Stochastic self-regulated gene expression model

Romain Yvinec [a] Michael C. Mackey [b] and Changjing Zhuge [c]

a) Institut Camille Jordan, Universite de Lyon 1, France

&

b) McGill University Montreal, QC CANADA

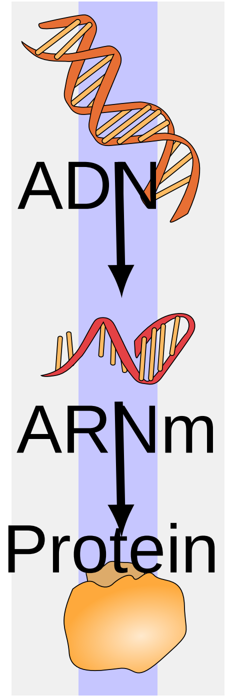
&

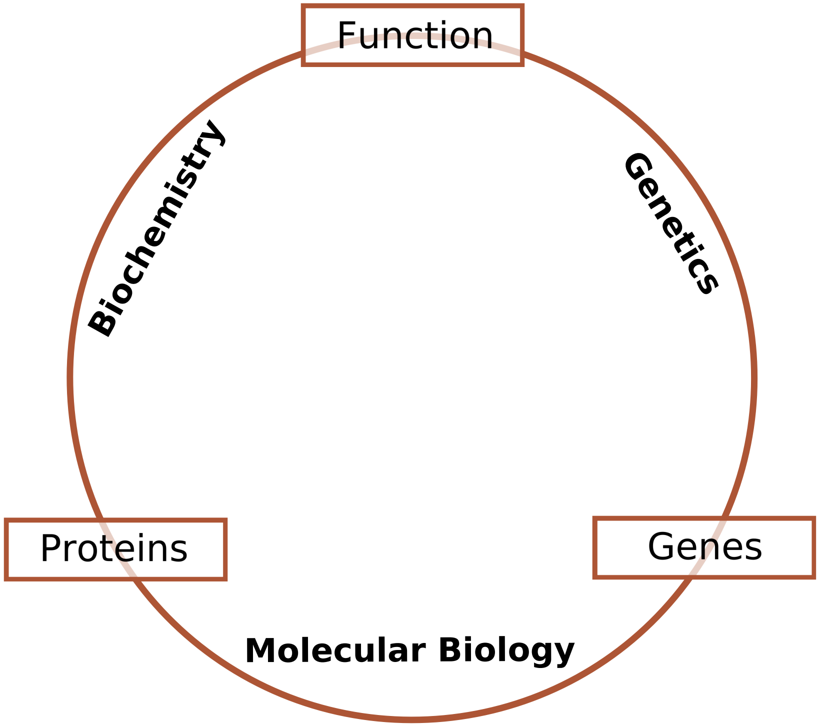
c) Zhou Pei-Yuan Center for Applied Mathematics, Tsinghua University, Beijing, China

Outline

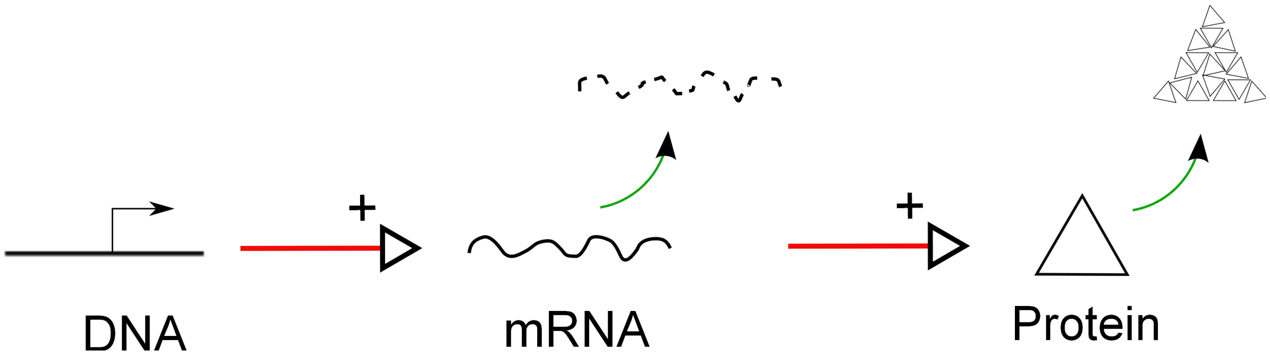
- Un brin de biologie moléculaire (et de philo)
- Quick review of the standard stochastic models of gene expression
- Can we explain the **Bursting phenomena** through an adiabatic reduction of the standard model?

Biologie moléculaire





Central dogma



Francis Crick: 1958

Jacques Monod, François Jacob et André Lwoff: 1965

Mise en évidence de la stochasticité en Biologie

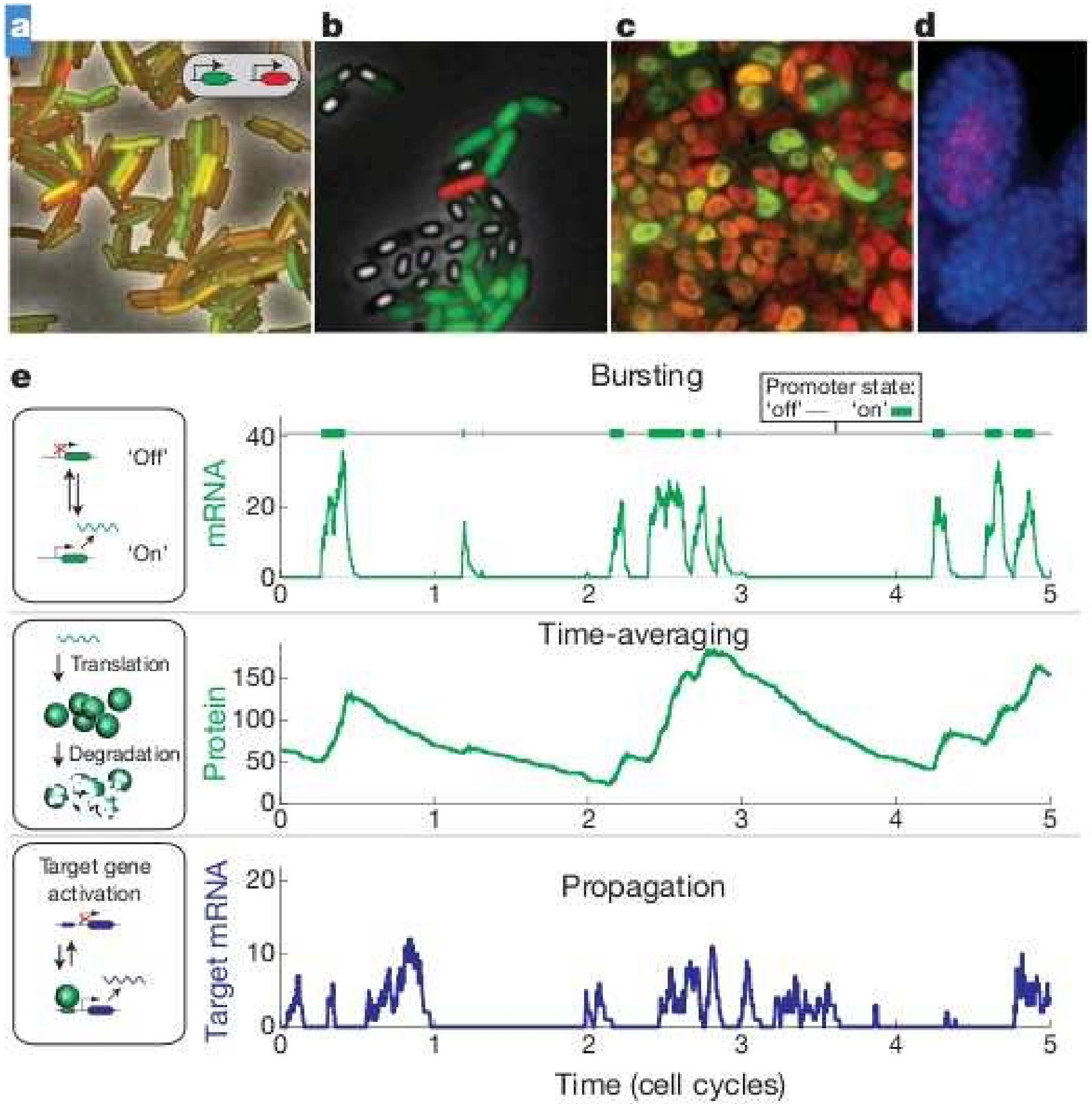


Figure 1 | Gene expression noise is ubiquitous, and affects diverse systems at several levels. a, *E. coli* expressing two identical promoters driving two different fluorescent proteins, in red and green, respectively. Because of

La stochasticité en Biologie

- Dogme central= vision cybernétique,“programme génétique”,“ordre dû”
- Laplace/Newton VS Poincaré
- Hilbert VS Gödel
- Shrödinger,C. Bernard VS ?? JJ Kupiec, G. Longo...

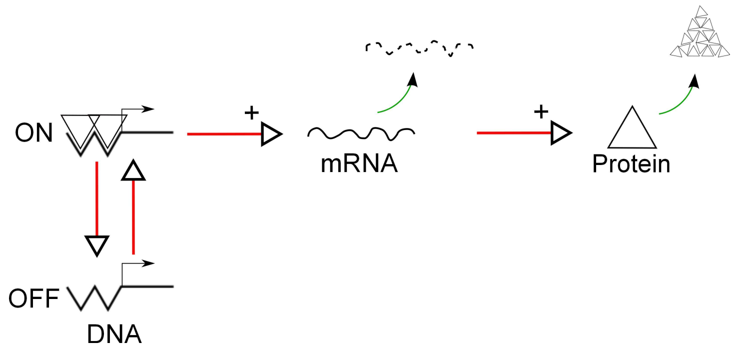
Paradigm shift

Composante essentielle
2 types d'aléatoire (cf Physique)

- Variabilité, multi-modalité, variation temporelle, oscillations... (individu)
- Bruit pour créer un signal à l'état macroscopique

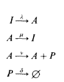
Darwinisme cellulaire (JJ Kupiec)

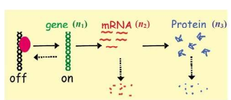
New Central dogma

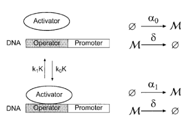


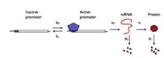
Studied either in a discrete state-space (Peccoud 95, Paulsson 05, Swain 08, Hornos 06) or in a continuous state-space (Kepler 01, Lipniacki 06, Mackey 2010)

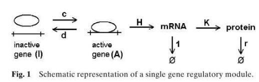
Problème: on ne sait pas résoudre ce modèle!!

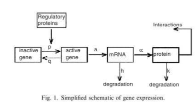




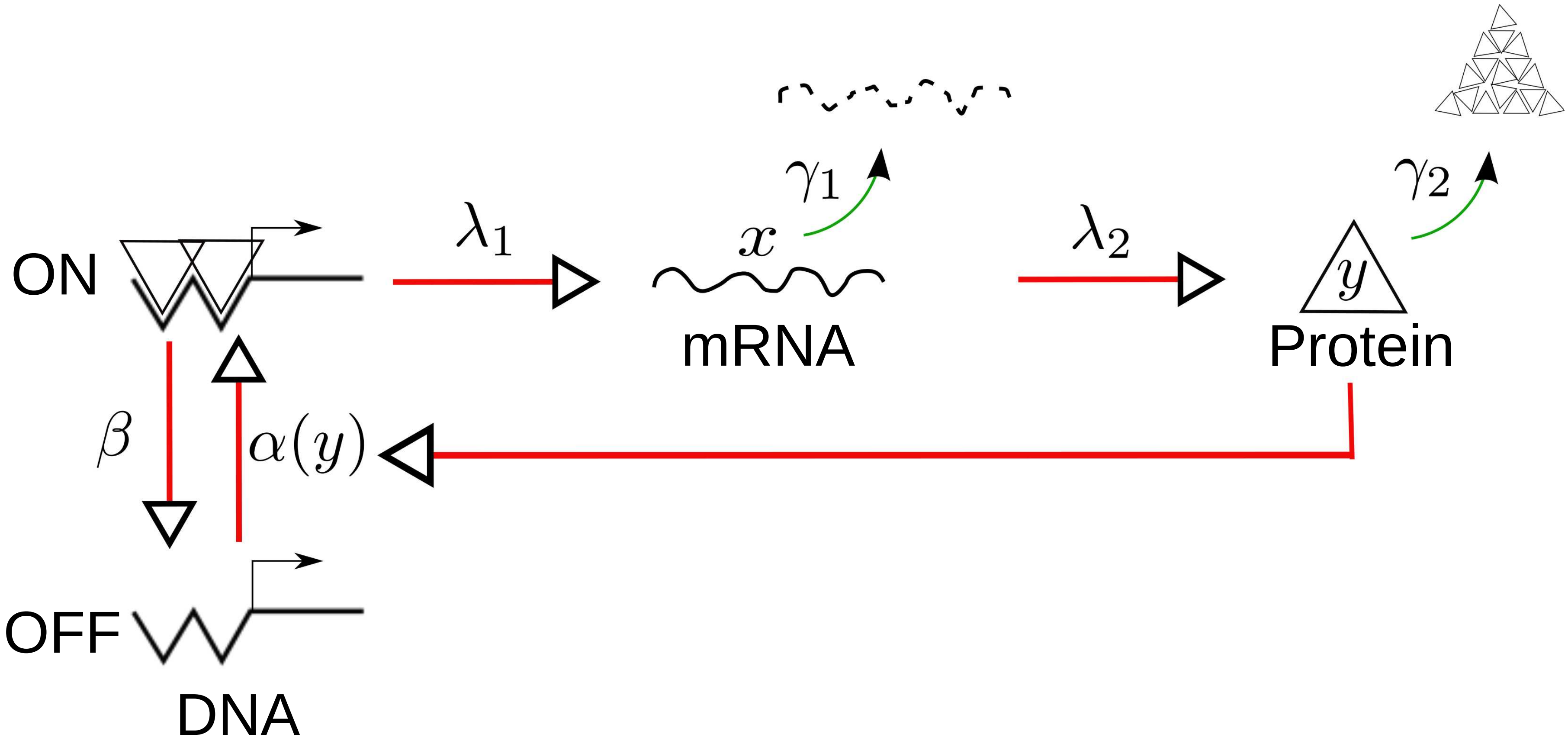




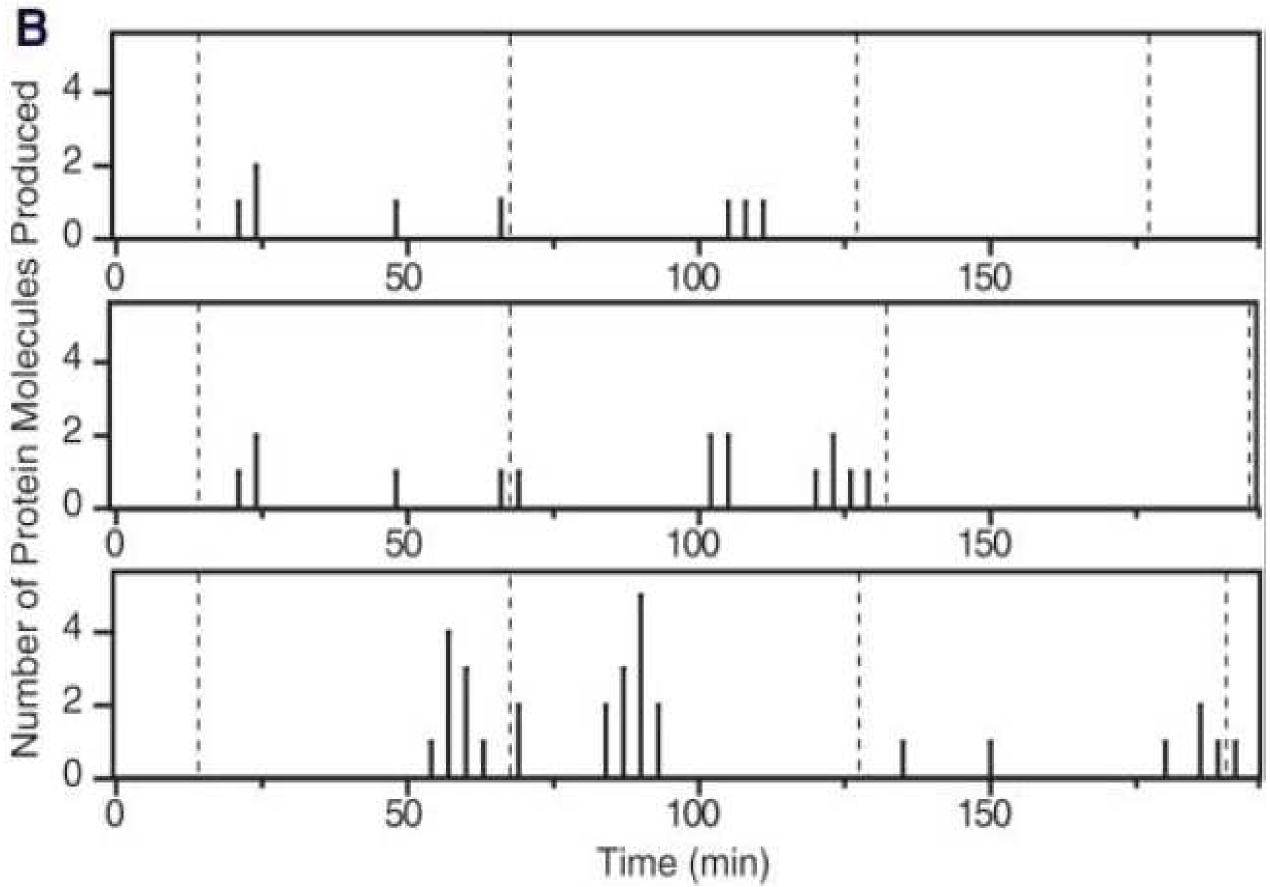




Self-regulated gene



The bursting phenomena



Yu et al. 06

Question 1) **Quand est-ce que le modèle stochastique prédit l'apparition de Burst?**
Question 2) **Que peut-on dire dans ces cas là?**

Méthode 1: Simulation

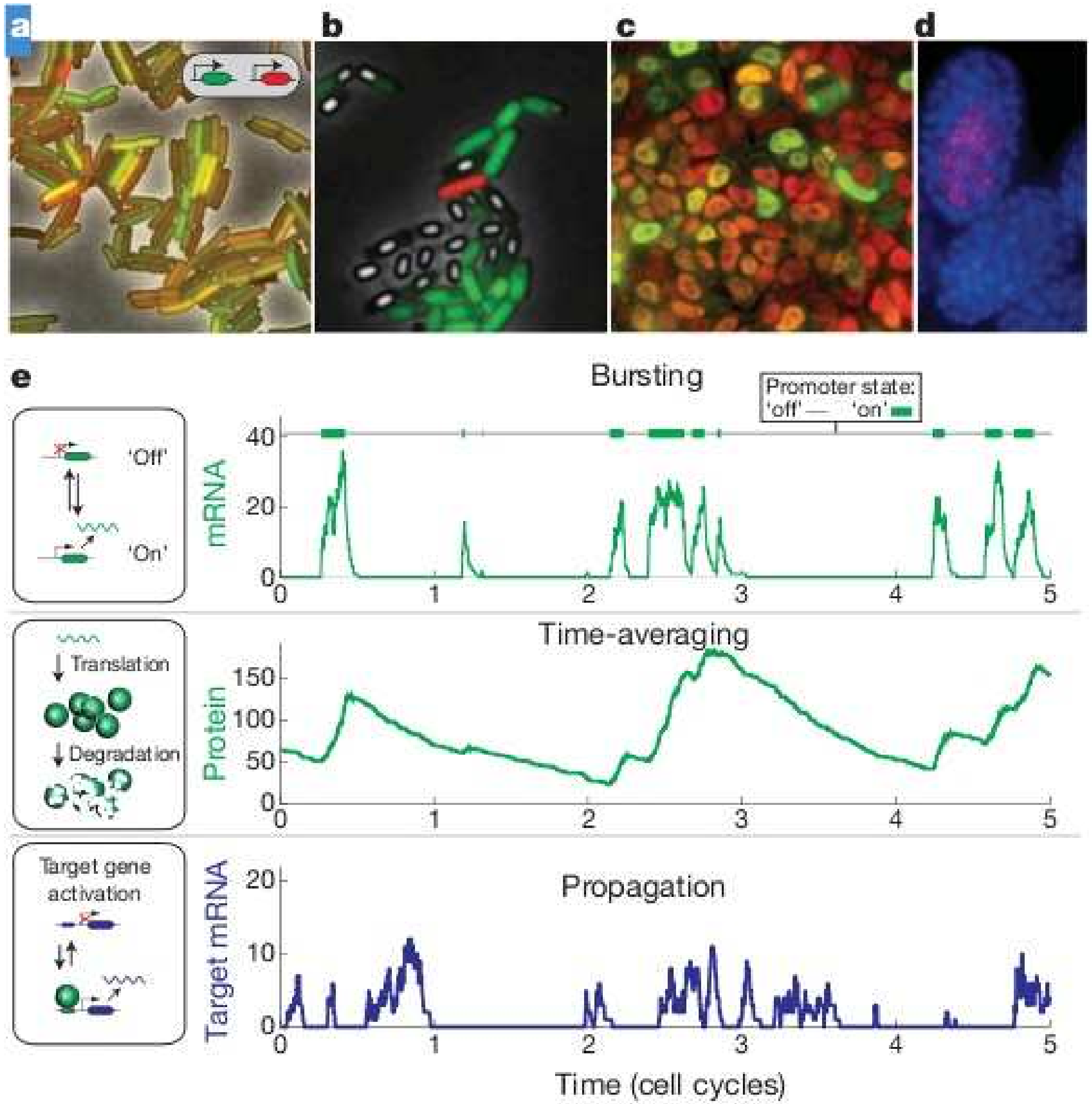


Figure 1 | Gene expression noise is ubiquitous, and affects diverse systems at several levels. a, *E. coli* expressing two identical promoters driving two different fluorescent proteins, in red and green, respectively. Because of

Méthode 2: Décomposition des variances

The normalized stationary variance in the number of protein molecules per cell follows:

$$\overbrace{\frac{\sigma_3^2}{\langle n_3\rangle^2}}^{\text{Total protein noise}} = \overbrace{\underbrace{\frac{1}{\langle n_3\rangle}}_{\text{Poisson}}}^{\text{From individual births and deaths}} + \overbrace{\underbrace{\frac{1}{\langle n_2\rangle}}_{\text{Poisson}}\;\underbrace{\frac{\tau_2}{\tau_3+\tau_2}}_{\text{One-step time-averaging}}}^{\text{From spontaneous mRNA noise}} + \overbrace{\underbrace{\frac{1-P_{\mathrm{on}}}{\langle n_1\rangle}}_{\text{Binomal}}\;\underbrace{\frac{\tau_2}{\tau_2+\tau_3}\,\frac{\tau_1}{\tau_1+\tau_3}\,\frac{\tau_1+\tau_3+\tau_1\tau_3/\tau_2}{\tau_1+\tau_2}}_{\text{Two-step time-averaging}}}^{\text{From forced mRNA noise, originating in gene activation-inactivation}}.$$

Paulsson et al. 05

$$\frac{\sigma_2}{<n>} = 1 + b$$

Méthode 3: Equations vérifié par la densité

En discret

$$\frac{dP_n}{dt} = \gamma_{n+1}P_{n+1} - \gamma_n P_n + \sum_{r=0}^{n} h_r \alpha_{n-r} P_{n-r} - \alpha_n P_n, \qquad (1)$$

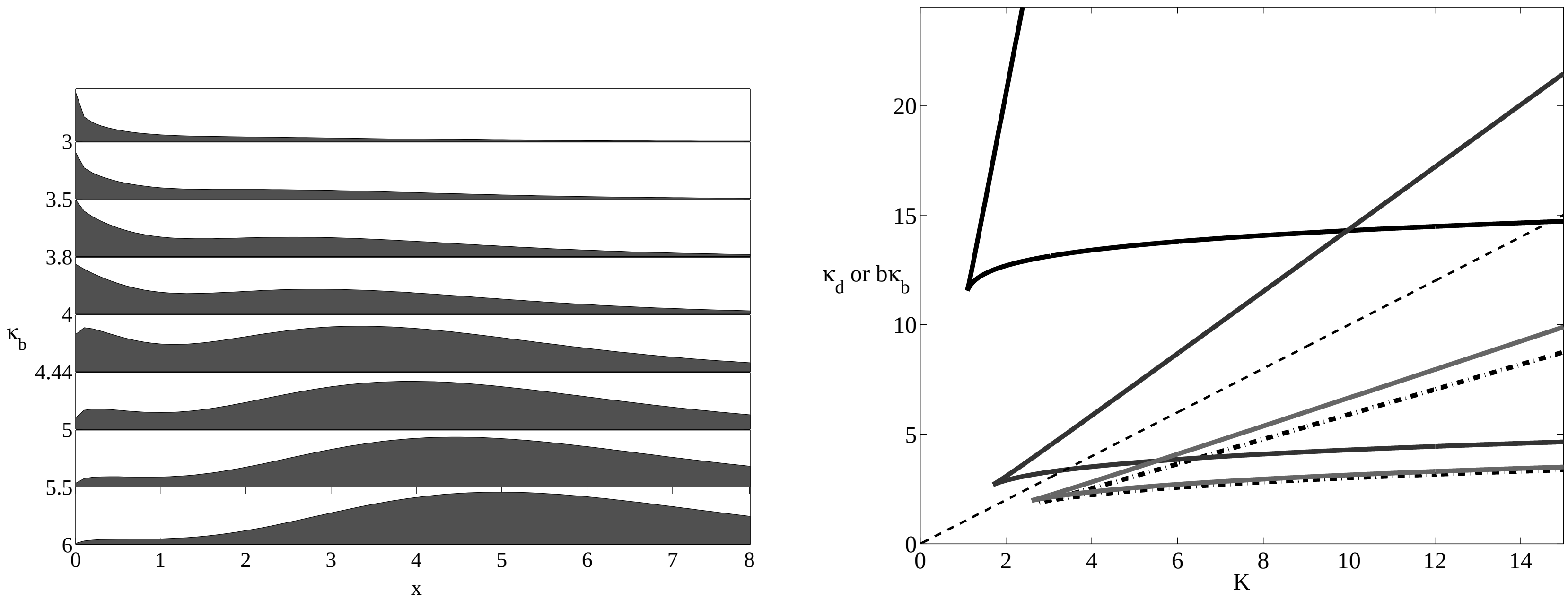
En continue

$$\frac{\partial u}{\partial t} = \frac{\partial(\gamma(x)u(t,x))}{\partial x} + \int_0^x \alpha(x)u(t,y)h(x-y)dy - \alpha(x)u(t,x). \qquad (2)$$

Une même solution stationnaire peut cacher plusieurs modèles!

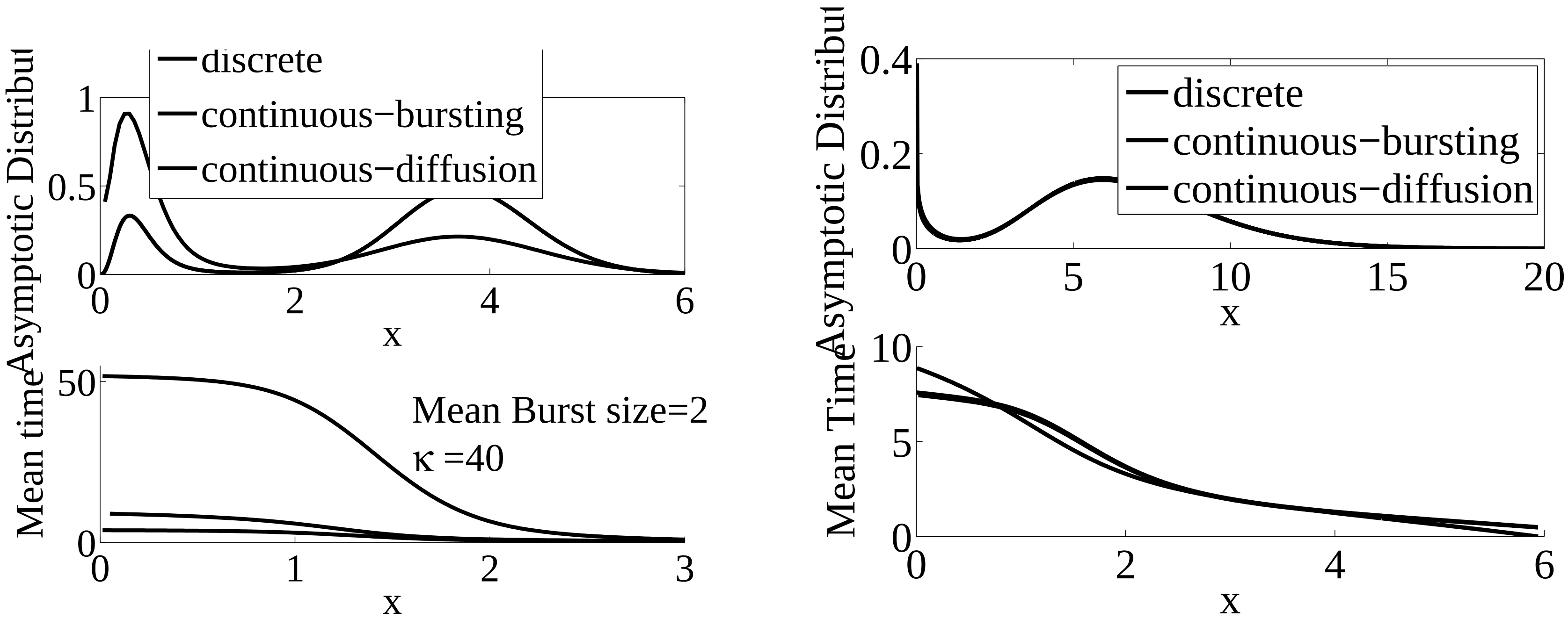
$$\frac{\partial u}{\partial t} = \frac{\partial(\gamma x u(t,x))}{\partial x} + \int_0^x \alpha(x)u(t,y)h(x-y)dy - \alpha(x)u(t,x). \quad (3)$$

$$\frac{\partial v}{\partial t} = \frac{\partial((\gamma x - \alpha(x))v(t,x))}{\partial x} + \frac{\partial^2(\sigma x v(t,x)}{\partial x^2} \quad (4)$$

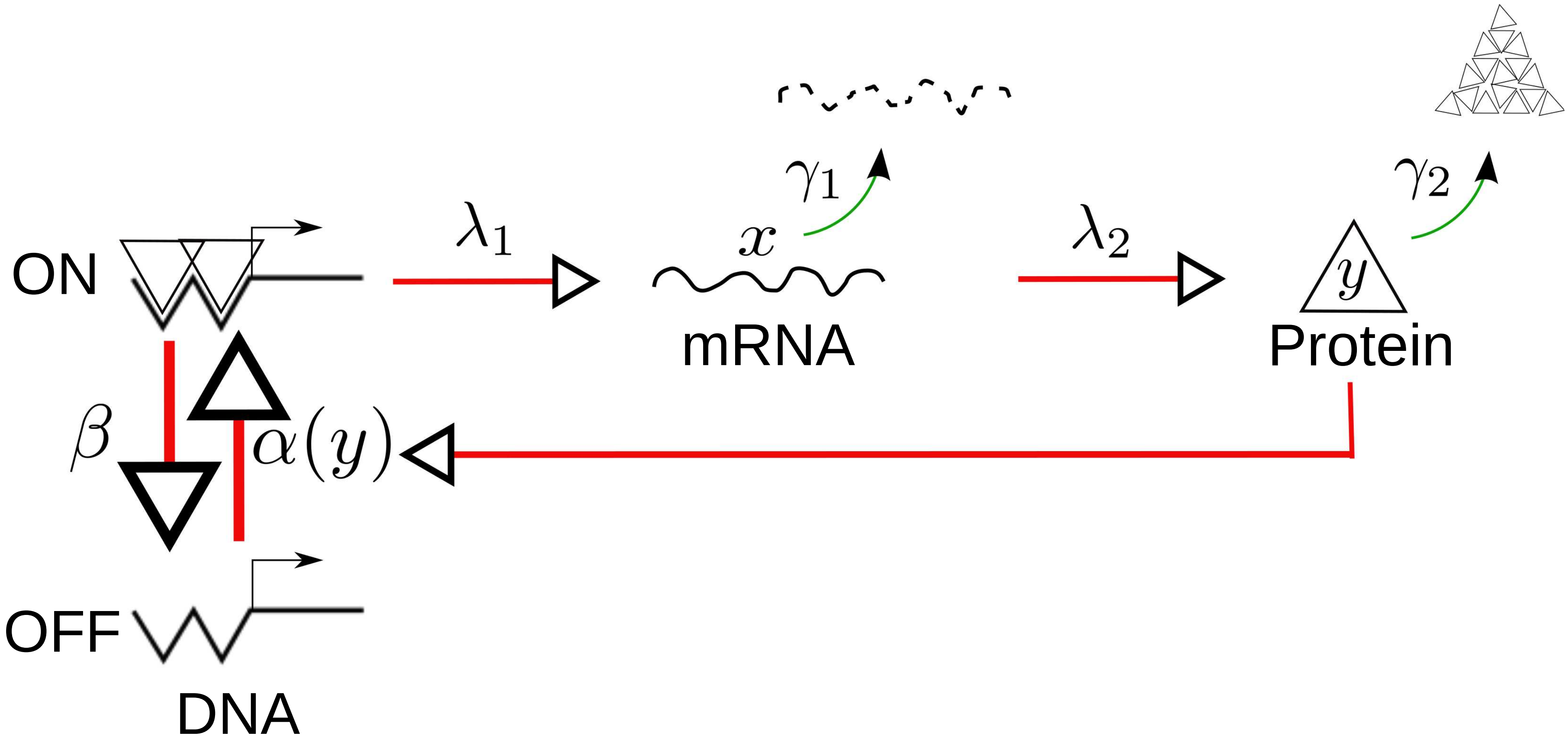


Une même solution stationnaire peut cacher plusieurs modèles!

Différentes dynamiques!

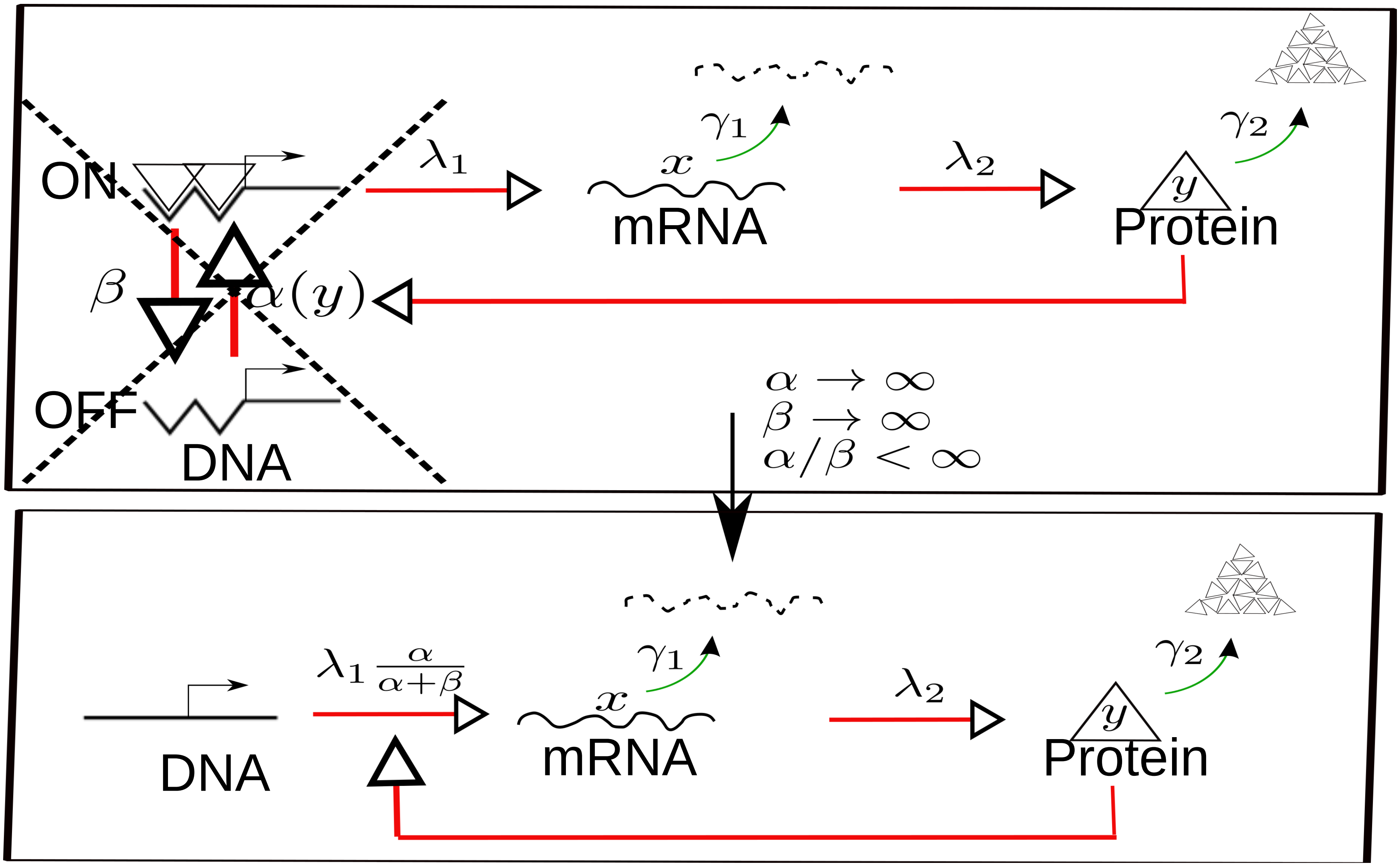


Méthode 4 : Reduction 1

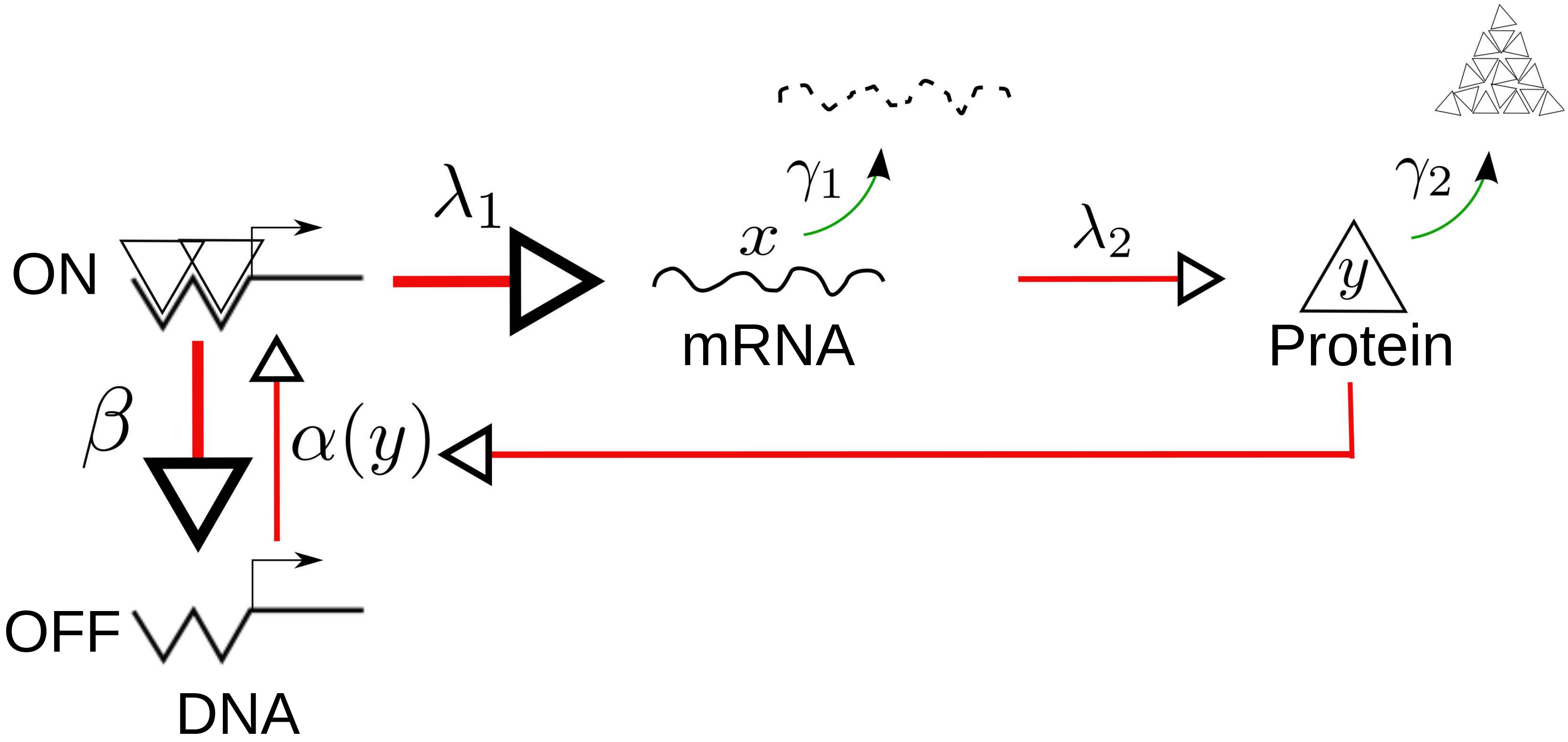


Bobrowski 06: Degenerate convergence of Semigroups.

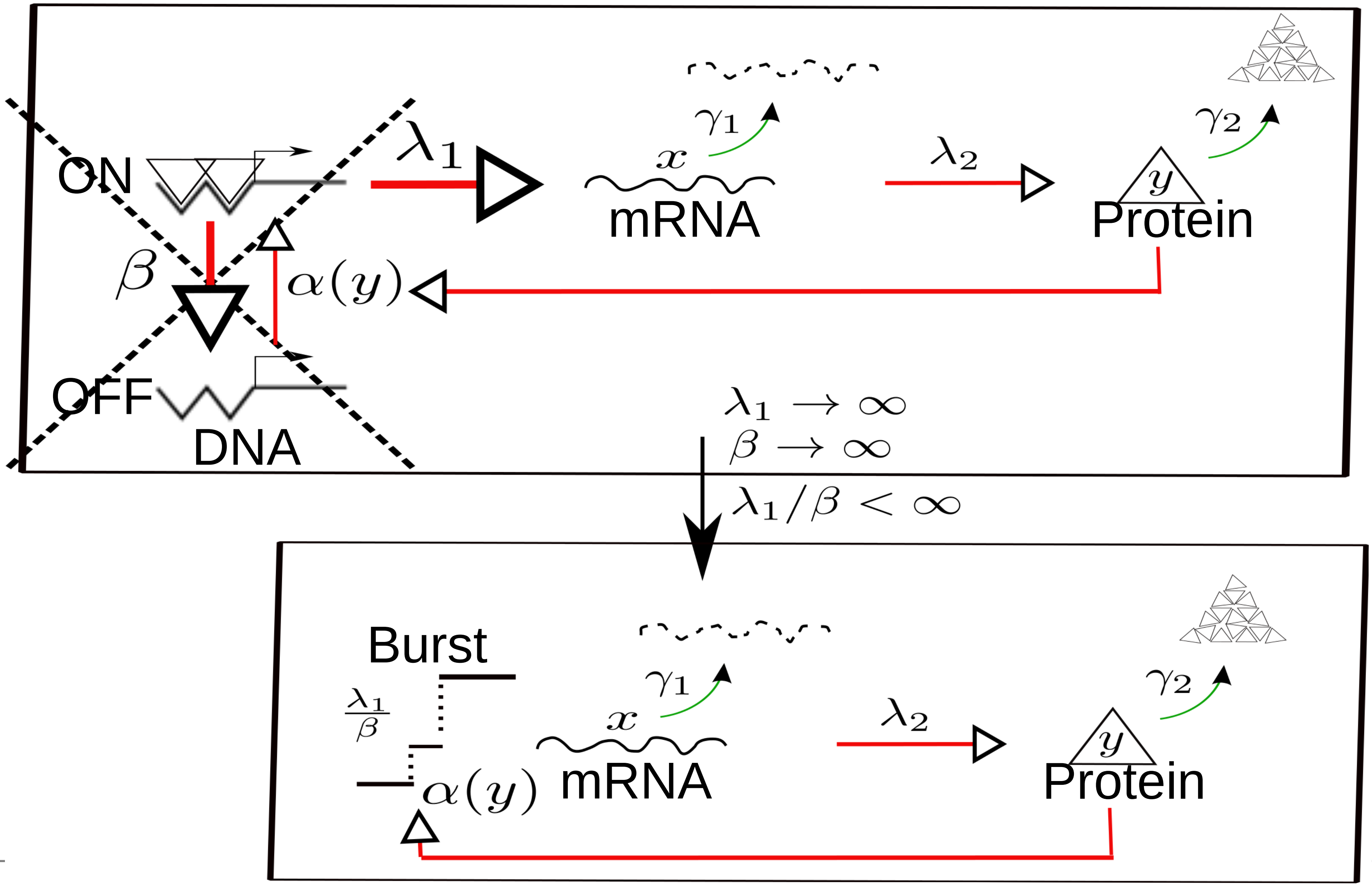
Méthode 4 : Reduction 1



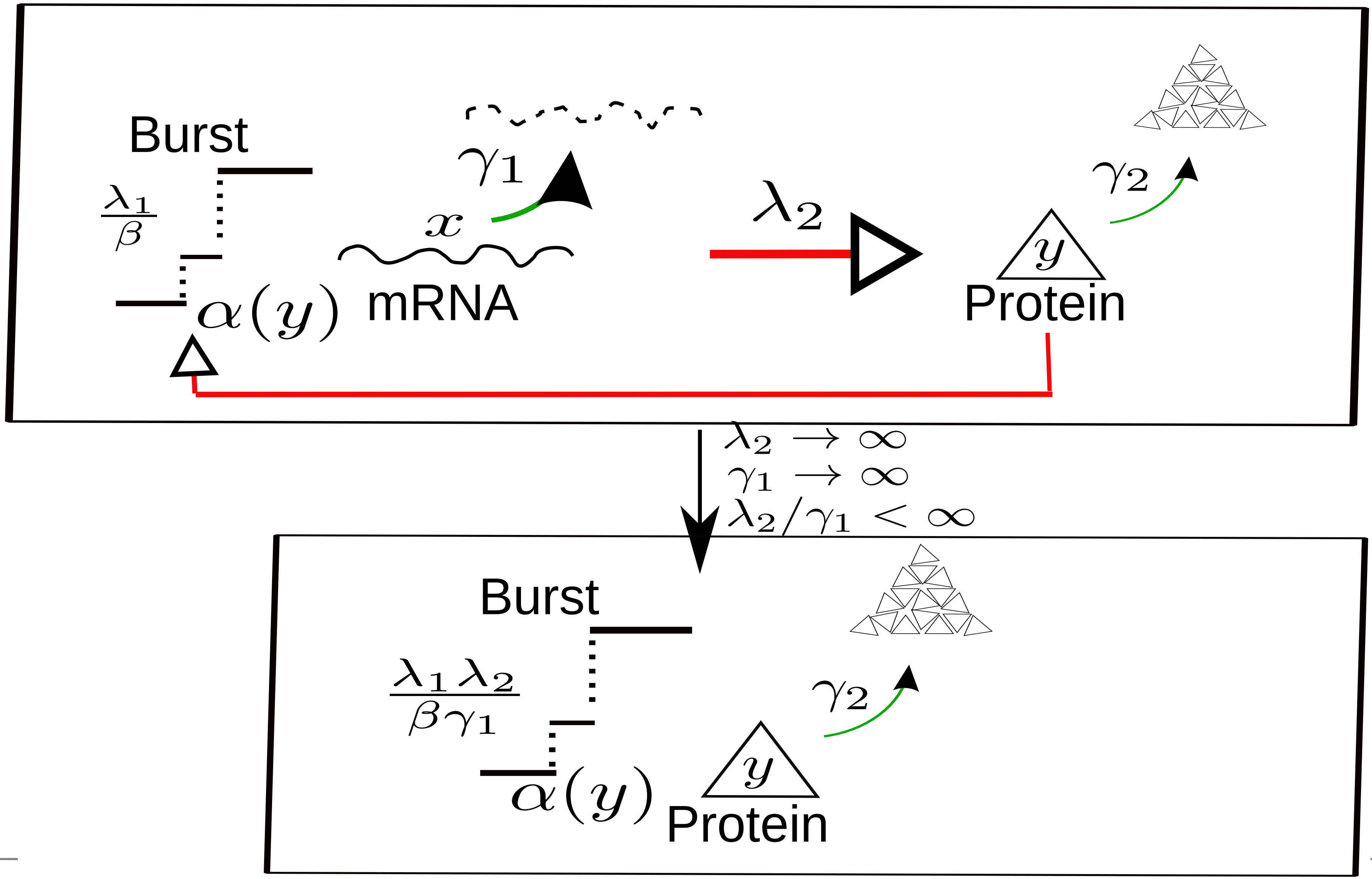
Méthode 4 : Reduction 2



Méthode 4 : Reduction 2



Méthode 4 : Reduction 3



Theoritical formulation

Initial process

$$\frac{dX}{dt} = \xi - \gamma X$$

where ξ is a dichotomous random process, which takes the values 0 or $\lambda > 0$ and switches at rate $\alpha(X)$ and β.

Final process, when $\beta \to \infty$, $\lambda \to \infty$, $\beta/\lambda \to b < \infty$

$$\begin{aligned} dX &= dN(\alpha(X), h) - \gamma X \\ X_t &= X_0 - \int_0^t \gamma X_{s^-} ds + \int_0^t \int_0^\infty \int_0^\infty z 1_{\{r \leq \alpha(X_{s-})\}} N(ds, dz, dr) \end{aligned}$$

where $N(ds, dz, dr)$ is a Poisson random measure on $(0, \infty) \times [0, \infty)^2$, of intensity $dsh(z)dzdr$, with h an exponential distribution of mean b.

Sample paths formulation

$$X_t = \begin{cases} \pi^0_{t-t_{k-1}} X_{t_{k-1}}, & t_{k-1} \le t < t_k, \\ \pi^1_{t-t_k} X_{t_k}, & t_k \le t < t_{k+1}, \end{cases}$$

where $t_k = t_{k-1} + \Delta^0 t_k$, $t_{k+1} = t_k + \Delta^1 t_k$ and

$$\begin{aligned} \Pr(\Delta^0 t_k \le t | Y_{t_{k-1}} = x) &= 1 - e^{-\int_0^t \alpha(\pi_s x) ds}, \quad t, x > 0. \\ \Pr(\Delta^1 t_k \le t | Y_{t_{k-1}} = x) &= 1 - e^{-\beta t}, \quad t, x > 0. \end{aligned}$$

$$X_t = \begin{cases} \pi^0_{t-t_{k-1}} X_{t_{k-1}}, & t_{k-1} \le t < t_k, \\ X_{t_{k-}} + e_k, & t = t_k, \end{cases}$$

where $t_k = t_{k-1} + \Delta^0 t_k$ and

$$\begin{aligned} \Pr(\Delta^0 t_k \le t | Y_{t_{k-1}} = x) &= 1 - e^{-\int_0^t \alpha(\pi_s x) ds}, \quad t, x > 0. \\ e_k &\sim h \end{aligned}$$

Semi-group proof

Let's define the semigroup $T_t f(x) = \mathbb{E}_x f(X_t)$ then the Dynkin's formula reads $T_t f(x) - f(x) = \int_0^t T_s \mathcal{A} f(X_s) ds$. For any test function f *in the domain of* A

$$\frac{d}{dt}\int_0^\infty p_t^0(x) f^0(x) + p_t^1(x) f^1(x) dx = -\gamma \int_0^\infty x p_t^0(x) \frac{df^0}{dx} dx$$
$$- \int_0^\infty (\gamma x - \lambda) p_t^1(x) \frac{df^1}{dx} dx + \int_0^\infty (\alpha(x) p_t^0(x) - \beta(x) p_t^1(x))(f^1(x) - f^0(x)) dx$$

For any test function such that $f^0(x) = f^1(x) = f(x)$,

$$\frac{d}{dt}\int_0^\infty (p_t^0(x) + p_t^1(x)) f(x) dx = - \gamma \int_0^\infty x (p_t^0(x) + p_t^1(x)) \frac{df}{dx} dx$$
$$+ \lambda \int_0^\infty p_t^1(x) \frac{df}{dx} dx$$

Trotter, Sova, Kurtz, Mackevicius

Theorem 1. *Let X and X^n be Feller processes in $(0, \infty)$ with semigroups (T_t) and $(T_{n,t})$ respectively. Then the two conditions are equivalent*

- $T_{n,t}f \to T_t f$ *for every* $f \in \mathcal{C}_0$*, uniformly in time for bounded* $t > 0$.
- *If* $X_0^n \to X$ *in distribution, then* $X^n \to X$ *in distribution.*

where $\mathcal{C}_0$ *is the class of continuous functions with* $f(y) \to 0$ *as* $y \to \infty$.

The two semi-group

Initial process

$$\frac{d}{dt}\int_0^\infty (p_t^0(x)+p_t^1(x))f(x)dx = - \gamma\int_0^\infty x(p_t^0(x)+p_t^1(x))\frac{df}{dx}dx + \lambda\int_0^\infty p_t^1(x)\frac{df}{dx}dx$$

Final process

$$\frac{d}{dt}\int_0^\infty p_t(x)f(x)dx = -\gamma\int_0^\infty xp_t(x)\frac{df}{dx}dx + \int_0^\infty \alpha(x)p_t(x)\Big(\int_x^\infty h(y-x)f(y)dy - f(x)\Big)dx$$

A particular choice for a test function

For any test function such that $f^0(x) = 0$, $f^1(x) = g(x)$,

$$\begin{aligned}\frac{d}{dt}\int_0^\infty p_t^1(x)g(x)dx &= -\gamma\int_0^\infty xp_t^1(x)\frac{dg}{dx}dx + \lambda\int_0^\infty p_t^1(x)\frac{dg}{dx}dx \\ &+ \int_0^\infty \alpha(x)p_t^0(x)g(x)dx - \beta\int_0^\infty p_t^1(x)g(x)dx\end{aligned}$$

One can perform a ***quasi steady-state approximation***, which gives

$$\begin{aligned}\lambda\int_0^\infty p_t^1(x)g(x)dx &= -\frac{\gamma\lambda}{\beta}\int_0^\infty xp_t^1(x)\frac{dg}{dx}dx + \frac{\lambda}{\beta}\int_0^\infty \alpha(x)p_t^0(x)g(x)dx \\ &+ \frac{\lambda^2}{\beta}\int_0^\infty p_t^1(x)\frac{dg}{dx}dx\end{aligned}$$

Iterating

Iterating the process, one can find,

$$\int_0^\infty \lambda p_t^1(x)\frac{df}{dx}dx = \sum_{i\geq 1}\left(\frac{\lambda}{\beta}\right)^i \int_0^\infty \alpha(x)(p_t^0(x)+p_t^1(x))\frac{d^i f}{dx^i}$$

$$- \quad \sum_{i\geq 1}\left(\frac{\lambda}{\beta}\right)^i \int_0^\infty \gamma x p_t^1(x)\frac{d^i f}{dx^i} - \sum_{i\geq 1}\left(\frac{\lambda}{\beta}\right)^i \int_0^\infty \alpha(x) p_t^1(x)\frac{d^i f}{dx^i}$$

The first sum give the jump kernel

$$\int_0^\infty \alpha(x)(p_t^0(x)+p_t^1(x))\Big(\int_x^\infty h(y-x)f(y)dy - f(x)\Big)dx,$$

where h is a distribution whose moments are given by

$$\mathcal{E}^i[h] = i!\left(\frac{\lambda}{\beta}\right)^i$$

The two others sum are shown to be arbitrary small.□

summary

- The same is working in a discrete formalism
- Powerful tools to perform adiabatic reduction in Hybrid systems
- The "full-model" can now be analysed according its different limit behaviour.
- Other scaling are to be investigated

summary

- Caractériser les oscillations dans le modèle de Bursting 2D (Vellela, Qian...)
- Modèle spatiale et/ou à retard (M. Chaplain, S.Bernard...)
- Variabilité à l'échelle d'une population de cellules (S. Méléard, N. Champagnat, V. Bansaye...)
- Modèles "darwinistes" de différentiation (JJ Kupiec, Kaern, Paulsson, Doumic, Perthame, Huang, Kauffman...)

Thank you for your attention!