



THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

pour le grade de
DOCTEUR DE L'UNIVERSITÉ DE RENNES 1
Mention : Mathématiques et Applications

Ecole doctorale MATISSE

présentée par

Alina Crudu

préparée à l'unité de recherche UMR CNRS 6625 -IRMAR
Institut de Recherche Mathématique de Rennes
U.F.R. de Mathématique

**Approximations
hybrides de
processus de Markov
à sauts multi-
échelles :
applications aux
modèles de réseaux
de gènes en biologie
moléculaire**

Thèse soutenue à Rennes
le 16 juillet 2009

devant le jury composé de :

Dimitri Pétritis

Professeur, Univ. Rennes1 / *président*

J. F. Yao

Professeur, Univ. Rennes1 / *examineur*

Elisabeth Remy

CR, IML (Marseille) / *examineur*

Arndt Benecke

CR, IHES & IRI (Lille) / *examineur*

Ovidiu Radulescu

MCF, Hdr, Univ. Rennes1 / *directeur de thèse*

Arnaud Debussche

Professeur, ENS Cachan / *directeur de thèse*

Table des matières

1	Introduction	1
I	Modélisation	7
2	Réseaux de réactions biochimiques	9
2.1	Modélisation déterministe	10
2.2	Les réseaux de Petri	12
3	Modélisation stochastique de réseaux biochimiques	15
3.1	Modèles de Markov à sauts pour systèmes de réaction chimiques .	17
3.1.1	Processus de Markov à sauts, quelques rappels	17
3.2	Algorithme de Gillespie	20
3.3	Algorithmes accélérés	21
4	Modèles stochastiques hybrides	25
4.1	Introduction	25
4.1.1	Processus déterministes par morceaux	26
4.1.2	Diffusions hybrides	28
4.2	Simplifications hybrides d'un processus de Markov à sauts purs . .	29
4.2.1	Partition des espèces et des réactions	30
4.2.2	Développement de Kramers-Moyal partiel	31
4.2.3	Approximation hybride avec switch	32
4.2.4	Approximation hybride avec breakage	33
4.2.5	Approximation hybride avec switch singulier	35
4.2.6	Justification de la limite singulière	37
4.3	Critères pratiques pour l'identification des paramètres petits et des super-réactions	40
4.4	Conclusion	41

5	Moyennisation de la cinétique chimique stochastique	43
5.1	Approximations hybrides moyennisées	44
5.2	Réduction de cycles linéaires	47
5.3	Réseaux bruités et règles de design pour hiérarchies de cycles . . .	52
6	Exemples d'applications	55
6.1	Méthodologie pour obtenir des simplifications hybrides	55
6.2	Modèle hybride avec switch : modèle de Cook	56
6.3	Modèle hybride avec breakage : neuroscience et site opérateur chez les bactéries	60
6.4	Premier exemple complexe : lambda-phage	66
6.5	Deuxième modèle complexe : fonctionnement en rafales d'un opéron réprimé chez les bactéries	72
6.5.1	Table 2 - Critères pratiques pour différentes approximations	81
II	Partie mathématique	87
7	Les processus déterministes par morceaux	89
7.1	Définition et construction	89
7.2	Le générateur et le problème de martingale associé	92
7.3	L'espace de Skohorod	99
8	Convergence des processus de Markov à sauts vers des Processus déterministes par morceaux	101
8.1	Introduction	101
8.2	Limite déterministe	102
8.3	Limite décrite par des PDP	106
8.3.1	PDP continu	108
8.3.2	PDP discontinu	117
8.4	PDP obtenus par moyennisation	121
8.5	PDP discontinu obtenus par "switching" singulier	129

Remerciements

Je tiens à remercier en tout premier lieu Arnaud Debbusche et Ovidiu Radulescu qui ont dirigé cette thèse dans la continuité de mon stage de Master 2. Malgré leur activités respectives chargées, ils se sont toujours montrés disponibles pour d'intenses et rationnelles discussions. Pour leur confiance, leur patience et leur générosité, je les remercie.

Je remercie les rapporteurs Ralph Blossey et Samy Tindel pour la rapidité avec laquelle ils ont lu mon manuscrit et l'intérêt qu'ils ont porté à mon travail. Je remercie également les membres du jury qui ont accepté de juger ce travail : Arndt Benecke, Elisabeth Remy, Dimitri Petritis et J.F. Yao.

A mes collègues et amis qui se rassemblent dans le bureau 434 et qui m'honorent de leur amitié, partagent avec moi les inéluctables pertes et coups durs de la vie, et continuent de me faire profiter de leur sagesse et de leur amour pour la vie : Nirmal Antonio Tamarasselvame, Adrien Saumard, Adrien Richou, Jonathan Marco, Victor Péron, Nicolas Bideau, Anne-Claire Bennis, Thomas Sierocinski, Audrey Houillier, Rodolphe Richard. Un grand merci à Ludo Goudenege et Jimmy Lamboley pour l'agréable voyage à Paris. Un grand merci à Maher Kachour et à Arnaud Jobin qui ont su me redonner le sourire.

Je tiens à remercier aussi Dan Goreac qui, malgré tout, m'a soutenu et encouragé pendant mes trois années de thèse.

Je remercie également les membres du Laboratoire et du Département de Mathématiques de l'Université de Rennes 1 pour leur accueil et disponibilité.

Dedic aceasta teza familiei mele care m-a sustinut neconditionat in momentele dificile.

Tin sa le multumesc prietenilor mei : Maria, Vlad, Brindusa, Andi, Calin, Bogdan pentru toate momentele petrecute impreuna. Sper ca aceste momente sa fie numeroase si in viitor.

Je remercie ceux ont le nom n'apparaît pas dans cette page et qui m'ont aidé d'une manière ou d'une autre.

Chapitre 1

Introduction

Le fonctionnement cellulaire au niveau moléculaire implique des interactions et de phénomènes très complexes. Par exemple, les événements et les composantes impliqués dans l'expression de gènes ont été étudiés en détail pour différents organismes : des protéines (activateurs) s'attachent à l'ADN et recrutent l'ARN polymérase pour un gène spécifique. Ensuite, une molécule de ARNm est créée. Cette dernière molécule est transformée dans une protéine fonctionnelle par les ribosomes. D'autre part, quelques protéines (represseurs ou activateurs) peuvent s'attacher à l'ADN et empêcher (ou au contraire, stimuler) le recrutement de la polymérase et l'initiation de la transcription. Ces protéines de régulation, activateurs et represseurs, sont elles-mêmes, produites par d'autres gènes. La capacité d'attachement des protéines à l'ADN peut être modifiée de multiples façons. Une protéine possède des sites permettant l'ajout de groupement fonctionnels (phosphoryl, acetyl, etc). Ces réactions sont catalysées par des molécules spécialisées (par exemple des kinases). Lors de la signalisation cellulaire, les modifications des protéines peuvent s'enchaîner en cascades, initiées par des signaux extérieurs à la cellule.

Simuler et comprendre de tels phénomènes représente un défi pour la modélisation. Assez souvent l'information sur les systèmes de régulation est incomplète et le modélisateur doit émettre des hypothèses. Aussi, prendre en compte toutes les interactions moléculaires connues par les biologiste peut s'avérer un mauvais choix. Beaucoup de ces détails compliquent inutilement le modèle et alourdissent les analyses et l'inversion des paramètres. L'art de la modélisation est donc de faire les simplifications nécessaires et suffisantes pour rendre le modèle le plus proche possible de la réalité et analysable.

Les modèles auxquelles nous nous intéressons dans cette thèse sont des réseaux

de réactions biochimiques. Ces modèles utilisent non seulement des interactions (activation, répression) mais aussi les mécanismes par lesquels les interactions sont mis en place. Reconstruire un mécanisme de réaction est une question difficile que nous n'abordons pas ici. Des techniques expérimentales, permettant d'enlever des composantes et d'enregistrer les effets produits peuvent servir à cette fin. D'autres techniques (par exemples l'inférence bayésienne) permet d'estimer les paramètres cinétiques en comparant la dynamique prédite par le modèle et la dynamique observée expérimentalement.

Suivant le type d'application, la dynamique des réseaux de réactions biochimiques peut être modélisée de plusieurs façons.

La modélisation purement déterministe des réseaux de réactions biochimiques donne lieu à la résolution de systèmes d'équations différentielles ordinaires. Cette approche prend en compte les effets macroscopiques mais néglige le caractère microscopique, stochastique (souvent très relevant) [19],[52],[82] des réactions. Du point de vue numérique, les différents solveurs déterministes permettent de calculer rapidement la solution des équations différentielles et, ainsi, de la dynamique macroscopique. Toutefois, cette méthode est essentiellement basée sur une limite thermodynamique (système de grande taille), hypothèse qui se trouve être loin de la réalité biologique pour une classe suffisamment importante de systèmes. Ainsi, pour les systèmes contenant des espèces en petit nombres, les molécules peuvent présenter des changements discontinus d'état.

Pour rendre compte de la nature discrète et aléatoire des phénomènes à l'échelle microscopique, des méthodes moléculaires et stochastiques ont été proposées. La dynamique moléculaire (Moléculaire Dynamics -MD) représente une manière précise pour décrire l'évolution temporelle d'un système. Cette approche décrit la position de chaque molécule dans le système et simule toute réaction. Ainsi, il est possible de suivre l'évolution temporelle et spatiale de chaque molécule. Le grand désavantage de cette méthode réside justement dans sa précision car pour l'algorithme est extrêmement long même pour les systèmes les plus simples. De même, les effets de l'environnement (agitation thermique, effets en solution) sont difficiles à représenter.

Une approximation importante consiste de considérer que le système est bien mélangé ou homogène spatialement. De plus, les interactions entre molécules sont traitées de manière stochastique en définissant les probabilités par unité de temps pour que différentes réactions se produisent. Dans cette approche les réactions sont considérées indépendantes dans leur ensemble. La méthode est connue sous le nom

de l'algorithme de Gillespie [44] (Stochastic Simulation Algorithm - SSA). Cette méthode a été largement utilisée en biologie moléculaire. Dans [10], les auteurs étudient le rôle des phénomènes stochastiques dans la bifurcation du développement du bactériophage lambda. (Garcia-Olivares, 2000) applique l'algorithme de Gillespie pour des simulations stochastiques de la dynamique du cycle glycolytique et gluconéogénétique. Le même algorithme est utilisé aussi pour donner la relation entre les fréquences initiales de traduction et transcription et la magnitude des fluctuations stochastiques dans l'expression de gènes procaryotes [69].

Au vue des nombreuses applications de l'algorithme de Gillespie, un logiciel pour la simulation des systèmes biochimiques a été développé par [10] **SIMULAC**. [70] écrit, à son tour, un logiciel **STOCKS** en langage C++ dédié à la simulation des processus cellulaires. Plus tard, [1] ont développé un logiciel proposant une interface graphique qui permet à l'utilisateur de spécifier le type de variable (continue ou discrète) pour chaque espèce présente dans le système.

L'inconvénient majeur de l'algorithme SSA réside dans le fait que son coût (en temps d'exécution) est élevé. Comme SSA simule une réaction à la fois, la simulation des réactions fréquentes nécessite un effort de calcul considérable. Par exemple, une enzyme 60 *kDa* avec une activité de 100 *U/mg* et une concentration de 1 μL in vivo produira, en moyenne, 60.000 molécules par seconde (voir, dans ce sens, [61]). D'autre part, la transcription se produira assez rarement. Un tel système, dans lequel les deux événements se produisent simultanément (bien que ayant des échelles différentes), peut être modélisé avec SSA. Toutefois, une telle approche s'avère extrêmement longue.

Des nombreuses améliorations de cet algorithme ont été proposées. On peut citer [46] qui propose une procédure d'approximation tau-leap avec une perte acceptable de la précision du résultat. Cet algorithme utilise une approximation de diffusion du processus moléculaire. Dans cette approximation les fluctuations dans les concentrations deviennent, en utilisant le théorème central limite, Gaussiennes et peuvent être modélisées par l'équation de Langevin.

Un autre algorithme pour la simulation des réactions biochimiques est Next Reaction Method [43]. Il utilise un seul nombre aléatoire pour chaque événement simulé. La gestion du calcul de taux de réaction permet d'avoir un temps de simulation qui est proportionnel au logarithme du nombre de réactions et pas au nombre de réactions lui-même comme pour l'algorithme de Gillespie. En plus, la méthode est adaptée pour inclure des taux de réaction dépendants du temps.

Pour prendre en compte la nature multi-échelle des processus, quelques algo-

rithmes hybrides ont été proposés. [88] proposent un mélange de tau-leap et de Next Reaction Method pour la simulation d'un système de réactions : Maximal Time Step Method . Tandis que les réactions « lentes » sont implémentées avec l'algorithme proposé par [43], les réactions « rapides » sont actualisées avec la méthode développée par [46]. Cet algorithme permet la simulation de systèmes contenant simultanément des réactions métaboliques intenses ainsi que des processus de régulation impliquant des molécules en petit nombre. D'autres algorithmes hybrides traitent des variables continues en utilisant l'équation de Langevin [98], ou équations différentielles ordinaires [116]. Le premier travail propose une partition du système en deux sous-ensembles : réactions « lentes » et réactions « rapides ». L'ensemble des réactions rapides est approché par un processus de Markov continu, en utilisant une équation Langevin. L'ensemble des réactions lentes est simulé avec l'algorithme stochastique Next Reaction. En 2008, [111],[116] proposent les processus Markov déterministes par morceaux comme une alternative raisonnable pour la modélisation des réseaux de régulation de gènes. Ils montrent que de tels processus sont bien adaptés pour révéler les effets du bruit intrinsèque.

Des méthodes similaires basées sur le partitionnement des réactions ont été développées par [51],[18],[66],[89],[53],[55],[64]. Des partitions dynamiques, adaptés ont été proposées par [47]. Dans cette méthode, par défaut, toutes les réactions sont discrètes. Une réaction est considérée comme rapide si elle est au-dessus d'un seuil actualisé de manière dynamique.

[96] décrivent l'application de l'algorithme hybride introduit par [6],[5] pour le système de signalisation du calcium.

Nombreuses synthèses de résultats et des applications en biologie moléculaires sont disponibles dans la littérature : [111],[85],[62].

Dans une perspective mathématique, la modélisation stochastique des systèmes de réactions biochimiques par des processus de Markov à sauts a été proposée par [32] et formalisée mathématiquement par [13], [95].

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. Sous ces conditions, les trajectoires des processus de Markov à sauts convergent, vers des trajectoires déterministes, solutions d'équations différentielles. En théorie des files d'attente cette limite est appelée limite fluide. En cinétique chimique elle est appelée limite thermodynamique. Les techniques nécessaires pour prouver ces résultats sont essentiellement des inégali-

tés de martingales [29],[71],[72],[81]. Il s'agit de la loi des grands nombres pour les processus de Markov. Sous les mêmes conditions, il existe aussi un théorème central limite qui permet l'approximation des processus de Markov par des diffusions [71],[7],[45].

Plus récemment, quelques résultats concernant la modélisation hybride par [51] ont été rendus rigoureux par [16].

L'objectif principal de cette thèse a été de développer des nouveaux outils mathématiques pour l'étude des phénomènes stochastiques en biologie moléculaire. Les modèles mathématiques pour la dynamique stochastique des réseaux de réactions biochimiques sont basés sur les processus de Markov à sauts. On propose des approximations hybrides pour les processus de Markov à sauts multi-échelles. En utilisant comme argument heuristique un développement limité du générateur du processus à sauts (procédé connu en chimie et en physique sous le nom de développement de Kramers-Moyal) nous identifions plusieurs types d'asymptotiques hybrides : processus déterministes par morceaux et diffusions hybrides. Le développement de Kramers-Moyal permet d'obtenir de manière systématique des modèles hybrides, qui sont simulés par la suite avec des algorithmes adaptés.

Les approximations déterministes par morceaux sont étudiées avec des méthodes mathématiques rigoureuses. On montre la convergence faible du processus de Markov à sauts vers deux types de processus déterministes par morceaux : avec et sans sauts dans les variables continues. Les approximations hybrides peuvent être simplifiées davantage en utilisant des méthodes de moyennisation. On propose aussi quelques résultats dans cette direction.

Ces résultats sont à la base d'un algorithme de réduction de modèles stochastiques qui a été testé sur plusieurs exemples issus du cadre biologique : un modèle de happloinsuffisance [27], un modèle de fonctionnement stochastique de neurones (Stein), un modèle de phage λ [52] et deux modèles pour la production de protéines dans les procaryotes [38],[70].

Cette thèse est composée de deux parties indépendantes : une partie modélisation et une partie mathématique.

Dans les deux premiers chapitres de la partie modélisation nous présentons plusieurs méthodes de modélisation des systèmes biologiques (chapitre 2 et chapitre 3). Cette approche nous permet notamment de mettre en évidence les points faibles et d'envisager une autre méthodologie dans l'étude des systèmes issus de la réalité biologique.

Dans le chapitre 4 nous mettons en évidence des phénomènes d'intermittence

pouvant apparaître dans des réseaux de réactions biochimiques, qui conduisent à un fonctionnement bruité. Ces phénomènes apparaissent lorsque plusieurs conditions sont satisfaites : des conditions sur les concentrations (certaines espèces sont en faible nombre) et des conditions sur les échelles de temps. Selon les différentes conditions discutées, on peut rencontrer plusieurs types de comportements hybrides : avec switch (sauts des taux de réaction) et avec breakage (sauts des concentrations).

Dans le chapitre 5 nous proposons une solution générale pour la simplification des réseaux de réaction. Cette méthode peut être appliquée pour simplifier les réseaux stochastiques et pour éliminer les cycles rapides. Dans le chapitre 6, les méthodes d'obtention d'approximations hybrides sont illustrées sur quelques modèles biologiques.

Dans la partie mathématique, nous montrons que les processus de saut correspondant à l'ensemble des réactions chimiques avant simplification convergent en loi vers les processus déterministes par morceaux associés aux divers modèles réduits. On justifie ainsi les modèles PDP continus - ou avec "switching" -, discontinus - ou avec "breaking". On justifie aussi le passage à la limite par moyennisation ou par "switching" singulier.

Première partie

Modélisation

Chapitre 2

Réseaux de réactions biochimiques

Plusieurs méthodes pour la représentation d'un système biologique sont disponibles. Les biologistes ont favorisé les diagrammes accompagnés d'explications concernant les mécanismes de fonctionnement. D'autre part, les mathématiciens préfèrent travailler avec systèmes d'équations différentielles ordinaires (EDO). Cette approche est plus complète du point de vue quantitatif et beaucoup plus précise. Toutefois, cette méthode présente nombreuses désavantages. Dans certains cas, les EDOs captent seulement une description trop évasive des mécanismes d'interaction entre les acteurs biologiques.

Entre les deux extrêmes, les systèmes de réactions biochimiques peuvent être représentés comme des réseaux des réactions chimiques couplés.

Les réseaux de réactions biochimiques offrent une manière simple et compréhensive pour la description des interactions entre les molécules. Pour utiliser ces modèles, on doit supposer que la système est bien mélangée, autrement dit, toute molécule se trouve avec probabilité égale à tout endroit dans le système. De façon équivalente, on pourra supposer que le taux de diffusion moléculaire est beaucoup plus important que les taux de réaction. Dans cette situation, l'état physique du système peut être décrit par un ensemble de variables intrinsèques de base :

- pour les modèles microscopiques, l'état est donné par le nombre de molécules de chaque type,
- pour les modèles macroscopiques, l'état est l'ensemble des concentrations de chaque type chimique.

La modélisation de la dynamique au niveau macroscopique est déterministe

et continue. Elle utilise les équations différentielles ordinaires (ODEs). Au niveau microscopique la dynamique est discrète et stochastique. Plusieurs techniques sont présentes dans la littérature : réseaux de Petri [87], modélisation stochastique [44], modélisation booléenne, pi-calcul etc.

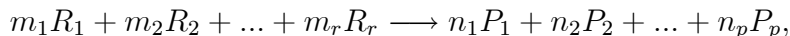
Dans cette thèse nous discuterons aussi un niveau mésoscopique, qui représente en fait le cadre habituel des processus cellulaires. A ce niveau, certaines variables sont des concentrations (pour les espèces en grand nombre), d'autres des nombres de molécules (pour les espèces en petit nombre).

Par la suite on présente deux modèles mathématiques les plus utilisés. Les réseaux de réactions stochastiques seront introduites dans le chapitre suivant.

2.1 Modélisation déterministe

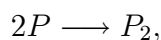
La modélisation déterministe est basée sur des systèmes d'équations différentielles qui sont obtenus à partir des réactions biochimiques.

Une réaction bio-chimique est une transformation de la forme suivante



où r représente le nombre de réactants et p le nombre de produits. $R_i, 1 \leq i \leq r$ représente la i ème molécule réactante et $P_j, 1 \leq j \leq p$ représente la j ème molécule produit. m_i est le nombre de molécules de R_i consommées dans une seule réaction et n_j est le nombre de molécules produites par une seule réaction. Les coefficients m_i et n_j sont appelés coefficients stoechiométriques. Les coefficients stoechiométriques sont, en général, des nombres entiers. Les constantes $R_i, 1 \leq i \leq r$ et $P_j, 1 \leq j \leq p$ peuvent avoir la même valeur et il est parfaitement raisonnable de considérer qu'une certaine molécule est consommée et produite par la même réaction. Autrement dit, une réaction décrit, de manière précise, quelles espèces chimiques interagissent et dans quelle proportion et aussi quelles espèces sont produites et dans quelle proportion.

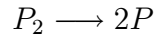
Pour donner un exemple, on considère la dimérisation d'une protéine P . Cela s'écrit de manière formelle



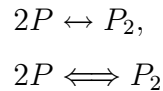
ce qui signifie que deux molécules de P interagissent pour produire une seule molécule de type P_2 .

La molécule P a une stoechiométrie d'ordre 2 et la molécule P_2 a une stoechio-

métrie d'ordre 1. Les constantes stoichiométriques d'ordre 1 ne sont pas écrites explicitement. De la même manière, la réaction pour décrire la dissociation du dimère s'écrit

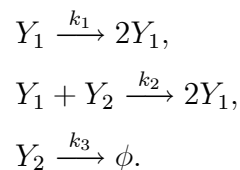


Une réaction peut se produire dans les deux directions et dans ce cas on appelle cette réaction réversible. Ainsi, on écrit une telle réaction avec une flèche double



Donc, une méthode convenable et flexible pour décrire un modèle est d'écrire une liste de réactions correspondant au système choisi. Toutefois, les réactions elles-mêmes donnent seulement la structure qualitative du modèle et doivent être complétées avec une information additionnelle. La description du modèle est complète si on spécifie les taux correspondants à chaque réaction ainsi que les quantités initiales existantes dans le système.

Soit le système de réactions suivantes



Le modèle ci-dessus nous indique que les nombres de molécules Y_1 et Y_2 changent par des quantités discrète quand un événement se produise (i.e. une réaction a lieu). De manière classique chaque molécule sera caractérisée par sa concentration, mesuré en M (moles per litre), qui peut varier de manière continue dans le temps. Par convention, la concentration d'une espèce X est notée $[X]$. Même si les molécules peuvent se rencontrer rapidement, les réactions biochimiques n'ont pas lieu instantanément. En effet, pour qu'une réaction biochimique puisse avoir lieu, une barrière de potentiel doit être dépassée. Ceci demande de l'énergie qui peut venir par exemple de l'énergie thermique. Les constantes de réaction $k_i, 1 \leq i < 3$ spécifient le temps nécessaire aux molécules pour interagir. Ces constantes dépendent de la température.

En général, les taux instantanés de réaction sont proportionnels avec les concentrations (i.e. avec la masse) de chaque réactant. Cette "loi" cinétique est appelée

loi d'action des masses. Pour le système introduit, la deuxième réaction se produit avec un taux proportionnel à $[Y_1][Y_2]$. Par conséquent, sous l'effet de cette réaction, $[Y_1]$ décroît avec le taux instantané $k_2[Y_1][Y_2]$ et $[Y_2]$ croît avec le même taux. En considérant toutes les trois réactions on peut écrire le système d'équations différentielles (ODEs) :

$$\begin{aligned}\frac{d[Y_1]}{dt} &= k_1[Y_1] - k_2[Y_1][Y_2], \\ \frac{d[Y_2]}{dt} &= k_2[Y_1][Y_2] - k_3[Y_2].\end{aligned}\tag{2.1}$$

Les trois taux de réaction k_1, k_2, k_3 et les concentrations initiales du système doivent être spécifiés. Ainsi, la dynamique du système est complètement déterminée et peut être étudiée soit analytiquement soit numériquement.

Le système (2.1) s'écrit sous forme matricielle

$$\frac{d}{dt} \begin{pmatrix} [Y_1] \\ [Y_2] \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} k_1[Y_1] \\ k_2[Y_1][Y_2] \\ k_3[Y_2] \end{pmatrix},$$

où la matrice 2×3 est la matrice stœchiométrique associée au système.

2.2 Les réseaux de Petri

Les réseaux de Petri (RdP), autres grands classiques des outils de modélisation, sont également mis à contribution pour modéliser les réseaux moléculaires. Inventés par Carl Adam Petri, en 1962 [87], ils représentent un formalisme graphique et mathématique qui permet de représenter des systèmes (informatiques, industriels) travaillant sur des variables discrètes.

Un RdP classique se compose de places, de transitions, d'arcs et de jetons. L'idée de Reddy et al. [112] est de faire correspondre une place à une population de molécules d'une espèce chimique, une transition à une réaction, un arc à un coefficient stœchiométrique, et finalement, un jeton à une molécule. Ainsi, il devient possible de décrire qualitativement un réseau moléculaire au niveau microscopique. L'état d'un système représenté par un RdP correspond à la position des jetons dans les places du réseau. En fonction de l'état, certaines transitions

peuvent être valides. Le franchissement d'une transition valide est un événement indivisible qui entraîne le déplacement, la création ou la suppression de jetons. Un franchissement d'étape correspond donc à un changement d'état. Lorsque plusieurs transitions sont simultanément valides, il est impossible de déterminer laquelle sera franchie. Le formalisme des RdP n'est pas déterministe, il y a plusieurs états futurs possibles. Si on étend le formalisme aux RdP stochastiques, il devient alors possible de prendre en compte les différentes probabilités de réaction entre les molécules et d'écrire quantitativement un réseau moléculaire biochimique stochastique au niveau microscopique.

Un réseau de Petri, N est un n -uplet $(P, T, Pre, Post, M)$, où $P = \{p_1, \dots, p_u\}$, ($u > 0$) est un ensemble fini des états, $T = \{t_1, \dots, t_v\}$, ($v > 0$) est un ensemble fini de transitions et $P \cap T = \emptyset$. Pre est une matrice $v \times u$ contenant les poids de chaque arc (l'élément (i, j) de cette matrice représente le poids de l'arc entre l'état j et la transition i). $Post$ est aussi une matrice $v \times u$ contenant les poids entre les transitions et les états (l'élément (i, j) est le poids de l'arc entre la transition i et l'état j). M est un vecteur u -dimensionnel représentant l'état courant du système.

L'état initial du réseau est désigné par M_0 . La forme de Pre et $Post$ garantit l'existence d'un arc entre nœuds de différents types. Ainsi, le réseau résultant est un graphe biparties. Une transition t_i , peut avoir lieu si $M_j \geq Pre_{ij}$, $j = 1, \dots, u$.

On considère le même système d'équations



Les taux stochastiques s'écrivent

$$\begin{aligned} h_1(y, c_1) &= c_1 y_1, \\ h_2(y, c_2) &= c_2 y_1 y_2, \\ h_3(y, c_3) &= c_3 y_2. \end{aligned}$$

Le réseau Petri correspondant au système (2.2) s'écrit

$$N = (P, T, Pre, Post, M, h, c),$$

où $P = (Y_1, Y_2)$, $T = (R_1, R_2, R_3)$,

$$Pre = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}, Post = \begin{pmatrix} 2 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix},$$

et $h(y, c) = (c_1 y_1, c_2 y_1 y_2, c_3 y_2)$. Il reste juste à spécifier l'état initial du système M et le vecteur de constantes pour finir la description du modèle.

Chapitre 3

Modélisation stochastique de réseaux biochimiques

Plusieurs observations nous conduisent à introduire une approche stochastique dans la dynamique des 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 [92]. 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 [27], mais aussi le cas des promoteurs transcriptionnels de faible activité [62]. 3) La production intermittente, par paquets, des protéines [83],[19], [103]; 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 [10], [52],[107]. De même, la différenciation cellulaire en morphogénèse peut être abordée de façon stochastique [76]. 4) Dans le cas du cancer, sous l'incidence du même stimulus, certaines cellules développent le cancer, d'autres non [57]. 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 être réalisée 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

[109]. Néanmoins, les exemples présentés plus haut suggèrent que les aspects stochastiques jouent un rôle important dans le fonctionnement des 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 [32] et a été formalisée plus tard par [13],[95]. 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 de 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 méthode de Gillespie [44]. 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 stoichiométriques 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 [46] propose aussi 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 [71],[72]. 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 [29],[30]. Dans les mêmes conditions, il existe aussi un théorème central limite qui permet l'approximation des processus

de Markov par des diffusions [72],[7],[45].

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 [20],[79],[31],[5], ou celui des diffusions à sauts [114], ou celui des diffusions à plusieurs échelles de temps [40]. 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 [5]. De même, il existe de nombreux résultats [79],[31] sur les propriétés des trajectoires qui peuvent s'appliquer en théorie du contrôle ou pour expliquer les observations expérimentales.

3.1 Modèles de Markov à sauts pour systèmes de réaction chimiques

3.1.1 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) $F_t^X = \sigma\{X(s), s \leq t\}$ et une fonction probabilité de transition $P : R \times E \times R \times B(E) \rightarrow [0, 1]$ ($B(E)$ est l'ensemble des boréliens), $P(s, x, t, A) = P[X(t) \in A | F_t^X] = P[X(t) \in A | X(s) = x]$, $t \geq s$, 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 [41] :

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

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

La fonction $\lambda : R \times E \rightarrow R_+$ 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 donnés par de temps aléatoires τ_i qui suivent des lois exponentielles et tels que $\lim_{n \rightarrow \infty} \sum_{i=1}^n \tau_i = \infty$ et $X(t)$ soit constant sur $\left[\sum_{i=1}^n \tau_i, \sum_{i=1}^{n+1} \tau_i \right]$ [41]. 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 $\mu : X(t) = X_{\nu(t)}$.

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 [65].

L'intérêt du choix markovien réside aussi dans la simplicité et la puissance des techniques mathématiques permettant leur étude. 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 $T(t)f(x) = E[f(X(t)) | X(0) = x]$. Son générateur est défini par $Af = \lim_{\tau \rightarrow 0} \frac{T(\tau)f - f}{\tau}$, pour toute fonction f dans le domaine $D(A)$ où cette limite existe. La relation suivante va être utilisée par la suite [37] :

$$\frac{dT(t)f}{dt} = AT(t)f = T(t)Af$$

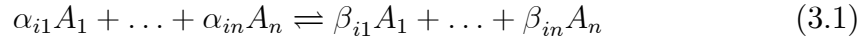
Un processus de Markov à sauts homogène peut être ainsi caractérisé par son générateur [37] :

$$Af(x) = \lambda(x) \int_E (f(z) - f(x)) \mu(x, dz).$$

Systèmes de réactions chimiques

Les systèmes de réactions chimiques sont modélisés par des processus de Markov à sauts homogènes [95],[13]. Lorsqu'il y a n espèces chimiques A_1, \dots, A_n l'espace d'états est $E = \mathbb{R}^n$, un état étant un vecteur X indiquant les nombres de molécules de chaque espèce.

A chaque réaction chimique du type



on associe un saut $\theta_i = \beta_i - \alpha_i \in^n, i = 1, n_r$ (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 :

$$\lambda(X) = \sum_{i=1}^{n_r} [V_i(X) + V_{-i}(X)] \quad (3.2)$$

$$\mu(X, \cdot) = \sum_{i=1}^{n_r} [q_i(X)\delta_{X+\theta_i}(\cdot) + q_{-i}(X)\delta_{X-\theta_i}(\cdot)] \quad (3.3)$$

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 [49]. 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 :

$$V_i(X) = \mathcal{V}v_i(x) = k_i \prod_{s=1}^n x_s^{\alpha_{is}}, \quad V_{-i}(X) = \mathcal{V}v_{-i}(x) = k_{-i} \prod_{s=1}^n x_s^{\beta_{is}} \quad (3.4)$$

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.

Le générateur du processus est :

$$\mathcal{A}f(X) = \sum_{\pm i \in [1, n_r]} [f(X + \gamma_i) - f(X)]V_i(X) \quad (3.5)$$

L'opérateur adjoint du générateur définit l'équation maîtresse (chemical master equation, CME) décrivant l'évolution temporelle de la distribution $p(X, t)$ du processus

$$\frac{\partial p}{\partial t}(X, t) = \mathcal{A}^*p(X, t) = -p(X, t)\lambda(X) + \sum_{\pm i \in [1, n_r]} V_i(X - \gamma_i)p(X - \gamma_i, t) \quad (3.6)$$

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).

3.2 Algorithme de Gillespie

L'algorithme de Gillespie (**Stochastic Simulation Algorithm-SSA**) décrit l'état du système comme un vecteur dont les composantes sont les nombres de molécules impliquées dans les réactions. L'algorithme a été conçu pour répondre à deux questions "quelle est la prochaine réaction à être effectuée et quand cette réaction a-t-elle lieu?". Deux nombres aléatoires sont générés. Le premier sera utilisé pour déterminer la réaction à exécuter et le deuxième pour déterminer le temps d'attente avant que cet événement se réalise. La simulation d'une réaction à la fois fait que cette méthode soit très lente.

Dans ce qui suit on présente de manière détaillée l'algorithme de Gillespie [44].

On considère $N \geq 1$ espèces moléculaires $\{S_1, \dots, S_n\}$. L'état du système est donné par le vecteur $X(t) = (X_1(t), \dots, X_N(t))$, où $X_i(t)$ représente le nombre de molécules de type S_i au temps t . L'état initial du système est donné par

$$X(t_0) = X_0. \quad (3.7)$$

Soit $\{R_1, \dots, R_M\}$, $M \geq 1$ réactions. Les vitesses de réactions sont $V_i(X)$ et dépendent de l'état X .

Pour tout j , $j = 1, \dots, M$, $V_j(X(t))dt$ est la probabilité que, sachant que $X(t) = X$, la réaction R_j est effectuée dans l'intervalle de temps $[t, t + dt)$. Soit soit $\lambda(X)$ la somme de toutes les $V_j(X(t))$

$$\lambda(X) = \sum_{j=1}^M V_j(X).$$

Pour toute réaction biochimique on construit le vecteur stoechiométrique associé. Par exemple, $(-1, -1, 1, 0, \dots, 0)$ est le vecteur stoechiométrique associé à la réaction chimique



A chaque fois qu'une réaction R_j est exécutée, $X(t)$ change d'état. La quantité

par laquelle X_i change est donnée par le vecteur stoechiométrique associé à la réaction R_j . La matrice $N \times M : (m_{ji})$, est appelé matrice stoechiométrique. Chaque colonne de cette matrice représente une réaction. En particulier si la réaction R_j est exécutée dans l'intervalle de temps $[t, t + dt)$, la j ème colonne γ_j de la matrice stoechiométrique est utilisée pour effectuer le changement d'état

$$X(t + dt) = X(t) + \gamma_j. \quad (3.8)$$

L'algorithme de Gillespie est décrit par les étapes suivantes :

1. On fixe la condition initiale $X(t_0)$ comme dans (3.7),
2. Trouver le premier temps τ après t lorsque la prochaine réaction sera exécutée. Ce temps est généré en utilisant une loi exponentielle,
3. Choisir de manière aléatoire une réaction à effectuer au temps $t + \tau$. Tirer un nombre aléatoire uniformément distribué entre 0 et 1. Si le nombre tombe entre 0 et V_1/λ la réaction 1 est choisie, entre V_1/λ et $(V_1 + V_2)/\lambda$ la réaction 2 est choisie et ainsi de suite,
4. Changer l'état de système $X(t)$ en $X(t + \tau)$ en utilisant la relation (3.8) avec $dt = \tau$,
5. Répéter à partir du pas 2) avec le nouveau état $X(t + \tau)$ jusqu'à ce que le temps $TMAX$ soit atteint.

3.3 Algorithmes accélérés

Des nombreuses améliorations de l'algorithme de Gillespie ont été proposées. Citons [46] qui propose une méthode approchée **tau-leap** plus rapide et avec une perte acceptable de la précision du résultat.

Quelques stratégies pour la sélection des paramètres de contrôle et pour l'amélioration de l'erreur de l'algorithme sont présentées.

Cette méthode utilise un temps τ prédéfini qui peut inclure plus d'un événement (i.e. plus d'une réaction). La valeur de τ est limitée par une condition : les changements introduit dans les propensités V_i pendant le temps τ doivent être petits. Soit $N_j(t, t + \tau)$ représente le nombre de réactions du type R_j qui vont se produire entre t et $t + \tau$. Soit $X(t) = x$. Ainsi, si τ est considéré tel que $V_j(X(t))$ est à peu près constant sur l'intervalle $[t, t + \tau]$, alors le nombre des réactions N_j suit une loi de Poisson $\mathcal{P}(V_j(x)\tau)$. Il y a deux version de tau-leap, une qui utilise des variables de Poisson, l'autre qui approche ces variables par des variables

normales.

Il vient,

$$X(t + \tau) = x + \sum_{j=1}^M P(V_j(x)\tau)\gamma_j.$$

ou $P(V_j(x)\tau)$ sont des variables de Poisson ou normales indépendantes.

Pour résumer, l'algorithme de simulation τ -leap est décrit par les étapes suivantes :

1. Dans l'état X , choisir τ tel que le changement dans toutes les fonctions de propensité dans $[t, t + \tau]$ est $\leq \varepsilon\lambda(x)$.
2. Générer le nombre k_j des réactions R_j à exécuter dans $[t, t + \tau]$, avec

$$k_j \sim \mathcal{P}(V_j(x)\tau), \quad j = 1, \dots, M,$$

3. Remplacer $t \leftarrow t + \tau$ et $x \leftarrow x + \sum_{j=1}^M k_j\gamma_j$

4. Répéter à partir de pas 1) avec le nouveau état (x, t) jusqu'à ce que le temps $TMAX$ est atteint.

Un autre algorithme pour la simulation des réactions biochimiques est **Next Reaction Method** [43]. Il utilise un seul nombre aléatoire pour chaque événement simulé.

L'algorithme de Gillespie n'est pas adapté à des réseaux à plusieurs échelles de temps. Si, par exemple, on rencontre des réactions R_j plus fréquentes que d'autres réactions ou quelques espèces X_i en nombre plus important par rapport aux autres, on doit utiliser ces caractéristiques particulières pour développer un algorithme hybride.

[88] propose un mélange de tau-leap et de Next Reaction Method pour la simulation d'un système de réactions : Maximal Time Step Method. Tandis que les réactions « lentes » sont implémentées en utilisant l'algorithme proposé par [43], les réactions « rapides » sont actualisées avec la méthode développée par [46]. Cet algorithme permet la simulation de systèmes contenant en même temps des réactions de forte propensité (réactions métaboliques) ainsi que des processus de régulation impliquant des molécules en petit nombre.

La méthode est implémentée pour modéliser le métabolisme central chez la bactérie *Escherichia Coli*. L'expression de gènes, la traduction et les activités enzymatiques sont intégrées dans le même modèle.

La présentation détaillée des modèles hybrides est donnée dans le chapitre

suivant.

Chapitre 4

Modèles stochastiques hybrides

4.1 Introduction

Les processus stochastiques hybrides sont des couples $(\theta(t), x(t))$, où $\theta(t)$ représente une composante discrète (constante par morceaux) et $x(t)$ est un vecteur avec composantes continues par morceaux. La composante discrète $\theta(t)$ peut être considérée comme une chaîne de Markov contrôlée. La matrice de transition associée à la chaîne de Markov dépend de la variable continue $x(t)$. La composante discrète θ prend valeurs dans l'ensemble fini ou dénombrable Θ . Par exemple, pour un système de réactions biochimiques, les variables discrètes sont les espèces moléculaires "rares", qui se trouvent en petits nombres dans le système. Soit r variables discrètes représentant des espèces moléculaires "rares". On suppose que le nombre de chaque espèce moléculaire ne dépasse pas une valeur N . Alors, $\Theta \subset \{0, \dots, N\}^r$ a, au plus, N^r états. D'autres exemples, c'est quand les valeurs de θ dénombrent les modifications possible d'une molécule. En particulier, une protéine avec r sites de phosphorylation possède 2^r états possibles (chaque site peut être phosphorylé ou pas). Finalement, une autre classe d'exemples est représentée par l'existence d'interactions moléculaires complexes. Ainsi, la transcription de l'ARNm peut être contrôlée par un ou plusieurs facteurs. Lorsque N facteurs peuvent s'attacher à r sites sur la molécule de ADN, le nombre d'états possibles est au plus $(N + 1)^r$.

Il y a plusieurs manières équivalentes pour décrire les transitions entre les états discrets. En général, les transitions sont données par la matrice stochastique $\lambda_{i,j}(x, t)$ où $\lambda_{i,j}dt + o(dt)$ est la probabilité de transition de l'état i à l'état j , dans l'intervalle $[t, t + dt)$. Toutefois, cette représentation n'est pas pratique quand les matrices ont de grandes dimensions. Dans ce cas, la description la plus appropriée

est de définir les caractéristiques du processus de Markov à sauts. Notamment, définir les vecteurs stoechiométriques γ_i correspond à spécifier les sauts possibles. La dynamique est complètement spécifiée par l'intensité (appelée propensité par les chimistes) $\lambda_i(\theta, x)$ représentant la probabilité par unité de temps de faire un saut de l'état θ à l'état $\theta + \gamma_i$ [44]. La probabilité de faire un saut quelconque entre t et $t + dt$ est l'intensité (propensité) totale $\lambda(\theta, x)dt + o(dt)$, où $\lambda = \sum_i \lambda_i$. L'inverse de l'intensité totale représente le temps moyen d'attente entre deux sauts successifs.

L'évolution des variables continues peut être de plusieurs types, chacun conduisant vers un classe différente de processus hybrides.

4.1.1 Processus déterministes par morceaux

L'état d'un processus déterministe par morceaux (piecewise deterministic process - PDP) est un couple $\zeta = (\theta, x)$ où $\theta \in \Theta$ représente la variable discrète et $x \in \mathbb{R}^n$ représente la variable continue. Un PDP est décrit par trois caractéristiques locales, notamment :

- un champs de vecteurs $\chi_\theta(x)$,
- l'intensité de saut $\lambda(\theta, x)$ et
- une mesure de transition Q^T définie telle que $Q^T(\theta, x)$ soit la distribution de probabilité des positions après saut sur $\Theta \times \mathbb{R}^n$. La distribution de sauts $Q^J(\cdot; \theta, x) = Q^T(\cdot + (\theta, x); \theta, x)$ est obtenue à partir de Q^T par translation dans l'espace des phases.

Un tel processus implique un mouvement déterministe interrompu par une séquence de temps d'attente aléatoires τ_i . Sur l'intervalle $[t_i, t_{i+1} = t_i + \tau_i]$, tandis que la variable discrète reste constante $\theta = \theta_i$, la variable continue évolue selon la dynamique donnée par l'équation différentielle :

$$\begin{cases} \frac{dx}{dt} = \chi_{\theta_i}(x), \\ x(0) = x_i. \end{cases} \quad (4.1)$$

Supposons que $\phi(t; (x_i, \theta_i))$ est la solution de (4.1). Le temps d'attente τ_i est une variable aléatoire

$$P[\tau_i > t] = F(t, \theta_i, x_i) = \exp[-\int_0^t \lambda(\theta_i, \phi(s; (x_i, \theta_i)))ds].$$

Après une évolution déterministe sur le $(i + 1)$ -ème intervalle, la variable discrète change d'état vers θ_{i+1} . La variable continue et la variable discret changent d'état selon la relation :

$$x_i \longrightarrow x_i + \gamma_i^C, \theta_i \longrightarrow \theta_i + \gamma_i^D$$

où γ_i^C et γ_i^D sont tirés avec la probabilité $Q^J(\cdot; \theta_i, x_i)$.

Supposons que $\gamma_i^C = 0$, pour tout i . Alors, la variable continue ne change pas lors des sauts. C'est uniquement sa vitesse de variation qui change via le changement de paramètre discret θ dans l'équation différentielle. Dans ce cas, les trajectoires des variables continues sont continues partout. On appelle cette possibilité "switch". Si γ_i^C n'est pas nul, alors la variable continue a des sauts et la trajectoire est continue uniquement par morceaux. On appelle cette possibilité "breakage". Dans la modélisation, on rencontre les deux cas.

Un PDP peut aussi être caractérisé par son générateur A , défini sur des fonctions de classe C^1 en x

$$Af(\theta, x) = \chi_\theta(x) \cdot \nabla_x f + \lambda(x, \theta) \int [f(\theta + \gamma^D, x + \gamma^C) - f(\theta, x)] Q^J(d\gamma^D, d\gamma^C; \theta, x)$$

L'opérateur adjoint du générateur est utilisé pour obtenir l'équation Fokker-Planck décrivant l'évolution temporelle de la distribution du processus

$$\begin{aligned} \frac{\partial p(\theta, x, t)}{\partial t} = A^* p(\theta, x, t) = & -\nabla_x \cdot [\chi_\theta(x) p] - \lambda(\theta, x) p(\theta, x, t) + \\ & + \int \lambda(\theta - \gamma^D, x - \gamma^C) p(\theta - \gamma^D, x - \gamma^C, t) Q^J(d\gamma^D, d\gamma^C; \theta - \theta^C, x - \gamma^C) \end{aligned}$$

Un processus PDP peut être simulé par l'algorithme suivant :

- (1) Choisir la condition initiale $x(t_0) = x_0, \nu = \nu_0$; imposer une durée maximale du processus t_{\max} ,
- (2) Générer une variable aléatoire u qui suit la loi uniforme $\mathcal{U}[0, 1]$,
- (3) Intégrer le système d'équations différentielles

$$\begin{cases} \frac{dx}{dt} = \chi_{\theta_i}(x), \\ \frac{dF}{dt} = -\lambda(t, \theta_i, x) F, \\ x(t_i) = x_i, F(t_i) = 1, \end{cases}$$

dans l'intervalle $[t_i, t_i + \tau_i]$ avec la condition d'arrêt $F(t_i + \tau_i, \nu_i, x_i) = u$,

(4) Générer une deuxième variable aléatoire v de loi $\mathcal{U}[0, 1]$ et l'utiliser pour choisir une réaction (la décision se fait de la même manière que pour l'algorithme Gillespie).

(5) Changer l'état du système $(x(t_i), \theta_i)$ en $(x(t_i) + \gamma_i^C, \theta_i + \gamma_i^D)$,

(6) Répéter le système à partir du pas 2) avec le nouveau état jusqu'au temps t_{\max} défini avant.

4.1.2 Diffusions hybrides

Une classe très générale de diffusions hybrides est introduite par [48]. Une version légèrement moins générale est suffisante pour étudier les approximations des systèmes de réactions biochimiques. On introduit :

- un champs de vecteurs $\chi(\theta, x)$,
- une matrice de diffusion $\sigma(\theta, x)$,
- une intensité de sauts et une mesure de transition définies de la même manière que pour un PDP.

Entre deux sauts consécutifs à t_i et à $t_i + \tau_i$, la variable continue $x(t)$ satisfait l'équation différentielle stochastique de Itô [58] :

$$dx_k(t) = \chi_k(\theta, x)dt + \sum_j \sigma_{ik}(\theta, x)dW_j(t) \quad (4.2)$$

où $W_j(t)$ sont des processus Browniens indépendants. Le temps d'attente est déterminé de la même manière que pour un processus PDP. Ainsi, on doit intégrer

$$\begin{cases} dx_k(t) = \chi_k(\theta, x) + \sum_j \sigma_{ik}(\theta, x)dW_j(t), \\ dF = -\lambda(\theta, x)Fdt, \end{cases} \quad (4.3)$$

avec la condition initiale donnée par $F(t_i) = 1$, $x(t_i) = x_i$ et avec la condition d'arrêt $F(t_i + \tau_i) = u$, où u est une variable aléatoire uniforme dans $[0, 1]$, indépendante de (θ_j, x_j) , $j \leq i$. Le générateur d'un tel processus est donné par :

$$\mathcal{A}f(\theta, x) = \chi_\theta(x) \cdot \nabla_x f + \frac{1}{2} \sigma(\theta, x) : \nabla_x^2 f + \sum_i \lambda_i [f(\theta + \gamma_i, x) - f(\theta, x)] \quad (4.4)$$

où : est le produit tensoriel contracté ($\sigma : \nabla_x^2 f = \sum_{ij} \sigma_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}$). Le lecteur est invité à remarquer qu'un PDP est un cas particulier de diffusion hybride lorsque $\sigma = 0$.

4.2 Simplifications hybrides d'un processus de Markov à sauts purs

Les processus de Markov à sauts pur ont été présentés dans le chapitre précédent. Ces processus, entièrement discrets, sont à la base du formalisme de Delbrück-Rényi-Bartholomay pour la dynamique stochastique de réactions biochimiques et peuvent être simulés par l'algorithme de Gillespie (SSA). Les inconvénients de cette méthode ont été discutés précédemment. Nous montrerons comment on peut approcher les processus de Markov à sauts purs par des processus hybrides. L'approximation hybride est bien adaptée aux systèmes de réactions biochimiques multi-échelles et permet la diminution du temps d'exécution des simulations ainsi que l'étude analytique du comportement intermittent des trajectoires.

Il existe plusieurs algorithmes hybrides pour la dynamique stochastique des réseaux de réactions biochimiques. Ces algorithmes cherchent à réduire le nombre de sauts à effectuer pour simuler une trajectoire, en traitant les réactions rapides en tant que variables continues [51],[16],[5],[6]. Malgré des avantages incontestables (facilité de l'attribution adaptative des degrés de liberté continus) ces algorithmes ne sont pas adaptés à rendre compte de phénomènes d'intermittence le plus souvent rencontrés en biologie moléculaire. En particulier, l'utilisation des nombres de réactions rapides en tant que variables continues conduit à des modèles réduits difficiles à interpréter. Une décomposition des variables basée sur des espèces moléculaires rend mieux compte de la nature des processus.

Ainsi, l'activité de certains promoteurs et la production des protéines correspondantes peut présenter le phénomène de rafales, caractérisé par courtes périodes d'activité intense suivies de périodes d'inactivité [115]. Dans ce cas, une croissance abrupte du nombre de transcrits est suivie par une dégradation exponentielle en général plus lisse. Dans cet exemple les variables continues sont naturellement les concentrations des transcrits ou les produits des gènes (protéines), tandis que l'état d'activité de l'ADN définie par la présence ou l'absence d'un facteur de transcription est la variable discrète. Dans les réseaux de gènes le bruit est pro-

duit au niveau microscopique des espèces en petit nombre et transmis par des phénomènes d'intermittence au niveau mésoscopique des produits de gènes. L'amplitude du bruit au niveau mésoscopique n'est pas reliée au nombre de molécules (à ce niveau les molécules sont en grand nombre) mais à la nature intermittente, discontinue du processus. Contraire à l'intuition il est possible d'avoir des amplitudes de bruit qui augmentent, au lieu de diminuer avec le nombre de transcrits dans une rafale.

L'utilisation des espèces moléculaires pour définir des variables lentes et rapides est proposé aussi dans [78]. Ce travail, très proche de nos approches propose une combinaison des perturbations singulières pour l'équation maîtresse et des développements Ω de Van Kampen pour obtenir des solutions approchées pour la dynamique stochastique [117]. Néanmoins, dans [78] on s'intéresse moins à l'intermittence (on discute uniquement la quasi-stationnarité qui dans notre approche représente un cas de moyennisation traité plus loin) et on ne couvre pas toutes les situations qu'on peut rencontrer lorsque des espèces en petits nombres transmettent le bruit à des espèces en grands nombres. De même, les approximations proposées ne sont pas caractérisées en tant que processus hybrides.

Notre méthode comporte les éléments suivants :

- Une partition binaire des espèces et ternaire des réactions.
- Un développement du générateur du processus Markov à sauts pur. Ce développement correspond en premier ordre au développement Ω de Van Kampen et en deuxième ordre au développement de Kramers-Moyal.
- Certains cas à plusieurs échelles de temps doivent être traités séparément par des techniques de perturbations singulières.

4.2.1 Partition des espèces et des réactions

Tout d'abord, les espèces sont partagées en deux sous-ensembles $X = (X_D, X_C)$, où X_D sont les espèces en petits nombres et X_C sont les espèces en grands nombres. Cette partition, basée sur les espèces, diffère d'autres partitions proposées dans la littérature [51],[6],[5],[105],[99],[47], qui sont basées sur des réactions. Les travaux cités se proposent comme objectif la réduction du temps de simulation. Par rapport à cet objectif, toutes les schémas qui évitent la simulation individuelle de réactions trop fréquentes en les regroupant dans des variables continues, sont plus ou moins équivalentes. Nous cherchons à satisfaire aussi un autre objectif, notamment trouver une représentation naturelle pour les processus simplifiées.

De ce point de vue, une partition basée sur des espèces est plus adaptée qu'une partition basée sur des réactions.

Dans notre approche, la partition des réactions est secondaire. Elle résulte du choix de la partition des espèces. Soit M l'ensemble de toutes les réactions dans le système. On définit la partition $M = \mathcal{R}_D \cup \mathcal{R}_{DC} \cup \mathcal{R}_C$. Les réactions \mathcal{R}_D impliquent uniquement les réactants et les produits en petits nombres (de X_D). De même, les réactions \mathcal{R}_C impliquent uniquement des composantes en grand nombres (de X_C). Les réactions \mathcal{R}_{DC} impliquent au moins un réactant ou un produit dans chacun des sous-ensembles X_D et X_C .

Le deuxième pas vers un processus hybride est de rescaler les variables continues en les divisant par le volume pour obtenir des concentrations $x_C = X_C/\mathcal{V}$. Les variables discrètes X_D resteront inchangées. Leur évolution est décrite par de transitions discrètes.

4.2.2 Développement de Kramers-Moyal partiel

Le troisième pas est de considérer un développement Taylor de l'équation maîtresse pour les variables continues x_C et par rapport aux quantités γ_i/\mathcal{V} . Ce développement peut être effectué de manière équivalente sur le générateur (3.5) ou sur l'équation maîtresse (3.6). Pour simplifier la présentation, on utilise le générateur. Les fonctions de test dépendent de deux variables (x_C, X_D) et on considère qu'elles sont deux fois différentiables par rapport aux variables x_C . On suppose que les taux des réactions \mathcal{R}_C sont proportionnels au volume (conséquence directe de la loi d'action des masses), $V_i = \mathcal{V}v_i(x_C)$, $i \in \mathcal{R}_C$. Le développement de Taylor au second ordre du (3.5) par rapport à x_C donne :

$$\begin{aligned} \mathcal{A}^{(2)}f(x_C, X_D) = & \sum_{i \in \mathcal{R}_D} [f(x_C, X_D + \gamma_i^D) - f(x_C, X_D)] V_i(X_D) + \\ & \sum_{i \in \mathcal{R}_{DC}} [f(x_C + \gamma_i^C/\mathcal{V}, X_D + \gamma_i^D) - f(x_C, X_D)] V_i(x_C, X_D) + \left[\sum_{i \in \mathcal{R}_C} \gamma_i^C v_i(x_C) \right] \nabla_{x_C} f(x_C, X_D) + \\ & \frac{1}{2\mathcal{V}} \left[\sum_{i \in \mathcal{R}_C} v_i(x_C) \gamma_i^C \otimes \gamma_i^C \right] : \nabla_{x_C, x_C}^2 f(x_C, X_D) \end{aligned} \quad (4.5)$$

où \otimes est le produit tensoriel ($(\gamma \otimes \gamma)_{kl} = \gamma_k \gamma_l$). Le générateur approché correspond à une diffusion hybride avec le champs (terme de dérive) $\chi(x_C) = \sum_{i \in \mathcal{R}_C} \gamma_i^C v_i(x_C)$ et de matrice de diffusion $\sigma(x_C) = \frac{1}{\mathcal{V}} \sum_{i \in \mathcal{R}_C} v_i(x_C) \gamma_i^C \otimes \gamma_i^C$.

On appelle Eq.(4.5) développement de Kramers-Moyal partiel à l'ordre deux.

Les termes de dérive et de diffusion pour les variables x_C ne dépendent pas des variables discrètes X_D . Ainsi, le switch n'est pas présent dans cette approximation et les variables discrètes peuvent être oubliées si elles ne sont pas observées. En plus, les sauts γ_i^C/\mathcal{V} disparaissent dans la limite $\mathcal{V} \rightarrow \infty$. Cela implique que les variables continues ne subissent pas des sauts (il y a absence de breakage). Toutefois, les deux types d'événements sont présents dans beaucoup d'exemples dans la biologie moléculaire. Pour introduire les possibilités de switch et breakage, on doit considérer que certaines réactions dans \mathcal{R}_{DC} ont des propriétés particulières. Ainsi, on considère deux types de réactions spéciales :

1. réactions dans \mathcal{R}_{DC} qui sont assez fréquentes pour affecter les taux de variation des variables continues et induire, ainsi, le switch,
2. réactions dans \mathcal{R}_{DC} avec un vecteur stoechiométrique assez grand pour introduire un breakage.

On appellera ces équations "super-réactions". Le premier type de super réactions correspond à l'ensemble $S_1 \subset \mathcal{R}_{DC}$ et le deuxième type de super réactions correspond $S_2 \subset \mathcal{R}_{DC}$. Comme conséquence, on distingue entre deux types d'approximation suivant un des deux types de super réactions.

4.2.3 Approximation hybride avec switch

Le premier type de processus approché est induit par les super-réactions $S_1 \subset \mathcal{R}_{DC}$ avec des vecteurs stoechiométriques d'ordre zero, mais avec des taux de réaction d'ordre un dans le volume : $V_i = \mathcal{V}v_i(X_D, x_c)$, $i \in S_1$. Les réactions dans S_1 induisent des transitions rapides dans les variables continues. L'ensemble S_2 est considéré vide. Pour cette approximation, même si les taux des réactions dépendent des variables discrètes, on suppose que les super réactions ne changent pas les variables discrètes. Autrement dit,

$$\gamma_i^D = 0, \text{ pour tout } i \in S_1.$$

Si cette condition n'est pas satisfaite, un troisième cas survient : quelques variables discrètes changent de manière rapide et on peut les moyenniser. Cette situation sera discutée par la suite. L'évolution des variables discrètes est déterminée par des sauts lents (réactions de \mathcal{R}_D). La composante x_C a des sauts rapides mais, leurs amplitudes sont relativement petites. Ainsi, entre deux transitions consécutives

de la variable discrète, les variables continues peuvent être approchées par une trajectoire déterministe. Les variables continues ont des trajectoires continues, mais des vitesses discontinues. Un développement au premier ordre de Kramers-Moyal correspond à une *approximation déterministe par morceaux* définie par le champ de vecteurs

$$\chi(X_D, x_C) = \sum_{i \in S_1} \gamma_i^C v_i^{DC}(X_D, x_C) + \sum_{i \in R_C} \gamma_i^C v_i^C(x_C), \quad (4.6)$$

l'intensité de saut

$$\lambda(X_D, x_C) = \sum_{i \in R_D} V_i^D(X_D) + \sum_{i \in R_{DC} \setminus S_1} V_i^{DC}(X_D, x_C), \quad (4.7)$$

et la probabilité de saut donnée par

$$P[X_D \longrightarrow X_D + \gamma_i^D] = \frac{1}{\lambda(X_D, x_C)} \times \begin{cases} V_i^D(X_D), & \text{if } i \in R_D, \\ V_i^{DC}(X_D, x_C), & \text{if } i \in R_{DC} \setminus S_1, \end{cases} \quad (4.8)$$

où γ_i^D, γ_i^C sont les projections de γ_i sur les composantes continues, respectivement sur les composantes discrètes. La relation (4.7) montre que l'intensité de saut dépend des variables continues et des variables discrètes. Néanmoins, seulement les variables discrètes sont changées par les sauts. Le développement à l'ordre deux de Kramers-Moyal correspond à une *approximation par diffusion hybride* avec la matrice de diffusion :

$$\sigma(X_D, x_C) = \frac{1}{V} [\sum_{i \in S_1} v_i^{DC}(X_D, x_C) \gamma_i^C \otimes \gamma_i^C + \sum_{i \in R_C} v_i^C(x_C) \gamma_i^C \otimes \gamma_i^C], \quad (4.9)$$

Pour des résultats rigoureux concernant ce processus voir section 8.3.1.

4.2.4 Approximation hybride avec breakage

Ce type d'approximation est induit par les super réactions avec un vecteur stoechiométrique significatif (i.e. réactions de type S_2). On suppose également que l'ensemble S_1 est vide. Autrement dit, pour avoir cette approximation, on suppose qu'au moins une réaction i dans l'ensemble R_{DC} possède un vecteur stoechiométrique assez grand : $(\gamma_i)_j/V$ est comparable avec $|x_j|$ pour $i \in S_2$ et pour des espèces $j \in C$. Une super réaction de S_2 peut changer les deux composantes : continues et discrètes. La différence majeure par rapport à S_1 consiste dans le fait

que les concentrations des composantes x_C sont changées de manière significative, dans un seul pas, par une réaction S_2 .

Les réactions de type S_2 ne contribuent pas au flot déterministe. Elles apparaissent comme des sauts. Comme avant, on écrit les caractéristiques locales pour un processus PDP (développement Kramers-Moyal au premier ordre), notamment le flot

$$\chi(x_C, t) = \sum_{i \in R_C} \gamma_i^C v_i^C(x_C), \quad (4.10)$$

l'intensité de saut

$$\lambda(X_D, x_C) = \sum_{i \in R_D} V_i^D(X_D) + \sum_{i \in R_{DC}} V_i^{DC}(X_D, x_C), \quad (4.11)$$

les probabilités de saut pour les variables discrètes

$$P[X_D \longrightarrow X_D + \gamma_r^D] = \frac{1}{\lambda(X_D, x_C)} \times \begin{cases} V_i^D(X_D), & \text{si } i \in R_D, \\ V_i^{DC}(X_D, x_C), & \text{si } i \in R_{DC} \setminus S_2, \end{cases} \quad (4.12)$$

et les probabilités de saut pour les variables continues

$$P[x_C \longrightarrow x_C + \gamma_i^C/V] = \frac{1}{\lambda(X_D, x_C)} V_i^D(X_D), \text{ pour tout } i \in S_2. \quad (4.13)$$

Si, pour les super-réactions $i \in S_2$ on a $\gamma_i^{(D)} \neq 0$, alors les sauts se produisent simultanément dans les variables continues et dans les variables discrètes. Le flot déterministe défini par le champs de vecteur $\chi_{X_D}(x_C)$ (équation 4.10) montre que seulement les variables continues contribuent à l'évolution déterministe. Dans le développement à l'ordre deux de Kramers-Moyal, seulement les réactions de \mathcal{R}_C contribuent à la diffusion :

$$\sigma(x_C) = \frac{1}{V} \sum_{i \in R_C} v_i^C(X_D, x_C) \gamma_i^C \otimes \gamma_i^C, \quad (4.14)$$

Si les deux ensemble S_1, S_2 sont non-vides, alors les deux types de discontinuités apparaissent dans la trajectoire de x_C :

- switching : discontinuité des vitesses (saut dans X_D), sans discontinuité de la trajectoire,
- breakage : discontinuité de la trajectoire (saut dans les variables x_C).

Tandis que le premier type de discontinuité est induit par les réactions de S_1 , le second cas est induit par les réactions de S_2 . Pour des résultats rigoureux

concernant ce processus voir section 8.3.2.

4.2.5 Approximation hybride avec switch singulier

Du point de vue mathématique le breakage et le switch sont distincts. Cependant, dans certains cas les deux situations peuvent être indiscernables. En effet, le saut dans la variable continue peut être déterminé par l'application répétée d'un ou plusieurs types de réactions appartenant à l'ensemble $S_3 \subset S_1$ et qui sont beaucoup plus rapides que les autres réactions dans S_1 . Pour éviter que le système explose, nous considérons que les réactions super-rapides dans S_3 n'agissent que pendant un très court intervalle quand certaines variables discrètes sont non nulles. Ainsi,

$$V_i = \begin{cases} \frac{\nu}{\epsilon} v_i(X_D, x), & \text{si } X_{D_i} \neq 0, \\ 0, & \text{si } X_{D_i} = 0. \end{cases} \quad (4.15)$$

avec $0 < \epsilon \ll 1$ un petit paramètre positif et où X_{D_i} , $i \in S_3$ représente des espèces discrètes (substrats des réactions i). De la même manière que pour les réactions appartenant à l'ensemble S_1 , on suppose que

$$\gamma_i^D = 0,$$

pour tout $i \in S_3$. Pour toute réaction $i \in S_3$ on considère deux sous-ensembles dans \mathcal{R}_D . Ainsi, dénote par $\mathcal{R}_{D_i}^+$ l'ensemble des réactions produisant les espèces X_{D_i} et par $\mathcal{R}_{D_i}^-$ l'ensemble des réactions qui consomment les espèces X_{D_i} . Soit

$$\begin{aligned} \mathcal{R}_D^- &= \bigcup_i \mathcal{R}_{D_i}^-, \\ \mathcal{R}_D^+ &= \bigcup_i \mathcal{R}_{D_i}^+. \end{aligned}$$

On considère que toutes les réactions dans $\mathcal{R}_{D_i}^-$ sont rapides :

$$V_j = \frac{1}{\epsilon} v_j(X_D), \text{ pour tout } j \in \mathcal{R}_{D_i}^-.$$

Le processus contient deux petits paramètres $1/\mathcal{V}$ et ϵ . Nous allons considérer successivement les limites $\mathcal{V} \rightarrow \infty$ et $\epsilon \rightarrow 0$. La première limite correspond au développement Kramers-Moyal partiel et permet d'obtenir un PDP avec switch

ayant le champ de vecteurs

$$\chi(X_D, x) = \chi_0(x) + \frac{1}{\epsilon} \chi_1(X_D, x), \quad (4.16)$$

où

$$\chi_0(x) = \sum_{i \in \mathcal{R}_C} \gamma_i^C v_i(x), \quad (4.17)$$

et

$$\chi_1(X_D, x) = \sum_{i \in \mathcal{S}_3} \gamma_i^C v_i(X_D, x). \quad (4.18)$$

L'intensité de saut du processus PDP est donnée par

$$\lambda(X_D, x) = \sum_{i \in \mathcal{R}_D \setminus \mathcal{R}_D^-} V_i(X_D) + \frac{1}{\epsilon} \sum_{j \in \mathcal{R}_D^-} v_j(X_D). \quad (4.19)$$

La deuxième limite est singulière et doit être traitée différemment. Pour la simplicité des calculs on considère que les espèces discrètes X_{D_i} , composantes des réactions de type \mathcal{S}_3 , prennent seulement deux valeurs $X_{D_i} \in \{0, 1\}$.

Lorsque $\epsilon \rightarrow 0$, le PDP converge vers le processus dont le générateur est :

$$\begin{aligned} \mathcal{A}f(X_D, x) = & \\ & \chi_0(x) \cdot \nabla_x f + \sum_{i \in \mathcal{R}_D \setminus (\mathcal{R}_D^- \cup \mathcal{R}_D^+)} V_i(X_D) [f(X_D + \gamma_i, x) - f(X_D, x)] + \\ & + \sum_{i \in \mathcal{R}_D^+} V_i(X_D) \sum_{j \in \mathcal{R}_D^-} \int_0^\infty [f(X_D + \gamma_i^D + \gamma_j^D, \Phi(s; x, X_D + \gamma_i^D)) - f(X_D, x)] \rho_{ij}(s) ds \end{aligned} \quad (4.20)$$

où $\rho_{ij}(s) = v_j(X_D + \gamma_i^D) \exp[-v_j(X_D + \gamma_i^D)s]$ est une densité de probabilité satisfaisant $\int_0^\infty \rho_{ij}(s) ds = 1$ et $\Phi(s; x, X_D)$ est la solution de l'équation différentielle suivante $\Phi_s = \chi_1(X_D, \Phi)$, $\Phi(0) = x$.

Ce résultat est justifié heuristiquement dans la sous-section suivante et prouvé rigoureusement dans la deuxième partie de la thèse.

Le processus limite est un PDP avec breakage. Contrairement au cas précédent, quand le breakage résulte de super réactions \mathcal{S}_2 et les tailles des sauts sont constantes, les sauts sont maintenant aléatoires et leurs tailles sont distribuées continûment. Les transitions $X_{D_i} = 0 \rightarrow X_{D_i} = 1$ et au retour disparaissent dans la limite singulière (les variables X_{D_i} restent avec probabilité un dans l'état $X_{D_i} = 0$). Pour des résultats rigoureux concernant ce processus voir section 8.5.

4.2.6 Justification de la limite singulière

On peut obtenir le processus limite par deux méthodes. La première méthode utilise un changement d'échelle, la deuxième méthode utilise un développement de l'équation Fokker-Planck.

L'état du système est (x, X_D, X'_D) où x sont les espèces continues, X_D les espèces discrète, et X'_D les espèces discrètes qui sont substrats des réactions du type \mathcal{S}_3 . On considère que X'_D prennent seulement deux valeurs $X'_D \in \{0, 1\}$. Si $X'_D = 1$, alors les espèces x sont produites rapidement par les réactions \mathcal{S}_3 , et si $X'_D = 0$ la production des espèces x s'arrête. Une fois arrivé dans l'état $X'_D = 1$ le système revient rapidement à l'état $X'_D = 0$ par les réactions rapides R_D^- .

Première méthode, changement d'échelle Considérons les processus dans un état $(X_D, X'_D = 0, x)$. Avec intensité $V_i(X_D)$ on saute dans l'état $(X_D + \gamma_i^D, X'_D = 1, x)$, où on reste un temps très court. Dans cet état les variables continues subissent la dynamique rapide :

$$\frac{dx_\epsilon}{dt} = \frac{1}{\epsilon} \chi_1(X_D + \gamma_i^D, x_\epsilon), \quad x_\epsilon(0) = x$$

pendant un temps aléatoire τ_ϵ qui satisfait

$$P[\tau_\epsilon > t] = \exp\left[-\frac{1}{\epsilon} v_j(X_D + \gamma_i^D) t\right] \quad (4.21)$$

jusqu'au saut $X_D + \gamma_i^D + \gamma_j^D, X'_D = 0$ induit par une réaction $j \in R_-$.

Soit $\Phi(s; x, X_D)$ la solution de l'équation

$$\frac{d\Phi}{ds} = \chi_1(\Phi, X_D), \quad \Phi(0) = x$$

En changeant l'échelle de temps $s = t/\epsilon$, on peut montrer facilement que la variation de la variable continue partant de x et après le temps τ_ϵ est :

$$\Delta x_\epsilon = x_\epsilon(\tau_\epsilon) - x = \Phi(\tilde{s}; x, X_D + \gamma_i^D) - x$$

où le temps aléatoire \tilde{s} satisfait

$$P[\tilde{s} > s] = \exp[-v_j(X_D + \gamma_i^D) s]$$

La variable aléatoire Δx_ϵ devient le saut (breakage) des variables continues. On remarque à partir de l'Eq.(4.21) que $\tau_\epsilon \rightarrow 0$ (presque sûrement). On obtient

le générateur du processus limite :

$$\begin{aligned}
Af(X_D, x) = & \\
& \chi_0(x) \cdot \nabla_x f + \sum_{i \in R_D \setminus (R_D^- \cup R_D^+)} V_i(X_D) [f(X_D + \gamma_i, x) - f(X_D, x)] + \\
& + \sum_{i \in R_D^+} V_i(X_D) \sum_{j \in R_D^-} \int_0^\infty [f(X_D + \gamma_i^D + \gamma_j^D, \Phi(s; x, X_D + \gamma_i^D)) - f(X_D, x)] \rho_{ij}(s) ds
\end{aligned} \tag{4.22}$$

où $\rho_{ij}(s) = v_j(X_D + \gamma_i^D) \exp[-v_j(X_D + \gamma_i^D)s]$ est une densité de probabilité $\int_0^\infty \rho_{ij}(s) ds = 1$.

A partir du générateur on peut obtenir l'équation de Fokker-Planck (on utilise l'opérateur adjoint). Dans l'adjoint, les trajectoires partant de x doivent être remplacées par des trajectoires arrivant en x , le temps s devient $-s$ et la densité doit être rénormalisée pour prendre en compte la transformation de volume dans l'espace de phase. On peut ainsi écrire l'équation de Fokker-Planck hybride :

$$\begin{aligned}
\frac{\partial p}{\partial t}(X_D, x, t) = & \\
& - \nabla_x \cdot (\chi_0(x) p(X_D, x, t)) + \sum_{i \in R_D \setminus (R_D^- \cup R_D^+)} p(X_D - \gamma_i, x, t) V_i(X_D - \gamma_i) \\
& - p(X_D, x, t) \sum_{i \in R_D \setminus (R_D^- \cup R_D^+)} V_i(X_D) + \\
& + \sum_{i \in R_D^+} \sum_{j \in R_D^-} V_i(X_D - \gamma_i^D - \gamma_j^D) \int_0^\infty p(X_D - \gamma_i^D - \gamma_j^D, \phi(-s; x, X_D - \gamma_j^D), t) \rho'_j(s) ds \\
& - p(X_D, x, t) \sum_{i \in R_D^+} V_i(X_D)
\end{aligned} \tag{4.23}$$

où $\rho'_j(s) = v_j(X_D - \gamma_j^D) \exp[-\int_0^s (\lambda_-(X_D - \gamma_j^D) + \nabla_x \cdot \chi_1(X_D - \gamma_j^D, \phi(-s'; x, X_D - \gamma_j^D))) ds']$.

Deuxième méthode : développement de l'équation de Fokker-Planck

Pour simplifier les formules on considère sans perte de généralité que $R_D = R_D^- \cup R_D^+$.

L'équation de Fokker-Planck pour le PDP avec switch singulier est :

$$\begin{aligned} \frac{\partial p}{\partial t}(X_D, 1, x, t) &= -\nabla_x \cdot [p(X_D, 1, x, t)(\chi_0(x) + \frac{1}{\epsilon}\chi_1(X_D, x))] + \\ &\sum_{i \in R_D^+} p(X_D - \gamma_i^D, 0, x, t)V_i(X_D - \gamma_i^D) - \frac{1}{\epsilon}\lambda_-(X_D)p(X_D, 1, x, t) \\ \frac{\partial p}{\partial t}(X_D, 0, x, t) &= -\nabla_x \cdot [(p(X_D, 0, x, t)\chi_0(x))] + \\ &\frac{1}{\epsilon}\sum_{j \in R_D^-} p(X_D - \gamma_j^D, 1, x, t)v_j(X_D - \gamma_j^D, 1) - \lambda_+(X_D)p(X_D, 0, x, t) \end{aligned} \quad (4.24)$$

où

$$\lambda_-(X_D) = \sum_{j \in R_D^-} v_j, \quad \lambda_+(X_D) = \sum_{j \in R_D^-} V_j.$$

Développons la solution en puissances de ϵ , $p(X_D, X'_D, x, t) = p_0(X_D, X'_D, x, t) + \epsilon p_1(X_D, X'_D, x, t) + \dots$

Par développement de Taylor dans l'équation de Fokker-Planck 4.24 il vient :

A l'ordre ϵ^{-1}

$$p_0(X_D, 1, x, t) = 0$$

A l'ordre ϵ^0

$$\left\{ \begin{array}{l} -\nabla_x [p_1(X_D, 1, x, t)\chi_1(X_D, x)] \\ + \sum_{i \in R_D^+} p_0(X_D - \gamma_i^D, 0, x, t)V_i(X_D - \gamma_i^D) - \lambda_-(X_D)p_1(X_D, 1, x, t) = 0 \\ \frac{\partial p_0}{\partial t}(X_D, 0, x, t) = -\nabla_x [p_0(X_D, 0, x, t)\chi_0(x)] \\ + \sum_{j \in R_D^-} p_1(X_D - \gamma_j^D, 1, x, t)v_j(X_D - \gamma_j^D, 1) - \lambda_+(X_D)p_0(X_D, 0, x, t) \end{array} \right. \quad (4.25)$$

A partir de la première équation Eqs.4.25 il vient

$$\begin{aligned} \frac{dp_1(*, 1, \Phi(-s, x, *, *))}{de} &= [\lambda_-(*) + \nabla_x \cdot \chi_1(X_D, \Phi(-s; x, **))]p_1(*, 1, \Phi(-s; x, **), t) - \\ &- \sum_{i \in R_D^+} p_0(* - \gamma_i^D, 0, \Phi(-s; x, **), t)V_i(* - \gamma_i^D) \end{aligned} \quad (4.26)$$

L'équation différentielle linéaire (4.26) a la solution :

$$\begin{aligned}
& p_1(*, 1, x, t) \\
&= p_1(*, 1, \Phi(0; x, **), t) = \\
&= \int_0^\infty \sum_{i \in R_D^+} p_0(* - \gamma_i^D, 0, \Phi(-s; x, **), t) V_i(* - \gamma_i^D) \times \\
&\quad \exp\left[-\int_0^s (\lambda_-(*) + \nabla \cdot \chi_1(X_D, \Phi(-s'; x, **))) ds'\right] ds
\end{aligned} \tag{4.27}$$

En remplaçant (4.27) dans la deuxième équation 4.25 on obtient de nouveau l'éq.(4.23).

4.3 Critères pratiques pour l'identification des paramètres petits et des super-réactions

La loi des grands nombres s'applique dans la limite ($\mathcal{V} \rightarrow \infty$). Cependant, dans le cadre de la biologie cellulaire, l'idée du volume infiniment grand doit être considérée avec précaution. Ainsi, on va remplacer cette condition par des critères concernant l'ordre de paramètres des modèles. Les critères proposés définissent le domaine d'applicabilité des approximations déterministes par morceaux.

Tous nos critères sont basés sur deux références de base. La première référence est le plus petit nombre N de molécules appartenant aux espèces continues. N doit être un grand nombre. La deuxième référence est un temps τ . Ce temps est le minimum entre le temps caractéristique de la dynamique déterministe des variables continues et le temps entre deux sauts successives de la variable discrète. Le temps τ peut être relié aux paramètres cinétiques par des méthodes proposées dans [50],[91].

L'approximation déterministe s'applique aux variables continues sous deux conditions. La première condition concerne l'amplitude des sauts dans la variable discrète. Ainsi, les sauts doivent être petits par rapport aux nombres de molécules, i.e. $\gamma_i^C \ll N$ pour tout $i \in R_C$. La deuxième condition concerne le nombre de sauts sur l'échelle de temps τ . Ce nombre doit être très grand, i.e. $V_i \tau \gg 1$ pour tout $i \in R_C$. Les deux conditions sont automatiquement satisfaites dans la limite $\mathcal{V} \rightarrow \infty$ car N, V_i ont le même ordre de grandeur que \mathcal{V} et γ_i^C, τ ne dépendent pas de \mathcal{V} . Une condition similaire doit être satisfaite par les super-réactions de

type 1 : $V_i \gg \tau^{-1}$ pour tout $i \in \mathcal{S}_1$.

Les super-réactions de type 2 ne satisfont pas la condition de sauts petits, i.e. $|\gamma_i^C|$ doit être comparable à N pour tout $i \in \mathcal{S}_2$.

Pour obtenir une approximation avec switch singulier on introduit un nouveau paramètre $0 < \epsilon \ll 1$ et les conditions suivantes :

a) Il existe de super-réactions de type 3, c'est à dire des super-réactions de type 1 très rapides.

b) Le temps d'action des super-réactions de type 3 est limité par les réactions $\mathcal{R}_{D_i}^-$. Ce $\tau_\epsilon = \epsilon\tau$, i.e. $V_i = (\epsilon\tau)^{-1} \gg \tau^{-1}$ pour au moins un $i \in \mathcal{R}_{D_i}^-$. Les réactions dans $\mathcal{R}_{D_i}^-$ qui inhibent les super-réactions de type 3 peuvent être aussi rapides que les réactions dans \mathcal{R}_C ou dans \mathcal{S}_1 (mais pas aussi rapides que les réactions dans \mathcal{S}_3).

c) Les super-réactions de type 3 sont suffisamment rapides pour produire un saut comparable avec N dans les variables continues, pendant le temps τ_ϵ , i.e. $V_i\epsilon\tau \sim N$ pour tout $i \in \mathcal{S}_3$.

Les conditions du type $V_i \gg \tau^{-1}$ peuvent être simplifiées si les réactions sont de premier ordre par rapport aux espèces discrètes. Dans ce cas $V_i = k_i X_D$ avec $X_D \approx 1$, ainsi la condition devient $k_i \gg \tau^{-1}$. Tous ces critères sont décrits dans le Table2 (voir page 93).

4.4 Conclusion

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

Nous avons montré que des phénomènes d'intermittence peuvent apparaître dans des réseaux de réactions biochimiques, conduisant à un fonctionnement bruité. Ces phénomènes apparaissent lorsque plusieurs conditions sont satisfaites : des conditions sur les concentrations (certaines espèces sont en petits nombres) et des conditions sur les échelles de temps. Selon les différentes conditions discutées on peut rencontrer plusieurs types de comportements hybrides : avec switch (sauts des taux de réaction) et avec breakage (sauts des concentrations).

Chapitre 5

Moyennisation de la cinétique chimique stochastique

Les performances des approximations hybrides peuvent être très mauvaises lorsque le mécanisme de réactions discrètes contient des cycles qui sont au moins aussi rapides que la dynamique déterministe des variables continues. En effet, dans ce cas, l'intervalle de temps entre deux réactions successives du cycle discret rapide est échantillonné inutilement par le solveur déterministe. De plus, le nombre de sauts des variables continues entre deux réactions du cycle discret peut ne pas être grand et l'approximation hybride n'est pas justifiée. Finalement, l'algorithme de simulation passe la plupart de son temps à générer les réactions du cycle rapide. Dans ce cas, utiliser l'algorithme hybride est un mauvais choix, car ses performances sont inférieures à celles de l'algorithme de Gillespie.

Il est donc important d'éliminer les cycles rapides des variables discrètes. Ceci est possible par moyennisation.

Les principes de moyennisation sont largement utilisés pour les systèmes dynamiques déterministes (le lecteur peut consulter les ouvrages classiques de [11],[100], ou les articles plus récents de [12]) et pour les systèmes dynamiques stochastiques (équations différentielles stochastiques [40], chaînes de Markov [114],[110].).

L'idée classique de la moyennisation est d'identifier des variables rapides ergodiques et de moyenner la dynamique des variables rapides par rapport à la distribution quasi-stationnaire des variables rapides. Pour un problème donné, la difficulté principale consiste dans l'identification des variables lentes et rapides. Pour des systèmes dynamiques hamiltoniens les variables action-angle offrent le cadre naturel pour la moyennisation [11].

Dans le cas de la cinétique chimique, il y a deux types de variables lentes qu'on doit considérer pour la moyennisation. Il y a d'abord les composantes lentes (composantes discrètes qui varient rarement ou des composantes continues qui ont un temps de relaxation très long) Le deuxième type de variables lentes sont les combinaisons linéaires de composantes conservées (les lois de conservation) par la dynamique des cycles rapides. Ces variables lentes, qui jouent un rôle similaire aux variables action dans les systèmes hamiltoniens, conduisent à des variables agrégées dans le système moyennisé.

L'agrégation d'états a déjà été utilisée pour les chaînes de Markov [114],[110]. L'agrégation d'espèces qu'on propose ici est adaptée au cas de la cinétique chimique. L'agrégation de variables pour quelques cas particuliers de modèles réactifs est utilisée dans certaines applications, notamment en écologie (voir par exemple [4]). Dans [50],[91] on propose une solution générale pour la simplification des réseaux de réactions, qui s'applique à tout mécanisme mono-moléculaire avec des constantes de réaction bien séparées. L'agrégation est une étape particulière dans cet algorithme qui contient d'autres étapes telles que la chirurgie des cycles. On montre ici que le même algorithme s'applique également pour simplifier les réseaux stochastiques et pour éliminer les cycles rapides.

5.1 Approximations hybrides moyennisées

Le cas le plus général d'approximations hybrides moyennisées implique deux paramètres. Un grand paramètre \mathcal{V} (représentant le volume) ou son inverse qui est un petit paramètre est utilisé pour le développement Kramers-Moyal. Un deuxième paramètre ϵ , petit, est utilisé pour séparer la dynamique rapide à l'intérieur du cycle de la dynamique lente entre les cycles. Dans certains cas, la moyennisation s'impose même avec un seul paramètre asymptotique.

On peut considérer les cas suivants :

- Cas 1 :** Cycles rapides discrets produisant des espèces continues. Dans ce cas il y a des super-réactions du type 1 qui agissent sur les espèces discrètes $\gamma_i^D \neq 0$ for $i \in \mathcal{S}_1$. On peut considérer un seul paramètre $\eta = 1/\mathcal{V}$, petit.
- Cas 2 :** Cycles rapides discrets. Dans ce cas, il y a des transitions rapides entre espèces discrètes. Le petit paramètre est ϵ le rapport entre les échelles de temps de la dynamique intra- et inter-cycles.
- Cas 3 :** Une combinaison de ces deux cas. Il y a deux petits paramètres, η and

ε.

Nous discutons ici uniquement les deux premiers cas.

Cas 1 Les cycles rapides agissent en même temps sur des variables discrètes et continues. Les espèces discrètes changent cycliquement leurs valeurs dans un ensemble fini, leurs nombres restant petits et bornés supérieurement. Les espèces continues sont produites continûment et leurs nombres peuvent croître sans bornes.

Soit $x_c = \eta X_c$ le vecteur de concentrations des espèces continues. Soit X_D^2 les espèces discrètes affectées par des super-réactions du type 1 et X_D^1 les autres espèces discrètes.

Partons de l'équation maîtresse :

$$\begin{aligned}
& \frac{\partial p}{\partial t}(x_c, X_D^1, X_D^2, t) \\
&= \sum_{i \in R_C} \eta^{-1} v_i(x_c - \eta \gamma_i^C) p(x_c - \eta \gamma_i^C, X_D^1, X_D^2, t) - \eta^{-1} p(x_c, X_D^1, X_D^2, t) \sum_{i \in R_C} v_i(x_c) + \\
&+ \sum_{i \in R_D} V_i(X_D^1 - \gamma_i^{D,1}, X_D^2 - \gamma_i^{D,2}) p(x_c, X_D^1 - \gamma_i^{D,1}, X_D^2 - \gamma_i^{D,2}, t) \quad (5.1) \\
&- p(x_c, X_D^1, X_D^2, t) \sum_{i \in R_D} V_i(X_D^1, X_D^2) \\
&+ \sum_{i \in R_{DC}} \eta^{-1} v_i(x_c - \eta \gamma_i^C, X_D^1, X_D^2 - \gamma_i^{D,2}) p(x_c - \eta \gamma_i^C, X_D^1, X_D^2 - \gamma_i^{D,2}, t) \\
&- \eta^{-1} p(x_c, X_D^1, X_D^2, t) \sum_{i \in R_{DC}} v_i(x_c, X_D^1, X_D^2)
\end{aligned}$$

Développons p en série des puissances de η :

$$p(x_c, X_D^1, X_D^2, t) = p_0(x_c, X_D^1, X_D^2, t) + \eta p_1(x_c, X_D^1, X_D^2, t) + \dots$$

Cette série donnera la solution en temps longs de la l'équation maîtresse. En utilisant un développement de Taylor de l'équation (5.1) nous obtenons aux ordres successifs :

A l'ordre η^{-1}

$$\sum_{i \in R_{DC}} v_i(x_c, X_D^1, X_D^2 - \gamma_i^{D,2}) p_0(x_c, X_D^1, X_D^2 - \gamma_i^{D,2}, t) - p_0(x_c, X_D^1, X_D^2, t) \sum_{i \in R_{DC}} v_i(x_c, X_D^1, X_D^2) = 0$$

Ceci signifie qu'à t, x_c, X_D^1 fixés, p_0 est la distribution quasi-stationnaire du

processus défini par les réactions rapides R_{DC} . Considérons que ce processus soit ergodique. Cette dernière condition assure l'unicité de la distribution quasi-stationnaire et on peut écrire :

$$p_0(x_c, X_D^1, X_D^2, t) = \psi(x_c, X_D^1, t)\rho(X_D^2) \quad (5.2)$$

où $\rho(X_D^2)$ est l'unique distribution stationnaire du processus rapide et $\psi(x_c, X_D^1, t)$ est la distribution dépendante du temps des autres variables.

A l'ordre η^0

$$\begin{aligned} \frac{\partial p_0}{\partial t} = & -\nabla_{x_c} \cdot [\chi_c(x_c)p_0] + \\ & + \sum_{i \in R_D} V_i(X_D^1 - \gamma_i^{D,1}, X_D^2 - \gamma_i^{D,2})p_0(x_c, X_D^1 - \gamma_i^{D,1}, X_D^2 - \gamma_i^{D,2}, t) \\ & - p_0(x_c, X_D^1, X_D^2, t) \sum_{i \in R_D} V_i(X_D^1, X_D^2) \\ & - \nabla_{x_c} \cdot \left[\sum_{i \in R_{DC}} v_i(x_c, X_D^1, X_D^2 - \gamma_i^{D,2})\gamma_i^C p(x_c, X_D^1, X_D^2 - \gamma_i^{D,2}, t) \right] + \\ & + \sum_{i \in R_{DC}} v_i(x_c, X_D^1, X_D^2 - \gamma_i^{D,2})p_1(x_c, X_D^1, X_D^2 - \gamma_i^{D,2}, t) \\ & - p_1(x_c, X_D^1, X_D^2, t) \sum_{i \in R_{DC}} v_i(x_c, X_D^1, X_D^2) \end{aligned} \quad (5.3)$$

où $\chi_c(x_c) = \sum_{i \in R_C} v_i \gamma_i^C$.

En utilisant (5.2) et en sommant (5.3) par rapport à toutes valeurs possibles de X_D^2 on obtient :

$$\begin{aligned} \frac{\partial \psi}{\partial t} = & -\nabla_{x_c} \cdot [(\chi_c(x_c) + \bar{\chi}_{DC}(x_c, X_D^1))\psi(x_c, X_D^1, t)] + \\ & + \sum_{i \in R_D} \bar{V}_i(X_D^1 - \gamma_i^{D,1})\psi(x_c, X_D^1 - \gamma_i^{D,1}, t) - \psi(x_c, X_D^1, t) \sum_{i \in R_D} \bar{V}_i(X_D^1) \end{aligned} \quad (5.4)$$

où $\bar{V}_i(X_D^1) = \sum_{X_D^2} V_i(X_D^1, X_D^2)\rho(X_D^2)$, $\bar{\chi}_{DC}(x_c, X_D^1) = \sum_{X_D^2} \sum_{i \in R_{DC}} v_i(x_c, X_D^1, X_D^2)\gamma_i^C$ sont des taux moyens.

Eq.(5.4) montre que l'approximation d'ordre zéro est un PDP moyennisé.

Cas 2 Considérons qu'il y a des réactions rapides R_2 et des réactions lentes R_1 . Considérons aussi que $V_i = \epsilon^{-1}v_i$ pour $i \in R_2$. Comme dans le Cas 1 on considère deux types d'espèces. X_D^2 sont les espèces affectées par les réactions rapides et X_D^1 les autres espèces.

Nous appliquons le même schéma que dans le cas précédent. Nous partons de l'équation maîtresse :

$$\begin{aligned}
\frac{\partial p}{\partial t}(X_D^1, X_D^2, t) &= \sum_{i \in R_1} V_i(X_D^1 - \gamma_i^{D,1}, X_D^2 - \gamma_i^{D,2}) p(X_D^1 - \gamma_i^{D,1}, X_D^2 - \gamma_i^{D,2}, t) \\
&\quad - p(X_D^1, X_D^2, t) \sum_{i \in R_1} V_i(X_D^1, X_D^2) + \\
&\quad + \sum_{i \in R_2} \epsilon^{-1} v_i(X_D^1 - \gamma_i^{D,1}, X_D^2 - \gamma_i^{D,2}) p(X_D^1 - \gamma_i^{D,1}, X_D^2 - \gamma_i^{D,2}, t) \\
&\quad - p(X_D^1, X_D^2, t) \sum_{i \in R_2} \epsilon^{-1} v_i(X_D^1, X_D^2)
\end{aligned} \tag{5.5}$$

Supposons l'ergodicité du processus défini par les réactions rapides. Il vient $p = \psi(X_D^1, t) \rho(X_D^2) + \epsilon p_1(X_D^1, X_D^2, t)$. En ordre zéro, $\rho(X_D^2)$ est la distribution quasi-stationnaire du processus rapide et $\psi(X_D^1, t)$ représente la dynamique moyennisée :

$$\frac{\partial \psi}{\partial t}(X_D^1, t) = \sum_{i \in R_1} \bar{V}_i(X_D^1 - \gamma_i^{D,1}) \psi(X_D^1 - \gamma_i^{D,1}, t) - \psi(X_D^1, t) \sum_{i \in R_1} \bar{V}_i(X_D^1) \tag{5.6}$$

où $\bar{V}_i(X_D^1) = \sum_{X_D^2} V_i(X_D^1, X_D^2) \rho(X_D^2)$ sont des taux moyens.

Lorsque le mécanisme R_2 est composé d'un seul cycle $A_1^2 \rightarrow A_2^2 \rightarrow \dots \rightarrow A_m^2 \rightarrow A_1^2$, sa distribution stationnaire peut être facilement calculée. Cette distribution est paramétrée par la masse totale du cycle $N = \sum_{i=1}^m X_i^2$, qui est une variable lente. Considérons que les constantes de réaction du cycle sont séparées et que la dernière réaction est limitante. Alors, la masse est concentrée avec une probabilité $1 - \sum_{i \neq m} k_{\text{lim}}/k_i$ proche de un, au début de la réaction limitante et avec petites probabilités k_{lim}/k_i au début de la réaction $i \neq m$. La moyenne de X_j^2 par rapport à la distribution stationnaire ρ est $\bar{X}_j^2 = N k_{\text{lim}}/k_j$. En conséquence, le taux moyen de la réaction $A_j^2 \rightarrow B$ de constante de réaction k où B est extérieur au cycle est $\bar{V}_j = k \bar{X}_j^2 = N k k_{\text{lim}}/k_j$.

5.2 Réduction de cycles linéaires

L'argument de la section précédente justifie un algorithme de réduction pour les cycles linéaires rapides d'un mécanisme autrement quelconque. Cette méthode est une extension pour le cas stochastique de la méthode utilisée dans [50],[91] pour le cas déterministe.

Il existe deux types de réseaux linéaires : réseaux de réactions mono-moléculaires et réseaux de réaction de premier ordre. Les réseaux de réactions mono-moléculaires peuvent être complètement définies par un graphe orienté. Dans ce graphe, les sommets sont des espèces chimiques A_i , les arrêts sont des réactions $A_i \rightarrow A_j$ dont les constantes cinétiques sont $k_{ji} > 0$. Pour chaque sommet, A_i , on définit une variable positive réelle c_i (concentration).

La dynamique déterministe d'un réseau de réactions mono-moléculaires est définie par

$$\frac{\partial c_i}{\partial t} = \sum_j k_{ij} c_j - \left(\sum_j k_{ji} \right) c_i, \quad (5.7)$$

Les réseaux de réactions de premier ordre, contiennent des réactions avec un seul réactant et dont les taux de réaction sont proportionnels à la concentration du réactant. Les réseaux de premier ordre peuvent contenir des réactions qui ne sont pas mono-moléculaires, par exemple des réactions du type $A \rightarrow A + B$, ou $A \rightarrow B + C$.

On se restreindra aux réactions de premier ordre pseudo-conservatives, c'est à dire aux réactions qui ne changent pas le nombre total de molécules dans le sous-mécanisme rapide. Les réactions $A \rightarrow A + B$ seront permises, à condition que B soit extérieur au sous-mécanisme ; de même les réactions $A \rightarrow B + C$ seront admises, à condition que B ou C soit extérieur au sous-mécanisme). Avec de telles contraintes, le nombre total de molécules dans le sous-mécanisme est conservé et les équations cinétiques sont les mêmes que celles dans le cas mono-moléculaire (5.7).

Les réactions de dégradation peuvent être étudiées dans ce cadre en incluant une composante spéciale (sink) qui collecte les molécules dégradées. D'autres extensions sont possibles. Par exemple le système peut être ouvert en permettant des termes source constants ou lentement variables dans Eq.(5.7). Ces termes changeront l'état stationnaire, mais n'aurons pas d'impact sur les temps de relaxation du système. Finalement, les sous-mécanismes linéaires peuvent être considérés comme parties d'un système non-linéaire plus grand avec lequel ils sont couplés par des entrées lentement variables (frontière). Par exemple, même pour des systèmes contenant des réactions binaires, on peut définir des réactions pseudo-monomoléculaires quand un des substrats de la réaction binaire ne change pas (ou change très lentement). Cette condition est satisfaite notamment si le substrat est en excès.

Nous allons considérer que les constantes de réactions sont bien-séparées, c'est à dire qu'on ait soit $k_i \ll k_j$, soit $k_i \gg k_j$, pour tout $i \neq j$.

La dynamique stochastique d'une seule molécule dans les réseaux linéaires pseudo-conservatifs est définie par la probabilité $p(j, t)$ que la molécule soit en A_j au moment t . On peut facilement montrer que l'équation maîtresse pour $p(j, t)$ et l'équation cinétique déterministe (5.7) sont identiques. Considérer une seule molécule ne restreint pas la généralité, car lorsque plusieurs molécules sont présentes dans un réseau linéaire, celles-ci se comportent de manière indépendante.

Ainsi, l'algorithme de réduction proposé pour des réseaux de réactions déterministes [91],[50] peut être aussi appliqué à des réseaux stochastiques. L'algorithme est basé sur un ensemble d'opérations transformant le graphe orienté de réactions dans un graphe acyclique sans bifurcation (il n'y a pas de réactions distinctes qui ont le même substrat). Pour ce type de graphe, les vecteurs propres et les valeurs propres de la matrice cinétique peuvent être calculés facilement. Ce calcul résout complètement le problème de la dynamique déterministe. On pourrait utiliser le même schéma pour simplifier et résoudre la dynamique stochastique. Toutefois, on n'utilisera que certaines étapes de l'algorithme, suffisantes pour réduire le temps de calcul de la méthode SSA (algorithme de Gillespie) et pour obtenir des approximations hybrides efficaces.

Rappelons ici l'algorithme introduit dans [91],[50] Cet algorithme utilise deux types d'opérations :

I. Construction d'un réseau auxiliaire (par dominance).

Pour chaque composante A_i du sous-mécanisme linéaire, définissons κ_i comme la plus grande constante cinétique des réactions $A_i \rightarrow A_j : \kappa_i = \max_j \{k_{ji}\}$. Pour la composante j correspondante on utilise la notation $\phi(i) : \phi(i) = \arg \max_j \{k_{ji}\}$. On dit que la plus rapide réaction domine toutes les autres réactions qui ont le même substrat. Le réseau de réactions auxiliaire \mathcal{W} est l'ensemble de réactions $A_i \rightarrow A_{\phi(i)}$ avec constantes cinétiques κ_i . Un réseau auxiliaire ne possède plus de bifurcation car parmi toutes les réactions avec le même substrat on ne garde que la plus rapide.

II. Réduction de cycles (par agrégation). Les cycles rapides sont remplacés par un seul noeud. Les réactions partant du cycle sont remplacées par des réactions avec des constantes rénormalisées selon la recette proposée par la moyennisation dans la section précédente.

Nous présentons l'algorithme de réduction à l'aide de deux exemples simples.

D'abord, considérons une chaîne de réactions $A_1 \rightarrow A_2 \rightarrow \dots A_m$. La constante

de la réaction $A_i \rightarrow A_{i+1}$ est k_i . Toutes les constantes de réaction sont bien séparées.

La plus petite constante de réaction dans la chaîne est appelée limitante et notée par k_{lim} . Si $1/k_{lim} \ll \tau$ (chaîne rapide), alors pendant l'échelle de temps τ toutes les molécules A_1 sont transformées en molécules A_m . On peut ainsi ignorer les réactions de la chaîne et considérer que toute la masse est dans A_m . Le nombre de pas pour transférer la masse au bout de la chaîne est limité par la longueur de la chaîne et par le nombre total de molécules. Ainsi, les chaînes transformant des variables discrètes (en petit nombre) peuvent très bien être simulées par la méthode SSA, sans perte de vitesse de calcul.

En tant que deuxième exemple, considérons le cycle C composé de réactions mono-moléculaires $A_1 \rightarrow A_2 \rightarrow \dots A_m \rightarrow A_1$. Considérons que les constantes des réactions sont bien séparées et soit k_{lim} la plus petite constante, comme dans l'exemple précédent. Ajoutons au cycle une réaction de sortie du cycle : celle-ci transforme une composante du cycle A_j en une composante B externe au cycle.

On considère les situations suivantes :

- (I) la réaction de sortie est $A_j \rightarrow B$ de constante k et $k \gg k_j$,
- (II) la réaction de sortie est $A_j \rightarrow B$ et $k \ll k_j$,
- (III) la réaction de sortie est $A_j \rightarrow A_j + B$, ou
- (IV) la réactions de sortie est $A_j \rightarrow A_{j+1} + B$ de constante k_j .

Dans la première situation (I), la réaction de sortie est rapide et domine la réaction $A_j \rightarrow A_{j+1}$ du cycle. Cette situation a été déjà étudiée [91],[50] comme un cas de "cycle brisé". La réaction $A_j \rightarrow A_{j+1}$ peut être éliminée et ainsi, le cycle est ouvert. La dynamique multi-échelle résultante est celle d'une chaîne.

Dans la situation (II), la réaction de sortie est beaucoup moins rapide que les réactions du cycle. Les molécules de l'intérieur de cycle ont des transformations rapides et la distribution de masse atteint le quasi-équilibre à l'intérieur du cycle.

Dans [91],[50], [49] on montre que le temps de relaxation d'un cycle avec des constantes séparées est l'inverse de la deuxième plus lente constante $k^{(2)} \gg k^{(1)} = k_{lim}$. Pour comprendre ce résultat, il faut s'imaginer les deux chemins équilibrer les différences de masse entre deux points sur un cycle, un chemin qui passe par la plus lente réaction et l'autre chemin qui passe par la deuxième constante la plus lente. Le chemin le plus rapide court-circuite l'autre. Ainsi, un cycle peut être considéré quasi-stationnaire si $k^{(2)} \gg 1/\tau$. Un cycle rapide non-moyennisé coûte cher en temps d'exécution si une molécule peut effectuer un nombre important d'étapes

sur le cycle dans l'échelle de temps déterministe τ . La condition correspondante implique le flux quasi-stationnaire du cycle (et non son temps de relaxation) et s'écrit $k^{(1)} = k_{lim} \gg 1/\tau$.

A la quasi-stationnarité un cycle perd lentement sa masse de manière stochastique, par les réactions de sortie. L'intensité du processus de perte peut être calculée en remplaçant X_j par sa moyenne. La moyenne de X_j par la distribution quasi-stationnaire du cycle est $\bar{X}_j = N(t)k_{lim}/k_j$, où $N(t)$ est la masse totale du cycle $N = \sum_{j=1}^m N_j$. On obtient l'intensité moyennée $\bar{\lambda} = k\bar{X}_j = N(t)kk_{lim}/k_j$.

Dans les situations (III) ou (IV) les intensités moyennées des réactions de sortie sont $k\bar{X}_j = N(t)kk_{lim}/k_j$ et $k_j\bar{X}_j = N(t)k_{lim}$, respectivement.

Toutes ces opérations servent à définir un algorithme de réduction qui élimine les cycles rapides non-brisés.

Algorithme de réduction

Entrée :

mécanisme G des réactions du premier ordre avec constantes bien séparées

Sortie :

un mécanisme simplifié de premier ordre de réactions

tant il y a des cycles rapides non-brisés

pour tout cycle dans G sans réaction de type (I) avec un flux suffisamment rapide ($k_{lim} \gg 1/\tau$) **do**

- remplacer le cycle par un seul nœud C ayant comme masse totale N
- remplacer la réaction de sortie de type (II) $A_j \rightarrow B$ avec le taux de réaction k par la réaction $C \rightarrow B$ de constante effective $k' = kk_{lim}/k_j$.
- remplacer la réaction de type (III) $A_j \rightarrow A_j + B$ de taux k par la réaction $C \rightarrow C + B$ de taux effectif $k' = kk_{lim}/k_j$.
- remplacer la réaction de type (IV) $A_j \rightarrow A_{j+1} + B$ de taux k_j par la réaction $C \rightarrow C + B$ de taux effectif $k' = k_{lim}$.

La condition d'agrégation portant sur le flux du cycle et non sur le temps de relaxation du cycle nous permet d'itérer l'algorithme jusqu'il n'y ait plus de cycle rapide non-brisé. Une condition sur le temps de relaxation créerait des complications. En effet, le temps de relaxation d'un cycle est donné par sa deuxième constante la plus lente. Cela peut cacher une possibilité surprenante : des cycles lents intégrés dans des cycles rapides. Par exemple, un cycle rapide ayant comme nœuds des cycles rapides peut avoir un nœud moins rapide comme origine de sa réaction limitante. Ainsi, des réactions internes de ce nœud sont sans doute plus rapides que la réaction limitante du cycle entier mais ils peuvent être moins ra-

pides que la deuxième constante qui donne le temps de relaxation du cycle entier. Les échelles lentes sont perdues par agrégation. Cette difficulté peut être résolue en complétant l'agrégation par une étape de récupération. Les techniques de récupération sont présentées dans [91],[49]. Dans cette opération, tous les cycles sont restaurés sans leurs réactions limitantes. Dans notre cas, cette opération n'est pas nécessaire car tous les cycles moyennisés sont plus rapides que l'échelle de temps déterministe τ .

5.3 Réseaux bruitées et règles de design pour hiérarchies de cycles

Le bruit intrinsèque des réseaux biochimiques est généré au niveau "microscopique" par les variables discrètes et il est visible au niveau "mésoscopique" des variables continues par des événements comme le "switch" ou "breakage". Ainsi, dans le cas où toutes les espèces moléculaires sont en grande nombre (i.e. un réseau sans variables discrètes) le bruit intrinsèque est relativement faible. De même, si les événements de switch sont plus rapides que l'échelle de temps déterministe, la moyennisation est applicable et le bruit n'est pas transmis aux variables continues. Ainsi, une approximation déterministe est applicable dans les deux situations pour des raisons différentes (grands nombres, transitions rapides).

Le seul moyen de transmettre le bruit par un événement de switch du niveau microscopique au niveau mésoscopique est par le phénomène d'intermittence. Cela implique une combinaison de réactions lentes qui changent les valeurs des espèces discrètes et des réactions rapides qui changent les espèces continues. Le transfert du bruit entre niveaux d'organisation demande des principes de design car seules la présence de molécules en petits nombre et d'échelles de temps lentes et rapides ne suffisent pas à un fonctionnement bruitée du réseau.

Une condition favorable à la possibilité de transfert de bruit est l'organisation hiérarchique du réseau. En particulier les hiérarchies de cycles sont particulièrement intéressantes dans ce contexte. Par hiérarchies de cycles on comprend des mécanismes de réaction contenant des cycles reliés par des réactions de sortie dans des cycles plus grands au niveau supérieur. Génériquement, dans les hiérarchies de cycles, les constantes effectives des réactions de sortie sont plus lentes que les réactions limitantes du cycle d'origine. Coupler des cycles de manière hiérarchique permet d'obtenir des phénomènes de relaxation de plus en plus lentes et

de l'intermittence.

Bien évidemment, certaines situations sont à éviter si on souhaite un fonctionnement bruité. Nous pouvons énoncer quelques règles de design. Ainsi, les réactions de sortie des types (II), (III), mais pas (IV) peuvent générer de l'intermittence lorsque la composante de sortie est au début d'une étape limitante d'un cycle non-brisé. En effet, à moins que k_j soit la constante limitante du cycle, on a $k_{\text{lim}}/k_j \ll 1$. Alors, l'intensité moyenne de la réaction de sortie du type (II) ou de type (III) est faible et représente une source d'intermittence dans le système. Cette situation doit être évitée pour diminuer le bruit ou crée, si on souhaite un fonctionnement bruité.

Chapitre 6

Exemples d'applications

6.1 Méthodologie pour obtenir des simplifications hybrides

Dans cette section on montre comment on peut obtenir, à partir des processus de Markov à sauts associés aux réseaux des gènes, des processus hybrides simplifiés. Les processus hybrides ainsi obtenus préservent les principales propriétés stochastiques des processus de départ.

La procédure de simplification comporte quatre étapes :

1. Identification des variables discrètes et continues ; partition des réactions ;
2. Moyennisation des cycles rapides non-brisés.
3. Identification des super-réactions.
4. Construction du modèle hybride simplifié.

Pour justifier l'utilité de notre approche on montre, à l'aide des exemples, que tous les types de processus hybrides discutés sont représentés dans le fonctionnement stochastique des réseaux des gènes.

Les notations suivantes seront utilisées pour tous les modèles illustrés. Si la i -ème réaction est réversible, alors R_i^+ est la i -ème réaction dans le sens directe et R_i^- représente la i -ème réaction dans le sens inverse. Les réactions irréversibles sont tout simplement représentées par R_i .

Les unités de mesure pour les taux des réactions de premier ordre sont en s^{-1} et en $s^{-1}(nM)^{-1}$ pour les réactions bi-moléculaires. Pour obtenir les taux de réaction pseudo-monomoléculaires à partir des taux des réactions bi-moléculaires

on utilise la formule suivante :

$$k_{mono}[s^{-1}] = k_{binaire}[s^{-1}(nM)^{-1}] \frac{X \text{ [molécules]}}{N_A \mathcal{V}[(nM)^{-1}]}$$

où N_A représente le nombre d'Avogadro. Pour une bactérie, $N_A \mathcal{V} \approx 1(nM)^{-1}$.

De la même manière, les taux de réaction sont obtenus par les formules :

$$V_{mono}[(\text{réactions})s^{-1}] = k_{mono}[s^{-1}] X \text{ [molécules]},$$

$$V_{binaire}[(\text{réactions})s^{-1}] = \frac{k_{binaire}[s^{-1}(nM)^{-1}]}{N_A \mathcal{V}[(nM)^{-1}]} X \text{ [molécules]} Y \text{ [molécules]}.$$

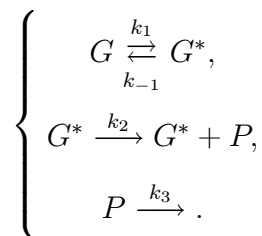
Par la suite on présente deux modèles simples, un modèle avec switch et un modèle avec breakage. Puis, deux modèles plus complexes sont présentés qui illustrent la moyennisation.

6.2 Modèle hybride avec switch : modèle de Cook

L'haploinsuffisance est un terme utilisé lorsque la présence d'un seul exemplaire actif d'un gène donné (ou d'un groupe de gènes) est insuffisante au fonctionnement normal de l'individu. L'haploinsuffisance a lieu quand un organisme diploïde a un seul gène fonctionnel (avec un autre gène inactivé par une mutation) et le fonctionnement d'un seul gène ne produit pas assez de protéine. Le fonctionnement intermittent, stochastique de la production de protéine a été modélisé dans [27].

Dans le modèle de Cook, on considère qu'un gène peut commuter au hasard entre l'état inactif (G) et l'état actif (G^*) suivant des réactions cinétiques de premier ordre ou k_1 est la constante d'activation et k_{-1} est la constante de désactivation qui détermine combien du temps le gène reste dans l'état actif. Quand le gène est actif, il exprime un produit (P) avec la vitesse k_2 . Le produit est dégradé selon une réaction de premier ordre de constante cinétique k_3 .

Le modèle de Cook pour la modélisation du phénomène de haploinsuffisance peut être décrit par les réactions suivantes :



Le système conserve la quantité $G + G^* = G_0$. Dans ce modèle, le régime d'haploinsuffisance correspond à $G_0 = 1$.

Soit $X = (X_D, X_C)$ le vecteur d'état. Dans ce cas on considère la partition suivante : les espèces en grand nombre sont $X_C = \{P\}$ et les espèces en petit nombre $X_D = \{G, G^*\}$. Cette partition des molécules impose naturellement, une partition des réactions. Avec les notations introduites on décrira l'ensemble de réactions impliquant des variables discrètes \mathcal{R}_D , l'ensemble des réactions impliquant les variables continues \mathcal{R}_C , respectivement l'ensemble de réactions impliquant des variables continues et discrètes \mathcal{R}_{DC}

$$\mathcal{R}_D = \{R_1^+, R_1^-\},$$

$$\mathcal{R}_C = \{R_3\},$$

$$\mathcal{R}_{DC} = \{R_2\}.$$

Ainsi, la famille des réactions est $\mathcal{R} = \{\mathcal{R}_D \cup \mathcal{R}_C \cup \mathcal{R}_{DC}\}$.

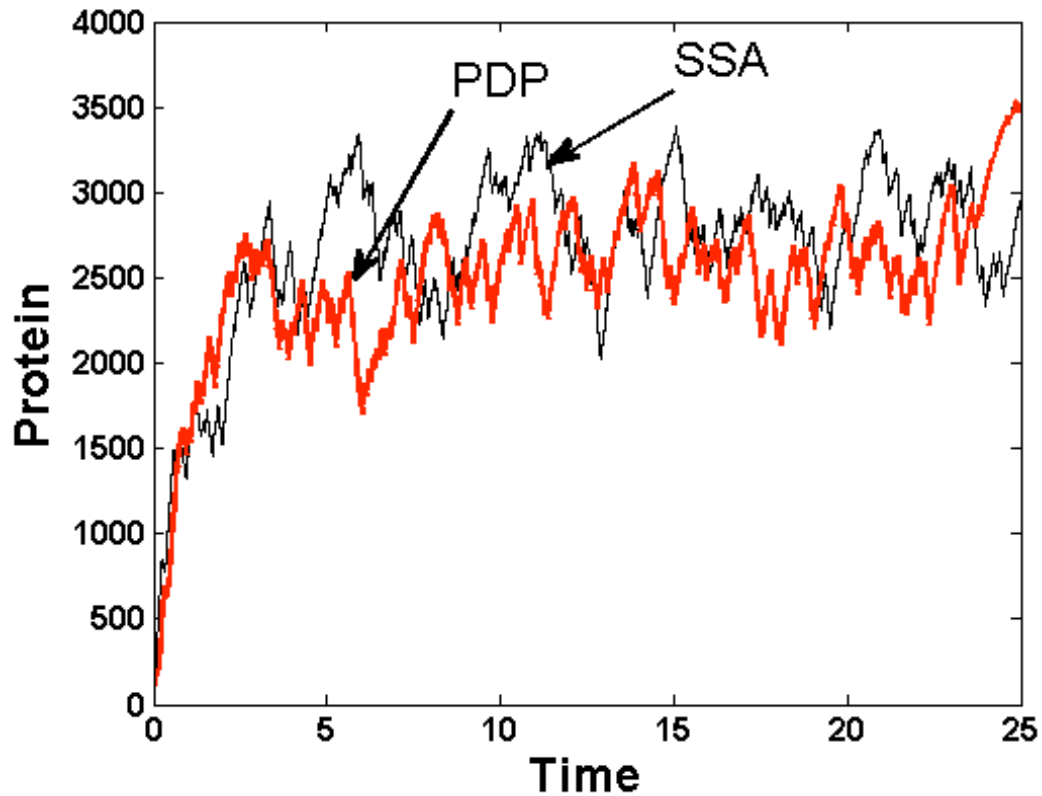


Figure 1a. Modèle de Cook pour haploinsuffisance. L'évolution temporelle de la concentration de la protéine avec la méthode Gillespie et l'approximation PDP

L'échelle de temps déterministe est $\tau = (k_3)^{-1}$. Les vecteurs stœchiométriques sont tous d'ordre un, beaucoup moins importants que le nombre de protéines N .

Pour les valeurs des paramètres utilisées dans [27] on a $k_2/k_3 \gg 1$ et R_2 est une super-réaction de type S_1 . De manière similaire, la réaction R_3 satisfait $V_i = Nk_3 \gg k_3$.

Avec un développement Kramers-Moyal de premier ordre on obtient une approximation PDP avec switch. On dénote par $x = [P]$ la variable continue et par $\theta = G^*$ la variable discrète. Le modèle approché est un PDP avec l'espace d'états $(\theta, x) \in \{0, 1\} \times R$.

Le champs définissant le flot déterministe $\chi(\theta, x)$ est défini par

$$\chi(\theta, x) = -k_3x + k_2\theta,$$

et l'intensité de sauts λ est

$$\lambda_\theta = k_{m1}(1 - \theta) + k_1\theta.$$

Le PDP résultant est identique au modèle ON-OFF utilisé en recherche opérationnelle [20]. Le même modèle a été proposé ailleurs en tant que modèle abstrait de production stochastique de protéine [111],[63].

En utilisant l'algorithme hybride on simule des trajectoires de l'approximation PDP. L'état initial du système utilisé pour illustrer les trajectoires est $x(t_0) = 0, 6$. En plus, on considère que le système est dans l'état inactif (i. e. $\theta = 0$). La même condition initiale est utilisée pour la simulation du modèle initiale avec l'algorithme SSA. Tandis que le temps moyen de simulation en utilisant l'approche stochastique PDP est de 2,6 secondes, le temps moyen de simulation avec l'algorithme SSA est de 14,5 secondes. Les trajectoires obtenues sont similaires du point de vue qualitatif : les intervalles de production sont suivis par des intervalles de dégradation. Toutefois, la variable continue ne présente pas de discontinuités (voir Figure 1a).

Pour tester, du point de vue quantitatif, l'exactitude de l'algorithme PDP, on calcule d'une part la distribution stationnaire du nombre de protéines à partir des trajectoires obtenues avec l'algorithme PDP et, d'autre part, la distribution stationnaire obtenue avec les trajectoires SSA. Cette démarche est justifiée par l'ergodicité du processus exact et de son approximation PDP.

La distribution théorique pour le modèle PDP est une loi beta. Plus précisément, la variable $x = P/(k_2/k_3)$ suit la loi beta $\mathcal{B}(k_1/k_3, k_{m1}/k_3) : p(x) = \frac{1}{B(k_1/k_3, k_{m1}/k_3)} x^{k_1/k_3-1} (1-x)^{k_{m1}/k_3-1}$, où B est la fonction beta [111]. La comparaison entre les différentes distributions est illustrée dans la Figure 1b.

Le modèle de Cook fonctionne dans le région $k_2 \gg k_{m1}$, ce qui correspond à un cycle brisé dans les variables discrètes. Si l'inégalité opposée est valable $k_2 \ll k_{m1}$, alors la moyennisation est nécessaire. Le cycle $G \xrightleftharpoons[k_{m1}]{k_1} G^*$ doit être remplacé par une nouvelle composante \bar{G} de masse $\bar{G} = G_0 = 1$ et la réaction de sortie R_2 dévient $\bar{G} \xrightarrow{k'_2} \bar{G} + P$, où $k'_2 = k_2 \frac{k_1}{k_1 + k_{m1}}$. Le modèle résultant est un modèle de naissance et de mort avec le taux effectif de naissance $k'_2 G_0$ et avec la constante du taux de mort k_3 .

Si $k_2/k_3 \gg 1$ (signifiant que R_2 reste super-réaction du type \mathcal{S}_1), la dynamique du modèle peut être approchée par une dynamique déterministe définie par le champs $\chi(x) = -k_3 x + k'_2 G_0$.

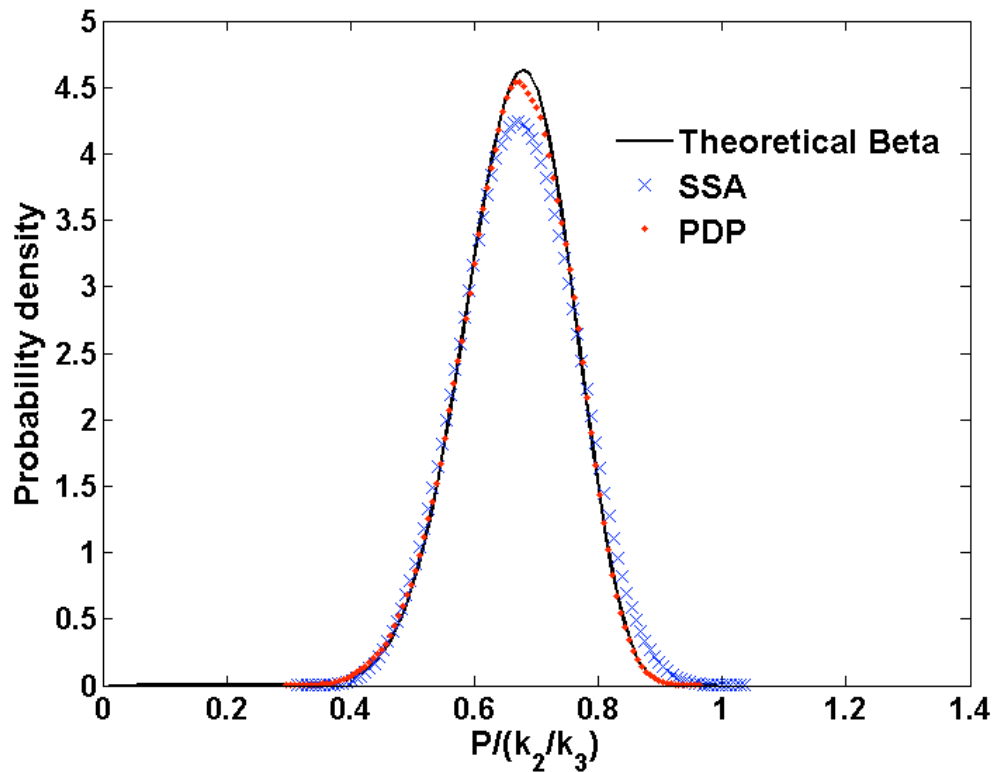


Figure 1b. Distribution de probabilité estimée à partir d'une trajectoire Gillespie vs distribution stationnaire estimée à partir d'une trajectoire PDP. La distribution théorique beta a été ajoutée.

6.3 Modèle hybride avec breakage : neuroscience et site opérateur chez les bactéries

Les neurones fonctionnent de manière stochastique. Ce type de fonctionnement facilite des phénomènes utiles pour leur comportement collectif tels que la synchronisation. Stein [104] propose un modèle simple pour l'évolution du potentiel de membrane neuronal.

Dans ce modèle, le comportement du potentiel de membrane d'un neurone est décrit par un processus de Markov déterministe par morceaux avec breakage. Les potentiel $V(t)$ subit des sauts avec une intensité constante λ et d'amplitude constante.

Dans le modèle proposé par Stein les sauts peuvent être de signe positif

ou négatif correspondant aux excitations ou inhibitions synaptiques. Dans notre exemple on considère seulement des sauts de signe positif. Si t est un instant de saut, alors, $V(t_0^+) - V(t_0^-) = a$. où a est l'amplitude de saut.

Entre deux sauts V décroît exponentiellement suivant la dynamique donnée par $\frac{dV}{dt} = -\alpha V$, où α est une constante.

Le générateur d'un tel processus est :

$$\mathcal{A}f(x) = -\alpha x f'(x) + \lambda(f(x+a) - f(x)).$$

La distribution de probabilité invariante du processus est la solution de l'équation différentielle avec délais suivante

$$\alpha x p'(x) + \alpha p(x) + \lambda(p(x-a) - p(x)) = 0.$$

et $p(x) = 0$, pour tout $x < 0$.

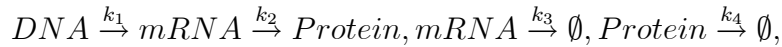
On ne connaît pas de solution analytique pour cette équation avec délais. Pour trouver numériquement la solution sur $[0, \infty)$, on utilise une méthode de différences finies.

La comparaison entre les distributions de probabilité, estimée par la méthode Monte-Carlo et calculée numériquement, est illustrée dans la Figure 2c,d. Dans la Figure 2a,b on montre quelques trajectoires du système.

Le modèle de Stein peut être utilisé aussi pour modéliser la production intermittente d'une composante moléculaire par un réseau de gène. Mieux encore, l'absence d'une période réfractaire, qui est un défaut du modèle en neuro-science, n'est plus un problème pour les réseaux de gènes. Des généralisations de ce modèle, permettant des distributions continues de tailles des sauts pourrait être utilisées pour modéliser l'activité intermittente d'un site promoteur.

Nous avons montré dans une section précédente que des modèles hybrides avec breakage peuvent résulter en tant que limites singulières de modèles hybrides avec switching. Nous allons montrer par des exemples comment passer de modèles moléculaire à des modèles PDP avec breakage. Un exemple typique est le fonctionnement d'un site opérateur chez les bactéries. Ce type de modèle a été discuté par des nombreux auteurs. Nous allons revenir plus loin sur le modèle moléculaire détaillé introduit par [69],[70] pour un site opérateur réprimé. Une version simplifiée de ce modèle a été proposée par [38]. Ce modèle simplifié est défini par les

réactions biochimiques suivantes :



La première réaction est d'ordre zéro, toutes les autres réactions sont de premier ordre. Les paramètres satisfont $k_1 = ak_4$, $k_2 = bk_3$, $k_4 = \epsilon k_3$. On considère que $b \gg 1$, $\epsilon \ll 1$.

A partir de l'aspect des trajectoires (celles-ci présentent des rafales), les auteurs [38] lancent l'hypothèse que le processus peut être approché par un PDP avec breakage. Ils résolvent l'équation de stationnarité et trouvent la distribution de probabilité invariante de la concentration de protéine.

Montrons maintenant que l'équation de Fokker-Planck hybride proposée dans [38] pour la distribution invariante, est une conséquence de notre approche.

Tout d'abord, on remarque que le processus de Markov à sauts peut être approché par un PDP avec switching. La partition des espèces est $X_D = mRNA$, $X_C = Protein$. La composante discrète nRNA suit un processus de naissance et mort avec intensité constante k_1 de naissance et intensité de mort $(k_3 + k_2) \times mRNA$. En utilisant l'équation maîtresse pour des processus de naissance et mort [23] et en considérant que $k_1/(k_3 + k_2) = a\epsilon/(1 + b) \ll 1$, on trouve que la probabilité d'avoir plus d'une molécule de mRNA est beaucoup plus petite que la probabilité d'avoir une seule molécule de mRNA dans le système.

La partition des réactions est $\mathcal{R}_D = \{R_1, R_3\}$, $\mathcal{R}_{DC} = \{R_2\}$, $\mathcal{R}_C = \{R_4\}$. L'échelle de temps de la variable continue est $\tau = (k_4)^{-1}$. $V_2 = k_4 X \gg \tau^{-1}$, où X est le nombre de protéines, à condition que $X \gg 1$. Cette condition, permettant l'application de l'approximation PDP, est satisfaite car $b \gg 1$ (on verra plus loin que la signification de b est le nombre moyen de protéines produites dans une rafale). Ainsi, la réaction R_2 est une super-réaction du type \mathcal{S}_3 .

Pour montrer que le switching est singulier il faut montrer aussi la satisfaction des conditions sur les réactions discrètes rapides. La réaction $R_D^+ = \{R_1\}$ correspond à la transition $0 \rightarrow 1$, tandis que la réaction $R_D^- = \{R_3\}$ correspond à la transition $1 \rightarrow 0$ de la composante discrète $mRNA$. L'intensité de la réaction R_3 satisfait $V_3 = (\epsilon\tau)^{-1}$. Ainsi, la durée moyenne d'une rafale est $\epsilon\tau$. Selon les critères présentés dans le chapitre antérieur, le processus approché est un PDP avec breakage.

La variable discrète $mRNA$ peut être considérée nulle presque tout le temps, sauf pendant la durée négligeable d'une rafale quand $mRNA = 1$.

L'évolution de l'unique variable continue x (concentration de protéine) est décrite par le générateur hybride suivant (obtenu après un simple changement de variables dans l'intégrale de l'Eq.(4.20)) :

$$\mathcal{A}f(x) = -k_4 x f'(x) + k_1 \int_0^\infty [f(x + \tilde{x}) - f(x)] \frac{\nu}{b} \exp[-\tilde{x} \frac{\nu}{b}] d\tilde{x}$$

On peut remarquer que la distribution des amplitudes de breakage est continue (il s'agit d'une loi exponentielle de paramètre $\frac{\nu}{b}$). Le paramètre b représente le nombre moyen de protéines produites dans une rafale (amplitude moyenne d'un breakage). Cette situation est différente du modèle de Stein quand l'amplitude de saut est constante. L'équation de Fokker-Planck stationnaire peut être résolue et la distribution invariante de la variable continue x est la loi Gamma $p(x) = \frac{1}{(b/\nu)^a \Gamma(a)} x^{a-1} e^{-x\nu/b}$ [38].

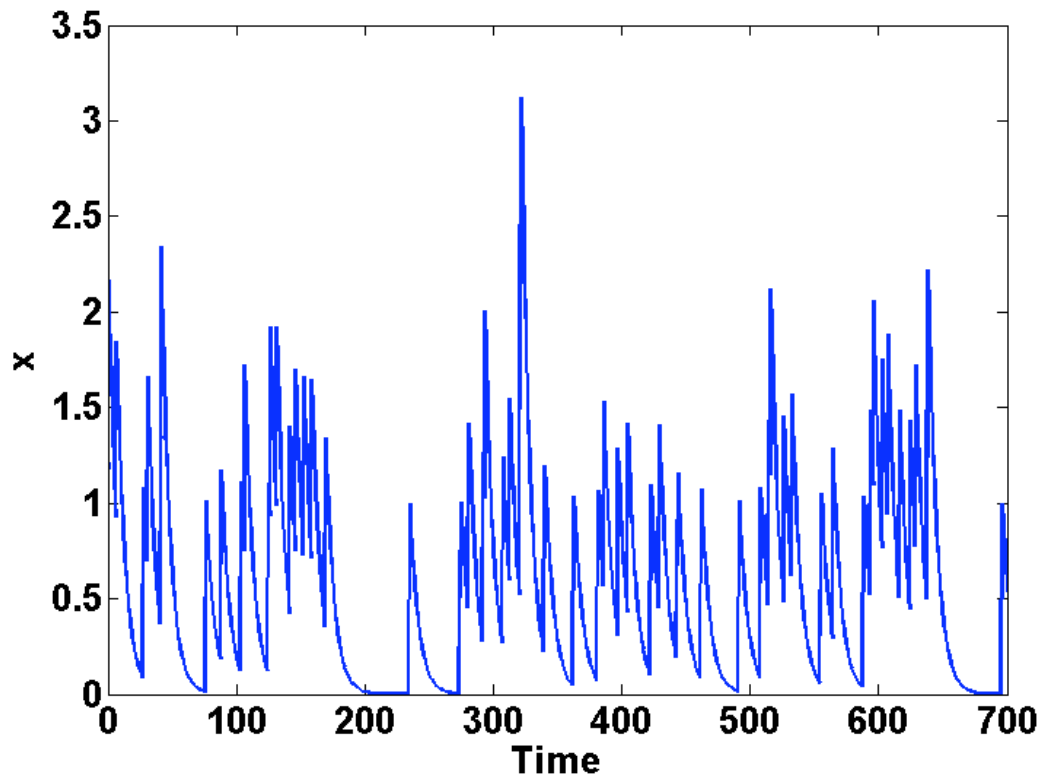


Figure 2a. Modèle de Stein. Trajectoires simulées avec les paramètres $a = 1, \alpha = 0.15, \lambda = 0.1$

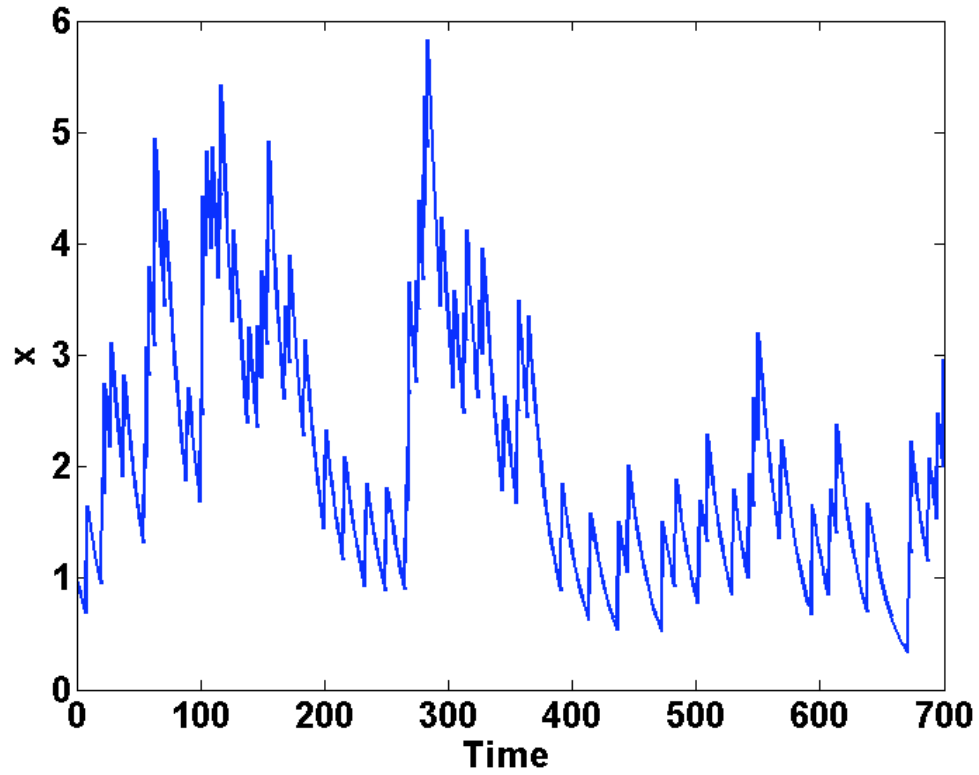


Figure 2b. Trajectoires simulées avec les paramètres $a = 1$, $\alpha = 0.05$, $\lambda = 0.1$.

L'erreur sera définie comme la différence entre la solution approximé et la solution analytique exacte. Deux sources d'erreur sont possibles dans la méthode des différences finies : la perte de précision due aux approximations sur l'ordinateur et l'erreur introduite avec la discrétisation de l'espace.

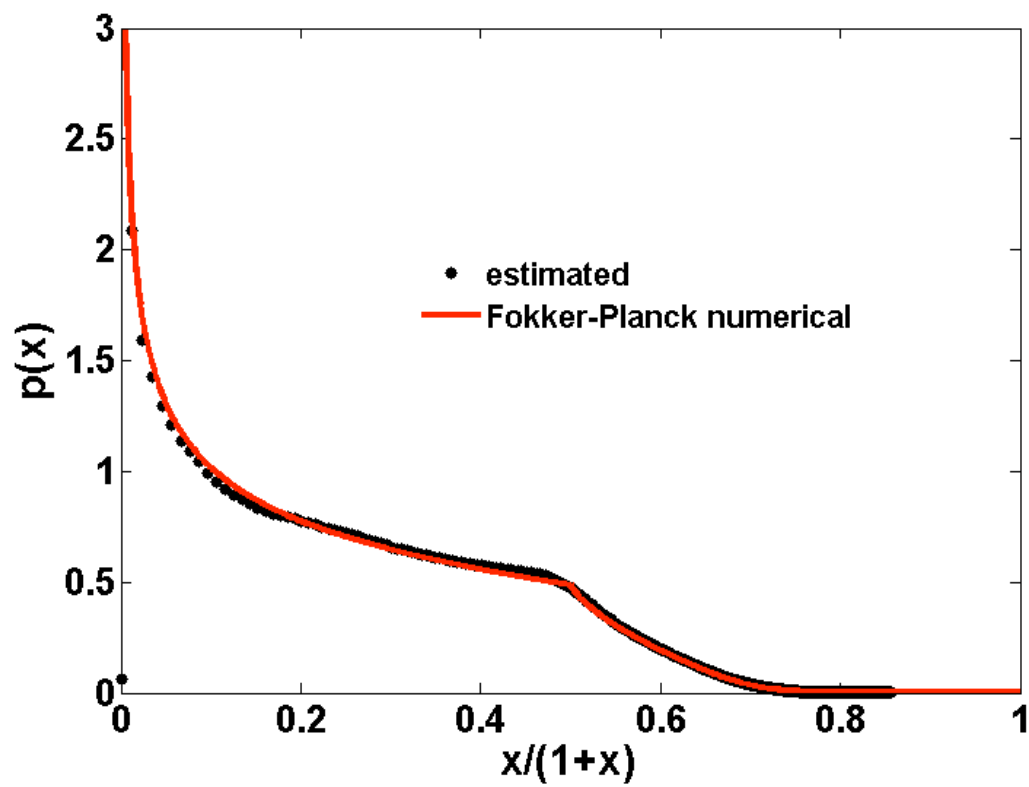


Figure 2c)

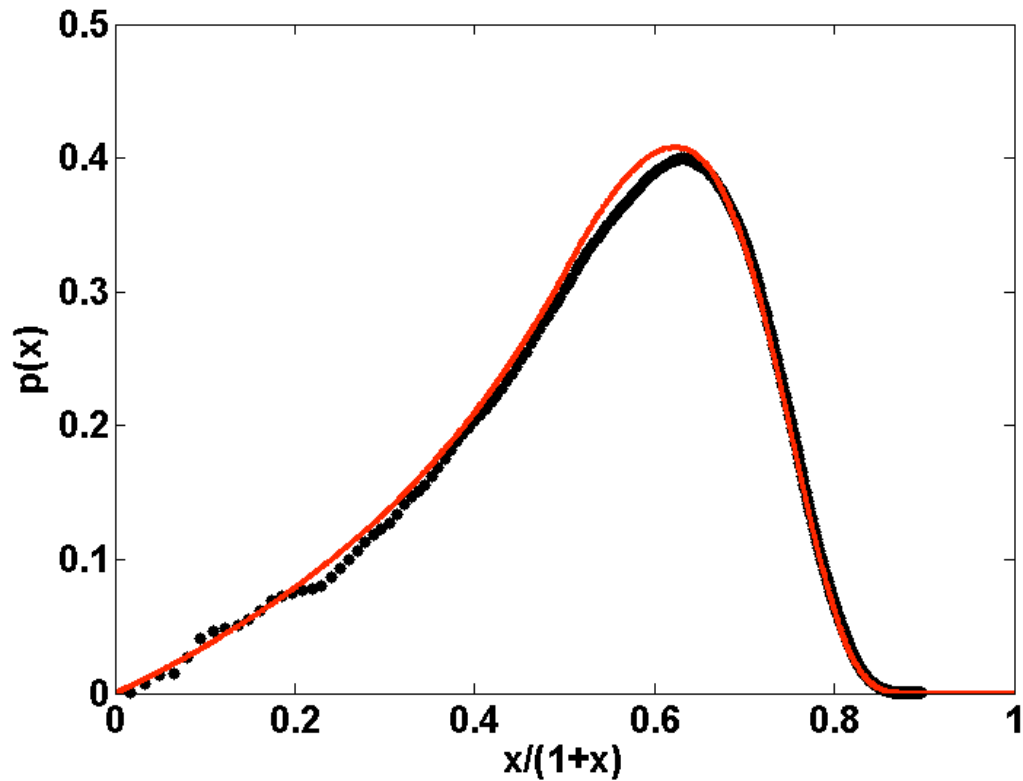


Figure 2 d. Distribution empirique vs distribution estimée pour le modèle de Stein.

Les paramètres utilisés sont c) $a = 1, \alpha = 1, \lambda = 0.1$; d) $a = 1, \alpha = 0.05, \lambda = 0.1$

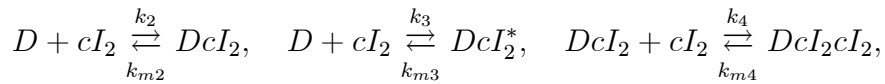
En rouge, on pourra distinguer la distribution de la dynamique du modèle de Stein obtenue à partir de la méthode des différences finies et en noir, la distributions empirique obtenue par les réalisations des trajectoires déterministes par morceaux. La condition initiale utilisée pour la simulation est $X(t_0) = 0.2$ et elle a été la même pour les deux méthodes.

6.4 Premier exemple complexe : lambda-phage

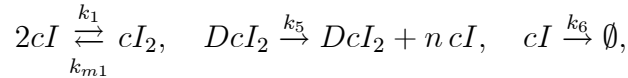
Dans cette section on présente un exemple classique de commutateur moléculaire. Le système biomoléculaire bistable basé sur le répresseur-cro (cI) a été étudié par un nombre considérable d'auteurs [94],[10],[52],[107], etc..

Le cycle de vie des phages peut prendre deux voies distinctes, la voie lysogénique quand le phage se multiplie en même temps que l'ADN de son hôte, et la voie lytique quand le virus produit des grandes quantités de son propre ARNm. Le passage d'un état à l'autre s'effectue suite à un stress extérieur et est contrôlé

par le niveau de protéine cI . La voie lytique correspond à des niveaux importants d'inhibiteur cI , tandis que la voie lysogène correspond à des niveaux faibles de cI . Un dimère cI peut s'attacher à un des trois sites opérateurs dans l'ordre OR1, OR2, OR3. Par coopérativité les sites OR1 et OR2 sont presque simultanément occupés par cI . Le troisième site OR3 sera occupé seulement lorsque la concentration de cI est suffisamment importante. Un modèle plus simple, sans le site OR1 a été proposé dans [52]. Soit cI , cI_2 et D le répresseur, le dimère répresseur et le promoteur non-occupé. La dynamique des sites opérateurs peut être décrite par le système suivant de réactions biochimiques :



où DcI_2 et DcI_2^* sont des complexes sur ADN signifiant l'attachement du dimère inhibiteur au site OR2 ou OR3, ou aux deux sites, respectivement. Les autres réactions représentent la production et la dégradation des molécules cI, cI_2 :



où n est le nombre de protéines traduites par $mRNA$ transcrit.

La propriété intéressante du système phage λ est sa bistabilité. La méthode naïve pour trouver les états stationnaires stables de ce système est de résoudre les équations de stationnarité du système déterministe. Ceci consiste à considérer l'équilibre des réactions réversibles R_i , $i \in [1, 4]$, ainsi que la stationnarité de l'espèce cI . A ces équations il faut ajouter la loi de conservation. On caractérise les solutions par les valeurs x de la concentration cI à stationnarité. $x = 0$ est une des états stationnaires stables, notamment l'attracteur lytique. Cet état est aussi état absorbant du processus de Markov (le système ne peut pas quitter l'état une fois qu'il y arrive). L'état lytique peut être rendu non-absorbant en incluant dans le modèle, une réaction de faible taux de production basale du facteur cI . Cette réaction perturbe faiblement la position de l'état stationnaire mais elle permet de sortir de cet état car elle fonctionne même si la concentration de cI est nulle. Nous n'avons pas considéré ce type de réaction dans notre approche, car le caractère absorbant ou non de l'état lytique n'est pas important pour illustrer les possibilités de notre approche. Lorsque $k_2/k_{m2} = k_3/k_{m3}$ on montre que l'état stationnaire doit satisfaire :

$$\sigma u^4 + 2u^2 - \alpha u + 1 = 0 \tag{6.1}$$

$$\text{où } u = \sqrt{\frac{k_1 k_2}{k_{m1} k_{m2}}} x, \alpha = \frac{n k_5 D_0}{k_6 N_A V} \sqrt{\frac{k_1 k_2}{k_{m1} k_{m2}}}, \sigma = \frac{k_4 k_{m2}}{k_2 k_{m4}}.$$

Rappelons que l'échec de l'approche naïve pour trouver les état stationnaires a été démontré dans le cas de certains systèmes non-linéaires par [78]. Le problème qui peut apparaître est du au fait qu'on impose des équations moyennes d'équilibre sur des espèces qui sont en petit nombre et dont la dynamique continue à la stationnarité aussi (de manière stochastique). Toutefois, on montre que ceci n'est pas un problème pour le modèle étudié ici. Les calculs de stationnarité pourront être justifiés par la moyennisation de cycles de réactions linéaires.

Selon l'approche déterministe la bistabilité apparaît quand l'équation quartique (6.1) possède deux racines positives réelles, ce qu'est le cas lorsque σ est petit et α grand. Plus précisément, le modèle est bistable lorsque $\alpha > \frac{4}{3} \frac{2 + \sqrt{1+3\sigma}}{(1 + \sqrt{1+3\sigma})^{1/4}}$. Comme dans [52] nous avons utilisé $\sigma = 5$. Pour avoir la bistabilité nous avons choisi $\alpha = 7$.

I. Partition des espèces.

Le vecteur d'état est $X = (X_D, X_C)$ où $X_D = \{D, DcI_2, DcI_2^*, DcI_2cI_2\}$, $X_C = \{cI, cI_2\}$. Le choix de variables discrètes est dicté ici, comme dans le premier exemple (modèle de Cook), par les lois de conservation. En effet, on a $D + DcI_2 + DcI_2^* + DcI_2cI_2 = D_0$. Si D_0 est petit (ce qui est le case car il y a un seul promoteur cI sur l'ADN), alors les nombres de $D, DcI_2, DcI_2^*, DcI_2cI_2$ sont petits aussi.

La partition des espèces induit la partition suivante des réactions :

$$\mathcal{R}_D = \emptyset, \quad \mathcal{R}_{DC} = \{R_2^+, R_2^-, R_3^+, R_3^-, R_4^+, R_4^-, R_5\}, \quad \mathcal{R}_C = \{R_1^+, R_1^-, R_6\}.$$

II. Moyennisation des cycles.

Le sous-mécanisme R_{DC} contient des cycles rapides non-brisés qui doivent être moyennisés avant tout autre approximation du système. Ces cycles rapides (voir Figure 3) correspondent à des attachements et déattachements rapides du dimère cI_2 à l'ADN.

On passe du modèle initial Modèle 1 au modèle moyennisé Modèle 2 (voir Figure 3) :

1.1 Le cycle DcI_2, DcI_2cI_2 est non-brisé. Il est agrégé en un seul noeud $\bar{D}cI_2$ dont la masse totale est la somme de DcI_2 et de DcI_2cI_2 .

1.2 L'étape limitante de ce cycle est $k_{lim} = k_{m4} \ll k_4 cI_2$.

1.3 La réaction de sortie $DcI_2 \rightarrow nX + DcI_2$ est remplacée par $\bar{D}cI_2 \rightarrow nX + \bar{D}cI_2$ de constante effective $k'_5 = \frac{k_{m4}}{k_4 cI_2} k_5$. La réaction de sortie $DcI_2 \rightarrow D + cI_2$ est remplacée par $\bar{D}cI_2 \rightarrow D + cI_2$ de constante effective $k'_{m2} = \frac{k_{m4}}{k_4 cI_2} k_{m2}$.

Après moyennisation, on reste avec deux autres cycles dans le Modèle 2. Néanmoins, les ordres des constantes cinétiques des réactions dans ces deux cycles ne sont pas séparés et ne permettent plus l'application des nos approches. De même, les réactions lentes $\bar{D}cI_2 \rightarrow D$, $DcI_2^* \rightarrow D$ produisent de l'intermittence et ne peuvent en aucun cas être moyennisées.

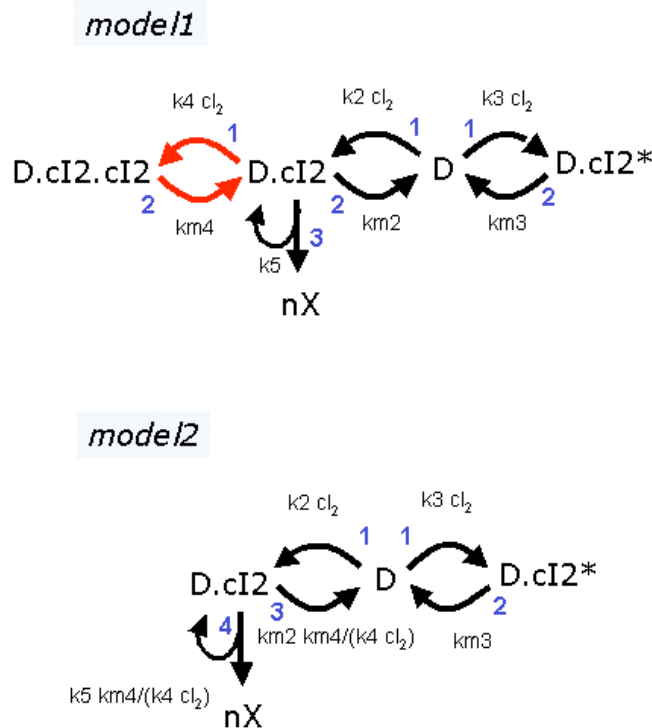


Figure3. Les cycles à moyenniser sont en rouge. Le modèle1 est le système de départ. Le modèle2 est obtenu à partir modèle1 par moyennisation. Les indices entiers représentent les ordres des constantes de réaction (1 indique les réactions les plus rapides).

III. Approximation hybride

Le processus moyennisé peut être approché par un PDP avec switching obtenu par un développement de Kramers-Moyal en premier ordre.

Les réactions R_C et la super-réaction $R_5 \in \mathcal{S}_1$ contribuent au flot déterministe

des variables continues, défini par les équations différentielles suivantes :

$$\begin{aligned}\frac{dx_1}{dt} &= -2k_1x_1^2 - k_6x_1 + 2k_{m_1}x_2 + k_5DcI_2 \\ \frac{dx_2}{dt} &= k_1x_1^2 - k_{m_1}x_2.\end{aligned}\quad (6.2)$$

où les variables continues x_1 et x_2 sont les concentrations de cI et de cI_2 respectivement.

Les autres réactions, produisent les sauts des variables discrètes (switching).

Dans la Figure 3b,c nous avons représenté les trajectoires s'approchant d'un attracteur lysogénique pour le modèle de Markov à sauts non-réduit (Modèle 1) et pour le PDP (Modèle 4). On peut remarquer le comportement intermittent des variables cI , cI_2 qui est bien représentée dans le PDP avec switching.

Les distributions de probabilité invariantes dans le voisinage de l'attracteur lysogénique sont représentées dans la Figure 3d, pour tous les modèles utilisés dans cette étude.

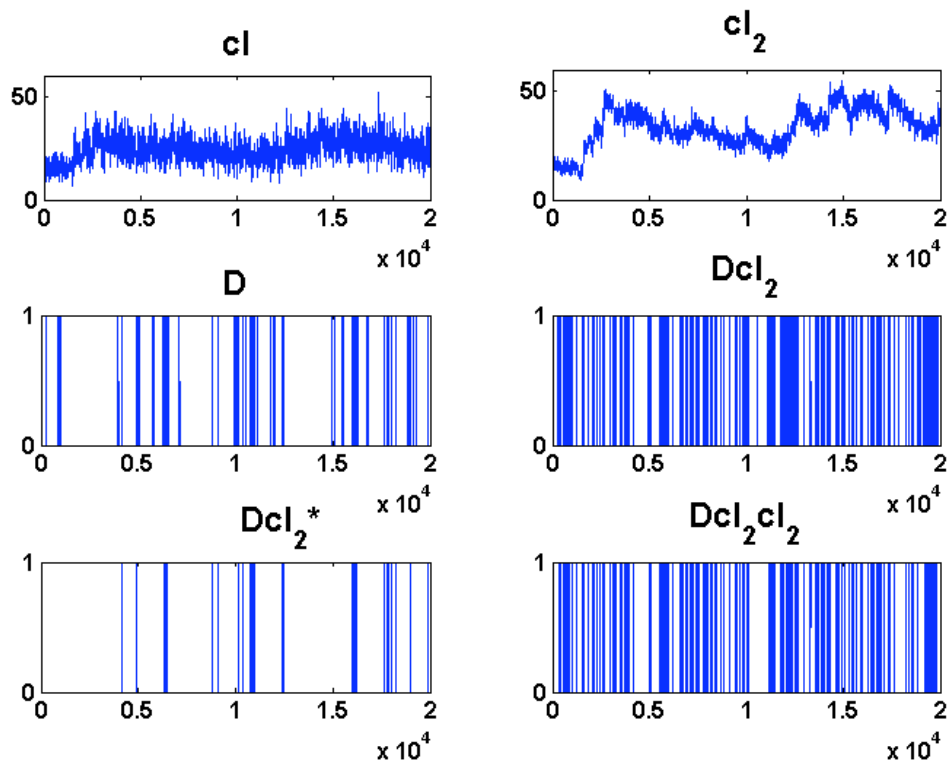


Figure 3b. Trajectoires Gillespie-SSA (Modèle 1)

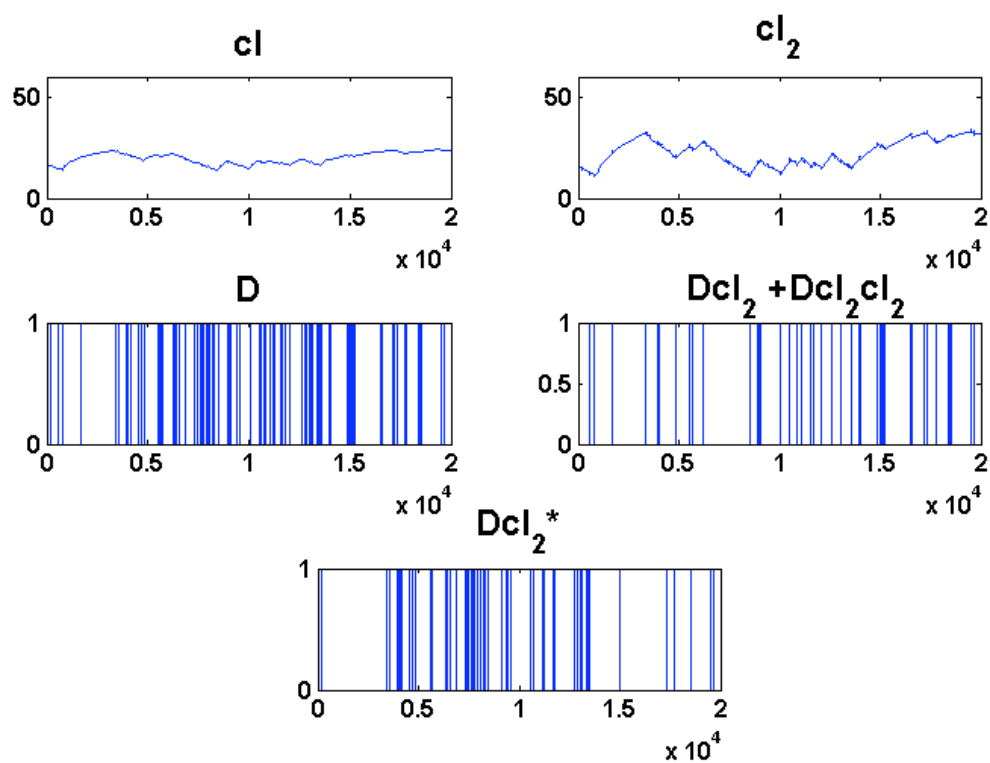


Figure 3c. Trajectories avec l'approximation PDP (modèle4)

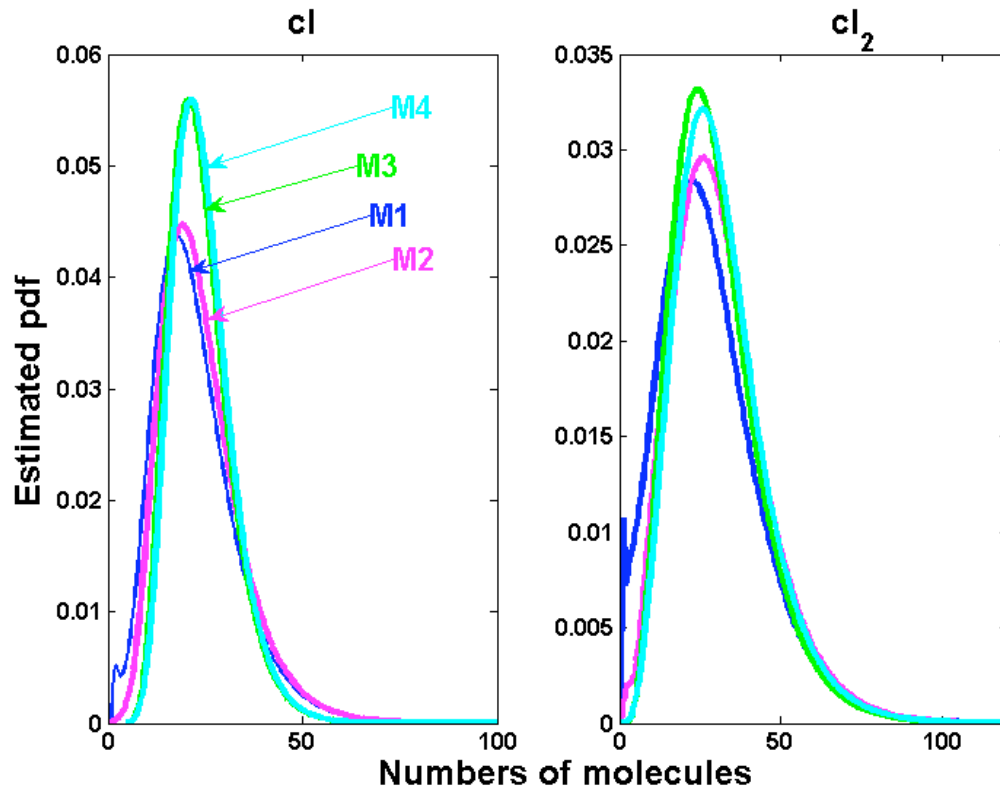


Figure 3d. Distributions stationnaires pour les modèles 1-4.

6.5 Deuxième modèle complexe : fonctionnement en rafales d'un opéron réprimé chez les bactéries

Sous forte répression, la production de protéine par un site opérateur chez une bactérie présente le phénomène de rafales. Un modèle stochastique pour la production d'une protéine chez les procaryotes a été proposé par [69]. Le comportement du site opérateur sous la régulation d'une molécule répresseur a été modélisé par [67]. Le modèle correspondant est présenté dans la Table 1 et dans

la Figure 4a.

Table 1 . Paramètres pour le modèle de [67]

Réaction	Paramètre
Attachement du répresseur et initialisation de la transcription/traduction	
1) $D + R \xrightleftharpoons[km1]{k1} D.R$	$k1 = 10^8 M^{-1} s^{-1}$ $km1 = 1 s^{-1}$
2) $D + RNAP \xrightleftharpoons[km2]{k2} D.RNAP$	$k2 = 10^8 M^{-1} s^{-1}$ $km1 = 10 s^{-1}$
3) $D.RNAP \xrightarrow{k3} TrRNAP$	$k3 = 0.1 s^{-1}$
4) $TrRNAP \xrightarrow{k4} RBS + D + RNAP$	$k4 = 0.3 s^{-1}$
Traduction	
5) $RBS \xrightarrow{k5}$	$k5 = 0.3 s^{-1}$
6) $RBS + Rib \xrightleftharpoons[km6]{k6} RibRBS$	$k6 = 10^8 M^{-1} s^{-1}$ $km6 = 2.25 s^{-1}$
7) $RibRBS \xrightarrow{k7} ElRib + RBS$	$k7 = 0.5 s^{-1}$
8) $ElRib \xrightarrow{k8} Protein$	$k8 = 0.015 s^{-1}$
Repliement et dégradation	
9) $Protein \xrightarrow{k9} FoldedProtein$	$k9 = (\ln 2 / 5400) s^{-1}$
10) $Protein \xrightarrow{k10}$	$k10 = 10^{-5} s^{-1}$
11) $FoldedProtein \xrightarrow{k11}$	$k11 = 10^{-5} s^{-1}$

Dans ce modèle, la bactérie est dans une phase de croissance exponentielle, ce qui signifie qu'elle se divise à des intervalles de temps régulières. La croissance cellulaire est simulée par une croissance linéaire du volume en fonction du temps. Pendant la phase de réplication du cycle cellulaire, le matériel nucléaire est doublé (les variables D , $D.R$, $D.RNAP$ doublent leur nombre). A la fin du cycle cellulaire, le matériel nucléaire est divisé par deux et les autres composantes sont divisées de manière aléatoire. On utilise une loi binomiale pour décrire la répartition du matériel cellulaire entre les deux cellules résultantes.

I. Partition des espèces et des réactions

Tenant compte de la loi de conservation $D + D.R + D.RNAP + TrRNAP = const.$, quelques composantes ($D, D.R, D.RNAP, TrRNAP$) sont en petit nombre. Deux composantes RBS et $RibRBS$ se dégradent vite (le temps de vie de ces composantes est d'environ $2s$) et ne peuvent pas s'accumuler en quantités importantes. Ainsi, les variables discrètes sont :

$$X_D = (D, D.R, D.RNAP, TrRNAP, RBS, RibRBS).$$

Les autres composantes du système sont en grands nombres et forment les variables continues :

$$X_C = (ElRib, protein, Foldedprotein).$$

Finalement, les composantes Rib , $RNAP$ et R sont pratiquement constantes. Cela justifie une approximation du premier ordre pour le mécanisme R_D . L'ensemble des espèces discrètes $D, D.R, D.RNAP$ et $RBS, RibRBS$ forment des cycles non-brisés et ainsi, doivent être moyennisés.

II. Moyennisation de cycles

On applique la procédure de moyennisation trois fois.

1.1 Le cycle $D, D.R$ est non-brisé. Il est remplacé par le nœud D^* ayant comme masse totale la masse de D et $D.R$.

1.2 La constante limitante du cycle est $k_{lim} = k_{m1} \ll k_1$.

1.3 La réaction de sortie $D \rightarrow D.RNAP$ est remplacée par $D^* \rightarrow D.RNAP$ ayant comme constante effective $k'_2 = \frac{k_{m1}}{k_1 R} k_2$.

2.1 Le cycle $D^*, D.RNAP$ est non-brisé. Il est remplacé par le nœud D^{**} ayant comme masse totale la masse de D et $D.R$ et $D.RNAP$.

2.2 Comme $k'_2 \ll k_{m2}$, la constante limitante du cycle est k'_2 .

2.3 La réaction de sortie $D.RNAP \rightarrow TrRNAP$ est remplacée par $D^{**} \rightarrow$

$TrRNAP$ ayant comme constante effective $k'_3 = \frac{k_{m1} k_2 RNAP}{k_{m2} k_{1R}} k_3$.

3.1 Le cycle $RBS, Rib.RBS$ est non-brisé. Il est remplacé par le nœud RBS^* ayant comme masse totale la masse de RBS et de $Rib.RBS$.

3.2 La constante limitante du cycle est $k_{m6} \ll k_6 Rib$.

3.3 La réaction de sortie $Rib.RBS \rightarrow ElRib + RBS$ est remplacée par $RBS^* \rightarrow ElRib + RBS^*$ avec la constante effective $k_7' \approx k_7$.

3.4 La réaction de sortie $RBS \rightarrow$ est remplacée par la réaction $RBS^* \rightarrow$ de constante effective $k_5^* = \frac{k_5 k_{m6} N_{AV}}{k_6 Rib}$.

Le troisième pas de moyennisation implique une perte de précision. La séparation des constante des réactions de sortie et des réactions du cycle n'est pas assez satisfaisante. Ainsi, en théorie on a besoin de $k_7/k_{m6} \ll 1$ alors que $k_7/k_{m6} \approx 0.22$.

III. Identification des super-réactions

Après la moyennisation, on reste avec la réaction lente $D^* \rightarrow TrRNAP$. Cette réaction induit de l'intermittence dans le fonctionnement stochastique du système. Le mécanisme réduit est représenté dans Figure 4b. La partition des espèces discrètes/ continues dans le mécanisme réduit est la même que dans le modèle initial. Dans le mécanisme réduit, la réaction $RBS^* \rightarrow RBS^* + ElRib$ est très rapide, encore plus rapide que les réactions continues $ElRib \rightarrow Prot$, $Prot \rightarrow FoldedProt$, $FoldedProt \rightarrow \emptyset$. En conséquence, on peut la considérer comme une super-réaction de type S_3 .

IV. Approximation hybride

Le développement de Kramers-Moyal au premier ordre appliqué au processus de Markov moyennisé conduit à un PDP avec switching. Un développement de second ordre peut être effectué si une précision supplémentaire est nécessaire (dans ce cas on obtient une diffusion hybride).

Les variables continues évoluent selon la dynamique :

$$\begin{aligned} \frac{d}{dt} ElRib &= k_7' - k_8 ElRib \\ \frac{d}{dt} protein &= k_8 ElRib - k_9 protein \\ \frac{d}{dt} Foldedprotein &= k_9 protein - k_{11} Foldedprotein \end{aligned} \quad (6.3)$$

Les trois variables discrètes ($D^*, TrRNAP, RBS^*$) forment un processus de Markov à sauts. Dans ce contexte, il existe une chaîne rapide de production de $TrRNAP$ passant par RBS^* et jusqu'à $ElRib$. La réaction de production pour

TrRNAP alimente cette chaîne et est extrêmement lente. Ainsi, la variable discrète RBS^* est différente de zéro seulement rarement et pour une période très courte. Ceci conduit à la production par rafale de *ElRib*. Les parties croissantes des trajectoires de *ElRib* sont raides et de courte durée, car la variable discrète RBS^* change très rapidement d'état et devient nulle par la réaction $RBS^* \rightarrow$. Les courts intervalles de croissance sont suivies d'une relaxation exponentielle plus lente. Ainsi, le comportement de *ElRib* est un exemple de PDP avec switching dans lequel la croissance peut être assimilée avec un breakage (see Figure 5a).

Pour conclure, on a traité cinq modèles : un modèle de Markov non-réduit, deux modèles de Markov à sauts réduits (le deuxième modèle obtenu après les étapes de moyennisation 1 et 2, le troisième modèle après l'application de toutes les trois étapes) et deux modèles PDP avec switching (obtenus chacun à partir d'un des deux modèles de Markov à sauts moyennisés). Tandis que, les deux premiers modèles sont simulés avec l'algorithme Gillespie, les modèles hybrides sont simulés avec l'algorithme PDP.

Pour comparer les performances des modèles (en termes de complexité temporelle) on représente les intensités de sauts totales comme fonctions du temps sur une trajectoire. Le modèle le plus rapide, du point de vue computationnel, est le modèle avec la plus petite intensité de sauts. Dans la Figure 5b, on note une décroissance de quelques ordres de magnitude de cette intensité pour le modèle moyennisé par rapport à l'intensité totale du modèle Gillespie non-réduit.

La même méthode peut fournir une estimation erronée pour le temps de calcul dans le cas de l'algorithme PDP. Ainsi, même si le temps de calcul sur l'ordinateur dépend du nombre d'opérations à faire, le nombre d'opérations à faire dans la partie déterministe de l'algorithme dépend du type de solveur utilisé. Si une précision élevée n'est pas demandée, on peut réduire le temps de calcul en choisissant des solveurs moins précis. Pour réduire le nombre d'opérations par pas de temps nous avons privilégié les schémas à un pas (même si ceci n'est important pour les grands systèmes où la principale difficulté est le calcul de la fonction définissant le flot) Nous avons également favorisé les solveurs implicites pour des problèmes raides, qui peuvent fonctionner avec des grands pas de temps. Tous les calculs ont été effectués avec le solveur *ode23* de MATLAB (un solveur implicite pour des problèmes raides qui utilise la formule de Rosenbrock d'ordre 2). Une comparaison du temps de calcul nécessaire pour calculer les trajectoires pour 20 cycles cellulaires est présentée dans la Table 3. Cette comparaison montre les avantages de chaque approximation. La moyennisation permet une réduction impressionnante

des temps de calcul. Ainsi, on a un temps de calcul de presque 50 fois moins important pour un répresseur en faible concentration et une réduction d'ordre 10^4 pour un répresseur en forte concentration (dans ce cas, les effets stochastiques sont plus fortes). Le développement de Kramers-Moyal, conduisant à une approximation PDP permet d'obtenir une réduction supplémentaire d'environ deux fois du temps de calcul. Cette réduction n'est valable que si le répresseur est présent dans une faible concentration et après la moyennisation des cycles rapides.

Un modèle hybride qui n'a pas été encore moyennisé ou qui a été moyennisé que partiellement peut avoir des performances très mauvaises. Par exemple, le modèle PDP M4 est moyennisé partiellement et la performance de ce modèle est moins bonne que celle du modèle Gillespie non-réduit. Une des raisons pour avoir une telle situation est liée à la discrétisation beaucoup trop fine de la partie déterministe.

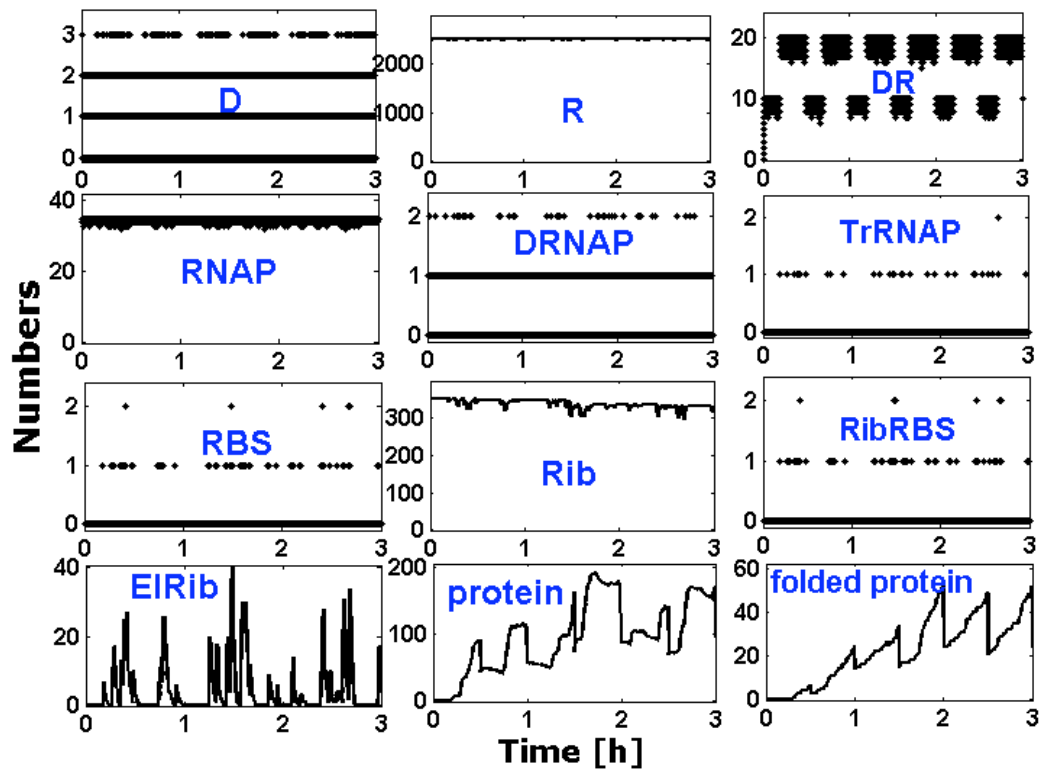


Figure 5a. Trajectoires dans le modèle d'expression de gènes pour un répresseur $R=2500$.

Pour des tests plus quantitatifs des approximations proposées on calcule la distribution stationnaire de la protéine et de la protéine repliée en utilisant les tra-

jectoires obtenues avec les cinq modèles. La distribution obtenue avec le modèle de Markov à sauts est considérée comme exacte. On note une meilleure précision du second modèle par rapport au troisième et quatrième modèle (Figure 5c). Toutefois, même si du point de vue quantitatif elles sont moins exacts, le troisième et le quatrième modèles fournissent une représentation qualitative correcte des caractéristiques stochastiques des trajectoires (rafales). Tous les modèles reproduisent correctement le phénomène de rafale du système initial. Un avantage important du modèle simplifié consiste dans la réduction du nombre de paramètres. Le modèle initial SSA contient 14 paramètres. Après deux pas de moyennisation, le nombre de paramètres est réduit à 9 et après encore un pas de moyennisation on arrive à un modèle avec seulement 7 paramètres. Le PDP a exactement le même nombre de paramètres que le modèle Markov correspondant, c'est-à-dire 9 et respectivement 7 paramètres. En plus, les paramètres des modèles simplifiés sont de monômes de paramètres du modèle non-réduit. Cette observation peut être utilisée dans l'analyse de la sensibilité. Ainsi, après l'identification de des monômes critiques pour une propriété d'intérêt il faut rebrousser le chemin vers les paramètres du modèle initial. L'influence du paramètre initial sur la propriété considérée est spécifiée par la puissance du paramètre correspondant dans le monôme critique. Cette méthode est décrite dans [91].

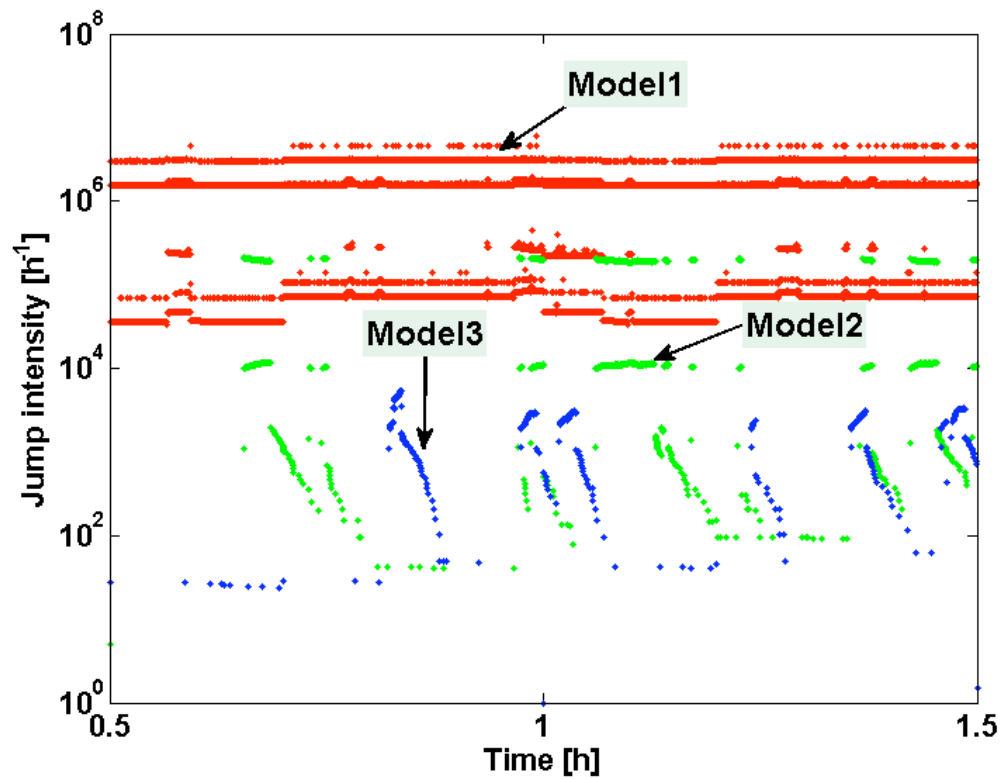


Figure 5b. Intensités de sauts pour les trois modèles SSA pour un répresseur $R=2500$. Modèle1 (sans moyenisation), Modèle2 (moyenisation D, D.R, D.RNAP), Modèle3 (moyenisation D, D.R, D.RNAP, RBS, RibRBS).

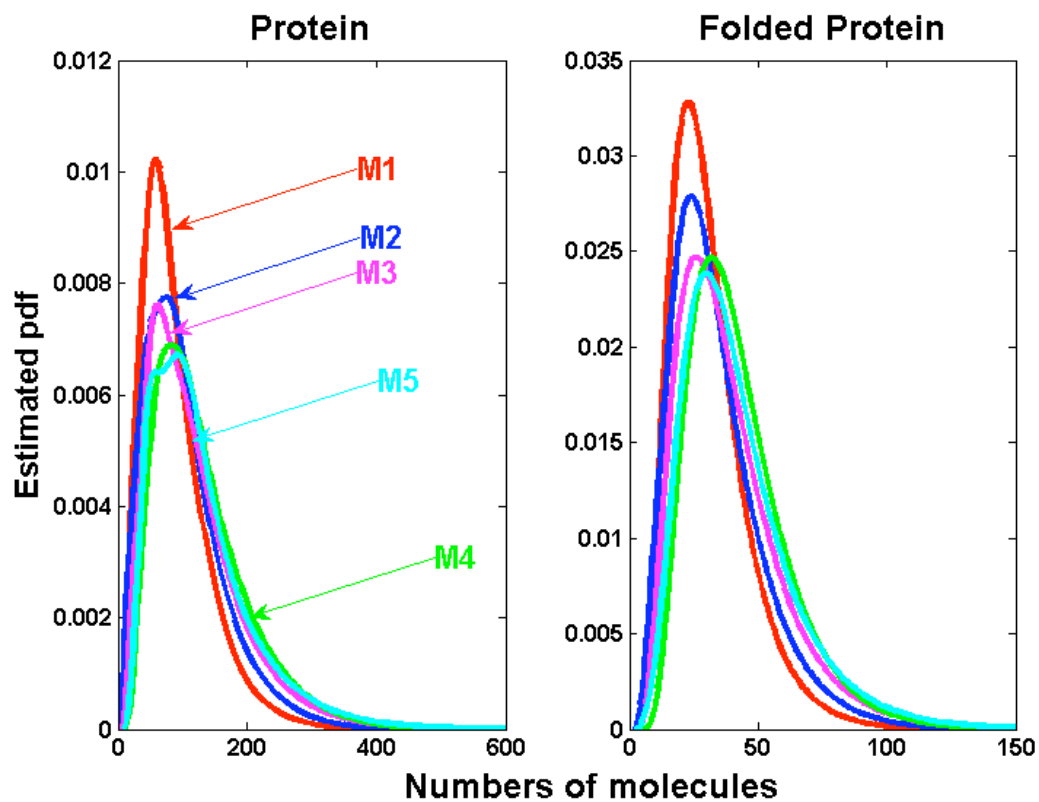


Figure5c. Distribution stationnaire pour la protéine et pour la protéine repliée avec les cinq versions de modèle, $R=2500$. M3 et M4 sont des approximations PDP obtenues des modèles M2 et M3.

6.5.1 Table 2 - Critères pratiques pour différentes approximations

Ensemble de réactions	Approximation	Conditions
R_C	hybride	$\gamma_i^C \ll N$ et $V_i \gg (\tau)^{-1}$
S_1	hybride avec switching	$\gamma_i^C \ll N, V_i \gg \tau^{-1}$
S_2	hybride avec breaking	$\gamma_i^C \sim N$
$R_{D_i}^-$	hybride avec breaking comme limite singulière	$V_i = (\epsilon\tau)^{-1} \gg \tau^{-1}$
S_3	hybride avec breaking comme limite singulière	$V_i \gg (\epsilon\tau)^{-1}$
cycle de réactions	hybride moyennisée	$k_{lim} \gg (\tau)^{-1}$
réactions de sortie	hybride moyennisée	$k \ll k_j, j$ cycling

Table 3 - Temps d'exécution [s] pour 20 cycles de division cellulaire pour les modèles M1-M5 ($3.6 \cdot 10^4$ s temps biologique).

Le modèle M4 est obtenu à partir du modèle M2 et le modèle M5 est obtenu à partir de modèle M3, avec un développement Kramers-Moyal.

R	M1-SSA	M2-averaging+SSA	M3-averaging+SSA	M4-PDP ode23s	M5-PDP ode23s
50	$7.4 \cdot 10^3$	$2.9 \cdot 10^3$	$1.5 \cdot 10^2$	$1.1 \cdot 10^4$	$6.7 \cdot 10^1$
100	$6.8 \cdot 10^3$	$6.4 \cdot 10^2$	$4.0 \cdot 10^1$	$3.7 \cdot 10^3$	$3.6 \cdot 10^1$
200	$5.4 \cdot 10^3$	$1.6 \cdot 10^2$	$1.4 \cdot 10^1$	$1.6 \cdot 10^3$	$1.9 \cdot 10^1$
300	$4.9 \cdot 10^3$	$8.4 \cdot 10^1$	8.5	$8.0 \cdot 10^2$	$1.3 \cdot 10^1$
2500	$3.7 \cdot 10^3$	3.8	$8.3 \cdot 10^{-1}$	$1.0 \cdot 10^2$	4.0
5000	$3.6 \cdot 10^3$	$1.5 \cdot 10^{-2}$	$2.5 \cdot 10^{-1}$	$4.7 \cdot 10^1$	3.0
10000	$3.6 \cdot 10^3$	$1.5 \cdot 10^{-2}$	$9.7 \cdot 10^{-2}$	$3.1 \cdot 10^1$	1.7

Annexes

Méthode numérique pour calculer la distribution de probabilité pour le modèle de Stein

Le générateur d'un tel processus est :

$$\mathcal{A}f(x) = -\alpha x f'(x) + \lambda(f(x+a) - f(x)).$$

L'opérateur adjoint associé au générateur est

$$\mathcal{A}^*f(x) = \alpha x f'(x) + \alpha f(x) + \lambda(f(x-a) - f(x)).$$

On sait par la théorie générale des processus de Markov que la distribution invariante est liée de la solution de l'équation intégro-différentielle suivante

$$\mathcal{A}^*f(x) = 0.$$

Pour trouver la solution sur $[0, \infty)$, on utilise une méthode numérique de différences finies. La méthode des différences finies consiste à remplacer les dérivées de l'équation différentielle par des différences finies et ensuite les approximer. On doit réduire le problème à une forme équivalente posée sur l'ensemble compact $[0, 1]$. On fait le changement de variable

$$z = \frac{x}{1+x},$$

et l'équation équivalente à résoudre s'écrit

$$\alpha g(z) + \alpha z(1-z)g'(z) + \lambda \left[g\left(\frac{z+a(1-z)}{1+a(1-z)}\right) - g(z) \right] = 0,$$

pour tout $z \in [0, 1]$. On utilise une discrétisation équiuniforme de l'intervalle $[0, 1]$ ($z_i = ip, p = 1/n, n \in \mathbb{N}^*$). On doit résoudre le système linéaire

$$\alpha i(1-ip) [g((i+1)p) - g(ip)] + \alpha g(ip) + \lambda \left[g\left(\frac{i + \frac{a}{p}(1-ip)}{1 + a(1-ip)}p\right) - g(ip) \right] = 0,$$

pour tout $0 \leq i \leq n-1$. On désigne par $X_i = g(ip)$ et le système à résoudre devient

$$[-\alpha i(1-ip) + a - \lambda]X_i + \alpha i(1-ip)X_{i+1} + \lambda X_j = 0,$$

pour tout $i \leq n-1$ (où l'indice j est la partie entière de la quantité réelle $\frac{i+\frac{\alpha}{p}(1-ip)}{1+a(1-ip)}$).

Méthode numérique pour estimer la distribution stationnaire

La distribution invariante peut être estimée à partir des réalisations du processus déterministe par morceaux étudié avec une méthode de type Monte Carlo. Tout d'abord, on considère un horizon de temps fixé $TMAX$. Une simulation pour obtenir une trajectoire est effectuée. Pour réduire l'erreur la première partie de la trajectoire $[0; TMIN]$ est éliminée. On considère une discrétisation de l'intervalle de temps $[TMIN, TMAX]$

$$t_0 = TMIN \leq t_1 < \dots < t_N \leq TMAX.$$

La trajectoire est constituée d'une série de moments et des états correspondantes x_0, \dots, x_N . Dans le cas où le processus est constant entre deux instants de sauts t_i et t_{i+1} la distribution peut être tracée par un histogramme de valeurs x_i avec les poids $t_{i+1} - t_i$. Cette procédure est convenable pour les processus de saut pur qui sont constants par morceaux. Pour le processus déterministe par morceaux cette hypothèse n'est plus satisfaite. Ainsi, dans le cas des processus déterministes par morceaux dans le cadre donné par la règle du trapèze x_i doit être remplacé par $(x_i + x_{i+1})/2$. Pour réduire le temps de calcul de l'ordinateur on choisit parfois de re-échantillonner la trajectoire en utilisant une discrétisation équidistante et une fonction d'interpolation spline.

Simulation cellulaire dans le modèle pour l'expression de gènes pro-caryotes

- un cycle cellulaire a une durée de $T = 1800s$ [67]
- Volume : croissance linéaire, $V = 10^{-15}L$
- Le niveau d'expression de gène trouve un équilibre après 12 générations

Condition Initiale :

- la simulation commence avec un seul promoteur D ;
- la condition initiale pour RNAP est de **35** et le nombre de ribosomes est fixe à **350**.
- on fixe toutes les autres molécules à zéro.

Protocol de Simulation

On utilise un algorithme de simulation Gillespie :

1. Simulation Gillespie jusqu'au temps T ,
2. Au temps T une division cellulaire est simulée. La simulation continue avec le nombre de molécules divisé par 2 et avec un seul promoteur. Le volume de la cellule est également divisé par 2.

Deuxième partie

Partie mathématique

Chapitre 7

Les processus déterministes par morceaux

Les processus déterministes par morceaux interviennent dans de nombreuses modélisations. Dans ce travail, nous nous intéressons à leurs applications à la biologie. On trouvera de nombreuses autres applications dans le livre [31]. Ce livre propose d'ailleurs une étude approfondie de ces processus. Ce chapitre présente divers aspects de cette théorie, nécessaires à notre étude.

Dans la suite, si F est un espace métrique, on notera $\mathcal{B}_b(F)$ (resp. $C_b(F)$) l'espace des fonctions boréliennes (resp. continues) et bornées sur F . Si F est un espace vectoriel normé, $C_b^k(F)$ est l'espace des fonctions bornées de classe C^k sur F telles que toutes les dérivées jusqu'à l'ordre k sont bornées.

7.1 Définition et construction

Nous considérons ici des processus déterministes par morceaux - PDP pour "piecewise deterministic process" dans la suite - de variables d'état

$$x = (y, \nu) \in E = \mathbb{R}^n \times \mathbb{N}^d.$$

Leur évolution est déterminée par une famille de champs de vecteurs $(\chi_\nu)_{\nu \in \mathbb{N}^d}$ de classe C^1 sur \mathbb{R}^n , une intensité de saut $\lambda : E \rightarrow \mathbb{R}^+$ et une mesure de transition Q définie sur E et à valeur dans les mesures de probabilités sur E .

Nous ferons systématiquement l'hypothèse suivante sur les caractéristiques du PDP.

Hypothèse 1 Pour tout $\nu \in \mathbb{N}^d$, χ_ν est de classe C^1 et l'équation différentielle

$$\begin{cases} \frac{dy}{dt} = \chi_\nu(y), \\ y(0) = y_0 \end{cases}$$

a une unique solution globale pour tout $y_0 \in \mathbb{R}^n$. L'intensité de saut λ est bornée sur E et de classe C^1 par rapport à la variable $y \in \mathbb{R}^n$. La mesure de transition Q vérifie

1. $x \mapsto Q(A, x)$ est mesurable pour tout borélien A de E .
2. $Q(\{x\}, x) = 0$ pour tout $x \in E$.
3. Pour tout $\psi \in \mathcal{B}_b(E)$ de classe C^1 par rapport à la variable y et $\nu \in \mathbb{N}^d$, $y \mapsto \lambda(y, \nu) \int_E \psi(z) Q(dz, (y, \nu))$ est de classe C^1

On note alors ϕ_ν le flot associé à cette équation différentielle et on sait que $\phi_\nu : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ est de classe C^1 .

Remarque 2 En pratique, c'est souvent l'évolution de y qui est intéressante à étudier.

Pour certains résultats nous aurons besoin de l'hypothèse plus forte suivante.

Hypothèse 3 Les fonctions χ_ν , $\nu \in \mathbb{N}^d$, λ et $x \mapsto \lambda(x) \int_E \psi(z) Q(dz, x)$, où $\psi \in \mathcal{B}_b(E)$ est de classe C^1 par rapport à la variable y , sont bornées ainsi que leurs dérivées par rapport à y .

Lorsque l'hypothèse 3 sera satisfaite, on notera

$$\begin{aligned} M &= \sup_{\nu \in \mathbb{N}^d} \|\chi_\nu\|_\infty, & L &= \sup_{\nu \in \mathbb{N}^d} \|D_y \chi_\nu\|_\infty, \\ M_\lambda &= \|\lambda\|_\infty, & L_\lambda &= \|D_y \lambda\|_\infty. \end{aligned}$$

et L_Q une constante telle que pour $\psi \in C_b^1(E)$:

$$\left\| D_y \left(\lambda(x) \int_E \psi(z) Q(dz, x) \right) \right\|_\infty \leq L_Q (\|\psi\|_\infty + L_\psi)$$

où L_ψ est la constante de Lipschitz de ψ par rapport à la variable y .

Un PDP de caractéristiques $(\chi_\nu)_{\nu \in \mathbb{N}^d}$, λ , Q et démarrant en $x_0 = (y_0, \nu_0)$ évolue de la façon suivante : y évolue suivant le flot ϕ_{ν_0} sur $[0, T_1)$ où T_1 désigne le premier temps de saut et a pour taux λ . A l'instant T_1 , la variable d'état saute suivant la mesure Q . Ensuite, on recommence la même évolution.

Pour construire mathématiquement un tel processus, on procède de la façon suivante. On se donne deux suites $(U_k)_{k \in \mathbb{N}}$ et $(V_k)_{k \in \mathbb{N}}$ de variables aléatoires indépendantes uniformes sur $[0, 1]$ sur l'espace de probabilité canonique $(\Omega, \mathcal{F}, \mathbb{P})$. Pour un point de départ $x_0 = (y_0, \nu_0) \in E$, on pose

$$F(t, x_0) = \exp \left(- \int_0^t \lambda(\phi_{\nu_0}(s, y_0), \nu_0) ds \right).$$

On définit ensuite pour $u \in [0, 1]$, $x \in [0, 1]$,

$$\psi_1(u, x) = \begin{cases} \inf\{t \geq 0 : F(t, x) \leq u\} \\ \infty, \text{ si } F(t, x) > u, \text{ pour tout } t > 0. \end{cases}$$

On prend alors le premier temps de saut

$$T_1 = \psi_1(U_1, x_0).$$

De cette façon, on a

$$\mathbb{P}(T_1 > t) = F(t, x_0).$$

On pose aussi

$$S_1 = T_1.$$

Pour construire la valeur du processus après le saut, on se donne une fonction mesurable

$$\psi_2 : [0, 1] \times E \rightarrow E$$

telle que, pour tout borélien A de E et $x \in E$,

$$m\{u \in [0, 1] : \psi_2(u, x) \in A\} = Q(A, x).$$

On a noté m la mesure de Lebesgue sur $[0, 1]$.

Alors sur $[0, T_1)$, la trajectoire du PDP est donnée par

$$x_t = (\phi_{\nu_0}(t, y_0), \nu_0).$$

Si $T_1 < \infty$, on a ensuite

$$x_{T_1} = \psi_2(V_1, (\phi_{\nu_0}(T_1, y_0), \nu_0)).$$

On recommence cette construction en partant de x_{T_1} . On pose

$$S_2 = \psi_1(U_2, x_{T_1}), \quad T_2 = T_1 + S_2.$$

Sur $[T_1, T_2)$

$$x_t = (\phi_{\nu_{T_1}}(t - T_1, y_{T_1}), \nu_{T_1}).$$

Si $T_2 < \infty$, on a ensuite

$$x_{T_2} = \psi_2(V_2, (\phi_{\nu_{T_1}}(S_2, y_{T_1}), \nu_{T_1})).$$

On construit ainsi le processus de manière récurrente sur chaque intervalle $[T_k, T_{k+1})$.

Notons qu'il peut être utile de considérer la chaîne de Markov suivante

$$Z_k = x_{T_k}, \quad k \in \mathbb{N}.$$

Si on pose

$$N_t = \sum_{k \in \mathbb{N}} \mathbf{1}_{t \geq T_k}, \quad t \geq 0,$$

qui représente le nombre de sauts avant l'instant t , une hypothèse de base et nécessaire pour la validité de nombreux résultats du livre [31] est que, pour tout point de départ $x_0 \in E$, on a

$$\mathbb{E}(N_t) < \infty, \quad t > 0.$$

Ceci sera vérifié dès que l'hypothèse 1 le sera. (Voir Proposition 24.6 de [31]).

Il est démontré dans [31] que le processus que nous venons de définir a la propriété de Markov forte.

7.2 Le générateur et le problème de martingale associé

Le générateur du processus de Markov construit dans le paragraphe précédent est décrit de manière très précise dans [31]. Nous n'avons pas besoin d'une

7.2 LE GÉNÉRATEUR ET LE PROBLÈME DE MARTINGALE ASSOCIÉ 93

description aussi précise. Nous utiliserons néanmoins la conséquence suivante du Théorème 26.14 de ce livre.

Définissons

$$\mathcal{E} = \{f : E \rightarrow \mathbb{R}, \text{ bornée et, pour tout } \nu \in \mathbb{N}^d, f(\cdot, \nu) \in C^1(\mathbb{R}^n)$$

et ses dérivées sont bornées uniformément en $\nu\}$.

On notera L_f la constante de Lipschitz par rapport à la variable y d'une fonction $f \in \mathcal{E}$:

$$L_f = \|D_y f\|_\infty, f \in \mathcal{E},$$

et on peut munir \mathcal{E} de la norme

$$\|f\|_{\mathcal{E}} = \|f\|_\infty + L_f$$

qui en fait un espace de Banach.

Pour $f \in \mathcal{E}$, on pose

$$\mathcal{A}f(x) = \chi_\nu(y) \cdot \nabla_y f(x) + \lambda(x) \int_E (f(z) - f(x)) Q(dz, x), x = (y, \nu) \in E.$$

Alors, si $(x_t)_{t \geq 0}$ est un PDP de caractéristiques $(\chi_\nu)_{\nu \in \mathbb{N}^d}$, λ , Q et démarrant en x ,

$$f(x_t) - f(x) - \int_0^t \mathcal{A}f(x_s) ds$$

est une martingale locale. Sous l'hypothèse 3, cette martingale locale est en fait bornée.

De façon équivalente, si on note $(\tilde{x}_t)_{t \geq 0} = (\tilde{y}_t, \tilde{\nu}_t)_{t \geq 0}$ le processus canonique sur l'espace de Skohorod $\mathbb{D}(\mathbb{R}^+; E)$ et P_x la loi de $(x_t)_{t \geq 0}$,

$$f(\tilde{x}_t) - f(x) - \int_0^t \mathcal{A}f(\tilde{x}_s) ds$$

est une P_x martingale pour tout $f \in \mathcal{E}$. Ce qui est encore équivalent à

$$\begin{aligned} & \mathbb{E}_x \left[\left(f(\tilde{x}_t) - f(x) - \int_0^t \mathcal{A}f(\tilde{x}_s) ds \right) \varphi(\tilde{x}_{t_1}, \dots, \tilde{x}_{t_\ell}) \right] \\ &= \mathbb{E}_x \left[\left(f(\tilde{x}_r) - f(x) - \int_0^r \mathcal{A}f(\tilde{x}_s) ds \right) \varphi(\tilde{x}_{t_1}, \dots, \tilde{x}_{t_\ell}) \right] \end{aligned}$$

pour tout $f \in \mathcal{E}$, $\varphi \in C_b\left(\left(\mathbb{R}^{n+d}\right)^\ell\right)$, $r \in [0, t]$ et $t_1, \dots, t_\ell \in [0, r]$.

On dit que P_x est solution du problème de martingale pour le PDP de caractéristiques $(\chi_\nu)_{\nu \in \mathbb{N}^d}$, λ , Q et de donnée initiale x .

Dans la suite, nous ne nous intéressons qu'aux distributions des PDP considérés. Nous ne ferons donc pas la distinction entre le PDP construit dans le paragraphe 1 et le processus canonique de loi P_x . Autrement dit nous omettrons de mettrons le symbole $\tilde{\cdot}$ pour le processus canonique.

Nous allons maintenant démontrer que, sous certaines hypothèses, P_x est l'unique solution du problème de martingale. Auparavant nous avons besoin de certains résultats préliminaires.

Soit $(P_t)_{t \geq 0}$ le semigroupe de transition associé au PDP $(x_t)_{t \geq 0}$. D'après le Théorème 27.6 de [31], on sait que si l'hypothèse 1 est vérifiée alors $(P_t)_{t \geq 0}$ vérifie la propriété de Feller. Nous utilisons le même type d'argument pour démontrer le résultat suivant.

Proposition 4 *Supposons l'hypothèse 3 vérifiée alors, pour tout $t \geq 0$, P_t envoie \mathcal{E} dans \mathcal{E} . De plus il existe des constantes c et K ne dépendant que des caractéristiques du PDP telle que, si $f \in \mathcal{E}$,*

$$|P_t f(y_1, \nu) - P_t f(y_2, \nu)| \leq c(\|f\|_\infty + L_f) e^{Kt} |y_1 - y_2|, \quad y_1, y_2 \in \mathbb{R}^n, \quad \nu \in \mathbb{N}^d.$$

7.2 LE GÉNÉRATEUR ET LE PROBLÈME DE MARTINGALE ASSOCIÉ 95

Preuve. On définit pour $f \in \mathcal{E}$ et $\Psi \in \mathcal{B}_b(\mathbb{R}^+ \times E)$:

$$\begin{aligned} G_f \psi(t, x) &= \mathbb{E}_x (f(x_t) \mathbf{1}_{t < T_1} + \psi(t - T_1, x_{T_1}) \mathbf{1}_{t \geq T_1}) \\ &= f(\phi_\nu(t, y), \nu) F(t, x) \\ &\quad + \int_0^t \int_E \psi(t - s, z) Q(dz, \phi_\nu(s, y), \nu) \lambda(\phi_\nu(s, y), \nu) F(s, x) ds \end{aligned}$$

pour $(t, x) \in \mathbb{R}^+ \times E$. Alors, grâce au Lemme 27.3 de [31] :

$$G_f^n \psi(t, x) = \mathbb{E}_x (f(x_t) \mathbf{1}_{t < T_n} + \psi(t - T_n, x_{T_n}) \mathbf{1}_{t \geq T_n})$$

et

$$\lim_{n \rightarrow \infty} G_f^n \psi(t, x) = P_t f(x).$$

On en déduit par convergence dominée

$$P_t f(x) = f(\phi_\nu(t, y), \nu) F(t, x) + \int_0^t \int_E P_{t-s} f(z) Q(dz, \phi_\nu(s, y), \nu) \lambda(\phi_\nu(s, y), \nu) F(s, x) ds.$$

Pour $T > 0$, définissons l'espace de Banach $\mathcal{E}_T = L^\infty(0, T; \mathcal{E})$ muni de la norme

$$\|\psi\|_{\mathcal{E}_T} = \sup_{t \in [0, T]} e^{-Kt} (\|\psi(t, \cdot)\|_\infty + L_{\psi(t, \cdot)})$$

où K sera fixé plus loin. Avec nos hypothèses, on vérifie que $q_t : (y, \nu) \mapsto f(\phi_\nu(t, y), \nu) F(t, y, \nu)$ est dans \mathcal{E} pour tout $t \geq 0$. De plus on montre facilement :

$$|D_y \phi_\nu(t, y)| \leq e^{Lt}, \quad (y, \nu) \in E,$$

$$\begin{aligned}
|D_y F(t, y, \nu)| &\leq \int_0^t |D_y \lambda(\phi_\nu(s, y), \nu)| ds \\
&\leq L_\lambda \int_0^t |D_y \phi_\nu(s, y)| ds \\
&\leq \frac{L_\lambda}{L} (e^{Lt} - 1), \quad (y, \nu) \in E
\end{aligned}$$

et

$$|F(t, y, \nu)| \leq 1, \quad (y, \nu) \in E.$$

On en déduit

$$|D_y (f(\phi_\nu(t, y), \nu) F(t, y, \nu))| \leq \left[L_f e^{Lt} + \|f\|_\infty \frac{L_\lambda}{L} e^{Lt} \right], \quad (y, \nu) \in E.$$

Donc $q : (y, \nu) \mapsto q_t(t, y, \nu)$ est dans \mathcal{E}_T et si $K \geq L$

$$\|q\|_{\mathcal{E}_T} \leq \left(L_f + \frac{L_\lambda}{L} \right)$$

Si $\psi \in \mathcal{E}_T$, on vérifie facilement que $G_f \psi \in \mathcal{E}_T$. De plus si $\psi_1, \psi_2 \in \mathcal{E}_T$, on a

$$|G_f \psi_1(t, x) - G_f \psi_2(t, \cdot)| \leq M_\lambda \int_0^t \|\psi_1(t-s, \cdot) - \psi_2(t-s, \cdot)\|_\infty ds$$

et donc

$$\|G_f \psi_1(t, \cdot) - G_f \psi_2(t, \cdot)\|_\infty \leq M_\lambda \int_0^t \|\psi_1(t-s, \cdot) - \psi_2(t-s, \cdot)\|_\infty ds. \quad (7.1)$$

De même

$$\begin{aligned}
&|D_y G_f \psi_1(t, y, \nu) - D_y G_f \psi_2(t, y, \nu)| \\
&\leq L_Q \int_0^t (\|\psi_1(t-s, \cdot) - \psi_2(t-s, \cdot)\|_\infty + L_{\psi_1(t-s, \cdot) - \psi_2(t-s, \cdot)}) e^{Ls} ds \\
&+ \frac{L_\lambda}{L} \int_0^t \|\psi_1(t-s, \cdot) - \psi_2(t-s, \cdot)\|_\infty e^{Ls} ds
\end{aligned}$$

7.2 LE GÉNÉRATEUR ET LE PROBLÈME DE MARTINGALE ASSOCIÉ 97

et on obtient après un calcul simple

$$\begin{aligned} \|G_f\psi_1 - G_f\psi_2\|_{\mathcal{E}_T} &\leq \kappa \sup_{t \in [0, T]} e^{-Kt} \left(\int_0^t e^{K(t-s)} ds + \int_0^t e^{K(t-s)} e^{Ls} ds \right) \|\psi_1 - \psi_2\|_{\mathcal{E}_T} \\ &\leq \kappa \left(\frac{1}{K} + \frac{1}{K-L} \right) \|\psi_1 - \psi_2\|_{\mathcal{E}_T} \end{aligned}$$

où κ ne dépend que des caractéristiques du PDP. Donc, en choisissant K suffisamment grand, on montre que G_f a un unique point fixe dans \mathcal{E}_T . En remarquant que (7.1) entraîne l'existence d'un unique point fixe dans les fonctions bornées, on en déduit que ces points fixes sont les mêmes et égaux à $P_t f$ qui est donc dans \mathcal{E}_T . ■

Nous introduisons maintenant le générateur infinitésimal $\bar{\mathcal{A}}$ de $(P_t)_{t \geq 0}$ qui peut être défini par

$$(\mu - \bar{\mathcal{A}})^{-1} f = \int_0^t e^{-\mu t} P_t f dt, \quad \mu > 0, \quad f \in \mathcal{B}_b(E).$$

En utilisant la Proposition 4, on montre que si l'hypothèse 3 est vérifiée, $(\mu - \bar{\mathcal{A}})^{-1}$ envoie \mathcal{E} dans lui-même si $\mu > K$. En effet, si $f \in \mathcal{E}$ alors $\varphi = (\mu - \bar{\mathcal{A}})^{-1} f$ est clairement borné. De plus, on montre que c'est une fonction de classe C^1 de y grâce aux théorèmes classiques de dérivation sous le signe intégrale. Enfin, si $(y_1, \nu), (y_2, \nu) \in E$, on a :

$$\begin{aligned} |\varphi(y_1, \nu) - \varphi(y_2, \nu)| &= \left| \int_0^t e^{-\mu t} (P_t f(y_1, \nu) - P_t f(y_2, \nu)) dt \right| \\ &\leq c(\|f\|_\infty + L_f) \int_0^\infty e^{(-\mu+K)t} dt |y_1 - y_2| \\ &\leq c(\|f\|_\infty + L_f) \frac{1}{\mu - K} |y_1 - y_2| \end{aligned}$$

Donc, si $f \in \mathcal{E}$ et $\varphi = (\mu - \bar{\mathcal{A}})^{-1} f$:

$$\varphi(x_t) - \varphi(x) - \int_0^t \mathcal{A}\varphi(x_s) ds$$

est une P_x martingale. On en déduit en particulier

$$\begin{aligned} |P_t\varphi(x) - \varphi(x)| &= \left| \int_0^t \mathbb{E}_x(\mathcal{A}\varphi(x_s)) ds \right| \\ &\leq t(M L_\varphi + 2M_\lambda \|\varphi\|_\infty). \end{aligned}$$

Donc

$$\frac{P_t\varphi - \varphi}{t}$$

reste borné quand $t \rightarrow 0$. De plus, comme $\mathcal{A}\varphi$ est continu et borné, on a

$$\frac{P_t\varphi(x) - \varphi(x)}{t} \rightarrow \mathcal{A}\varphi(x), \quad x \in E.$$

D'autre part, comme $\varphi = (\mu - \bar{\mathcal{A}})^{-1}f$, on a $\varphi \in \mathcal{D}(\bar{\mathcal{A}})$ et en particulier :

$$\frac{P_t\varphi(x) - \varphi(x)}{t} \rightarrow \bar{\mathcal{A}}\varphi(x), \quad x \in E.$$

On en déduit

$$\mathcal{A}\varphi(x) = \bar{\mathcal{A}}\varphi(x), \quad x \in E.$$

Nous pouvons maintenant énoncer et démontrer le résultat donnant l'unicité du problème de martingale.

Théorème 5 *Si l'Hypothèse 3 est vérifiée, alors pour tout $x \in E$, P_x est l'unique solution du problème de martingale associé au PDP de caractéristiques $(\chi_\nu)_{\nu \in \mathbb{N}^d}$, λ , Q .*

Preuve. Soit \tilde{P}_x une autre solution du problème de martingale. Comme ci-dessus, on prend $f \in \mathcal{E}$ et, pour $\mu > K$, $\varphi = (\mu - \bar{\mathcal{A}})^{-1}f$. Alors $\varphi \in \mathcal{E}$ et

$$\varphi(x_t) - \varphi(x) - \int_0^t \mathcal{A}\varphi(x_s) ds$$

est une P_x martingale. En particulier

$$\tilde{\mathbb{E}}_x \left(\varphi(x_t) - \int_0^t \mathcal{A}\varphi(x_s) ds \right) = \varphi(x).$$

Multiplions cette identité par $\mu e^{-\mu t}$ et intégrons sur $[0, \infty)$. On obtient après un

calcul élémentaire

$$\tilde{\mathbb{E}}_x \left(\int_0^\infty e^{-\mu t} (\mu \varphi(x_t) - \mathcal{A}\varphi(x_s)) ds \right) = \varphi(x).$$

Or, nous avons vu que $\bar{\mathcal{A}}\varphi = \mathcal{A}\varphi$, donc $(\mu - \mathcal{A})\varphi = f$ et

$$\tilde{\mathbb{E}}_x \left(\int_0^\infty e^{-\mu t} f(x_s) ds \right) = \varphi(x) = \int_0^\infty e^{-\mu t} P_t f(x).$$

Cette identité est vraie pour tout $\mu > K$ et l'injectivité de la transformée de Laplace nous permet d'affirmer

$$\tilde{\mathbb{E}}_x(f(x_t)) = P_t f(x), \quad t \geq 0.$$

La marginale à t fixé de \tilde{P}_x coïncide donc avec celle de P_x . Il est classique que cela implique que $\mathbb{P}_x = P_x$. ■

7.3 L'espace de Skohorod

Les processus déterministes par morceaux ne sont en général pas continus. Ils sont par contre cadlag. On rappelle ici la définition de l'espace de Skohorod et quelques unes de ses propriétés.

Si F est un espace métrique et I est un intervalle de \mathbb{R} , on note $\mathbb{D}(I; F)$ l'espace des fonctions cadlag de I dans F , c'est à dire l'espace des fonctions continues à droite et ayant une limite à gauche en tout point $t \in I$.

On rappelle qu'une suite $(f_n)_{n \in \mathbb{N}}$ converge vers f dans $\mathbb{D}([0, T]; F)$ si et seulement si il existe une suite de fonctions $(\lambda_n)_{n \in \mathbb{N}}$ de fonctions continues strictement croissantes de $[0, T]$ dans lui-même telle que

- $\lim_{n \rightarrow \infty} \sup_{t \in [0, T]} |\lambda_n(t) - t| = 0.$
- $\lim_{n \rightarrow \infty} \sup_{t \in [0, T]} d_F(f_n(\lambda_n(t)), f(t)) = 0.$

De façon équivalente, on peut introduire la métrique d_T telle que $d_T(f, g)$ est l'infimum des ϵ telle qu'il existe λ continue strictement croissante de $[0, T]$ dans lui-même vérifiant

- $\sup_{t \in [0, T]} |\lambda(t) - t| \leq \epsilon.$
- $\sup_{t \in [0, T]} d_F(f(\lambda(t)), g(t)) \leq \epsilon.$

Une suite $(f_n)_{n \in \mathbb{N}}$ converge vers f dans $\mathbb{D}(\mathbb{R}^+; F)$ si et seulement si elle converge dans $\mathbb{D}([0, T]; F)$ pour tout $T \geq 0$.

L'espace de Skohorod peut être muni d'une métrique qui définit la même topologie que ci-dessus et qui le rend complet.

Nous utiliserons d'autres propriétés de cet espace. Elles seront rappelées lorsque nous en aurons besoin.

Chapitre 8

Convergence des processus de Markov à sauts vers des Processus déterministes par morceaux

8.1 Introduction

Dans ce chapitre, nous justifions rigoureusement les modèles obtenus formellement dans les chapitres précédents. Nous montrons que les processus de saut correspondant à l'ensemble des réactions chimiques avant simplification convergent en loi vers les processus déterministes par morceaux - PDP dans la suite - associés aux divers modèles réduits. On justifie ainsi les modèles PDP continus - ou avec "switching" -, discontinus - ou avec "breaking". On justifie aussi le passage à la limite par moyennisation ou celui par "switching" singulier.

On part d'un ensemble de réactions R_r , $r \in \mathcal{R}$. Elles font intervenir des espèces dans un ensemble $S = 1, \dots, M$. Le nombre de molécules de i sera noté X_i et $X \in \mathbb{N}^M$ est le vecteur formé par ces X_i . Une réaction chimique fait intervenir certaines espèces dont le nombre de molécules est modifié quand la réaction se produit. Chaque réaction R_r est caractérisée par un taux $\lambda_r(X)$ qui dépend de l'état du système X et induit le changement $X \rightarrow X + \gamma_r$, où $\gamma_r \in \mathbb{N}^M$.

L'évolution du système est décrite mathématiquement par un processus de saut. Chaque réaction induit un saut de la variable d'état. Elles se produisent à

des instants $(T_i)_{i \geq 0}$ définis par

$$T_0 = 0, \quad T_i = \tau_1 + \cdots + \tau_i,$$

où $(\tau_k)_{k \geq 1}$ est une suite de variables aléatoires indépendantes de loi exponentielles. Sur chaque intervalle $[T_i, T_{i+1})$, $i \geq 0$, X est constant. La variable aléatoire τ_{i+1} a pour paramètre $\lambda(X(T_i)) = \sum_{r \in \mathcal{R}} \lambda_r(X(T_i))$:

$$\mathbb{P}(\tau_{i+1} \geq t) = \exp\left(-\sum_{r \in \mathcal{R}} \lambda_r(X(T_i))t\right).$$

A l'instant T_{i+1} , on choisit la réaction $r \in \mathcal{R}$ avec la probabilité $\lambda_r(X(T_i))/\lambda(X(T_i))$ et l'état du système change de la façon suivante :

$$X(T_{i+1}^+) = X(T_i) + \gamma_r.$$

Ce processus de Markov a pour générateur :

$$\mathcal{A}f(X) = \sum_{r \in \mathcal{R}} [f(X + \gamma_r) - f(X)] \lambda_r(X).$$

Nous n'avons pas besoin d'une description précise du domaine de ce générateur. On peut prendre f borélienne bornée sur \mathbb{N}^M .

8.2 Limite déterministe

La situation la plus simple et étudiée depuis longtemps est celle où toutes les espèces sont en grande quantité et les taux de réactions sont grands. On fait alors un changement d'échelle pour se ramener à des variables qui restent de l'ordre de l'unité. On suppose que l'on peut isoler un paramètre N tel que si on pose

$$x = \frac{1}{N}X$$

et

$$\tilde{\lambda}_r(x) = \frac{1}{N} \lambda_r(X), \quad r \in \mathcal{R}, \quad X \in \mathbb{N}^M,$$

alors x et $\tilde{\lambda}_r$ sont de l'ordre de l'unité.

Remarque 6 *Il est très important pour ce qui suit que les facteurs d'échelles de X et de λ_r soient du même ordre. Nous les prenons exactement égaux, ce qui*

peut paraître restrictif. Il est très facile de voir que la situation apparemment plus générale où on pose $\tilde{\lambda}_r(x) = \alpha_N \lambda_r(X)$, $r \in \mathcal{R}$, $X \in \mathbb{N}^M$, avec $N\alpha_N \rightarrow 1$ se traite exactement de la même manière. Par contre si $N\alpha_N \rightarrow 0$ ou ∞ , les résultats présentés ici ne s'appliquent pas. On considérera cependant une situation de ce type au paragraphe 8.5.

Après changement d'échelle, le générateur devient

$$\tilde{\mathcal{A}}_N f(x) = N \sum_{r \in \mathcal{R}} \left[f\left(x + \frac{\gamma_R}{N}\right) - f(x) \right] \tilde{\lambda}_r(x).$$

La limite $N \rightarrow \infty$ a été étudiée par Kurtz ([71], [72], voir aussi [29], [30]). On peut résumer les résultats essentiels de la façon suivante.

Théorème 7 Soit (x_N) une suite de processus homogènes de Markov à sauts d'espaces d'états $E_N \subset \mathbb{R}^k$, intensités $\lambda_N(x)$ et lois de saut $\mu_N(x, dy)$. Soit

$$F_N(x) = \lambda_N(x) \int_{E_N} (z - x) \mu_N(x, dz)$$

la fonction débit et supposons qu'il existe $F : \mathbb{R}^k \rightarrow \mathbb{R}^k$ Lipschitzienne et $x_0 \in \mathbb{R}^k$ telle que :

1. $\lim_{N \rightarrow \infty} \sup_{x \in E_N} |F_N(x) - F(x)| = 0$.
2. $\lim_{N \rightarrow \infty} |x_N(0) - x_0| = 0$, p.s.
3. $\lim_{N \rightarrow \infty} \sup_{x \in E_N} \lambda_N(x) \int_{E_N} |z - x|^2 \mu_N(x, dz) = 0$. Soit $(x(t))_{t \geq 0}$ l'unique solution du système d'équations différentielles

$$\begin{cases} \frac{dx}{dt} = F(x), \\ x(0) = x_0. \end{cases} \quad (8.1)$$

Alors

$$\lim_{N \rightarrow \infty} \mathbb{P} \left(\sup_{t \in [0, T]} |x_N(t) - x(t)| > \delta \right) = 0$$

pour tout $\delta > 0$ et $T > 0$.

Définissons la matrice de diffusion

$$G_N^{ij}(x) = \lambda(x) \int_{E_N} (z_i - x_i)(z_j - x_j) \mu_N(x, dz)$$

Si il existe $\alpha_N \rightarrow \infty$, $\epsilon \rightarrow 0$ et $G : \mathbb{R}^k \rightarrow M_k(\mathbb{R})$ tels que

1. $\lim_{N \rightarrow \infty} \sup_{x \in E_N} \|\alpha_N^2 G_N(x) - G(x)\| = 0$
2. $\lim_{N \rightarrow \infty} \sup_{x \in E_N} \alpha_N^2 \lambda_N(x) \int_{|z-x| > \epsilon_N / \alpha_N} |z-x|^2 \mu_N(dx) = 0$

alors $\alpha_N(x_N - x)$ converge en loi vers le processus d'Ito défini par $V(t) = \int_0^t G(x(s))^{1/2} d\beta(s)$.

Remarque 8 L'hypothèse que F est Lipschitzienne peut paraître restrictive pour les applications. En fait, elle n'est pas nécessaire et il est facile de la remplacer par l'hypothèse plus naturelle que le système d'équations différentielles (8.1) a une unique solution globale.

La première partie du théorème est la plus importante. Elle montre que le praticien peut simuler le système différentiel (8.1) à la place du système complexe de réactions chimiques. L'ingrédient principal de la preuve est l'inégalité de martingale suivante :

$$\begin{aligned} \mathbb{P} \left(\sup_{t \in [0, T]} \left| x_N(t) - x_N(0) - \int_0^t F_N(x_N(s)) ds \right| > \delta \right) \\ \leq \frac{t}{\delta^2} \sup_{x \in E_N} \lambda_N(x) \int_{E_N} |z-x|^2 \mu_N(x, dz). \end{aligned}$$

La deuxième partie est un théorème central limite et quantifie la vitesse de convergence. Des inégalités de concentration (voir [30]) ont aussi été utilisées.

Ce théorème s'applique à notre cadre en posant

$$\lambda^N = N \sum_{r \in \mathcal{R}} \tilde{\lambda}_r(x)$$

et

$$\mu_N(x, dz) = \sum_{r \in \mathcal{R}} \delta_{x + \frac{\gamma_r}{N}}(dz) \frac{\tilde{\lambda}_r(x)}{\sum_{r \in \mathcal{R}} \tilde{\lambda}_r(x)}.$$

Alors

$$F_N(x) = \sum_{r \in \mathcal{R}} \gamma_r \tilde{\lambda}_r(x)$$

et

$$G_N(x) = \frac{1}{N} \sum_{r \in \mathcal{R}} \tilde{\lambda}_r(x) \gamma_r^i \gamma_r^j.$$

Les hypothèses du Théorème 7 sont vérifiées si on suppose par exemple que les taux de réactions $\tilde{\lambda}_r$, $r \in \mathcal{R}$ sont bornés et Lipschitziens. Comme nous l'avons déjà dit, cette hypothèse peut être remplacé par une hypothèse plus naturelle.

Nous détaillons maintenant une autre méthode pour justifier le modèle limite (8.1). Elle est aussi due à Kurtz (voir [73], [37]). Nous généraliserons cette méthode dans la suite de ce travail. On suppose encore que les taux de réactions $\tilde{\lambda}_r$, $r \in \mathcal{R}$ sont bornés et Lipschitziens. Soit $(Y_r)_{r \in \mathcal{R}}$ une famille de processus de Poisson indépendants. Par le Théorème 4.1, chapitre 6 de [37], on sait qu'il existe des processus $(\tilde{x}_N)_{N \in \mathbb{N}}$ dans $\mathbb{D}(\mathbb{R}^+, \mathbb{R}^M)$ tels que

$$\tilde{x}_N(t) = x_N(0) + \sum_{r \in \mathcal{R}} \frac{\gamma_r}{N} Y_r \left(N \int_0^t \tilde{\lambda}_r(\tilde{x}_N(s)) ds \right).$$

De plus, pour tout $N \in \mathbb{N}$, \tilde{x}_N a la même loi que le processus x_N décrivant les réactions chimiques et de générateur $\tilde{\mathcal{A}}_N$.

Posons

$$F(x) = \sum_{r \in \mathcal{R}} \gamma_r \tilde{\lambda}_r(x).$$

Alors

$$\begin{aligned} & \tilde{x}_N(t) - x_N(0) - \int_0^t F(\tilde{x}_N(s)) ds \\ &= \sum_{r \in \mathcal{R}} \frac{\gamma_r}{N} \left[Y_r \left(N \int_0^t \tilde{\lambda}_r(\tilde{x}_N(s)) ds \right) - N \int_0^t \tilde{\lambda}_r(\tilde{x}_N(s)) ds \right] \end{aligned}$$

Utilisons maintenant la loi des grands nombres suivante

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T]} \left| \frac{1}{N} Y_r(Nt) - t \right| = 0,$$

Comme nous avons supposé que les taux de réactions $\tilde{\lambda}_r$, $r \in \mathcal{R}$ sont bornés, on peut écrire

$$\begin{aligned} & \lim_{N \rightarrow \infty} \sup_{t \in [0, T]} \left| \tilde{x}_N(t) - x_N(0) - \int_0^t F(\tilde{x}_N(s)) ds \right| \\ & \leq \lim_{N \rightarrow \infty} \sum_{r \in \mathcal{R}} \sup_{t \in [0, MT]} \left| \frac{1}{N} Y_r(Nt) - t \right| \\ & = 0 \end{aligned} \tag{8.2}$$

presque sûrement. Ici, M est un majorant des $\tilde{\lambda}_r$.

Soit $(x(t))_{t \geq 0}$ la solution de

$$\begin{cases} \frac{dx}{dt} = F(x), \\ x(0) = x_0. \end{cases} \quad (8.3)$$

Alors

$$\begin{aligned} \tilde{x}_N(t) - x(t) &= x_N(0) - x_0 + \int_0^t F(\tilde{x}_N(s)) - F(x(s)) ds \\ &\quad + \tilde{x}_N(t) - x_N(0) - \int_0^t F(\tilde{x}_N(s)) ds \end{aligned}$$

et

$$\begin{aligned} |\tilde{x}_N(t) - x(t)| &\leq |x_N(0) - x_0| + L_F \int_0^t |\tilde{x}_N(s) - x(s)| ds \\ &\quad + \left| \tilde{x}_N(t) - x_N(0) - \int_0^t F(\tilde{x}_N(s)) ds \right|. \end{aligned}$$

Il suffit maintenant d'utiliser le Lemme de Gronwall pour obtenir

$$|\tilde{x}_N(t) - x(t)| \leq e^{L_F t} \left(|x_N(0) - x_0| + \sup_{s \in [0, T]} \left| \tilde{x}_N(s) - x_N(0) - \int_0^s F(\tilde{x}_N(\tau)) d\tau \right| \right).$$

En utilisant (8.2), on obtient

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T]} |\tilde{x}_N(t) - x(t)| = 0$$

presque sûrement si $|\tilde{x}_N(0) - x_0| \rightarrow 0$ presque sûrement. On retrouve donc la première partie du théorème précédent. Ces arguments peuvent être prolongés pour obtenir aussi un théorème centrale limite et retrouver la deuxième partie des résultats (voir [37] chapitre 11).

Ces résultats sont encore vrais si on suppose seulement que (8.3) a une unique solution globale, il faut pour cela introduire des troncatures. Nous détaillerons cette méthode sur les cas plus complexes traiter ci-dessous.

8.3 Limite décrite par des PDP

Comme on l'a vu dans les chapitres précédents, la modélisation n'est pas réaliste pour certaines applications. Les molécules intervenant dans le système de réaction peuvent être en nombre présentant plusieurs échelles. Certaines molécules sont en faible nombre et le passage à la limite ci-dessus n'est pas valable.

On est amené à séparer l'ensemble des espèces en deux classes D et C de de M_D et M_C . Les molécules dans la classe D (resp. C) sont en faible (resp. grand) nombre. Cela induit la décomposition du vecteur X représentant les nombres de molécules pour chaque espèce en $X = (X_D, X_C)$. De même, les vecteurs de saut sont décomposés en $\gamma_r = (\gamma_r^D, \gamma_r^C)$. Le vecteur X_D est de l'ordre de l'unité tandis que X_C est grand et on suppose que l'on peut isoler un paramètre N tel que si on définit

$$x_C = \frac{1}{N} X_C$$

alors x_C est de l'ordre de l'unité. On pose

$$x = (x_C, X_D).$$

L'ensemble des réactions est lui aussi divisé en plusieurs sous ensembles :

$$\mathcal{R} = \mathcal{R}_D \cup \mathcal{R}_C \cup \mathcal{R}_{DC}$$

Les réactions dans \mathcal{R}_D (resp. \mathcal{R}_C) ne font intervenir que des espèces de type D (resp. \mathcal{R}_C).

Une réaction dans \mathcal{R}_{DC} a au moins un réactant dans D et un autre dans C .

De même que dans le paragraphe précédent, les taux des réactions dans \mathcal{R}_C sont grands et doivent être mis à l'échelle. on pose

$$\tilde{\lambda}_r = \frac{1}{N} \lambda_r, \quad r \in \mathcal{R}_C.$$

Les autres réactions ont a priori un taux de l'ordre de l'unité.

Les taux des réactions dans \mathcal{R}_D (resp. \mathcal{R}_c ne dépendent que de X_D (resp. x_C). Par contre le taux d'une réaction dans \mathcal{R}_{DC} dépend en général de X_D et de x_C .

Après le changement d'échelle, le générateur a la forme suivante

$$\begin{aligned}\tilde{\mathcal{A}}_N f(x_C, X_D) &= \sum_{r \in \mathcal{R}_C} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D) - f(x_C, X_D) \right] N \tilde{\lambda}_r(x_C) \\ &+ \sum_{r \in \mathcal{R}_{DC}} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D + \gamma_r^D) - f(x_C, X_D) \right] \lambda_r(x_C, X_D) \\ &+ \sum_{r \in \mathcal{R}_D} [f(x_C, X_D + \gamma_r^D) - f(x_C, X_D)] \lambda_r(X_D)\end{aligned}$$

Pour deviner le système limite lorsque $N \rightarrow \infty$, on peut prendre formellement la limite du générateur. On obtient pour f assez régulière :

$$\begin{aligned}\tilde{\mathcal{A}}_\infty f(x_C, X_D) &= \left(\sum_{r \in \mathcal{R}_C} \gamma_r^C \tilde{\lambda}_r(x_C) \right) \cdot \nabla_{x_C} f(x_C, X_D) \\ &+ \sum_{r \in \mathcal{R}_{DC}} [f(x_C, X_D + \gamma_r^D) - f(x_C, X_D)] \lambda_r(x_C, X_D) \\ &+ \sum_{r \in \mathcal{R}_D} [f(x_C, X_D + \gamma_r^D) - f(x_C, X_D)] \lambda_r(X_D).\end{aligned}$$

On observe que les évolutions des types d'espèces se découplent. La variable continue x_C évolue par un système d'équations différentielles comme dans le paragraphe précédent. La variable discrète X_D n'a aucune influence sur celle-ci. Il serait facile de justifier rigoureusement ce passage à la limite en utilisant les mêmes argument que dans le paragraphe suivant.

8.3.1 PDP continus

Nous étudions maintenant un premier cas où les comportements limites des variables x_C et X_D ne sont pas découplés.

Dans certaines applications, certaines réactions dans un sous ensemble S_1 de \mathcal{R}_{DC} ont un taux de réactions grand et du même ordre de grandeur que les réac-

tions dans \mathcal{R}_C . On est donc amené à poser

$$\tilde{\lambda}_r = \frac{1}{N} \lambda_r, \quad r \in S_1.$$

On suppose que pour ces réactions, la variable discrète n'est pas modifiée :

$$\gamma_r^D = 0, \quad r \in S_1. \quad (8.4)$$

Si cette hypothèse n'est pas vérifiée, le comportement limite est différent et doit être étudié par une méthode de moyennisation. Nous étudierons ce cas dans le paragraphe 8.4.

Le générateur a maintenant la forme :

$$\begin{aligned} \tilde{\mathcal{A}}_N f(x_C, X_D) &= \sum_{r \in \mathcal{R}_C} \left[f\left(x_C + \frac{1}{N} \gamma_r^C, X_D\right) - f(x_C, X_D) \right] N \tilde{\lambda}_r(x_C) \\ &+ \sum_{r \in S_1} \left[f\left(x_C + \frac{1}{N} \gamma_r^C, X_D\right) - f(x_C, X_D) \right] N \tilde{\lambda}_r(x_C, X_D) \\ &+ \sum_{r \in \mathcal{R}_{DC} \setminus S_1} \left[f\left(x_C + \frac{1}{N} \gamma_r^C, X_D + \gamma_r^D\right) - f(x_C, X_D) \right] \lambda_r(x_C, X_D) \\ &+ \sum_{r \in \mathcal{R}_D} \left[f(x_C, X_D + \gamma_r^D) - f(x_C, X_D) \right] \lambda_r(X_D). \end{aligned} \quad (8.5)$$

Formellement le générateur limite est donné par

$$\begin{aligned}
\tilde{\mathcal{A}}_\infty f(x_c, X_D) &= \left(\sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C) \gamma_r^C \right) \cdot \nabla_{x_C} f(x_C, X_D) \\
&+ \left(\sum_{r \in S_1} \tilde{\lambda}_r(x_C, X_D) \gamma_r^C \right) \cdot \nabla_{x_C} f(x_C, X_D) \\
&+ \sum_{r \in \mathcal{R}_{DC} \setminus S_1} [f(x_C, X_D + \gamma_r^D) - f(x_C, X_D)] \lambda_r(x_C, X_D) \\
&+ \sum_{r \in \mathcal{R}_D} [f(x_C, X_D + \gamma_r^D) - f(x_C, X_D)] \lambda_r(X_D).
\end{aligned}$$

C'est le générateur d'un PDP. Celui-ci aura des trajectoires continues pour la variable x_C . En effet, seule la composante discrète X_D présente des sauts. Par contre, la dérivée en temps de x_C n'est en général pas continu aux instants où X_D change car le champs de vecteurs dirigeant x_C change. On dit qu'il y a "switching".

Le résultat suivant établi rigoureusement la convergence vers ce PDP.

On suppose que les taux de réactions $\tilde{\lambda}_r$, $r \in \mathcal{R}_C \cup S_1$, et λ_r , $r \in \mathcal{R}_{DC} \setminus S_1$, sont des fonctions C^1 de la variable x_c . On pose

$$F(x_c, X_D) = \left(\sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C) \gamma_r^C + \sum_{r \in S_1} \tilde{\lambda}_r(x_C, X_D) \gamma_r^C \right).$$

C'est donc une fonction C^1 de x_C . On suppose que, pour tout $X_D \in \mathbb{N}^{M_D}$ et tout $x_{C,0} \in \mathbb{R}^{M_C}$, le système différentiel

$$\begin{cases} \frac{dx_C}{dt} = F(x_C, X_D), \\ x_C(0) = x_{C,0}. \end{cases} \quad (8.6)$$

a une unique solution globale. Enfin, on suppose que λ_r est borné pour $r \in (\mathcal{R}_{DC} \setminus S_1) \cup \mathcal{R}_D$. Ainsi, le PDP de générateur $\tilde{\mathcal{A}}_\infty$ vérifie l'Hypothèse 1.

Nous définissons aussi un ensemble de fonctions tests pour lesquels les géné-

rateurs sont bien définis et que nous utiliserons dans la preuve :

$$\mathcal{E} = \{f : \mathbb{R}^{M_C} \times \mathbb{N}^{M_D} \rightarrow \mathbb{R}, \text{ bornée, pour tout } X_D, f(\cdot, X_D) \in C_b^1(\mathbb{R}^{M_C}; \mathbb{R}) \\ \text{et les dérivées sont bornées uniformément en } X_D\}$$

Théorème 9 *Sous les hypothèses ci-dessus, soit la famille de processus $x^N = (x_C^N, X_D^N)$ de générateur $\tilde{\mathcal{A}}_N$ démarrant en $x^N(0) = (x_C^N(0), X_D^N(0))$, décrivant l'évolution du système de réactions chimiques. Si*

$$x^N(0) = (x_C^N(0), X_D^N(0)) \rightarrow x_0 = (x_{C,0}, X_{D,0})$$

en distribution dans $\mathbb{R}^{M_C} \times \mathbb{N}^{M_D}$, alors $x^N = (x_C^N, X_D^N)$ converge en distribution dans $\mathbb{D}(\mathbb{R}^{M_C} \times \mathbb{N}^{M_D})$ vers $x = (x_c, X_D)$, le PDP de générateur $\tilde{\mathcal{A}}_\infty$ démarrant en $x_0 = (x_{C,0}, X_{D,0})$.

Preuve. Comme dans le paragraphe précédent, on utilise le fait que les processus x^N ont la même loi que les processus $\tilde{x}^N = (\tilde{x}_C^N, \tilde{X}_D^N)$ vérifiant

$$\begin{aligned} \tilde{x}_C^N(t) &= x_C^N(0) + \sum_{r \in \mathcal{R}_C} \frac{1}{N} \gamma_r^C Y_r \left(N \int_0^t \tilde{\lambda}_r(\tilde{x}_C^N(s)) ds \right) \\ &+ \sum_{r \in S_1} \frac{1}{N} \gamma_r^C Y_r \left(N \int_0^t \tilde{\lambda}_r(\tilde{x}_C^N(s), \tilde{X}_D^N(s)) ds \right) \\ &+ \sum_{r \in \mathcal{R}_{DC} \setminus S_1} \frac{1}{N} \gamma_r^C Y_r \left(\int_0^t \lambda_r(\tilde{x}_C^N(s), \tilde{X}_D^N(s)) ds \right) \end{aligned}$$

et

$$\begin{aligned} \tilde{X}_D^N(t) &= X_D^N(0) + \sum_{r \in \mathcal{R}_{DC} \setminus S_1} \gamma_r^D Y_r \left(\int_0^t \lambda_r(\tilde{x}_C^N(s), \tilde{X}_D^N(s)) ds \right) \\ &+ \sum_{r \in \mathcal{R}_D} \gamma_r^D Y_r \left(\int_0^t \lambda_r(\tilde{X}_D^N(s)) ds \right) \end{aligned}$$

Où $(Y_r)_{r \in \mathcal{R}}$ est une famille de processus de Poisson indépendants. Etant donné que l'on ne s'intéresse qu'aux distributions des procesus, nous ne ferons pas dans

la suite de la preuve la distinction entre \tilde{x}^N et x^N . Autrement dit, nous omettons le symbole tilde. Cela ne créera aucune confusion.

Dans un premier temps, nous faisons l'hypothèse que les tous les taux de réactions sont bornés. De plus, $\tilde{\lambda}_r$, $r \in \mathcal{R}_C$ et, pour chaque $X_D \in \mathbb{N}^{M_D}$, $\tilde{\lambda}_r(\cdot, X_D)$, $r \in S_1$, et $\lambda_r(\cdot, X_D)$, $r \in \mathcal{R}_{DC} \setminus S_1$ sont dans $C_b^1(\mathbb{R}^{M_C})$ et leurs dérivées sont bornées uniformément par rapport à X_D . De cette façon, le PDP limite vérifie l'Hypothèse 3. Cette hypothèse sera levée à la fin de la preuve.

Etape 1 : Tension de la famille $(x^N)_{N \in \mathbb{N}}$ dans $\mathbb{D}(\mathbb{R}^+, \mathbb{R}^{M_C} \times \mathbb{N}^{M_D})$.

On commence par montrer la tension $(X_D^N)_{N \in \mathbb{N}}$ dans $\mathbb{D}(\mathbb{R}^+, \mathbb{N}^{M_D})$. Définissons

$$X_r^N(t) = Y_r \left(\int_0^t \lambda_r(x_C^N(s), X_D^N(s)) ds \right), \quad r \in \mathcal{R}_{RD} \setminus S_1$$

et

$$X_r^N(t) = Y_r \left(\int_0^t \lambda_r(X_D^N(s)) ds \right), \quad r \in \mathcal{R}_D.$$

Il suffit clairement de montrer que chaque X_r^N est tendu dans $\mathbb{D}(\mathbb{R}^+; \mathbb{N})$.

Rappelons qu'un sous ensemble $A \subset \mathbb{D}(\mathbb{R}^+, \mathbb{N})$ est relativement compact si et seulement si

1. Pour tout $t \geq 0$, il existe $M_t \geq 0$ tel que $|y_t| \leq M_t$ pour tout $y \in A$.
2. Pour tout $T > 0$, $\lim_{\delta \rightarrow 0} \sup_{y \in A} w'(y, \delta, T) = 0$.

Dans 1), il suffit de prendre $t \in \mathbb{Q}$. Cette remarque est en fait peu utile. Dans la pratique, M_t peut même souvent être le même sur tout intervalle borné $[0, T]$.

Dans 2), w' est une généralisation du module de continuité :

$$w'(x, \delta, T) = \inf_{(t_i)} \max_i \sup_{s, t \in (t_i, t_{i+1})} |x(t) - x(s)|,$$

où $x \in \mathbb{D}(\mathbb{R}^+, \mathbb{N}^{M_D})$, $\delta > 0$ et l'infimum est pris sur toutes les subdivisions : $0 = t_0 < t_1 < \dots < t_n = T$ de pas supérieur à δ : $\min_i |t_{i+1} - t_i| > \delta$. Clairement w' est une fonction croissante de δ .

Soit M_1 un majorant des taux de réaction de $(\mathcal{R}_{RD} \setminus S_1) \cup \mathcal{R}_D$, alors

$$\sup_{t \in [0, T]} |X_r^N(t)| \leq \sup_{t \in [0, M_1 T]} |Y_r(t)|$$

et

$$w_r'^N, \delta, T \leq w'(Y_r, M_1 \delta, M_1 T).$$

La tension est maintenant immédiate.

Pour $(x_C^N)_{N \in \mathbb{N}}$, on adapte l'argument présenté dans le paragraphe 2. On écrit :

$$\begin{aligned} x_C^N(t) &= x_C^N(0) + \sum_{r \in \mathcal{R}_C} \frac{1}{N} \gamma_r^C \bar{Y}_r \left(N \int_0^t \lambda_r(x_C^N(s)) ds \right) \\ &+ \sum_{r \in S_1} \frac{1}{N} \gamma_r^C \bar{Y}_r \left(N \int_0^t \lambda_r(x_C^N(s), X_D^N(s)) ds \right) \\ &+ \int_0^t F(x_C^N(s), X_D^N(s)) ds \\ &+ \sum_{r \in \mathcal{R}_{DC} \setminus S_1} \frac{1}{N} \gamma_r^C Y_r \left(\int_0^t \lambda_r(x_C^N(s), X_D^N(s)) ds \right), \end{aligned}$$

où maintenant $\bar{Y}_r(u) = Y_r(u) - u$, $r \in \mathcal{R}_C \cup S_1$, est un processus de Poisson centré.

Il est classique que pour tout $K > 0$:

$$\sup_{u \in [0, K]} \frac{1}{N} \bar{Y}_r(Nu) \rightarrow 0, \text{ a.s.}$$

On en déduit, grâce à l'hypothèse de bornitude des taux de réactions, que pour tout $T > 0$:

$$\begin{aligned} \sup_{t \in [0, T]} \left| \sum_{r \in \mathcal{R}_C} \frac{1}{N} \gamma_r^C \bar{Y}_r \left(N \int_0^t \lambda_r(x_C^N(s)) ds \right) \right. \\ \left. + \sum_{r \in S_1} \frac{1}{N} \gamma_r^C \bar{Y}_r \left(N \int_0^t \lambda_r(x_C^N(s), X_D^N(s)) ds \right) \right| \rightarrow 0, \text{ a.s.} \end{aligned}$$

lorsque $N \rightarrow \infty$. De plus

$$\begin{aligned} & \sup_{t \in [0, T]} \left| \sum_{r \in \mathcal{R}_{DC} \setminus S_1} \frac{1}{N} \gamma_r^C Y_r \left(\int_0^t \lambda_r(x_C^N(s), X_D^N(s)) ds \right) \right| \\ & \leq \frac{1}{N} \sum_{r \in \mathcal{R}_{DC} \setminus S_1} \gamma_r^C \sup_{t \in [0, M_1 T]} |Y_r(t)| \\ & \rightarrow 0, \text{ a.s.} \end{aligned}$$

On peut donc écrire

$$x_C^N(t) = x_C^N(0) + \int_0^t F(x_C^N(s), X_D^N(s)) ds + \varepsilon_N(t)$$

avec

$$K_N = \sup_{t \in [0, T]} |\varepsilon_N(t)| \rightarrow 0, n \rightarrow \infty, \text{ a.s.}$$

Soit M_2 un majorant de F :

$$|F(x_C, X_D)| \leq M_2, (x_C, X_D) \in \mathbb{R}^{M_C} \times \mathbb{N}^{M_D}.$$

On a alors

$$|x_C^N(t)| \leq |x_C^N(0)| + M_2 t + K_N, \text{ a.s.}$$

De plus, pour $t_1, t_2 \in [0, T]$

$$|x_C^N(t_1) - x_C^N(t_2)| \leq M_2 |t_1 - t_2| + 2K_N, \text{ a.s.}$$

La tension de $(x_C^N)_{N \in \mathbb{N}}$ dans $\mathbb{D}([0, \infty); \mathbb{R}^{M_C})$ en résulte. De plus, toutes les lois limites sont portées par $C([0, \infty); \mathbb{R}^{M_C})$.

Etape 2 : Identification des points limites et convergence dans le cas borné

On suppose toujours que les taux de réactions sont bornés et dans $C_b^1(\mathbb{R}^{M_C})$ par rapport à la variable x_C .

On note $(x_t)_{t \geq 0}$ le processus canonique sur $\mathbb{D}(\mathbb{R}^+; \mathbb{R}^M)$ et, pour $N \in \mathbb{N}$ $P_{x, N}$ la

loi de $(x_N(t))_{t \geq 0}$ dans $\mathbb{D}(\mathbb{R}^+; \mathbb{R}^M)$. Alors pour $\varphi \in \mathcal{E}$

$$\varphi(x_t) - \varphi(x) - \int_0^t \tilde{\mathcal{A}}_N \varphi(x_s) ds$$

est une $P_{x,N}$ martingale. De manière équivalente, pour tout $n \in \mathbb{N}$, $t_1, \dots, t_n \in [0, r]$, $t \geq r \geq 0$, $\psi \in (C_b(\mathbb{R}^M))^n$ et $\varphi \in \mathcal{E}$

$$\begin{aligned} & \mathbb{E}_{P_{N,x}} \left(\left(\varphi(x_t) - \varphi(x_0) - \int_0^t \tilde{\mathcal{A}}_N \varphi(x_s) ds \right) \psi(x_{t_1}, \dots, x_{t_n}) \right) \\ &= \mathbb{E}_{P_{N,x}} \left(\varphi(x_r) - \varphi(x_0) - \int_0^r \tilde{\mathcal{A}}_N \varphi(x_s) ds \right). \end{aligned} \quad (8.7)$$

Soit $(P_{N_k})_k$ une suite qui converge faiblement et P_x un point limite. On sait que x is P_x presque sûrement continu en chaque $t \in \mathbb{R}^+$ sauf sur un ensemble dénombrable $D(x)$. De plus, pour t_1, \dots, t_n n'appartenant pas à $D(x)$ la loi de $(x_{N_k}(t_1), \dots, x_{N_k}(t_n))$ converge vers $(x(t_1), \dots, x(t_n))$ où $(x(t))$ est un processus de loi P_x . On peut maintenant prendre la limite $k \rightarrow \infty$ dans (8.7) selon la sous suite N_k et obtenir pour t, t_1, \dots, t_n n'appartenant pas à $D(x)$:

$$\begin{aligned} & \mathbb{E}_{P_x} \left(\left(\varphi(x_t) - \varphi(x_0) - \int_0^t \mathcal{A}_\infty \varphi(x_s) ds \right) \psi(x_{t_1}, \dots, x_{t_n}) \right) \\ &= \mathbb{E}_{P_x} \left(\varphi(x_r) - \varphi(x_0) - \int_0^r \mathcal{A}_\infty \varphi(x_s) ds \right), \end{aligned} \quad (8.8)$$

Si $t \in D(x)$, on prend une suite (t^k) en dehors de $D(x)$ telle que $t^k \rightarrow t$ et $x(t^k)$ converge en distribution vers $x(t)$. Alors, en utilisant (8.8) avec t^k à la place de t et en prenant $k \rightarrow \infty$, on montre que (8.8) est encore vrai pour $t \in D(x)$. En procédant de même, on montre que l'on peut prendre t_1, \dots, t_n dans $D(x)$.

On a montré que P_x est solution du problème de martingale associé au PDP de générateur $\tilde{\mathcal{A}}_\infty$ où les fonctions tests sont prises dans \mathcal{E} .

Or, avec les hypothèses faites sur les taux de réactions, on peut appliquer le Théorème 5 et en déduire qu'il y a une unique solution à ce problème de martingale. On en déduit que toute la suite $(P_N)_{N \in \mathbb{N}}$ converge vers la loi du PDP de générateur $\tilde{\mathcal{A}}_\infty$.

Etape 3 : Conclusion

Il reste à montrer que l'hypothèse supplémentaire sur les taux de réactions n'est pas nécessaire. Pour cela, on utilise une fonction de troncature $\theta \in C_0^\infty(\mathbb{R})$

telle que

$$\begin{cases} \theta(x) = 1, & x \in [0, 1], \\ \theta(x) = 0, & x \in [2, \infty), \end{cases}$$

et pour $K \geq 1$ on définit

$$\theta_K(x) = \theta\left(\frac{|x|^2}{K^2}\right), \quad x \in \mathbb{R}^M.$$

Enfin, pour $r \in \mathcal{R}$, on pose

$$\tilde{\lambda}_r^K(x) = \theta_K(x)\lambda(x).$$

Alors, les $\tilde{\lambda}_r$ vérifient les hypothèses des étapes précédentes. On définit $\tilde{\mathcal{A}}_{N,K}$ et $\tilde{\mathcal{A}}_{\infty,K}$ les générateurs où les taux de réactions ont été tronqués, c'est à dire où les $\tilde{\lambda}_r$ ont été remplacé par $\tilde{\lambda}_r^K$. On considère aussi des processus $x_K^N = (x_{C,K}^N, X_{D,K}^N)$, $N \in \mathbb{N}$ et $x_K = (x_{C,K}, X_{D,K})$ associés à ces générateurs. D'après ce qui précède, on sait que pour tout $K \in \mathbb{N}$ $(x_K^N)_{N \in \mathbb{N}}$ converge en loi vers (x_K) dans $\mathbb{D}(\mathbb{R}^+; \mathbb{R}^{M_C} \times \mathbb{N}^{M_D})$. Donc $((x_K^N)_{K \in \mathbb{N}})_{N \in \mathbb{N}}$ converge en loi vers $(x_K)_{K \in \mathbb{N}}$ dans $(\mathbb{D}(\mathbb{R}^+; \mathbb{R}^{M_C} \times \mathbb{N}^{M_D}))^{\mathbb{N}}$. Par le théorème de Skohorod, on peut donc supposer, après avoir pris un nouvel espace de probabilité, que pour tout $K \in \mathbb{N}$ (x_K^N) converge presque sûrement vers (x_K) dans $\mathbb{D}(\mathbb{R}^+; \mathbb{R}^{M_C} \times \mathbb{N}^{M_D})$.

Définissons le temps d'arrêt

$$\tau^K = \inf\{t \geq 0, |x_K(t)| \geq K\}.$$

Alors on vérifie alors facilement que pour $K, L \in \mathbb{N}$

$$x_K(t) = x_L(t), \quad t \in [0, \tau^K \wedge \tau^L]$$

et donc τ^K est croissant.

De plus si x le PDP associé à $\tilde{\mathcal{A}}_{\infty}$, alors

$$x_K(t) = x(t), \quad t \in [0, \tau^K).$$

De même, si

$$\tau_N^K = \inf\{t \geq 0, |x^N(t)| \geq K\}.$$

alors

$$x_K^N(t) = x^N(t), \quad t \in [0, \tau_N^K].$$

On remarque que si $\tau^{K-1} > T$ et $d_T(x_K^N, x_K) < \varepsilon$, où d_T est la distance sur l'espace de Skohorod $\mathbb{D}([0, T]; \mathbb{R}^{M_C} \times \mathbb{N}^{M_D})$, alors pour ε assez petit :

$$\sup_{t \in [0, T]} |x_K^N(t)| \leq K$$

et donc $\tau_N^K > T$ et $x_K^N = x^N$ sur $[0, T]$. Comme $\tau^K \geq \tau^{K-1} > T$; on a aussi $x_K = x$ sur $[0, T]$ et donc

$$d_T(x^N, x) < \varepsilon.$$

Il en ressort que

$$\mathbb{P}(d_T(x^N, x) \geq \varepsilon) \leq \mathbb{P}(\tau^{K-1} \leq T) + \mathbb{P}(d_T(x_K^N, x_K) \geq \varepsilon).$$

On prend K grand puis ensuite N grand pour en déduire que x^N converge en probabilité vers x . ■

Remarque 10 *En utilisant la preuve que nous venons d'exposer et le Corollaire 6.4 chapitre 4 de [37], on montre que le PDP limite a une solution unique au sens des martingales.*

8.3.2 PDP discontinus

Maintenant, nous considérons le cas où il existe aussi des réactions dans un sous ensemble S_2 de \mathcal{R}_{DC} telles que γ_r^C est grand et du même ordre que X_C . On est alors amené à poser $\tilde{\gamma}_r^C = \frac{1}{N} \gamma_r^C$ pour $r \in S_2$. On définit $S = S_1 \cup S_2$.

On obtient ainsi le générateur

$$\begin{aligned}
\tilde{\mathcal{A}}_N f(x_C, X_D) &= \sum_{r \in \mathcal{R}_C} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D) - f(x_C, X_D) \right] N \tilde{\lambda}_r(x_C) \\
&+ \sum_{r \in S_2} [f(x_C + \tilde{\gamma}_r^C, X_D + \gamma_r^D) - f(x_C, X_D)] \lambda_r(x_C, X_D) \\
&+ \sum_{r \in S_1} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D + \gamma_r^D) - f(x_C, X_D) \right] N \tilde{\lambda}_r(x_C, X_D) \\
&+ \sum_{r \in \mathcal{R}_{DC} \setminus S} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D + \gamma_r^D) - f(x_C, X_D) \right] \lambda_r(x_C, X_D) \\
&+ \sum_{r \in \mathcal{R}_D} [f(x_C, X_D + \gamma_r^D) - f(x_C, X_D)] \lambda_r(X_D).
\end{aligned} \tag{8.9}$$

Formellement, pour $N \rightarrow \infty$, on obtient le générateur limite

$$\begin{aligned}
\tilde{\mathcal{A}}_\infty f(x_C, X_D) &= \left(\sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C) \gamma_r^C + \sum_{r \in S_1} \tilde{\lambda}_r(x_C, X_D) \gamma_r^C \right) \cdot \nabla_{x_C} f(x_C, X_D) \\
&+ \sum_{r \in S_2} [f(x_C + \tilde{\gamma}_r^C, X_D + \gamma_r^D) - f(x_C, X_D)] \lambda_r(x_C, X_D) \\
&+ \sum_{r \in \mathcal{R}_{DC} \setminus S} [f(x_C, X_D + \gamma_r^D) - f(x_C, X_D)] \lambda_r(x_C, X_D) \\
&+ \sum_{r \in \mathcal{R}_D} [f(x_C, X_D + \gamma_r^D) - f(x_C, X_D)] \lambda_r(X_D).
\end{aligned}$$

On voit que le PDP limite peut présenter des discontinuités dans la variable x_C . On dit alors qu'il y a "breakage".

La justification rigoureuse de ce passage à la limite résulte directement du travail déjà fait dans le paragraphe précédent. Pour voir cela, introduisons le système artificiel faisant intervenir les variables $(x_C^{1,N}, x_C^{2,N}, X^D)$ dans $\mathbb{R}^{M_C} \times \mathbb{R}^{M_C} \times \mathbb{N}^{M_D}$:

- Si $r \in \mathcal{R}_C$, $(x_C^{1,N}, x_C^{2,N}, X^D) \rightarrow (x_C^{1,N} + \frac{1}{N}\gamma_C^r, x_C^{2,N}, X^D)$.
- Si $r \in S_2$, $(x_C^{1,N}, x_C^{2,N}, X^D) \rightarrow (x_C^{1,N}, x_C^{2,N} + \gamma_C^r, X^D)$.
- Si $r \in \mathcal{R}_{DC} \setminus S_2$, $(x_C^{1,N}, x_C^{2,N}, X^D) \rightarrow (x_C^{1,N} + \frac{1}{N}\gamma_C^r, x_C^{2,N}, X^D + \gamma_D^r)$.
- Si $r \in \mathcal{R}_D$, $(x_C^{1,N}, x_C^{2,N}, X^D) \rightarrow (x_C^{1,N}, x_C^{2,N}, X^D + \gamma_D^r)$.

Les taux de réactions correspondant sont donnés par $\lambda_r(x_C^{1,N}, x_C^{2,N}, X^D) = \lambda_r(x_C^{1,N} + x_C^{2,N}, X^D)$. Si on prend pour donnée initiale $(x_C^{1,N}(0), x_C^{2,N}(0), X_D^N(0)) = (x_C^N(0), 0, X_D^N(0))$ alors les deux systèmes sont équivalents. En effet, partant du système auxiliaire, on retrouve les valeurs du système original en posant

$$(x_C^N, X_D^N) = (x_C^{1,N} + x_C^{2,N}, X_D).$$

Réciproquement, si on a le processus de saut (x_C^N, X_D^N) , alors pour N assez grand on peut construire $(x_C^{1,N} + x_C^{2,N}, X_D)$ en isolant les grands sauts de x_C^N .

Le système auxiliaire correspond au générateur suivant

$$\begin{aligned}
& \mathcal{A}_{aux} f(x_C^1, x_C^2, X^D) \\
&= \sum_{r \in \mathcal{R}_C} \left[f(x_C^1 + \frac{1}{N} \gamma_r^C, x_C^2, X_D) - f(x_C^1, x_C^2, X_D) \right] N \tilde{\lambda}_r(x_C^1 + x_C^2) \\
&+ \sum_{r \in S_2} [f(x_C^1, x_C^2 + \tilde{\gamma}_r^C, X_D + \gamma_r^D) - f(x_C^1, x_C^2, X_D)] \lambda_r(x_C^1 + x_C^2, X_D) \\
&+ \sum_{r \in S_1} \left[f(x_C^1 + \frac{1}{N} \gamma_r^C, x_C^2, X_D) - f(x_C^1, x_C^2, X_D) \right] N \tilde{\lambda}_r(x_C^1 + x_C^2, X_D) \\
&+ \sum_{r \in \mathcal{R}_{DC} \setminus S} \left[f(x_C^1 + \frac{1}{N} \gamma_r^C, x_C^2, X_D + \gamma_r^D) - f(x_C^1, x_C^2, X_D) \right] \lambda_r(x_C^1 + x_C^2, X_D) \\
&+ \sum_{r \in \mathcal{R}_D} [f(x_C^1, x_C^2, X_D + \gamma_r^D) - f(x_C^1, x_C^2, X_D)] \lambda_r(X_D).
\end{aligned} \tag{8.10}$$

Ce générateur est presque de la même forme que dans le paragraphe précédent en prenant pour variable discrète (x_C^2, X_D) . Les seules différences sont que les taux de réactions de \mathcal{R}_C font intervenir la variable discrète et que les taux des réactions de S_2 , qui sont maintenant vues comme des réactions sur la variable discrète, dépendent de la variable continue. Il est cependant immédiat de voir que la preuve ci-dessus se généralise immédiatement et on obtient la convergence en

loi de $(x_C^{1,N}, x_C^{2,N}, X_D)$ vers le PDP (x_C^1, x_C^2, X_D) de générateur

$$\begin{aligned}
& \mathcal{A}_{aux, \infty} f(x_C^1, x_C^2, X_D) \\
&= \left(\sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C^1 + x_C^2, X_D) \gamma_r^C + \sum_{r \in \mathcal{S}_1} \tilde{\lambda}_r(x_C^1 + x_C^2) \gamma_r^C \right) \cdot \nabla_{x_C^1} f(x_C^1, x_C^2, X_D) \\
&+ \sum_{r \in \mathcal{S}_2} [f(x_C^1, x_C^2 + \tilde{\gamma}_r^C, X_D + \gamma_r^D) - f(x_C^1, x_C^2, X_D)] \lambda_r(x_C^1 + x_C^2, X_D) \\
&+ \sum_{r \in \mathcal{R}_{DC} \setminus \mathcal{S}} [f(x_C^1, x_C^2, X_D + \gamma_r^D) - f(x_C^1, x_C^2, X_D)] \lambda_r(x_C^1 + x_C^2, X_D) \\
&+ \sum_{r \in \mathcal{R}_D} [f(x_C^1, x_C^2, X_D + \gamma_r^D) - f(x_C^1, x_C^2, X_D)] \lambda_r(X_D).
\end{aligned}$$

Pour cela, il faut supposer que les hypothèses suivantes sont vérifiées

- Pour $r \in \mathcal{S}_2 \cup (\mathcal{R}_{DC} \setminus \mathcal{S}) \cup \mathcal{R}_D$, les taux de réactions λ_r sont bornés, de classe C^1 par rapport à x_C .
- Pour tout $x_{C,0}$, le système (8.6) a une unique solution globale.

En effet, traduites en termes de (x_C^1, x_C^2, X_D) , on a exactement les hypothèses du Théorème 9.

En posant $x^C = x_C^1 + x_C^2$, on obtient finalement la convergence en loi des processus (x_C^N, X_D^N) vers le PDP de générateur \mathcal{A}_∞ .

8.4 PDP obtenus par moyenisation

Lorsque l'hypothèse (8.4) n'est pas vérifiée les résultats précédents ne sont plus valables. On étudie maintenant ce qui se passe dans ce cas. On suppose que (8.4) est vraie pour une partie des variables discrètes mais pas nécessairement pour toutes. On est donc amenés à décomposer X_D en

$$X_D = (X_D^1, X_D^2) \in \mathbb{N}^{M_{D,1}} \times \mathbb{N}^{M_{D,2}}.$$

Cela induit une décomposition des sauts en

$$\gamma_D = (\gamma_r^{D,1}, \gamma_r^{D,2}).$$

L'hypothèse (8.4) est remplacée par

$$\gamma_r^{D,1} = 0, r \in S_1.$$

La variable discrète X_D^2 change très rapidement. Ses sauts sont fréquents et non petits. Leur fréquence est de l'ordre de N , leur taille de l'ordre de l'unité. On ne s'attend pas à ce que le processus décrivant l'évolution de cette variable ait une limite. Sous des hypothèses adéquates d'ergodicité, les autres variables dont l'évolution est plus lente ne verront l'effet de cette variable qu'en moyenne.

On va donc chercher le comportement limite des processus décrivant l'évolution de (x_C, X_D^1) .

On a vu au paragraphe 8.3.2 que les réactions de type S_2 peuvent être traitées simplement en introduisant un système auxiliaire. Afin de ne pas alourdir la présentation, nous nous plaçons dans le cas où il n'y a pas de réactions de type S_2 .

Le générateur du processus a la forme suivante

$$\begin{aligned}
& \tilde{\mathcal{A}}_N f(x_C, X_D^1, X_D^2) \\
&= \sum_{r \in \mathcal{R}_C} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D^1, X_D^2) - f(x_C, X_D^1, X_D^2) \right] N \tilde{\lambda}_r(x_C) \\
&+ \sum_{r \in S_1} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D^1, X_D^2 + \gamma_r^{D,2}) - f(x_C, X_D^1, X_D^2) \right] N \tilde{\lambda}_r(x_C, X_D^1, X_D^2) \\
&+ \sum_{r \in \mathcal{R}_{DC} \setminus S_1} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D^1 + \gamma_r^{D,1}, X_D^2 + \gamma_r^{D,2}) - f(x_C, X_D^1, X_D^2) \right] \lambda_r(x_C, X_D^1, X_D^2) \\
&+ \sum_{r \in \mathcal{R}_D} \left[f(x_C, X_D^1 + \gamma_r^{D,1}, X_D^2 + \gamma_r^{D,2}) - f(x_C, X_D^1, X_D^2) \right] \lambda_r(X_D^1, X_D^2).
\end{aligned} \tag{8.11}$$

Contrairement aux cas déjà étudiés, on ne peut pas prendre directement la limite formelle dans le générateur. Comme on ne s'intéresse qu'au comportement des variables x_C et X_D^1 , il est naturel de considérer des fonctions test ne dépendant que de celles-ci. Le problème vient du fait que le générateur appliqué à de telles fonctions fait intervenir des termes d'ordre 1 dépendant de la variable X_D^2 . Cette difficulté est classique pour les problèmes de moyennisation. Une façon de la surmonter est de prendre des "fonctions tests perturbées" (voir [74], [84]).

Encore une fois, nous commençons par travailler formellement. Soit f une fonction de (x_C, X_D^1) , on cherche $f_N(x_C, X_D^1, X_D^2)$ de la forme

$$f_N(x_C, X_D^1, X_D^2) = f(x_C, X_D^1) + \frac{1}{N} f^1(x_C, X_D^1, X_D^2)$$

telle que $\tilde{\mathcal{A}}_N f_N$ ait une limite ne dépendant que des variables (x_C, X_D^1) . La fonction f^1 sera choisie de façon à éliminer les termes d'ordre 1 dépendant de la variable

X_D^2 . On a

$$\begin{aligned}
& \tilde{\mathcal{A}}_N f_N(x_C, X_D^1, X_D^2) \\
&= \sum_{r \in \mathcal{R}_C} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D^1) - f(x_C, X_D^1) \right] N \tilde{\lambda}_r(x_C) \\
&+ \sum_{r \in \mathcal{R}_C} \left[f^1(x_C + \frac{1}{N} \gamma_r^C, X_D^1, X_D^2) - f^1(x_C, X_D^1, X_D^2) \right] \tilde{\lambda}_r(x_C) \\
&+ \sum_{r \in \mathcal{S}_1} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D^1) - f(x_C, X_D^1) \right] N \tilde{\lambda}_r(x_C, X_D^1, X_D^2) \\
&+ \sum_{r \in \mathcal{S}_1} \left[f^1(x_C + \frac{1}{N} \gamma_r^C, X_D^1, X_D^2 + \gamma_r^{D,2}) - f^1(x_C, X_D^1, X_D^2) \right] \tilde{\lambda}_r(x_C, X_D^1, X_D^2) \\
&+ \sum_{r \in \mathcal{R}_{DC} \setminus \mathcal{S}_1} \left[f(x_C + \frac{1}{N} \gamma_r^C, X_D^1 + \gamma_r^{D,1}) - f(x_C, X_D^1) \right] \lambda_r(x_C, X_D^1, X_D^2) \\
&+ \frac{1}{N} \sum_{r \in \mathcal{R}_{DC} \setminus \mathcal{S}_1} \left[f^1(x_C + \frac{1}{N} \gamma_r^C, X_D^1 + \gamma_r^{D,1}, X_D^2 + \gamma_r^{D,2}) - f^1(x_C, X_D^1, X_D^2) \right] \lambda_r(x_C, X_D^1, X_D^2) \\
&+ \sum_{r \in \mathcal{R}_D} [f(x_C, X_D^1 + \gamma_r^{D,1}, X_D^2) - f(x_C, X_D^1)] \lambda_r(X_D^1, X_D^2) \\
&+ \frac{1}{N} \sum_{r \in \mathcal{R}_D} [f^1(x_C, X_D^1 + \gamma_r^{D,1}, X_D^2 + \gamma_r^{D,2}) - f^1(x_C, X_D^1, X_D^2)] \lambda_r(X_D^1, X_D^2).
\end{aligned} \tag{8.12}$$

En prenant $N \rightarrow \infty$, on obtient

$$\begin{aligned}
& \tilde{\mathcal{A}}_\infty f(x_C, X_D^1) \\
&= \left(\sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C) \gamma_r^C + \sum_{r \in S_1} \tilde{\lambda}_r(x_C, X_D^1, X_D^2) \gamma_r^C \right) \cdot \nabla_{x_C} f(x_C, X_D^1) \\
&+ \sum_{r \in S_1} [f^1(x_C, X_D^1, X_D^2 + \gamma_r^{D,2}) - f^1(x_C, X_D^1, X_D^2)] \tilde{\lambda}_r(x_C, X_D^1, X_D^2) \quad (8.13) \\
&+ \sum_{r \in \mathcal{R}_{DC} \setminus S_1} [f(x_C, X_D^1 + \gamma_r^{D,1}) - f(x_C, X_D^1)] \lambda_r(x_C, X_D^1, X_D^2) \\
&+ \sum_{r \in \mathcal{R}_D} [f(x_C, X_D^1 + \gamma_r^{D,1}, X_D^2) - f(x_C, X_D^1)] \lambda_r(X_D^1, X_D^2).
\end{aligned}$$

On va prendre f^1 telle que

$$\sum_{r \in S_1} [f^1(x_C, X_D^1, X_D^2 + \gamma_r^{D,2}) - f^1(x_C, X_D^1, X_D^2)] \tilde{\lambda}_r(x_C, X_D^1, X_D^2) = g_{x_C, X_D^1}(X_D^2) \quad (8.14)$$

où g permet d'éliminer la dépendance en X_D^2 . Pour que cela soit possible, il faut que certaines hypothèses soient satisfaites par les taux de réactions et par la fonction g . Pour tout (x_C, X_D^1) , on définit le générateur

$$\mathcal{A}_{x_C, X_D^1} h(X_D^2) = \sum_{r \in S_1} [h(X_D^2 + \gamma_r^{D,2}) - h(X_D^2)] \tilde{\lambda}_r(x_C, X_D^1, X_D^2).$$

On suppose que pour tout (x_C, X_D^1) , le processus associé est ergodique, de mesure invariante ν_{x_C, X_D^1} . De plus on suppose que pour tout $g \in \mathcal{B}_b(\mathbb{N}^{M_D, 2})$ telle que

$$\bar{g} = \int_{\mathbb{N}^{M_D, 2}} g(X_D^2) \nu_{x_C, X_D^1}(dX_D^2) = 0$$

l'équation

$$\mathcal{A}_{x_C, X_D^1} h(X_D^2) = g$$

a une unique solution, notée $\mathcal{A}_{x_C, X_D^1}^{-1}g$. Ces hypothèses sont classiques et naturelles.

On peut maintenant prendre

$$\begin{aligned}
& g_{x_C, X_D^1}(X_D^2) \\
&= \left(\sum_{r \in S_1} \bar{\lambda}_r(x_C, X_D^1) - \tilde{\lambda}_r(x_C, X_D^1, X_D^2) \right) \cdot \nabla_{x_C} f(x_C, X_D^1) \\
&+ \sum_{r \in \mathcal{R}_{DC} \setminus S_1} [f(x_C, X_D^1 + \gamma_r^{D,1}) - f(x_C, X_D^1)] (\bar{\lambda}_r(x_C, X_D^1) - \lambda_r(x_C, X_D^1, X_D^2)) \\
&+ \sum_{r \in \mathcal{R}_D} [f(x_C, X_D^1 + \gamma_r^{D,1}, X_D^2) - f(x_C, X_D^1)] (\bar{\lambda}_r(x_C, X_D^1) - \lambda_r(X_D^1, X_D^2))
\end{aligned} \tag{8.15}$$

et f^1 définit par (8.14). Où on a noté

$$\begin{aligned}
\bar{\lambda}_r(x_C, X_D^1) &= \int_{N^{M_D, 2}} \tilde{\lambda}_r(x_C, X_D^1, X_D^2) \nu_{x_C, X_D^1}(dX_D^2), \quad r \in \mathcal{R}_C \cup (\mathcal{R}_{DC} \setminus S_1), \\
\bar{\lambda}_r(x_C, X_D^1) &= \int_{N^{M_D, 2}} \tilde{\lambda}_r(X_D^1, X_D^2) \nu_{x_C, X_D^1}(dX_D^2), \quad r \in \mathcal{R}_D.
\end{aligned} \tag{8.16}$$

Alors, on a le générateur limite :

$$\begin{aligned}
& \tilde{\mathcal{A}}_\infty f(x_C, X_D^1) \\
&= \left(\sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C) + \sum_{r \in S_1} \bar{\lambda}_r(x_C, X_D^1) \right) \cdot \nabla_{x_C} f(x_C, X_D^1) \\
&+ \sum_{r \in (\mathcal{R}_{DC} \setminus S_1) \cup \mathcal{R}_D} [f(x_C, X_D^1 + \gamma_r^{D,1}, X_D^2) - f(x_C, X_D^1)] \bar{\lambda}_r(x_C, X_D^1).
\end{aligned} \tag{8.17}$$

C'est le générateur d'un PDP dont les coefficients ont été moyennés par rapport à la variable rapide X_D^2 .

Nous aurons besoin d'hypothèses supplémentaires pour justifier tout cela. Nous

supposons que les taux moyennés définis en (8.16) sont de classe C^1 par rapport à x_C . De plus, si les taux initiaux sont à dérivées bornées uniformément par rapport aux variables X_D^1 et X_D^2 alors les taux moyennés sont à dérivées uniformément bornées par rapport à X_D^1 . Nous supposons aussi que si $g_{x_C, X_D^1}(X_D^2)$ est bornée par rapport aux trois variables, il en est de même de $h_{x_C, X_D^1}(X_D^2)$ la solution de

$$\mathcal{A}_{x_C, X_D^1} h_{x_C, X_D^1}(X_D^2) = g_{x_C, X_D^1}.$$

De plus, si $g_{x_C, X_D^1}(X_D^2)$ est de classe C^1 par rapport à x_C , il en est de même pour $h_{x_C, X_D^1}(X_D^2)$. Enfin si les dérivées de $g_{x_C, X_D^1}(X_D^2)$ par rapport à x_C sont bornées par rapport aux trois variables, il en est de même pour $h_{x_C, X_D^1}(X_D^2)$.

Remarque 11 *Ces hypothèses seront vérifiées par exemple si X_D^2 vit dans un ensemble fini et*

- i) les taux des réactions de S_1 sont bornés, C^1 par rapport à x_C et de dérivées bornées*
- ii) la première valeur propre non nulle de \mathcal{A}_{x_C, X_D^1} est uniformément éloigné de 0.*

Dans i), on pourrait en fait affaiblir les hypothèses de bornitude. Il suffit en effet qu'elles soient vérifiées pour les taux de réactions tronqués (voir Etape 3 de la preuve du Théorème 9).

Nous supposerons aussi que les taux des réactions de $\mathcal{R}_{DC} \setminus S_1$ et de \mathcal{R}_D sont bornés, que tous les taux de réactions sont de classe C^1 par rapport à x_C et enfin que si on définit

$$\bar{F}(x_C, X_D^1) = \sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C) + \sum_{r \in S_1} \bar{\lambda}_r(x_C, X_D^1)$$

alors le système différentiel

$$\begin{cases} \frac{dx_C}{dt} = \bar{F}(x_C, X_D^1), \\ x_C(0) = x_{C,0}, \end{cases}$$

a une unique solution globale pour tout X_D^1 .

Nous énonçons et démontrons le résultat rigoureux pour ce passage à la limite.

Théorème 12 *Sous les hypothèses précédentes, considérons la famille de processus $x^N = (x_C^N, X_D^{N,1}, X_D^{N,2})$ de générateur $\tilde{\mathcal{A}}_N$ et démarrant en $x^N(0) = (x_C^N(0), X_D^{N,1}(0), X_D^{N,2}(0))$. Supposons*

$$(x_C^N(0), X_D^1(0)) \rightarrow (x_C, X_{D,0}^1)$$

en distribution dans $\mathbb{R}^{M_C} \times \mathbb{N}^{M_{D,1}}$, alors $(x_C^N, X_D^{N,1})$ converge en distribution dans $\mathbb{D}(\mathbb{R}^{M_C} \times \mathbb{N}^{M_{D,1}})$ vers (x_C, X_D^1) , le PDP de générateur $\tilde{\mathcal{A}}_\infty$ démarrant en $(x_C(0), X_D^1(0))$.

Preuve. On procède de la même manière que dans la preuve du Théorème 9. On commence par supposer que tous les taux de réactions sont bornés et à dérivées bornées.

La preuve de la tension des lois des processus $(x_C^N, X_D^{N,1})$ s'obtient de la même façon. Comme on ne s'intéresse qu'aux lois, on réécrit $(x_C^N, X_D^{N,1})$ sous la forme

$$\begin{aligned} x_C^N(t) &= x_C^N(0) + \sum_{r \in \mathcal{R}_C} \frac{1}{N} \gamma_r^C Y_r \left(N \int_0^t \tilde{\lambda}_r(x_C^N(s)) ds \right) \\ &+ \sum_{r \in \mathcal{S}_1} \frac{1}{N} \gamma_r^C Y_r \left(N \int_0^t \tilde{\lambda}_r(x_C^N(s), X_D^{N,1}(s), X_D^{N,2}(s)) ds \right) \\ &+ \sum_{r \in \mathcal{R}_{DC} \setminus \mathcal{S}_1} \frac{1}{N} \gamma_r^C Y_r \left(\int_0^t \lambda_r(x_C^N(s), X_D^{N,1}(s), X_D^{N,2}(s)) ds \right) \end{aligned}$$

et

$$\begin{aligned} X_D^{N,1}(t) &= X_D^{N,1}(0) + \sum_{r \in \mathcal{R}_{DC} \setminus \mathcal{S}_1} \gamma_r^{D,1} Y_r \left(\int_0^t \lambda_r(x_C^N(s), X_D^{N,1}(s), X_D^{N,2}(s)) ds \right) \\ &+ \sum_{r \in \mathcal{R}_D} \gamma_r^{D,1} Y_r \left(\int_0^t \lambda_r(X_D^{N,1}(s), X_D^{N,2}(s)) ds \right) \end{aligned}$$

Les processus $X_D^{N,2}$ vérifient une équation similaire et $(Y_r)_{r \in \mathcal{R}}$ est une famille de processus de Poisson indépendants. En utilisant les mêmes arguments que précédemment, on obtient la tension.

Soit maintenant $f(x_C, X_D^1)$ bornée, de classe C^1 par rapport à x_C et à dérivées bornées uniformément en X_D^1 . On définit f^1 par (8.14) et (8.15) et $f_N = f + \frac{1}{N} f^1$. Alors, si $P_{x,N}$ est la loi de $(x_C^N, X_D^{N,1}, X_D^{N,2})$,

$$f_N(x_C(t), X_D^1(t), X_D^2(t)) - f_N(x_C(0), X_D^1(0), X_D^2(0))$$

$$- \int_0^t \tilde{\mathcal{A}}_N f_N(x_C(s), X_D^1(s), X_D^2(s)) ds$$

est une $P_{x,N}$ martingale.

Avec les hypothèses faites, il est facile de voir que f^1 est bornée, de classe C^1 par rapport à x_C et à dérivées bornées uniformément en X_D^1 et X_D^2 . On en déduit que f_N converge uniformément vers f et que $\tilde{\mathcal{A}}_N f_N$ converge uniformément vers $\tilde{\mathcal{A}}_\infty f$. En procédant comme dans la preuve du Théorème 9, on montre que tous les points d'accumulation de la suite $(\tilde{P}_{x,N})$ des lois de $(x_C^N, X_D^{N,1})$ sont solution du problème de martingale associé à $\tilde{\mathcal{A}}_\infty$.

De plus, nos hypothèses et le Théorème 5 implique que le PDP associé à ce générateur est l'unique solution à ce problème de martingale et donc toute la suite $(\tilde{P}_{x,N})$ converge vers la loi de ce PDP.

Finalement on reproduit l'étape 3 de la preuve du Théorème 9 pour enlever l'hypothèse supplémentaire faite sur les taux de réactions. ■

8.5 PDP discontinus obtenus par "switching" singulier

Nous traitons maintenant une situation différente où le comportement du système n'a pas la même échelle de temps suivant la valeur d'une des variables discrète. Nous l'appellons θ et supposons qu'elle ne prend que deux valeurs 0 et 1. Pour simplifier la présentation, on suppose qu'il n'y a pas d'autre variable discrète.

On suppose toujours que les nombres de molécules des espèces dans C sont l'ordre de N et on travaille avec la variable $x_C = \frac{1}{N} X_C$.

Lorsque $\theta = 0$, les vitesses de réactions de $\mathcal{R}_C \cup \mathcal{R}_{DC}$ sont l'ordre de N mais elles sont beaucoup plus grandes lorsque $\theta = 1$. On introduit donc un nouveau paramètre ε et on suppose

$$\lambda_r(x_C, 0) = O(N), \quad r \in \mathcal{R}_C \cup \mathcal{R}_{DC},$$

$$\lambda_r(x_C, 1) = O\left(\frac{N}{\varepsilon}\right), \quad r \in \mathcal{R}_C \cup \mathcal{R}_{DC}.$$

Il est immédiat - et cela ressortira de la preuve ci-dessous - de voir que ce système ne peut avoir de limite que si on suppose que le système ne peut pas rester plus longtemps qu'un temps de l'ordre de ε avec $\theta = 1$. Autrement dit la vitesse de réaction pour passer de $\theta = 1$ à $\theta = 0$ doit être de l'ordre de $\frac{1}{\varepsilon}$:

$$\lambda_r(x_C, 1) = O\left(\frac{1}{\varepsilon}\right), \quad r \in \mathcal{R}_D.$$

On note

$$\tilde{\lambda}_r(x_C, 0) = \frac{1}{N} \lambda_r(x_C, 0), \quad r \in \mathcal{R}_C \cup \mathcal{R}_{DC},$$

$$\tilde{\lambda}_r(x_C, 1) = \frac{\varepsilon}{N} \lambda_r(x_C, 1), \quad r \in \mathcal{R}_C \cup \mathcal{R}_{DC},$$

et

$$\lambda(x_C, 0) = \sum_{r \in \mathcal{R}_D} \lambda_r(x_C, 0),$$

$$\lambda(x_C, 1) = \varepsilon \sum_{r \in \mathcal{R}_D} \lambda_r(x_C, 1).$$

Le générateur a alors la forme

$$\tilde{\mathcal{A}}_{N,\varepsilon} f(x_C, 0) = \sum_{r \in \mathcal{R}_C} \left[f\left(x_C + \frac{1}{N} \gamma_r^C, 0\right) - f(x_C, 0) \right] N \tilde{\lambda}_r(x_C, 0)$$

$$+ [f(x_C, 1) - f(x_C, 0)] \lambda(x_C, 0),$$

$$\tilde{\mathcal{A}}_{N,\varepsilon} f(x_C, 1) = \frac{1}{\varepsilon} \sum_{r \in \mathcal{R}_C} \left[f\left(x_C + \frac{1}{N} \gamma_r^C, 1\right) - f(x_C, 1) \right] N \tilde{\lambda}_r(x_C, 1)$$

$$+ \frac{1}{\varepsilon} [f(x_C, 0) - f(x_C, 1)] \lambda(x_C, 1).$$

Pour comprendre vers quel comportement limite converge ces processus, prenons

d'abord la limite $N \rightarrow \infty$. On obtient le générateur :

$$\begin{aligned} \tilde{\mathcal{A}}_{\infty,\varepsilon} f(x_C, 0) &= \left(\sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C, 0) \gamma_r^C \right) \cdot \nabla_{x_C} f(x_C, 0) \\ &\quad + [f(x_C, 1) - f(x_C, 0)] \lambda(x_C, 0), \\ \tilde{\mathcal{A}}_{\infty,\varepsilon} f(x_C, 1) &= \frac{1}{\varepsilon} \left(\sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C, 1) \gamma_r^C \right) \cdot \nabla_{x_C} f(x_C, 1) \\ &\quad + \frac{1}{\varepsilon} [f(x_C, 0) - f(x_C, 1)] \lambda(x_C, 1). \end{aligned}$$

A la limite $\varepsilon \rightarrow 0$, la variable θ ne prend plus la valeur 1 et on cherche le comportement limite de la seule variable discrète x_C . De la même façon que dans le paragraphe 8.4, on va chercher un générateur limite sur des fonctions de x_C , la méthode est cependant un peu différente. Etant donné une fonction $f(x_C)$, on pose

$$f(x_C, 0) = f(x_C)$$

ce qui est assez naturel car la variable θ est identiquement égale à 0 à la limite. Pour obtenir un générateur qui reste borné, il est naturel de choisir $f(\cdot, 1)$ tel que

$$F_1(x_C) \cdot \nabla_{x_C} f(x_C, 1) - \lambda(x_C, 1) f(x_C, 1) = \lambda(x_C, 1) f(x_C, 0) \quad (8.18)$$

où on a noté

$$F_1(x_C) = \sum_{r \in \mathcal{R}_C} \gamma_r^C \tilde{\lambda}_r(x_C, 1). \quad (8.19)$$

Formellement, (8.18) est vérifié si on prend

$$f(x_C, 1) = \int_0^\infty f(\phi_1(t, x_C), 0) \lambda(\phi_1(t, x_C), 1) e^{-\int_0^t \lambda(\phi_1(s, x_C), 1) ds} dt, \quad (8.20)$$

où ϕ_1 est le flot associé au champs de vecteur F_1 . Alors, on a

$$\tilde{\mathcal{A}}_{\infty,\varepsilon}f(x_C, 1) = 0$$

et

$$\begin{aligned} \tilde{\mathcal{A}}_{\infty,\varepsilon}f(x_C, 0) &= \left(\sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C, 0) \gamma_r^C \right) \nabla_{x_C} f(x_C, 0) \\ &+ \lambda(x_C, 0) \int_0^\infty (f(\phi_1(t, x_C), 0) - f(x_C, 0)) \lambda(\phi_1(t, x_C), 1) e^{-\int_0^t \lambda(\phi_1(s, x_C), 1) ds} dt. \end{aligned}$$

On a ainsi deviné le générateur du processus limite

$$\begin{aligned} \tilde{\mathcal{A}}_\infty f(x_C) &= \left(\sum_{r \in \mathcal{R}_C} \tilde{\lambda}_r(x_C, 0) \gamma_r^C \right) \nabla_{x_C} f(x_C) \\ &+ \lambda(x_C, 0) \int_0^\infty (f(\phi_1(t, x_C)) - f(x_C)) \lambda(\phi_1(t, x_C), 1) e^{-\int_0^t \lambda(\phi_1(s, x_C), 1) ds} dt \end{aligned}$$

C'est encore le générateur d'un PDP mais il y a une discontinuité de la variable continue qui n'est pas du au même phénomène qu'au paragraphe 8.3.2. Il y a en fait "switching" mais sur un temps très court et cela engendre la discontinuité.

Théorème 13 *Supposons que tous les taux de réactions sont bornés, C^1 par rapport à x_C et à dérivées bornées et que $\lambda(\cdot, 1)$ est minoré :*

$$\lambda(x_C, 1) \geq \alpha > 0, \quad x_C \in \mathbb{R}^{M_C}.$$

Soit $(x_C^{\varepsilon, N}, \theta^{\varepsilon, N})$ un processus à valeurs dans $\mathbb{R}^{M_C} \times \{0, 1\}$ de générateur $\tilde{\mathcal{A}}_{N,\varepsilon}$ et démarrant en $(x_C^{\varepsilon, N}(0), \theta^{\varepsilon, N}(0))$. Supposons aussi que $x_C^{\varepsilon, N}(0)$ converge en distribution dans \mathbb{R}^{M_C} vers $x_C(0)$ quand $\varepsilon \rightarrow 0$ et $N \rightarrow \infty$.

Alors la suite de processus $(x_C^{\varepsilon, N})$ converge en distribution dans $L^p(0, T; \mathbb{R}^{M_C})$ pour tout $1 \leq p < \infty$ vers le PDP de générateur $\tilde{\mathcal{A}}_\infty$ et démarrant en $x_C(0)$.

Remarque 14 *En prenant d'abord la limite d'abord $N \rightarrow \infty$ puis $\varepsilon \rightarrow 0$ comme ci-dessus, on voit que le processus associé à $\tilde{\mathcal{A}}_{\infty,\varepsilon}$ est à trajectoires continues. Ce qui n'est pas le cas du processus associé à $\tilde{\mathcal{A}}_\infty$. Il y a donc création de discontinuité*

8.5 PDP DISCONTINUS OBTENUS PAR "SWITCHING" SINGULIER 133

dans le passage à la limite $\epsilon \rightarrow 0$. La convergence dans l'espace de Skorohod ne permet pas la création de discontinuité, c'est pour cela que nous obtenons un résultat de convergence pour la topologie L^p . Notons aussi que nous montrons la convergence quand on prend simultanément la limite sur ϵ et N .

Preuve. Pour $N \geq 1$, $\epsilon > 0$, on prend $(x_C^{\epsilon,N}, \theta^{\epsilon,N})$, processus à valeurs dans $\mathbb{R}^{M_C} \times \{0, 1\}$ de générateur $\tilde{\mathcal{A}}_{N,\epsilon}$.

On montre d'abord la tension de la famille des lois de ses processus.

On note M un majorant des taux de réaction et de leurs dérivées par rapport à x_C .

Choisissons d'abord la fonction test

$$f(x_C, 1) = 1, f(x_C, 0) = 0, x_C \in \mathbb{R}^{M_C}.$$

Alors, comme $\tilde{\mathcal{A}}_{N,\epsilon} f(x_C, \theta) = \lambda(x_C, 0) \mathbf{1}_{\theta=0} - \frac{1}{\epsilon} \lambda(x_C, 1) \mathbf{1}_{\theta=1}$, on sait que

$$\begin{aligned} & \mathbf{1}_{\theta^{\epsilon,N}(t)=1} - \mathbf{1}_{\theta^{\epsilon,N}(0)=1} \\ & - \int_0^t \lambda(x_C^{\epsilon,N}(s), 0) \mathbf{1}_{\theta^{\epsilon,N}(s)=0} - \frac{1}{\epsilon} \lambda(x_C^{\epsilon,N}(s), 1) \mathbf{1}_{\theta^{\epsilon,N}(s)=1} ds \end{aligned}$$

est une martingale. On en déduit

$$\frac{\alpha}{\epsilon} \mathbb{E} \left(\int_0^t \mathbf{1}_{\theta^{\epsilon,N}(s)=1} ds \right) \leq MT + 1. \quad (8.21)$$

Et donc, $\theta^{\epsilon,N} \rightarrow 0$ dans $L^1(\Omega \times [0, T]; \{0, 1\})$. On écrit maintenant

$$\begin{aligned} x_C^{\epsilon,N}(t) &= x_C^{\epsilon,N}(0) \\ &+ \sum_{r \in \mathcal{R}_C} \frac{\gamma_r^C}{N} Y_r \left(\int_0^t N \tilde{\lambda}_r(x_C^{\epsilon,N}(s), 0) \mathbf{1}_{\theta^{\epsilon,N}(s)=0} + \frac{N}{\epsilon} \tilde{\lambda}_r(x_C^{\epsilon,N}(s), 1) \mathbf{1}_{\theta^{\epsilon,N}(s)=1} ds \right) \end{aligned}$$

Alors

$$|x_C^{\epsilon,N}(t)| \leq |x_C^{\epsilon,N}(0)| + \sum_{r \in \mathcal{R}_C} \frac{|\gamma_r^C|}{N} Y_r \left(NM \left(t + \frac{1}{\epsilon} \int_0^t \mathbf{1}_{\theta^{\epsilon,N}(s)=1} ds \right) \right) \quad (8.22)$$

Comme $\sup_{[0,T]} \frac{1}{N} Y_r(Nt)$ est borné presque sûrement, en utilisant (8.21) et (8.22), on montre que pour tout $K \geq 0$

$$\mathbb{P}(\sup_{[0,T]} |x_C^\epsilon, N(t)| \geq K) \leq \varepsilon(K)$$

où $\varepsilon(K) \rightarrow 0$ quand $K \rightarrow \infty$. De même si $0 = t_0 < t_1 < \dots < t_n = T$ est une subdivision de $[0, T]$, on obtient en utilisant la croissance des processus Y_r :

$$\begin{aligned} & \sum_{i=0}^{n-1} |x_C^{\epsilon, N}(t_{i+1}) - x_C^{\epsilon, N}(t_i)| \\ & \leq \sum_{i=0}^{n-1} \sum_{r \in \mathcal{R}_C} \frac{|\gamma_r^C|}{N} \left[Y_r \left(\int_0^{t_{i+1}} N \tilde{\lambda}_r(x_C^{\epsilon, N}(s), 0) \mathbf{1}_{\theta^{\epsilon, N}(s)=0} + \frac{N}{\epsilon} \tilde{\lambda}_r(x_C^{\epsilon, N}(s), 1) \mathbf{1}_{\theta^{\epsilon, N}(s)=1} ds \right) \right. \\ & \quad \left. - Y_r \left(\int_0^{t_i} N \tilde{\lambda}_r(x_C^{\epsilon, N}(s), 0) \mathbf{1}_{\theta^{\epsilon, N}(s)=0} + \frac{N}{\epsilon} \tilde{\lambda}_r(x_C^{\epsilon, N}(s), 1) \mathbf{1}_{\theta^{\epsilon, N}(s)=1} ds \right) \right] \\ & = \sum_{r \in \mathcal{R}_C} \frac{|\gamma_r^C|}{N} Y_r \left(\int_0^T N \tilde{\lambda}_r(x_C^{\epsilon, N}(s), 0) \mathbf{1}_{\theta^{\epsilon, N}(s)=0} + \frac{N}{\epsilon} \tilde{\lambda}_r(x_C^{\epsilon, N}(s), 1) \mathbf{1}_{\theta^{\epsilon, N}(s)=1} ds \right) \end{aligned}$$

On obtient donc que les processus $(x_C^{\epsilon, N})$ sont bornés en probabilité dans $BV(0, T; \mathbb{R}^{M_C}) \cap L^\infty(0, T; \mathbb{R}^{M_C})$ et leur lois sont donc tendues dans $L^p(0, T; \mathbb{R}^{M_C})$ pour $1 \leq p < \infty$.

En utilisant le théorème de Skohorod, quitte à changer d'espace de probabilité, on a donc une sous suite $(x_C^{\epsilon_k}, N_k, \theta^{\epsilon_k, N_k})$ qui converge presque sûrement dans $L^p(0, T; \mathbb{R}^{M_C}) \times L^1(0, T; \{0, 1\})$ vers (x_C, θ) où θ est constant égal à 0.

Soit φ une fonction test dans $C_b^1(\mathbb{R}^{M_C})$. Comme on prend la limite simultanée sur N et ϵ , on construit f de manière différente de ci-dessus. On pose

$$f(x_C, 0) = \varphi(x_C), \quad x_C \in \mathbb{R}^{M_C}$$

et on prend $f(x_C, 1)$ vérifiant

$$\sum_{r \in \mathcal{R}_C} \left[f(x_C + \frac{1}{N} \gamma_r^C, 1) - f(x_C, 1) \right] N \tilde{\lambda}_r(x_C, 1) - \lambda(x_C, 1) f(x_C, 1) \tag{8.23}$$

$$= \lambda(x_C, 1) f(x_C, 0).$$

Pour construire $f(x_C, 1)$, on considère le processus $y_N(\cdot, y^0)$ de générateur

$$A\psi(y) = \sum_{r \in \mathcal{R}_C} \left[\psi(y + \frac{1}{N} \gamma_r^C) - \psi(y) \right] N \tilde{\lambda}_r(y, 1)$$

et démarrant en y^0 . On définit aussi le semi-groupe $(S_t^N)_{t \geq 0}$ sur $\mathcal{B}_b(\mathbb{R}^{M_C})$ défini par

$$S_t^N \psi(y^0) = \mathbb{E} \left(\psi(y_N(t, y^0)) e^{-\int_0^t \lambda(y_N(s, y^0), 1) ds} \right).$$

Alors, on peut prendre $f(x_C, 1)$ défini par

$$f(x_C, 1) = \int_0^\infty S_t^N (\lambda(\cdot, 1) \varphi) (x_C) dt.$$

Il est classique de vérifier que $f(x_C, 1)$ est bien défini grâce à l'hypothèse de minoration sur $\lambda(x_C, 1)$ et que (8.23) est bien vérifié. De plus, si $\varphi \in C_b^1(\mathbb{R}^{M_C})$, il en est de même pour $f(\cdot, 1)$.

On a alors

$$\tilde{\mathcal{A}}_{N,\epsilon} f(x_C, 1) = 0$$

et

$$\tilde{\mathcal{A}}_{N,\epsilon} f(x_C, 0) = \sum_{r \in \mathcal{R}_C} \left[f(x_C + \frac{1}{N} \gamma_r^C, 0) - f(x_C, 0) \right] N \tilde{\lambda}_r(x_C, 0)$$

$$+ \left[\int_0^\infty S_t^N (\lambda(\cdot, 1) \varphi(\cdot, 0)) (x_C) dt - f(x_C, 0) \right] \lambda(x_C, 0).$$

En utilisant les résultats de la section 8.2, on montre que pour tout $x_C \in \mathbb{R}^{M_C}$

$$S_t^N (\lambda(\cdot, 1) \varphi(\cdot, 0)) (x_C) \rightarrow f(\phi_1(t, x_C), 0) e^{-\int_0^t \lambda(\phi_1(s, x_C), 1) ds}$$

presque partout en $t \geq 0$ et en restant majoré par $|f(\cdot, 0)|_\infty e^{-\alpha t}$. (Rappelons que

ϕ_1 est le flot associé au champs de vecteur $F_1(x_C) = \sum_{r \in \mathcal{R}_C} \gamma_r^C \tilde{\lambda}_r(x_C, 1)$. On en déduit que

$$\tilde{\mathcal{A}}_{N,\epsilon} f(x_C, 0) \rightarrow \tilde{\mathcal{A}}_\infty \varphi(x_C)$$

pour tout x_C en restant borné.

Ecrivons maintenant que pour $0 \leq t_1, \dots, t_\ell \leq t$ et $\psi \in C_b(\mathbb{R}^{M_C \times \ell})$

$$\begin{aligned} & \mathbb{E} \left(\left[f(x_C^{\epsilon,N}(t), \theta^{\epsilon,N}(t)) - f(x_C^{\epsilon,N}(0), \theta^{\epsilon,N}(0)) \right. \right. \\ & \left. \left. - \int_0^t \tilde{\mathcal{A}}_{N,\epsilon} f(x_C^{\epsilon,N}(s), 0) \mathbb{1}_{\theta^{\epsilon,N}(s)=0} ds \right] \psi(x_C^{\epsilon,N}(t_1), \dots, x_C^{\epsilon,N}(t_\ell)) \right) = 0. \end{aligned}$$

La convergence presque sûr de $x_C^{\epsilon,N}$ vers x_C dans $L^p(0, T; \mathbb{R}^{M_C})$ permet d'affirmer que, quitte à extraire une nouvelle sous suite,

$$x_C^{\epsilon_k, N_k}(t) \rightarrow x_C(t)$$

presque partout par rapport à t et à ω . On peut donc prendre la limite dans l'égalité ci-dessous et obtenir

$$\mathbb{E} \left(\left[\varphi(x_C(t)) - \varphi(x_C(0)) - \int_0^t \tilde{\mathcal{A}}_\infty \varphi(x_C(s)) ds \right] \psi(x_C(t_1), \dots, x_C(t_\ell)) \right) = 0$$

pour presque tout t, t_1, \dots, t_ℓ . On remarque maintenant que la preuve du théorème 5 se généralise immédiatement au cas où une solution du problème de martingale dans ce sens affaibli est considéré. On en déduit que x_C est en fait à trajectoire dans $\mathbb{D}(\mathbb{R}^+, \mathbb{R}^{M_C})$ (où plus exactement qu'il existe une version de x_C à trajectoires dans $\mathbb{D}(\mathbb{R}^+, \mathbb{R}^{M_C})$).

On conclut donc comme dans les preuves précédentes que toute la suite $x_C^{\epsilon,N}$ converge en loi vers x_C dans $L^p(0, T; \mathbb{R}^{M_C})$. ■

Remarque 15 *Comme nous obtenons la convergence en loi dans $L^p(0, T; \mathbb{R}^{M_C})$, et non dans l'espace de Skohorod, nous n'avons pas de bornes a posteriori uniformes en temps et nous ne pouvons pas appliquer le raisonnement utilisé auparavant pour enlever les hypothèses de bornitudes sur les taux de réactions.*

Bibliographie

- [1] Adalsteinsson D., McMillen D., Elston T.C., *Biochemical Network Stochastic Simulator (BioNetS) : software for stochastic modeling of biochemical networks*, BMC Bioinformatics, 2004,
- [2] Aldous D., *Deterministic and Stochastic Models for Coalescence (Aggregation, Coagulation) : A Review of the Mean-Field Theory for Probabilists*, Bernoulli, vol. 5, p. 3-48, 1999,
- [3] Aurell E., Brown S., Johanson J., Sneppen Kim, *Stability puzzles in phage λ* , The American Physical Society, 2002,
- [4] Auger P, de la Para RB, Poggiale JC, Sanchez E, Huu TN, *Aggregation of variables and applications to population dynamics*, Structured Population Models in Biology and Epidemiology, LNM 1936, Mathematical Biosciences Subseries. Edited by Magal P, Ruan S, Berlin : Springer 2008 :209-263.
- [5] Alfonsi, A., Cances E., Turinici, G., Di Ventura, B., Huisinga W., *Exact simulation of hybrid and deterministic models for biochemical systems*, Inria, 2004,
- [6] Alfonsi A, Cances E, Turinici G, Di Ventura B, Huisinga W : *Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems*, ESAIM Proceedings, Volume 14 2005 :1,13,
- [7] 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, p. 891-894, 1976,
- [8] Allen M., Tildesley D., *Computer simulation of liquids*, Clarendon Press, New York, 1988,
- [9] 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, p. 129-138, 2004,
- [10] Arkin, A., Ross, J., McAdams, *Stochastic Kinetic Analysis of Developmental Pathway Bifurcation in Phage λ -Infected Escherichia coli Cells*, Genetics 1998,
- [11] Arnold V, *Supplementary chapters to the theory of ordinary differential equations*, Russian Nauka, Moscow, 1978.
- [12] Artstein Z, *Averaging of time-varying differential equations revisited*, Journal of Differential Equations 2007, 243(2) :146-167.
- [13] Bartholomay A.F., *Statistical Fluctuations in Auticatalytic reactions*, Phd thesis, 1957,
- [14] Balding D., Green N., *Diffusion-controlled reactions in one dimension : Exact solutions and deterministic approximations*, Phys.Rev.A, vol. 40, p. 4585-4592, 1989,
- [15] 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, p. 11250-11255, 2000,
- [16] Ball K., Kurtz T., Popovic L., Rempala G., *Asymtotic analysis of multiscales aproximations to reactions networks*, Annals of applied probability, to appear,
- [17] Berglund N., Gentz B., *Geometric singular perturbation theory for stochastic differential equations*, J. Differential Equations, vol. 191, p. 1-54, 2003,
- [18] Bentele M., Eils R., *General Stochastic Hybrid Method for the Simulation of Chemical Reaction Processes in Cells*,
- [19] BlakeW., Kaern M., Cantor C., Collins J., *Noise in eukaryotic gene expression*, Nature, vol.422, p. 633-637, 2003.
- [20] 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, p. 1-14, 2005,
- [21] Bortolussi L., Policriti A., *Hybrid approximation of stochastic process algebras for systems biology*, Proceedings of the 17th World Congress The International Federation of Automatic Control, 2008,
- [22] Burrage K., Tian T., Burrage P., *A multi-scale approach for simulating chemical reaction systems*, Progr.Biophys.Mol.Bio., vol. 85, p. 217-234, 2004,

- [23] Barucha-Reid A : *Elements of the Theory of Markov Processes and their Applications*, McGraw-Hill Book Co., New York 1960,
- [24] Capasso V., Morale D., *Stochastic modelling of tumour-induced angiogenesis*, Journal. Mathematical Biology, 2008,
- [25] Cassandras C.G., Lygeros J., *Stochastic Hybrid Systems*,
- [26] Cercignani C., Gerasimenko V., Petrina D., *Many-Particle Dynamics and Kinetic Equations*, Kluwer, Dordrecht, 1997,
- [27] Cook D. L., Gerber A. N., Tapscott S. J., *Modeling stochastic gene expression : Implications for haploinsufficiency*, Proc. Natl. Acad. Sci. USA, vol. 95, p. 15641-15646, 1998,
- [28] Cristini V., Frieboes H.B., Gatenby R., Caserta S., Ferrari M., Sinek J., *Clinical Cancer Research*, 2005,
- [29] R. Darling, *Fluid Limits of Pure Jump Markov Processes : A Practical Guide*, man- uscript, 2002.
- [30] R. Darling, J. Norris, *Structure of large random hypergraphs*, Ann.Appl.Prob., vol. 15, p. 125-152, 2005.
- [31] M.H.A. Davis, *Markov models and optimization*. Chapman & Hall, London, 1993.
- [32] Delbrück M., *Statistical Fluctuations in Autocatalytic Reactions*, J.Chem.Phys., vol. 8, p. 120-124, 1940,
- [33] Dufour F., Costa OLV., *Stability of piecewise-deterministic Markov processes*, Decision and Control (Proceedings), 1999,
- [34] Costa OLV., Dufour F., *Stability and Ergodicity of Piecewise Deterministic Markov Processes*, Siam Journal of Control and Optimisation, 2008,
- [35] Elf J., Ehrenberg M., *Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases*, Syst.Biol., vol. 1, p. 230-236, 2004,
- [36] Elowitz M., Surette M., Wolf P.-E., Stock J., Leibler S., *Protein Mobility in the Cytoplasm of Escherichia Coli*, J. Bacteriology, vol. 181, p. 197-203, 1999,
- [37] S. N. Ethier, T. G. Kurtz, *Markov processes, characterization and convergence*. Wiley, New York, 1986.

- [38] Friedmann Nir, Long Cai, Xie Sunney, *Liking Stochastic Dynamics to Poulaiton Distribution : An Analytical Framdework of Gene Expression*, Physical Review Letter, 2006,
- [39] Fisher A., *Eugenics*, 1937,
- [40] Freidlin M., Wentzell A., *Random perturbations of dynamical systems*, Springer, New York,1984.
- [41] Gikhman I., Skorokhod A., *Introduction to the Theory of Random Processes*, W.B.Saunders Co., Philadelphia, 1969.
- [42] Gibson, M.A., *Computational Methods for Stochastic Biological Systems* (thesis), 2000,
- [43] Gibson, M.A., Bruck J., *A probabilistic model of a prokaryotic gene and its regulation*,
- [44] Gillespie D. T.J.Comput.Phys., vol. 22, p. 403, 1976,
- [45] Gillespie D. T., *The Chemical Langevin equation*, J.Chem.Phys, vol. 113, p. 297-306, 2000,
- [46] Gillespie, D.T., *Approximate accelerated stochastic simulation of chemically reacting systems*, J.Chem. Phys.,vol. 115, p.1716-733,2001,
- [47] Griffith M., Courtney T., Peccoud J., SandersW.H., *Dynamic partitioning for hybrid simulation of the bistable HIV-1 transactivation network*, Bioinformatics, 2006,
- [48] Gosh M.K., Bagchi A., *Modeling stochastic hybrid systems*,
- [49] Gorban AN, Radulescu O : *Dynamical robustness of biological networks with hierarchical distribution of time scales*, IET Systems Biology 2007, 1 :238-246,
- [50] Gorban AN, Radulescu O, *Dynamic and static limitation in multiscale reaction networks*, revisited, *Advances in Chemical Engineering : Mathematics and Chemical Engineering and Kinetics*, Volume 34. Edited by Marin G, West D, Yablonsky G, Academic Press 2008 :103-173,
- [51] Haseltine E. L., Rawlings J. B., *Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics*, J.Chem.Phys., vol. 117, p. 6959-6969, 2002,
- [52] Hasty, J., Pradines, J., Dolnik, M., Collins, J.J., *Noised-based switches and amplifiers for gene expression*, PNAS, 2000,

- [53] Hespanhan J.P., Singhz A., *Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems*, International Journal of robust and nonlinear control, 2005,
- [54] Hellander A., Lötstedt P., *Hybrid method for the chemical master equation*, Journal of Computational Physics, 2007,
- [55] Hoops S., Sahle S., Gauges R. , Lee C., Pahle J., Simus N., Singhal M., Xu L., Mendes P., and Kummer U., *COPASI—a COmplex PATHway SIMulator*, Bioinformatics, 2006,
- [56] Higham D.J., *Modeling and Simulating Chemical Reactions*, Siam Review, 2008,
- [57] Iwasa Y., Michor F., Nowak M. A., *Stochastic tunnels in evolutionary dynamics*, Genetics, vol. 166, p. 1571-1579, 2004,
- [58] Ikeda.Watanabe, *Stochastic Differential equations and diffusion process*. North Holland, Kodansha (1981).
- [59] Julius J.A., Halász A., Sakar M.S., Rubin H., Kumar V., Fellow, Pappas G.J., *Stochastic Modeling and Control of Biological Systems : The Lactose Regulation System of Escherichia Coli*, Systems Biology, 2008,
- [60] J..., Lee S., *Excluded Volume Effect on the Diffusion-Influenced Bimolecular Reactions*, J. Phys.Chem., vol. 101, p. 5255-5261, 1997,
- [61] Kaznessis Y.N., *Multi-scale models for gene network engineering*, Chemical Engineering Science, 2005,
- [62] Kaern M., A.Elston T., Blake W. J., Collins J. J., *Stochasticity in gene expression : from theories to phenotypes*, Nature Rev.Genet., vol. 6, p. 451-464, 2005,
- [63] Karmakar R., Bose I., *Graded and binary responses in stochastic gene expressions*, arXiv :qbio :OT/0411012v1, 2004,
- [64] Kouretas P., Koutroumpas K., Lygeros, J., Lygerou Z., *Stochastic Hybrid Modeling of Biochemical Processes*, Stochastic Hybrid Systems, 2006,
- [65] Korolyuk V., Swishchuk A., *Semi-Markov Random Evolutions*, Kluwer, Dordrecht, 1995,
- [66] Kiehl T.R., Mattheyses R.M., Simmons M.K., *Hybrid simulation of cellular behavior*, Bioinformatics, 2004,

- [67] Krishna S., Banerjee B., Ramakrishnan T., Shivashankar G.V., *Stochastic simulations of the origins and implications of a long-tailed distributions in gene expression*, PNAS, 2005,
- [68] Krishna S., Banerjee B., Ramakrishnan T., Shivashankar G.V. , *Supplementary material*, 2005,
- [69] Kierzek A., Zaim Jolanta, Zielenkiewicz P., *The Effect of Transcription and Translation Initiation Frequencies on the Stochastic Fluctuations in Prokaryotic gene Expression*, The journal of Biological Chemistry, 2001,
- [70] Kierzek A., *STOCKS : STOChastic Kinetic Simulations od biochemical systems with Gillespie algorithm*, Bioinformatics, 2002,
- [71] T.G. Kurtz, *Solutions of ordinary differential equations as limits of pure jump Markov process*, J.Appl.Prob., vol. 7, p. 49-58, 1970,
- [72] T.G. Kurtz, *Limit theorems for sequences of jump Markov processes approximating ordinary differential processes*, J.Appl.Prob., vol. 8, p. 344-356, 1971,
- [73] T.G. Kurtz, *Strong approximation theorems for density dependent Markov chains*, Stoch. Proc. Appl., 6, pp 223-240, 1978.
- [74] H.J. Kushner, *Approximation and weak convergence methods for random processes*, MIT Press, Cambridge, 1994.
- [75] Kuzovkov V., Kotomin E., *Kinetics of bimolecular reactions in condensed media*, Rep.Prog.Phys., vol. 51, p. 1479-1524, 1988,
- [76] Laforge B., Guez D., Martinez M., Kuipec 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, p. 93-120, 2005,
- [77] 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, p. 147-150, 2004,
- [78] Masty E.A., Haseltine E.L., Rawlings, J.B., *Two classes of quasi-steady-state model reductions for stochastic kinetics*, The Journal of Chemical Physics, 2007,
- [79] Meyn S. P., Tweedie R. L., *Markov Chains and Stochastic Stability*, Springer-Verlag, New York, 1993.

- [80] Morton-Firth C., *Stochastic simulation of cell signalling pathways*, Phd, Cambridge, 1998,
- [81] Muller A., *Modélisation stochastique des réseaux de gènes*, Rapport de stage, Université de Rennes 1, septembre, 2003.
- [82] Ochiai, H., Inoue, Y., Takeya, M., Sasaki, A. and Kaku, H., *Genome sequence of Xanthomonas oryzae pv. oryzae suggests contribution of large numbers of effector genes*, 2005,
- [83] Ozbudak E.M., *Regulation of noise in the expression of a single gene*, Nature Genetic, 31(1), 69-73, 2002,
- [84] G. Papanicolaou, D.W. Strook, S.R.S. Varadhan, *Martingale approach to some limit theorem dans "Statistical Mechanics ans Dynamical systems*, D. Ruelle ed.. Duke Turbulence Conf., Duke Univ. Math. Series III, part IV, pp1-120, 1976,
- [85] Paulsson, J, *Summing uo the noise in gene networks*, Nature, 2004, 427 :415-418,
- [86] Peleš S., Munsky B., Khammash M., *Reduction and solution of the chemical master equation using time scale separation and finite state projection*, The journal of Chemical Physics, 2006,
- [87] Petri, C. A. (1962), *Fundamentals of a theory of asynchronous information flow*, Proc. IFIP Congress, N-H.
- [88] Puchalka J., Kierzek A.M., *Bridging the Gap between Stochastic and Deterministic Regimes in the Kinetic Simulations of the Biochemical Reaction Networks*, Biophysical Journal, 2004,
- [89] Plyasunov S., *On Hybrid Simulation schemes for Stochastic Reaction Dynamics*, 2005,
- [90] Pola G, Bujorianu M, Lygeros J, Di Benedetto M, *Stochastic hybrid models : An overview*, Proceedings IFAC Conference on Analysis and Design of Hybrid Systems 2003,
- [91] Radulescu O, Gorban AN, Zinovyev A, Lilienbaum A : *Robust simplifications of multiscale biochemical networks*, BMC Systems Biology 2008, 2 :86.
- [92] 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, p. 4999-5010, 2003,

- [93] Rao C. V., Wolf D. M., Arkin A. P., *Control, exploitation and tolerance of intracellular noise*, Nature, vol. 420, p. 231-237, 2002,
- [94] 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, p. 47-56, 1995,
- [95] Rényi A., *Betrachtung chemischer Reaktionen mit Hilfe der Theorie der stochastischen Prozesse*, Magyar Tud.Akad.Alkalm.Mat.Int.Közl., vol. 2, p. 93-101, 1954,
- [96] Rudiger, S., Shuai, J.W., Huisinga, W., Nagaiah C., G. Warnecke, Parker I., Falckey M., *Hybrid Stochastic and Deterministic Simulations of Calcium Blips*, Biophysical Journal, 2007,
- [97] Rossinelli D., Bayati B., Koumoutsakos P., *Accelerated stochastic and hybrid methods for spatial simulations of reaction-diffusion systems*, Chemical Physics Letters, 2008,
- [98] Salis H., Kaznessis Y.N., *Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions*, The Journal of Chemical Physics, 2005,
- [99] Salis, H., Sotiropoulos, Kaznessis, Y.N., *Multiscale Hy3S : hybrid stochastic simulation for supercomputers*, BMC Bioinformatics, 2006,
- [100] Sanders J, Verhulst F : *Averaging methods in nonlinear dynamical systems*, Springer 1985,
- [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, p. 10322-10324, 2000,
- [102] Skorokhod A., *Asymptotic methods in the theory of stochastic differential equations*, AMS, Providence, 1989,
- [103] Swain P., Elowitz M., Siggia E., *Intrinsic and extrinsic contributions to stochasticity in gene expression*, Proc.Natl.Acad.Sci.USA, vol. 99, p. 12795-12800, 2002,
- [104] Stein R, Gossen E, Jones K : *Neuronal variability : noise or part of the signal ?*, Nature Reviews Neuroscience 2005, 6(5) :389,
- [105] Surovtsova I, Sahle S, Pahle J, Kummer U : *Approaches to complexity reduction in a systems biology research environment (SYCAMORE)*, Proceedings

- of the 37th conference on Winter simulation, Winter Simulation Conference 2006 :1683, 1689.
- [106] Takahashi K., Arjunan S., Tomita M., *Space in systems biology of signalling pathways - towards intracellular molecular crowding in silico*, FEBS Lett., vol. 579, p. 1783-1788, 2005,
- [107] Tian T., Burrage K., *Bistability and switching in the lysis/lysigeny genetic regulatory network of bacteriophage λ* , J.Theor.bio., vol. 227, p. 229-237, 2004,
- [108] Togashi Y., Kaneko K., *Discreteness-induced stochastic steady state in reaction diffusion systems : self-consistent analysis and stochastic simulations*, Physica D, vol. 205, p. 87-99, 2005,
- [109] 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, p. 221-231, 2003,
- [110] Yin G., *On Limit Results for a Class of Singularly Perturbed Switching Diffusions*, J.Theor.Probab., vol. 14, p. 673-697, 2001,
- [111] Radulescu O., Muller A., Crudu A., *Theoremes limites por les processus de Markov a sauts*, TSI, 2007,
- [112] Reddy V, Liebman M., Mavrovouniotis M., *Qualitative analysis of biochemical reaction systems. Computers in biology and medicine*, 26 (1) : 9-24, 1996,
- [113] Tuckwell, H.C., *Stochastic Processes in the Neurosciences*, SIAM, 1989,
- [114] Yin G., *On Limit Results for a Class of Singularly Perturbed Switching Diffusions*, Journal of Theoretical Probability, 2001,
- [115] Yu J, Xiao J, Ren X, Lao K, Xie XS : *Probing Gene Expression in Live Cells, One Protein Molecule at a Time*. Science 2006, 311 :1600{1603.
- [116] Zeiser S., Franz U., Wittich O., Liebscher V., *Simulation of genetic networks modelled by piecewise deterministic Markov processes*, IET System Biology, 2008.
- [117] Van Kampen N, *Stochastic processes in physics and chemistry*, third edition, North Holland 2007.