

Théorèmes limites pour les processus de Markov à sauts

Synthèse de résultats et applications en biologie moléculaire

Ovidiu Radulescu* — Aurélie Muller — Alina Crudu***

**IRMAR UMR CNRS 6625*

Université de Rennes1, Campus de Beaulieu

F-35042 Rennes

Ovidiu.Radulescu@univ-rennes1.fr

***CEMAGREF*

3 bis quai Chauveau CP 220

F-69336 Lyon cedex 09

RÉSUMÉ. Les processus de Markov à sauts permettent la modélisation des phénomènes stochastiques en biologie moléculaire. Néanmoins, il y a peu de résultats mathématiques sur la dynamique de ces processus. De plus, leur simulation sur ordinateur rencontre des difficultés dues au temps d'exécution. Nous présentons des résultats permettant de réduire la complexité de la dynamique stochastique. Ces méthodes utilisent des théorèmes limites probabilistes.

ABSTRACT. Markov jump processes can model stochastic phenomena in molecular biology. Nevertheless, there are few mathematical results on the dynamics of these processes in their full generality. Furthermore, their computer simulation is time expensive. We present here some mathematical results allowing to reduce the complexity of stochastic dynamics. These methods employ probabilistic limit theorems.

MOTS-CLÉS : processus de Markov, théorèmes limites, réseaux de gènes.

KEYWORDS: Markov processes, limit theorems, gene networks.

1. Introduction

Plusieurs observations nous conduisent à introduire une approche stochastique dans la dynamique de réseaux de gènes en biologie moléculaire. 1) La variabilité épigénétique : dans les mesures par des méthodes de fluorescence des niveaux d'expression du même gène dans des cellules du même type, on constate des distributions statistiques des niveaux d'expression, dépendantes des conditions externes (Rao *et al.*, 2002). 2) Les phénomènes de faible dosage : certains gènes de faible activité peuvent être exprimés par certaines cellules et ne pas l'être par d'autres cellules du même tissu. C'est le phénomène de haploinsuffisance (Cook *et al.*, 1998), mais aussi le cas des promoteurs transcriptionnels de faible activité (Kaern *et al.*, 2005). 3) La production intermittente, par paquets, des protéines (Ozbudak *et al.*, 2002; Blake *et al.*, 2003; Swain *et al.*, 2002); 4) Les phénomènes de bistabilité : dans le cas de la bistabilité du phage- λ , des distributions bimodales du niveau du gène lytique cro sont observées et attribuées à la dynamique stochastique (Arkin *et al.*, 1998; Hasty *et al.*, 2000; Tian *et al.*, 2004). De même, la différenciation cellulaire en morphogénèse peut être abordée de façon stochastique (Laforge *et al.*, 2005). 4) Dans le cas du cancer, sous l'incidence du même stimulus, certaines cellules développent le cancer, d'autres non (Iwasa *et al.*, 2004). Ces différents phénomènes montrent l'importance d'une approche stochastique pour l'étude du fonctionnement des systèmes moléculaires complexes.

La compréhension des phénomènes de réponse aux signaux, différents en fonction du contexte cellulaire, ne peut s'appréhender que par une approche systémique des réseaux de signalisation et des réseaux de gènes en interaction. L'hypothèse courante de modélisation est de considérer que la dynamique de tels systèmes est essentiellement déterministe et que le bruit est une perturbation sans conséquence (Tyson *et al.*, 2003). Néanmoins, les exemples présentés plus haut suggèrent que les aspects stochastiques jouent un rôle important dans le fonctionnement de réseaux complexes rencontrés en biologie moléculaire. Il est donc nécessaire d'approfondir la modélisation et les techniques d'analyse de la dynamique stochastique de ces réseaux. Il est aussi nécessaire de développer des outils informatiques permettant l'étude du comportement des réseaux biologiques à grande échelle tels que les réseaux de gènes et les réseaux de signalisation.

La manière classique de modéliser le bruit dans des systèmes de réactions chimiques a été introduite par (Delbrück, 1940) et a été formalisée plus tard par (Bartholomay, 1957; Rényi, 1954). Dans ces modèles, l'évolution d'un système chimique ou biochimique est markovienne et consiste à produire différents types de réactions chimiques aléatoires, avec des intensités qui dépendent du type de réaction et de l'état du système. L'ensemble des types des réactions définit le mécanisme de réaction. L'état du système est un vecteur aléatoire dont les composantes sont des nombres entiers, représentant le nombre de molécules des différentes espèces. Les fluctuations relatives des nombres de molécules sont en général plus importantes pour les espèces de faible concentration.

Cette représentation markovienne des systèmes de réactions chimiques a été appliquée dans un algorithme de simulation connu sous le nom de la méthode de Gillespie (Gillespie, 1976). L'algorithme de Gillespie simule un processus de Markov à sauts dont l'état est le vecteur des nombres de molécules des différentes espèces et dont les sauts sont représentés par les vecteurs stoechiométriques des différentes réactions. Lorsque les nombres de molécules sont importants, l'effet d'une seule réaction est négligeable. Il faut beaucoup de réactions pour changer l'état du système. Ce phénomène conduit à des temps d'exécution importants. Au lieu de générer les réactions les unes après les autres, séparées par des intervalles de temps aléatoires, Gillespie (Gillespie, 2001) propose des algorithmes accélérés, dans lesquels plusieurs réactions sont générées en une seule étape avec une discrétisation déterministe du temps. Le choix du pas de discrétisation peut être difficile.

Les trajectoires des processus de Markov à sauts convergent vers des trajectoires déterministes, solutions d'équations différentielles, lorsque le nombre de chaque type de réaction par unité de temps est important (Kurtz, 1970; Kurtz, 1971). Dans cette limite, appelée limite fluide par analogie avec le même concept en recherche opérationnelle, les phénomènes stochastiques disparaissent. Il s'agit de la loi des grands nombres pour les processus de Markov (Darling, 2002; Darling *et al.*, 2005). Dans les mêmes conditions, il existe aussi un théorème central limite qui permet l'approximation des processus de Markov par des diffusions (Kurtz, 1971; Allain, 1976; Gillespie, 2000).

Lorsqu'une ou plusieurs espèces moléculaires ont des faibles concentrations tandis que toutes les autres espèces ont des fortes concentrations, on ne peut plus appliquer la limite fluide, ni l'approximation par diffusions. Le cadre décrivant cette situation est celui des processus de Markov déterministes par morceaux (Boxma *et al.*, 2005; Meyn *et al.*, 1993; Davis, 1993), ou celui des diffusions à sauts (Yin, 2001), ou celui des diffusions à plusieurs échelles de temps (Freidlin *et al.*, 1984). Les processus déterministes par morceaux sont particulièrement faciles à simuler et offrent une économie importante en temps de calcul par rapport à la méthode de Gillespie lorsque leur application est justifiée (Alfonsi *et al.*, 2004). De même, il existe de nombreux résultats (Meyn *et al.*, 1993; Davis, 1993) sur les propriétés des trajectoires qui peuvent s'appliquer pour expliquer les observations expérimentales. La section 2 définit les modèles de réactions chimiques ; les quatre sections suivantes présentent des résultats mathématiques sur la convergence des processus de Markov à sauts vers un processus déterministe, puis la vitesse de convergence, et enfin les processus déterministes par morceaux pour des processus de Markov ne vérifiant pas les conditions de convergence vers un processus déterministe. La section 8 indique quelques applications en biologie moléculaire.

2. Modélisation par des systèmes de réactions chimiques, processus de Markov à sauts

2.1. Processus de Markov à sauts, quelques rappels

A tout processus de Markov $X(t) \in E$, E étant un espace métrique, on associe sa filtration naturelle (son passé et présent) $\mathcal{F}_t^X = \sigma\{X(s), s \leq t\}$ et une fonction probabilité de transition $P : \mathbb{R} \times E \times \mathbb{R} \times \mathcal{B}(E) \rightarrow [0, 1]$ ($\mathcal{B}(E)$ est l'ensemble des boréliens), $P(s, x, t, A) = \mathbb{P}[X(t) \in A | \mathcal{F}_s^X] = \mathbb{P}[X(t) \in A | X(s) = x], t \geq s$, la dernière égalité étant la propriété de Markov (le futur d'un processus de Markov dépend du passé uniquement par le présent). P satisfait les propriétés suivantes :

- 1) $P(s, x, t, A)$ est mesurable en x et représente une mesure de probabilité en A .
- 2) $P(t, x, t, \cdot) = \delta_x$ (masse de Dirac en x).
- 3) $P(t_1, x, t_3, A) = \int P(t_1, x, t_2, dy) P(t_2, y, t_3, A)$ (équation de Chapman-Kolmogorov).

Les processus de Markov à sauts sont définis par l'existence de la limite suivante (Gikhman *et al.*, 1969) :

$$\lim_{t \searrow s} \frac{P(s, x, t, A) - \delta_x(A)}{t - s} = \lambda(s, x) \mu(s, x, A) \quad [1]$$

quel que soit $t \in \mathbb{R}, x \in E, A \in \mathcal{B}(E)$.

La fonction $\lambda : \mathbb{R} \times E \rightarrow \mathbb{R}_+$ est l'intensité (nombre moyen de sauts par unité de temps). $\mu(s, x, A)$ est une mesure de probabilité en A nommée loi des sauts. Elle représente la probabilité qu'on choisisse de sauter de x en A . Pour des processus homogènes en temps, l'intensité et la loi des sauts ne dépendent pas du temps s .

Lorsque E est séparable (ce qui est le cas de \mathbb{R}^n), tout processus de Markov à sauts est équivalent à un processus càd-làg (continu à droite et possédant une limite à gauche) et il existe presque sûrement une suite d'intervalles de temps aléatoires τ_i de lois exponentielles tels que $\lim_{n \rightarrow \infty} \sum_{i=1}^n \tau_i = \infty$ et $X(t)$ est constant sur $[\sum_{i=1}^n \tau_i, \sum_{i=1}^{n+1} \tau_i[$ (Gikhman *et al.*, 1969). Ainsi, un processus de Markov à sauts peut être défini d'une façon équivalente à partir d'un processus de comptage $\nu(t)$ (d'intensité λ) et d'une chaîne de Markov X_n de noyau de transition $\mu : X(t) = X_{\nu(t)}$.

Cette dernière description d'un processus de Markov à sauts est utilisée dans l'algorithme de simulation de Gillespie (Gillespie, 1976) :

- 1) $X = X_0, t = 0$.
- 2) Générer un intervalle de temps aléatoire τ_n de loi exponentielle de paramètre $\lambda(X, t)$.
- 3) Incrémenter le temps $t = t + \tau_n$.
- 4) Effectuer un saut $X \rightarrow Y$ où Y est généré avec la loi $\mu(X, t)$.
- 5) Si $t < t_{max}$, retour à l'étape 2.

Notons l'importance de la loi exponentielle des intervalles de temps dans la description markovienne. En effet, remplacer la loi exponentielle par une autre loi avec mémoire, conduirait à la perte du caractère markovien de la dynamique. Ceci est néanmoins possible dans des modèles semi-markoviens où les intervalles de temps peuvent être choisis quelconques (Korolyuk *et al.*, 1995).

L'intérêt du choix markovien réside aussi dans la simplicité et la puissance des techniques mathématiques disponibles pour ce type de processus. Ainsi à tout processus de Markov homogène on peut associer un semi-groupe d'évolution, c'est-à-dire un opérateur sur l'ensemble $B(E)$ des fonctions bornées sur E , $T(t) : B(E) \rightarrow B(E)$ défini par $T(t)f(x) = \mathbb{E}[f(X(t))|X(0) = x]$. Son générateur est défini par $\mathcal{A}f = \lim_{t \rightarrow 0} \frac{T(t)f - f}{t}$, pour toute fonction f dans le domaine $\mathcal{D}(\mathcal{A})$ où cette limite existe. La relation suivante va être utilisée par la suite (Ethier *et al.*, 1986) :

$$\frac{dT(t)f}{dt} = \mathcal{A}T(t)f = T(t)\mathcal{A}f \quad [2]$$

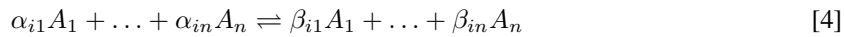
Un processus de Markov à sauts homogène peut être ainsi caractérisé par son générateur (Ethier *et al.*, 1986) :

$$\mathcal{A}f(x) = \lambda(x) \int_E (f(z) - f(x))\mu(x, dz) \quad [3]$$

2.2. Systèmes de réactions chimiques

Les systèmes de réactions chimiques sont modélisés par des processus de Markov à sauts homogènes (Rényi, 1954; Bartholomay, 1957). Lorsqu'il y a n espèces chimiques A_1, \dots, A_n l'espace d'états est $E = \mathbb{Z}^n$, un état étant un vecteur X indiquant les nombres de molécules de chaque espèce.

A chaque réaction chimique du type



on associe un saut $\theta_i = \beta_i - \alpha_i \in \mathbb{Z}^n$, $i = 1, n_r$ (n_r est le nombre de réactions). On considère également la réaction inverse ce qui correspond au saut $-\theta_i$. L'intensité et la loi des sauts sont :

$$\lambda(X) = \sum_{i=1}^{n_r} [V_i(X) + V_{-i}(X)] \quad [5]$$

$$\mu(X, \cdot) = \sum_{i=1}^{n_r} [q_i(X)\delta_{X+\theta_i}(\cdot) + q_{-i}(X)\delta_{X-\theta_i}(\cdot)] \quad [6]$$

V_i, V_{-i} sont les vitesses des réactions (nombre moyen de réactions par unité de temps) dans un sens et dans l'autre. Les probabilités $q_i = V_i / \sum_{j=1}^{n_r} (V_j + V_{-j})$ satisfont $\sum_{i=1}^{n_r} (q_{-i} + q_i) = 1$.

Plusieurs choix sont possibles pour les vitesses de réaction (Gorban *et al.*, 2000). Le choix courant est d'utiliser la loi d'action des masses. Selon cette loi, les vitesses de réaction sont des fonctions polynomiales de concentrations :

$$V_i(X) = \mathcal{V}v_i(x) = k_i \prod_{s=1}^n x_s^{\alpha_{is}}, \quad V_{-i}(X) = \mathcal{V}v_{-i}(x) = k_{-i} \prod_{s=1}^n x_s^{\beta_{is}} \quad [7]$$

où \mathcal{V} est le volume de réaction, qui en biologie moléculaire est de l'ordre du volume d'une cellule ou plus petit, $x = X/\mathcal{V}$ est le vecteur des concentrations et v_i sont les vitesses de réaction par unité de volume.

Ce choix de modélisation s'applique sous les conditions suivantes :

- diffusion rapide, système bien mélangé,
- chaos moléculaire (absence de corrélation entre les positions des molécules ou entre les réactions).

La validité de la première hypothèse dépend généralement de la comparaison entre une longueur de diffusion \mathcal{L}_D et l'échelle spatiale de l'observation expérimentale \mathcal{L}_O . Si D est le coefficient de diffusion spatiale d'une molécule et si τ est un temps caractéristique aux réactions bio-chimiques, la condition de diffusion rapide s'écrit :

$$\mathcal{L}_D = \sqrt{D\tau} \ll \mathcal{L}_O \quad [8]$$

Les valeurs du coefficient de diffusion intra-cellulaire dépendent de beaucoup de facteurs. Parmi ces facteurs, l'encombrement joue un rôle important (Takahashi *et al.*, 2005). Ce facteur peut limiter la diffusion à des valeurs aussi petites que $10^{-8} \text{cm}^2 \text{s}^{-1}$ (Elowitz *et al.*, 1999). Pour $D \sim 10^{-8} \text{cm}^2 \text{s}^{-1}$, $\tau = 1 \text{s}$, on a $\mathcal{L}_D = 10^{-4} \text{cm}$ ce qui est comparable aux dimensions de la cellule d'un procaryote. On peut donc considérer qu'une cellule de procaryote est bien mélangée sur des échelles de temps supérieures à 1s. Ce temps doit être multiplié par 100 pour les cellules eucaryotes, qui ont des diamètres 10 fois plus grands. Pour certains processus confinés dans des compartiments de taille beaucoup plus petite que celle de la cellule (noyau, mitochondries), ou pour les phénomènes de transport actif cette estimation est sans doute trop drastique. Reste l'importante catégorie des processus de transport cytoplasmique passif pour laquelle le temps de diffusion n'est pas du tout négligeable par rapport aux échelles de temps de secondes ou de dizaines de secondes. Le cas d'un syncytium est pire encore. Un syncytium est une région contenant un nombre important de noyaux non séparés par des membranes cellulaires (cette structure apparaît par exemple dans les premiers stades de développement de *Drosophila melanogaster* (Reinitz *et al.*, 1995)). Le syncytium de *Drosophila* d'une taille d'environ $5 \cdot 10^{-2} \text{cm}$ est beaucoup trop grand pour être considéré bien mélangé sur une échelle de temps, de signification physiologique quelconque. Dans de tels cas, le modélisateur peut décomposer l'espace en sous-volumes bien mélangés, et introduire la diffusion entre

sous-volumes. Cette solution est utilisée dans des algorithmes informatiques tels que SmartCell (Ander *et al.*, 2004), NSM (*Next Subvolume Method*) (Elf *et al.*, 2004).

Un traitement rigoureux du cas où l'hypothèse 1) n'est pas satisfaite n'est plus possible en gardant la représentation des états en termes de nombre de molécules de chaque espèce. Les molécules doivent être traitées individuellement et leurs positions spatiales doivent être prises en compte. La complexité d'une telle description est très grande, mais des approches mathématiques existent pour des cas particuliers (processus de fragmentation/coagulation (Aldous, 1999), équation de Boltzmann ou des hiérarchies BBGKY tronquées à bas ordre pour le traitement approché des corrélations entre molécules (Cercignani *et al.*, 1997; Kuzovkov *et al.*, 1988)).

La perte de la propriété de bon mélange conduit à des nouveaux phénomènes. A) Localisation : les produits peu mobiles ou avec un temps de vie relativement court (c'est le cas de l'ARN messenger, dont le temps de vie est court) restent localisés dans le voisinage de leurs sources. La localisation peut modifier les vitesses de réaction (en modifiant la loi d'action des masses) et conduire à des déplacements de l'équilibre (Togashi *et al.*, 2005; Scherb *et al.*, 2000). B) Contrôle des vitesses de réaction par la diffusion : lorsque la longueur de diffusion est plus petite que la distance moyenne entre molécules (condition plus forte que [8]) la cinétique est contrôlée par la diffusion. Des approches du type champ moyen permettent de retrouver la loi de l'action de masse avec des coefficients cinétiques variables dans le temps et dépendants du coefficient de diffusion (Balding *et al.*, 1989).

L'hypothèse du chaos moléculaire s'applique en théorie cinétique des gaz (Cercignani *et al.*, 1997), mais il n'est pas du tout évident qu'elle s'applique aussi aux milieux encombrés tels que l'intérieur de la cellule. Gillespie (Gillespie, 1976) a formulé une condition qui pourrait conduire au chaos moléculaire : le nombre de collisions non réactives doit être beaucoup plus important que le nombre de collisions réactives. Les corrélations entre molécules et entre réactions en cinétique chimique sont assez peu étudiées à l'exception de l'effet de volume exclu ; ce dernier effet diminue la compétition entre molécules semblables, diminue les fluctuations et accélère la cinétique des réactions binaires (Jung *et al.*, 1997). Un algorithme informatique qui peut rendre compte des corrélations entre réactions et effets spatiaux est StochSim (Morton-Firth, 1998). En utilisant une représentation individuelle de chaque molécule, cet algorithme est mieux adapté pour des simulations des milieux encombrés.

3. Limite déterministe thermodynamique, loi des grands nombres pour processus de Markov

La dynamique des processus de Markov à sauts est simplifiée lorsque les sauts sont petits et rapides. Il s'agit d'un passage à la limite dans lequel l'intensité des sauts converge vers l'infini, la taille des sauts converge vers zéro et le produit entre la taille et l'intensité reste constant. Si on imagine les sauts comme des quantités de substance, ce produit représente un débit. En théorie des files d'attente cette limite est

appelée limite fluide. En cinétique chimique elle est appelée limite thermodynamique (voir plus loin). Dans cette limite, le phénomène dynamique devient déterministe. Les techniques nécessaires pour prouver ces résultats sont essentiellement des inégalités de martingales (Kurtz, 1970; Kurtz, 1971; Darling, 2002; Darling *et al.*, 2005; Muller, 2003).

Les deux théorèmes suivants dus à (Kurtz, 1971) sont particulièrement importants :

Théorème 1 Soit $(X_n)_{n \in \mathbb{N}}$ une suite de processus de Markov à sauts homogènes, d'espaces d'états, intensités et lois de sauts $E_n \subset \mathbb{R}^k$, $\lambda_n(x)$, $\mu_n(x, dy)$.

Soit $F_n(x) = \lambda_n(x) \int_{E_n} (z - x) \mu_n(x, dz)$ une fonction "débit",

$Z_n(t) = X_n(t) - X_n(0) - \int_0^t F_n(X_n(u)) du$, l'accroissement compensé du processus $X(t)$.

Alors,

$$\mathbb{P}^x \left[\sup_{0 \leq s \leq t} |Z_n(s)| > \delta \right] \leq \frac{t}{\delta^2} \sup_{x \in E_n} \lambda_n(x) \int_{E_n} |z - x|^2 \mu_n(x, dz) \quad [9]$$

Théorème 2 Soit $(x_n)_{n \in \mathbb{N}}$ une suite de processus de Markov à sauts homogènes qui satisfait :

i) $\lim_{n \rightarrow \infty} \sup_{x \in E_n} |F_n(x) - F(x)| = 0$.

ii) $\lim_{n \rightarrow \infty} |x_n(0) - x_0| = 0$ presque sûrement.

iii) $\lim_{n \rightarrow \infty} \sup_{x \in E_n} D_n(x) = 0$, où $D_n(x) = \lambda_n(x) \int_{E_n} |z - x|^2 \mu_n(x, dz)$.

où $F : \mathbb{R}^k \rightarrow \mathbb{R}^k$ est une fonction uniformément lipschitzienne.

Soit $x(s)$ l'unique solution du système d'équations différentielles :

$$\frac{dx(s)}{ds} = F(x(s)), x(0) = x_0 \quad [10]$$

Alors,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\sup_{0 \leq s \leq t} |x_n(s) - x(s)| > \delta \right] = 0, \forall \delta > 0 \quad [11]$$

REMARQUE 1. – Le théorème 2 signifie que lorsque le moment d'ordre deux des sauts (qui représente le coefficient de diffusion «radial» dans l'espace des concentrations) converge vers zéro uniformément, le processus à sauts s'approche de la solution d'un système d'équations différentielles.

REMARQUE 2. – Dans la condition iii), \mathbb{R}^k pourrait être remplacé par un compact \mathcal{K} invariant.

Application aux systèmes de réactions chimiques

Pour les systèmes de réactions chimiques la limite fluide correspond à la limite thermodynamique. Considérons une suite (X_n) de processus de Markov à sauts correspondant à des systèmes de réactions chimiques. Les types de réactions chimiques sont les mêmes pour chaque système de la suite, mais le volume de réaction est de plus en plus grand, $\mathcal{V}_n \rightarrow \infty$. Considérons les processus des concentrations $x_n = X_n/\mathcal{V}_n$. Dans cette section nous allons re-numéroter les réactions de manière à éviter l'utilisation des indices \pm qui, pour les réactions réversibles, font la distinction entre les deux directions. Ainsi, chaque réaction réversible contribuera avec deux vecteurs de sauts de signes opposés. Selon la loi de l'action des masses les vitesses de réaction satisfont la loi d'échelle :

$$V_i(X_n) = \mathcal{V}_n v_i(x_n) \quad [12]$$

L'intensité, la loi de sauts, la fonction débit et le coefficient de diffusion radial du processus x_n sont respectivement :

$$\lambda_n(x) = \mathcal{V}_n \sum_{i=1}^{n_r} v_i(x) \quad [13]$$

$$\mu_n(x, dz) = \sum_{i=1}^{n_r} \delta_{x + \frac{\theta_i}{\mathcal{V}_n}}(dz) p_i(x), \text{ où } p_i(x) = v_i(x) / \left(\sum_{i=1}^{n_r} v_i(x) \right) \quad [14]$$

$$F_n(x) = F(x) = \sum_{i=1}^{n_r} v_i(x) \theta_i, \quad D_n(x) = \frac{1}{\mathcal{V}_n} \sum_{i=1}^{n_r} v_i(x) |\theta_i|^2 \quad [15]$$

Dans ce cas on a $\lim_{n \rightarrow \infty} D_n(x) = 0$. Dans la limite thermodynamique, la dynamique converge en probabilité vers la solution $x(t)$ du système d'équations différentielles [10].

4. Vitesse de convergence, théorème limite central

L'inégalité de martingales (théorème 1) borne supérieurement la vitesse de convergence, d'une façon analogue au théorème de Tchebychev pour des suites de variables aléatoires indépendantes :

$$\mathbb{P} \left[\sup_{0 \leq s \leq t} |x_n(s) - x(s)| > \delta \right] \leq \frac{Ct}{\mathcal{V}\delta^2} \sup_x \left[\sum_{i=1}^{n_r} |\theta_i|^2 v_i(x) \right], C > 0$$

Cette borne n'est pas optimale et peut être améliorée par des inégalités de concentration (Darling *et al.*, 2005) ou par le théorème limite central. Pour approfondir l'étude du comportement de la différence $x_n(s) - x(s)$, utilisons le théorème limite central suivant (Kurtz, 1971) :

Théorème 3 Soit une suite de réels $\alpha_n > 0, \alpha_n \rightarrow \infty$. Soit $X_n(t)$ une suite de processus de Markov à sauts partant de X_0 à $t = 0$. Définissons des fonctions "débit" $F_n(x)$ comme pour les théorèmes 1,2. De plus, définissons la "matrice de diffusion" suivante $G_n^{ij}(x) = \lambda_n(x) \int_{E_n} (z_i - x_i)(z_j - x_j) \mu_n(x, dz)$. Supposons que les conditions suivantes sont satisfaites :

- i) $\lim_{n \rightarrow \infty} \sup_{x \in E_n} |F_n(x) - F(x)| = 0$.
- ii) $\lim_{n \rightarrow \infty} \sup_{x \in E_n} \|\alpha_n^2 G_n(x) - G(x)\| = 0$.
- iii) il existe une suite $\epsilon_n \rightarrow 0$ telle que

$$\lim_{n \rightarrow \infty} \sup_{x \in E_n} \alpha_n^2 \lambda_n(x) \int_{|z-x| > \epsilon_n / \alpha_n} |z - x|^2 \mu_n(x, dz) = 0$$

Alors $W_n(t) = \alpha_n(x_n(t) - x(t))$ converge en loi vers le processus de Itô $V(t) = \int_0^t [G(X(s))]^{1/2} dB_s$, où $X(t)$ est la solution du système $\frac{dX}{dt} = F(X(t))$, $X(0) = X_0$

Application aux systèmes de réactions chimiques

Comme dans la section précédente, soit $x_n(t)$ la suite des processus des concentrations. Choisissons $\alpha_n = \sqrt{V_n}$. Les éléments de la matrice de diffusion pour le processus $x_n(t)$ sont :

$$G_n^{ij}(x) = \frac{1}{V} \sum_{k=1}^{n_r} v_k(x) \theta_k^i \theta_k^j \quad [16]$$

La condition ii) est satisfaite pour

$$G^{ij}(x) = \sum_{k=1}^{n_r} v_k(x) \theta_k^i \theta_k^j \quad [17]$$

De plus, la condition iii) est satisfaite également car lorsqu'on choisit $\epsilon_n = \frac{1}{\sqrt{V_n}} \sup_k |\theta_k|$:

$$\alpha_n^2 \lambda_n(x) \int_{|z-x| > \epsilon_n / \alpha_n} |z - x|^2 \mu_n(x, dz) = \sum_{|\theta_k| > \epsilon_n \sqrt{V_n}} v_k(x) |\theta_k|^2 = 0$$

REMARQUE 1. – Pour $n \rightarrow \infty$ le processus $x_n(t)$ converge vers la solution déterministe $u(t)$. Le théorème nous dit ce qu'il se passe à des échelles fines $1/\alpha_n = 1/\sqrt{V_n}$ dans l'espace des concentrations. A cette échelle, la diffusion n'est pas négligeable. Néanmoins, cette échelle devient arbitrairement petite, donc on peut considérer que la diffusion s'effectue dans un voisinage de la trajectoire déterministe suffisamment petit pour pouvoir négliger les termes de dérive. Pour un volume fini ou pour des temps

importants cette approximation peut être insuffisante. Gillespie (Gillespie, 2000) justifie par des arguments heuristiques une meilleure approximation de diffusion qui rend compte aussi de la dérive. Cette approximation est à la base de l'algorithme accéléré nommé τ -leap (Gillespie, 2001). Au lieu de simuler un nombre important d'intervalles aléatoires entre réactions qui se succèdent rapidement, l'algorithme τ -leap discrétise le temps de manière déterministe et simule une seule variable aléatoire normale (le nombre de réactions) par type de réaction et par intervalle de temps.

REMARQUE 2. – Ce théorème permet d'estimer les fluctuations par rapport à la trajectoire déterministe. En effet, en utilisant l'isométrie de Itô, on obtient :

$$\mathcal{V}_n \mathbb{E}[(x_n(t) - u(t))(x_n(t) - x(t))^t] \rightarrow \int_0^t G(u(s))ds \quad [18]$$

5. Limites singulières

Les théorèmes sur la limite déterministe thermodynamique ne sont valables qu'asymptotiquement, sous condition que les nombres de molécules de chaque espèce soient très grands et proportionnels au volume. En réalité, toute cellule a un volume fini et certaines espèces moléculaires peuvent se trouver en petits nombres à l'intérieur de la cellule. Deux situations typiques en cinétique chimique conduisent à des faibles nombres de molécules :

1) Existence de composants transitoires (radicaux, complexes enzymatiques, etc.) qui n'ont des concentrations significatives que pendant des courtes durées lors des déplacements d'équilibre.

2) Existence de lois de conservation avec coefficients positifs imposant des bornes supérieures pour les nombres de certaines molécules. Lorsque ces bornes sont basses, les molécules restent en faibles nombres pendant toute la durée de la cinétique.

Remarquons que les deux situations sont dynamiquement différentes :

– la première situation implique des temps transitoires courts ; nous verrons que dans la deuxième situation les temps transitoires sont remplacés par des basculements de régime lorsque les espèces rares subissent des réactions,

– la première situation est une situation typique à deux échelles de temps : une échelle rapide pendant laquelle les concentrations des espèces transitoires atteignent des niveaux faibles et une échelle lente pendant laquelle les autres espèces évoluent vers un équilibre. La deuxième situation n'implique pas nécessairement des échelles de temps séparées.

Les deux cas correspondent à la limite $\epsilon \rightarrow 0$, où ϵ représente la concentration des espèces rares. Bien que des limites de ce type aient été abordées par plusieurs auteurs (Haseltine *et al.*, 2002; Rao *et al.*, 2003; Burrage *et al.*, 2004; Plyasunov, 2005), à notre connaissance il n'existe pas une présentation unitaire de ce sujet. Nous allons donner ici quelques arguments heuristiques permettant de guider le lecteur.

Soit $X = (X^f, X^r)$ une décomposition du vecteur des nombres de molécules en espèces fréquentes et rares, respectivement. Considérons le processus de Markov suivant :

$$x_{\epsilon, \mathcal{V}}(t) = (x^f(t), x^r(t)) = \left(\frac{X^f(t)}{\mathcal{V}}, \frac{X^r(t)}{\epsilon \mathcal{V}} \right) \quad [19]$$

où \mathcal{V} est le volume.

Le générateur de ce processus est :

$$\mathcal{A}_{\epsilon, \mathcal{V}} f(x^f, x^r) = \sum_{i=1}^r [f(x^f + \frac{\theta_i^f}{\mathcal{V}}, x^r + \frac{\theta_i^r}{\epsilon \mathcal{V}}) - f(x^f, x^r)] V_i \quad [20]$$

Notons que $\epsilon \mathcal{V}$ représente l'ordre de grandeur du nombre de molécules rares. Alors, les deux situations citées plus haut correspondent à $\mathcal{V} \rightarrow \infty, \epsilon \rightarrow 0, \epsilon \mathcal{V} \rightarrow \infty$ et à $\mathcal{V} \rightarrow \infty, \epsilon \rightarrow 0, \epsilon \mathcal{V} \rightarrow 1$, respectivement.

Dans la première situation nous avons une limite thermodynamique déterministe, mais celle-ci conduit à une dynamique à deux échelles de temps. De façon heuristique, on emploie un développement de Taylor au deuxième ordre dans [20] et la loi d'action de masses, pour obtenir une approximation de diffusion du générateur du processus $x_{\epsilon, \mathcal{V}}(t)$:

$$\begin{aligned} \mathcal{A}_{\epsilon, \mathcal{V}}^D f(x^f, x^r) &= f_{x^f} \cdot F^f(x^f, x^r) + \frac{1}{2\mathcal{V}} f_{x^f, x^f} : G^f(x^f, x^r) \\ &+ \frac{1}{\epsilon} f_{x^r} \cdot F^r(x^f, x^r) + \frac{1}{2\epsilon^2 \mathcal{V}} f_{x^r, x^r} : G^r(x^f, x^r) \end{aligned} \quad [21]$$

F^f, F^r, G^f, G^r sont les composantes des fonctions vectorielle F et matricielle G définies par les équations [15],[17], $f_{x^f, r}$ sont des gradients, $f_{x^f, r, x^f, r}$ sont des matrices hessiennes, \cdot est le produit scalaire et $:$ est le produit tensoriel deux fois contracté.

Le générateur dans [21] correspond au système d'équations différentielles stochastiques :

$$\begin{aligned} dx^f &= F^f(x^f, x^r)dt + \frac{1}{\sqrt{\mathcal{V}}} \sigma^f(x^f, x^r)dB^f \\ dx^r &= \frac{1}{\epsilon} F^r(x^f, x^r)dt + \frac{1}{\epsilon \sqrt{\mathcal{V}}} \sigma^r(x^f, x^r)dB^r \end{aligned} \quad [22]$$

où $\sigma^{f,r}(\sigma^{f,r})^t = G^{f,r}$ et où $B^{f,r}$ sont des browniens indépendants.

La limite singulière $\epsilon \rightarrow 0$ dans le système [22] est caractérisée par le principe d'homogénéisation (Freidlin *et al.*, 1984; Skorokhod, 1989). Soit $y(t)$ le processus

satisfaisant l'équation $dy = F^r(x^f, y)dt + \sigma^r(x^f, y)dB$. Le processus x^f à lui seul converge vers un processus de Markov qui satisfait l'équation différentielle stochastique homogénéisée $dx^f = \bar{F}^f(x^f)dt + \frac{1}{\sqrt{\mathcal{V}}} \bar{\sigma}^f(x^f)dB^f$, où \bar{F}^f , $\bar{\sigma}^f$ sont obtenus à partir de F^f , σ^f en moyennant la dépendance en x^r sur une trajectoire de $y(t)$ (Freidlin *et al.*, 1984) ou sur la loi de y lorsque ceci est ergodique (Skorokhod, 1989). (Haseltine *et al.*, 2002; Rao *et al.*, 2003) montrent aussi des principes d'homogénéisation à l'aide de l'équation maîtresse (équation de Kolmogorov). Mentionnons aussi qu'un point de vue différent à l'utilisation du principe d'homogénéisation est possible et consiste à étudier la localisation des trajectoires dans un voisinage de la variété lente définie par l'équation $F^r(x^f, x^r) = 0$ (Berglund *et al.*, 2003).

Considérons maintenant la deuxième situation. Supposons que les vitesses de réactions satisfont :

$$\begin{aligned} V_i &= \tilde{V}_i\left(\frac{X^f}{\mathcal{V}}, \frac{X^r}{\epsilon\mathcal{V}}\right), \forall i \quad \text{t.q.} \quad \theta_i^r \neq 0 \\ V_i &= \mathcal{V}\tilde{v}_i\left(\frac{X^f}{\mathcal{V}}, \frac{X^r}{\epsilon\mathcal{V}}\right), \forall i \quad \text{t.q.} \quad \theta_i^r = 0 \end{aligned} \quad [23]$$

Ces conditions expriment le fait que les espèces rares n'obéissent pas à la loi d'action de masses. Dans la limite thermodynamique, les vitesses des réactions qui produisent ou qui consomment des espèces rares ne dépendent pas du volume.

Le même argument heuristique appliqué à [20] suggère que le processus de Markov $x_{\epsilon, \mathcal{V}}(t)$ converge vers un processus de générateur :

$$\begin{aligned} \mathcal{A}^{DM} f(x^f, X^r) &= f_{x^f} \cdot F^f(x^f, X^r) + \sum_{i=1}^{n_r} [f(x^f, X^r + \theta_i^r) - \\ &\quad f(x^f, X^r)] \tilde{V}_i(x^f, X^r) \end{aligned} \quad [24]$$

Ce processus est déterministe par morceaux. Pendant le temps entre deux réactions impliquant une molécule rare (par exemple apparition ou disparition d'une molécule rare), la dynamique est déterministe. Le changement produit par une telle réaction induit un changement de paramètres de la dynamique déterministe.

6. Processus déterministes par morceaux

Un modèle à deux états (présence et absence d'une molécule rare), utilisé en recherche opérationnelle, est particulièrement adapté pour décrire des situations variées en biologie moléculaire.

Soit $(X(t), \gamma(t))$ un processus de Markov à espace d'états $E = [0, \infty[\times \{0, 1\}$. $X(t)$ évolue selon une dynamique différentielle paramétrée par la variable discrète $\gamma(t)$:

$$\frac{dX}{dt} = \begin{cases} -r_0(X) & \text{si } \gamma(t) = 0 \\ r_1(X) & \text{si } \gamma(t) = 1 \end{cases} \quad [25]$$

$r_0(X) \geq 0$ est un terme de consommation tandis que $r_1(X) \geq 0$ est un terme de production. On suppose également qu'il n'y a pas consommation lorsque $X = 0$, $\gamma = 0$, *i.e.* $r_0(0) = 0$, et qu'il n'y a pas de production lorsque $X = 1$, $\gamma = 1$, *i.e.* $r_1(1) = 0$. Avec ces deux conditions on montre que si $X(0) \in [0, 1]$, alors $X(t) \in [0, 1]$, $t \geq 0$. De plus, γ saute de 0 à 1 avec l'intensité $\lambda_0(X)$ et de 1 à 0 avec l'intensité $\lambda_1(X)$. Pour que les états $(0, 0)$, $(1, 1)$ ne soient pas absorbants on impose $\lambda_1(0) > 0$ et $\lambda_0(1) > 0$.

Le générateur du processus $(X(t), \gamma(t))$ est :

$$\mathcal{A}f(x, 0) = -r_0(x)f'(x, 0) + \lambda_0(x)(f(x, 1) - f(x, 0)) \quad [26]$$

$$\mathcal{A}f(x, 1) = r_1(x)f'(x, 1) + \lambda_1(x)(f(x, 0) - f(x, 1)) \quad [27]$$

Soit π la loi invariante de (X_t, γ_t) . Une forme suffisamment générale de π correspond à :

$$\begin{aligned} \mathbb{E}_\pi f(X(t), \gamma(t)) = & p_0 f(0, 0) + p_1 f(1, 1) + c_0 \int_0^1 g_0(x) f(x, 0) dx + \\ & c_1 \int_0^1 g_1(x) f(x, 1) dx \end{aligned} \quad [28]$$

où $p_0, p_1, c_0, c_1 \in [0, 1]$, $p_0 + p_1 + c_0 + c_1 = 1$, $\int_0^1 g_i(x) dx = 1$.

La signification des fonctions et paramètres est la suivante :

– $p_i + c_i$ définit la loi marginale (discrète) de γ , $p_i + c_i = \mathbb{P}[\gamma = i]$. De plus, $p_i = \mathbb{P}[X = i, \gamma = i]$, $i = 0, 1$.

– $\frac{c_i}{p_i + c_i} g_i(x)$ est la densité de la partie continue de la loi conditionnelle de X , $\frac{c_i}{p_i + c_i} g_i(x) dx = \mathbb{P}[X(t) \in dx | \gamma(t) = i]$.

– la partie continue de la loi marginale de X a la densité $p(x) = c_0 g_0(x) + c_1 g_1(x)$.

En utilisant la propriété du générateur et l'invariance de π on obtient :

$$\mathbb{E}_\pi \mathcal{A}f(X(t), \gamma(t)) = \frac{d}{dt} \mathbb{E}_\pi f(X(t), \gamma(t)) = 0 \quad [29]$$

L'équation d'équilibre [29] est valable pour toute fonction f dans le domaine du générateur. En considérant des fonctions différentiables en x et qui satisfont $f(x, 1) = f(x, 0) = 0$ pour $x = 0, 1$, on obtient l'équation de Kolmogorov (rétrograde) :

$$\frac{d[c_0 r_0(x) g_0(x)]}{dx} - c_0 g_0(x) \lambda_0(x) + c_1 g_1(x) \lambda_1(x) = 0 \quad [30]$$

$$\frac{d[c_1 r_1(x) g_1(x)]}{dx} - c_0 g_0(x) \lambda_0(x) + c_1 g_1(x) \lambda_1(x) = 0 \quad [31]$$

De [30],[31] il suit que $c_0 r_0(x) g_0(x) = c_1 r_1(x) g_1(x) + \text{const.}$. En réalité, la constante est nulle car en appliquant [29] aux fonctions différentiables en x et qui satisfont $f(x, 1) = f(x, 0)$ pour tout x , on obtient :

$$c_0 r_0(x) g_0(x) = c_1 r_1(x) g_1(x) \quad [32]$$

L'équation [32] représente une condition de balance détaillée partielle en x : pour chaque x les flux entre les états $(x, 0)$ et $(x, 1)$ dans un sens et dans l'autre sont égaux. A noter que le processus $(X(t), \gamma(t))$ est entièrement réversible.

En adaptant un résultat de (Boxma *et al.*, 2005) qui utilise [30], [32] on obtient le théorème suivant :

Théorème 4 *Supposons que la loi invariante π existe et qu'elle a la forme indiquée. Si les intégrales $I_1^\epsilon = \lim_{x \rightarrow 0} \int_x^\epsilon [\frac{\lambda_1(u)}{r_1(u)} - \frac{\lambda_0(u)}{r_0(u)}] du = -\infty$ pour un $\epsilon > 0$ et $I_2^\epsilon = \lim_{x \rightarrow 1} \int_\epsilon^x [\frac{\lambda_1(u)}{r_1(u)} - \frac{\lambda_0(u)}{r_0(u)}] du = \infty$ pour un $\epsilon < 1$, alors $p_0 = p_1 = 0$ et $g_i(x) = \frac{1}{\alpha_i r_i(x)} \exp[-\Phi(x)]$, où le potentiel $\Phi(x)$ obéit à l'équation différentielle*

$$\frac{d\Phi}{dx} = \frac{\lambda_1(x)}{r_1(x)} - \frac{\lambda_0(x)}{r_0(x)}$$

$$\text{et } \alpha_i = \int_0^1 \frac{1}{r_i(x)} \exp[-\Phi(x)] dx, \quad c_i = \frac{\alpha_i}{\alpha_0 + \alpha_1}.$$

REMARQUE 1. – I_1^ϵ peut être fini lorsque $\tau_\epsilon = \int_x^\epsilon \frac{du}{r_0(u)}$ et $\lambda_0(0)$ sont finis. τ_ϵ est le temps nécessaire pour arriver en $X = 0$ en partant de $X = \epsilon$ lorsque $\gamma = 0$. $1/\lambda_0(0)$ est le temps moyen de séjour en $(0, 0)$. La loi invariante charge l'état $(0, 0)$ seulement si on peut y arriver en temps fini et si on ne quitte pas cet état immédiatement après y arriver. On ne rencontre pas une telle situation dans nos exemples car la relaxation de X vers 0 est exponentielle et le temps pour arriver dans l'état $(0, 0)$ est infini. Une remarque analogue s'applique à l'état $(1, 1)$.

7. Algorithmes hybrides, complexité

Pour l'algorithme de Gillespie, le temps de calcul est proportionnel au nombre de réactions à générer, donc il est proportionnel au paramètre \mathcal{V} (volume) du modèle.

Dans la limite déterministe, ou déterministe par morceaux, ce temps devient infini. Lorsque le processus qu'on étudie possède une limite déterministe par morceaux il est plus approprié d'utiliser un algorithme hybride. Dans un algorithme hybride, la dynamique des molécules rares est discrète et simulée selon la méthode de Gillespie. Entre deux réactions successives impliquant des espèces rares, la dynamique des concentrations d'espèces fréquentes est calculée par un solveur d'équations différentielles ordinaires. Pour le modèle à deux états de la section précédente l'algorithme hybride est le suivant :

- 1) $X = X_0, \gamma = \gamma_0, t = 0$.
- 2) Générer un intervalle de temps aléatoire τ_n de loi exponentielle de paramètre $\lambda_i(X)$ lorsque $\gamma = i, i = 0, 1$. Les fonctions $\lambda_i(X)$ résultent du passage à la limite $\mathcal{V} \rightarrow \infty$ ($\lambda_i = \tilde{V}_i$ dans [24]) et ne dépendent pas de \mathcal{V} .
- 3) Propager $X(t) \rightarrow X(t + \tau)$ avec un solveur déterministe (paramétré par γ). Changer γ de 0 en 1, ou de 1 en 0.
- 4) Incrémenter le temps $t = t + \tau_n$.
- 5) Si $t < t_{max}$, retour à l'étape 2.

L'utilisation d'approximations déterministes par morceaux permet un gain en temps de calcul. En effet, lorsque le paramètre \mathcal{V} augmente, le temps de calcul pour une simulation déterministe par morceaux ne diverge pas (car le nombre de sauts par unité de temps est fini et indépendant de \mathcal{V}) ; ce temps dépend seulement de la longueur temporelle de la trajectoire.

8. Exemples

Nous allons traiter deux exemples. Le premier est un modèle de régulation des états lytique et lysogénique du bactériophage λ , le deuxième est un modèle de haploinsuffisance.

8.1. Méthodes informatiques

Toutes les simulations qui illustrent cette section utilisent l'algorithme de Gillespie décrit dans la section 2.1. L'algorithme a été implémenté en Matlab version 7.1. Dans le cas du modèle de Cook, pour illustrer le gain en temps de calcul on a également implémenté l'algorithme hybride décrit dans la section 7. Pour les simulations effectuées en Matlab, l'algorithme hybride est 5 fois plus rapide que l'algorithme de Gillespie (il faut moins d'une seconde pour simuler une trajectoire par l'algorithme hybride et 5 secondes par l'algorithme de Gillespie).

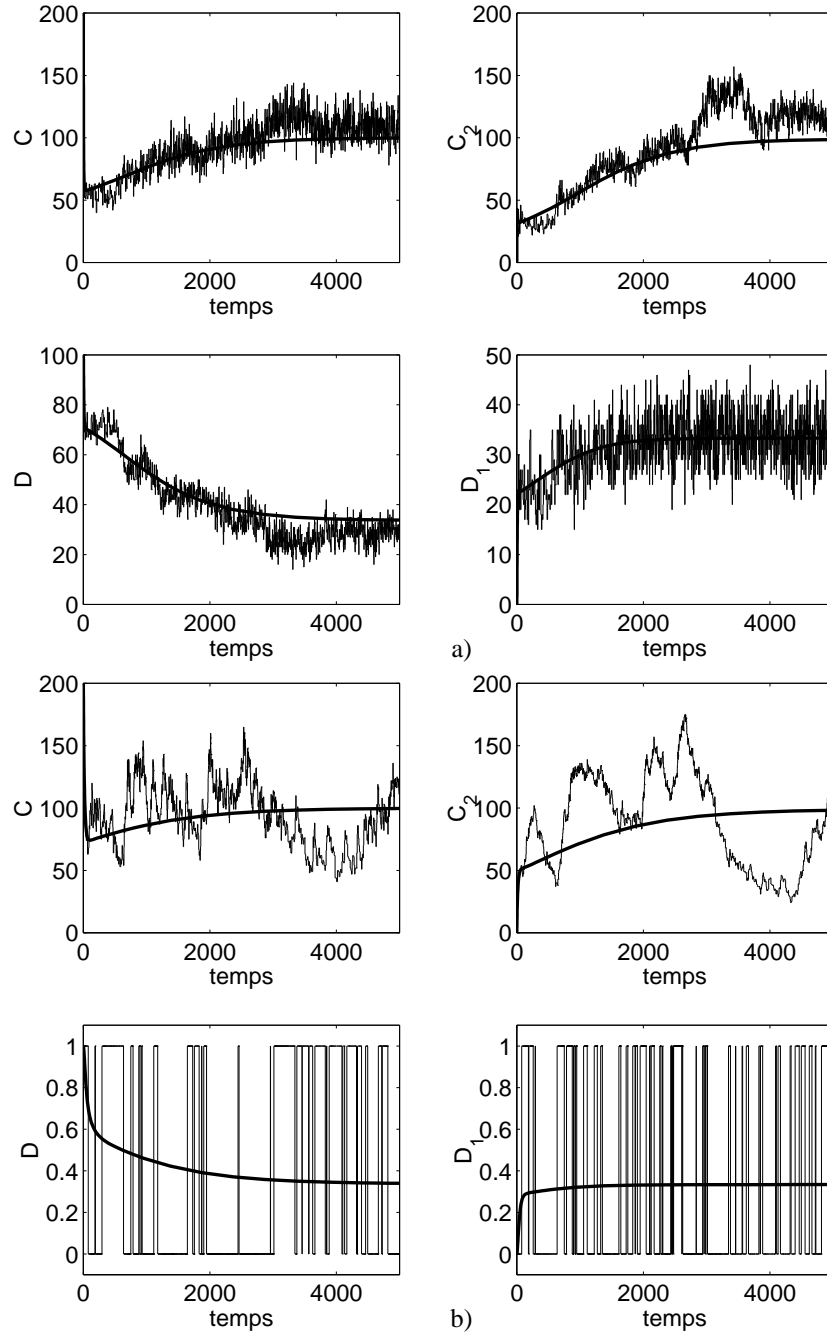


Figure 1. *Modèle de phage λ , la limite déterministe est marquée en trait continu. a) molécules et sites en grands nombres ; b) un seul site. Les paramètres sont a) $k_1^\pm = k_2^\pm = k_3^\pm = 0.1, k_4 = 0.006, k_5 = 0.01$ b) $k_1^\pm = k_3^\pm = 0.01, k_4 = 0.3, k_5 = 0.005$. Les trajectoires stochastiques sont représentées tous les a) 100 b) 10 sauts*

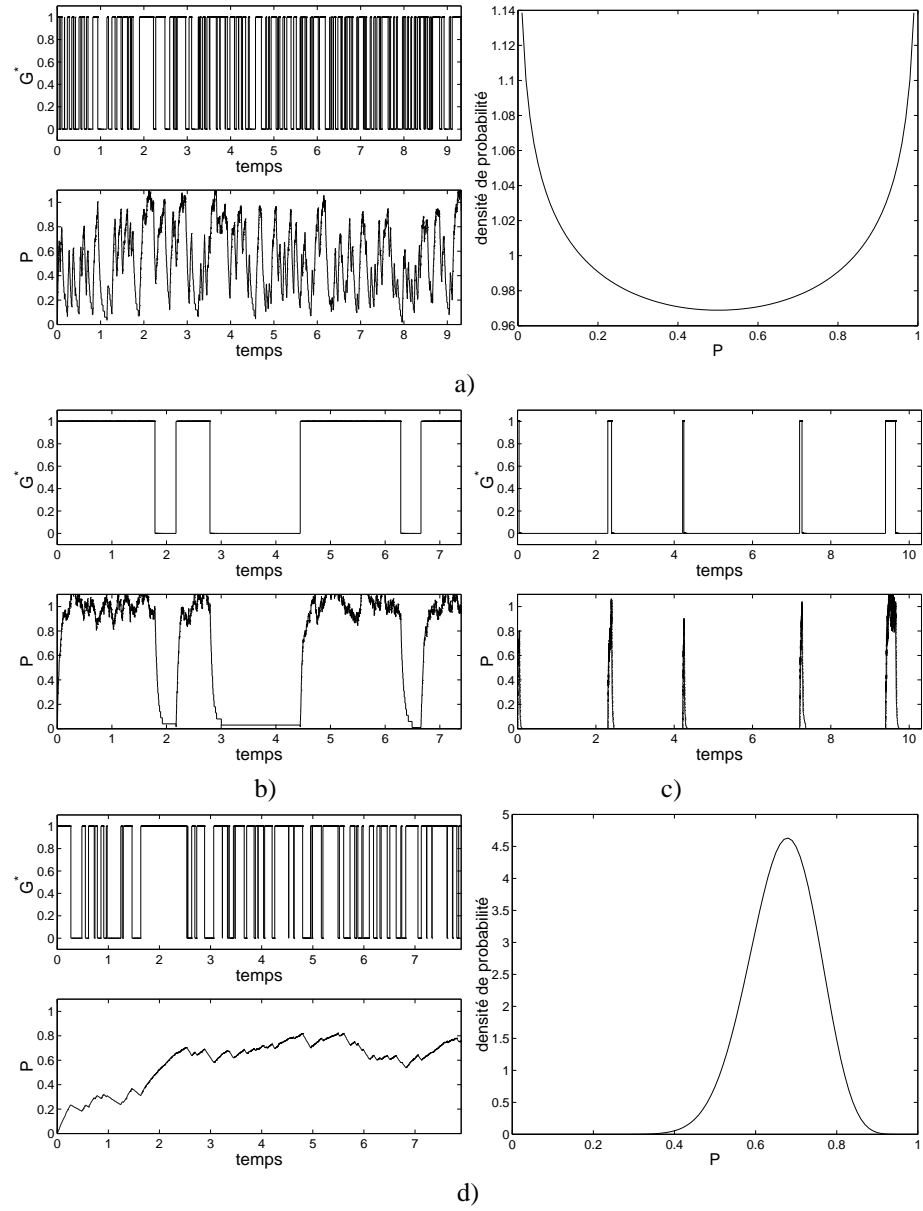
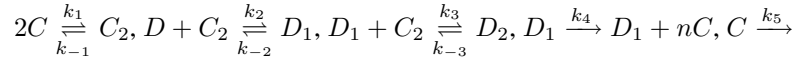


Figure 2. *Modèle de Cook. a) Fluctuations rapides, loi invariante bimodale* $k_1 = 19, k_{-1} = 19, k_3 = 20, k_2 = 4000$; *b) intermittence* $k_1 = 1, k_{-1} = 1, k_3 = 20, k_2 = 4000$; *c) pics stochastiques* $k_1 = 0.5, k_{-1} = 10 \frac{P^2}{0.01 + P^2}, k_3 = 50, k_2 = 4000$; *d) fluctuations rapides, loi invariante unimodale*, $k_1 = 20, k_{-1} = 10, k_3 = 1, k_2 = 4000$

8.2. Phage λ

Le bactériophage λ qui parasite E.Coli est l'exemple classique de système biomoléculaire bistable. Il possède deux états d'équilibre stable : l'état lysogénique quand le virus est dormant dans l'ADN de la bactérie, et l'état lytique quand le virus se multiplie de façon indépendante et finit par détruire son hôte. Le passage d'un état à l'autre s'effectue suite à un stress extérieur et est contrôlé par deux gènes : CI et CRO. CI est un inhibiteur de CRO qui induit l'état lytique. On présente ici une version légèrement simplifiée d'un modèle proposé par (Hasty *et al.*, 2000). Dans ce modèle, la bistabilité est entièrement assurée par CI qui s'autorégule. L'action de CI sur lui-même se fait par l'attachement des homodimères de CI successivement sur plusieurs sites dans la région promotrice ; peu de molécules accolées stimulent CI, beaucoup de molécules accolées répriment CI. Notons C le nombre de molécules de ARN de CI. Soient D, D_1, D_2 les nombres de régions promotrices non occupées, avec simple et avec double occupation respectivement. Seul l'état D_1 déclenche la transcription de CI. Le modèle est décrit par les réactions suivantes (dans les réactions, C, C_2, D, D_1, D_2 désignent respectivement les variables ARN de CI, dimère de ARN de CI, site non occupé, site en simple occupation et site en double occupation ; le nombre entier n est la taille du paquet de transcription/traduction) :



L'état du système est le vecteur $X = (C, C_2, D, D_1, D_2)$. Les vecteurs stoechiométriques correspondant aux cinq types de réactions sont $\theta_1 = (-2, 1, 0, 0, 0)$, $\theta_2 = (0, -1, -1, 1, 0)$, $\theta_3 = (0, -1, 0, -1, 1)$, $\theta_4 = (n, 0, 0, 0, 0)$, $\theta_5 = (-1, 0, 0, 0, 0)$. Il y a donc huit vecteurs de sauts $\pm\theta_1, \pm\theta_2, \pm\theta_3, \theta_4, \theta_5$ (les réactions réversibles dans un sens et dans un autre sont considérées indépendamment).

Notons la loi de conservation suivante :

$$D(t) + D_1(t) + D_2(t) = D_0 \quad [33]$$

8.2.1. Limite déterministe thermodynamique

L'évolution déterministe, dans la limite thermodynamique est décrite par [10] pour une fonction de débit $F(x) = (k_1 c^2 - k_{-1} c_2, k_2 d c_2 - k_{-2}, k_3 d_1 c_2 - k_{-3} d_2, k_4 d_1, k_5 c)$. Les états stationnaires sont les solutions de l'équation $F(x) = 0$. Il y a deux états stationnaires stables. Soit $\chi_i = k_i/k_{-i}, i = 1, \dots, 3$. Lorsque $\alpha = nk_4 d_0/k_5 = 3$ et $\chi_2 = \chi_3$ on a des expressions simples pour les deux attracteurs lytique $x_1 = (0, 0, d_0, 0, 0)$ et lysogénique $x_2 = A(1, \chi_1 A, d_0/3, d_0/3, d_0/3)$, où $A = 1/\sqrt{\chi_1 \chi_2}$.

Les fluctuations autour des trajectoires déterministes qui convergent vers l'attracteur lysogénique sont illustrées dans la figure 1a.

8.2.2. Échec de la limite déterministe

Pour obtenir la limite déterministe on doit supposer que toutes les molécules sont en grand nombre dans le volume de réaction. En réalité, cette condition n'est pas satisfaite. Effectivement, une fois que la bactérie est infectée, il n'y a pas de surinfection : $D_0 = 1$. Ceci signifie que la limite déterministe dans laquelle on remplace le nombre total de sites D_0 par une densité macroscopique de sites d_0 a uniquement un sens mathématique. L'occupation de sites est un phénomène essentiellement stochastique.

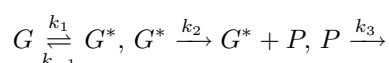
Dans la figure 1b nous avons simulé des trajectoires qui évoluent vers l'attracteur lysogénique pour $D_0 = 1$. On remarque le caractère discret des variables D, D_1, D_2 . Les processus déterministes par morceaux sont des bonnes approximations pour les trajectoires de C, C_2 . En effet, les molécules C, C_2 sont en grand nombre, on peut leur appliquer la limite fluide entre deux événements «site non occupé ou en double occupation» et «site en simple occupation». Ceci n'est pas le cas pour le nombre de sites D, D_1 qui ne prend que deux valeurs 0, 1 puisqu'il n'y a qu'un seul site qui peut être non occupé, en simple occupation ou en double occupation. L'évolution de (D, D_1) peut être décrite par un processus à sauts sur l'ensemble $\{0, 1\}^2$. Lorsque le site est non occupé ou en double occupation $D = 1, D_1 = 0$ ou $D = 0, D_1 = 0$, il n'y a pas de production de C qui tend à s'équilibrer avec son dimère (les trajectoires en projection sur la direction de C sont décroissantes). Lorsque le site est en simple occupation $D = 0, D_1 = 1$ il y a production de C (les trajectoires en projection sur la direction C sont croissantes).

Notons que le critère heuristique de la section 5 d'identification de la limite déterministe par morceaux s'applique ici. En effet, le petit paramètre du problème est $\epsilon = d_0 = 1/\mathcal{V}$. De plus lorsque $\epsilon \rightarrow 0$, $k_4 = \mathcal{O}(1/\epsilon) = \mathcal{O}(\mathcal{V})$ pour maintenir $\alpha = 3$. On peut alors aisément vérifier les lois d'échelle [23] conduisant lorsque $\epsilon \rightarrow 0$ vers un processus déterministe par morceaux.

En conclusion, dans le modèle de régulation des états lytique et lisogénique du bactériophage λ , les variables D et D_1 doivent être modélisées par un processus de sauts à deux états, et ne convergent pas vers une limite déterministe. Le vecteur (D, D_1, C, C_2) est markovien. De plus, pour des valeurs fixées de D, D_1 , les molécules C sont produites ou détruites en grands nombres, ce qui permet d'approcher l'évolution de C par un modèle déterministe par morceaux. Il en est de même pour C_2 . Le caractère déterministe par morceaux de la dynamique des variables C, C_2 ne signifie pas que les fluctuations de ces variables sont absentes. Au contraire, ces variables peuvent être soumises à des fluctuations importantes (figure 1b). Cette possibilité conduit à une conclusion biologique : la source des fluctuations n'est pas nécessairement la petitesse du nombre moyen de molécules observées, mais pourrait être la petitesse du nombre d'autres molécules non observées, autrement dit cachées (dans cet exemple l'occupation de sites).

8.3. Modèle de Cook

Le modèle introduit par (Cook *et al.*, 1998) pour modéliser les phénomènes de haploinsuffisance (invalidité de la moitié du nombre total de copies d'un gène) peut être décrit par le système de réactions suivant :



Dans le modèle de Cook G, G^* sont des versions dormante et active d'un gène, P est la protéine traduite.

Ce système de réactions conserve la quantité $G + G^* = G_0$. Le régime de haploinsuffisance est défini par une faible valeur G_0 . Pour simplifier considérons $G_0 = 1$, ce qui signifie une seule copie valide du gène. Dans cette situation $G, G^* \in \{0, 1\}$. Plus généralement, le modèle de Cook pourrait décrire d'autres cas d'activité intermittente en biologie moléculaire. Par exemple, dans le fonctionnement de voies de signalisation, G, G^* peuvent être considérées comme des variables cachées à deux valeurs paramétrisant la dynamique d'un système moléculaire dans deux situations (présence et absence d'une « molécule-clé »). Des extensions sont possibles à des variables cachées à plusieurs valeurs discrètes (le comportement du système pourrait dépendre de la présence et de l'absence de certaines molécules).

Les vitesses de réaction dans ce modèle sont $(V_1, V_{-1}, V_2, V_3) = (k_1 G, k_{-1} G^*, k_2 G^*, k_3 P)$. Le petit paramètre du problème est $\epsilon = 1/\mathcal{V}$. Pour obtenir la limite déterministe par morceaux il faut considérer que $k_2 = \mathcal{O}(\mathcal{V})$ (dans ce cas on peut vérifier [23]). Cette condition signifie une grande efficacité de la production. Dès que le gène est actif, la protéine P est produite en grande quantité et sa concentration tend vers une valeur d'équilibre $P_{eq} = k_2/k_3$. Cette concentration est atteinte si G reste longtemps actif, mais elle n'est pas nécessairement atteinte lorsque G retourne immédiatement à l'état inactif et reste longtemps inactif. Puisque le nombre de molécules P est grand, on peut considérer que la dynamique de P est déterministe pour une valeur de G^* fixée. On arrive au modèle déterministe par morceaux suivant :

$$\frac{dP}{dt} = -k_3 P + k_2 G^*(t)$$

où $G^*(t)$ est un processus de Markov à espace d'états $E = \{0, 1\}$ et fonction d'intensité :

$$\lambda(G^*) = \begin{cases} k_1 & \text{si } G^* = 0 \\ k_{-1} & \text{si } G^* = 1 \end{cases}$$

La loi des sauts est triviale, consistant à des passages avec probabilité un de $G^* = 1$ vers $G^* = 0$ et vice-versa. Notons que dans ce cas simple, non seulement le couple (G^*, P) mais aussi G^* est markovien.

On peut maintenant appliquer le théorème 4 avec :

$$X = P/(k_2/k_3), \quad \gamma = G^*, \quad r_0(x) = k_3 x, \quad r_1(x) = k_3(1-x), \quad \lambda_1 = k_{-1}, \quad \lambda_0 = k_1$$

Dans ce cas, le potentiel est

$$\phi(x) = -\ln x^{k_1/k_3} (1-x)^{k_{-1}/k_3}$$

et les fonctions densité sont

$$c_0 g_0(x) = \frac{1}{B(k_1/k_3, k_{-1}/k_3)} x^{k_1/k_3-1} (1-x)^{k_{-1}/k_3}$$

$$c_1 g_1(x) = \frac{1}{B(k_1/k_3, k_{-1}/k_3)} x^{k_1/k_3} (1-x)^{k_{-1}/k_3-1}$$

où B est la fonction beta.

On peut alors calculer la loi de X , qui est la loi beta de paramètres $r = k_1/k_3$, $s = k_{-1}/k_3$ et de densité :

$$p(x) = \frac{1}{B(r, s)} x^{r-1} (1-x)^{s-1}$$

Ce résultat est obtenu aussi par (Karmakar *et al.*, 2004) par une méthode qui consiste à intégrer l'équation de Kolmogorov rétrograde.

Rappelons que pour tout processus de Markov qui possède une loi invariante, la densité de cette loi dans un état quelconque est proportionnelle à la fraction du temps total passé dans cet état. Analysons les situations possibles selon la forme de la fonction densité de la loi invariante.

C1) Régime bimodal : $r = k_1/k_3 < 1$, $s = k_{-1}/k_3 < 1$

Pour $r, s < 1$, le potentiel ϕ a deux minima locaux et la densité de la loi beta a deux maxima locaux en $x = 0$ et $x = 1$. En conséquence les trajectoires passent beaucoup de temps dans les états $P = 0$ et $P = k_2/k_3$. Ce type de régime élargit la notion de multi-stationnarité introduite pour des systèmes biologiques déterministes. En effet, ce phénomène purement aléatoire est limité à des faibles valeurs de G_0 . Il ne subsisterait pas dans le régime déterministe (grandes valeurs de G_0), car tout système linéaire déterministe possède au plus un équilibre stable.

La loi invariante ne dit rien sur la façon dont les longueurs des intervalles de séjour sont distribuées. Les passages dans un sens et dans l'autre entre $P = 0$ et $P = k_2/k_3$ peuvent s'enchaîner rapidement ou au contraire être très rares. Évidemment ce sont les constantes k_1, k_{-1} qui contrôlent le temps de commutation. De même que Boxma (Boxma *et al.*, 2005), pour quantifier le temps de commutation définissons $\tau_r(x)$ comme le temps entre deux passages consécutifs au niveau $P = x$, le premier passage s'effectuant dans le sens de P croissant, et le deuxième dans le sens de P décroissant. Notons que le temps de retour ainsi défini n'est pas nécessairement égal au temps de retour dans l'état $G^* = 0$ après être passé par l'état $G^* = 1$. En effet, on peut activer et désactiver rapidement le gène G sans avoir un effet important sur P (voir figure 2b).

On peut adapter aisément les estimations de (Boxma *et al.*, 2005) pour obtenir :

$$\mathbb{E}[\tau_r(x)] = \exp[\phi(x)](\alpha_0 + \alpha_1) \int_x^1 p(y) dy \quad [34]$$

Pour le modèle de Cook ceci implique, lorsque $x \rightarrow 0$:

$$\mathbb{E}[\tau_r(x)] = \frac{B(k_1/k_3, k_{-1}/k_3)}{k_3} x^{-k_1/k_3} (1 + \mathcal{O}(x)) \quad [35]$$

Il est intéressant de remarquer le comportement du temps de retour avec k_1 . On a $\mathbb{E}[\tau_r(x)] \sim 1/k_1$ si k_1 est faible (temps de retour contrôlé par le basculement de

G, G^*) et $\mathbb{E}[\tau_r(x)] \sim \exp(A^{k_1})$, $A > 0$ si k_1 est grand (le temps de retour diverge avec k_1 car dans cette limite on revient au régime déterministe).

En fonction des valeurs du temps de retour et des valeurs de la densité de la loi invariante près des états $P = k_2/k_3$ et $P = 0$ on peut avoir plusieurs sous-régimes :

Oscillations rapides : $k_1/k_3 < 1, k_{-1}/k_3 < 1$, mais k_1, k_{-1} restent grands, figure 2a.

Dans ce régime le temps de retour est petit. Les trajectoires passent globalement un temps grand près des états $P = 0$ et $P = k_2/k_3$ mais basculent rapidement entre ces deux états.

Intermittence : $k_1, k_{-1} \ll k_3$, figure 2b.

Dans ce régime (bimodal) le temps de retour est grand : les trajectoires basculent rarement. Si le temps de retour est plus grand que le temps d'observation ou plus grand que le temps de vie d'une cellule, l'observation expérimentale du phénomène est statique plutôt que dynamique. Dans la littérature, on appelle cela bruit extrinsèque (Swain *et al.*, 2002). La distribution des valeurs de P dans une population de cellules étant bimodale, on observe alors un phénotype fortement différent d'une cellule à une autre. Ceci pourrait être une caractéristique de la haploinsuffisance ((Cook *et al.*, 1998) associe l'hétérogénéité pigmentaire de la peau ou des cheveux à un phénomène de haploinsuffisance dans les mélanocytes).

Pics stochastiques : $k_1 \ll k_3$, mais $k_{-1} = k_\infty \frac{P^n}{k + P^n}$ est variable, figure 2c. L'utilisation d'une fonction de Hill pour décrire la dépendance de k_{-1} en P , traduit l'auto-inhibition de la protéine P au dessus d'un seuil. Ce cas est caractérisé par un court temps de séjour en état actif $G^* = 1$. Son étude analytique est possible, par le théorème 4 et l'équation [34]. La production de protéine comporte des pics, séparée par des intervalles aléatoires. La ressemblance avec l'activité pulsée de certains facteurs de transcription (par exemple le facteur suppresseur de tumeurs P53 (Lahav *et al.*, 2004)) est remarquable : le modèle stochastique est une alternative intéressante aux modèles déterministes proposés pour rendre compte de ce phénomène (Bar-Or *et al.*, 2000). Néanmoins, le caractère markovien de notre description implique une distribution exponentielle (sans mémoire) des intervalles entre pics. Une telle distribution n'exclut pas des pics très rapprochés l'un de l'autre, ce qui semble contraster aux observations. Pour borner inférieurement les intervalles entre les pics il faut sortir du cadre markovien, par exemple en incluant après chaque pic une période réfractaire pendant laquelle il n'y a pas de transcription.

C2) Régime unimodal : $r = k_1/k_3 > 1, s = k_{-1}/k_3 > 1$, figure 2d.

Lorsque $r, s > 1$ la loi invariante est unimodale, on retrouve l'unicité de l'équilibre déterministe. Le potentiel ϕ possède un unique minimum et la loi beta est unimodale. Les trajectoires fluctuent autour d'un niveau d'équilibre stochastique, mais l'amplitude des fluctuations décroît avec les valeurs de r, s . Les valeurs de k_1, k_{-1} étant grandes, les sauts sont rapides. La concentration de P à l'équilibre est contrôlée par le rapport r/s (l'espérance d'une variable aléatoire de loi beta vaut $r/(r+s)$). La limite $r, s \rightarrow \infty, r/s \rightarrow C$ est de nouveau une limite déterministe. Dans cette limite la loi beta peut être approchée par une loi normale dont la variance est $\frac{C}{s(1+C)^3}$.

9. Conclusions

Des théorèmes limites permettent d'approcher les processus de Markov à sauts utilisés dans la modélisation des phénomènes moléculaires en génétique par des processus déterministes ou déterministes par morceaux.

De plus, on peut estimer les fluctuations autour des trajectoires déterministes, à l'aide d'un théorème central limite.

En pratique, lorsque le nombre de molécules est important, les systèmes de réactions chimiques peuvent être approchés par des processus déterministes.

Dans le cas où le nombre de certaines espèces moléculaires est faible, deux types de dynamiques sont les plus représentatifs. Le premier type est la dynamique à deux échelles de temps ; les échelles de temps rapides peuvent être homogénéisées. Le deuxième type est la dynamique déterministe par morceaux que nous avons choisie pour une étude plus détaillée.

Dans la dynamique déterministe par morceaux les molécules en faible nombre sont modélisées par des chaînes de Markov à petit nombre d'états. Des modèles déterministes par morceaux, même très simples, couvrent une large gamme des comportements. Nous avons classé ces comportements selon la loi invariante du processus et selon le temps de retour. Dans le cas du modèle de Cook nous avons pu mettre en évidence un régime bimodal avec oscillations rapides, intermittence, ou pics stochastiques, et un régime unimodal dans lequel on retrouve la limite déterministe. Nous avons présenté des résultats dans le cas de variables à deux états. Le cas à plus de deux états pourrait être également important, car il permet de considérer des processus non réversibles qui ne respectent pas la balance détaillée.

Ces résultats peuvent être utilisés pour comprendre et guider les observations expérimentales.

Remerciements

Ce travail a été soutenu par le programme ACI IMPBio. Nous remercions D. Ropers, C. Soulé, ainsi que les rapporteurs anonymes pour leur commentaires très utiles.

10. Bibliographie

- Aldous D., « Deterministic and Stochastic Models for Coalescence (Aggregation, Coagulation) : A Review of the Mean-Field Theory for Probabilists », *Bernoulli*, vol. 5, 1999, p. 3-48.
- Alfonsi A., Cancès E., Turinici G., Ventura B. D., Huisinga W., Exact simulation of hybrid stochastic and deterministic models for biochemical systems, Rapport de recherche n° 5435, INRIA, décembre, 2004.

- Allain M., « Approximation par un processus de diffusion, des oscillations, autour d'une valeur moyenne, d'un processus de Markov de saut pur », *C.R.Acad.Sc.Paris*, vol. t.282, 1976, p. 891-894.
- Ander M., Beltrao P., Ventura B. D., Ferkinghoff-Borg J., Foglierini M., Kaplan A., Lemerle C., Tomas-Oliveira I., Serrano L., « SmartCell, a framework to simulate cellular processes that combine stochastic approximation with diffusion and localisation : analysis of simple networks », *Syst.Biol.*, vol. 1, 2004, p. 129-138.
- Arkin A., Ross J., McAdams H. H., « Stochastic Kinetic Analysis of Developmental Pathway Bifurcation in Phage λ -infected *Escherichia Coli* Cells », *Genetics*, vol. 149, 1998, p. 1633-1648.
- Balding D., Green N., « Diffusion-controlled reactions in one dimension : Exact solutions and deterministic approximations », *Phys.Rev.A*, vol. 40, 1989, p. 4585-4592.
- Bar-Or R. L., Maya R., Segel L. A., Alon U., Levine A. J., Oren M., « Generation of oscillations by the p53-Mdm2 feedback loop : A theoretical and experimental study », *Proc.Natl.Acad.Sci. USA*, vol. 97, 2000, p. 11250-11255.
- Bartholomay A., *A Stochastic Approach to Chemical Reaction Kinetics*, Phd thesis, Harvard University, 1957.
- Berglund N., Gentz B., « Geometric singular perturbation theory for stochastic differential equations », *J. Differential Equations*, vol. 191, 2003, p. 1-54.
- Blake W., Kaern M., Cantor C., Collins J., « Noise in eukaryotic gene expression », *Nature*, vol. 422, 2003, p. 633-637.
- Boxma O., Kaspi H., Kella O., Perry D., « On/Off Storage Systems with State-Dependent Input, Output and Switching Rates », *Probability in the Engineering and Informational Sciences*, vol. 19, 2005, p. 1-14.
- Burrage K., Tian T., Burrage P., « A multi-scale approach for simulating chemical reaction systems », *Progr.Biophys.Mol.Bio.*, vol. 85, 2004, p. 217-234.
- Cercignani C., Gerasimenko V., Petrina D., *Many-Particle Dynamics and Kinetic Equations*, Kluwer, Dordrecht, 1997.
- Cook D. L., Gerber A. N., Tapscott S. J., « Modeling stochastic gene expression : Implications for haploinsufficiency », *Proc. Natl. Acad. Sci. USA*, vol. 95, 1998, p. 15641-15646.
- Darling R., « Fluid Limits of Pure Jump Markov Processes : A Practical Guide », *manuscript*.
- Darling R., Norris J., « Structure of large random hypergraphs », *Ann.Appl.Prob.*, vol. 15, 2005, p. 125-152.
- Davis M., *Markov Models and Optimization*, Chapman and Hall, London, 1993.
- Delbrück M., « Statistical Fluctuations in Autocatalytic Reactions », *J.Chem.Phys.*, vol. 8, 1940, p. 120-124.
- Elf J., Ehrenberg M., « Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases », *Syst.Biol.*, vol. 1, 2004, p. 230-236.
- Elowitz M., Surette M., Wolf P.-E., Stock J., Leibler S., « Protein Mobility in the Cytoplasm of *Escherichia Coli* », *J. Bacteriology*, vol. 181, 1999, p. 197-203.
- Ethier S. N., Kurtz T. G., *Markov Processes*, John Wiley & sons, New York, 1986.
- Freidlin M., Wentzell A., *Random perturbations of dynamical systems*, Spinger, New York, 1984.

- Gikhman I., Skorokhod A., *Introduction to the Theory of Random Processes*, W.B.Saunders Co., Philadelphia, 1969.
- Gillespie D. T., *J.Comput.Phys.*, vol. 22, 1976, p. 403.
- Gillespie D. T., « The Chemical Langevin equation », *J.Chem.Phys.*, vol. 113, 2000, p. 297-306.
- Gillespie D. T., « Approximate accelerated stochastic simulation of chemically reacting systems », *J.Chem.Phys.*, vol. 115, 2001, p. 1716-1733.
- Gorban A., Karlin I., Zmievskii V., Dymova S., « Reduced description in the reaction kinetics », *Physica A*, vol. 275, 2000, p. 361-379.
- Haseltine E. L., Rawlings J. B., « Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics », *J.Chem.Phys.*, vol. 117, 2002, p. 6959-6969.
- Hasty J., Pradines J., Dolnik M., Collins J., « Noise-based switches and amplifiers for gene expression », *Proc. Natl. Acad. Sci. USA*, vol. 97, 2000, p. 2075-2080.
- Iwasa Y., Michor F., Nowak M. A., « Stochastic tunnels in evolutionary dynamics », *Genetics*, vol. 166, 2004, p. 1571-1579.
- Jung Y., Lee S., « Excluded Volume Effect on the Diffusion-Influenced Bimolecular Reactions », *J. Phys.Chem.*, vol. 101, 1997, p. 5255-5261.
- Kaern M., A.Elston T., Blake W. J., Collins J. J., « Stochasticity in gene expression : from theories to phenotypes », *Nature Rev.Genet.*, vol. 6, 2005, p. 451-464.
- Karmakar R., Bose I., « Graded and binary responses in stochastic gene expressions », *arXiv :q-bio :OT/0411012v1*.
- Korolyuk V., Swishchuk A., *Semi-Markov Random Evolutions*, Kluwer, Dordrecht, 1995.
- Kurtz T. G., « Solutions of ordinary differential equations as limits of pure jump Markov process », *J.Appl.Prob.*, vol. 7, 1970, p. 49-58.
- Kurtz T. G., « Limit theorems for sequences of jump Markov processes approximating ordinary differential processes », *J.Appl.Prob.*, vol. 8, 1971, p. 344-356.
- Kuzovkov V., Kotomin E., « Kinetics of bimolecular reactions in condensed media », *Rep.Prog.Phys.*, vol. 51, 1988, p. 1479-1524.
- Laforge B., Guez D., Martinez M., Kuipiec J.-J., « Modeling embryogenesis and cancer : and approach based on the equilibrium between the autostabilization of stochastic gene expression and the interdependence of cells for proliferation », *Prog.Biophys.Mol.Biol.*, vol. 89, 2005, p. 93-120.
- Lahav G., Rosenfeld N., Sigal A., Geva-Zatorsky N., Levine A. J., Elowitz M. B., Alon U., « Dynamics of the p53-Mdm2 feedback loop in individual cells », *Nature Gen.*, vol. 36, 2004, p. 147-150.
- Meyn S. P., Tweedie R. L., *Markov Chains and Stochastic Stability*, Springer-Verlag, New York, 1993.
- Morton-Firth C., *Stochastic simulation of cell signalling pathways*, Phd, Cambridge, 1998.
- Muller A., *Modélisation stochastique des réseaux de gènes*, Rapport de stage, Université de Rennes 1, septembre, 2003.
- Ozbudak E., Thattai M., Kurtser I., Grossman A., van Oudenaarden A., « Regulation of noise in the expression of a single gene », *Nature Genet.*, vol. 31, 2002, p. 69-73.
- Plyasunov S., « On hybrid simulation schemes for stochastic reaction dynamics », *arXiv :math :ST/0504477v1*.

- Rao C. V., Arkin A. P., « Stochastic chemical kinetics and the quasi-steady-state assumption : Application to the Gillespie algorithm », *J.Chem.Phys.*, vol. 118, 2003, p. 4999-5010.
- Rao C. V., Wolf D. M., Arkin A. P., « Control, exploitation and tolerance of intracellular noise », *Nature*, vol. 420, 2002, p. 231-237.
- Reinitz J., Mjolsness E., Sharp D., « Model for cooperative control of positional information in *Drosophila* by bicoid and maternal hunchback », *J. Exp. Zool.*, vol. 271, 1995, p. 47-56.
- Rényi A., « Betrachtung chemischer Reaktionen mit Hilfe der Theorie der stochastischen Prozesse », *Magyar Tud.Akad.Akalm.Mat.Int.Közl.*, vol. 2, 1954, p. 93-101.
- Scherb N. M., Louzoun Y., Bettelheim E., Solomon S., « The importance of being discrete : Life always wins on the surface », *Proc.Natl.Acad.Sci.USA*, vol. 97, 2000, p. 10322-10324.
- Skorokhod A., *Asymptotic methods in the theory of stochastic differential equations*, AMS, Providence, 1989.
- Swain P., Elowitz M., Siggia E., « Intrinsic and extrinsic contributions to stochasticity in gene expression », *Proc.Natl.Acad.Sci.USA*, vol. 99, 2002, p. 12795-12800.
- Takahashi K., Arjunan S., Tomita M., « Space in systems biology of signalling pathways - towards intracellular molecular crowding in silico », *FEBS Lett.*, vol. 579, 2005, p. 1783-1788.
- Tian T., Burrage K., « Bistability and switching in the lysis/lysigeny genetic regulatory network of bacteriophage λ », *J.Theor.bio.*, vol. 227, 2004, p. 229-237.
- Togashi Y., Kaneko K., « Discreteness-induced stochastic steady state in reaction diffusion systems : self-consistent analysis and stochastic simulations », *Physica D*, vol. 205, 2005, p. 87-99.
- Tyson J. J., Chen C., Novák B., « Sniffers, buzzers, toggles and blinkers : dynamics of regulatory and signaling pathways in the cell », *Curr. Opinion Cell Biol.*, vol. 15, 2003, p. 221-231.
- Yin G., « On Limit Results for a Class of Singularly Perturbed Switching Diffusions », *J.Theor.Probab.*, vol. 14, 2001, p. 673-697.

Article reçu le 5 octobre 2005

Article accepté le 18 mai 2006

Ovidiu Radulescu est maître de conférences en mathématiques à l'Université de Rennes 1. Ses travaux portent sur la modélisation des systèmes complexes en physique et en biologie. Il s'intéresse à plusieurs aspects du fonctionnement des systèmes biologiques : l'aléatoire, la robustesse et la modularité.

Aurélien Muller a obtenu son doctorat en statistiques appliquées en cotutelle entre l'Université de Montpellier 2 et le CEMAGREF de Lyon. Ses travaux portent sur le comportement des valeurs extrêmes de pluies en hydrologie, ainsi que sur la modélisation de réseaux de gènes en biologie moléculaire.

Alina Crudu prépare un doctorat en mathématiques à l'Université de Rennes 1. Ses travaux portent sur les phénomènes stochastiques en biologie moléculaire.