Supplement S1

Daniel E Tufcea, Paul François*

* E-mail: Corresponding paulf@physics.mcgill.ca

August 29, 2014

1 Parameters for 2D model

Rates		
	N	O
Production α	5	5
Degradation γ	1	1

Hill Coefficients h_{ij} and h_{iM}			
	N	O	M
N	-	1	-1
O	5	-	-1

Treshholds θ_{ij} and θ_{iM}			
	N	O	M
N	-	1	1
O	1	-	1

Table 1: Parameters used in the simplified model from Balaskas et al. [1] in matrix form.

1.1 Parameter for 2D model with the notation of Balaskas et al.

Parameters	Matrix Notation	Description	Value
α	-	Maximum rate of P	3
β	α_O	Maximum rate of O	5
γ	α_N	Maximum rate of N	5
h_1	-	Hill coef of N on P	6
h_2	-	Hill coef of O on P	2
h_3	h_{ON}	Hill coef of N on O	5
h_4	h_{NO}	Hill coef of O on N	1
h_5	h_{NP}	Hill coef of P on N	1
k_1	-	Degradation rate of P	1
k_2	γ_O	Degradation rate of O	1
k_3	γ_N	Degradation rate of N	1
O_{critP}	-	Threshold of O on P	1
N_{critP}	-	Threshold of N on P	1
O_{critN}	θ_{NO}	Threshold of O on N	1
N_{critO}	θ_{ON}	Threshold of N on O	1
P_{critN}	-	Threshold of P on N	1
n	h_{OS}	Hill coef of S on O	-1
m	h_{NS}	Hill coef of S on N	-1
-	θ_{OS}	Threshold of S on O	1
-	θ_{NS}	Threshold of S on N	1

Table 2: Corresponding notation for the parameters used in the simplified model of Balaskas et al. [1]

With the following definition of the Hill function, applied multiplicatively for each repressor.

$$\text{Hill}(G, \theta, h) = \frac{1}{1 + (G/\theta)^h} = \begin{cases} \frac{\theta^h}{\theta^h + G^h} & \text{if } h > 0, \theta > 0 \\ \frac{G^h}{\theta^h + G^h} & \text{if } h < 0, \theta > 0 \\ 1 & \text{if } h = 0, \theta = 0 \end{cases} \tag{1}$$

2 Generating Travelling Waves

Assume an input M_1 mediating, under the quasi-static assumption, the output M_2 through the Hill function

$$M_2 = \alpha \frac{M_1^H}{M_1^H + T^H} \tag{2}$$

Let M_1 be an decaying exponentially gradient of the form

$$M_1(x,t) = C_1 \exp(-C_2 x - C_3 t) + C_4 \tag{3}$$

then $M_2(x,t) = M_2(M_1(x,t))$ is the traveling wave shown in figure S1. The necessary condition $\alpha C_1 \gg T$, corresponds to a system initially past the saturation point of the Hill function and ensures that cells initially start with the on-phase of the traveling wave. The constant C_4 sets the minimum concentration of the traveling wave and allowing for non-zero steady-states in the case of activating morphogens. The parameters of figure S1 are given in table S3

Parameter	Value
α	1
H	5
T	.1
C_1	5
C_2	.25
C_3	.5
C_4	0

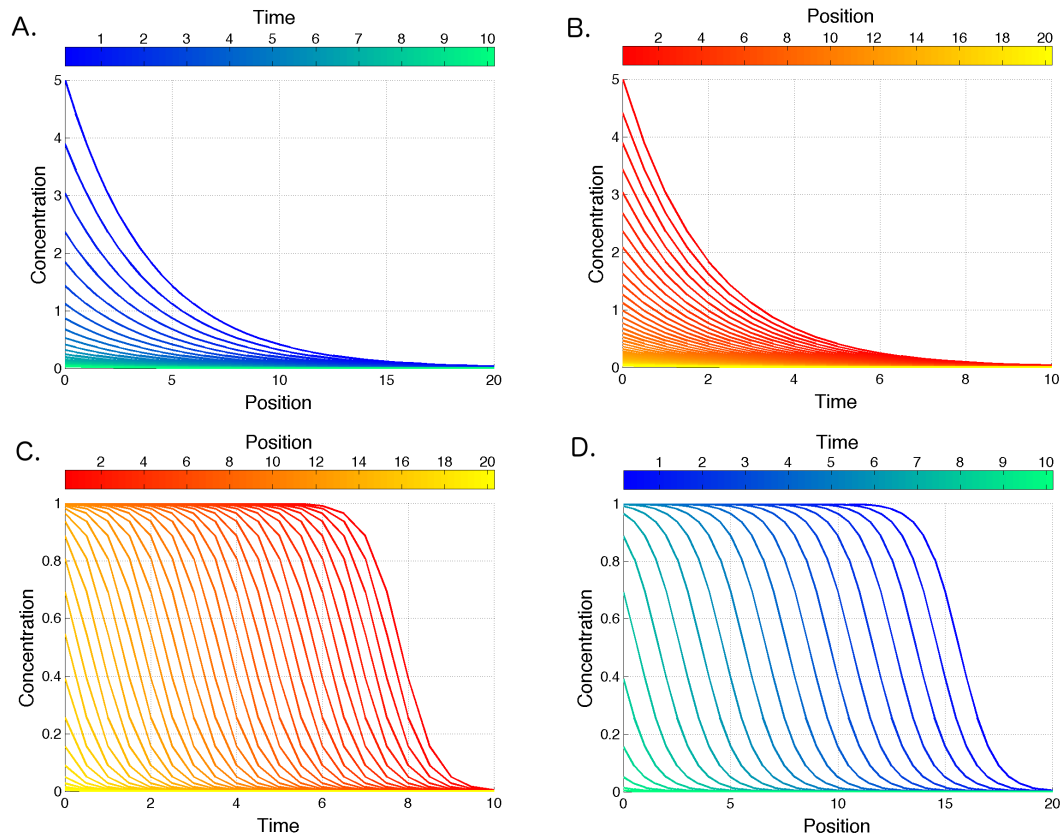Table 3: Parameters used to generate the traveling wave of figure S1

Figure S 1: **Generating travelling waves from decaying exponential gradients.**
Decaying exponential gradients of an input M_1 generate traveling waves of the output M_2
when M_1 activate M_2 through Michaelis-Menten kinetics. A. Snapshots of an exponential
gradient M_1 as a function of position for different times. The blue-green color gradient
labels time from early to late. B. Associated concentration profile to (A) as a function
of time for different positions. The red-yellow color gradient labels the position from
anterior to posterior. C. The output M_2 is a traveling Hill function as a function of
time. The position red-yellow gradient labels different position. D. The M_2 traveling
wave associated to (C) as a function of position. The blue-green gradient labels the time
for each snapshot.

3 Piecewise fit

Figure 3 F. of the paper shows a piecewise fit of the time course. The yellow points highlight point which satisfy the valley condition roughly within 2%. That is

$$|\vec{v}_+ - \dot{\vec{x}}| < 0.02$$

and the gray points correspond to an exponentially decaying gene $G_{i-1} = e^{-t+t_0}$ and the two consecutive genes slaved to the decay by the equations (with Hill replaced by the shorter notation H)

$$t_0 = \text{Time when } G_i \text{ reaches } \theta_{i-1,i} = 0.4$$

$$t_1 = \text{Time when } G_{i-1} \sim 0.0 \text{ and the valley starts}$$

$$G_{i-1}(t) = e^{-t+t_0}$$

$$\dot{G}_i(t) = \alpha_i H(M, \theta_{iM}, h_{iM}) H(G_{i-1}, \theta_{i,i-1}, h_{i,i-1}) H(G_{i+1}(t_1), \theta_{i,i+1}, h_{i,i+1}) - \gamma_i G_i$$

$$\dot{G}_{i+1}(t) = \alpha_{i+1} H(M, \theta_{i+1,M}, h_{i+1,M}) H(G_{i-1}, \theta_{i+1,i-1}, h_{i+1,i-1}) H(G_i(t_1), \theta_{i+1,i}, h_{i+1,i}) - \gamma_{i+1} G_{i+1}$$

Because $G_{i-1}(t)$ is a known function and we replace the full time dependence of G_i on G_{i+1} and G_{i+1} on G_i by their values at time t_1, we obtain two first order ODE in one variable, the solution of which is an hypergeometric function.

Because the valley lives on a 2D plane between G_i and G_{i+1}, not all slow points lie on the valley. As the system approaches the valley, G_{i-1} is strongly repressed by G_i and therefore decays exponentially following e^{-t+t_0} where t_0 is a constant which shifts the curve in time according to which peak is fit. This exponential decay happens perpendicularly to the plane of the valley. As G_{i-1} reaches 0, the points agree with the valley condition. This exponential decay is modeled by the gray lines and it is the transition between two valleys. On the valley, G_i slowly decreases while G_{i+1} slowly increases. Once G_{i+1} reaches $\theta_{i,i+1} = 0.4$ it strongly represses G_i which in turn exponentially decays following e^{-t+t_0}. The cycle thus continues.

4 Stochastic analysis

The cell environment is intrinsically noisy because only a finite number of molecules, η, are interacting. To have a useful model, it needs to be robust to noise, or at least we must offer an argument for how it can be made more robust. There is an obvious problem when considering a system that evolves in time stochastically: time compounds noise, especially in system with sharp transitions such as those we present. The fact that there is a bifurcation with a time scale finely controlled by the proximity to a critical value also means that there is a trade off between noise robustness and the ability to produce domains of variable length. Using the theory of the chemical Langevin equation which we use to numerically simulate noise using the τ Leaping algorithm [2], we show how the introduction of a cell coupling rescues patterning even in very noisy environments ($\eta \sim 500$).

Figures S2 and S3 show ensemble averages of 100 independent trials at a given noise level, controlled by η, the number of molecules in a given volume. The average time course resembles the deterministic time course and in the $\eta \to \infty$ we retrieve the deterministic solution. However, for smaller η, the peaks' amplitudes of gene expression are decaying as a function of time, as the variance increase. As time goes on, the loss of predictive power makes it impossible to predict which state will be expressed at some later time t. Due to the trade off between the critical behaviour at the bifurcation, this worsens as T_{iM} is adjusted to lengthen the domains. To remedy this problem, we therefore propose a model by which cells interact with each other to reduce noise.
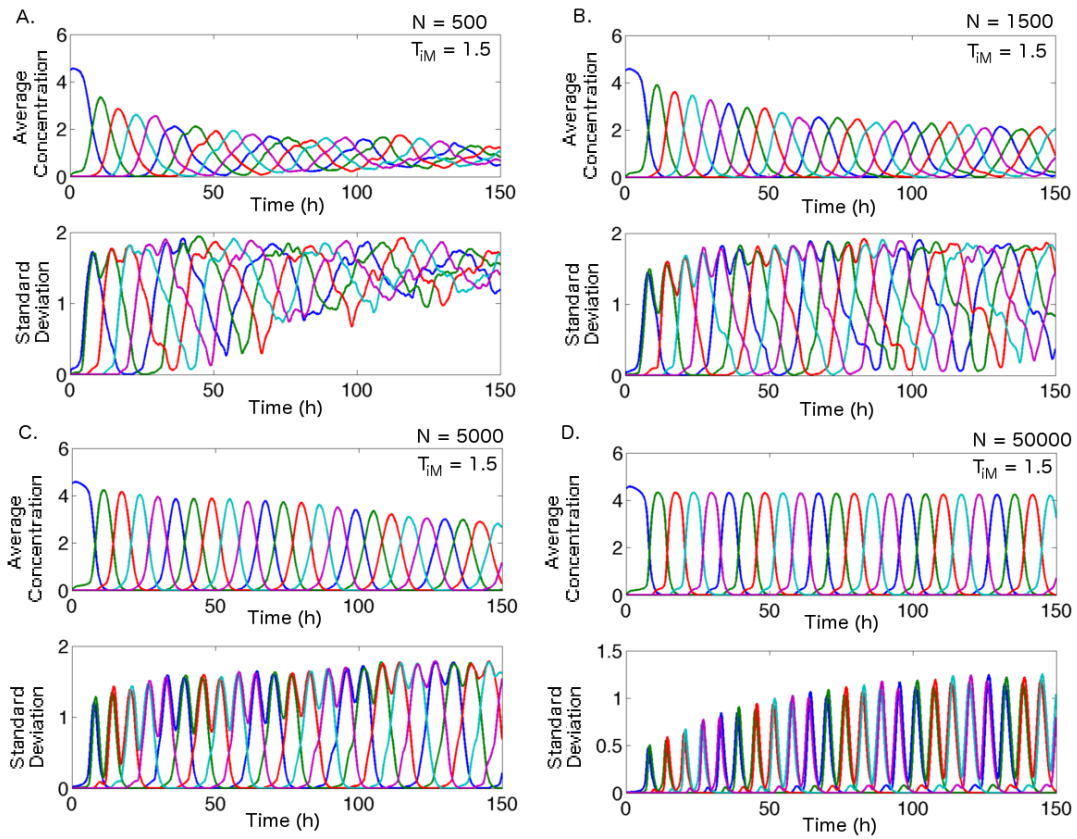
Figure S 2: **Noise sensitivity for fast timescales.** Stochastic ensemble average of the oscillating 5D GRN with a fast timescale, $\theta_{iM} = 1.5$ and different η. The higher η the closer to the deterministic result and the variance becomes constrained to the transition between two genes. For smaller η, the variance increases quickly and the predictive power is lost. A.-D. represent the values of $\eta = 500, 1500, 5000, 50000$.
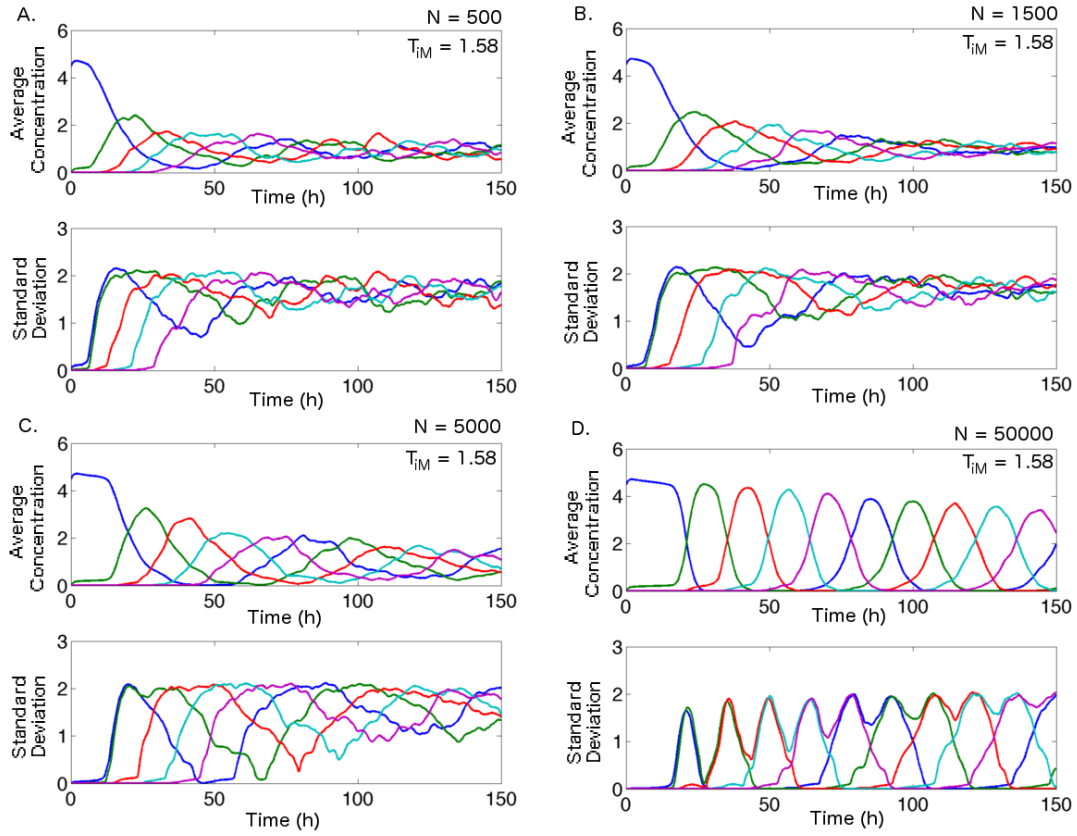
Figure S3: **Noise sensitivity for slow timescales.** Stochastic ensemble average of the oscillating 5D GRN with a slow timescale, $\theta_{iM} = 1.58$ and different η. The GRN with slow timescale is much more sensitive to the noise than the fast one. The higher η the closer to the deterministic result and the variance becomes constrained to the transition between two genes. For smaller η, the variance increases quickly and the predictive power is lost. A.-D. represent the values of $\eta = 500, 1500, 5000, 50000$.

4.1 Cell averaging

We simulate cell development with a two dimensional cell array. Each cell C_{mn} is exposed to a concentration of morphogen M_{mn} which varies in time and space. We model M_{mn} as a traveling wave propagating along the x axis (n index) so that each cell along the y axis (m index) is exposed to the same concentration of M at a given time.

We postulate cell-to-cell interactions. To model cell interactions we will add interaction rates $\mathbf{R}$ to our set of coupled ODEs. The resulting system is not cell autonomous and requires keeping track of every cell individually. To motivate the form of our interaction term, we note that it is known that cells interact with each other through many means such as simple diffusion, endocytosis, the use of transport proteins such proteoglycans or signaling pathways (e.g. Notch). Rather than modeling a particular form of cell-to-cell interaction, we take a phenomenological effective term. We model interactions between two cells using the rate function

$$\mathbf{R}(C_{mn}[G_i] - C_{pq}[G_i], \xi_{int}, \sigma_{int}) = -\frac{\sigma_{int}(C_{mn}[G_i] - C_{pq}[G_i])}{\xi_{int}} e^{-\left(\frac{C_{mn}[G_i]-C_{pq}[G_i]}{\xi_{int}}\right)^2} \quad (4)$$

where $C_{mn}[G_i], C_{pq}[G_i]$ are the concentration of the gene G_i in the two different cells indexed by $(m,n), (p,q)$ in the cell array, ξ_{int} is the length of interaction sets a concentration scale[1] past which the cells do not interact and σ_{int} is the strength of interaction which plays the role of the rate at which the interaction happens.

The purpose of this interaction is to average the difference between two interacting cells. When $C_{mn}[G_i] - C_{pq}[G_i]$ is zero, the function returns zero and there is no interaction. When $|C_{mn}[G_i] - C_{pq}[G_i]| \gg \xi_{int}$ the cells are deemed dissimilar and they do not interact (Figure 4C of main text). If $|C_{mn}[G_i] - C_{pq}[G_i]| \lesssim \xi_{int}$, the cells interact in such a way as to return to their mean. If $C_{mn}[G_i] > C_{pq}[G_i]$, $\mathbf{R}$ will be an effective degradation rate and if $C_{mn}[G_i] < C_{pq}[G_i]$, $\mathbf{R}$ will be an effective production rate. The function $\mathbf{R}$ is plotted in figure 4D of the main text. By construction, for reasonable values of ξ_{int}, σ_{int} , $\mathbf{R}$ cannot change the deterministic steady state because it vanishes at zero and for values much greater than ξ_{int}.

To obtain a local interaction at the level of the embryo, we assume all cells interact with each other with an interaction strength that falls off exponentially with their distance. Define

$$d(C_{mn}, C_{pq}) = \sqrt{(m-p)^2 + (n-q)^2} \quad (5)$$

to be the cartesian distance between the two cells at position (m,n) and (p,q) respectively, in the array. Then they interact with an interaction strength $\tilde{\sigma}_{int} = \sigma_{int} e^{-\frac{1-d}{2}}$ such that neighbor cells for which $d = 1$ interact with strength σ_{int} and cells separated separated by $d > 10$ interact very little.

---

[1] By interaction length, we do not mean physical distance between cell, but difference between gene concentrations.

$$M_{mn}(t) = L_t \frac{T_{mn}^h}{T_{mn}^h + t^h}$$

$$\frac{dC_{mn}[G_i]}{dt} = \alpha_i \text{Hill}(S_{mn}, \theta_{iM}, h_{iM}) \prod_{h_{ij} \leq 0} \text{Hill}(C_{mn}[G_j], \theta_{ij}, h_{ij}) - \lambda_i C_{mn}[G_i] \quad (6)$$

$$+ C_{mn}[G_i] \sum_{p \neq m} \sum_{q \neq n} \mathbf{R}(C_{mn}[G_i] - C_{pq}[G_i], \xi_{int}, \sigma_{int} \, e^{-\frac{d(C_{mn}, C_{pq})-1}{2}})$$

Finally, the deterministic coupled ODEs that represent these interactions is given by Eq. 6 where T_{mn} is the exposure time which depends on cell position (m, n) and $\text{Hill}(G, \theta, h)$ is the Hill function as defined in eq. 1. The chemical Langevin equation is implemented according to the τ-leaping treating the production and degradation terms as noisy reactions. The interaction term is assumed to be exact.

4.2 *In Silico* Evolution selects for ξ_{int}, σ_{int}

The interaction parameters ξ_{int}, σ_{int} have to be carefully chosen to counteract two opposing forces. If the parameters are too small, the averaging effect will be too weak and there will be little improvement in noise robustness. If the parameters are too big, then cells which have been exposed for shorter period of time and have therefore achieved steady state will bias the expression of cells which are still deciding their fate.

To find suitable values for the interaction parameters, we run the an evolution algorithm. To do so, we create a population of embryos (i.e. cell arrays equipped with the GRN) each with different parameters ξ_{int}, σ_{int}. We integrate the the embryo once deterministically without interaction terms to get $\tilde{C}_{mn}[G_i](t)$ and multiple times stochastically with the interaction term to get each time $C_{mn}[G_i](t)$. Embryos are then ranked according to their fitness given by Eq. 7 which measures the RMS deviation from the deterministic timecourse. Survival of the fittest dictates that embryos with bad fitness are replaced by embryos with better fitness. The population is then mutated by slightly changing ξ_{int} and σ_{int} and keeping all other parameters constant.

$$F = -\frac{1}{N_{rows}} \sum_m \frac{1}{N_{cols}} \sum_n \frac{1}{N_{genes}} \sum_i \sqrt{\frac{1}{N_t} \sum_t (\tilde{C}_{mn}[G_i](t) - C_{mn}[G_i](t))^2}$$

$$F_{avg} = \frac{1}{N_{tries}} \sum_{\text{trials}} F \quad (7)$$

For the simulation, we used $\eta = 500, N_{rows} = 100, N_{cols} = 30, N_{genes} = 5, N_t = 5000, N_{tries} = 5$. We set $\theta_{iS} = 1.59$ which gives a long timescale. The best parameters were found to be $\xi_{int} = 0.0588$ and $\sigma_{int} = 0.0300$. Figure 4 of the main text shows the ensemble average of the 30 cells per row for different columns for both the uncoupled and the coupled system. We conclude that cell averaging preserves the timescale and the predictive power which would otherwise be lost to noise.

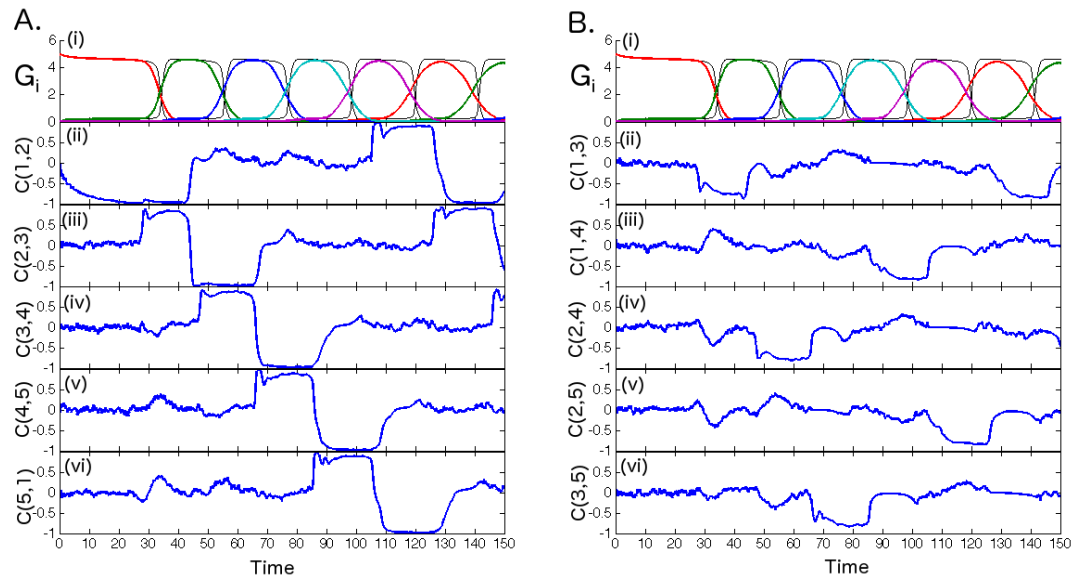# 5 Criticality through the correlations as functions of time

Figure S 4: **Correlations between genes along the time course.** A. Correlations between adjacent genes i and $i + 1$. (i): deterministic time course and average (over 500 runs) stochastic time course with $N = 250000$ molecules and $\theta_{iS} = 1.59$. (ii)-(vi): Correlations between pairs of adjacent genes B. Correlations between non-adjacent genes i and $i + 2$. (i): deterministic time course and average (over 500 runs) stochastic time course with $N = 250000$ molecules and $\theta_{iS} = 1.59$. (ii)-(vi): Correlations between pairs of non-adjacent genes

References

[1] Balaskas N, Ribeiro A, Panovska J, Dessaud E, Sasai N, et al. (2012) Gene Regulatory Logic for Reading the Sonic Hedgehog Signaling Gradient in the Vertebrate Neural Tube. Cell 148: 273–284.

[2] Gillespie DT (2007) Stochastic simulation of chemical kinetics. Annual review of physical chemistry 58: 35–55.