

Phénomènes stochastiques en biologie moléculaire

O.Radulescu

Biologie moléculaire

Le dogme central de la génétique : ADN (gène) \Rightarrow ARN \Rightarrow Protéines

Pourtant la complexité d'un organisme vivant n'est pas donnée par le nombre de gènes, mais par les relations entre eux. L'organisme humain possède un nombre modeste de 2.10^4 gènes.

La même propriété dans l'organisation du cerveau : les connexions entre neurones produisent la complexité.

Depuis une dizaine d'années des techniques expérimentales telles que les puces d'ADN, la microscopie confocale et la spectroscopie des protéines ont permis l'étude des interactions. On parle maintenant de:

Réseaux de gènes (transcriptome)

Réseaux de protéines (protéome)

Projet MathResoGen, ACI IMPBio

nov.2003-nov.2006

Equipes : IRMAR (O.R., A.Debussche, D.Grigoriev, D.Petritis), IRISA (A.Siegel, M.le Borgne), INSERM Rennes (N.Theret), INRA Rennes (S.Laggarigue,M.Douaire), CNRS et CHU Salpetriere (A.Lilienbaum), IMAB U.Bourgogne (E.Pécou), INRIA Rhone-Alpes (F.Grognard)

Problèmes biologiques :

- Mécanismes de signalisation du $\text{NF}\kappa\text{B}$
- Cancer du foie et $\text{TGf}\beta$
- Métabolisme des lipides chez le poulet

Problèmes mathématiques :

Réduction de modèles
Equilibres et leurs déplacements
Dynamique de réseaux
Stochasticité

Informatique :

Aide à la création de modèles
Analyse qualitative et diagnostic
Théorie des signaux et dynamique

Le stochastique en biologie moléculaire

L'origine supposée du bruit en biologie moléculaire est la nature particulière des interactions.

En fait, comme E. Shrödinger le remarquait la naissance de l'ordre à partir des combinaisons aléatoires d'un petit nombre de molécules est une spécificité du vivant qui tient presque du domaine du miracle.

Les processus moléculaires ont lieu dans un volume petit (le noyau, les ribozomes pour les eucaryotes, la proximité de l'ADN pour le procaryotes).

Ces processus impliquent des interactions protéine/protéine, protéine/ADN, protéine/ARN (complexes et soumises à des perturbations aléatoires)

Exemples

- transcriptionels
 - di ou multimérisation,
 - attachement,
 - répression,
 - activation,
 - remodelage de la chromatine,
 - polymérisation,
 - moteurs moléculaires
- traductionels
 - attachement,
 - synthèse protéique,
 - repliement
- réactions biochimiques du métabolisme

Les nombres de molécules des différentes espèces sont des variables aléatoires.

En aval des régulations elles sont supposées du type Poisson (analogie avec les files d'attente?)

D'autres lois : binomiale négative.

Approximation de la dynamique par des réactions

Les interactions entre molécules sont des événements discrets : réactions chimiques.

Temps entre réactions \ll durée des réactions.

Réactions binaires : le nombre des réactants ≤ 2 .

Exemple : transcription

- 1) $TF + TF \rightleftharpoons TFD$ (dimère)
- 2) $TFD + PROM \rightleftharpoons TFD PROM$
- 3) $PROM + ACT \rightleftharpoons PROM ACT$
- 4) $TFD + PROM ACT \rightleftharpoons TFD PROM ACT$
- 5) $TFD PROM + RNAP \Rightarrow TFD PROM + RNAP^*$
- 6) $TFD PROM ACT + RNAP \Rightarrow TFD PROM ACT + RNAP^*$
- 7) $RNAP \Rightarrow RNA + RNAP$

Observations expérimentales

- Nombres de copies petits et fluctuants dans eucaryotes: 50% des gènes de *S.Cerevisae* (levure) sont exprimés à moins de un exemplaire en moyenne par cellule. La transcription est intermittente. Il semblerait aussi que le bruit est source de variabilité épigénétique : il contribue de façon significative à l'hétérogénéité au sein d'une population de cellules clones.
- Bistabilité des procaryotes: Le bacteriophage λ de *E.Coli* peut exister en deux états différents: lytique (multiplication du virus) et lysogénique (virus dormant). Les deux états se distinguent par l'expression de gènes différents (*cro* et *cl*) dont les statistiques cellule par cellule sont disponibles.
- Évolution des distributions empiriques de *relA* dans des fibroblastes suite à l'induction par interleukine-1.

Trois questions physiques

1. Le bruit, peut-il renseigner sur les réseaux de gènes?
2. Le bruit, est-il nuisible ou indispensable au fonctionnement?
3. Le bruit, se manifeste-t-il dans le phenotype? ou
Comment se propage le bruit dans les cascades des régulations?

Description mathématique

Processus de Markov à sauts $X(t) \in \mathbb{Z}_+^n$

- Ensemble de sauts $S = \{\theta_i, i = 1, k\} \subset \mathbb{Z}^n$
(colonnes de la matrice stoechiométrique)
- Paramètres infinitésimaux

$$q_{X, X+\theta_i} = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \mathbb{P}[X(t+\tau) = X+\theta_i | X(t) = X]$$

Exemple :

$$A + B \rightleftharpoons C, \theta = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

$$q_{X, X+\theta} = V k_1 [A][B]$$

$$q_{X, X-\theta} = V k_2 [C]$$

Algorithmes de simulation

Processus de Markov à sauts : temps d'attente $\exp(\lambda(X))$, loi de sortie loi discrète $P_i(X)$ sur S .

$$\lambda(X) = \sum_i q_{X, X+\theta_i}, \quad P_i(X) = (\lambda(X))^{-1} q_{X, X+\theta_i}$$

Algorithme de Gillespie:

1. Générer $\tau \sim \exp(\lambda(X))$
2. $t = t + \tau$
3. Générer $i \sim P_i(X)$
4. $X = X + \theta_i$
5. goto 1

Algorithme accéléré, Gillespie 2001

Choisir une durée fixée déterministe τ t.q.
 $q_{X, X+\theta_i} \approx \text{const.}$

Le nombre de réactions d'un type i est $N_i \sim \mathcal{P}(q_{X, X+\theta_i}\tau)$ ou même $\mathcal{N}(q_{X, X+\theta_i}\tau, q_{X, X+\theta_i}\tau)$

1. Choisir τ suffisamment petit.
2. $t=t+\tau$
3. Générer $N_i \sim \mathcal{P}(q_{X, X+\theta_i}\tau)$
4. $X = X + \sum_i N_i \theta_i$
5. goto 1

Approximation de diffusion

Cadre: vitesses de toutes les reactions sont grandes

Vecteur de concentrations $x = X/V$

$q_{X, X+\theta_i} = V f_i(x)$, $f_i(x)$ vitesse de réaction

entre t et $t + dt$ le nombre de réactions i est une variable de Poisson de paramètre $\Lambda = V f_i(x)dt$. l'approximation de diffusion consiste de l'approcher par une variable normale $\mathcal{N}(\Lambda, \Lambda)$. les réactions sont considérées indépendantes.

$$x(t + dt) = x(t) + F(x)dt + \frac{1}{\sqrt{V}} \sum_{i=1}^m \theta_i \sqrt{f_i(x)} dB_i$$

$$\text{où } F(x) = \sum_i \theta_i f_i(x)$$

N.B.: m Browniens indépendants, un pour chaque réaction

dynamique déterministe

$$\frac{dx}{dt} = F(x) = \sum_i \theta_i f_i(x)$$

Fluctuations autour de l'équilibre

Paulsson (Nature 2004)

Processus de naissance et mort à deux composantes:

$(n_1, n_2) \Rightarrow (n_1 \pm 1, n_2)$ avec paramètres $R_1^\pm(n_1)$.

$(n_1, n_2) \Rightarrow (n_1, n_2 \pm 1)$ avec paramètres $R_2^\pm(n_1, n_2)$.

$$\frac{\sigma_2^2}{\langle n_2 \rangle^2} = \frac{1}{\langle n_2 \rangle H_{22}} + \frac{\sigma_1^2}{\langle n_1 \rangle^2} \frac{H_{21}^2}{H_{22}^2} \frac{H_{22}/\tau_2}{H_{11}/\tau_1 + H_{22}/\tau_2}$$

Paulsson met cette relation à la base d'une "Théorie Stochastique des Systèmes Biochimiques".

Bruit comme source d'information

Plus modestement, on peut remarquer que cette relation est une conséquence de la relation de fluctuation-dissipation (à la base de la théorie de la réponse linéaire en physique)

$$JV + VJ^t = -\sigma\sigma^t$$

qu'on peut trouver dans la limite thermodynamique de l'équation maitresse ou en utilisant le processus de Ornstein-Uhlenbeck.

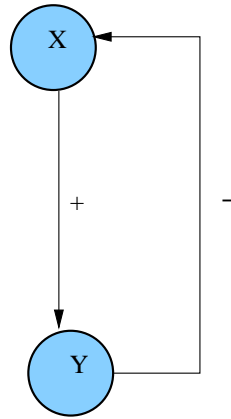
Les fluctuations renseignent sur le jacobien de $J = \nabla F$. Par un choix astucieux des cibles en fluorescence on peut emporter des informations supplémentaires ou tester un modèle.

Bruit nuisible

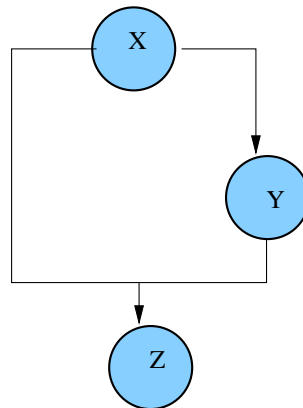
Le bruit peut être nuisible dans plusieurs situations :

- 1) Systèmes trop sensibles, par exemples systèmes bistables proches de la bifurcation col-noeud.
- 2) Horloges biologiques : perte de synchronicité.

Mécanismes de reduction du bruit



Boucle de retroaction négative.



Detecteur de persistance

Bruit indispensable

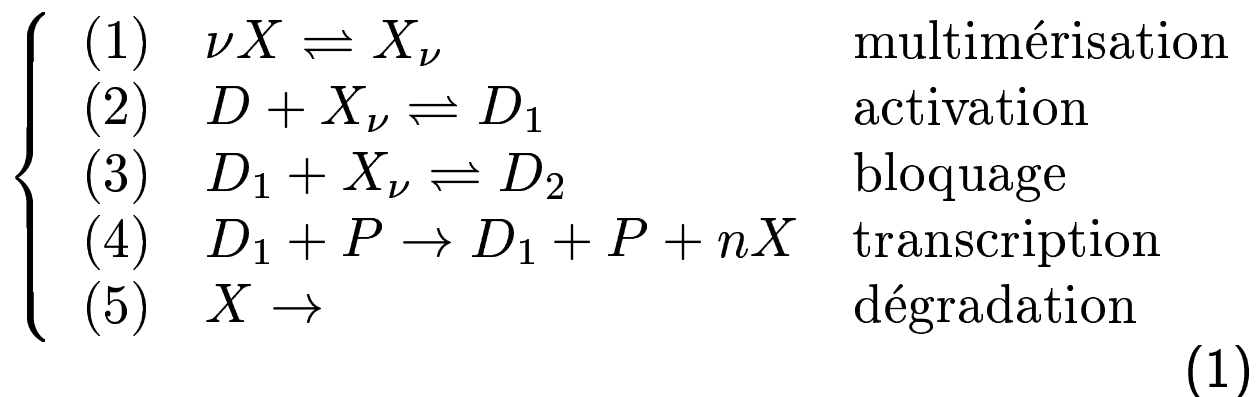
Le bruit augmente la variabilité isogénique.

Le bruit facilite les transitions.

Le bruit augmente la sensibilité : résonance stochastique.

Bruit et multistationarité

λ -phage, E.Coli. Une bactérie infectée peut être dans un des deux états: lytique, permettant la multiplication du virus et conduisant à sa propre destruction, ou lysogénique, permettant l'incorporation passive de l'ADN viral dans son propre ADN, sans multiplication. Soumise à un stress, par exemple des irradiations UV ou un changement de température, le virus dormant en état lysogénique peut basculer en état lytique, détruire et quitter la cellule.



Quelques résultats

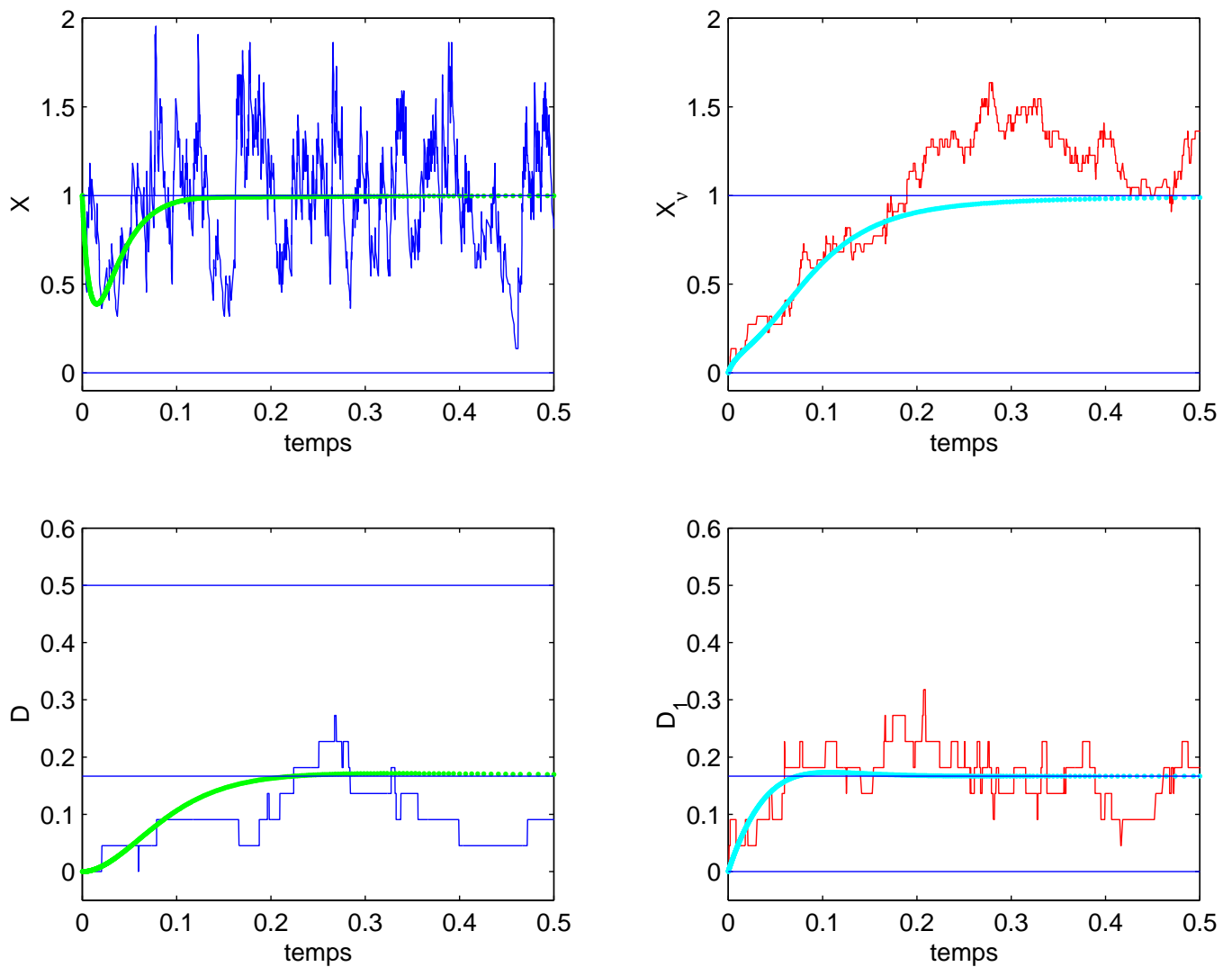


Figure 1: Evolution déterministe et stochastique vers l'attracteur ($y = 1$) partant de l'état $(1, 0, 0, 0)$ (ce n'est pas un point stable). Les valeurs de (X, X_v, D, D_1) correspondant aux deux attracteurs sont marquées par des lignes horizontales

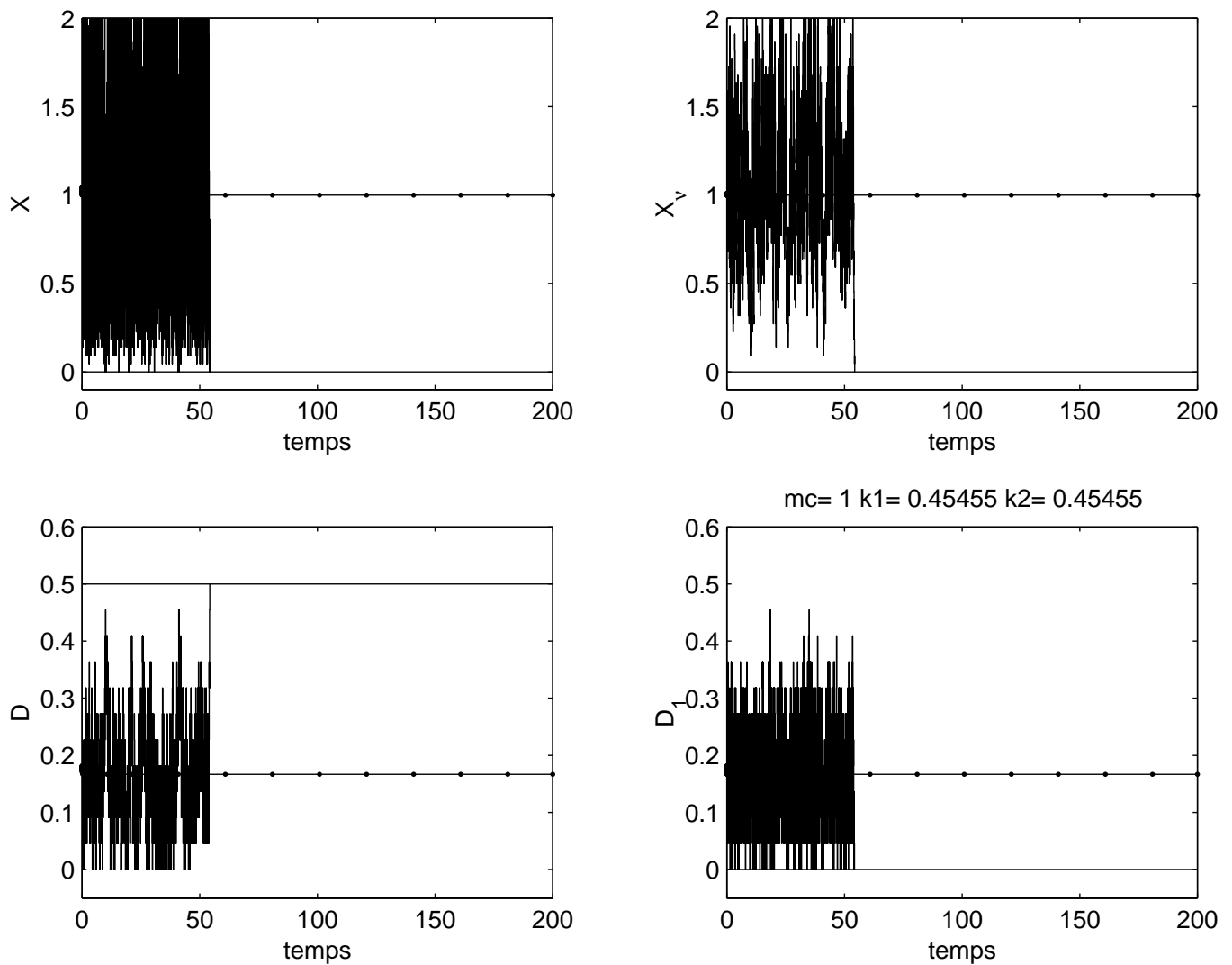


Figure 2: Sortie de l'attracteur ($y = 1$) et absorption dans le deuxième attracteur en ($y = 0$). C'est le phénomène de basculement. On remarque que X , D_1 atteignent des valeurs nulles plusieurs fois mais le basculement se fait uniquement lorsque $X_v = 0$ aussi.

Le bruit, se manifeste-t-il dans le phenotype?

50 % de gènes dans la levure ne sont pas utilisés en conditions normales.

Ces gènes sont exprimés en nombre très faible de copies (la transcription est un processus de Poisson de faible intensité).

Ces gènes arrivent-ils être traduits?

Comment le bruit se propage dans les cascades de regulation?

Problèmes mathématiques et de modélisation

Etude fine des processus de Markov à sauts (petites et grandes déviations)

Etude des diffusions avec sauts : deux classes de reactions (lentes et rapides). Ito (1951), Gihman et Skorohod (1972)

Dynamique à deux échelles de temps, perturbations singulières. Algorithmes de simulation : Rao et Arkin (J.Chem.Phys, en cours de publication). Fluctuations des trajectoires : Berglund and Gentz (2003)

Cascades de régulation : cas du $NF_{\kappa}B$.

Biblio: <http://name.math.univ-rennes1.fr/ovidiu.radulescu/bib/bib.htm>