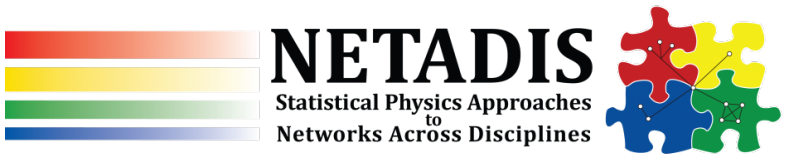


Inference of dynamical Gene Regulatory Networks and comparative genome analysis

Silvia Grigolon

Supervisors: Prof. *Silvio Franz* & Prof. *Olivier Martin*

LPTMS - Laboratoire de Physique Théorique et Modèles Statistique

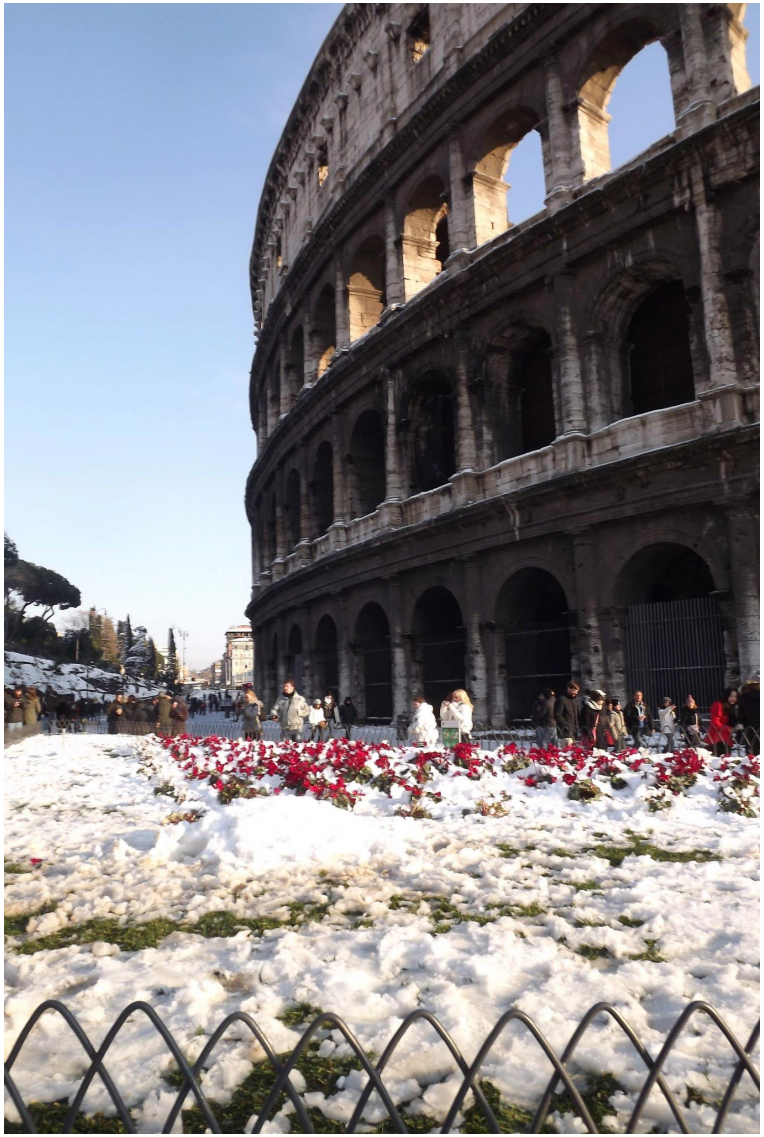
—Netadis Kick - Off Meeting—

Turin, February 3rd – 6th, 2013

Outline

1. Where I come from...
   - My Origins
   - Past Scientific Background & Academic Interests
   - Brief Excursion into my Master's Thesis
2. What I do...
   - Current Scientific Background
   - Current work
3. What I am going to do...

My Origins



Past Scientific Background & Academic Interests

✓ November 2009, Laurea Triennale in Physics (BS) at *Sapienza, Universitá di Roma*;

✓ March 2012, Laurea Magistrale in Physics (MS) at *Sapienza, Universitá di Roma*:

- specialization in Theoretical Physics – Statistical Mechanics;
- Master's Degree Thesis on:

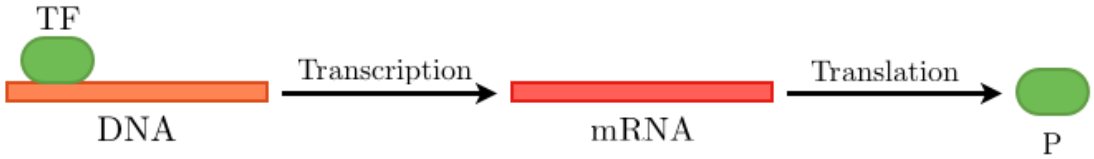
MicroRNA – based Networks and Circuits: a Noise Buffering Analysis,

under the supervision of Prof. *Enzo Marinari* and *Francesca Di Patti*, Ph.D.

I am interested in:

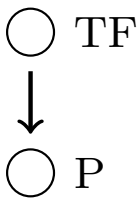
* Systems Biology;
* Statistical Mechanics (Classical and Quantum) and Disordered Systems;
* Computational Physics;
* Mathematical Physics and Dynamical Systems (especially what concerns chaotic systems and turbulence).

The role of Noise in Regulatory Processes

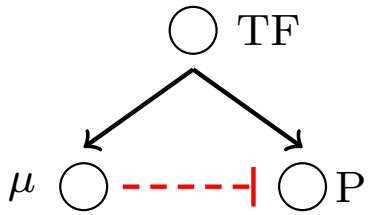


- **GRN**→ set of genes (nodes of the network) interacting with each other through their produced proteins.
- **What is microRNA?** Small RNA molecule which inhibits genic expression.
- **What did we analyze?** **3 genetic circuits** involving microRNA to understand if this inhibitor behaves as a **genetic fine tuner**.

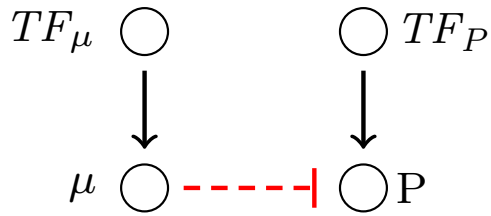
TF



FFL



Open



The degree of noise buffering depends on the structure of the circuit

- **How did we study the problem?**

Theoretically

van Kampen's Linear Noise Expansion

$$n_i \equiv \mathcal{N}\phi_i + \mathcal{N}^{\frac{1}{2}}\xi_i,$$

where:
* n_i is the number of molecules of the $i-$species;
* ϕ_i is the density of molecules in absence of fluctuations;
* ξ_i is a random noise;
* $\mathcal{N}$ is the size of the system.

This permits to avoid all the three points correlations resulting from the second degree reaction involving microRNA.

Numerically

Gillespie's Algorithm

a *self-consistent algorithm on the time variable*:

it generates by itself the temporal step given as input all the reaction rates.

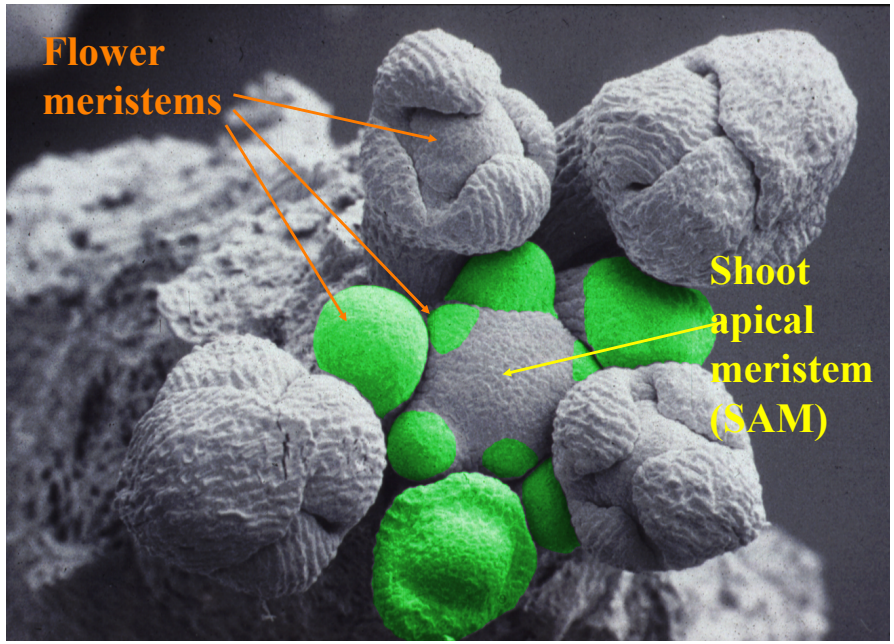
In the end, we found:

- a way to develop a **good theoretical framework** in agreement with numerical results;
- microRNA actually reduces protein fluctuations acting as a **fine tuner**.

S.Grigolon, F. Di Patti, A. De Martino, E. Marinari, *to be published.*

What is the genetic program for Floral Morphogenesis?

Arabidopsis Thaliana → **Model Organism in plant biology and genetics**



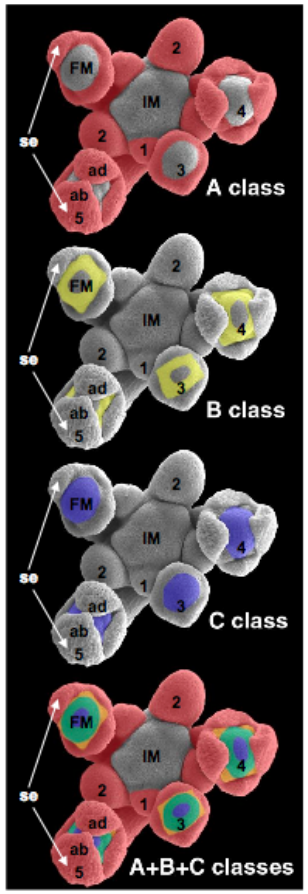
The overall conservation of flower structure suggests the existence and the persistence through evolution of **robust** **GRN modules** controlling the basic features of flower development.

Flower Development: the ABC model (Cohen & Meyerowitz, 1991)

4 Main and Highly Specified Structures in the Flower Meristem also known as **Floral Organs**:

- **Sepals**;
- **Petals**;
- **Stamens**;
- **Carpels**.

ABC Model:
The genes involved in flower development are divided into **3 classes** named **A**, **B**, **C** that give rise to organs through these **FOS–GRN Interactions**:



$$
\begin{aligned}
A &\rightarrow \text{Sepals},\\
A + B &\rightarrow \text{Petals},\\
B + C &\rightarrow \text{Stamens},\\
C &\rightarrow \text{Carpels}.
\end{aligned}
$$

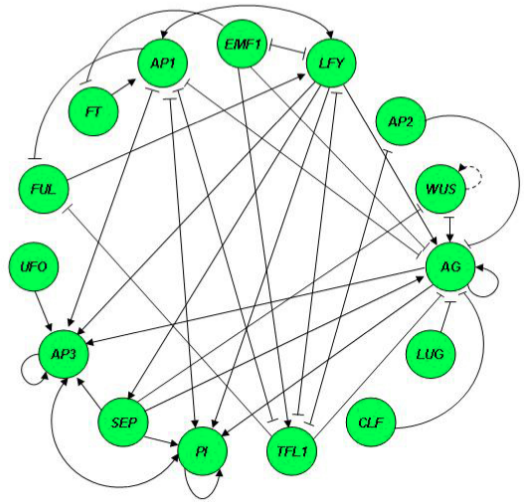
A first approach: a Boolean GRN (1)

E. R. Álvarez-Buylla et al., **PLoS One**, 3, 11, 2008

GRN extending the ABC model. It is aimed to obtain the gene expression profile of the steady states of the network.

$$\underline{x}(t) \equiv \text{state of the GRN, } \mathcal{N} - \text{components vector,}$$

where $\mathcal{N}(= 15)$ is the number of genes in the GRN.



Stochastic Time-Evolution:

$$\forall n = \{1, ..., \mathcal{N}\}, \frac{dx_n(t)}{dt} = \begin{cases} \alpha \cdot [F_n(\hat{x}_{n_1}, ..., \hat{x}_{n_k}) - x_n(t)], & \eta \\ \alpha \cdot [1 - F_n(\hat{x}_{n_1}, ..., \hat{x}_{n_k}) - x_n(t)], & 1 - \eta \end{cases},$$

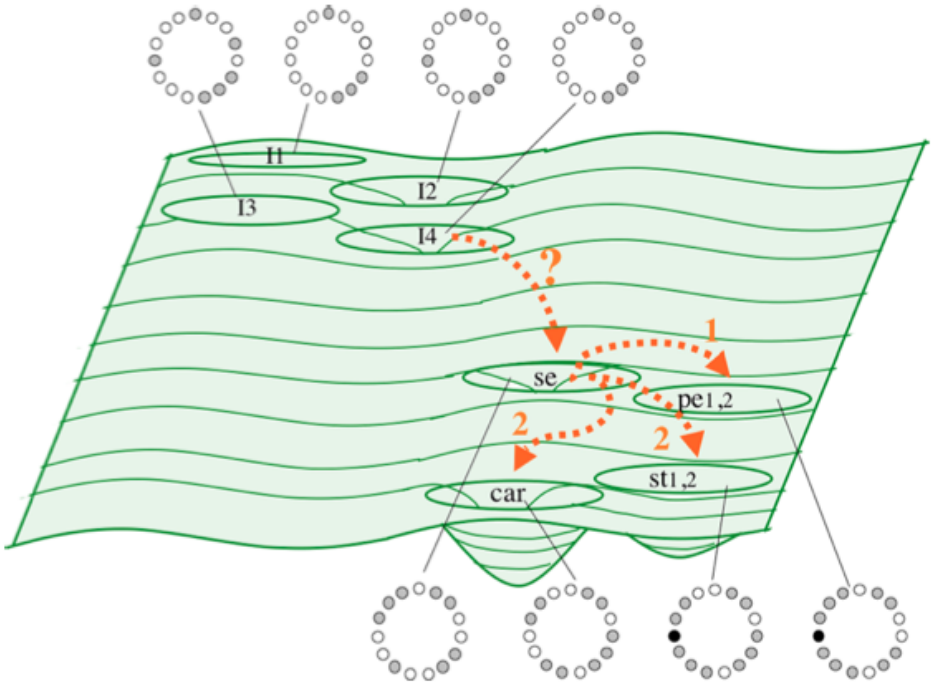
where:

- $F_n(\hat{x}_{n_1}, ..., \hat{x}_{n_k})$ is a boolean variable-function guessed from microarray experiments;
- $\hat{x}_n$ is a boolean variable connected to x_n (continuous variable) through the relation $\hat{x}_n = H(x_n - \theta_n)$;
- α is a constant connected to the relaxation time of the gene expression profile, τ ;
- η is a probability, defined to introduce the stochasticity in the model.

A first approach: a Boolean GRN (2)

This system of 15 differential equations has **10 attractors**:

6 corresponding to the gene expression profiles of sepals, petals (1,2), stamens(1,2) and carpels;
4 corresponding to the gene expression profiles of inflorescence configurations.



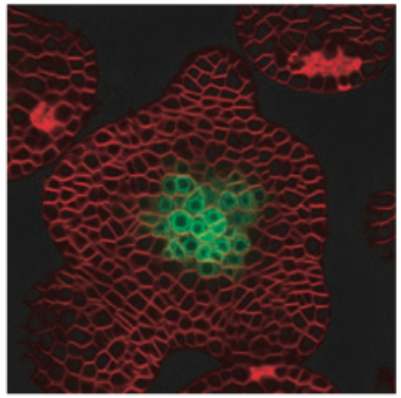
They obtained in particular the probability transitions among these attractors, caused by the presence of the noise η.

Why another model?

The previous model does not treat space and so there is no *spatio–temporal dynamics* → unable to explain the highly regular phyllotactic patterns?

Very recent biological studies:

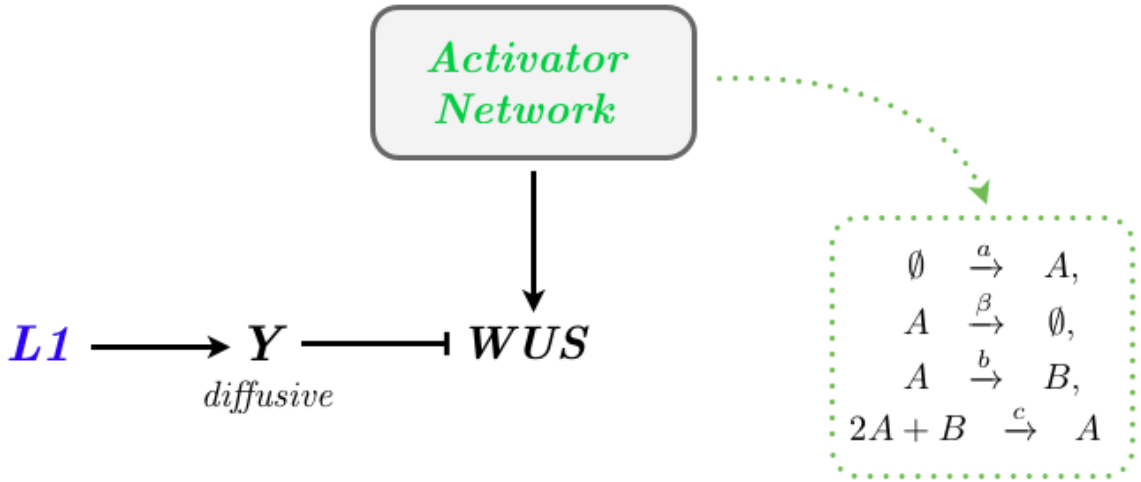
∗ F. Besnard, Ph.D. Thesis under the supervision of Prof. T. Vernoux, École Normale Supérieure de Lyon, 2011



- in floral morphogenesis a key compound regulating developmental processes: the **hormone auxin**, trigger of the initiation of floral primordia, depleted around them once they have been formed (*local inhibitory fields*);
- (*synergetic local inhibitory fields*) between Auxin and other genetic expression products (e.g., *WUS*);
- the time between two consecutive organ initiations is **NOT FIXED**;
- the observed patterns are **ROBUST** under environmental noise;
- co-initiations of organs could not be caused by a noise on the size of the involved fields but it depends on the noise of activations' thresholds.

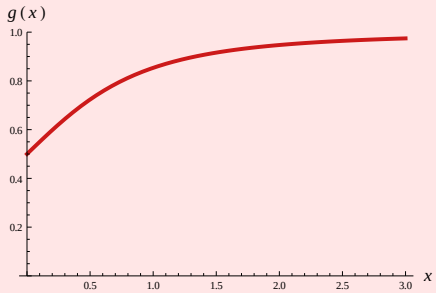
A second approach: a reaction–diffusion system

H. Jönsson et al., **Bioinformatics**, 21:1, i232, 2005



Activation Function:

$$g(x) = \frac{1}{2}\left(1 + \frac{x}{\sqrt{1+x^2}}\right)$$



L1 Layer:

$$L_1(x,y) = \begin{cases} 1, & (x,y) \in \partial\mathcal{L} \\ 0, & \text{otherwise} \end{cases}$$

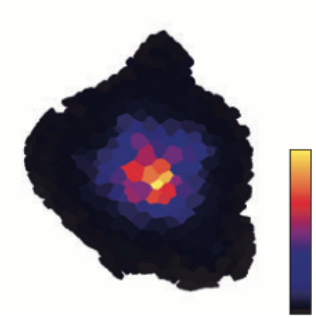
A and **B** → Brusselator molecules!

Boundary Conditions:

$$\nabla \cdot C_i = 0 \text{ on } \partial\mathcal{L}$$

A second approach: a reaction–diffusion system (2)

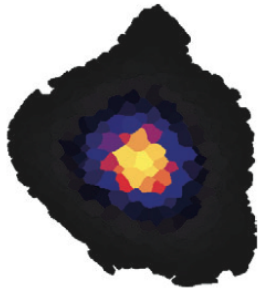
Experimental Results



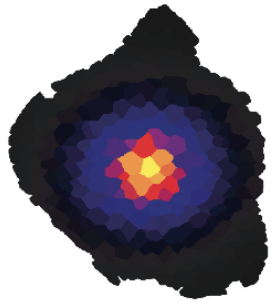
WUSCHEL relative concentration

Numerical Results

5th Order Runge–Kutta Method with Adaptive Step Size



WUSCHEL concentration



A concentration

L1 Layer

Main Problem:
The simulated system in not robust to small parameters changes!

A simple diffusive model with thresholds

Main Assumption: Floral Morphogenesis is driven by **Auxin** (A) and **Wuschel** (W).

Basic Features:

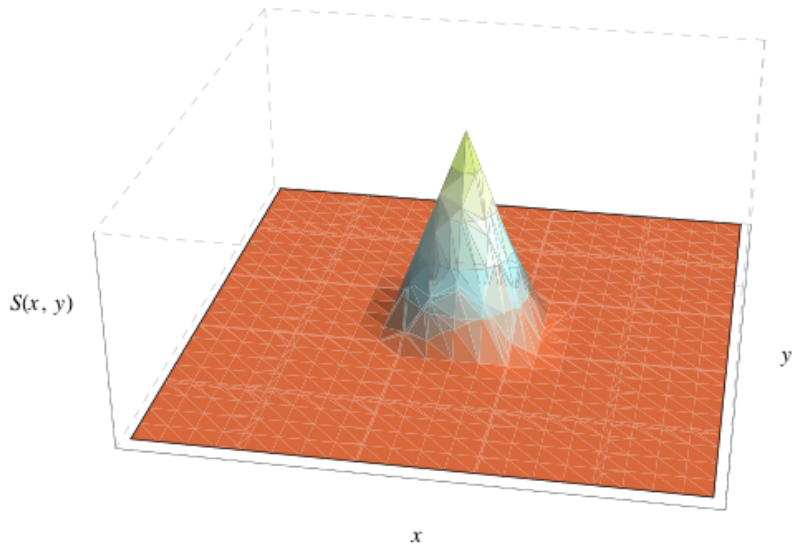
* the phenomenon takes place in a 2D square lattice $\mathcal{L}$ where the origin is fixed in the center of the lattice;
* Auxin and Wuschel can diffuse $\rightarrow D_A$, D_W;
* they can be degraded $\rightarrow \lambda_A$, λ_W;
* they are produced through a source term with polar symmetry:

$$S(r) = \mathcal{A}_i \left(1 - \frac{r}{r_i}\right) \mathcal{H}(r_i - r),$$

$$i = A, W,$$

where:

- $\mathcal{A}_i$ is the maximum concentration that can be produced;
- r_i is the *decay length* of the source;
- $\mathcal{H}(r_i - r)$ is the Heaviside function.



A simple diffusive model with thresholds (2)

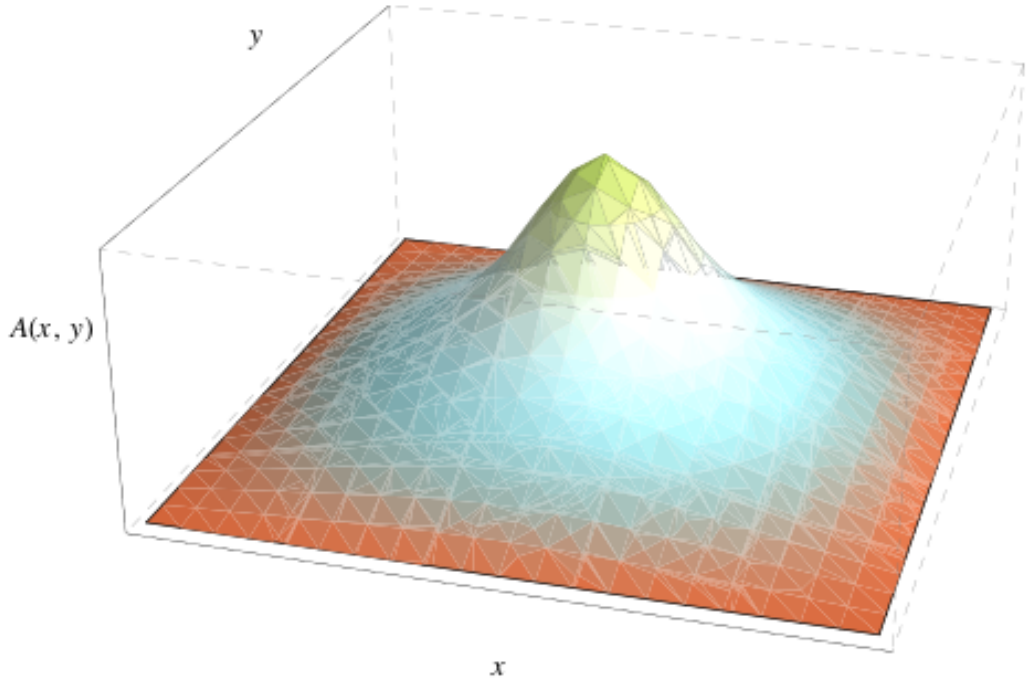
The dynamics

$$\begin{cases} \frac{\partial C_A}{\partial t} = S(r)\big|_{i=A} + D_A \nabla^2 C_A - \lambda_A C_A \\ \frac{\partial C_W}{\partial t} = S(r)\big|_{i=W} + D_W \nabla^2 C_W - \lambda_W C_W \end{cases}$$

with as BCs: $C_i(\partial\mathcal{L}) = 0, \forall i = A, W.$

How to integrate these equations?

Numerically, with a Runge–Kutta V order algorithm with adaptive step size.



A simple diffusive model with thresholds (2)

How to reproduce Floral Organ Specification?

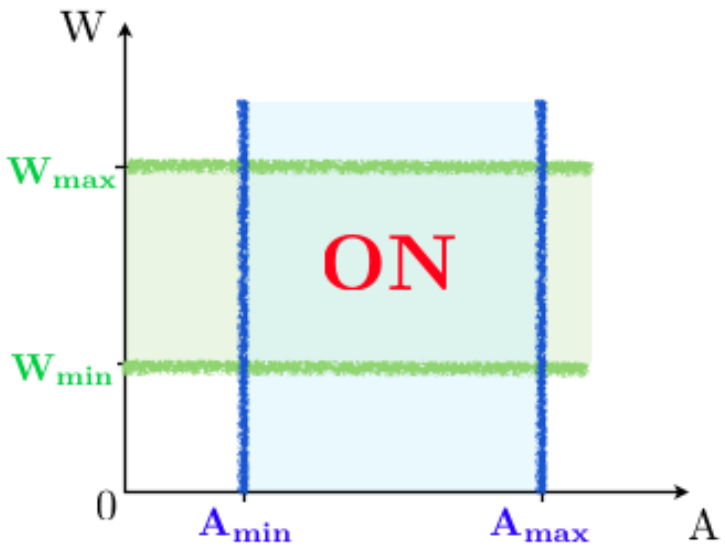
Lattice $\mathcal{L}$ as a set of cells: every cell can be **differentiated** OR **undifferentiated**.

differentiated $\equiv$ gene state changes!

* **How can we achieve this?**

According to Auxin and Wuschel concentration in cells!

If this condition is satisfied, hence $A > A_{min}$ AND $W < W_{max}$, the cell becomes **differentiated** and so on...



* **And the inhibitory fields?**

To reproduce this, it must be remembered that Auxin's transport in cells is given by **DIFFUSION + ACTIVE TRANSPORT**.

$\rightarrow$ the more Auxin there is, the more Auxin arrives! + Poissonian Noise on Flux

Secondments & Further Projects

About the model...

- development of sepal formation and then focus on the other organs;
- study the stochastic version of the system;
- search for adjusted parameters in the model via MCMC.

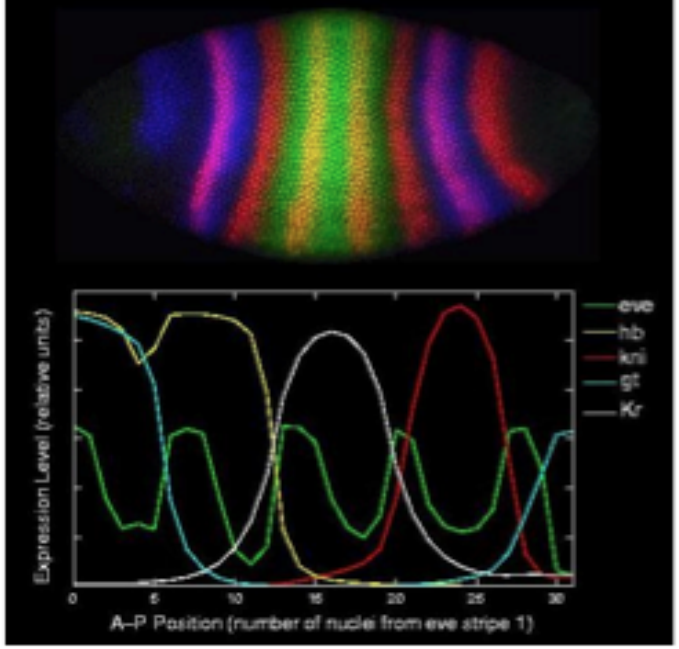
Secondments...

- *King's College of London* (UK);
- *Technische Universitaet Berlin*, Berlin, Germany.

Other Projects...

- inference on phylogenetic trees for *Influenza A* virus with Prof. Silvio Franz.

$$\frac{dv_i^a}{dt} = R_a g_a \Big(\sum_{b=1}^{N} T^{ab} v_i^b + m^a v_i^{bcd} + h^a\Big) + D(n)[(v_{i-1}^a - v_i^a) + (v_{i+1}^a - v_i^a)] - \lambda_a v_i^a$$



Thank you
for your attention!