha_m = 3\text{min}^{-1}$ e.g. $\gamma_i = 1000$ (repression)
Global Reactions	Description	Parameter value
$m_i \xrightarrow{\mu_m} \emptyset$	degradation of mRNA	$\mu_m = 0.06\text{min}^{-1}$
$p_i \xrightarrow{\mu_p} \emptyset$	degradation of protein	$\mu_p = 0.03\text{min}^{-1}$

Table 1: The reaction processes and their accompanying parameter values used throughout this investigation. The colours indicate where in the domain the reactions take place - green in the cytoplasm, red within the promoter voxel and black globally.

2.3. Simulations

At any given time the state of the system is described by the number of each chemical species within the domain. Changes to the state will either be by the chemical reactions at the voxel level or the movement (diffusion jumps) of a molecule between neighbouring voxels - see Table 1. The temporal evolution of the probability distribution of each state in the state space is governed by the reaction diffusion master equation (RDME) - see Appendix A. We complete the model set-up with zero-flux boundary conditions at the cell membrane, while we impose continuity of flux on the nuclear membrane. Note, in reality nucleocytoplasmic transport is a complex process; the translocation of proteins from the cytoplasm to the nucleus, for example, requires proteins to bind with importins to navigate the nuclear pore complex. In this paper, however, we limit transport by diffusion and require that proteins must first make it to the voxel containing the promoter in order to start the transcription process. The model can easily be extended to include this process and has been done for the deterministic case (see, Sturrock et al., 2011). For initialisation, we suppose that there is only a single free promoter within each gene/promoter voxel.

We solve the spatial-stochastic system on the 3D domain given in Figure 1, through use of the URDME (Unstructured-mesh Reaction-Diffusion Master Equation) software framework. URDME is implemented through a Matlab interface which couples the RDME (with reaction propensities written in an ANSI C file) to the geometry and tetrahedral mesh created by the finite-element package COMSOL. COMSOL determines the diffusion rates d_{ijk} for

each species in each voxel. URDME uses a computational solver which is an efficient implementation of the next subvolume method, NSM (Gibson and Bruck, 2000). For a precise description of the URDME framework and how it is implemented we refer the reader to the original articles where this software is first described (Cullhed et al., 2008; Engblom et al., 2009; Drawert et al., 2012).

3. Repressilators

3.1. The Hes1 System

The Hes1 protein (a basic helix-loop-helix (bHLH) transcription factor) is a useful starting point for a simple negative feedback GRN, since it is known to repress the transcription of its own gene through direct binding to regulatory sequences in the Hes1 promoter (Hirata et al., 2002). As such we refer to it as a one-gene repressilator; a schematic of the Hes1 system is shown in Figure 2. It is known that periodically changing levels of Hes1 protein controls embryonic development, specifically in correctly timed somite segmentation (see, for example, Kageyama et al., 2007). Mathematical models have sought to reproduce this fluctuating expression. For a PDE model, Chaplain et al. (2015) rigorously proved that the diffusion parameter controls whether or not the system oscillates. Macnamara and Chaplain (2016) further indicated the importance of spatial aspects showing that the variation of molecular concentrations over time is governed by the combination of diffusion coefficients and locations of transcription and translation within a cell. The investigation of Sturrock et al. (2013) showed, using wavelet analysis, that periods comparable with those found experimentally could be discerned from the

spatial-stochastic model results. We model the Hes1 system using the reac-

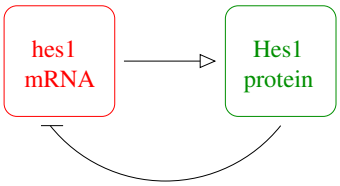

Figure 2: Simple schematic of the Hes1 gene regulatory system. Hes1 protein is produced from hes1 mRNA via translation, but inhibits the production of hes1 mRNA (represses or down-regulates transcription). The colours correspond to where the molecular reactions take place, i.e. red for at the promoter site and green in the cytoplasm.

tions given in Table 1 where $n = 1$. We take $\gamma = 1000$ which is significantly high enough that repression is extremely efficient and effectively the gene is turned off.

3.2. Varying the Diffusion Coefficients

In this Section we look at varying the diffusion coefficient, D, and present the results of the spatial-stochastic model for a one-gene repressilator (e.g. Hes1). We place the single promoter at the origin and simulate the behaviour over the spherical domain depicted in Figure 1 for 1600 minutes (capturing just over a days worth of data). The results are presented in Figure 3.

For a gene site at the origin we note that there is a minimum diffusion coefficient, $D_{syn} \gtrsim 1 \times 10^{-13}\text{m}^2\text{min}^{-1}$, such that protein can be synthesised; if the diffusion coefficient is equal to or lower than this, mRNA either fails to reach the cytoplasm or does not survive long enough once it reaches the

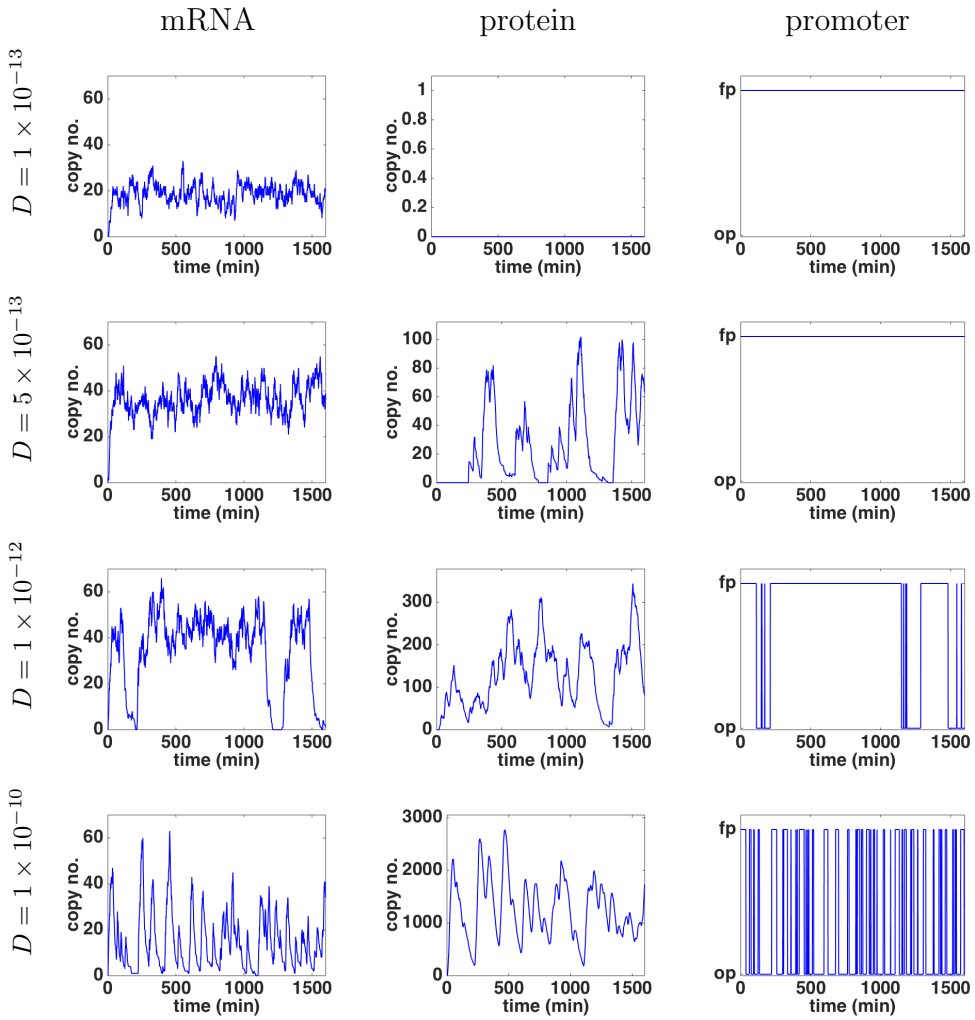

Figure 3: Simulation results for the Hes1 model showing the mRNA and protein copy numbers along with the promoter behaviour (left to right, respectively) as we vary the diffusion coefficient (top to bottom). Units of the diffusion coefficient, D, are $\mathrm{m}^2\mathrm{min}^{-1}$. The promoter is either free, fp, or occupied, op.

cytoplasm for protein to be synthesised. This is shown clearly in the top panel of Figure 3; at no time is there any protein and the promoter is always free. For a diffusion coefficient in the range, $D_{syn} < D < D_{bind}$, where $D_{bind} \gtrsim 5 \times 10^{-13}\text{m}^2\text{min}^{-1}$ mRNA survives long enough for protein to be synthesised but in turn the protein does not survive long enough to diffuse back to the gene site in order to occupy the promoter (as shown in the second panel of Figure 3; protein is produced but the promoter remains free for all time). Since the promoter always remains free, dynamics observed in the mRNA and protein levels are purely stochastic for a diffusion coefficient in this range.

For a diffusion coefficient higher than D_{bind}, protein is both synthesised and able to diffuse back to the promoter site of the gene within the nucleus and binds with the promoter such that the promoter fluctuates between being free and occupied. The resulting negative feedback from the protein-promoter binding is apparent in both the mRNA and protein levels. As we increase the diffusion coefficient there is more likelihood that a protein molecule will be available to bind with the promoter, so we see more frequent changes in the promoters occupancy (observe the differences in promoter behaviour for $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ and $D = 1 \times 10^{-10}\text{m}^2\text{min}^{-1}$). This corresponds to greater amplitude fluctuations in both mRNA and protein copy numbers. At the same time the protein exhibits higher copy numbers. For these latter regimes it is possible to investigate potential periodic behaviour, which we examine in Section 3.4. Note, it is unlikely that the diffusion rate varies over the full range of values we have selected here; however, we show the differ-

ences in behaviour from a theoretical point of view. Consistent with previous models (e.g. Sturrock et al., 2013; Macnamara and Chaplain, 2016) it is likely that a typical diffusion coefficient is approximately $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$, i.e. a diffusion coefficient which gives rise to possible periodic fluctuations. Throughout this paper we will focus on this diffusion coefficient regime, whilst also considering cases where $D = 1 \times 10^{-10}\text{m}^2\text{min}^{-1}$ for comparison. It is likely that in reality the diffusion of mRNA and protein may differ from each other. Previous work (e.g. Sturrock et al., 2011, 2013; Macnamara and Chaplain, 2016) has varied the diffusion coefficients and the underlying dynamics remain qualitatively unchanged over wide range. Furthermore, we find, for example, no difference in the dynamics and subsequent analysis if we increase the diffusion rate of protein to $D_p = 1 \times 10^{-10}\text{m}^2\text{min}^{-1}$ while keeping the diffusion rate of mRNA as $D_m = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$. As such we proceed with keeping the diffusion rates the same for both species.

3.3. Varying the Promoter Location

In this section we consider how the position of the promoter site affects the behaviour of mRNA and protein copy numbers. In Figure 4 we show the mRNA, protein and promoter behaviour for three different defined promoter locations. Specifically we place the promoter site at three locations $(p_x, 0, 0)$, where $p_x = \{0.5\mu\text{m}, 1.5\mu\text{m}, 2.5\mu\text{m}\}$, and in all cases we fix $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$. As the promoter location is moved closer to the nuclear-cytoplasm membrane the likelihood that a protein binds with the promoter increases and the promoter fluctuates more frequently between free and occupied. This leads to more frequent fluctuations in both mRNA and protein levels. However, when the promoter location is very close to the nuclear-

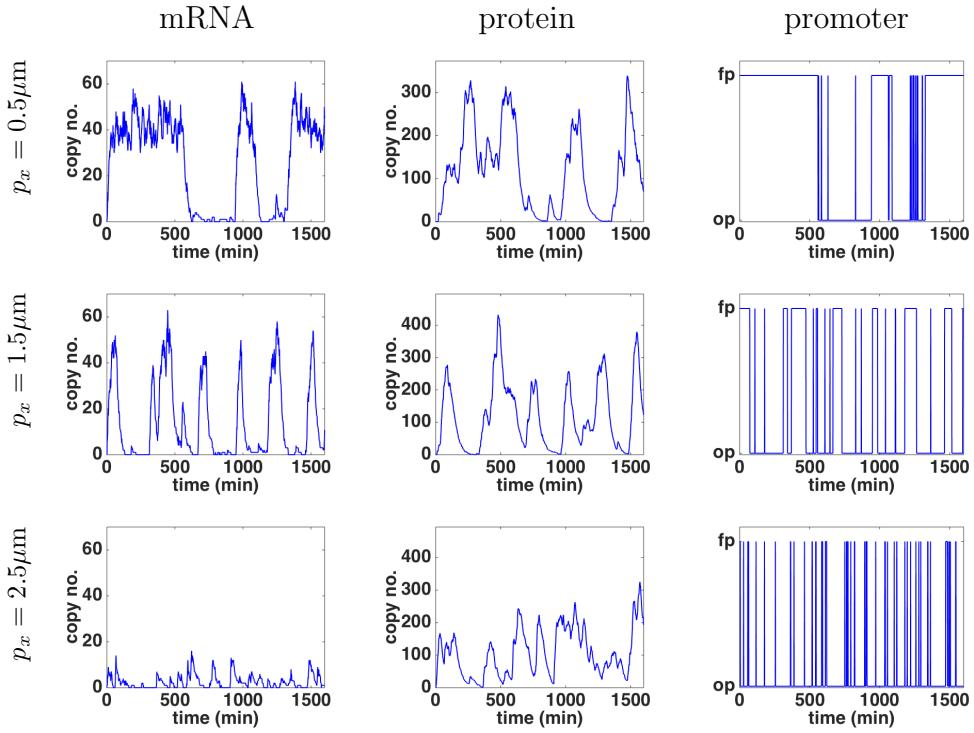

Figure 4: Simulation results for the Hes1 model showing the mRNA and protein copy numbers along with the promoter behaviour (left to right, respectively) as we vary the location of the promoter (top to bottom). The diffusion coefficient used is $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$; the promoter is placed at $(p_x, 0, 0)$, where $p_x = \{0.5\mu\text{m}, 1.5\mu\text{m}, 2.5\mu\text{m}\}$. The promoter is either free, fp, or occupied, op.

cytoplasm membrane the promoter is occupied more often than it is free. The knock-on effect is far lower copy numbers of mRNA which is repressed by the protein-promoter complex and subsequently lower protein copy numbers.

3.4. Determining Periodic Behaviour

We are interested in determining the presence of periodic behaviour in the levels of the molecular species. Following Sturrock et al. (2013) we estimate the period(s) of oscillations using a Morlet continuous time wavelet transform (CWT) as implemented by a MATLAB toolbox called WAVOS, please see Harang et al. (2012) for details. Given the highly oscillatory and noisy nature of our trajectories, the use of standard Fourier techniques can lead to inaccurate estimates of the period, as Fourier analysis assumes stationarity of the signal and that its basis functions are unbounded in time (Mallat, 1998). Wavelets, in contrast, are localised in both time and frequency. This localises the analysis, allowing the changes in signal properties to be tracked over time (Torrence and Compo, 1998). The instantaneous period is calculated for each time step and so varies for a single simulation, we make use of gaussian edge elimination to minimise artefacts in the approximation of the period. For example, in Figure B.22, to be found in Appendix B we give the trajectories of five individual simulations of the Hes1 model with $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ and the promoter site located at the origin; we show the behaviour of the mRNA and protein copy numbers and the instantaneous period calculated from the protein trajectory in each case. We also show on the plots of the instantaneous periods lines which indicate the mean and mode of the period data. As can be seen in Figure B.22 in several cases neither provides a good representation of the behaviour of the period data; potentially no true period found for a trajectory may lie on the mean period line and the mode may not capture the full profile of the data in which there may be more than one dominating period. In the analysis which follows we

refer to period "modes" which are taken to be statistically significant periods which dominate the data. To calculate these period modes we group the period data into 5 minute intervals and find the proportion of periods found in each - we then use the MATLAB function `findpeaks` to determine maxima of this derived data, we stipulate that a maxima should be at least 5% of the data series and that maxima must be separated by at least 45 minutes. For any given simulation we observe that there may be any number from a single period mode to four distinct period modes. In Figure 5 we determine, for each of 100 runs of the simulation for both $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ and $D = 1 \times 10^{-10}\text{m}^2\text{min}^{-1}$, how many period modes there are.

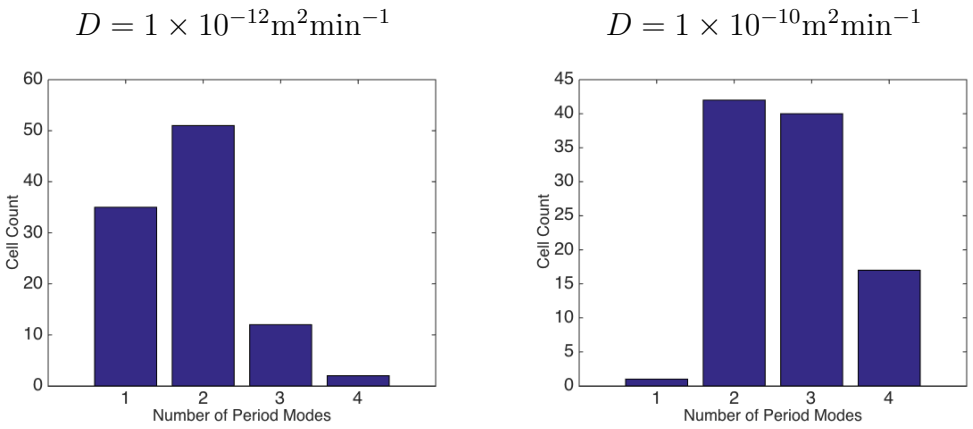

Figure 5: Histogram indicating how the number of period modes determined for 100 simulations of the Hes1 model for different diffusion coefficient regimes.

In Figure 6 we display the results from 50 runs of the simulation (for both $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ and $D = 1 \times 10^{-10}\text{m}^2\text{min}^{-1}$); we show the range

of periods detected (black lines), the mean of those detected periods (red squares) and any period modes (blue diamonds). The majority of detected periods are found to lie between 100 and 500 minutes. Note that the horizontal dashed line indicates the maximum period which can be ascertained from the time series data; it may be the case that when only this maximum period is discovered (e.g. in simulation 11) the WAVOS tool has failed to detect a meaningful period, periods which fall on this line must be treated with caution. For $D = 1 \times 10^{-10}\text{m}^2\text{min}^{-1}$ we typically observe a far greater range of periods; equally, it is more likely that there will be two or more distinct period modes. However, a high number of period modes sit on the maximum line, possibly for such simulations longer periods could be detected if we increased the number of time steps, however, since the Hes1 system typically displays periods of around two hours (Hirata et al., 2002) such longer periods are biologically unrealistic and irrelevant to this current investigation. Increasing the diffusion coefficient enables mRNA to travel to the cytoplasm more readily and hence for protein synthesis to occur at a higher rate (compare the protein behaviour in Figure 3). In turn this enables protein to travel to the nucleus, consequently, there will be more protein available to bind to the promoter, reflected in the significantly higher rate of binding of protein to the promoter (compare the promoter behaviour in Figure 3) switching it between free and occupied. This increased fluctuation turning repression on and off frequently leads to more complex periodic behaviour in the protein copy numbers, with bi-modal and multi-modal periods detected.

We repeated this period analysis with $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ and the pro-

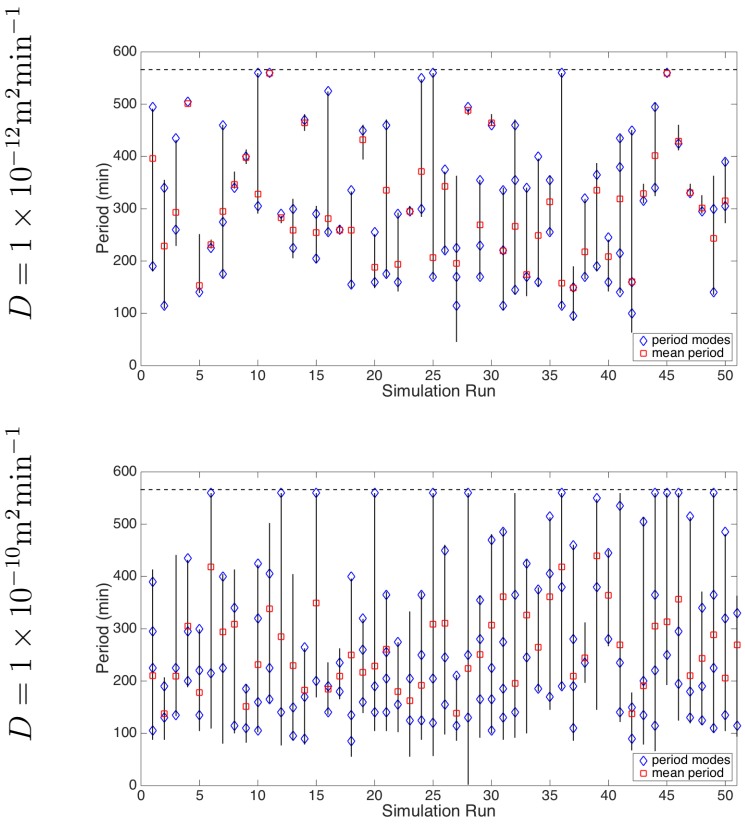

Figure 6: Period data for 50 simulations of the Hes1 model for different diffusion coefficient regimes. The red squares are the mean periods, the black lines indicate the range of periods, the blue diamonds indicate the value of the period mode(s).

moter site located at $(2.5\mu\mathrm{m}, 0, 0)$, the results are shown in Figure 7. We typically observe shorter mean periods dominating, with the majority of detected periods laying in the range 100 to 400 minutes. In addition, for this regime, typically more than one period mode is detected. Moving the promoter closer to the nucleus-cytoplasm membrane leads to increased fluctuations in

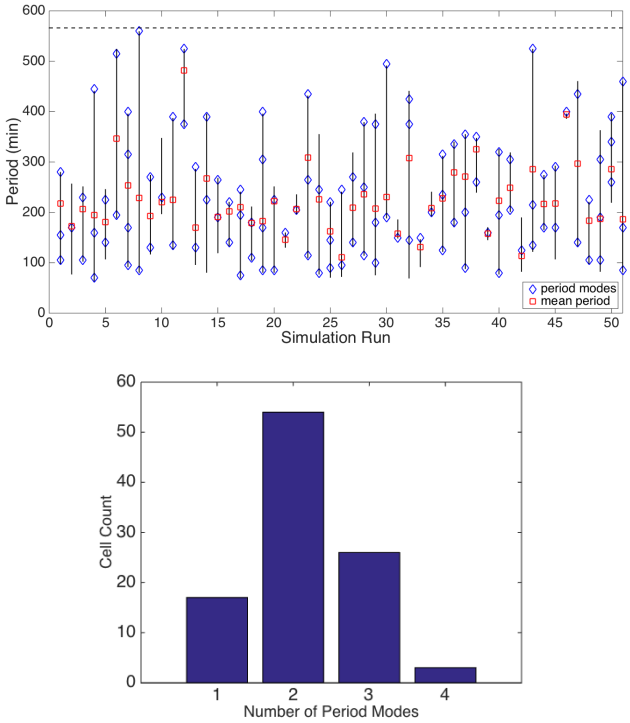

Figure 7: Top plot: period data for 50 simulations of the Hes1 model: the red squares are the mean periods, the black lines indicate the range of periods, the blue diamonds indicate the value of the period mode(s). Bottom plot: histogram indicating the number of period modes determined for 100 simulations. $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ and the promoter is located at $(2.5\mu\text{m}, 0, 0)$.

the promoter status turning repression on and off frequently and resulting in more complex periodic behaviour in the protein copy numbers. However, when the promoter is very close to the membrane the frequently occupied promoter serves to provoke more rapid mRNA and protein fluctuations and hence shorter periods.

3.5. *Two-gene Repressilator*

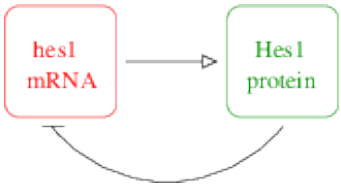

Figure 8: Simple schematic of the two-gene repressilator system. The mRNA of each species produces its own protein. The protein of one species inhibits the production of mRNA of the other species. The colours correspond to where the molecular reactions take place, i.e. red for at the promoter site and green in the cytoplasm.

We extend the work of Sturrock et al. (2013) to consider a two-gene (or species) repressilator system. For a two-gene repressilator, each of the two genes inhibits the other, i.e. the protein from one species inhibits the production of mRNA from the other. A simple schematic of a generic two-gene repressilator is shown in Figure 8.

We simulate the spatial-stochastic model, with reactions (and associated parameters) given in Table 1, where $n = 2$. We place two individual gene sites within the inner sphere (nucleus). In Figure 9 we show the mRNA and protein behaviour for both gene species for two different diffusion coefficient regimes with promoter sites at $(\pm 0.5\mu\text{m}, 0, 0)$. Overall the behaviour for each species in a two gene repressilator can be compared to the behaviour of a single one gene repressilator species, with mRNA, protein copy numbers and promoter status behaviours being roughly similar (compare to the bottom two panels in Figure 3).

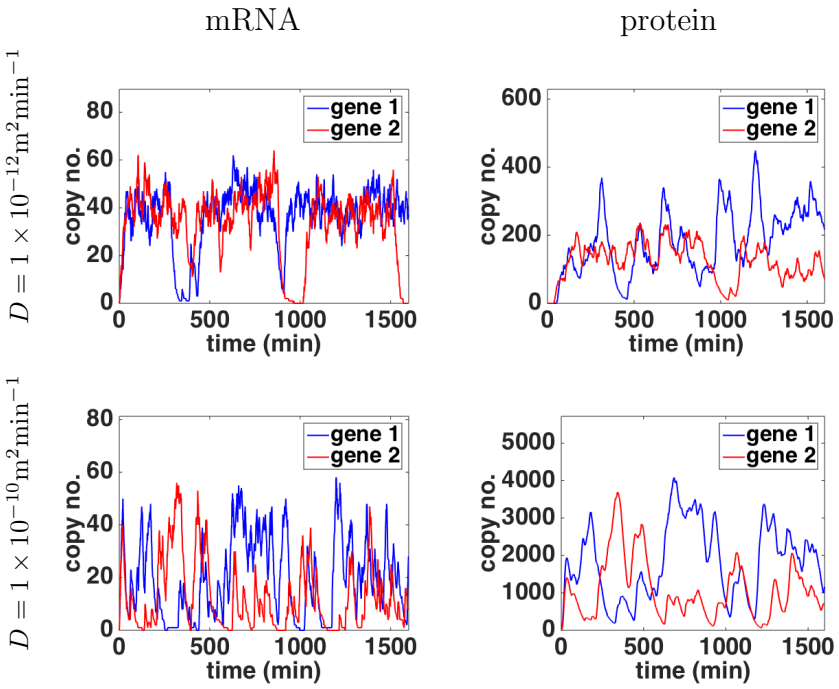

Figure 9: Simulation results for the two-gene repressilator model showing the behaviour of the mRNA and protein copy numbers as we vary the diffusion coefficient. The promoters are located at $(\pm 0.5\mu\text{m}, 0, 0)$.

Differences in behaviour between the one and two-gene repressilators may be noted when the promoter locations are moved closer to the nuclear-cytoplasm membrane. In Figure 10 we show the mRNA and protein behaviour for both gene species when $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ with promoter sites at $(\pm 2.5\mu\text{m}, 0, 0)$. We note significantly different behaviour when compared with the bottom panel of Figure 4. In this case protein levels in particular are more consistent over time, remaining elevated. Equally, copy numbers of

mRNA and protein are higher. To investigate these discrepancies further we consider the spatial behaviour in the following section.

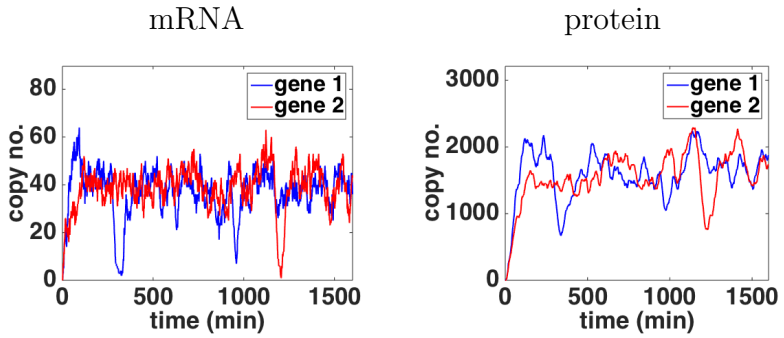

Figure 10: Simulation results for the two-gene repressilator model showing the behaviour of the mRNA and protein copy numbers when $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ and the promoters are located at $(\pm 2.5\mu\text{m}, 0, 0)$.

3.5.1. Spatial Behaviour

In Figure 11 we show snapshots at distinct times of the spatial distribution of the protein species within the spherical domain. Theses plots correspond to the temporal behaviour given by the top right panel of Figure 9, i.e. where $D = 1\times10^{-12}\text{m}^2\text{min}^{-1}$ and the promoters are located at $(\pm 0.5\mu\text{m}, 0, 0)$. Each plot displays a single slice through the 3D domain ($z = 0$) and the colour indicates the protein copy number in each visualised part-voxel. For movies showing the behaviour for the full range of times please see the supplementary material (these movies show the behaviour over three 2D slices through the 3D domain). We observe the fluctuating behaviour of the protein copy number with times when little protein is noted and times when it appears

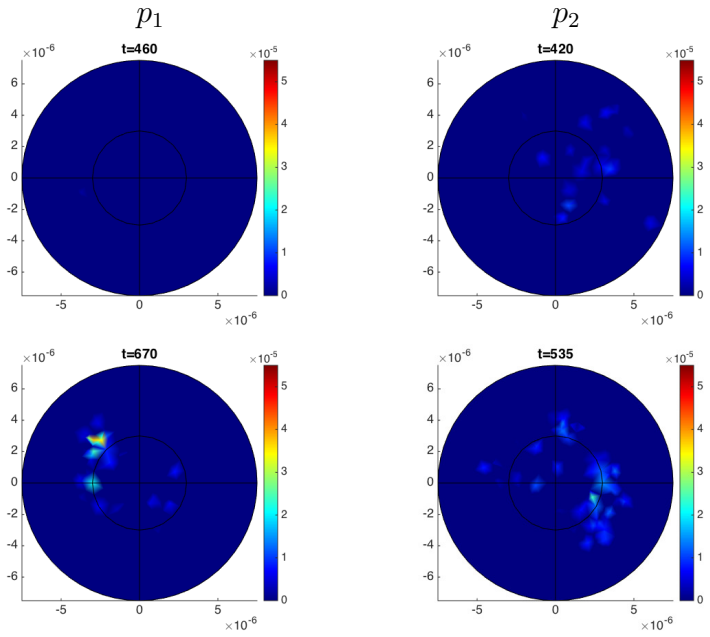

Figure 11: Simulation results for the two-gene repressilator model showing the spatial distribution of the protein copy numbers at the times given. $D = 1 \times 10^{-12}\mathrm{m}^2\mathrm{min}^{-1}$ and the promoter are located at $(\pm 0.5\mu\mathrm{m}, 0, 0)$.

in several voxels throughout the domain. Although protein copy numbers do not appear to be particularly concentrated about any particular voxel(s) the highest peaks of protein are to be found close to the nuclear-cytoplasm membrane.

By way of comparison, in Figure 12, we show the spatial behaviour for simulations where $D = 1 \times 10^{-12}\mathrm{m}^2\mathrm{min}^{-1}$ and the promoter sites are located close to the nuclear membrane at $(\pm 2.5\mu\mathrm{m}, 0, 0)$, corresponding to the righthand plot of Figure 10. Again each plot displays a single slice through

the 3D domain ($z = 0$) and the colour indicates the protein copy number in each part-voxel. For movies showing the behaviour for the full range of times please see the supplementary material. We note some variation of the levels with time, however, the main observation is that high protein levels are localised close to the nuclear-cytoplasm membrane and fixed close to its associated gene promoter site, with protein levels rarely seen in the opposite half of the domain. Since the protein of one gene must bind with the promoter of the other gene to activate the negative feedback, this accounts for the differences observed for a two-gene repressilator compared to a one-gene repressilator. We note that in this case each promoter is rarely occupied (promoter behaviour not shown here), the protein molecules not being in the correct part of the domain to bind with the appropriate promoter. As such the frequent lack of a protein-promoter complexes leads to infrequent repression of the mRNAs which in turn translates into consistent and high levels of protein. For the Hes1 system, with only one gene, protein localisation within the domain would not have this affect since the protein binds with the promoter of its own gene. Instead since the protein is localised close to its promoter this accounts for the observation that the promoter is frequently occupied for the one-gene system since there is always protein available to bind.

For completeness in Figure 13 we show snapshots at distinct times of the spatial distribution of the protein species which correspond to the temporal behaviour given by the bottom right panel of Figure 9, i.e. where $D = 1 \times 10^{-10}\mathrm{m}^2\mathrm{min}^{-1}$ and the promoters are located at $(\pm 0.5\mu\mathrm{m}, 0, 0)$.

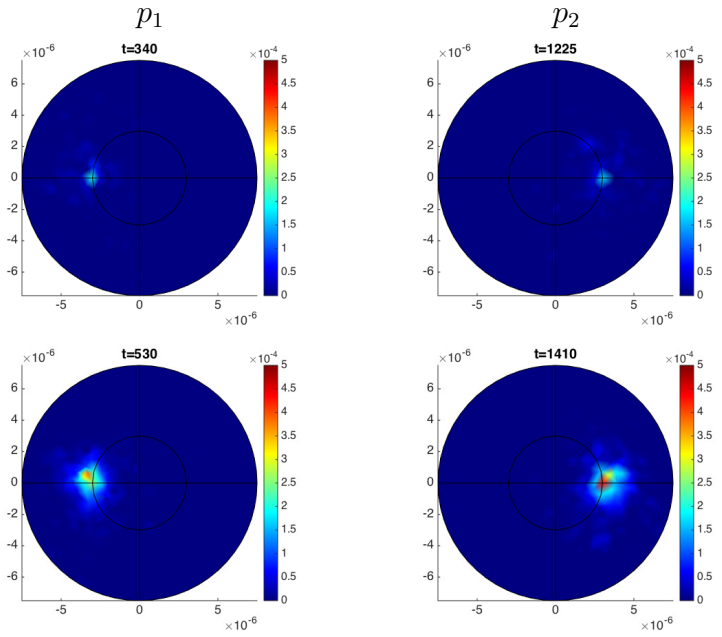

Figure 12: Simulation results for the two-gene repressilator model showing the spatial distribution of the protein copy numbers at the times given. $D = 1 \times 10^{-12}\mathrm{m}^2\mathrm{min}^{-1}$ and the promoter are located at $(\pm 2.5\mu\mathrm{m}, 0, 0)$.

Each plot displays three 2D slices through the 3D domain and the colour indicates the protein copy number in each part-voxel on each slice. For movies showing the behaviour for the full range of times please see the supplementary material. For both protein species fluctuations are observed, with times when very little protein is noted throughout the domain and times when it is widespread. The increase in diffusion coefficient can be clearly noted by the speed and extent to which protein spreads throughout the domain.

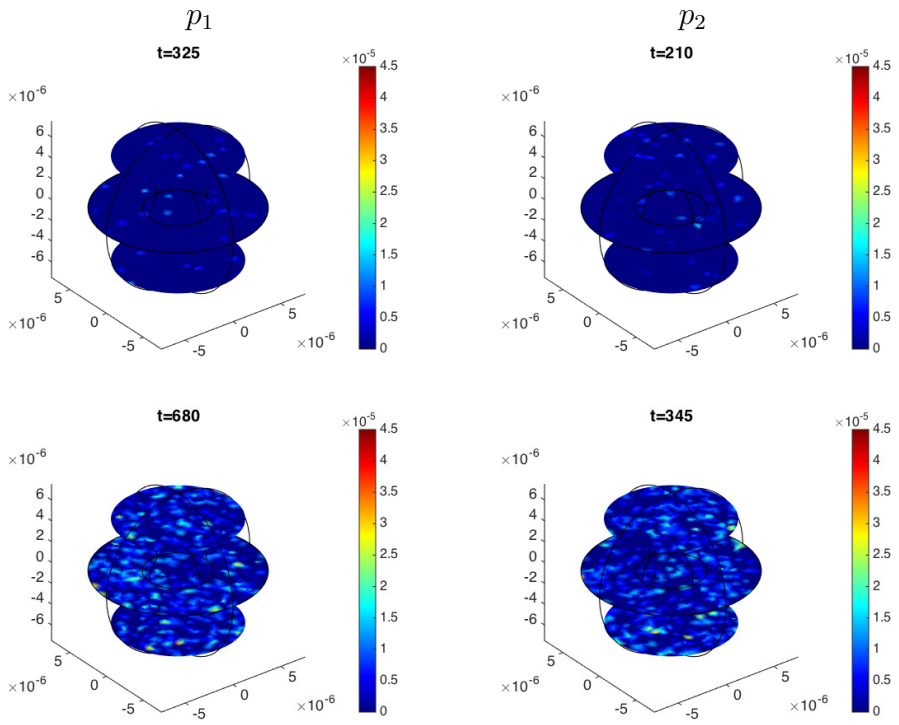

Figure 13: Simulation results for the two-gene repressilator model showing the spatial distribution of the protein copy numbers at the times given. $D = 1 \times 10^{-10}\mathrm{m}^2\mathrm{min}^{-1}$ and the promoter are located at $(\pm 0.5\mu\mathrm{m}, 0, 0)$.

3.5.2. Period Analysis

We carried out period analysis with $D = 1\times10^{-12}\mathrm{m}^2\mathrm{min}^{-1}$ and the promoters defined at $(\pm 0.5\mu\mathrm{m}, 0, 0)$, the results are shown in Figures 14 and 15. In general the species of a two gene repressilator show period behaviour as for a one gene repressilator with either one or two distinct period modes. However, for individual simulations, it does not follow that the behaviour of the species is equivalent. For example, one species may exhibit a single

period mode while the other species exhibits two. In this case of a two gene repressilator we see quite a number of cases in which WAVOS only detects a single period equal to the maximum possible period, we should be wary in such cases whether a true and realistic period has been detected. For the deterministic case in Macnamara and Chaplain (2016) it was noted that a two-gene repressilator was a weak oscillator compared to the equivalent one-gene repressilator. This may well be the case here too but is not a focus of this investigation.

4. Activator-Repressor Systems

We can easily extend our investigation to examine activator-repressor systems, rather than repressilator-only systems. We do this by considering the parameter γ_i, as described by Sturrock et al. (2013). For cases of repression $\gamma_i > 1$, in order that the production of mRNA is reduced from its baseline value, α_m, when the promoter is occupied by the appropriate protein. If we wish a specific protein to promote rather than repress the production of mRNA, we require that $0 < \gamma_i < 1$.

4.1. Two-gene Activator-Repressor

Specifically we will consider a two-gene activator-repressor system; a simple schematic of which is shown in Figure 16; this can be directly compared to Figure 8. We note that now the protein of species 1 promotes rather than inhibits the production of species 2 mRNA.

We simulate the spatial-stochastic model, with reactions (and associated

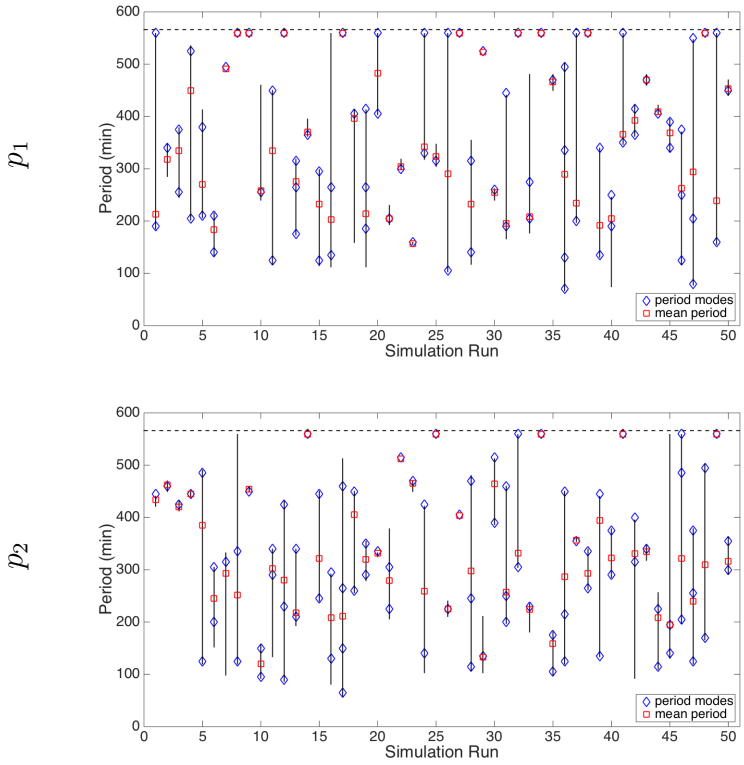

Figure 14: Period data for 50 simulations of the two gene repressilator model; the red squares are the mean periods, the black lines indicate the range of periods, the blue diamonds indicate the value of the period mode(s). $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ and the promoters are located at $(\pm 0.5\mu\text{m}, 0, 0)$.

parameters) given in Table 1 (with the sole exception being γ_i; we choose $\gamma_1 = 10$ and $\gamma_2 = 0.1$) where $n = 2$. We place two individual gene sites within the inner sphere (nucleus), specifically at $(\pm 0.5\mu\text{m}, 0, 0)$. In Figure 17 we show the mRNA and protein behaviour for both gene species for two different diffusion coefficient regimes. For this activator-repressor system we notice that although protein fluctuations are seen in both species the amplitudes of

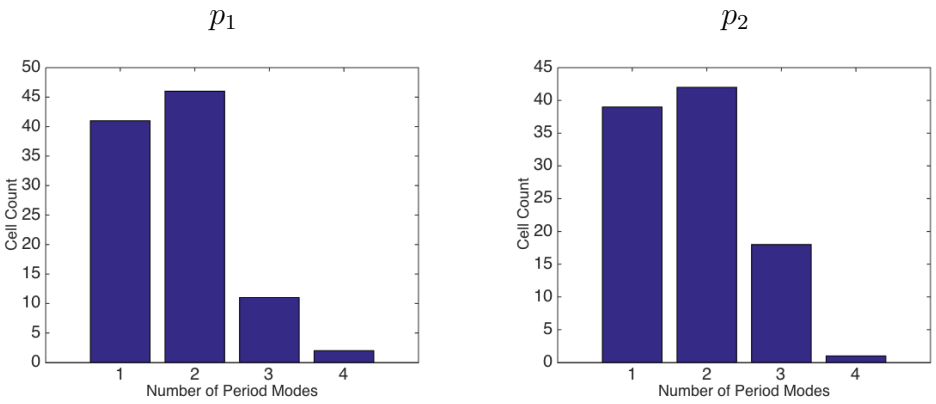

Figure 15: Histogram indicating the number of period modes determined for 100 simulations of the two gene repressilator model. $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ and the promoters are located at $(\pm 0.5\mu\text{m}, 0, 0)$.

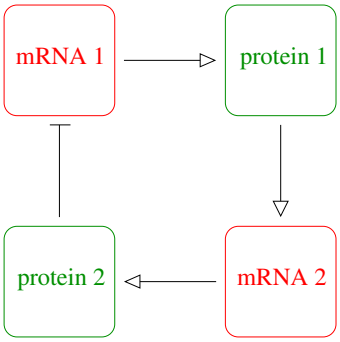

Figure 16: Simple schematic of the two-gene activator-repressor system. The mRNA of each species produces its own protein. The protein of species 1 promotes the production of species 2 mRNA, while the protein of species 2 inhibits the production of species 1 mRNA. The colours correspond to where the molecular reactions take place, i.e. red for at the promoter site and green in the cytoplasm.

these fluctuations are typically far greater for species 2 (the promoted species) than for species 1 (the inhibited species). For $D = 1 \times 10^{-12}\mathrm{m}^2\mathrm{min}^{-1}$, the promoter behaves in a very similar way for both species, whereas for $D = 1 \times 10^{-10}\mathrm{m}^2\mathrm{min}^{-1}$ we notice long periods when the promoter for species 1 is switched off (promoter behaviour not shown here) corresponding to very low although still fluctuating levels of mRNA and protein. As for the repressilator systems we observe that increasing the diffusion coefficient increases the frequency of promoter-protein binding, this leads to more frequent fluctuations in the mRNA and protein copy numbers.

4.2. Spatial Behaviour

In Figure 18 we show snapshots at distinct times of the spatial distribution of the protein species within the spherical domain, when $D = 1 \times 10^{-12}\mathrm{m}^2\mathrm{min}^{-1}$ and the promoters are located at $(\pm 0.5\mu\mathrm{m}, 0, 0)$ in Figure 19 the behaviour is shown when $D = 1 \times 10^{-10}\mathrm{m}^2\mathrm{min}^{-1}$. Each plot displays three 2D slices through the 3D domain and the colour indicates the protein copy number in each part-voxel on each slice. For movies showing the behaviour for the full range of times please see the supplementary material. For both diffusion coefficient regimes the levels of protein of species 1 (the inhibited species) fluctuate far less than the levels of protein of species 2 (the activated species). When $D = 1 \times 10^{-12}\mathrm{m}^2\mathrm{min}^{-1}$ the highest levels of protein of either species typically appear close to the nuclear-cytoplasm membrane whereas when $D = 1 \times 10^{-10}\mathrm{m}^2\mathrm{min}^{-1}$ protein is found more uniformly throughout the domain.

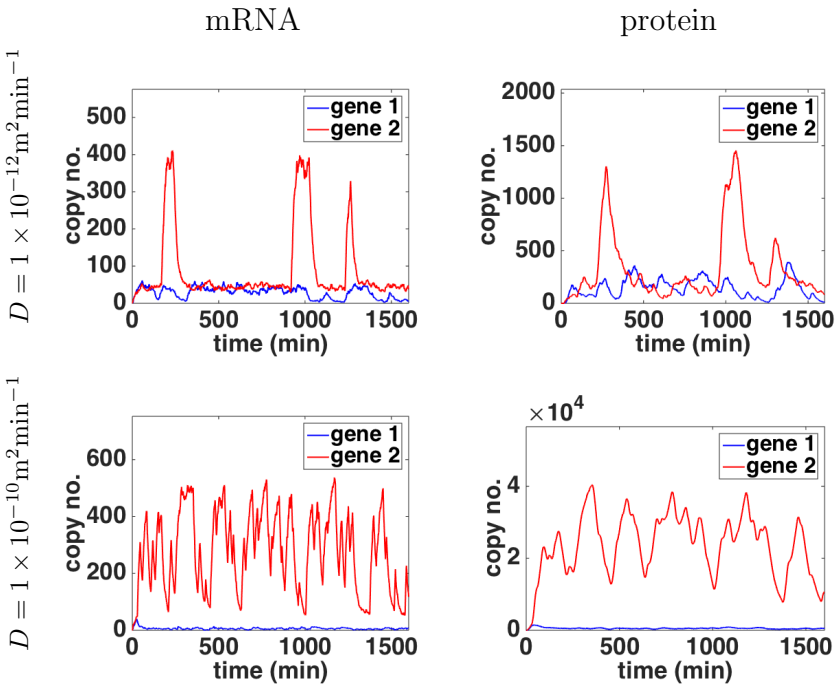

Figure 17: Simulation results for the two-gene activator-repressor model showing the behaviour of the mRNA and protein copy numbers. The promoters are located at $(\pm 0.5\mu\mathrm{m}, 0, 0)$ and parameters are as given in Table 1 bar γ_i, here $\gamma_1 = 10$ and $\gamma_2 = 0.1$.

4.2.1. Period Analysis

We repeat the period analysis and the results are shown in Figures 20 and 21. We observe differences in the period behaviour of the two species. The inhibited species (species one) typically exhibits one or two distinct periods. The promoted species is more likely to only exhibit a single period. In actual fact it is likely that we are unable to find true and realistic periods for this species for many of the simulations; notice the high proportion of simula-

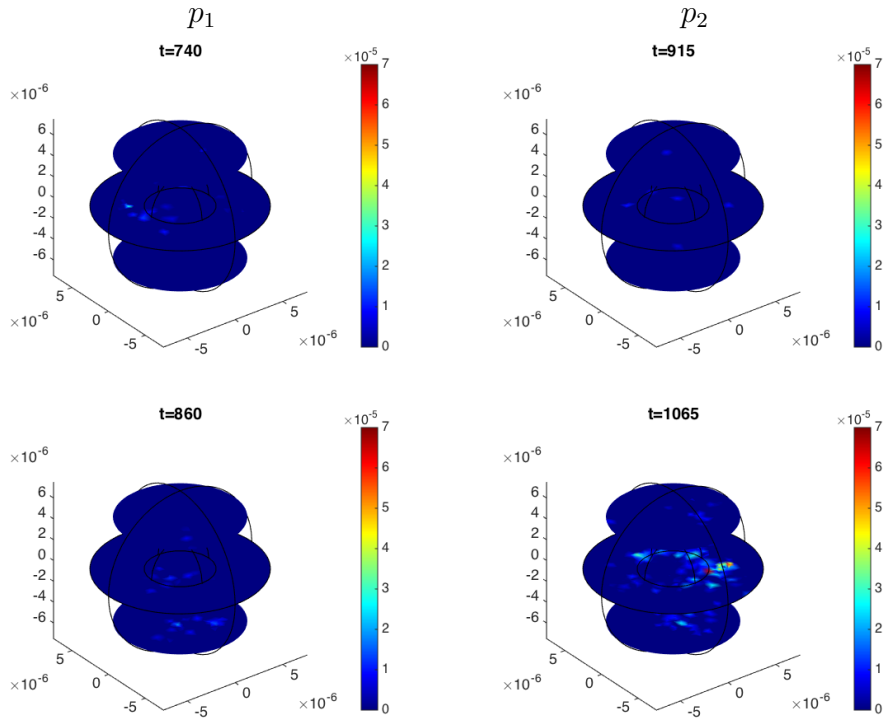

Figure 18: Simulation results for the two-gene activator-repressor model showing the spatial distribution of the protein copy numbers at the times given. $D = 1 \times 10^{-12}\mathrm{m}^2\mathrm{min}^{-1}$ and the promoters are located at $(\pm 0.5\mu\mathrm{m}, 0, 0)$.

tions which result in single period modes lying on the maximal period line (indicated by the dashed black line). While interesting dynamics are clearly observed both temporally and spatially for this activator-repressor system true periodic behaviour maybe elusive. This appears in agreement with the corresponding deterministic model of Macnamara and Chaplain (2016) which was found to be a weak oscillator.

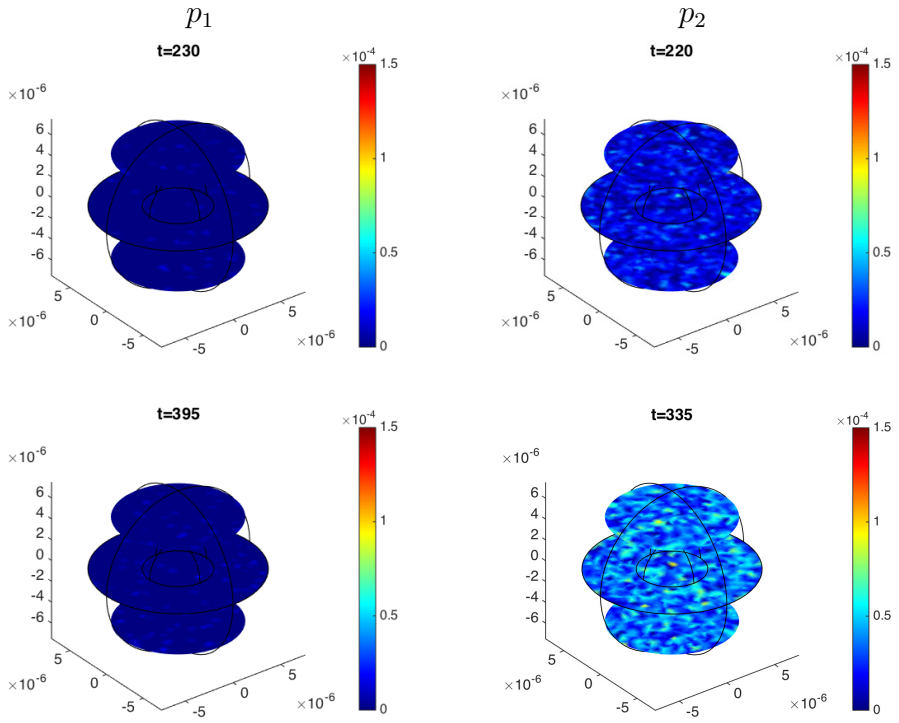

Figure 19: Simulation results for the two-gene activator-repressor model showing the spatial distribution of the protein copy numbers at the times given. $D = 1 \times 10^{-10}\mathrm{m}^2\mathrm{min}^{-1}$ and the promoters are located at $(\pm 0.5\mu\mathrm{m}, 0, 0)$.

5. Discussion and Conclusions

In this paper we have developed spatial-stochastic models of gene regulatory networks, focussing on repressilators and activator-repressor systems, and explored the effects of altering the diffusion coefficients of the molecules and the precise location of the promoter region in the nucleus on the spatio-temporal behaviour of such systems. Before exploring synthetic GRNs, we

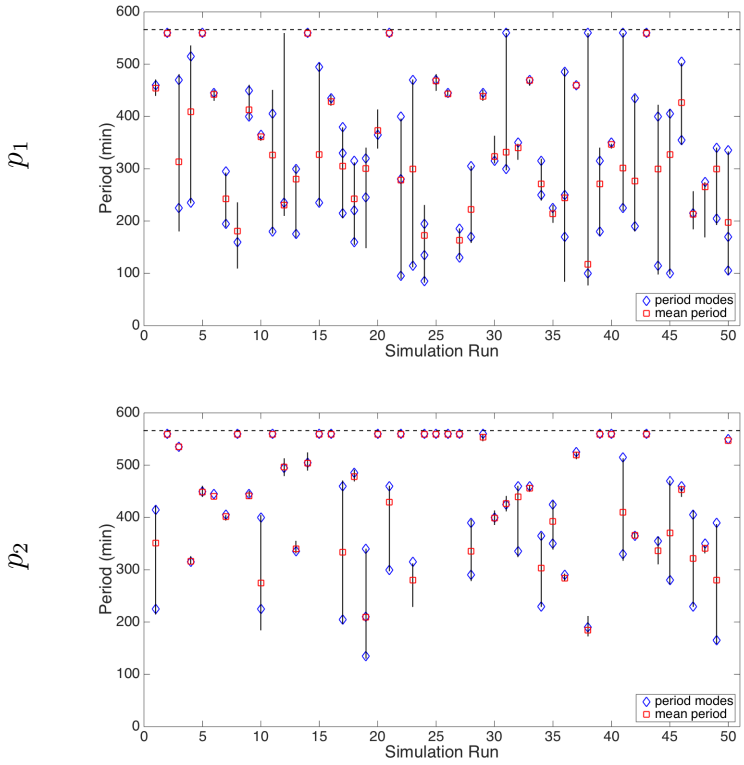

Figure 20: Period data for 50 simulations of the two gene activator-repressor model; the red squares are the mean periods, the black lines indicate the range of periods, the blue diamonds indicate the value of the period mode(s). $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$ and the promoters are located at $(\pm 0.5\mu\text{m}, 0, 0)$.

further investigated the Hes1 system (which may be classified as a one-gene repressilator) previously modelled by Sturrock et al. (2013). We then extended this work to focus on a two-gene repressilator system before also exploring an activator-repressor system.

Our investigation into the Hes1/one-gene repressilator system showed that,

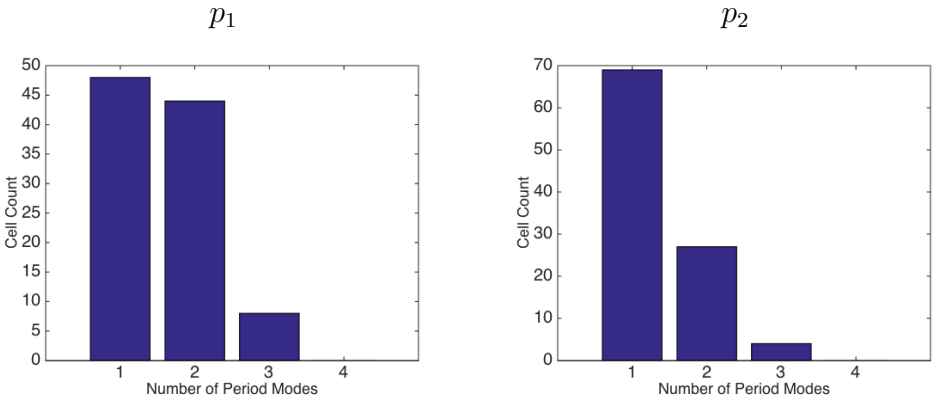

Figure 21: Histogram indicating the number of period modes determined for 100 simulations of the two gene activator-repressor model. $D = 1 \times 10^{-12}\mathrm{m}^2\mathrm{min}^{-1}$ and the promoters are located at $(\pm 0.5\mu\mathrm{m}, 0, 0)$.

as for the PDE model (Macnamara and Chaplain, 2016), spatial aspects play a crucial role in determining the spatio-temporal behaviour of such systems. Spatio-temporal dynamics are governed by diffusion coefficients and promoter locations. This is apparent in both the time series data - where we note increased protein copy numbers and fluctuations for increased diffusion coefficient, and the spatial distributions of the molecules - where we observe localisation of protein levels as the promoter is moved closer to the nucleus-cytoplasm membrane. Our computational simulation results show that diffusion coefficients of the molecules must lie within some appropriate range for oscillations to occur. If the diffusion of molecules is too slow they will not be able to reach the appropriate locations for transcription/translation processes to occur. If diffusion occurs too quickly and molecules spread too efficiently throughout the domain promoter-binding occurs too frequently

leading to unrealistic periodic behaviour. Equally, the positioning of promoter sites requires careful consideration so as to achieve spatio-temporal behaviour which is biologically realistic and relevant. Moving the promoter location closer to the nuclear-cytoplasm membrane has a similar effect to increasing the diffusion coefficient. However, if the promoter is positioned very close to the membrane a frequently occupied promoter leads to low mRNA and protein copy numbers and more frequent low amplitude fluctuations. We note that results for the Hes1 system are comparative to those determined from the PDE model, which captures the essential features of the behaviour of such gene regulatory systems. However, the stochastic model implemented here would allow a direct comparison with experimental data such as seen in Hirata et al. (2002), for example.

Our investigations into synthetic systems provide further corroboration of the importance of considering spatial aspects. Both the two-gene repressilator and activator-repressor systems show considerably different dynamics depending on diffusion coefficient and promoter location. Differences are also observed between the two systems and the one-gene Hes1 system. While a two-gene repressilator may behave in a similar way to two one-gene repressilators the spatio-temporal dynamics of the system are entirely dependant on the level of diffusion and placement of promoter locations. Negative feedback within a two-gene repressilator may be switched on or off depending on whether the protein molecules are able to travel to the appropriate promoter sites for binding. Only with negative feedback activated may periodic behaviour be observed. Two-gene systems behave very differently when coupled

by both positive and negative feedback as in the activator-repressor system; the inhibited species fluctuating more rapidly and at much lower copy numbers such that periodic behaviour is less frequently observed. However, while this system may not exhibit true and realistic periodic behaviour interesting spatio-temporal behaviour is observed which is dependant on the spatial terms within the model set-up. Again the behaviour of the stochastic systems corroborate to an extent with deterministic models.

In summary, deterministic spatio-temporal models provide informative qualitative results about these GRN systems, however, the stochastic models developed and implemented here effectively offer *in silico* single cell experiments which can be compared to the single cell experiments of biologists. Such a comparison can be done for both real and synthetic systems alike. The purpose of this paper is to introduce the canonical form of a model which may be adapted to fit any cyclic system of genes coupled by any combination of positive and negative feedback, and in particular we discuss the behaviour of two-gene repressilator and activator repressor systems. These do not link directly with known biological systems at this time but rather serve as an example of the modelling framework (see, Purcell et al., 2010; O'Brien et al., 2012). In Szymanska et al. (2018) we model the NF-κB system using a similar spatial-stochastic approach, and the computational simulation results obtained may be directly compared to experimental data such as in Nelson et al. (2004) and Ashall et al. (2009). A future aim of the models developed in this paper would be to provide a full spatial-stochastic model of the p53-Mdm2 system and compare it to experimental data such as in

Lahav et al. (2004). p53 is an important intracellular protein and as such it has garnered much interest since its discovery almost 40 years ago (Lane and Crawford, 1979) and subsequent role as a tumour suppressor and its ability to control apoptosis (Lane, 1992). If the dynamics of the p53-Mdm2 GRN were understood more deeply, this would improve the ability to design anti-cancer drugs/therapies which target the appropriate part of the pathway.

The study of synthetic GRNs is of great relevance given the recent growth in the field of synthetic biology. Interdisciplinary teams of biologists and mathematicians build models of such systems to analyse and gain a deeper understanding of the underlying biology of complex intracellular systems (see, for example, Balagadde et al., 2008; Becskei and Serrano, 2000; Elowitz and Leibler, 2000; Purcell et al., 2010; Chen et al., 2012; O'Brien et al., 2012; Yordanov et al., 2014). At the same time, the ability to simulate such models *in silico* alleviates financial and potentially ethical costs associated with *in vitro* and *in vivo* experiments. The findings in this paper reinforce the message that molecular movement must be taken into account when trying to design such systems. How quickly molecules are able to move through a domain and access the precise locations with transcription/translation processes occur has a critical effect on the overall spatio-temporal behaviour of the system.

Acknowledgements

MAJC and CKM gratefully acknowledge support of EPSRC Grant No. EP/N014642/1 (EPSRC Centre for Multiscale Soft Tissue Mechanics - With Application to

Heart & Cancer).

Appendix A. The Reaction Diffusion Master Equation

The model discussed in Section 2 is governed by a reaction-diffusion master equation (RDME) for $p(\mathbf{x}, t)$ - the probability that the system can be found in state $\mathbf{x}$ at time t. The RDME is

$$\frac{\mathrm{d}}{\mathrm{dt}}p(\mathbf{x},t) = \underbrace{\mathscr{M}p(\mathbf{x},t)}_{\text{reactions}} + \underbrace{\mathscr{D}p(\mathbf{x},t)}_{\text{diffusion}}. \tag{A.1}$$

While we consider a spherical domain which globally is not well-mixed (in order to account for spatial inhomogeneity) we discretise it into K non-overlapping voxels inside which we do assume a well-stirred system in which species are uniformly distributed and under thermal equilibrium. By doing this we can model the reactions, which occur within voxels, as a stochastic continuous-time discrete space Markov process. The reactions are governed by a chemical master equation (CME) for M reactions concerning G genes in K voxels,

$$\begin{aligned}\mathscr{M}p(\mathbf{x},t) &= \sum_{j=1}^{K}\sum_{r=1}^{M}\omega_{jr}(\mathbf{x}_{j\cdot}-\mu_{jr})p(\mathbf{x}_{1\cdot},\ldots,\mathbf{x}_{j\cdot}-\mu_{jr},\ldots,\mathbf{x}_{K\cdot},t)\\ &\quad -\sum_{j=1}^{K}\sum_{r=1}^{M}\omega_{jr}(\mathbf{x}_{j\cdot})p(\mathbf{x},t)\end{aligned} \tag{A.2}$$

where the state of the system, $\mathbf{x}$, is described by a $K \times N$ matrix which at anyone time holds the copy number of each species $s = 1, \ldots, N$ within each voxel $j = 1, \ldots, K$ (note for the systems we consider $N = 4G$ as there are 4 species per gene). The propensity function $\omega_{jr}(\mathbf{x}_{j\cdot})$ describing the rate of reaction, r, depends on the copy number of the species within the jth voxel,

ψ_j (note $\mathbf{x}_{j\cdot}$ is the jth row of the $K \times N$ matrix). The $1 \times N$ stoichiometry vector, μ_{jr} specifies the change in copy number of the reactants and products for reaction, r in voxel, ψ_j.

For example, for the single gene Hes1 system there are four species fp, op, m and p and seven reactions:

$$
\begin{aligned}
r_1 &: \quad m \xrightarrow{\alpha_p} p \ (+\, m) \\
r_2 &: \quad fp + p \xrightarrow{k_1} op \\
r_3 &: \quad op \xrightarrow{k_2} fp + p \\
r_4 &: \quad fp \xrightarrow{\alpha_m} m \ (+\, fp) \\
r_5 &: \quad op \xrightarrow{\alpha_m/\gamma} m \ (+\, op) \\
r_6 &: \quad m \xrightarrow{\mu_m} \emptyset \\
r_7 &: \quad p \xrightarrow{\mu_p} \emptyset.
\end{aligned}
$$

Since we assume mass action kinetics the propensity function for reaction, r_1, for example is $\alpha_p[m]$ where $[m]$ is the copy number of mRNA. The stoichiometry vector for this reaction is $[0, 0, 0, 1]$ - there is a net gain of one protein molecule per reaction.

While reactions are summed over voxels to add to the RDME we model diffusion as linear jumps governed by the following master equation

$$
\begin{aligned}
\mathscr{D}p(\mathbf{x}, t) \;=\; & \sum_{s=1}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K} a(\mathbf{x}_{\cdot s} - \nu_{sjk})p(\mathbf{x}_{\cdot 1}, \ldots, \mathbf{x}_{\cdot s} - \nu_{sjk}, \ldots, \mathbf{x}_{\cdot N}, t) \\
& - \sum_{s=1}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K} a(\mathbf{x}_{\cdot s})p(\mathbf{x}, t) \qquad \text{(A.3)}
\end{aligned}
$$

The propensity function $a(\mathbf{x}_{\cdot s}) = d_{ijk}[S_{ij}]$ (where $[S_{ij}]$ is the copy number of the ith species in the jth voxel) describes the "rate" of diffusion, i.e. the probability of species, S_{ij} in voxel ψ_j moving to one of the immediate neighbour voxels ψ_k (note $\mathbf{x}_{\cdot s}$ is the sth column of the $K \times N$ matrix). The $K \times 1$ stoichiometry vector, v_{sjk}, specifies the change in copy number of species s and so has all components zero except for $\nu_{sjk}(j) = -1$ and $\nu_{sjk}(k) =$ 1 - a molecule leaves the jth voxel and enters the kth voxel. The jump rate constants, d_{ijk} vary depending on the size and shape of the voxels, for a detailed description of the theory for how to obtain correct rates, see Engblom et al. (2009). In brief, however, the value $1/(d_{ijk}[S_{ij}])$ is the expected time for the first molecule of species S_i to leave voxel ψ_j and become well-mixed in voxel ψ_k. A reasonable requirement on these rate constants is that they are chosen in such a way that the diffusion process converges to the diffusion equation in the thermodynamic limit.

Appendix B. Period Analysis of the Hes1 Time Series

We display the mRNA and protein copy numbers along with the instantaneous period derived from the changes in protein copy number using WAVOS for five runs of the simulation of the Hes1 system with $D = 1 \times 10^{-12}\text{m}^2\text{min}^{-1}$. We observe quite disparate results; for some runs the instantaneous period determined remains roughly constant (notably panel five of Figure B.22), whereas for other runs there are clearly two distinct periods derived from the stochastic data for different time ranges (notably panels two and three of Figure B.22).

References

Ashall, L., Horton, C. A., Nelson, D. E., Paszek, P., Harper, C. V., Sillitoe, K., S., R., Spiller, D. G., Unitt, J. F., Broomhead, D. S., B., K. D., Rand, D. A., Sée, V., White, M. R. H., 2009. Pulsatile stimulation determines timing and specificity of nf-kb–dependent transcription. Science 324, 242–246.

Balagadde, F. K., Song, H., Ozaki, J., Collins, C. H., Barnet, M., Arnold, F. H., Quake, S. R., You, L., 2008. A synthetic escherichia coli predator–prey ecosystem. Mol. Syst. Biol. 4:187.

Becskei, A., Serrano, L., 2000. Engineering stability in gene networks by autoregulation. Nature 405, 590–593.

Bernard, S., Čajavec, B., Pujo-Menjouet, L., Mackey, M. C., Herzel, H., 2006. Modeling transcriptional feedback loops: The role of gro/tle1 in hes1 oscillations. Philos. Trans. A. Math. Phys. Eng. Sci. 15, 1155–1170.

Busenberg, S., Mahaffy, J. M., 1985. Interaction of spatial diffusion and delays in models of genetic control by repression. J. Math. Biol. 22, 313–333.

Cangiani, A., Natalini, R., 2010. A spatial model of cellular molecular trafficking including active transport along microtubules. J. Theor. Biol. 267, 614–625.

Chaplain, M. A. J., Ptashnyk, M., Sturrock, M., 2015. Hopf bifurcation in a gene regulatory network model: Molecular movement causes oscillations. Math. Mod. Meth. Appl. S. 25 (6), 1179–1215.

665 Chen, Y. Y., Galloway, K. E., Smolke, C. D., 2012. Synthetic biology: ad-
666 vancing biological frontiers by building synthetic systems. Genome Biol.
667 13:240.

668 Cullhed, J., Engblom, S., Hellander, A., 2008. The urdme manual ver-
669 sion 1.0. technical report 2008-022. Tech. rep., Department of Inform-
670 ation Technology, Uppsala University, Uppsala, Sweden, Available from
671 http://www.it.uu.se/research.

672 Dimitrio, L., Clairambault, J., Natalini, R., 2013. A spatial physiological
673 model for p53 intracellular dynamics. J. Theor. Biol. 316, 69–24.

674 Drawert, B., Engblom, S., Hellander, A., 2012. Urdme: a modular frame-
675 work for stochastic simulation of reaction-transport processes in complex
676 geometries. BMC Syst. Biol. 6 (76).

677 Eliaš, J., Clairambault, J., 2014. Reaction-diffusion systems for spatio-
678 temporal intracellular protein networks: a beginner's guide with two ex-
679 amples. Comp. Struct. Biotechnol. J. 10, 14–22.

680 Eliaš, J., Dimitrio, L., Clairambault, J., Natalini, R., 2014a. Modelling p53
681 dynamics in single cells: physiologically based ode and reaction-diffusion
682 pde models. Phys. Biol. 11, 045001.

683 Eliaš, J., Dimitrio, L., Clairambault, J., Natalini, R., 2014b. The p53 protein
684 and its molecular network: modelling a missing link between dna damage
685 and cell fate. BBA Proteins Proteom. 1844, 232–247.

686 Elowitz, M. B., Leibler, S., 2000. A synthetic oscillatory network of tran-
687 scriptional regulators. Nature 403, 335–338.

688 Engblom, S., Ferm, L., Hellander, A., L̀otstedt, P., 2009. Simulation of
689 stochastic reaction–diffusion processes on unstructured meshes. SIAM J.
690 Sci. Comput. 31 (3), 1774–1797.

691 Gibson, M. A., Bruck, J., 2000. Efficient exact stochastic simulation of chem-
692 ical species and many channels. J. Phys. Chem. 104, 1876–1889.

693 Glass, L., Kauffman, S. A., 1970. Co-operative components, spatial localiza-
694 tion and oscillatory cellular dynamics. J. Theor. Biol. 34, 219–237.

695 Goodwin, B. C., 1965. Oscillatory behaviour in enzymatic control processes.
696 Adv. Enzyme Regul. 3, 425–428.

697 Griffith, J. S., 1968. Mathematics of cellular control processes. i. negative
698 feedback to one gene. J. Theor. Biol. 20, 202–208.

699 Harang, R., Bonnet, G., Petzold, L. R., 2012. Wavos: a matlab toolkit for
700 wavelet analysis and visualization of oscillatory systems. BMC Res. Notes
701 5, 163.

702 Hirata, H., Yoshiura, S., Ohtsuka, T., Bessho, Y., Harada, T., Yoshikawa,
703 K., Kageyama, R., 2002. Oscillatory expression of the bhlh factor hes1
704 regulated by a negative feedback loop. Science 298, 840–843.

705 Jensen, M. H., Sneppen, J., Tiana, G., 2003. Sustained oscillations and time
706 delays in gene expression of protein hes1. FEBS Lett. 541, 176–177.

707 Kageyama, R., Ohtsuka, T., Kobayashi, T., 2007. The hes gene family:
708 repressors and oscillators that orchestrate embryogenesis. Development
709 134, 1243–51.

Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A. J., Elowitz, M. B., Alon, U., 2004. Dynamics of the p53-mdm2 feedback loop in individual cells. Nature Genet. 36, 147–150.

Lane, D. P., 1992. p53, guardian of the genome. Nature 358, 15–16.

Lane, D. P., Crawford, L. V., 1979. T antigen is bound to a host protein in sv40-transformed cells. Nature 278, 261–263.

Lewis, J., 2003. Autoinhibition with transcriptional delay: A simple mechanism for the zebrafish somitogenesis oscillator. Curr. Bio. 13, 1398–1408.

Mackey, M. C., Glass, L., 1977. Oscillation and chaos in physiological control systems. Science 197, 287–289.

Macnamara, C. K., Chaplain, M. A. J., 2016. Diffusion driven oscillations in gene regulatory networks. J. Theor. Biol. 407, 51–70.

Mahaffy, J. M., 1988. Genetic control models with diffusion and delays. Math. Biosci. 90, 519–533.

Mahaffy, J. M., Pao, C. V., 1984. Models of genetic control by repression with time delays and spatial effects. J. Math. Biol. 20, 39–57.

Mallat, S. A., 1998. A wavelet tour of signal processing. Academic Press.

Momiji, H., Monk, N. A. M., 2008. Dissecting the dynamics of the hes1 genetic oscillator. J. Theor. Biol. 254, 784–798.

Monk, N. A. M., 2003. Oscillatory expression of hes1, p53, and nf-κb driven by transcriptional time delays. Curr. Biol. 13, 1409–1413.

731 Naqib, F., Quail, T., Musa, L., Vulpe, H., Nadeau, J., Lei, J., Glass, L.,
732 2012. Tunable oscillations and chaotic dynamics in systems with localized
733 synthesis. Phys. Rev. E 85, 046210.

734 Nelson, D. E., Ihekwaba, A. E. C., Elliott, M., Johnson, J. R., Gibney, C. A.,
735 Foreman, B. E., Nelson, G., See, V., Horton, C. A., Spiller, D. G., Edwards,
736 S. W., McDowell, H. P., Unitt, J. F., Sullivan, E., Grimley, R., Benson, N.,
737 Broomhead, D., Kell, D. B., White, M. R. H., 2004. Oscillations in nf-.b
738 signaling control the dynamics of gene expression. Science 306, 704–708.

739 O'Brien, E. L., Itallie, E. V., Bennett, M. R., 2012. Modeling synthetic gene
740 oscillators. Math. Biosci. 236, 1–15.

741 Purcell, O., Savery, N. J., Grierson, C. S., di Bernardo, M., 2010. A com-
742 parative analysis of synthetic genetic oscillators. J. R. Soc. Interface 7,
743 1503–1524.

744 Shymko, R. M., Glass, L., 1974. Spatial switching in chemical reactions with
745 heterogeneous catalysis. J. Chem. Phys. 60, 835–841.

746 Smolen, P., Baxter, D. A., Byrne, J. H., 1999. Effects of macromolecular
747 transport and stochastic fluctuations on the dynamics of genetic regulatory
748 systems. Am. J. Physiol. 277, C777–C790.

749 Smolen, P., Baxter, D. A., Byrne, J. H., 2001. Modeling circadian oscillations
750 with interlocking positive and negative feedback loops. J. Neurosci. 21,
751 6644–6656.

752 Smolen, P., Baxter, D. A., Byrne, J. H., 2002. A reduced model clarifies the

role of feedback loops and time delays in the drosophila circadian oscillator. Biophys. J. 83, 2349–2359.

Sturrock, M., Hellander, A., Matzavinos, A., Chaplain, M. A. J., 2013. Spatial stochastic modelling of the hes1 gene regulatory network: intrinsic noise can explain heterogeneity in embryonic stem cell differentiation. J. R. Soc. Interface 10, 20120988.

Sturrock, M., Terry, A. J., Xirodimas, D. P., Thompson, A. M., Chaplain, M. A. J., 2011. Spatio-temporal modelling of the hes1 and p53-mdm2 intracellular signalling pathways. J. Theor. Biol. 273, 15–31.

Sturrock, M., Terry, A. J., Xirodimas, D. P., Thompson, A. M., Chaplain, M. A. J., 2012. Influence of the nuclear membrane, active transport, and cell shape on the hes1 and p53-mdm2 pathways: insights from spatio-temporal modelling. Bull. Math. Biol. 74, 1531–1579.

Szymanska, Z., Cytowski, M., Mitchell, E., Macnamara, C. K., Chaplain, M. A. J., 2018. Computational modelling of cancer development and growth: modelling at multiple scales and multiscale modelling. Bull. Math. Biol. 80 (5), 1366–1403.

Szymańska, Z., Parisot, M., Lachowicz, M., 2014. Mathematical modeling of the intracellular protein dynamics: The importance of active transport along microtubules. J. Theor. Biol. 363, 118–128.

Tiana, G., Jensen, M. H., Sneppen, K., 2002. Time delay as a key to apoptosis induction in the p53 network. Eur. Phys. J. B 29, 135–140.

775 Torrence, C., Compo, G. P., 1998. A practical guide to wavelet analysis. Bull.
776 Am. Meteorol. Soc. 79, 61–78.

777 Yordanov, B., Dalchau, N., Grant, P. K., Pedersen, M., Emmott, S., Haseloff,
778 J., Phillips, A., 2014. A computational method for automated character-
779 ization of genetic components. ACS Synth. Biol. 3, 578–588.

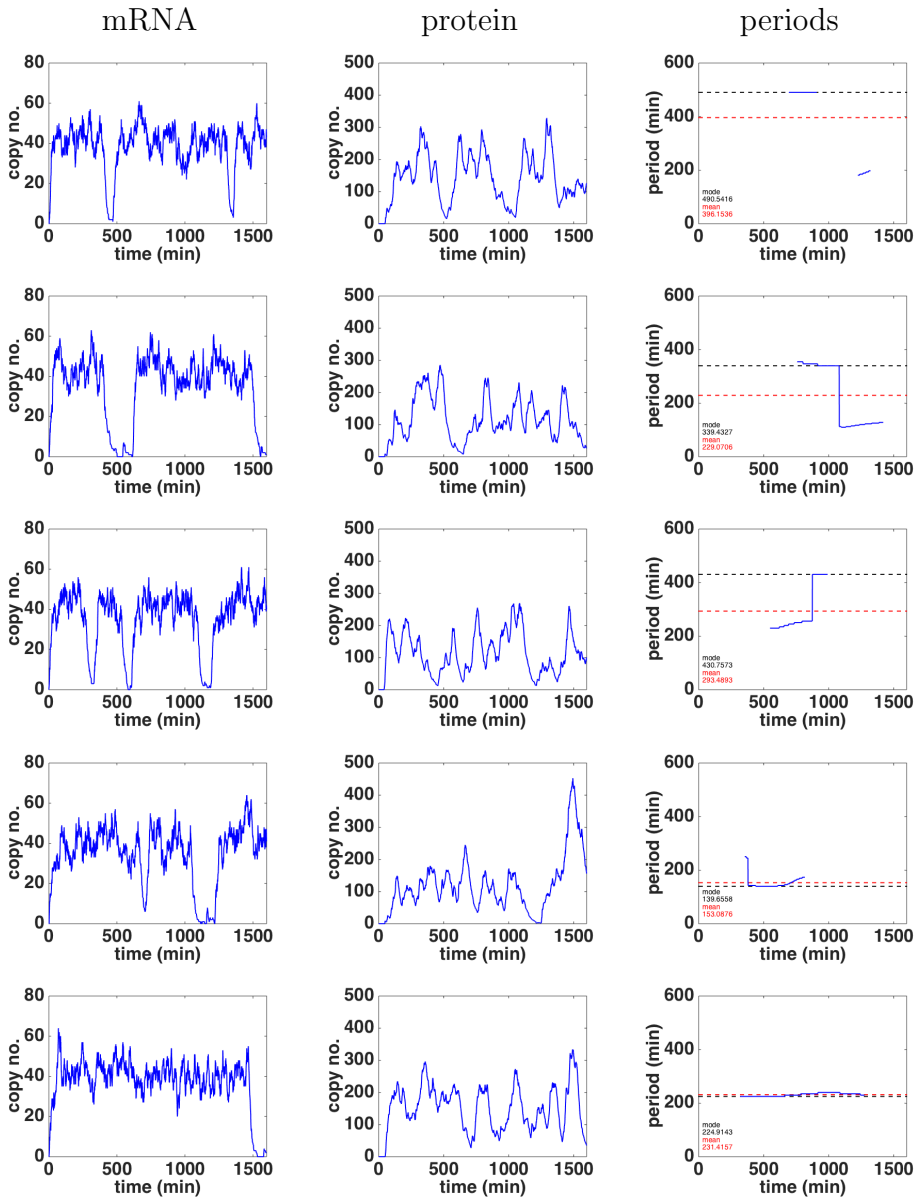

Figure B.22: Time Series for 5 simulations of the Hes1 model when $D = 1 \times 10^{-12}\mathrm{m}^2\mathrm{min}^{-1}$. We indicate the behaviour of the mRNA and protein copy numbers along with the derived instantaneous period from the WAVOS analysis.