# Studying the effect of cell division on expression patterns of the segment polarity genes

Madalena Chaves* and Réka Albert†

**Abstract**

The segment polarity gene family, and its gene regulatory network, is at the basis of *Drosophila* embryonic development. The network's capacity for generating and robustly maintaining a specific gene expression pattern has been investigated through mathematical modeling. The models have provided several useful insights by suggesting essential network links, or uncovering the importance of the relative timescales of different biological processes in the formation of the segment polarity genes' expression patterns. But the developmental pattern formation process raises many other questions. Two of these questions are analyzed here: the dependence of the signaling protein sloppy paired on the segment polarity genes, and the effect of cell division on the segment polarity genes' expression patterns. This study suggests that cell division increases the robustness of the segment polarity network with respect to perturbations in biological processes.

## 1 Introduction

During the initial stages of development of the fruit fly *Drosophila melanogaster*, three families of genes are successively activated [1]: the gap genes, the pair-rule genes and the segment polarity genes. The resulting gene expression patterns contribute to gradually break the symmetry of the fertilized egg and accompany its transformation into a segmented embryo. Around stages 6 and 7 of embryonic development (i.e., around three hours after fertilization), the family of genes known as the segment polarity genes is activated: their expression patterns will define the position of the parasegmental grooves, the boundaries of the segments which form the body of the fruit fly. The segment polarity genes refine and maintain their expression through the network of intra- and intercellular regulatory interactions shown in Fig. 1. The stable expression pattern of these genes (specifically the expression of *wingless* and *engrailed*) defines and maintains the borders between different parasegments and contributes to subsequent developmental processes, including the formation of denticle patterns and of appendage primordia [2, 3].

To try to understand and study this network and its properties, a first mathematical model was proposed by von Dassow et al. [4]. Some improvements to this model, including an alternative mathematical description [5], and analysis of its properties [6, 7] have recently been presented. A common conclusion from these studies is that the network structure plays a fundamental role: that is, the interconnections among the genes and proteins that constitute the segment polarity network are the crucial factor for the robustness of the expression pattern with respect to biological ("small") perturbations.

In [5], Albert and Othmer proposed a Boolean version of the continuous model described in [4]. Boolean models provide a qualitative representation of a system, consisting of the nodes of the network (whose values are either 0="OFF" or 1="ON") and a set of logical rules to describe the regulatory links among them (activation or inhibition interactions). The choice of Boolean modeling in this context is very natural, as

*COMORE, INRIA, 2004 Route des Lucioles - BP 93, 06902 Sophia Antipolis, France, mchaves@sophia.inria.fr

†Department of Physics and Huck Institutes for the Life Sciences, Pennsylvania State University, University Park, PA 16802 USA, ralbert@phys.psu.edu

many genetic regulatory functions are known, but hardly any kinetic or binding parameters are available for the segment polarity network. Advantages of using Boolean rules include a clear modeling of the structure of interactions (that is, the links among the system's nodes) and very intuitive qualitative representation of the system and its behavior. In addition, various analytical methods can be used to study Boolean models [8, 9, 10, 11].

The Boolean model proposed in [5] improves on [4] by adding an activating signal that initiates expression of the segment polarity genes. This is the protein sloppy paired (SLP) which is part of the pair-rule gene family. (Actually, there are two different proteins encoded by two genes *sloppy paired*, but they are known to have similar roles, and thus are referred from now on as only one protein, SLP.) Further analysis of this model has provided many useful insights and contributed to better understanding the segment polarity network [11, 7], namely, it has shown the importance of considering different timescales for different biological processes (for instance, transcription and translation are generally slower than protein conformational changes [12]) in the correct formation of the segment polarity genes' expression. It has also raised further questions, so in the present work we have chosen to focus on two important issues: (i) the SLP activating signal has been assumed constant; however, evidence shows [13] that it can also depend on the expression of some segment polarity genes. One of our goals is to introduce a Boolean rule to describe SLP, and thus obtain a more autonomous module; (ii) most cited models of the segment polarity network assume four cell wide segments. However, evidence shows that there are rounds of cell division at stages 8 and 10 [3, 14]. We will investigate the effect of segment width on the robustness properties of the segment polarity network.

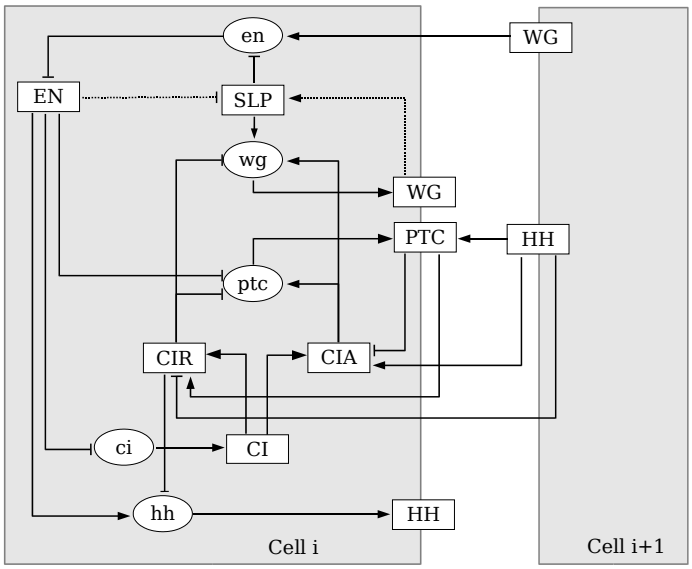


Figure 1: The diagram of interactions for the segment polarity network. Squares represent proteins and elipses represent mRNAs. Cell-to-cell communication is considered among nearest neighbours, through the wingless and hedgehog proteins. Only neighbouring cell $i+1$ is depicted, but both cells $i-1$ and $i+1$ have a similar effect on cell $i$. This model was developed in [5], except for the activation and inhibition links on protein SLP, which are introduced in the current work. See main text for more details.

## 2 Boolean models for genetic networks

In the model, each mRNA or protein is represented by a node of a network, and the interactions between them are encoded as directed edges (see Figure 1). The state of each node is $1$ or $0$, depending on whether

the corresponding substance is present or not. The states of the nodes can change in time, and the next state of node $\ell$ is determined by a Boolean (logical) function $F_\ell$ of its state and the states of those nodes that have edges incident on it. In general, a Boolean or logical function is written as a statement acting on the inputs using the logical operators "and", "or" and "not" and its output is 1 (0) if the statement is true (false).

The functions determining the state of each node are constructed from the interactions between nodes (such as displayed in Fig. 1) according to the following rules:

(i) mRNAs/proteins are synthesized if their transcriptional activators/mRNAs are present;

(ii) the effect of transcriptional activators and inhibitors is never additive, but rather, inhibitors are dominant;

(iii) mRNAs decay in the next updating step if not transcribed;

(iv) transcription factors and proteins undergoing post-translational modification decay if their mRNA is not present.

Consider a regulatory network with $N$ nodes, $X_1, \ldots, X_N$. The expression of each node along time can be computed by iterating the Boolean rules, to obtain a discrete sequence:

$$X_\ell(0), X_\ell(1), \ldots, X_\ell(k), \ldots,$$

where $X_\ell(k)$ denotes the expression of node $\ell$ at time instant $kT$, where $T$ represents a (fixed) time unit. To compute $X_\ell(k)$, using $X$ at the previous instants, several algorithms are available to iterate the Boolean rules. Some of these are briefly described next. The standard synchronous algorithm assumes that all $N$ nodes are simultaneously updated, that is:

$$X_\ell(k+1) = F_\ell(X_1(k), \ldots, X_N(k)), \quad \ell = 1, \ldots, N.$$

Asynchronous algorithms allow different nodes to be updated at different times, for example according to a random order. Assuming that all nodes are updated exactly once during each time unit $T$, an asynchronous algorithm can be constructed by randomly assigning an updating order at each iteration. For example, if $P = (P_1, \ldots, P_N)$ is a permutation of $\{1, \ldots, N\}$, let node $\ell$ be the $P_\ell$th node to be updated in the $k$th iteration of the rules. Then

$$X_\ell(k+1) = F_\ell(X_1(\tau_{1,k}), \ldots, X_N(\tau_{N,k})), \quad \ell = 1, \ldots, N.$$

where $\tau_{1,k} = k$ if $P_1 > P_\ell$ and $\tau_{1,k} = k+1$ if $P_1 < P_\ell$, that is, use $X_1(k)$ if node 1 should be updated later that node $\ell$, and use $X_1(k+1)$ if node 1 was updated already. The order of node updating may be different at each iteration $k$, since a new permutation $P_k$ can be randomly generated once all nodes have been updated exactly $k-1$ times. Other asynchronous algorithms can be developed, for example by choosing the permutation $P_k$ according to some criteria. In [11], one of the criteria consisted of choosing permutations where all protein nodes are updated first, and then all the mRNA nodes.

The steady states of a Boolean model ($\bar{X}$) are fixed points of the vector function $F$, and consist of patterns which do not change with model updating. They can be found by solving the equations

$$\bar{X}_\ell = F_\ell(\bar{X}_1, \ldots, \bar{X}_N), \quad \ell = 1, \ldots, N.$$

It is not difficult to check that both synchronous and asynchronous updating schemes have the same steady states. However, note that different asynchronous algorithms may lead the system from the same initial condition to different steady states.

### 2.1 Analysis of Boolean networks

Useful techniques are available for the analysis of discrete logical models. In particular, Glass [15] introduced a class of piecewise linear differential equations that combine logical rules for the synthesis of gene products with linear (free) decay by describing each node with two variables, one discrete ($X_\ell$) and one continuous ($\hat{X}_\ell$). For each node, a specific timescale ($\alpha_\ell > 0$) is also assigned. In a first approach, in equation (1), $\alpha_\ell$ represents both a degradation and a synthesis rate – or a turnover rate. The model can be extended to allow distinct synthesis and degradation rates. From the set of Boolean rules $F_\ell(X)$, $\ell = 1, \ldots, N$ a piecewise linear model can be obtained in the form:

$$\frac{d\,\hat{X}_\ell}{dt} = \alpha_\ell(-\hat{X}_\ell + F_\ell(X_1, X_2, \ldots, X_N)), \quad \ell = 1, \ldots, N, \tag{1}$$

At each instant $t$, the discrete variable $X_\ell$ is defined as a function of the continuous variable according to a threshold value:

$$X_\ell(t) = \begin{cases} 0, & \hat{X}_\ell(t) \leq \theta_\ell \\ 1, & \hat{X}_\ell(t) > \theta_\ell \,, \end{cases} \tag{2}$$

where $\theta_\ell \in (0, 1)$ defines the fraction of "maximal concentration" necessary for a protein or mRNA to regulate its successor nodes. As detailed in [7], the parameters $\alpha_\ell$ represent different timescales for different biological processes (transcription, translation or post-translational modifications). Note that $\alpha_\ell$ is also a scaling factor of the differential equation for $\hat{X}_\ell$. In fact, since solutions are piecewise increasing or decreasing exponentials, the evolution of $\hat{X}_\ell$ is governed by the term $\exp(-\alpha_\ell t)$. So, higher values of $\alpha_\ell$ indicate that the variation rate of $\hat{X}_\ell$ is higher. It is easy to see that the steady states of the piecewise linear equations (1) are still those of the Boolean model, since:

$$\frac{d\,\hat{X}_\ell}{dt} = 0 \;\Leftrightarrow\; \hat{X}_\ell = X_\ell = F_\ell(X_1, X_2, \ldots, X_N),\; \ell = 1, \ldots, N,$$

independently of $\theta_\ell$.

To study the effect of biological perturbations on the system, a natural way to proceed is to randomly assign values to time scaling parameters $\alpha_\ell$, and numerically solve equations (1). Variations in $\alpha_\ell$ represent perturbations in the relative timescales of each process: for instance, if $\alpha_{P1} > \alpha_{P2}$, then the total rate of production of protein P1 is faster than that of protein P2. Different combinations of $\{\alpha_1, \ldots, \alpha_N\}$ represent different scenarios for the segment polarity genes. By allowing $\alpha_\ell$ to take values in an interval $A_\ell$, it is possible to explore the space of biological fluctuations.

## 3 A Boolean model for the segment polarity network

As depicted in Fig. 1, the main segment polarity genes are *wingless* (*wg*), *engrailed* (*en*), *hedgehog* (*hh*), *patched* (*ptc*), and *cubitus interruptus* (*ci*) (see, for instance, [1, 16, 17]). These code for their corresponding proteins (which will be respectively represented by the symbols WG, EN, HH, PTC, and CI). The protein cubitus interruptus may be converted into a transcriptional activator (CIA), or may be cleaved to form a transcriptional repressor (CIR). The proteins EN, CIA and CIR are transcription factors, while WG and HH are secreted proteins, and PTC is a transmembrane receptor protein.

The pair-rule gene product SLP activates *wg* transcription and represses *en* transcription. The WG protein is secreted from the cells that synthesize it [3, 18] and initiates a signaling cascade leading to the transcription of *en* [19]. EN promotes the transcription of the *hh* gene [20] and represses the transcription of *ci* [21] and possibly *ptc* [22, 23]. The HH protein is also secreted, and binds to the HH receptor PTC on a neighboring cell [24]. The intracellular domain of PTC forms a complex with SMO [25] in which SMO is

inactivated by a post-translational conformation change [26]. Binding of HH to PTC removes the inhibition of SMO, and activates a pathway that results in the modification of CI [26]. The CI protein can be converted into one of two transcription factors, depending on the activity of SMO. When SMO is inactive, CI is cleaved to form CIR, a transcriptional repressor that represses *wg, ptc* [27] and *hh* transcription [28, 29]. When SMO is active, CI is converted to a transcriptional activator, CIA, that promotes the transcription of *wg* and *ptc* [27, 29, 30, 31].

The expression pattern of these genes and proteins is repeated periodically along the embryo, and defines the parasegmental grooves. In wild type embryos, the boundaries of the parasegments form between two consecutive cells, with one cell expressing *wingless* immediately anterior (to the left) to a cell expressing *engrailed* [3]. At stages 6-7 of embryonic development, the parasegments are about four cells wide, and *wingless* mRNA is expressed in one of four cells, *engrailed* and *hedgehog* also in one of four cells, immediately posterior (to the right) to the cells expressing *wg*. *Cubitus* and *patched* mRNA are typically expressed in all cells but those expressing *en*. The corresponding proteins will later follow these patterns. Typically, CIA is present and CIR absent in cells expressing *wg*, and either CIA is absent or CIR is present in cells not expressing *wg*.

### 3.1 Notation

Before proceeding, it is useful to introduce some notation conventions. Our model of the segment polarity network (and others mentioned), describes the evolution of a family of mRNAs and proteins, in each cell of a parasegment of the embryo. The length of each parasegment is denoted by $M$. At stages 6-7 of development $M = 4$, but at later stages, parasegments with $M \geq 4$ cells will be considered. The notations for the model's variables encode the node name, a spatial coordinate (cell number) and a time instant. For example $wg_i(k)$, $i = 1, \ldots, M$ denotes the *wingless* mRNA concentration at time instant $k$ in the $i$th cell of a parasegment; it is also convenient to write in short $wg(k)$, to denote the vector $(wg_1(k), \ldots, wg_M(k))$. Similar notation is adopted for the all the other mRNAs and proteins that form the segment polarity network (as listed above).

Periodic boundary conditions are assumed, meaning that: $node_{M+1} = node_1$ and $node_{1-1} = node_M$.

### 3.2 A new rule for SLP

The logical rule adopted for SLP in [5] summarizes, in a simple but effective way, experimental observations on the regulatory activity of sloppy paired protein in the segment polarity network:

$$\text{SLP}_i(k+1) = \begin{cases} 0 & \text{if } i \in \{1,2\} \\ 1 & \text{if } i \in \{3,4\} \end{cases} \tag{3}$$

One other possible rule for SLP was studied in [7], to include recent evidence of engrailed protein inhibiting *sloppy paired* transcription [13]. Additional regulation (again in the form of a constant input, RX) was represented by a combination of several possible effects from the pair-rule genes, namely *runt, opa* and Factor X [32] and of *slp* autoregulation:

$$\begin{aligned} \text{RX}_i(k+1) &= \begin{cases} 0, & \text{if} \quad i \in \{1,2\} \\ 1, & \text{if} \quad i \in \{3,4\} \end{cases} \\ slp_i(k+1) &= \text{RX}_i(k) \text{ and not } \text{EN}_i(k) \\ \text{SLP}_i(k+1) &= slp_i(k)\ , \end{aligned} \tag{4}$$

for $i = 1, \ldots, 4$. It is clear that both (3) and (4) impose somewhat strong conditions on the network, by assuming constant values on SLP or RX. To alleviate this constraint, and incorporate the feedback of the segment polarity network on the sloppy paired protein, a reasonable hypothesis seems to include activation

by *wingless*, as well as the observed inhibition by *engrailed*. This hypothesis is supported by some of the results reported in [33, 34]. A similar modeling approach was already considered in [35] and [6]. However, solely assuming activation by *wingless* and inhibition by *engrailed* does not explain why the domain of expression of SLP is wider than the domain of *wingless* and narrower than the domain where *engrailed* is absent, thus it is necessary to include the possibility of maintaining an initial prepattern. Thus the following rule for SLP is proposed:

$$\text{SLP}_i(k+1) = (\text{WG}_i(k) \text{ and not } \text{EN}_i(k)) \text{ or } \text{SLP}_i(k), \quad i = 1, \ldots, M. \tag{5}$$

To test feasibility of this rule, the resulting segment polarity model will be analyzed, and predictions for several scenarios will be given, to be compared to biological observations. The Boolean rules for the other nodes are unchanged from those in [5], and represent the network of interactions described above. A graphical representation of the model is given in Fig. 1, and the equations are summarized in Table 1. Note that SMO doesn't appear in the equations. But its expression is given by [5]:

$$\text{SMO}_i(k) = \text{not PTC}_i(k) \text{ or } \text{HH}_{i-1}(k) \text{ or } \text{HH}_{i+1}(k) \text{ or } \textit{hh}_{i-1}(k) \text{ or } \textit{hh}_{i+1}(k),$$

and substituted directly into the Cubitus activator and repressor proteins:

$$\text{CIA}_i(k+1) = \text{CI}_i(k) \text{ and } \text{SMO}_i(k), \;\; \text{CIR}_i(k+1) = \text{CI}_i(k) \text{ and not } \text{SMO}_i(k).$$

The model now represents an autonomous network, with no external signals apart from the initial condition, which represents the known expression pattern at initiation of the segment polarity genes. It is now possible to investigate the mechanisms leading to the final segment polarity genes expression pattern (as observed for stages 8-10).

Table 1: Regulatory functions governing the states of segment polarity gene products in the model. Each node is labeled by its biochemical symbol and subscripts signify cell number in a segment. The dynamics of the system is evaluated according to $X(k+1) = F(X(k))$, for $k = 0, 1, \ldots$. For a system with $M$ cells there are thus $N = 13M$ variables, and $X$ is a vector in $\{0,1\}^N$.

| Node | Boolean updating function |
|---|---|
| $\text{SLP}_i$ | $\text{SLP}_i(k+1) = (\text{WG}_i(k) \text{ and not } \text{EN}_i(k)) \text{ or } \text{SLP}_i(k)$ |
| $\textit{wg}_i$ | $\textit{wg}_i(k+1) = (\text{CIA}_i(k) \text{ and } \text{SLP}_i(k) \text{ and not } \text{CIR}_i(k))$ or $[\textit{wg}_i(k) \text{ and } (\text{CIA}_i(k) \text{ or } \text{SLP}_i(k)) \text{ and not } \text{CIR}_i(k)]$ |
| $\text{WG}_i$ | $\text{WG}_i(k+1) = \textit{wg}_i(k)$ |
| $\textit{en}_i$ | $\textit{en}_i(k+1) = (\text{WG}_{i-1}(k) \text{ or } \text{WG}_{i+1}(k)) \text{ and not } \text{SLP}_i(k)$ |
| $\text{EN}_i$ | $\text{EN}_i(k+1) = \textit{en}_i(k)$ |
| $\textit{hh}_i$ | $\textit{hh}_i(k+1) = \text{EN}_i(k) \text{ and not } \text{CIR}_i(k)$ |
| $\text{HH}_i$ | $\text{HH}_i(k+1) = \textit{hh}_i(k)$ |
| $\textit{ptc}_i$ | $\textit{ptc}_i(k+1) = \text{CIA}_i(k) \text{ and not } \text{EN}_i(k) \text{ and not } \text{CIR}_i(k)$ |
| $\text{PTC}_i$ | $\text{PTC}_i(k+1) = \textit{ptc}_i(k) \text{ or } (\text{PTC}_i(k) \text{ and not } \text{HH}_{i-1}(k)$ and not $\text{HH}_{i+1}(k))$ |
| $\textit{ci}_i$ | $\textit{ci}_i(k+1) = \text{not } \text{EN}_i(k)$ |
| $\text{CI}_i$ | $\text{CI}_i(k+1) = ci_i(k)$ |
| $\text{CIA}_i$ | $\text{CIA}_i(k+1) = \text{CI}_i(k) \text{ and [not } \text{PTC}_i(k) \text{ or } \text{HH}_{i-1}(k)$ or $\text{HH}_{i+1}(k) \text{ or } \textit{hh}_{i-1}(k) \text{ or } \textit{hh}_{i+1}(k)]$ |
| $\text{CIR}_i$ | $\text{CIR}_i(k+1) = \text{CI}_i(k) \text{ and } \text{PTC}_i(k) \text{ and not } \text{HH}_{i-1}(k)$ and not $\text{HH}_{i+1}(k) \text{ and not } \textit{hh}_{i-1}(k) \text{ and not } \textit{hh}_{i+1}(k)$ |

The initial wild type pattern (stage 7) is shown in Table 2. Starting from this initial condition, the final wild type pattern achieved by the segment polarity genes (stages 8-10) is of the form WT in Table 3, which

is indeed a steady state of the Boolean model, that is, this pattern satisfies $X = F(X)$. The mathematical model admits several other steady states, some of which can be identified with mutant patterns (these are listed in Tables 7 to 10, and some patterns are illustrated in Fig. 6 of the Appendix; see also [5, 7]). For instance, there is a non-segmented pattern (NS, *en* mutants), a state with a broad *wg* stripe (BS, *ptc* mutants), or an ectopic state (EC), where the boundary of the parasegment is displaced and inverted.

Table 2: The initial wild type pattern (stage 7). Nodes not indicated are set to zero.

| Node | initial pattern |
|---|---|
| SLP$(0)$ | 0011 |
| *wg*$(0)$ | 0001 |
| *en*$(0)$ | 1000 |
| *hh*$(0)$ | 1000 |
| *ptc*$(0)$ | 0111 |
| *ci*$(0)$ | 0111 |

The challenge is then to study robustness of the convergence of the system $X(k+1) = F(X(k))$ to the desired wild type state pattern WT, in Table 3. For the system with the simpler SLP rule (3), Albert and Othmer [5] verified that synchronous updating rules starting from the initial condition in Table 2, always lead to convergence of the state of the system to WT (Table 3). However, introducing totally asynchronous updating rules (which can be viewed as random perturbations to the timescales of the system) we have shown in [11] that there is a significant probability (about 40%) that the final state of the system is one of the "mutant" states depicted in Table 3. But if the asynchronous updating rules satisfy some ordering criteria, for instance in each round of iterations all proteins are upated first and then the mRNAs (corresponding to a separation of timescales of the various biological processes, with the protein binding and other kinetic events happening faster than transcription or translation), the convergence to the correct wild type pattern is always higher than $87.5\%$. Therefore, the system robustly generates and maintains the segment polarity pattern, for a large range of biological perturbations. In Sections 5 and 6, similar results are stated for the new, extended network, with SLP rule (5).

Table 3: Mathematical steady state patterns for the Boolean model with $M = 4$.

| Node | WT | NS | BS | EC |
|---|---|---|---|---|
| SLP$(\infty)$ | 0011 | 0011 | 0011 | 0011 |
| *wg*$(\infty)$ = WG$(\infty)$ | 0001 | 0000 | 0011 | 0010 |
| *en*$(\infty)$ = EN$(\infty)$ | 1000 | 0000 | 1100 | 0100 |
| *hh*$(\infty)$ = HH$(\infty)$ | 1000 | 0000 | 1100 | 0100 |
| *ptc*$(\infty)$ | 0101 | 0000 | 0011 | 1010 |
| PTC$(\infty)$ | 0111 | 1111 | 0011 | 1011 |
| *ci*$(\infty)$ = CI$(\infty)$ | 0111 | 1111 | 0011 | 1011 |
| CIA$(\infty)$ | 0101 | 0000 | 0011 | 1010 |
| CIR$(\infty)$ | 0010 | 1111 | 0000 | 0001 |

## 4 Rounds of cell division

The four-cell initial condition in Table 2 is representative for up to the beginning of stage 8. There is evidence showing that several rounds of cell divison occur during stages 8 to 10, summarized in Table 4 (for more details, see [5] and references therein). During stage 8 (3h10m-3h40m after fertilization) there is an asynchronous round of cell division, and a second round of cell division in stage 10 (4h20m-5h20m after

fertilization), according to Hooper and Scott [3]. Slightly different observations by Gonzalez et al. [14], indicate that the parasegment is about six cells wide in stage 8-9, and about eight cells wide at stage 10.

Table 4: Approximate timeframe for segment polarity genes' activation.

| Developmental stage | Hours from fertilization | Cells per parasegment | Expression pattern | Refs. |
|---|---|---|---|---|
| 7 | 3h00m | 4 | initial WT, Table 2 | [3] |
| 8 | 3h10m-3h40m | 4-5 | | [3] |
| 9 | 3h40m-4h20m | 5-6 | final WT, Table 6 | [14] |
| 10 | 4h20m-5h20m | 6-8 | final WT, Table 6 | [3],[14] |

These observations lead us to investigate the model in the case when the parasegment is not four, but five, six or eight cells wide. We will not consider the case of a growing parasegment, which would require the introduction of spatial variables in the model, and hence other mathematical tools. Instead, we will take advantage of the fact that the rules governing the segment polarity model (depicted in Table 1) are valid for *any number of cells in a parasegment*, as long as that number is fixed along time. Thus, we study time trajectories, from an initial condition to a steady state, for a fixed number of cells. We explore different initial conditions that represent the parasegment at different stages of cell division. Our goal is to check whether an analogous pattern is generated, independently of the fixed number of cells in a parasegment. If analogous patterns are generated, then one can check their robustness as a function of the number of cells. For instance, we will try to answer the question of whether a parasegment with six or eight cells is more or less vulnerable to environmental fluctuations than a parasegment with four cells.

The first question to be addressed is the initial condition: how to represent the initial stripes (the pattern in Table 2) in a wider parasegment? As a first approach, either of the four cells may divide in the first round, and it may be expected that the new daugther cell retains the expression levels of its mother cell. There are thus four possible initial conditions to consider, depending on whether the first, second, third or fourth cell divides from the state in Table 2. One should also consider that, when the first round of cell division starts,

Table 5: Possible initial conditions after first round of cell division.

| Dividing cell | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| SLP$(0)$ | 00011 | 00011 | 00111 | 00111 |
| *wg*$(0)$ | 00001 | 00001 | 00001 | 00011 |
| *en*$(0)$ | 11000 | 10000 | 10000 | 10000 |
| *hh*$(0)$ | 11000 | 10000 | 10000 | 10000 |
| *ptc*$(0)$ | 00111 | 01111 | 01111 | 01111 |
| *ci*$(0)$ | 00111 | 01111 | 01111 | 01111 |

the final pattern may be already (partly or fully) established (see Table 4). Thus, another possibility is to start from pattern WT in Table 3, and again study the four cases arising from division of each cell. Using the analysis method described in Section 2.1, the results for these two limiting situations are shown in Fig. 2.

For establishing the final pattern, some indications can be found in the literature. According to [36], during stages 8-10, the SLP stripe is adjacent and anterior to *en*, overlapping *wg* and extending anterior. At stage 10 (when the parasegment may be about 8 cells), the SLP stripe is 3-4 cells wide. This suggests that SLP is expressed at least in the last two cells and at most in the last half of the parasegment. According to the embryo stains shown in [13], at stage 11 (when presumably there are more than eight and up to 16 cells per parasegment), the ratio of "expressing" to "not expressing" cells for *slp* is 1:1.5. Similar ratios for other segment polarity genes include *en* and *hh* at 1:4, *wg* at 1:6, and *ci* at 3:1.

Putting together all these observations, it is reasonable to consider (in a five cell segment) single cell bands for *wg*, *en* and *hh*. For SLP one may consider two alternatives: either a two cell band or a three cell band. Indeed, for both a two or a three cell SLP band, there are steady state solutions of the system depicted in Table 1 with $i = 1, \ldots, 5$ (that is, patterns satisfying $X = F(X)$), which correspond to wild type expression patterns. All the mathematical steady states for the Boolean model which are reachable from initial conditions with SLP $= 0 \cdots 011$ or SLP $= 0 \cdots 0111$, for any $M \geq 4$ are given in Tables 7 to 10 of the Appendix. For $M = 5$, the wild type patterns are depicted as WT (I) and (II) in Table 6.

Table 6: Some of the mathematical steady state patterns for the Boolean model with $M = 5$.

| Node | WT (I) | NS (I) | BS (I) | WT (II) | NS (II) | BS (II) |
|---|---|---|---|---|---|---|
| SLP($\infty$) | 00011 | 00011 | 00011 | 00111 | 00111 | 00111 |
| *wg*($\infty$) = WG($\infty$) | 00001 | 00000 | 00011 | 00001 | 00000 | 00101 |
| *en*($\infty$) = EN($\infty$) | 10000 | 00000 | 10100 | 10000 | 00000 | 11000 |
| *hh*($\infty$) = HH($\infty$) | 10000 | 00000 | 10100 | 10000 | 00000 | 11000 |
| *ptc*($\infty$) | 01001 | 00000 | 01011 | 01001 | 00000 | 00101 |
| PTC($\infty$) | 01111 | 11111 | 01011 | 01111 | 11111 | 00111 |
| *ci*($\infty$) = CI($\infty$) | 01111 | 11111 | 01011 | 01111 | 11111 | 00111 |
| CIA($\infty$) | 01001 | 00000 | 01011 | 01001 | 00000 | 00101 |
| CIR($\infty$) | 00110 | 11111 | 00000 | 00110 | 11111 | 00010 |

Other patterns compatible with the steady states of the five cell-wide model include five-cell versions of the non-segmented state (NS), the state with a broad *wg* stripe (BS), or the ectopic state (EC) (not shown). (Compare Tables 3 and 6.) Again, note that there are two alternative forms for each pattern, depending on the width of the SLP band (marked I and II, respectively). Observe that, while the WT and NS patterns differ only in the SLP expression, the effect on BS is more complex, with the *en* and *wg* bands forming in different ways, and thus also affecting expression of the other genes and proteins.

It is interesting to note that comparison of the two (mathematical) BS steady state patterns with experimental data suggests that the patterns with a shorter SLP band (two cells) are biologically more relevant. In fact, the BS pattern is typically seen in *ptc* mutant embryos, with the obvious difference that *ptc* and PTC are not expressed in *ptc* mutants, while they are in the BS pattern. As observed in the work of Martinez Arias et al. [37], in the *ptc* mutant embryos *wingless* is expressed in a broad domain which occupies half the parasegment. After cell division, *engrailed* is observed to form an extra stripe just anterior to *wingless*. This is further supported by work on *ptc* mutants by Ingham et al. [38], who report that the expression domain of *wg* broadens after gastrulation, and that *en* transcription is induced in the cells anterior to the *wg* domain. This experimental evidence points to the state of the form I as the more plausible BS pattern. Hence, more generally, states with SLP $= 00011$ represent the biological system more faithfully.

The patterns in Table 6 extend (in an analogous form) as steady states of an $M$ cell-wide parasegment model (see Tables 7 to 10 in the Appendix). As in the five cell model, we only consider two or three cell SLP bands.

# 5 Results

The method described in Section 2.1 was used to analyze the evolution of the segment polarity pattern, after the first and second rounds of cell division during stages 8 to 10 of embryonic development. The model described in Table 1 was solved numerically with $M$ cells (and $i = 1, \ldots, M$), for $M = 4, 5, 6, 8$. For each biological case studied, a given initial condition was chosen (as indicated in the figure captions). To answer robustness questions, the timescale constants $\alpha_\ell$ were randomly chosen from a uniform distribution

in intervals of the form:

$$\alpha_\ell \in \left[\frac{1}{n}, n\right],$$

where the case $n = 1$ is equivalent to the synchronous Boolean algorithm, and larger values $n$ =2,3,4,6,10,20 (as indicated in the $x$-axis of Figs. 2, 3, 4) represent different orders of magnitude of the perturbations. Therefore, $n = 2$ represents a system where the relative timescales of different processes all have values between half and twice a normalized time unit (here 1), leading to small fluctuations. In contrast, $n = 10$ represent a system where very large fluctuations are allowed, between one tenth and ten times the normalized unit.

For each case studied (that is for each $M$ and each $n$), $j = 1, \ldots, 1000$ numerical experiments were performed. For each numerical experiment $j$, a set of timescales $\{\alpha_\ell^j,\ \ell = 1, \ldots, N\}$ was chosen from a uniform distribution on the interval $[1/n, n]$, the threshold values were set to $\theta_\ell = 0.5$ for all $\ell$, and system (1) was numerically solved. The solution was observed to reach a steady state $\bar{X}(j)$ (from those indicated in Tables 7 to 10 of the Appendix), and this steady state was registered (either of the form WT, BS, NS, EC). Each experiment $j$ represents a certain biological "scenario": the values of $\alpha_\ell^j$ indicate which nodes evolve at faster rates than others (see discussion in Section 2.1). The distribution of $\alpha_\ell^j$ can be viewed as the result of biological perturbations due to, for instance, temperature changes or other stresses, which may induce delays in expression of some genes, or prompt faster signaling processes. Since the values $\alpha_\ell^j$ are randomly chosen, no specific scenario is modeled, but instead all possible relative changes are explored. The goal is to measure the response of the system to "worst case" disruptions, by verifying how frequently the final steady state deviates from the correct wild type pattern.

Results shown in the figures represent the probability of the system reaching a given steady state, that is the frequency of each steady state over the 1000 replicate experiments:

$$P_{\text{WT}} = \frac{1}{1000} \sum_{j=1}^{1000} H(\bar{X}(j), \text{WT})$$

where $H(\bar{X}(j), \text{WT}) = 1$ if $\bar{X}(j) = \text{WT}$, and $H(\bar{X}(j), \text{WT}) = 0$ otherwise (here WT is either of form I or II). Similar expressions were used to compute the probabilities of the other patterns.

The results for the first round of cell division are shown in Fig. 2. As described in Section 4, two limiting cases were studied: either cell division occurs early, and system starts from one of initial conditions in Table 5; or the final pattern is already established, and the system starts from one of the four possible conditions arising by doubling each cell in the pattern WT (Table 3).

It is clear that the expression pattern is not greatly disrupted by cell division, if it was already established with only four cells. It is interesting to observe that no broad striped (BS) pattern is formed at this stage, that is, any perturbation either has no effect on groove formation (with probability higher than 90%), or it completely prevents groove formation (probability less than 10%). In contrast, if the pattern was not yet established with four cells, then there is still a significant possibility that strong perturbations to the system cause pattern disruption. In particular, division of the SLP-expressing cells (the two last) is more vulnerable, as the outcome may be the broad striped pattern (corresponding to *ptc* mutants) with probability of about 10%, or the non-segmented pattern (corresponding to *en* mutants) with probability between 10% and 20%.

The second round of cell division most likely occurs at stages 9-10 of embryonic development, when the wild type pattern of the segment polarity network is already fully established in segments composed of five cells (Table 4). Thus, as a starting condition for the six cell parasegment model, we take the WT (I) pattern in Table 6 and double each cell in turn – thus generating five possible initial conditions. Fig. 3 shows that the network appears quite robust at this stage, in the sense that the pattern extends correctly into six cells with a probability higher than 95%. Moreover, no cell position is more vulnerable than another (in contrast to the extension from four to five cells).

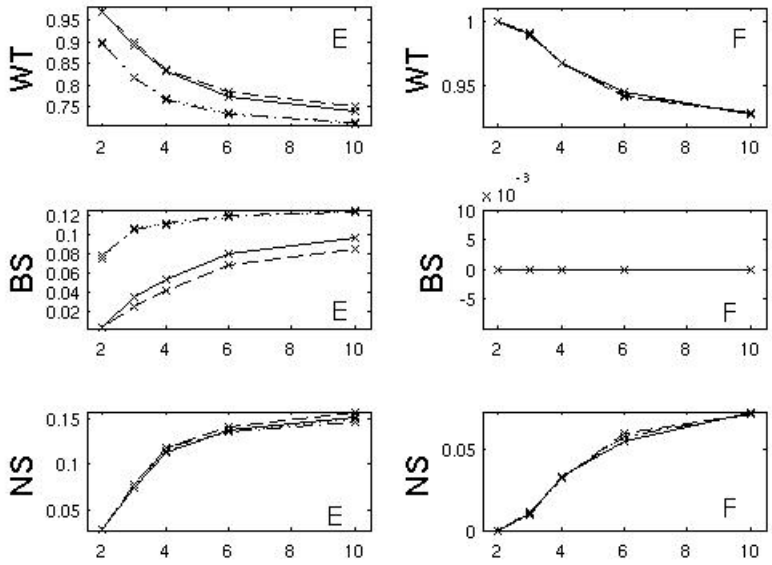


Figure 2: Rate of convergence to wild type (WT) and "mutant" patterns, broad striped (BS) or non-segmented (NS), after one round of cell division. The parasegment is now five cells wide, resulting from division of the first (solid line), second (dashed), third (dash-dotted), or fourth (dotted) cell. The $x$-axis lists the magnitude of the timescale perturbations, such that $\alpha_\ell \in [1/n, n]$. The left column (E) corresponds to cell division at an early time during stage 8 (initial conditions are as in Table 5), while results on the right column (F) correspond to cell division after the full pattern is established. For cases (E), the curves corresponding to division of third and fourth cells have very similar values, for both WT and BS, and thus overlap (dash-dotted and dotted). For cases (F), all curves are very similar.

Finally, to compare the robustness of four, five, six or eight cell wide parasegments, under similar conditions, we have analyzed the evolution of the system starting from a minimal pre-pattern:

$$\mathit{en}_1(0) = 1, \;\; \mathit{wg}_M(0) = 1, \;\; \mathrm{SLP}_{M-1,M}(0) = 1, \tag{6}$$

that is only *en* (first cell), *wg* (last cell), and SLP (last two cells) are expressed. Taking these initial conditions, the capacity of $M$ cell-wide parasegments to generate the segment polarity pattern was verified, again considering several magnitudes of timescale perturbations. Results are shown in Fig. 4. Interestingly, robustness to perturbations increases with the length of the parasegment, with a very significant increase from four to five cells. This indicates that a first round of cell division during stage 8 of embryonic development could be a very desirable event. Moreover, if the dividing cell is either the first or second in the parasegment (see the discussion of Fig. 2), cell division contributes to an increased robustness of the segment polarity network.

To further understand the mechanism responsible for higher robustness in wider parasegments, we explored the relative timescales for patched and cubitus interruptus proteins in the first cell expressing SLP, i.e., $\alpha_{\mathrm{PTC_{FS}}}$ and $\alpha_{\mathrm{CI_{FS}}}$. It was shown in [7] that $\alpha_{\mathrm{PTC_{FS}}} > \alpha_{\mathrm{CI_{FS}}}$ ($M = 4$, $_{\mathrm{FS}} = 3$) is a necessary condition to obtain 100% convergence to the wild type pattern (this result is now generalized in Section 6). Indeed, we verify that, among all $\{\alpha_X^j\}$ parameter combinations leading to the BS pattern (say $N_{\mathrm{BS}}$), more than $65\%$ break that condition, satisfying instead $\alpha_{\mathrm{PTC_{FS}}} < \alpha_{\mathrm{CI_{FS}}}$ (Fig. 5).

In addition, computing the average

$$q_{\mathrm{FS}} = \frac{1}{N_{\mathrm{BS}}} \sum_{k=1}^{N_{\mathrm{BS}}} \frac{\alpha^k_{\mathrm{CI_{FS}}}}{\alpha^k_{\mathrm{PTC_{FS}}}},$$

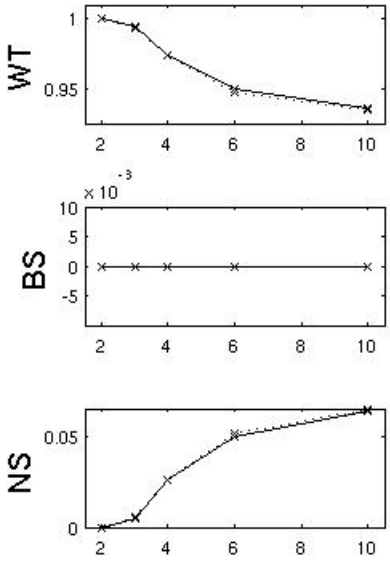


Figure 3: Rate of convergence to wild type (WT) and broad striped (BS) or non-segmented (NS) patterns, after two rounds of cell division. The parasegment is six cells wide, resulting from division of one of the five cells in the previous stage. The $x$-axis lists the magnitude of the timescale perturbations, such that $\alpha_\ell \in [1/n, n]$. The initial condition for the six cell model is WT (I), as in Table 6, with either of the five cells doubled (corresponding to cell division after full pattern is established). The results are very similar for all the five possible initial conditions, all curves overlapping. Shown here are the curves for doubling the first (solid) and fourth (dotted) cells.

we see that $q_{\text{FS}} > 1$ (Fig. 5), and is clearly lower for parasegments with four cells. In other words, a sharper difference between $\alpha_{\text{PTC}_{\text{FS}}}$ and $\alpha_{\text{CI}_{\text{FS}}}$ is needed to disrupt normal pattern development in wider parasegments (presumably a less likely situation).

# 6 Generating the wild type pattern with timescale separation

As in [7], we will analyze the behaviors of trajectories of systems of the form (1), assuming that trajectories are well-defined. Since the right-hand sides of equations of this type are discontinuous, it is very difficult to give general existence and uniqueness theorems for solutions of inital-value problems. One must impose additional assumptions, insuring that only a finite number of switches can take place on any finite time interval, and often tools from the theory of differential inclusions must be applied, see for instance [39] and [40] for more discussion.

For a segment of $M$ cells, SLP is typically expressed in the posterior part of the parasegment, in a group of adjacent cells (for instance in the last two or three cells of a five cell-wide parasegment). Then *wingless* is expressed in the last cell expressing SLP, while *engrailed* and *hedgehog* are expressed in the first cell of the parasegment. Thus we will adopt the notation $_{\text{FS}}$ (resp., $_{\text{LS}}$) to denote the first cell (resp., last cell) expressing SLP, and consider the following "generic" initial condition:

$$\text{SLP}_{\text{FS},\ldots,\text{LS}}(0) = 1;\ \ wg_{\text{LS}}(0) = 1;\ \ en_1(0) = 1;\ \ hh_1(0) = 1;\ \ ptc_{2,\ldots,\text{LS}}(0) = 1;\ \ ci_{2,\ldots,\text{LS}}(0) = 1, \qquad (7)$$

where, typically, one expects to see $_{\text{LS}} = M$ and $_{\text{FS}} > M/2$.

It was shown in [7] that a timescale separation assumption guaranteed convergence of the piecewise linear system (1) to the wild type pattern WT (Table 3), for a four cell-wide parasegment. We next state that

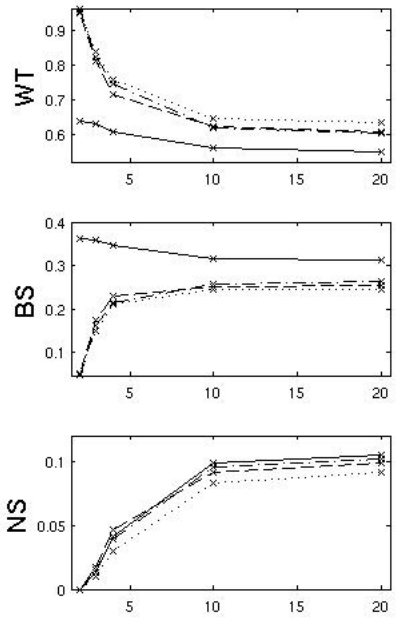


Figure 4: Rate of convergence to wild type (WT) and "mutant" patterns, broad striped (BS) or non-segmented (NS), starting from minimal patterning (6). The $x$-axis lists the magnitude of the timescale perturbations, such that $\alpha_\ell \in [1/n, n]$. Each of the four cases corresponds to four (solid), five (dashed), six (dash-dotted), or eight (dotted) cells-wide parasegments.

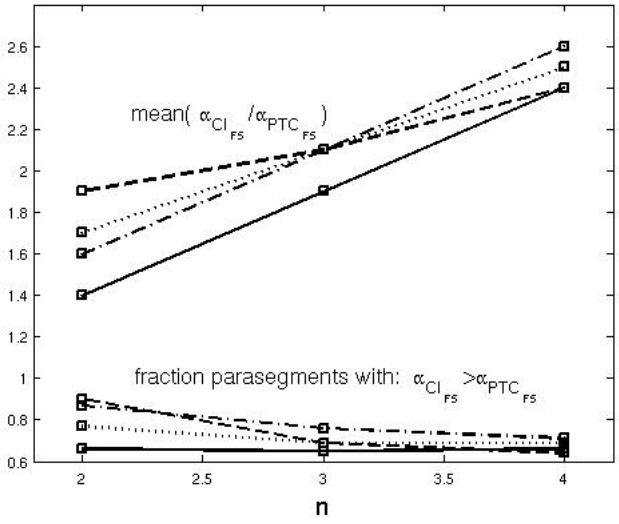


Figure 5: The average ratios $\alpha_{\text{CI}_{\text{FS}}}/\alpha_{\text{PTC}_{\text{FS}}}$, and percentage of (BS) parasegments satisfying $\alpha_{\text{CI}_{\text{FS}}} > \alpha_{\text{PTC}_{\text{FS}}}$, as a function of the magnitude of the timescale perturbations, $n$ (such that $\alpha_\ell \in [1/n, n]$). Curves correspond to four (solid), five (dashed), six (dash-dotted), or eight (dotted) cells-wide parasegments.

a similar result holds for wider parasegments. Let $A_{mRNA}$ and $A_{Prot}$ denote intervals for the scaling factors ($\alpha_\ell$) of the system's mRNAs or proteins (respectively). Assume that $A_{mRNA}$ and $A_{Prot}$ do not overlap, and satisfy:

(A1) For all $a \in A_{mRNA}$ and $b \in A_{Prot}$: $0 < 2a < b$.

Assume also that

(A2) $\theta_i = \theta$, for all $i$, and $\theta \leq 1/2$;

(A3) $\alpha_{\mathrm{PTC}_{\mathrm{FS}}} > \alpha_{\mathrm{CI}_{\mathrm{FS}}}$.

Such hypotheses seem quite reasonable from the biological point of view, as (A1) reflects the fact that post-translational modifications (such as protein conformational changes, which happen on a scale of thousandths of a second) are usually faster than transcription or translation (which happen on a scale of hundreds of seconds) (see, for instance, [12] and references therein). Hypothesis (A2) states that concentrations of protein (or mRNA) corresponding to less than 50% of its maximal value are sufficient to initiate an activation or inhibition by that protein. Hypothesis (A3) stems from observation of the numerical simulations (see discussion of ratio $q_{\mathrm{FS}} > 1$ and Fig. 5), but the following theoretical results show that indeed it guarantees convergence to the wild type steady state. It is thus a prediction of our model. Note that the next theorems are stated for any length $M \geq 4$ of the parasegment.

**Theorem 1.** Consider system (1) with initial condition (7). Suppose that assumptions (A1) to (A3) are satisfied. Then $wg_{\mathrm{FS}}(t) = 0$ for all $t$. ∎

This shows that the steady state representing the broad stripes pattern (BS I or II) cannot ever be reached in system (1) from the initial condition (7), under assumptions (A1) to (A3).

**Theorem 2.** Consider system (1) with initial condition (7). Suppose that assumptions (A1) to (A3) are satisfied. Then $wg_{\mathrm{LS}}(t) = 1$ and $\mathrm{PTC}_1(t) = 0$ for all $t$. ∎

This shows that the steady states represented by the no segmentation, or ectopic patterns (see Tables 3 and 6) cannot ever be reached in system (1) from the initial condition (7).

From Theorems 1 and 2 we conclude that, under the timescale separation assumption (and for appropriate $\theta$ values), the piecewise linear model (1) can never converge to its steady states corresponding to mutant patterns such as BS (I or II), NS or EC, for *any length of the parasegment*. Thus, since the only steady states reachable from initial condition (7) are WT, BS, NS, or EC (see Appendix) the wild type pattern can be expected when starting from the initial condition (7).

The proofs are very similar to those in [7]. One needs to consider first the modifications due to the dynamic rule for SLP, but then the arguments are analogous. For the four cell-wide parasegment model, the proofs are based on the initial conditions and Boolean rules for the third and fourth cells, that is the "first and last cells expressing SLP". The results can be easily adapted to deal more generally with SLP expressed in a group of adjacent cells, labeled FS to LS: indeed, to prove Theorem 1, follow the proof given in [7], but substituting "$node_3$" by "$node_{FS}$" and "$node_2$" by "$node_{FS-1}$"; to prove Theorem 2, simply replace "$node_4$" by "$node_{LS}$". For completeness, the proofs are included in the Supplementary material.

We will now analyze the effect of the SLP rule on the evolution of the network.

**Proposition 6.1.** Consider system (1) with initial condition (7). Then:

$$\mathrm{SLP}_i(t) = 0, \;\; wg_i(t) = \mathrm{WG}_i(t) = 0, \;\; \forall t > 0, \;\; i = 1, \ldots, \mathrm{FS} - 1, \tag{8}$$

and also

$$\begin{aligned} \mathrm{SLP}_i(t) = 1, \;\; en_i(t) = \mathrm{EN}_i(t) = 0, \;\; hh_i(t) = \mathrm{HH}_i(t) = 0, \\ ci_i(t) = \mathrm{CI}_i(t) = 1, \qquad \forall t > 0, \;\; i = \mathrm{FS}, \ldots, \mathrm{LS}. \end{aligned} \tag{9}$$

*Proof*: Since (at $t = 0$) $\text{SLP}_i(0) = wg_i(0) = \text{WG}_i(0) = 0$, for all $i = 1, \ldots, \text{FS} - 1$, to argue by contradiction, suppose that $t_1 > 0$ is the ***first*** instant for which $\text{SLP}_i(t_1) = 1$. First, observe that (since $wg_i$ starts at zero) $wg_i$ can only get activated after $\text{SLP}_i$ becomes 1. So, it holds that $wg_i(t) = 0$ for all $t \leq t_1$, for all $i = 1, \ldots, \text{FS} - 1$. For any $i = 1, \ldots, \text{FS} - 1$, note that $\text{SLP}_i(t_1) = 1$ implies (looking at the initial conditions and the SLP rule):

$$\text{there exists } \; t_a \in (0, t_1) : \;\; \text{WG}_i(t_a) = 1 \;\text{ and }\; \text{EN}_i(t_a) = 0.$$

And then $\text{WG}_i(t_a) = 1$ implies:

$$\text{there exists } \; t_b \in (0, t_a) : \;\; wg_i(t_b) = 1.$$

Now this implies (looking at the $wg_i$ rule):

$$\text{there exists } \; t_c \in (0, t_b) : \;\; wg_i(t_c) = 1 \;\text{ or }\; \text{SLP}_i(t_c) = 1.$$

But, finally, observe that this is impossible, a contradiction to our assumption, since both $\text{SLP}_i$ and $wg_i$ must be zero at $t_c < t_1$. This proves statement (8). The proof of (9) follows by direct application of the Boolean rules for SLP, *en*, EN, *hh*, and HH. ∎

Note that this result holds for every trajectory, even when none of the hypotheses (A1) to (A3) are satisfied. In fact, this Proposition guarantees that the anterior/posterior polarity in the parasegment is maintained throughout time, one of the first and crucial steps towards generating the wild type pattern. Indeed, note that the absence of SLP strictly prevents *wingless* expression, while the presence of SLP strictly prevents *engrailed* expression. It is known that parasegment boundaries are defined between a cell expressing *engrailed* and another expressing *wingless*. In adition, SLP is expressed in bands of 2 or more cells (for instance, $0 \cdots 011$ or $0 \cdots 0111$). Therefore, parasegment boundaries can only form on the first (FS) or last (LS) cells expressing sloppy paired. This result is interesting for several reasons. First, it indicates that not all cells of the parasegment are important, and incorrect expression in several of the cells can be corrected during development. Second, if given the information that two cells are essential, most Drosophila scientists would expect that those two key cells are the first and last cell of the parasegment, or the (first) cell expressing *engrailed* and the (last) cell expressing *wingless*. Only one of the two cells that we find coincides with this expectation: the last cell of sloppy paired, that is also the (last) cell expressing *wingless*. The first cell expressing sloppy paired does not coincide with *engrailed* expression, on the contrary, it determines the boundary of *engrailed* expression. This result suggests that the pattern of the segment polarity genes, and specifically the *wingless-engrailed* boundary, depends critically on the sloppy paired expression. As we can see from Tables 7 to 10 in the Appendix, wild type or mutant expression patterns evolve essentially around cells FS and LS. In wild type, sloppy paired protein counteracts the symmetry of the *en* $\rightarrow$ *wg* positive feedback loop to generate a single asymmetric *wingless - engrailed* boundary around LS. In mutants, an extra *wingless - engrailed* boundary may also form around FS.

# 7 Conclusion

Two questions were discussed in this paper, and possible answers proposed, leading to improvements in modeling and understanding the properties of Drosophila segment polarity network. The new dynamical rule for regulation of sloppy paired is composed of activation by wingless and inhibition by engrailed proteins, together with a maintenance term. Thus the network operates as an autonomous module with no "strong" external inputs. Starting from an initial expression pattern corresponding to development stage 7, the network can robustly generate and maintain an expression pattern which faithfully reproduces the segment polarity genes pattern at stages 9-10. Moreover, with the natural hypothesis of timescale separation, it is shown that the system cannot converge to any of the model's steady states corresponding to some

"mutant" patterns. This result further shows that generating the wild type segment polarity expression pattern depends essentially on the dynamics of the network at the first and last cells expressing sloppy paired protein. Rounds of cell division (and hence wider parasegments) will not disrupt the pattern: in fact, our results suggest that, as the number of cells in each parasegment increases, the system becomes more robust to perturbations in biological processes. One possible interpretation is that new cells in a parasegment strengthen the regulatory conditions definining the boundaries between parasegments, in such a way that those boundaries become much more difficult to break. The boundaries between parasegments are defined mainly by the relative positions of *en* and *wg*. According to the model, *en* can only be expressed in a cell immediately posterior or anterior to a cell expressing SLP. The results show that adding new cells between the first cell of the parasegment and the first cell expressing SLP leads to lower probability of formation of an extra *en* stripe just anterior to SLP. Hence a system with higher number of cells per parasegment is more likely to generate the desired wild type pattern.

## Acknowledgements

Part of this study was done while both authors were visiting the Kavli Institute for Theoretical Physics, Santa Barbara, USA during July 2007. The authors acknowledge fruitful discussions with Boris Shraiman. Boolean modeling in RA's group is funded by NSF grants MCB-0618402 and CCF-0643529 (CAREER), NIH grants 1R55AI065507 - 01A2 and 1 R01 GM083113-01, HFSP grant RGP20/2007 and USDA grant 2006-35100-17254.

## References

[1] B. Sanson. Generating patterns from fields of cells. examples from *Drosophila* segmentation. *EMBO Reports*, 21:1083–1088, 2001.

[2] L. Wolpert, R. Beddington, J. Brockes, T. Jessell, P. Lawrence, and E. Meyerowitz. *Principles of Development*. Current Biology Ltd., London, 1998.

[3] J. E. Hooper and M.P. Scott. The molecular genetic basis of positional information in insect segments. In W. Hennig, editor, *Early Embryonic Development of Animals*, pages 1–49. Springer, Berlin, 1992.

[4] G. von Dassow, E. Meir, E.M. Munro, and G.M. Odell. The segment polarity network is a robust developmental module. *Nature*, 406:188–192, 2000.

[5] R. Albert and H. G. Othmer. The topology of the regulatory interactions predicts the expression pattern of the *Drosophila* segment polarity genes. *J. Theor. Biol.*, 223:1–18, 2003.

[6] N.T. Ingolia. Topology and robustness in the Drosophila segment polarity network. *PLoS Biology*, 2:0805–0815, 2004.

[7] M. Chaves, E.D. Sontag, and R. Albert. Methods of robustness analysis for boolean models of gene control networks. *IEE Proc. Syst. Biol.*, 153:154–167, 2006.

[8] R. Thomas. Boolean formalization of genetic control circuits. *J. Theor. Biol.*, 42:563–585, 1973.

[9] L. Glass and S.A. Kauffman. The logical analysis of continuous, nonlinear biochemical control networks. *J. Theor. Biol.*, 39:103–129, 1973.

[10] R. Edwards and L. Glass. Combinatorial explosion in model gene networks. *Chaos*, 10:691–704, 2000.

[11] M. Chaves, R. Albert, and E.D. Sontag. Robustness and fragility of boolean models for genetic regulatory networks. *J. Theor. Biol.*, 235:431–449, 2005.

[12] J.A. Papin, T. Hunter, B.O. Palsson, and S. Subramaniam. Reconstruction of cellular signalling networks and analysis of their properties. *Nature Rev. Mol. Cell Biol.*, 6:99–111, 2005.

[13] C. Alexandre and J. P. Vincent. Requirements for transcriptional repression and activation by engrailed in *Drosophila* embryos. *Development*, 130:729–739, 2003.

[14] F. Gonzalez, L.S. Swales, A. Bejsovec, H. Skaer, and A. Martinez Arias. Secretion and movement of wingless protein in the epidermis of the *Drosophila* embryo. *Mech. Dev.*, 35:43–54, 1991.

[15] L. Glass. Classification of biological networks by their qualitative dynamics. *J. Theor. Biol.*, 54:85–107, 1975.

[16] J. E. Hooper and M.P. Scott. The drosophila *patched* gene encodes a putative membrane protein required for segmental patterning. *Cell*, 59:751–765, 1989.

[17] P. Aza-Blanc, F.A. Ramirez-Weber, M.-P. Leget, C. Schwartz, and T.B. Kornberg. Proteolysis that is inhibited by hedgehog targets Cubitus interruptus proteins to the nucleus and converts it to a repressor. *Cell*, 89:1043–1053, 1997.

[18] S. Pfeiffer and J.-P. Vincent. Signaling at a distance: Transport of Wingless in the embryonic epidermis of *drosophila*. *Cell & Dev. Biol.*, 10:303–309, 1999.

[19] K.M. Cadigan and R. Nusse. Wnt signaling: a common theme in animal development. *Genes & Dev.*, 11:3286–3305, 1997.

[20] T. Tabata, S. Eaton, and T.B. Kornberg. The *Drosophila hedgehog* gene is expressed specifically in posterior compartment cells and is a target of *engrailed* regulation. *Genes & Dev.*, 6:2635–2645, 1992.

[21] S. Eaton and T.B. Kornberg. Repression of ci-d in posterior compartments of drosophila by *engrailed*. *Genes & Dev.*, 4:1068–1077, 1990.

[22] A. Hidalgo and P. Ingham. Cell patterning in the *drosophila* segment: spatial regulation of the segment polarity gene *patched*. *Development*, 110:291–301, 1990.

[23] A.M. Taylor, Y. Nakano, J. Mohler, and P.W. Ingham. Contrasting distributions of patched and hedgehog proteins in the *drosophila* embryo. *Mechanisms of Development*, 42:89–96, 1993.

[24] P.W. Ingham and A. P. McMahon. Hedgehog signaling in animal development: paradigms and principles. *Genes & Dev.*, 15:3059–3087, 2001.

[25] M. van den Heuvel and P.W. Ingham. *smoothened* encodes a receptor-like serpentine protein required for *hedgehog* signaling. *Nature*, 382:547–551, 1996.

[26] P.W. Ingham. Transducing hedgehog: the story so far. *EMBO J.*, 17:3505–3511, 1998.

[27] P. Aza-Blanc and T. B. Kornberg. Ci a complex transducer of the hedgehog signal. *Trends in Genetics*, 15:458–462, 1999.

[28] J.T. Ohlmeyer and D. Kalderon. hedgehog stimulates maturation of cubitus interruptus into a labile transcriptional activator. *Nature*, 396:749–753, 1998.

[29] N. Méthot and K. Basler. Hedgehog controls limb development by regulating the activities of distinct transcriptional activator and repressor forms of cubitus interruptus. *Cell*, 96:819–831, 1999.

[30] A. Jacinto C. Alexandre and P. W. Ingham. Transcriptional activation of *hedgehog* target genes in *drosophila* is mediated directly by the Cubitus interruptus protein, a member of the GLI family of zinc finger DNA-binding proteins. *Genes & Dev.*, 10:2003–2013, 1996.

[31] T. von Ohlen and J.E. Hooper. Hedgehog signaling regulates transcription through Gli/Ci binding sites in the *wingless* enhancer. *Mechanisms of Development*, 68:149–156, 1997.

[32] D. Swantek and J. P. Gergen. Ftz modulates runt-dependent activation and repression of segment -polarity gene transcription. *Development*, 131:2281–2290, 2004.

[33] K.M. Bhat, E.H. van Beers, and P. Bhat. Sloppy paired acts as the downstream target of Wingless in the *drosophila* CNS and interaction between *sloppy paired* and gooseberry inhibits *sloppy paired* during neurogenesis. *Development*, 127:655–665, 2000.

[34] H.-H. Lee and M. Frasch. Wingless effects mesoderm patterning and ectoderm segmentation events via induction of its downstream target *sloppy paired*. *Development*, 127:5497–5508, 2000.

[35] G. von Dassow and G.M. Odell. Design and constraints of the *drosophila* segment polarity modude: robust spatial patterning emerges from intertwined cell state switches. *J. Exp. Zool. (Mol. Dev. Evol.)*, 294:179–215, 2002.

[36] K.M. Cadigan, U. Grossniklaus, and W.J. Gehring. Localized expression of sloppy paired protein maintains the polarity of *Drosophila* parasegments. *Genes & Dev.*, 8:899–913, 1994.

[37] A. Martinez Arias, N. Baker, and P.W. Ingham. Role of segment polarity genes in the definitiona and maintenance of cell states in the *Drosophila* embryo. *Development*, 103:157–170, 1988.

[38] P.W. Ingham, A.M. Taylor, and Y. Nakano. Role of *Drosophila patched* gene in positional signaling. *Nature*, 353:184–187, 1991.

[39] T. Gedeon. Attractors in continuous-time switching networks. *Communications on Pure and Applied Analysis*, 2:187–209, 2003.

[40] R. Casey, H. de Jong, and J.L. Gouzé. Piecewise linear models of genetic regulatory networks: equilibria and their stability. *J. Math. Biol.*, 52:27–56, 2006.

[41] A. Gallet, C. Angelats, S. Kerridge, and P.P. Thérond. Cubitus interruptus-independent transduction of the hedgehog signal in *Drosophila*. *Development*, 127:5509–5522, 2000.

[42] C. Schwartz, J. Locke, C. Nishida, and T.B. Kornberg. Analysis of *cubitus interruptus* regulation in *drosophila* embryos and imaginal disks. *Development*, 121:1625–1635, 1995.

[43] S. DiNardo, E. Sher, J. Heemskerk-Jongens, J.A. Kassis, and P.H. O'Farrell. Two-tiered regulation of spatially patterned *engrailed* gene expression during *drosophila* embryogenesis. *Nature*, 332:45–53, 1988.

[44] A. Bejsovec and E. Wieschaus. Segment polarity gene interactions modulate epidermal patterning in *drosophila* embryos. *Development*, 119:501–517, 1993.

# Appendix

## A Steady states of the mathematical model

The steady states of the Boolean model depicted in Table 1 can be computed by solving the equations: $X = F(X)$. It was shown that these steady states are also those of the piecewise continuous model (1).

In our analysis we focused on initial conditions of the form (7), with $_{FS} = M - 1$ or $_{FS} = M - 2$. Initial conditions of this form imply certain restrictions on the dynamics of both the Boolean and piecewise continuous model, namely that expression of SLP will not vary throughout time. This is established in Proposition 6.1, and follows roughly from the fact that, in each cell $i$, SLP$_i$ can only become expressed if *wg*$_i$ has been expressed at some previous instant but, conversely, *wg*$_i$ can only become expressed if SLP$_i$ was expressed at some previous time. Thus, if both SLP$_i$ and *wg*$_i$ are absent at time 0, neither can become expressed at later times.

Thus we compute only the steady states that can possibly be reached if the system starts from initial condition (7) (recall that different asynchronous updating strategies may lead the system from the same initial condition to different steady states). These are given in Tables 7, 8 (case $_{FS} = M-1$), and Tables 9, 10 (case $_{FS} = M - 2$). All steady states can be derived from the expression of *wingless* in the first and last cells expressing SLP. Thus there are four possible distinct steady states, associated with the four distinct combinations of the wingless values in the cells $_{FS}$ and $_{LS}$, respectively, $W_a$ and $W_b$.

The steady state corresponding to the case $(W_a, W_b) = (0, 1)$ represents the wild type (WT) expression pattern of the segment polarity genes. There are two (very similar) variants for this pattern, differing only on the value of PTC$_1$, which can be either $1 - 0 * 1 = 1$ or $1 - 1 * 1 = 0$.

The steady state corresponding to the case $(W_a, W_b) = (1, 1)$ represents the broad striped (BS) expression pattern, observed in *ptc* mutants [20]. These mutants express *wingless* in a stripe which is broader than that observed in wild type. For the case $_{FS} = M - 1$, the BS pattern expresses *wg* in two consecutive cells. For the case $_{FS} = M - 2$, *wg* is expressed in the first and last cells expressing SLP, but not in the middle cell. For this reason, we believe that initial conditions with SLP $= 0 \cdots 011$ represent the biological system more faithfully.

The steady state corresponding to the case $(W_a, W_b) = (0, 0)$ represents the non-segmented (NS) expression pattern, observed in *en* mutants [20]. Here, patched and cubitus repressor proteins are ubiquitously expressed, and neither wingless nor engrailed are expressed.

Finally, the steady state corresponding to the case $(W_a, W_b) = (1, 0)$ represents an ectopic (EC) expression pattern which, to our knowledge, has not been observed experimentally. The relative positions of *en* and *wg* are inverted, that is a segment could be formed, but with the boundary defined by a cell expressing engrailed to the left of a cell expressing wingless. In previous studies [11, 7], only a very small fraction of perturbations ($< 1\%$) to the mathematical model lead the system to this ectopic steady state. Here again there are two possible variants, differing only on the value of PTC$_{FS-1}$, which can be either $1 - 0 * 1 = 1$ or $1 - 1 * 1 = 0$.

Some of these patterns are illustrated in Fig. 6 (this figure can also be found in [11]). The wild type, broad striped and no-segmentation mathematical patterns are shown for a four-cell wide parasegment, and compared to the corresponding wild type or mutant embryos.

Table 7: Steady states for initial conditions of the form (7), FS $= M-1$ for $M = 4, 5$. The patterns for $M = 5$ are represented in the five consecutive cells labeled 1,2,3,$M$-1,$M$, while the patterns for $M = 4$ are represented in the four consecutive cells $M$-1,$M$,1,2. The symbols $W_a$, $W_b$, and $d$ represent digits in $\{0, 1\}$. $W_a$ and $W_b$ denote *wingless* expression in the FS and LS $= M$ cells, respectively. The functions $f$ and $g$ are defined as $f(W_a, W_b, d) = 1 - (d + (1 - d)W_a)W_b$ and $g(W_a, W_b) = (1 - W_a)(1 - W_b)$.

| Node \ cell | 1 | 2 | 3 | $M-1$ | $M$ | 1 | 2 |
|---|---|---|---|---|---|---|---|
| SLP | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| WG $=$ *wg* | 0 | 0 | 0 | $W_a$ | $W_b$ | 0 | 0 |
| EN $=$ *en* | $W_b$ | 0 | $W_a$ | 0 | 0 | $W_b$ | $W_a$ |
| HH $=$ *hh* | $W_b$ | 0 | $W_a$ | 0 | 0 | $W_b$ | $W_a$ |
| *ptc* | 0 | 1-$g(W_a, W_b)$ | 0 | $W_a$ | $W_b$ | (1-$W_b$)$W_a$ | (1-$W_a$)$W_b$ |
| PTC | 1-$dW_b$ | 1 | 1-$dW_a$ | 1 | 1 | $f(W_a, W_b, d)$ | $f(W_b, W_a, d)$ |
| CI $=$ *ci* | 1-$W_b$ | 1 | 1-$W_a$ | 1 | 1 | 1-$W_b$ | 1-$W_a$ |
| CIA | 0 | 1-$g(W_a, W_b)$ | 0 | $W_a$ | $W_b$ | (1-$W_b$)$W_a$ | (1-$W_a$)$W_b$ |
| CIR | 1-$W_b$ | $g(W_a, W_b)$ | 1-$W_a$ | 1-$W_a$ | 1-$W_b$ | $g(W_a, W_b)$ | $g(W_a, W_b)$ |

Table 8: Steady states for initial conditions of the form (7), FS $= M-1$ for $M \geq 6$ (the patterns for $M = 6$ are obtained from the columns 1,2,$M$-3,$M$-2,$M$-1,$M$). The symbols $W_a$, $W_b$, and $d$ represent digits in $\{0, 1\}$. $W_a$ and $W_b$ denote *wingless* expression in the FS and LS $= M$ cells, respectively.

| Node \ cell | 1 | 2 | 3 | $\cdots$ | M-5 | M-4 | M-3 | M-2 | M-1 | M |
|---|---|---|---|---|---|---|---|---|---|---|
| SLP | 0 | 0 | 0 | $\cdots$ | 0 | 0 | 0 | 0 | 1 | 1 |
| WG $=$ *wg* | 0 | 0 | 0 | $\cdots$ | 0 | 0 | 0 | 0 | $W_a$ | $W_b$ |
| EN $=$ *en* | $W_b$ | 0 | 0 | $\cdots$ | 0 | 0 | 0 | $W_a$ | 0 | 0 |
| HH $=$ *hh* | $W_b$ | 0 | 0 | $\cdots$ | 0 | 0 | 0 | $W_a$ | 0 | 0 |
| *ptc* | 0 | $W_b$ | 0 | $\cdots$ | 0 | 0 | $W_a$ | 0 | $W_a$ | $W_b$ |
| PTC | 1-d$W_b$ | 1 | 1 | $\cdots$ | 1 | 1 | 1 | 1-d$W_a$ | 1 | 1 |
| CI $=$ *ci* | 1-$W_b$ | 1 | 1 | $\cdots$ | 1 | 1 | 1 | 1-$W_a$ | 1 | 1 |
| CIA | 0 | $W_b$ | 0 | $\cdots$ | 0 | 0 | $W_a$ | 0 | $W_a$ | $W_b$ |
| CIR | 1-$W_b$ | 1-$W_b$ | 1 | $\cdots$ | 1 | 1 | 1-$W_a$ | 1-$W_a$ | 1-$W_a$ | 1-$W_b$ |

Table 9: Steady states for initial conditions of the form (7), FS $= M-2$ for $M = 5, 6$. The patterns for $M = 6$ are represented in the six consecutive cells labeled 1,2,3,$M$-2,$M$-1,$M$, while the patterns for $M = 5$ are represented in the five consecutive cells $M$-2,$M$-1,$M$,1,2. The symbols $W_a$, $W_b$, and $d$ represent digits in $\{0, 1\}$. $W_a$ and $W_b$ denote *wingless* expression in the FS and LS $= M$ cells, respectively. The functions $f$ and $g$ are defined as $f(W_a, W_b, d) = 1 - (d + (1 - d)W_a)W_b$ and $g(W_a, W_b) = (1 - W_a)(1 - W_b)$.

| Node \ cell | 1 | 2 | 3 | M-2 | M-1 | M | 1 | 2 |
|---|---|---|---|---|---|---|---|---|
| SLP | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| WG $=$ *wg* | 0 | 0 | 0 | $W_a$ | 0 | $W_b$ | 0 | 0 |
| EN $=$ *en* | $W_b$ | 0 | $W_a$ | 0 | 0 | 0 | $W_b$ | $W_a$ |
| HH $=$ *hh* | $W_b$ | 0 | $W_a$ | 0 | 0 | 0 | $W_b$ | $W_a$ |
| *ptc* | 0 | 1-$g(W_a, W_b)$ | 0 | $W_a$ | 0 | $W_b$ | (1-$W_b$)$W_a$ | (1-$W_a$)$W_b$ |
| PTC | 1-$dW_b$ | 1 | 1-$dW_a$ | 1 | 1 | 1 | $f(W_a, W_b, d)$ | $f(W_a, W_b, d)$ |
| CI $=$ *ci* | 1-$W_b$ | 1 | 1-$W_a$ | 1 | 1 | 1 | 1-$W_b$ | 1-$W_a$ |
| CIA | 0 | 1-$g(W_a, W_b)$ | 0 | $W_a$ | 0 | $W_b$ | (1-$W_b$)$W_a$ | (1-$W_a$)$W_b$ |
| CIR | 1-$W_b$ | $g(W_a, W_b)$ | 1-$W_a$ | 1-$W_a$ | 1 | 1-$W_b$ | $g(W_a, W_b)$ | $g(W_a, W_b)$ |

Table 10: Steady states for initial conditions of the form (7), FS $= M-2$ for $M \geq 7$ (the patterns for $M = 7$ are obtained from the columns 1,2,$M$-4,$M$-3,$M$-2,$M$-1,$M$). The symbols $W_a$, $W_b$, and $d$ represent digits in $\{0, 1\}$. $W_a$ and $W_b$ denote *wingless* expression in the FS and LS $= M$ cells, respectively.

| Node \ cell | 1 | 2 | 3 | $\cdots$ | M-5 | M-4 | M-3 | M-2 | M-1 | M |
|---|---|---|---|---|---|---|---|---|---|---|
| SLP | 0 | 0 | 0 | $\cdots$ | 0 | 0 | 0 | 1 | 1 | 1 |
| WG = *wg* | 0 | 0 | 0 | $\cdots$ | 0 | 0 | 0 | $W_a$ | 0 | $W_b$ |
| EN = *en* | $W_b$ | 0 | 0 | $\cdots$ | 0 | 0 | $W_a$ | 0 | 0 | 0 |
| HH = *hh* | $W_b$ | 0 | 0 | $\cdots$ | 0 | 0 | $W_a$ | 0 | 0 | 0 |
| *ptc* | 0 | $W_b$ | 0 | $\cdots$ | 0 | $W_a$ | 0 | $W_a$ | 0 | $W_b$ |
| PTC | 1-d$W_b$ | 1 | 1 | $\cdots$ | 1 | 1 | 1-d$W_a$ | 1 | 1 | 1 |
| CI = *ci* | 1-$W_b$ | 1 | 1 | $\cdots$ | 1 | 1 | 1-$W_a$ | 1 | 1 | 1 |
| CIA | 0 | $W_b$ | 0 | $\cdots$ | 0 | $W_a$ | 0 | $W_a$ | 0 | $W_b$ |
| CIR | 1-$W_b$ | 1-$W_b$ | 1 | $\cdots$ | 1 | 1-$W_a$ | 1-$W_a$ | 1-$W_a$ | 1 | 1-$W_b$ |

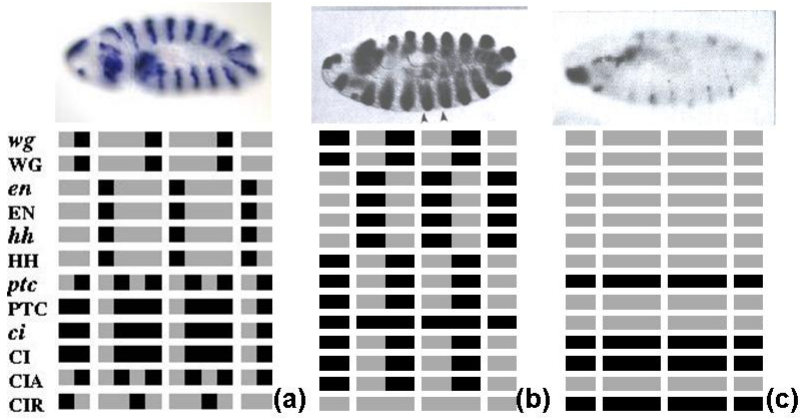


Figure 6: a) Top: Illustration of the gene expression pattern of *wingless* on a gastrulating (stage 9) embryo. The parasegmental furrows form at the posterior border of the *wg*-expressing cells [2]. Bottom: Wild type expression pattern of the Boolean model. Left corresponds to anterior and right to posterior in each parasegment. Horizontal rows correspond to the pattern of individual nodes - specified at the left side of the row - over two full and two partial parasegments. Each parasegment is assumed to be four cells wide. A black (gray) box denotes a node that is (is not) expressed. b) Top: *wingless* expression pattern in a *patched* knock-out mutant embryo at stage 11 [20]. The *wingless* stripes broaden, and secondary furrows appear at the middle of the parasegment, indicating a new *en*-*wg* boundary. Bottom: Broad striped steady state of the Boolean model. This state is obtained when *wg*, *en* or *hh* are initiated in every cell. A variant of this state is obtained when *patched* is kept off; the difference is in the fact that *ptc* and PTC are not expressed in the mutant state. This steady state agrees with all experimental observations on *ptc* mutants and heat-shocked genes [20, 41, 37, 42, 43, 38, 44]. c) Top: *wingless* expression pattern in an *engrailed* knock-out mutant embryo at stage 11 [20]. The initial periodic pattern is disappearing, and gives rise to a non-segmented, embryonic lethal phenotype. Bottom: Non-segmented steady state of the Boolean model. This steady state agrees with all experimental observations on *wg, en, hh* mutants [20, 43, 42, 22, 41]. Gene expression images obtained from http://www.fruitfly.org (a) and [20] (b,c).

# Supplementary material
## Studying the effect of cell division on expression patterns of the segment polarity genes

M. Chaves and R. Albert

## Proof of Theorems 1 and 2

As mentioned in the main text, the proofs are similar to those given in [7], with appropriate index changes. For easier reference and completeness, we provide the proofs next.

First, some observations regarding the solutions of system (1). Let $X$ denote any of the nodes in the network, and $\alpha$ its time rate. Since equations (1) are either of the form $d\hat{X}/dt = \alpha(-\hat{X}+1)$ or $d\hat{X}/dt = -\alpha\hat{X}$, their solutions are continous functions, piecewise combinations of:

$$\hat{X}^1(t) = 1-(1-\hat{X}^1(t_0))\,e^{-\alpha(t-t_0)} \quad (1)$$

$$\hat{X}^0(t) = \hat{X}^0(t_0)\,e^{-\alpha(t-t_0)} \quad (2)$$

$\hat{X}^1(t)$ (resp. $\hat{X}^0(t)$) is monotonically increasing (resp. decreasing). In addition, note that discrete variables $X$ can only switch between 0 and 1 at those instants when $\hat{X}(t_{\text{switch}}) = \theta$, that is:

$$t^1_{\text{switch}} = t_0 + \frac{1}{\alpha}\,\ln\,\frac{(1-\hat{X}(t_0))}{1-\theta} \quad (3)$$

$$t^0_{\text{switch}} = t_0 + \frac{1}{\alpha}\,\ln\,\frac{\hat{X}(t_0)}{\theta} \quad (4)$$

From Proposition 6.1 we can immediately conclude:

$$\widehat{wg}_{1,\ldots,\text{FS}-1}(t) = \widehat{\text{WG}}_{1,2}(t) = 0, \quad (5)$$

$$\widehat{en}_{\text{FS},\ldots,\text{LS}}(t) = \widehat{\text{EN}}_{\text{FS},\ldots,\text{LS}}(t) = 0,$$

$$\widehat{hh}_{\text{FS},\ldots,\text{LS}}(t) = \widehat{\text{HH}}_{\text{FS},\ldots,\text{LS}}(t) = 0, \quad (6)$$

$$\widehat{ci}_{\text{FS},\ldots,\text{LS}}(t) = 1 \text{ and } \widehat{\text{CI}}_{\text{FS},\ldots,\text{LS}}(t) = 1 - e^{-\alpha_{\text{CI}_{\text{FS},\ldots,\text{LS}}}\,t}. \quad (7)$$

**Lemma A.1.** Let $0 \le t_0 < t_3 \le t_1$ and $0 \le t_2 < t_3$. Define $\delta = \ln\frac{1}{1-\theta}/\max_{1,\ldots,N}\alpha_i$. Assume $\text{CIA}_{\text{FS}}(t) = 0$ for $t \in (t_2, t_3)$, and $wg_{\text{FS}}(t) = 0$ for $t \in [0, t_3)$. Then

(a) $wg_{\text{FS}}(t) = 0$ for $t \in [0, t_3+\delta)$;

(b) $\text{WG}_{\text{FS}}(t) = 0$ for $t \in [0, t_3+\delta)$;

(c) $en_{\text{FS}-1}(t) = \text{EN}_{\text{FS}-1}(t) = 0$ for $t \in [0, t_3+\delta)$;

(d) $hh_{\text{FS}-1}(t) = \text{HH}_{\text{FS}-1}(t) = 0$ for $t \in [0, t_3+\delta)$.

Assume further that $\text{PTC}_{\text{FS}}(t) = 1$ for $t \in (t_0, t_1)$. Then

(e) $\text{PTC}_{\text{FS}}(t) = 1$ for all $t \in (t_0, t_3+\delta)$.

(f) $\text{CIA}_{\text{FS}}(t) = 0$ for all $t \in (t_2, t_3+\delta)$.

*Proof:* Part (a) follows directly from the fact that $F_{wg_{\text{FS}}}(t) = 0$ on $[0, t_3)$, and from (3).

To prove parts (b), (c), and (d), first note that initial conditions together with $wg_{\text{FS}}(t) = 0$ for $t \in [0, t_3)$ imply

$$\widehat{\text{WG}}_{\text{FS}}(t) = 0,\ \widehat{en}_{\text{FS}-1}(t) = \widehat{\text{EN}}_{\text{FS}-1}(t) = 0,\ \widehat{hh}_{\text{FS}-1}(t) = \widehat{\text{HH}}_{\text{FS}-1}(t) = 0,$$

for $t \in [0, t_3]$. Then, from equations (1) to (4) we conclude that the corresponding discrete variables cannot switch from 0 to 1 during an interval of the form $[0, t_3 + \frac{1}{\alpha_j} \ln \frac{1}{1-\theta})$. Taking the largest common interval yields the desired results.

To prove parts (e) and (f), assume also that $\text{PTC}_{\text{FS}}(t) = 1$ for $t \in (t_0, t_1)$. From (6) and part (d), it follows that function $F_{\text{PTC}_{\text{FS}}}$ does not switch in the interval $(t_0, t_3 + \delta)$ and in fact $\text{PTC}_{\text{FS}}(t) = 1$ for all $t$ in this interval. This, together with (6) and part (d) yield $F_{\text{CIA}_{\text{FS}}}(t) = 0$ for $(t_0, t_3 + \delta)$, so that $\widehat{\text{CIA}}_{\text{FS}}$ cannot increase in this interval and the discrete level satisfies $\text{CIA}_{\text{FS}}(t) = 0$ for all $t \in (t_2, t_3 + \delta)$, as we wanted to show. ∎

**Corollary A.2.** Let $0 \le t_0 < t_3 \le t_1$ and $0 \le t_2 < t_3$. If $\text{PTC}_{\text{FS}}(t) = 1$ for $t \in (t_0, t_1)$, $\text{CIA}_{\text{FS}}(t) = 0$ for $t \in (t_2, t_3)$, and $wg_{\text{FS}}(t) = 0$ for $t \in [0, t_3)$, then $wg_{\text{FS}}(t) = 0$ for all $t$.

*Proof:* Applying Lemma A.1 we conclude that, given any $k \ge 0$:

$$\begin{aligned}
&\text{CIA}_{\text{FS}}(t) = 0, \ \text{ for } \ t \in (t_2, t_3 + k\delta) \\
&wg_{\text{FS}}(t) = 0, \ \text{ for } \ t \in [0, t_3 + k\delta) \\
&\text{PTC}_{\text{FS}}(t) = 1 \ \text{ for } \ t \in (t_0, t_3 + k\delta)
\end{aligned}$$

imply

$$\begin{aligned}
&\text{CIA}_{\text{FS}}(t) = 0, \ \text{ for } \ t \in (t_2, t_3 + (k+1)\delta) \\
&wg_{\text{FS}}(t) = 0, \ \text{ for } \ t \in [0, t_3 + (k+1)\delta) \\
&\text{PTC}_{\text{FS}}(t) = 1 \ \text{ for } \ t \in (t_0, t_3 + (k+1)\delta).
\end{aligned}$$

Since $\delta$ is finite, we conclude by induction on $k$ that $wg_{\text{FS}}(t) = 0$ for all $t$. ∎

*Proof of Theorem 1:* The rule for $\text{CIA}_{\text{FS}}$ may be simplified to (by (6))

$$F_{\text{CIA}_{\text{FS}}} = \text{CI}_{\text{FS}} \text{ and } [\text{notPTC}_{\text{FS}} \text{ or } hh_{\text{FS}-1} \text{ or } \text{HH}_{\text{FS}-1}].$$

From equation (7), we have that

$$\text{CI}_{\text{FS}}(t) = 1, \ \text{ for all } t > \frac{1}{\alpha_{\text{CI}_{\text{FS}}}} \ln \frac{1}{1-\theta}. \tag{8}$$

On the other hand, since $ptc_{\text{FS}}(0) = 1$, by continuity of solutions $ptc_{\text{FS}}(t) = 1$ for all $t < \frac{1}{\alpha_{ptc_{\text{FS}}}} \ln \frac{1}{\theta}$. This implies that the Patched protein satisfies

$$\widehat{\text{PTC}}_{\text{FS}}(t) = 1 - e^{-\alpha_{\text{PTC}_{\text{FS}}} t}, \ \ 0 \le t \le \frac{1}{\alpha_{ptc_{\text{FS}}}} \ln \frac{1}{\theta}$$

and therefore

$$\text{PTC}_{\text{FS}}(t) = \begin{cases} 0, & 0 \le t \le \frac{1}{\alpha_{\text{PTC}_{\text{FS}}}} \ln \frac{1}{1-\theta} \\ 1, & \frac{1}{\alpha_{\text{PTC}_{\text{FS}}}} \ln \frac{1}{1-\theta} < t < \frac{1}{\alpha_{ptc_{\text{FS}}}} \ln \frac{1}{\theta}. \end{cases} \tag{9}$$

By assumption, $\alpha_{\text{PTC}_{\text{FS}}} > \alpha_{ptc_{\text{FS}}}$ and also $\ln\frac{1}{1-\theta} \leq \ln\frac{1}{\theta}$, defining a nonempty interval where $\text{PTC}_{\text{FS}}$ is expressed. Now let $t_c = \frac{1}{\alpha_{\text{CI}_{\text{FS}}}} \ln\frac{1}{1-\theta}$ and $t_p = \frac{1}{\alpha_{\text{PTC}_{\text{FS}}}} \ln\frac{1}{1-\theta}$. $\widehat{\text{CIA}}_{\text{FS}}(t)$ starts at zero and must remain so while $\text{CI}_{\text{FS}} = 0$, so that

$$\text{CIA}_{\text{FS}}(t) = 0 \ \text{ for } 0 < t < t_c.$$

In the case $t_c > t_p$, letting $t_0 = t_p$, $t_1 = \frac{1}{\alpha_{ptc_{\text{FS}}}} \ln\frac{1}{\theta}$, $t_2 = 0$, and $t_3 = t_c$ in Corollary A.2, obtains $wg_{\text{FS}}(t) = 0$ for all $t$. This proves item (b) of the theorem, and part of (a).

To finish the proof of item (a), we assume that $(1-\theta)^2 < \theta$ and must now consider the case $t_c \leq t_p$. Then

$$\widehat{\text{CIA}}_{\text{FS}}(t) = \begin{cases} 0, & 0 \leq t \leq t_c \\ 1 - e^{-\alpha_{\text{CIA}_{\text{FS}}}(t-t_c)}, & t_c < t \leq t_p \\ \widehat{\text{CIA}}_{\text{FS}}(t_p)\, e^{-\alpha_{\text{CIA}_{\text{FS}}}(t-t_p)}, & t_p < t \leq \frac{1}{\alpha_{ptc_{\text{FS}}}} \ln\frac{1}{\theta}, \end{cases}$$

Following equation (3) with $t_0 = t_c$ and $\widehat{\text{CIA}}_{\text{FS}}(t_0) = 0$, $\text{CIA}_{\text{FS}}$ might become expressed at time $t_c < t_a < t_p$:

$$t_a = t_c + \frac{1}{\alpha_{\text{CIA}_{\text{FS}}}} \ln\frac{1}{1-\theta},$$

but it would then become zero again at (equation (4) with $t_0 = t_p$)

$$t_b = t_p + \frac{1}{\alpha_{\text{CIA}_{\text{FS}}}} \ln\, \frac{\widehat{\text{CIA}}_{\text{FS}}(t_p)}{\theta}.$$

Finally, we show that, even if $\text{CIA}_{\text{FS}}(t) = 1$ for $t \in (t_a, t_b)$, $wg_{\text{FS}}$ cannot become expressed in this interval. In this interval, $\widehat{wg}_{\text{FS}}$ evolves according to $\widehat{wg}_{\text{FS}}(t) = 1 - e^{-\alpha_{wg_{\text{FS}}}(t-t_a)}$, and $wg_{\text{FS}}$ can switch to 1 at time

$$t_w = t_a + \frac{1}{\alpha_{wg_{\text{FS}}}} \ln\frac{1}{1-\theta}.$$

We will show that $t_w > t_b$, so $wg_{\text{FS}}(t) = 0$ in the interval $[0, t_b)$. Writing

$$\ln\, \frac{\widehat{\text{CIA}}_{\text{FS}}(t_p)}{\theta} = \ln\, \frac{\widehat{\text{CIA}}_{\text{FS}}(t_p)}{1-\theta} \frac{1-\theta}{\theta} = \ln\, \frac{\widehat{\text{CIA}}_{\text{FS}}(t_p)}{1-\theta} + \ln\, \frac{1-\theta}{\theta} \leq \ln\frac{1}{1-\theta} + \ln\frac{1}{1-\theta}$$

where we have used $\widehat{\text{CIA}}_{\text{FS}}(t_p) \leq 1$ and the assumption on $\theta$: $\frac{1-\theta}{\theta} \leq \frac{1}{1-\theta}$. Therefore

$$t_b \leq t_p + \frac{2}{\alpha_{\text{CIA}_{\text{FS}}}} \ln\frac{1}{1-\theta} < \frac{1}{\alpha_{wg_{\text{FS}}}} \ln\frac{1}{1-\theta} + \frac{1}{\alpha_{\text{CIA}_{\text{FS}}}} \ln\frac{1}{1-\theta} < t_w$$

where we have used the timescale separation assumption (A1). Letting $t_0 = t_p$, $t_2 = 0$, and $t_1 = t_3 = \min\{t_b, \alpha^{-1}_{ptc_{\text{FS}}} \ln\frac{1}{\theta}\}$ in the Corollary, obtains $wg_{\text{FS}}(t) = 0$ for all $t$. ∎

We will next show that if $wg_{\text{LS}}(t) = 1$ in a given interval $[0, T)$, then in fact $wg_{\text{LS}}(t)$ remains expressed for a longer time, up to $T + \delta$, with $\delta > 0$. This is mainly due to assumption (A1), which says that mRNAs take longer than proteins to update their discrete values, because they have longer half-lives: $\alpha^{-1}_{mRNA} > \alpha^{-1}_{Prot}$. This allows the initial signal "$wg_{\text{LS}} = 1$" to travel down the network, sequencially affecting the wingless protein, *engrailed*, *hedgehog* and CIA, and feed back into *wingless* allowing $wg_{\text{LS}}$ to remain expressed for a further time interval.

**Lemma A.3.** Let $T \geq \frac{1}{\alpha_{wg_{\text{LS}}}} \ln\frac{1}{\theta}$ and define

$$\delta = \frac{1}{\alpha_{\text{WG}_{\text{LS}}}} \ln\, \frac{(1 - e^{-\frac{\alpha_{\text{WG}_{\text{LS}}}}{\alpha_{wg_{\text{LS}}}} \ln\frac{1}{\theta}})}{\theta}. \tag{10}$$

If $wg_{\text{LS}}(t) = 1$ for $0 \leq t < T$, then

(a) $\mathrm{WG_{LS}}(t) = 1$ for $t \in (\frac{1}{\alpha_{\mathrm{WG_{LS}}}} \ln \frac{1}{1-\theta}, T + \delta)$;

(b) $en_1(t) = 1$ for $t \in [0, T + \delta)$;

(c) $\widehat{\mathrm{EN}}_1(t) = 1 - e^{-\alpha_{\mathrm{EN}_1} t}$ for $t \in [0, T + \delta)$, and $\mathrm{EN}_1(t) = 1$ for $(\frac{1}{\alpha_{\mathrm{EN}_1}} \ln \frac{1}{1-\theta}, T + \delta)$;

(d) $ci_1(t) = 0$, $\mathrm{CI}_1(t) = 0$, $\mathrm{CIA}_1(t) = 0$, and $\mathrm{CIR}_1(t) = 0$ for $t \in [0, T + \delta)$;

(e) $hh_1(t) = 1$, for $t \in [0, T + \delta)$;

(f) $\mathrm{CIA_{LS}}(t) = 1$, for $t \in (\frac{1}{\alpha_{\mathrm{CI_{LS}}}} \ln \frac{1}{1-\theta} + \frac{1}{\alpha_{\mathrm{CIA_{LS}}}} \ln \frac{1}{1-\theta}, T + \delta)$, and $\mathrm{CIR_{LS}}(t) = 0$, for $t \in [0, T + \delta)$;

(g) $wg_{\mathrm{LS}}(t) = 1$ for $t \in [0, T + \delta)$.

*Proof:* Let $T \geq \frac{1}{\alpha_{wg_{\mathrm{LS}}}} \ln \frac{1}{\theta}$, and assume that $wg_{\mathrm{LS}}(t) = 1$ for $0 \leq t < T$. To prove part (a), note that $\widehat{\mathrm{WG}}_{\mathrm{LS}}(t)$ is of the form (1) (with $t_0 = 0$, and $\widehat{\mathrm{WG}}_{\mathrm{LS}}(0) = 0$) and the corresponding discrete variable is $\mathrm{WG_{LS}}(t) = 1$, for $t \in (\frac{1}{\alpha_{\mathrm{WG_{LS}}}} \ln \frac{1}{1-\theta}, T)$. Moreover, suppose that $wg_{\mathrm{LS}}(t) = 0$ for $t > T$, then

$$\widehat{\mathrm{WG}}_{\mathrm{LS}}(t) = (1 - e^{-\alpha_{\mathrm{WG_{LS}}} T}) e^{-\alpha_{\mathrm{WG_{LS}}}(t-T)}, \quad t > T.$$

But $\mathrm{WG_{LS}}$ remains 1 until the switching threshold is attained, that is up to time

$$T + \frac{1}{\alpha_{\mathrm{WG_{LS}}}} \ln \frac{(1 - e^{-\alpha_{\mathrm{WG_{LS}}} T})}{\theta} \geq T + \frac{1}{\alpha_{\mathrm{WG_{LS}}}} \ln \frac{(1 - e^{-\alpha_{\mathrm{WG_{LS}}} \frac{1}{\alpha_{wg_{\mathrm{LS}}}} \ln \frac{1}{\theta}})}{\theta} \equiv T + \delta.$$

Thus we conclude that $\mathrm{WG_{LS}}(t) = 1$ in the desired interval.

To prove part (b), observe that $F_{en_1}(t) = \mathrm{WG_{LS}}(t)$ for all $t$, from (5), and recall that $en_1(0) = 1$. From part (a), $F_{en_1}(t) = 1$ for $t \in (\frac{1}{\alpha_{\mathrm{WG_{LS}}}} \ln \frac{1}{1-\theta}, T + \delta)$. On the other hand, $en_1$ can only switch from 1 to 0 at $t = \alpha_{en_1}^{-1} \ln \frac{1}{\theta}$ which is larger than $\alpha_{\mathrm{WG_{LS}}}^{-1} \ln \frac{1}{1-\theta}$. So, in fact, $en_1(t) = 1$ for all $0 \leq t < T + \delta$.

Part (c) follows immediately by integration of the $\widehat{\mathrm{EN}}_1$ equation.

To prove part (d), first recall $F_{ci_1} = \text{not } \mathrm{EN}_1$ and the initial conditions $ci_1(0) = 0 = \mathrm{CI}_1(0) = \mathrm{CIA}_1(0) = \mathrm{CIR}_1(0)$. Therefore $\widehat{ci}_1(t)$ increases up to $t = \frac{1}{\alpha_{\mathrm{EN}_1}} \ln \frac{1}{1-\theta}$ and then decreases in $\alpha_{\mathrm{EN}_1}^{-1} \ln \frac{1}{1-\theta} < t < T + \delta$. Now note that the discrete variable $ci_1(t)$ remains 0 in the whole interval $[0, T + \delta)$. This is because $\widehat{ci}_1$ never reaches the $\theta$ threshold: this would be attained at some $t \geq \alpha_{ci_1}^{-1} \ln \frac{1}{1-\theta}$ but, since $\alpha_{ci_1}^{-1} \ln \frac{1}{1-\theta} > \alpha_{\mathrm{EN}_1}^{-1} \ln \frac{1}{1-\theta}$, the function $\widehat{ci}_1$ starts decreasing before it could reach the value $\theta$. Finally, from the rules of the Cubitus proteins it is immediate to see that $\mathrm{CI}_1(t) = \mathrm{CIA}_1(t) = \mathrm{CIR}_1(t) = 0$ for $t \in [0, T + \delta)$.

To prove part (e), recall that $F_{hh_1} = \mathrm{EN}_1$ and not $\mathrm{CIR}_1$. From part (a), it follows that $F_{hh_1}(t) = 0$ in the interval $[0, \alpha_{\mathrm{EN}_1}^{-1} \ln \frac{1}{1-\theta})$ and $F_{hh_1}(t) = 1$ in the interval $(\alpha_{\mathrm{EN}_1}^{-1} \ln \frac{1}{1-\theta}, T + \delta)$. Since $hh_1(0) = 1$, $\widehat{hh}_1(t)$ decreases in the interval $[0, \alpha_{\mathrm{EN}_1}^{-1} \ln \frac{1}{1-\theta})$ but increases in $(\alpha_{\mathrm{EN}_1}^{-1} \ln \frac{1}{1-\theta}, T + \delta)$. The discrete value is $hh_1(t) = 1$ in the whole interval, since $\widehat{hh}_1(t)$ remains above the $\theta$ threshold. (The justification is similar to the case of $ci_1(t)$ in part (d).)

To prove part (f), note that part (e) and then the use of (7), allows us to simplify $F_{\mathrm{CIA_{LS}}}$:

$$F_{\mathrm{CIA_{LS}}}(t) = \mathrm{CI_{LS}}(t) \text{ and } hh_1(t) = 1, \quad t \in \left( \frac{1}{\alpha_{\mathrm{CI_{LS}}}} \ln \frac{1}{1-\theta}, T + \delta \right).$$

Thus

$$\widehat{\mathrm{CIA}}_{\mathrm{LS}}(t) = \begin{cases} 0, & 0 \leq t \leq \frac{1}{\alpha_{\mathrm{CI_{LS}}}} \ln \frac{1}{1-\theta} \\ 1 - e^{-\alpha_{\mathrm{CIA_{LS}}} \left( t - \frac{1}{\alpha_{\mathrm{CI_{LS}}}} \ln \frac{1}{1-\theta} \right)}, & \frac{1}{\alpha_{\mathrm{CI_{LS}}}} \ln \frac{1}{1-\theta} < t \leq T + \delta, \end{cases}$$

and $\text{CIA}_{\text{LS}}(t) = 1$ for $t \in [\frac{1}{\alpha_{\text{CI}_{\text{LS}}}} \ln \frac{1}{1-\theta} + \frac{1}{\alpha_{\text{CIA}_{\text{LS}}}} \ln \frac{1}{1-\theta}, T+\delta)$. Observe that this interval is indeed nonempty, by assumption (A1). Finally, $F_{\text{CIR}_{\text{LS}}}(t) = \text{CI}_{\text{LS}}(t)$ and not$hh_1(t) = 0$, and hence $\text{CIR}_{\text{LS}}(t) = 0$ for $t \in [0, T+\delta)$.

To prove part (g), we note that (from part (f))

$$F_{wg_{\text{LS}}}(t) = 1, \quad t \in \left( \frac{1}{\alpha_{\text{CI}_{\text{LS}}}} \ln \frac{1}{1-\theta} + \frac{1}{\alpha_{\text{CIA}_{\text{LS}}}} \ln \frac{1}{1-\theta}, T+\delta \right),$$

implying that $\widehat{wg}_{\text{LS}}(t)$ increases in this interval. On the other hand, we know that $\widehat{wg}_{\text{LS}}(t) \geq \theta$ and $wg_{\text{LS}}(t) = 1$ up to at least $t = \frac{1}{\alpha_{wg_{\text{LS}}}} \ln \frac{1}{\theta} > \frac{1}{\alpha_{\text{CI}_{\text{LS}}}} \ln \frac{1}{1-\theta} + \frac{1}{\alpha_{\text{CIA}_{\text{LS}}}} \ln \frac{1}{1-\theta}$. This shows that in fact $wg_{\text{LS}}(t) = 1$ for all $t \in [0, T+\delta)$. ∎

*Proof of Theorem 2:* Since $wg_{\text{LS}}(0) = 1$, from equations (3), (4), we know that the earliest possible switching time from 1 to 0 is $\alpha_{wg_{\text{LS}}}^{-1} \ln \frac{1}{\theta}$. Applying Lemma A.3 with $T = \alpha_{wg_{\text{LS}}}^{-1} \ln \frac{1}{\theta}$ establishes that $wg_{\text{LS}}(t) = 1$ for $t \in [0, T+\delta)$, with $\delta$ given by (10). Next, applying Lemma A.3 with $T = \alpha_{wg_{\text{LS}}}^{-1} \ln \frac{1}{\theta} + k\delta$, $k \in \mathbb{N}$, shows that $wg_{\text{LS}}(t) = 1$ for $t \in [0, T + (k+1)\delta)$. Since $\delta$ is finite, we can conclude by induction that $wg_{\text{LS}}(t) = 1$ for all $t \geq 0$.

To prove that $\text{PTC}_1(t) \equiv 0$, note that $\text{CIA}_1(t) \equiv 0$ (Lemma A.3, with $T = +\infty$) implies $ptc_1(t) \equiv 0$. Since $\text{PTC}_1(0) = 0$ and $\text{PTC}_1$ cannot become expressed unless $ptc_1$ is first expressed, the desired result follows. ∎